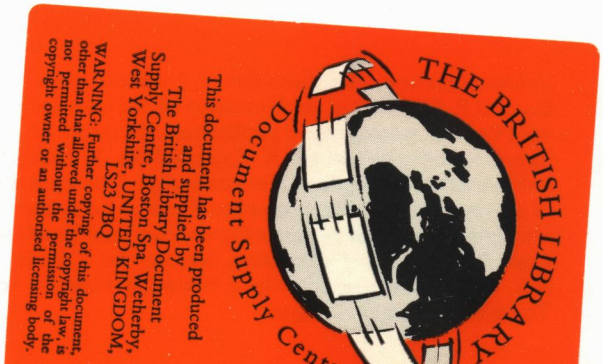


pathway in *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 86:8778–8782

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Sass P, Field J, Nikawa J, Toda T, Wigler M (1986) Cloning and characterization of the high-affinity cAMP phophodiesterase of *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 83:9303–9307

Shin DY, Matsumoto K, Iida H, Uno I, Ishikawa T (1987) Heat shock response of *Saccharomyces cerevisiae* mutants altered in cyclic AMP-dependent protein phosphorylation. Mol Cell Biol 7:244–250

Smith HE, Mitchell AP (1989) A transcriptional cascade governs entry into meiosis in *Saccharomyces cerevisiae*. Mol Cell Biol 9:2142–2152

Suzuki Y, Nogi Y, Abe A, Fukasawa T (1988) *GAL11* protein, an auxiliary transcription activator for genes encoding galactose-metabolizing enzymes in *Saccharomyces cerevisiae*. Mol Cell Biol 8:4991–4999

Tanaka K, Matsumoto K, Toh-e A (1988) Dual regulation of the expression of the polyubiquitin gene by cyclic AMP and heat shock in yeast. EMBO J 7:495–502

Tanaka K, Matsumoto K, Toh-e A (1989) *IRA1*, an inhibitory regulator of the RAS-cyclic AMP pathway in *Saccharomyces cerevisiae*. Mol Cell Biol 9:757–768

Tanaka K, Nakafuku M, Tamanoi F, Kaziro Y, Matsumoto K, Toh-e A (1990) *IRA2*, a second gene of *Saccharomyces cerevisiae* that encodes a protein with a domain homologous to mammalian ras GTPase-activating protein. Mol Cell Biol 10:4303–4313

Thompson-Jaeger S, Francios J, Gaughran JP, Tatchell K (1991) Deletion of *SNF1* affects the nutrient response of yeast and resembles mutations which activate the adenylate cyclase pathway. Genetics 129:697–706

Toda T, Uno I, Ishikawa T, Powers S, Kataoka T, Broek D, Cameron S, Broach J, Matsumoto K, Wigler M (1985) In yeast, RAS proteins are controlling elements of adenylate cyclase. Cell 40:27–36

Toda T, Cameron S, Sass P, Zoller M, Scott JD, McMullen B, Hurwitz M, Krebs EG, Wigler M (1987a) Cloning and characterization of *BCY1*, a locus encoding a regulatory subunit of the cyclic AMP-dependent protein kinase in *Saccharomyces cerevisiae*. Mol Cell Biol 7:1371–1377

Toda T, Cameron S, Sass P, Zoller M, Wigler M (1987b) Three different genes in *S. cerevisiae* encode catalytic subunits of the cAMP-dependent protein kinase. Cell 50:277–287

Toda T, Cameron S, Sass P, Wigler M (1988) *SCH9*, a gene of *Saccharomyces cerevisiae*, that encodes a protein distinct from, but functionally and structurally related to, cAMP-dependent protein kinase catalytic subunits. Genes Dev 2:517–527

Uno I, Matsumoto K, Adachi K, Ishikawa T (1983) Genetic and biochemical evidence that trehalase is a substrate of cAMP-dependent protein kinase. J Biol Chem 258:10867–10872

Wilson RB, Brenner AA, White TB, Engler MJ, Gaughran JP, Tatchell K (1991) The *saccharomyces cerevisiae SRK1* gene, a suppressor of *bcy1* and *ins1*, may be involved in protein phosphatase function. Mol Cell Biol 11:3369–3373

Yoshihisa T, Anraku Y (1989) Nucleotide sequence of *AMS1*, the structure gene of vacuolar α-mannosidase of *Saccharomyces cerevisiae*. Biochem Biophys Res Commun 163:908–915

Note added in proof

The accession number for the sequence reported in this paper is D13715.

Mol Gen Genet (1993) 238:17–25

MGG

The plant hormone abscisic acid mediates the drought-induced expression but not the seed-specific expression of *rd22*, a gene responsive to dehydration stress in *Arabidopsis thaliana*

Kazuko Yamaguchi-Shinozaki, Kazuo Shinozaki

Laboratory of Plant Molecular Biology, Tsukuba Life Science Center, The Institute of Physical and Chemical Research (RIKEN), 3-1-1 Koyadai, Tsukuba, Ibaraki 305, Japan

Received: 13 June 1992 / Accepted: 19 October 1992

Abstract. Nine cDNA clones, corresponding to genes that are responsive to dehydration (named RD), have been isolated from *Arabidopsis thaliana*. The sequence of a putative protein encoded by one of the RD cDNA clones, RD22, exhibits considerable homology to an unidentified seed protein (USP) of *Vicia faba*. Northern analysis showed that RD22 mRNA is induced by salt stress as well as by water deficit but not by cold or heat stress. RD22 mRNA appeared after the application of abscisic acid (ABA), an indication that transcription of RD22 mRNA is induced by endogenous ABA, the production of which is triggered by drought and salt stress. The induction of RD22 mRNA by ABA was inhibited by cycloheximide. Thus, it appears that protein synthesis is required for the induction of this mRNA by ABA. By contrast, protein synthesis was not required for the ABA-responsive induction RD29 mRNA, which corresponds to another dehydration-responsive gene of *A. thaliana*. These results suggest that there are at least two mechanisms for the induction of dehydration-responsive genes by ABA. RD22 mRNA was also expressed during the early and middle stages of seed development, showing a pattern of expression similar to that of USP. The seed-specific expression of RD22 seems not to be regulated by ABA. Structural analysis of the RD22 genomic clone revealed that the structural gene (designated *rd22*) contains three introns, and only a single copy of the gene is present in the *A. thaliana* genome, while the gene for USP from *V. faba* is actually a family of genes with 10 to 20 members. The site of initiation of transcription was determined by primer extension. Possible *cis*-acting elements involved in the expression of *rd22* are discussed.

Key words: *Arabidopsis thaliana* – Drought stress – Abscisic acid – Gene structure – Seed protein

The nucleotide sequence data reported in this paper will appear in the DDBJ, EMBL and Gene Bank Nucleotide Sequence Databases under the accession number D01113

Communicated by K. Isono

Correspondence to: K. Shinozaki

Introduction

Plants respond to drought or water stress with physiological changes as well as biochemical changes including the production of a number of proteins. Recently, many genes have been shown to be responsive to water deficit at the transcriptional level (Bartels et al. 1990; Guerrero et al. 1990; reviewed by Skriver and Mundy 1990). Most of the genes that have been studied to date are also induced by the plant hormone abscisic acid (ABA; Skriver and Mundy 1990; Mundy and Chua 1988; Close et al. 1989; Piatkowski et al. 1990; Bray 1991). Thus, plant water deficit triggers the production of ABA, which in turn induces the expression of various genes.

Many ABA-responsive genes are also expressed in late embryogenesis (Skriver and Mundy 1990; Bray 1991). The products of these genes are thought to function in protecting cells from dehydration (Dure III et al. 1989). Recently, several genes that are induced by water deficit have been proven to be unresponsive to ABA (Guerrero et al. 1990; Yamaguchi-Shinozaki et al. 1992), results that indicate the existence of ABA-independent as well as ABA-responsive signal transduction cascades between the initial signal of water stress and the expression of particular genes.

In order to learn more about the signal transduction pathways between water stress and gene expression, we have cloned and characterized genes that are responsive to dehydration in *Arabidopsis thaliana*. We have identified nine independent cDNA clones and named them RD (responsive to dehydration). Northern analyses revealed wide variation in the timing of induction of RD genes, and showed that two of the RD genes are responsive to ABA while three are not (Yamaguchi-Shinozaki et al. 1992). We have been interested in clarifying the role of ABA in the induction of RD genes, and further analysis revealed that six of the nine RD genes are responsive to exogenous ABA. However, the timing of induction of the ABA-responsive RD genes differed from gene to gene, suggesting that more than one signal pathway is operative between the production of ABA induced by

water stress and the expression of RD genes. Another possibility is that different ABA-inducible RD genes differ in sensitivity to the final step of the signal pathway.

In this report we describe the characterization and structural analysis of one of the ABA-responsive RD genes, *rd22*. We examined the effect of various stresses, such as cold, salt and heat stress, on the expression of this gene. Protein synthesis seems to be required for the induction of *rd22* by ABA. By contrast, tissue-specific expression of *rd22* at the early and middle stages of seed development seems not to be mediated by ABA.

Materials and methods

Plant growth. A. thaliana (Columbia ecotype) was grown under continuous illumination of approximately 2500 lx at 22° C for 3 weeks and harvested prior to bolting. Plants were grown on GM (germination medium) agar plates (Valvekens et al. 1988). The EN3 mutant (*abi3-3;* Nambara et al. 1992) was kindly provided by Drs. E. Nambara and P. McCourt and grown in soil as described (Yamaguchi-Shinozaki et al. 1992).

Dehydration, ABA, cycloheximide, cold and heat treatments. Whole plants were removed from GM agar plates and left to dehydrate in plastic dishes at 22° C and 60% humidity under dim light for up to 24 h. The relative water content (RWC) of the dehydrated plants is shown in Fig. 1A. Plants that were to be treated with ABA, NaCl or cycloheximide were removed from the agar plates and grown hydroponically in 100 µM ABA, 250 mM NaCl or 10 µM cycloheximide. Plants subjected to cold or heat stress were grown on GM agar plates at 22° C then transferred to 4° or 40° C under continuous light. At various times after the start of each treatment, the plants were frozen in liquid nitrogen.

RNA and DNA analyses. Total RNA was isolated by the method of Nagy et al. (1988). Poly(A)$^+$ RNA and genomic DNA were prepared by the methods of Maniatis et al. (1982). Genomic DNA and RNA blot hybridizations were performed as previously described (Maniatis et al. 1982; Yamaguchi-Shinozaki et al. 1989).

Generation of the rd22 *gene from a genomic library.* An *A. thaliana* genomic DNA library (Clontech, Palo Alto, Calif., USA) was screened by the plaque hybridization method of Maniatis et al. (1982). A 1.9 kb *Bam*HI fragment containing the entire coding sequence of the RD22 cDNA clone was labeled with ^{32}P by use of a random primer kit in accordance with the manufacturer's instructions (Boehringer Mannheim, Mannheim, FRG) and used as a DNA probe for the hybridization. Positive clones were plaque-purified and their DNA was isolated by phenol extraction and CsCl gradient centrifugation (Maniatis et al. 1982).

Subcloning and sequencing of fragments containing the rd22 *gene.* DNA fragments containing the *rd22* gene were subcloned into the pBluescript II vector (Stratagene, La Jolla, Calif., USA). A 5.5 kb *Eco*RI fragment was used for subcloning. An Exo/mung deletion kit (Stratagene) was used to make a deletion series in order to determine the DNA sequence of the 5.5 kb *Eco*RI fragment. A Sequenase version 2.0 kit (United States Biochemical Corporation, Cleveland, Ohio, USA) was used. Nucleotide and amino acid sequences were analyzed with the GENETYX software system (Software Development Co., Tokyo, Japan).

Primer extension analysis. Primer extension was performed as described (Yamaguchi-Shinozaki et al. 1989) using the following 24 base oligonucleotide, which corresponds to the complementary sequence of the coding region of *rd22:* 5′-CCGGTGTTAAATCAGCCGCAATCG-3′ (position +212 through position +189; Fig. 5).

Results

Northern analysis of expression of the rd22 *gene*

We have isolated nine independent cDNAs, designated RD, corresponding to drought-induced genes from *A. thaliana* by differential screening (Yamaguchi-Shinozaki et al. 1992). Induction of the corresponding RD genes by water deficiency was analyzed by RNA blot analysis. Whole *A. thaliana* plants were dehydrated for various lengths of time up to 24 h then total RNA was isolated and subjected to Northern analysis. Expression of all nine genes corresponding to the RD clones was induced by water deficit, but the timing of induction varied between the RD genes. Figure 1A shows the accumulation of RD22 mRNA in *A. thaliana* in response to dehydration and a number of other stress conditions, as determined by Northern blot analysis. The level of RD22 mRNA was very low in fully hydrated plants grown on GM agar plates. However, the mRNA was detected within 2 h after the onset of dehydration and it accumulated to high levels at 5 h. The transcript remained at an elevated level for at least 24 h. The RWC of 24 h dehydrated plants went down to 19% (Fig. 1A). However, all the dehydrated plants recovered even after 24 h dehydration. When plants were exposed to a water deficit for 8 h and then rehydrated by transfer to water, they recovered within 4 h and continued to grow. After rehydration had begun, mRNA corresponding to RD22 was still detected for at least 4 h but it became undetectable within 24 h, indicating that expression of the gene is induced by water deficiency (Fig. 1B).

We also examined the effect of cold and salt stress on the expression of RD22 mRNA in *A. thaliana* (Fig. 1A). When plants were transferred from normal growth conditions at 22° to 4° C under continuous light, we did not observe any increase in the level of RD22 mRNA. Under the same conditions, other genes corresponding to RD29, RD17, RD2, RD26 and RD20 are induced (Yamaguchi-Shinozaki and Shinozaki, unpublished observation). These observations indicate that the expression of RD22 is not induced by cold stress. To examine the effect of salt stress on the expression of RD22, we analyzed

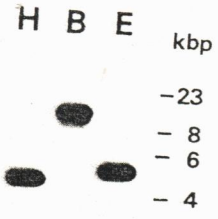


Fig. 3. Southern blot analysis of genomic DNA from *Arabidopsis thaliana*. Nuclear DNA from *A. thaliana* (5 µg) was digested to completion with *Hin*dIII (H), *Bam*HI (B) and *Eco*RI (E) and fractionated on a 0.7% agarose gel. The fragments were blotted onto a filter and hybridized to the RD22 cDNA

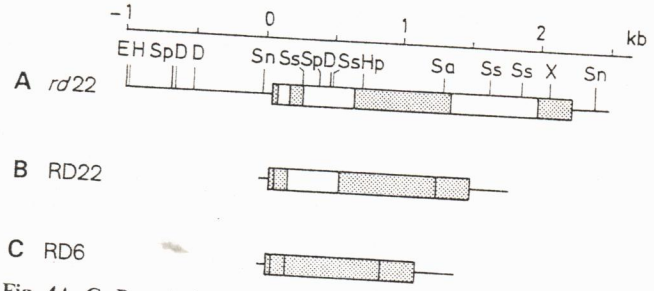


Fig. 4A–C. Restriction maps of *rd22*, RD22 and RD6 clones. The restriction maps of *rd22* genomic (**A**), RD22 cDNA (**B**) and RD6 cDNA (**C**) clones are shown. The DNA inserts in these clones were 5.5 kb *Eco*RI, 1.9 kb *Bam*HI and 1.5 kb *Bam*HI fragments, respectively. The *shaded* and *open boxes* represent exons and introns, respectively. The restriction sites are indicated as D (*Dra*I), E (*Eco*RI), H (*Hin*dIII), Hp (*Hap*I), Sa (*Sal*I), Sn (*Sna*BI), Sp (*Spe*I), Ss (*Ssp*I) and X (*Xho*I)

blot analysis with hybridization to the 1.9 kb *Bam*HI fragment of the RD22 cDNA (Fig. 3). Only one hybridizing band was identified in each digest, indicating that there is only one *rd22* gene in the *A. thaliana* genome. Thirteen genomic clones of the *rd22* gene were obtained by screening a genomic library of *A. thaliana* with the cDNA probe. Extensive restriction mapping of these clones showed that they all contained the same *rd22* gene, supporting the results of the DNA blot analysis.

Figure 4A shows a restriction map of the DNA fragment that contains the *rd22* gene. Figures 4B, C shows restriction maps of two cDNA clones, RD22 and RD6, respectively. We have sequenced all three of these clones and found that RD22 contains the second intron of the *rd22* gene that is present before splicing. Northern analysis revealed the presence of two species of RD22 RNAs of 1.9 and 1.5 kb in length (Figs. 1A, B). Thus, the second intron appears to be spliced out inefficiently, an observation similar to one reported for the gene for phytochrome in pea (Sato 1988).

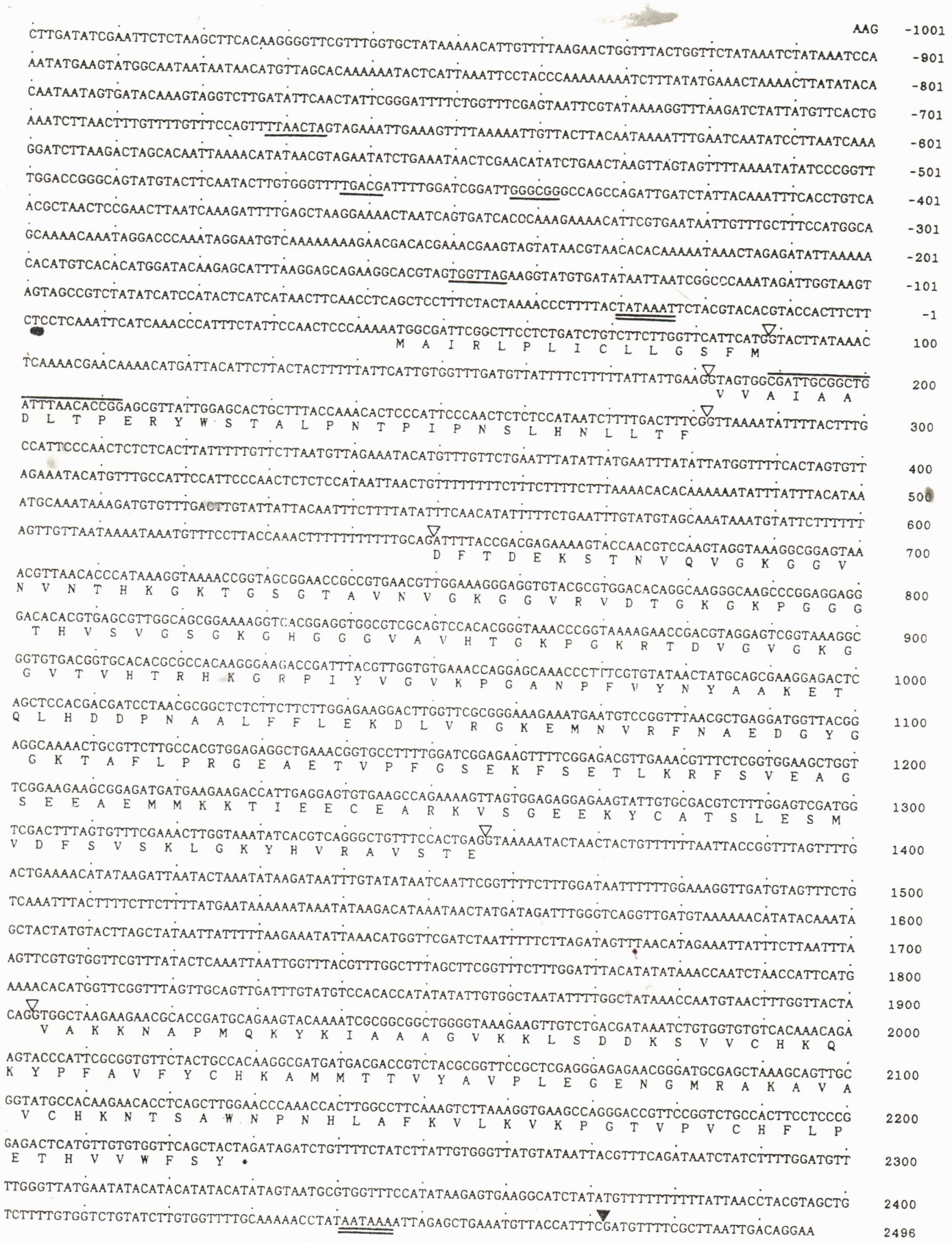
Figure 5 shows the nucleotide sequence of the *rd22* gene. The coding region and introns in this gene were determined on the basis of the nucleotide sequences of the RD6 and RD22 cDNA clones. The *rd22* gene contains three introns of 93, 369 and 457 bp, respectively. All three introns and the 5′ and 3′ untranslated regions are AT-rich while the coding regions have a high GC content. The 5′ end of the transcript was determined by primer extension (Fig. 6) and is indicated in the nucleotide sequence in Fig. 5 by position +1. The initiation codon is 44 nucleotides downstream from the site of initiation of transcription. This ATG triplet is flanked by a sequence that is similar to the consensus sequence frequently found in plants (AACA ATG GC; Joshi 1987). A typical TATA box sequence is located at position −30 (TATAAA). A number of sequence motifs have been identified as playing a role in the regulation of transcription of plant genes. We searched the upstream region of *rd22* for these *cis*-acting motifs, which include as1 (TGACG at −463; Lam et al. 1989) and sp1 (GGGCGG at −443; Briggs et al. 1986). We found two sequences that resembled *myb* recognition elements (TGGTTAG, the complementary strand of CTAACCA at −144, TTAACTA at −666) and two bHLH (basic helix-loop-helix) recognition elements (CACATG at −200 and at −191). Consensus sequences for the recognition sites of the myb and bHLH proteins are YAAC(G/T)G or (C/R)(Y/A)CAACYR (Biedenkapp et al. 1988; Nakagoshi et al. 1990) and CANNTG (Cai and Davis 1990), respectively. The functions of these sequences have not yet been elucidated. We could find neither a typical G-box sequence (ACGTGG/T; Schindler et al. 1992) nor an ABRE (ABA-responsive element) sequence (T/CACGTGGC; Yamaguchi-Shinozaki et al. 1989; Marcotte et al. 1989) in the 5′ flanking DNA of *rd22*.

The 3′ terminus of the *rd22* mRNA is located at position +2471 (Fig. 5). A typical polyadenylation signal is located at position +2441 (AATAAA).

Structure of the RD22 protein

The amino acid sequence of the RD22 protein was deduced from the nucleotide sequence of the RD22 clone (Figs. 5 and 7). The putative RD22 protein is 392 amino acids long with a predicted molecular weight of 42259 Da. We searched a protein data base, SWISS PROT, for homologies to RD22 and found that the carboxyl-terminal half of the RD22 protein exhibits sequence homology with an unidentified seed protein (USP) of the field bean (*Vicia faba* L; Bassüner et al. 1988) as shown in Fig. 7. USP mRNA is abundant during the early and middle stages but does not accumulate in

Fig. 5. Nucleotide sequence of the *rd22* gene. The nucleotide sequence of the noncoding strand, including the 5′ and 3′ flanking regions, is shown. Numbers refer to the nucleotides beginning at the 5′ end of the *rd22* mRNA. The amino acid sequence of the coding region is shown below the nucleotide sequence in the singler-letter code. A putative TATA box and polyadenylation signal, AATAAA, are *doubly underlined*. Sequences homologous to reported *cis*-acting elements are *underlined* (see text). The sequence used as the oligonucleotide for primer extension has a *line drawn above it*. *Open triangles* indicate junction sites between introns and exons. A *closed triangle* indicates the 3′ end of RD22 cDNA


AAG -1001
CTTGATATCGAATTCTCTAAGCTTCACAAGGGGTTCGTTTGGTGCTATAAAAACATTGTTTTAAGAACTGGTTTACTGGTTCTATAAATCTATAAATCCA -901
AATATGAAGTATGGCAATAATAATAACATGTTAGCACAAAAAATACTCATTAAATTCCTACCCAAAAAAAATCTTTATATGAAACTAAAACTTATATACA -801
CAATAATAGTGATACAAAGTAGGTCTTGATATTCAACTATTCGGGATTTTCTGGTTTCGAGTAATTCGTATAAAAGGTTTAAGATCTATTATGTTCACTG -701
AAATCTTAACTTTGTTTTGTTTCCAGTTTTAACTAGTAGAAATTGAAAGTTTTAAAAATTGTTACTTACAATAAAATTTGAATCAATATCCTTAATCAAA -601
GGATCTTAAGACTAGCACAATTAAAACATATAACGTAGAATATCTGAAATAACTCGAACATATCTGAACTAAGTTAGTAGTTTTAAAATATATCCCGGTT -501
TGGACCGGGCAGTATGTACTTCAATACTTGTGGGTTTTGACGATTTTGGATCGGATTGGGCGGGCCAGCCAGATTGATCTATTACAAATTTCACCTGTCA -401
ACGCTAACTCCGAACTTAATCAAAGATTTTGAGCTAAGGAAAACTAATCAGTGATCACCCAAAGAAAACATTCGTGAATAATTGTTTGCTTTCCATGGCA -301
GCAAAACAAATAGGACCCAAATAGGAATGTCAAAAAAAAGAACGACACGAAACGAAGTAGTATAACGTAACACACAAAAATAAACTAGAGATATTAAAAA -201
CACATGTCACACATGGATACAAGAGCATTTAAGGAGCAGAAGGCACGTAGTGGTTAGAAGGTATGTGATATAATTAATCGGCCCAAATAGATTGGTAAGT -101
AGTAGCCGTCTATATCATCCATACTCATCATAACTTCAACCTCAGCTCCTTTCTACTAAAACCCTTTTACTATAAATTCTACGTACACGTACCACTTCTT -1
CTCCTCAAATTCATCAAACCCATTTCTATTCCAACTCCCAAAAATGGCGATTCGGCTTCCTCTGATCTGTCTTCTTGGTTCATTCATGGTACTTATAAAC 100
M A I R L P L I C L L G S F M
TCAAAACGAACAAAACATGATTACATTCTTACTACTTTTTATTCATTGTGGTTTGATGTTATTTTCTTTTTATTATTGAAGGTAGTGGCGATTGCGGCTG 200
V V A I A A
ATTTAACACCGGAGCGTTATTGGAGCACTGCTTTACCAAACACTCCCATTCCCAACTCTCTCCATAATCTTTTGACTTTCGGTTAAAATATTTTACTTTG 300
D L T P E R Y W S T A L P N T P I P N S L H N L L T F
CCATTCCCAACTCTCTCACTTATTTTTGTTCTTAATGTTAGAAATACATGTTTGTTCTGAATTTATATTATGAATTTATATTATGGTTTTCACTAGTGTT 400
AGAAATACATGTTTGCCATTCCATTCCCAACTCTCTCCATAATTAACTGTTTTTTTTCTTTCTTTTCTTTAAAACACACAAAAAATATTTATTTACATAA 500
ATGCAAATAAAGATGTGTTTGACTTGTATTATTACAATTTCTTTTATATTTCAACATATTTTTCTGAATTTGTATGTAGCAAATAAATGTATTCTTTTTT 600
AGTTGTTAATAAAATAAATGTTTCCTTACCAAACTTTTTTTTTTTTGCAGATTTTACCGACGAGAAAAGTACCAACGTCCAAGTAGGTAAAGGCGGAGTAA 700
D F T D E K S T N V Q V G K G G V
ACGTTAACACCCATAAAGGTAAAACCGGTAGCGGAACCGCCGTGAACGTTGGAAAGGGAGGTGTACGCGTGGACACAGGCAAGGGCAAGCCCGGAGGAGG 800
N V N T H K G K T G S G T A V N V G K G G V R V D T G K G K P G G G
GACACACGTGAGCGTTGGCAGCGGAAAAGGTCACGGAGGTGGCGTCGCAGTCCACACGGGTAAACCCGGTAAAAGAACCGACGTAGGAGTCGGTAAAGGC 900
T H V S V G S G K G H G G G V A V H T G K P G K R T D V G V G K G
GGTGTGACGGTGCACACGCGCCACAAGGGAAGACCGATTTACGTTGGTGTGAAACCAGGAGCAAACCCTTTCGTGTATAACTATGCAGCGAAGGAGACTC 1000
G V T V H T R H K G R P I Y V G V K P G A N P F V Y N Y A A K E T
AGCTCCACGACGATCCTAACGCGGCTCTCTTCTTCTTGGAGAAGGACTTGGTTCGCGGGAAAGAAATGAATGTCCGGTTTAACGCTGAGGATGGTTACGG 1100
Q L H D D P N A A L F F L E K D L V R G K E M N V R F N A E D G Y G
AGGCAAAACTGCGTTCTTGCCACGTGGAGAGGCTGAAACGGTGCCTTTTGGATCGGAGAAGTTTTCGGAGACGTTGAAACGTTTCTCGGTGGAAGCTGGT 1200
G K T A F L P R G E A E T V P F G S E K F S E T L K R F S V E A G
TCGGAAGAAGCGGAGATGATGAAGAAGACCATTGAGGAGTGTGAAGCCAGAAAAGTTAGTGGAGAGGAGAAGTATTGTGCGACGTCTTTGGAGTCGATGG 1300
S E E A E M M K K T I E E C E A R K V S G E E K Y C A T S L E S M
TCGACTTTAGTGTTTCGAAACTTGGTAAATATCACGTCAGGGCTGTTTCCACTGAGGTAAAAATACTAACTACTGTTTTTTAATTACCGGTTTAGTTTTG 1400
V D F S V S K L G K Y H V R A V S T E
ACTGAAAACATATAAGATTAATACTAAATATAAGATAATTTGTATATAATCAATTCGGTTTTCTTTGGATAATTTTTTGGAAAGGTTGATGTAGTTTCTG 1500
TCAAATTTACTTTTCTTCTTTTATGAATAAAAAATAAATATAAGACATAAATAACTATGATAGATTTGGGTCAGGTTGATGTAAAAAACATATACAAATA 1600
GCTACTATGTACTTAGCTATAATTATTTTTAAGAAATATTAAACATGGTTCGATCTAATTTTTCTTAGATAGTTTAACATAGAAATTATTTCTTAATTTA 1700
AGTTCGTGTGGTTCGTTTATACTCAAATTAATTGGTTTACGTTTGGCTTTAGCTTCGGTTTCTTTGGATTTACATATATAAACCAATCTAACCATTCATG 1800
AAAACACATGGTTCGGTTTAGTTGCAGTTGATTTGTATGTCCACACCATATATATTGTGGCTAATATTTTGGCTATAAACCAATGTAACTTTGGTTACTA 1900
CAGGTGGCTAAGAAGAACGCACCGATGCAGAAGTACAAAATCGCGGCGGCTGGGGTAAAGAAGTTGTCTGACGATAAATCTGTGGTGTGTCACAAACAGA 2000
V A K K N A P M Q K Y K I A A A G V K K L S D D K S V V C H K Q
AGTACCCATTCGCGGTGTTCTACTGCCACAAGGCGATGATGACGACCGTCTACGCGGTTCCGCTCGAGGGAGAGAACGGGATGCGAGCTAAAGCAGTTGC 2100
K Y P F A V F Y C H K A M M T T V Y A V P L E G E N G M R A K A V A
GGTATGCCACAAGAACACCTCAGCTTGGAACCCAAACCACTTGGCCTTCAAAGTCTTAAAGGTGAAGCCAGGGACCGTTCCGGTCTGCCACTTCCTCCCG 2200
V C H K N T S A W N P N H L A F K V L K V K P G T V P V C H F L P
GAGACTCATGTTGTGTGGTTCAGCTACTAGATAGATCTGTTTTCTATCTTATTGTGGGTTATGTATAATTACGTTTCAGATAATCTATCTTTTGGATGTT 2300
E T H V V W F S Y *
TTGGGTTATGAATATACATACATATACATATAGTAATGCGTGGTTTCCATATAAGAGTGAAGGCATCTATATGTTTTTTTTTTATTAACCTACGTAGCTG 2400
TCTTTTGTGGTCTGTATCTTGTGGTTTTGCAAAAACCTATAATAAAATTAGAGCTGAAATGTTACCATTTCGATGTTTTCGCTTAATTGACAGGAA 2496

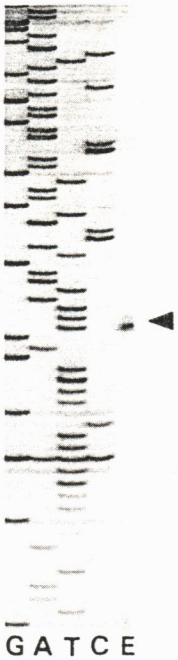


Fig. 6. Determination of the 5′ end of the transcript of the *rd22* gene by primer extension. An oligonucleotide primer (indicated in Fig. 4) was used for reverse transcription with 10 µg of total RNA from whole plants that had been dehydrated for 24 h. A *triangle* indicates the position of the major extension product. A portion of the sequence of the coding strand of the *rd22* gene is shown next to the sequence of the extension product (E)

the late stage of seed development in the field bean. However, in contrast to seed storage proteins, such as globulin, vicillin and legumin, neither USP nor its mRNA accumulate in mature seeds. A stretch of hydrophobic amino acids, typically found in signal peptides, is found at the N-terminus of the USP from *V. faba* (Bassüner et al. 1988). We found a similar hydrophobic region at the N-terminus of the RD22 protein (Fig. 7), which suggests that RD22, like USP, is translocated across the membrane of the endoplasmic reticulum. The RD22 protein contains a stretch of 109 amino acids with five tandemly repeated sequences in the N-terminal half of the polypeptide (underlined in Fig. 7). The consensus sequence in these repeats is T-V-VG-GGV------KGK, and, according to predictions of protein structure, it contains a beta sheet – turn – beta sheet structure. The function of this repeated sequence remains to be elucidated.

The structural similarities between RD22 and USP prompted us to perform Northern blotting to examine whether *rd22* is expressed in developing seeds of *A. thaliana*. Figure 8a shows an RNA blot of *rd22* mRNA in developing seeds of *A. thaliana*. We detected *rd22* mRNA during the early and middle stages (3 to 6 days after anthesis) of seed development of *A. thaliana*. The level of this mRNA decreased at the late stage (9 to 12 days after anthesis) and it did not accumulate in mature seeds (15 days after anthesis). This pattern is similar to the temporal pattern of expression of the mRNA for USP (Bassüner et al. 1988). Most of the LEA genes, which are expressed at the late stage of embryogenesis, are responsive to ABA (Baker et al. 1988). ABA is produced during the middle and late stages of seed development (Skriver and Mundy 1990). We examined the role of ABA in the expression of *rd22* using developing seeds of the EN3 mutant *(abi3-3)* of *A. thaliana*, which is insensitive to ABA with respect to germination and in which development of seeds is abnormal (Nambara et al. 1992). Developing seeds of EN3 remain green until maturity; they do not tolerate desiccation and do not become dormant. Desiccated seeds of the EN3 mutant do not survive, an indication of a null or very hypomorphic mutation. Figure 8b shows a Northern blot from developing seeds of the EN3 mutant. The pattern of temporal expression of the *rd22* gene in the EN3 mutant was similar to that of the wild type. By contrast, the levels of mRNAs transcribed from the 2S (at2S1) and 12S (CRA) genes that encode seed reserve proteins were significantly reduced in the EN3 mutant at the early and middle stages of seed development (Nambara et al. 1992). The level of 2S and 12S proteins in mature seeds in EN3 was also lower than that in the wild type (Nambara et al. 1992). These observations suggest that ABA is not involved in the seed-specific expression of *rd22*. In vegetative tissues of EN3

```
RD22    MAIRLPLICLLGSFMVVAIAADLTPERYWSTALPNTPIPNSLHNLLTFDFTDEKSTNVQV   60
USP92   MEFAHLTVLSLFCLAFVGITATSSEEDYWQSIWPNTPLPKTFSDLLIPSGKTN         53

RD22    GKGGVNVNTHKGKTGSGTAVNVGKGGVRVDTGKGKPGGGTHVSVGSGKGHGGGVAVHTGK  120

RD22    PGKRTDVGVGKGGVTVHTRHKGRPIYVGVKPGANPFVYNYAAKETQLHDDPNAALFFLEK  180

RD22    DLVRGKEMNVRFNAEDGYGGKTAFLPRGEAETVPFGSEKFSETLKRFSVEAGSEEAEMMK  240
USP92   SLPIKSEELKQYSTLFFEHDLHPRKNFILGNTNSVGSIIRPFTKSRQGVTDSIWLANKEK  113

RD22    KTIEE-CEARKVSGEEKYCATSLESMVDFSVSKLGKYHVRAVSTEVAKKNAPMQKYKIAA  299
USP92   QSFEDFCYSPTAIAEHKHCVSSLKSMIDQVISHFG----STKIKAISSNFAPYQD-QYVV  167

RD22    AGVKKLSDDKSVVCHKQKYPFAVFYCHKAMMTTVYAVPLEGENGMRAKAVAVCHKNTSAW  359
USP92   EDVKKVGDN-AVMCHRLNFEKVVFNCHQVRETTAYVVSLVASDGTKTKALTVCHHDTRGM  226

RD22    NPNHLAFKVLKVKPGTVPVCHFLPETHVVWFSY                             392
USP92   NP-ELLYEALEVTPGTVPVCHFIGNKAAAWVPNHTADNLCVM                    267
```

Fig. 7. Comparison of the amino acid sequence of the RD22 protein with that of the unidentified seed protein (USP) from *Vicia faba*. *Asterisks* denote amino acids that are identical in the two sequences and *dots* show conserved amino acids or neutral changes in amino acids. *Arrows* indicate the positions of introns. *Triangles* indicate the putative sites of processing of the signal peptides. The five repeated sequences in the RD22 protein are *underlined*

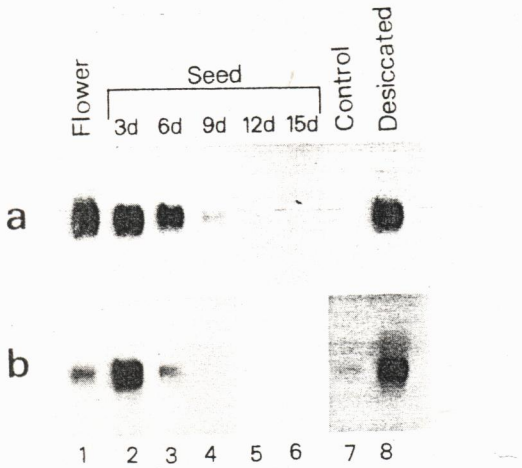


Fig. 8a, b. Northern blot analysis of expression of RD22 mRNA during seed development in the wild type (**a**), and the EN3 mutant (**b**). Each lane contained 30 μg of total RNA prepared from flowers (lane 1), developing seeds (including siliques) harvested the indicated number of days after anthesis (lane 2–6), fully hydrated rosette plants (lane 7) and rosette plants dehydrated for 10 h (lane 8). Northern hybridization was performed as described in Materials and methods

plants, RD22 mRNA also accumulated during water stress (Fig. 8). In vegetative tissues, such as leaves, roots and flowers, the level of ABA and the sensitivity to ABA in the EN3 mutant are similar to those in wild type as the effects of the *abi3* mutation are limited to the stage of seed development.

Discussion

We have cloned nine independent cDNAs of *A. thaliana* for genes that are induced by dehydration (RD clones) by differential screening (Yamaguchi-Shinozaki et al. 1992). The mRNAs of all the nine RD clones accumulate during water stress. We sequenced one of the RD clones, RD22, and, on examining the deduced amino acid sequence, we found sequences homologous to those of USP of *V. faba* (Bassüner et al. 1988) in the carboxy-terminal half of the protein (Fig. 7A). The first 22 amino acids at the amino-terminus are also homologous to the analogous region of the USP from *V. faba* and this region is highly hydrophobic. This region in the USP has been shown by translocation in vitro to be a signal peptide, which assists in the translocation of the protein across the endoplasmic membrane. It seems likely that the amino-terminus of the RD22 protein functions in a similar manner. The USP mRNA is abundant during the early and middle stages of seed development (Bassüner et al. 1988). Northern analysis revealed that RD22 mRNA is expressed during the early and middle stages of seed development, with a pattern of expression similar to that of USP. The structural gene corresponding to RD22, *rd22* appears to be present as a single copy in the *A. thaliana* genome, while the gene for USP in *V. faba* is actually a family of genes with roughly 10 to 20 members (Bäumlein et al. 1991). Genes for USP from *V. faba* have not been examined in terms of desiccation-responsive expression in vegetative plants. Some members of the family of USP genes may respond to water stress. The *rd22* gene contains three introns while the USP gene has only two (Bäumlein et al. 1991). The first two introns are at very similar positions in both genes (Fig. 6), and their similar locations support the hypothesis that *rd22* and the genes for USP have a common ancestor.

RD22 mRNA begins to appear within 2 h of the start of dehydration and is strongly expressed by 5 h. RD22 mRNA is induced by salt stress as well as by water deficit. Measurement of endogenous levels of ABA induced by water deficiency in *A. thaliana* revealed that ABA is detectable 2 h after dehydration and reaches a maximum level at 10 h (Kiyosue and Shinozaki, unpublished observations). RD22 mRNA was induced within 2 h after the application of exogenous ABA. Therefore, endogenous ABA, which is induced by these stresses, appears to be involved in expression of the *rd22* gene. Since cycloheximide inhibited the induction of RD22 mRNA (Fig. 2), protein synthesis is required in the induction of expression of the *rd22* gene by ABA. The ABA-mediated induction of the gene for proteinase inhibitor II (PI-II) of potato and tomato is also blocked by inhibitors of protein synthesis (Peña-Cortés et al. 1989). These results indicate that the production of protein factor(s), such as transcription factors, is required for the induction of *rd22* as well as of the gene for PI-II. By contrast, the induction of the *rab16A* gene of rice by ABA is not affected by various inhibitors of protein synthesis, such as cycloheximide, anisomycin and chloramphenicol (Mundy and Chua 1988). Protein synthesis is also not required for induction of expression of *rd29* by ABA, (Fig. 2). Therefore, the transcription factor(s) that regulate the expression of the *rab16A* or *rd29* genes must be present before the induction of transcription of these genes. The different requirements of translation processes in ABA-mediated gene expression support the hypothesis that there are at least two mechanisms for the induction of genes by ABA. Cold stress induces the synthesis of ABA (Chen et al. 1983). ABA has been shown to play a central role in the acclimation of *A. thaliana* (Nordin et al. 1991; Gilmour and Thomashow 1991). The *rd22* gene is not induced by cold stress. However, the level of ABA seems not to be high enough to induce transcription of *rd22*. Another possibility is that the synthesis of protein(s) required for expression of *rd22* does not proceed at low temperature.

The RD22 mRNA was detected at the early and middle stages of seed development but not at the late stage (Fig. 8A). However, ABA is produced during the middle stage of seed development (Skriver and Mundy 1990). The level of RD22 mRNA in developing seeds of the *abi3-3* mutant (EN3; Nambara et al. 1992), an ABA-insensitive mutant with respect to seed development, was the same as that in the wild type (Fig. 8B). The effects of the *abi3* mutation are similar to those of the *viviparous-1 (vp-1)* mutation in maize (Nambara et al. 1992). The *Vp-1* gene has been cloned by transposon tagging in maize and has been shown to encode a transcriptional activator that controls a number of ABA-inducible genes

during seed development (McCarty et al. 1991). Recently, the *abi3* gene has been shown to encode a *Vp-1* homolog (Giraudat et al. 1992). These results indicate that tissue-specific expression of *rd22* in developing seeds is not mediated by ABA. It is likely that ABA-independent *cis*-acting motif(s) are involved in the seed-specific expression of *rd22*.

An ABA-responsive element (ABRE) has been identified in the *rab16* gene of rice (Yamaguchi-Shinozaki et al. 1989; Mundy et al. 1990) and the *Em* gene of wheat (Marcotte et al. 1989), and it has a consensus sequence of RYACGTGGYR. Nuclear proteins that bind to this conserved sequence have also been identified (Mundy et al. 1990; Guiltinan et al. 1990). Recently, a cDNA clone encoding a DNA-binding protein(s) that interacts with an ABRE was isolated and the protein(s) was shown to be a member of the bZip family with a leucine zipper and a conserved region of basic amino acids (Guiltinan et al. 1990; Oeda et al. 1991). A double CATG sequence (CATGCATG; Sph element) was found in the promoters of the *C1*, *Em* and *rab17* genes (Paz-Ares et al. 1987; Hattori et al. 1992; Marcotte et al. 1989; Vilardell et al. 1990). The Sph element and GTGTC repeated sequence function in ABA-responsive induction of the *C1* gene of maize (Hattori et al. 1992). We found a tandemly repeated CATG motif (CACATGTCACACATG) in a region −200 bp upstream from the site of initiation of transcription of *rd22*. In this sequence, there are two bHLH recognition sequences (CANNTG) and pyrimidine-purine repeats. Four repeats of the hex3 sequence (Lam and Chua 1991), which is a mutant variant of the sequence of the histone hex1 sequence, have been shown to function as an ABRE in transgenic tobacco. We could find neither typical ABRE nor hex3 sequences in the promoter region of *rd22*.

The induction of *rd22* by ABA requires translation while that of *rd29* and rice *rab16A* does not (Fig. 2; Mundy and Chua 1988). Thus, the *cis*-acting element(s) that functions in the induction of *rd22* could be different from those found at the site of initiation of transcription of *rd29* and *rab16A*. Recently, we cloned a gene for a homolog of *myb* in *A. thaliana*, which is induced by the application of ABA, as well as by dehydration treatment (Urao, Yamaguchi-Shinozaki and Shinozaki, unpublished observation). In the upstream region of *rd22*, we found two *myb*-like recognition elements (Fig. 5), which could be possible *cis*-acting elements involved in the induction of *rd22* by ABA. Further analysis of these sequences is required to determine their function in the *rd22* promoter.

Acknowledgements. We thank Drs. E. Nambara of Nagoya University and P. McCourt of the University of Toronto for providing us with the EN3 mutant before publication. We thank Ms. S. Urao for technical assistance and Ms. K. Keith and Ms. Y. Takagi for editing and assistance in preparing this manuscript. This work is supported in part by the Special Coordination Fund of the Science and Technology Agency of the Japanese Government and by a Grant-in-Aid from the Ministry of Education, Science and Culture, Japan, to K.S. K.Y.-S. was supported by a fellowship from the Science and Technology Agency of Japan.

References

Baker J, Steele C, Dure L III (1988) Sequence and characterization of 6 *Lea* proteins and their genes from cotton. Plant Mol Biol 11:277–291

Bartels D, Schneider K, Terstappen G, Piatkowski D, Salamini F (1990) Molecular cloning of abscisic acid-modulated genes which are induced during desiccation of the resurrection plant *Craterostigma plantagineum*. Planta 181:27–34

Bassüner R, Bäumlein H, Huth A, Jung R, Wobus U, Rapoport TA, Saalbach G, Müntz K (1988) Abundant embryonic mRNA in field bean (*Vicia faba* L.) codes for a new class of seed proteins: cDNA cloning and characterization of the primary translation product. Plant Mol Biol 11:321–334

Bäumlein H, Boerjan W, Nagy I, Bassüner R, van Montagu M, Inze D, Wobus U (1991) A novel seed protein gene from *Vicia faba* is developmentally regulated in transgenic tobacco and *Arabidopsis thaliana* plants. Mol Gen Genet 225:459–467

Biedenkapp H, Borgemeyer U, Sippel, Klempnauer K-H (1988) Viral *myb* oncogene encodes a sequence-specific DNA-binding activity. Nature 335:835–837

Bray EA (1991) Regulation of gene expression by endogenous ABA during drought stress. In: Davies WJ, Jones HG (eds) Abscisic acid: physiology and biochemistry. Bios Scientific Publishers, Oxford, pp 81–98

Briggs MR, Kadonaga JT, Bell SP, Tjian R (1986) Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. Science 234:47–52

Cai M, Davis RW (1990) Yeast centromere-binding protein CBF-1, of helix-loop-helix protein family, is required for chromosome stability and methionine prototrophy. Cell 61:437–446

Chen HH, Li PH, Brenner ML (1983) Involvement of abscisic acid in potato cold acclimation. Plant Physiol 71:362–365

Close TJ, Kortt AA, Chandler PM (1989) A cDNA-based comparison of dehydration-induced proteins (dehydrins) in barley and corn. Plant Mol Biol 13:95–108

Dure L III, Crouch M, Harada J, Ho THD, Mundy J, Quatrano RS, Thomas T, Sung ZR (1989) Common amino acid sequence domains among the LEA proteins of higher plants. Plant Mol Biol 12:475–486

Gilmour SJ, Thomashow MF (1991) Cold acclimation and cold-regulated gene expression in ABA mutants of *Arabodopsis thaliana*. Plant Mol Biol 17:1233–1240

Giraudat J, Hauge BM, Valon C, Smalle J, Parcy F, Goodman HM (1992) Isolation of the Arabidopsis *ABI3* gene by position cloning. Plant Cell 4:1251–1261

Guerrero FD, Jones JT, Mullet JE (1990) Turgor-responsive gene transcription and RNA levels increase rapidly when pea shoots are wilted. Sequence and expression of three inducible genes. Plant Mol Biol 15:11–26

Guiltinan MJ, Marcotte WR, Quatrano RS (1990) A plant leucine-zipper protein that recognizes an abscisic acid response element. Science 250:267–271

Hattori T, Vasil V, Rosenkrans L, Hannah C, McCarty D (1992) The *Viviparous-1* gene and abscisic acid activate the *C1* regulatory gene for anthocyanin biosynthesis during seed maturation in maize. Gene Dev 6:609–618

Joshi CP (1987) An inspection of the domain between putative TATA boxes and translation start sites in 79 plant genes. Nucleic Acids Res 15:6643–6653

Lam E, Chua N-H (1991) Tetramer of a 21-base pair synthetic element confers seed expression and transcriptional enhancement in response to water stress and abscisic acid. J Biol Chem 266:17131–17135

Lam E, Benfey PN, Gilmartin PM, Fang RX, Chua N-H (1989) Site-specific mutations alter in vitro factor binding and change promoter expression pattern in transgenic plants. Proc Natl Acad Sci USA 87:7891–7894

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, pp 187–198

Marcotte WR Jr, Russell SH, Quatrano RS (1989) Abscisic acid response sequences from the *Em* gene of wheat. Plant Cell 1:969–976

McCarty DR, Hattori T, Carson CB, Vasil V, Vasil IK (1991) The *viviparous-1* developmental gene of maize encodes a novel transcription factor. Cell 66:895–905

Mundy J, Chua NH (1988) Abscisic acid and water-stress induce the expression of a novel rice gene. EMBO J 2279–2286

Mundy J, Yamaguchi-Shinozaki K, Chua N-H (1990) Nuclear proteins bind conserved elements in the abscisic acid-responsive promoter of rice *rab* genes. Proc Natl Acad Sci USA 87:406–410

Nagy F, Kay SA, Chua N-H (1988) Analysis of gene expression in transgenic plants. In: Gelvin SV, Schilperoort RA (eds) Plant molecular biology manual, B4. Kluwer Academic Publishers, Dordrecht, pp 1–29

Nakagoshi H, Nagase T, Kanei-Ishii C, Ueno Y, Ishii S (1990) Binding of the *c-myb* proto-oncogene product to the simian virus 40 enhancer stimulates transcription. J Biol Chem 265:3479–3483

Nambara E, Naito S, McCourt P (1992) A mutant of *Arabidopsis* which is defective in seed development and storage protein accumulation is a new *abi3* allele. Plant J 2:435–441

Nordin K, Heino P, Palva ET (1991) Separate signal pathways regulate the expression of a low-temperature-induced gene in *Arabidopsis thaliana*. (L.) Heynh. Plant Mol Biol 16:1061–1071

Oeda K, Salinas J, Chua N-H (1991) A tobacco pZip transcription activator (TAF-1) binds to a G-box-like motif conserved in plant genes. EMBO J 10:1793–1802

Paz-Ares JD, Gohsal D, Wienand U, Peterson PA, Saedler H (1987) The regulatory *C1* locus of *Zea mays* encodes a protein with homology to *myb* proto-oncogene products and with structural similarities to transcriptional activators. EMBO J 6:3553–3558

Peña-Cortés H, Sánchez-Serrano J, Mertens R, Willmitzer L, Prat S (1989) Abscisic acid is involved in the wound-induced expression of the proteinase inhibitor II gene in potato and tomato. Proc Natl Acad Sci USA 86:9851–9855

Piatkowski D, Schneider K, Salamini F, Bartels D (1990) Characterization of five abscisic acid-responsive cDNA clones isolated from the desiccation-tolerant plant *Craterostigma plantagineum* and their relationship to other water stress genes. Plant Physiol 94:1682–1688

Sato N (1988) Nucleotide sequence and expression of the phytochrome gene in *Pisum sativum:* differential regulation by light of multiple transcripts. Plant Mol Biol 11:679–710

Schindler U, Terzaghi W, Beckmann H, Kadesch T, Cashmore AR (1992) DNA-binding site preferences and transcriptional activation properties of the *Arabidopsis* transcription factor GBF1. EMBO J 11:1275–1289

Skriver K, Mundy J (1990) Gene expression in response to abscisic acid and osmotic stress. Plant Cell 2:502–512

Valvekens D, Montagu MV, Lijsebettens MV (1988) *Agrobacterium tumefaciens*-mediated transformation of *Arabidopsis thaliana* root explants by using kanamycin selection. Proc Natl Acad Sci USA 85:5536–5540

Vilardell J, Goday A, Freire MA, Torrent M, Martinez C, Torne JM, Pages M (1990) Gene sequence, developmental expression and protein phosphorylation of RAB17 in maize. Plant Mol Biol 14:423–432

Yamaguchi-Shinozaki K, Mundy J, Chua N-H (1989) Four tightly linked *rab* genes are differentially expressed in rice. Plant Mol Biol 14:29–39

Yamaguchi-Shinozaki K, Koizumi M, Urao S, Shinozaki K (1992) Molecular cloning and characterization of 9 cDNAs for genes that are responsive to desiccation in *Arabidopsis thaliana:* sequence analysis of one cDNA clone that encodes a putative transmembrane channel protein. Plant Cell Physiol 33:217–224