ABA-REGULATED GENE EXPRESSION: CIS-ACTING SEQUENCES AND TRANS-ACTING FACTORS

William R. Marcotte, Jr.+, Mark J. Guiltinan* and Ralph S. Quatrano
Department of Biology, University of North Carolina, CB# 3280, Coker Hall
Chapel Hill, NC 27599-3280

*present address: Department of Horticulture, The Pennsylvania State University, University Park, PA 16802).

+ To whom all correspondence should be sent

Send all correspondence to: William R. Marcotte, Jr.
Department of Biology
University of North Carolina
Chapel Hill, NC 27599-3280
(919) 962-0778 (phone)
(919) 962-0778 (fax)

BIOCHEMICAL SOCIETY TRANSACTIONS

(1992) 20: 93-97

Introduction

During plant development, phytohormones influence such diverse processes as seed maturation, germination and fruit ripening. The basis of phytohormone action has been a subject of intense interest to plant physiologists for many years, however, the molecular events involved have not been fully elucidated for any given response. While not all responses to hormones involve alterations in gene expression (e.g., stomatal closure in response to abscisic acid [ABA], see Hetherington and Quatrano, 1991 for review), it has been demonstrated that the expression of specific sets of RNA and protein molecules can be correlated with many of the physiological responses of plants to these compounds [2-6]. These sets of products may be unique to one response or there may be overlap in the sets between several responses.

The expression of hormone-regulated genes may be regulated at several levels, including transcription [7]. The transcription machinery interacts with the DNA primarily in the 5' flanking or promoter region of the gene and the pattern of expression of a single gene or set of genes is determined by these interactions. As such, a fundamental understanding of transcriptional regulation will be an important step in elucidating the molecular basis of phytohormone action. In this paper, we discuss recent findings which identify at least some of the cis sequences involved in ABA-mediated gene regulation [8-10] and a protein factor which interacts specifically with those sequences [10]. With this central point in the ABA signal transduction pathway identified, it should now be possible to follow the response chain in both directions ultimately identifying the initial receptor for ABA, the molecular mechanism(s) involved in gene activation and the role(s) of the proteins produced in response to this hormone.

Identification of an ABA Response Element (ABRE)

During late seed development in a wide variety of plants, both monocot and dicot, a set of abundant RNAs and their respective proteins is produced coincident with increases in endogenous ABA levels [11-13]. The accumulation of these products is correlated with the acquisition of desiccation tolerance of the seed embryo [14,15]. These same RNAs and proteins

can be precociously induced in embryos of earlier developmental stages through *in vitro* culture in the presence of ABA [16,17]. In the absence of ABA, these products do not accumulate and the embryos precociously germinate into completely normal seedlings.

In wheat, one of these RNAs encodes the embryo maturation protein, Em [6]. Early work with the Em gene demonstrated that approximately 1800bp of the 5' flanking region was sufficient to direct the induction of expression of an Em-ß-glucuronidase fusion gene (Em-GUS) in rice protoplasts in response to ABA [8]. The response was rapid, specific for ABA, and proportional to the amount of ABA added. A similar response was directed by a shorter 652bp fragment indicating that the shorter fragment contained all the information necessary for ABA-mediated induction. In addition, transcript mapping experiments demonstrated that the mRNA produced by the fusion in the rice protoplasts initiated at the same place as the wildtype Em transcript found in wheat seeds. This same 652bp promoter fragment, when used for the production of transgenic Em-GUS tobacco plants, was capable of directing the appropriate temporal and spatial expression of enzyme activity [9], i.e. only in the embryos of seeds and at a time when endogenous ABA levels normally increase during development. This activity was also inducible precociously in earlier stages of development by *in vitro* culture of excised embryos in the presence of ABA. As such, the results obtained using the transient expression system have been taken to be physiologically relevant, to accurately predict the response *in planta* and have been exploited to rapidly identify the sequences involved.

Deletion of 5' flanking sequence has been routinely used to identify sequences involved in directing a hormone-mediated response in many systems including animals [18,19] and plants [20,21]. Marcotte et al. (1989) demonstrated that deletion of Em 5' promoter sequences between -1800 and -168, while lowering the absolute levels of expression, had no qualitative effect on the ABA-mediated induction of GUS enzyme activity. The lower levels of expression observed when these sequences were deleted was attributed to the presence of several blocks of A/T-rich sequences. Similar sequences have been associated with quantitative levels of expression in a variety of genes possibly as a result of interactions with high mobility group

(HMG) proteins [22,23]. In contrast, however, deletion of an additional 62bp of 5' flanking region drastically reduced the induction phenomenon [9]. As such, at least part of an ABRE must be present within this 62bp.

Within this region of the Em promoter, there are several sequence motifs which appear in other seed protein gene 5' flanking regions. One of these, the Em1 box, is found in the promoter regions of all ABA-regulated genes for which sequence data is available. This sequence is present twice (as an imperfect repeat) in the Em promoter proximal to the 62bp deletion which abolished ABA induction. Guiltinan, et al. (1990) showed that a 76bp oligonucleotide containing the Em1 boxes (a and b), in the context of the -90 cauliflower mosaic virus (CaMV) 35S promoter, was sufficient for ABA induction of GUS expression in the transient assay, whereas the same -90 promoter fragment is by itself uninducible. Furthermore, they went on to show that mutation of the Em1a box reduced or eliminated the induction phenomenon. Marcotte, et al. (this work) have gone on to demonstrate that a 20bp oligonucleotide containing Em1a, in the context of the -90 CaMV 35S promoter, is sufficient for ABA-mediated induction (Table 1).

ABRE-Specific DNA-Protein Interaction

Gene transcription is controlled by the interaction of specific proteins, transcription factors, with specific DNA sequences. The nature of these interactions determines the functional state of the promoter, i.e. induced or repressed. The modulation of gene expression by external and internal stimuli, such as environmental and developmental cues, is believed to be accomplished by alterations in these interactions. As part of an ongoing effort to dissect the components of the ABA signal transduction pathway, we have identified a protein in nuclear extracts from wheat which is capable of binding to the ABA response element [10]. This interaction is specific for the ABRE identified above as the protein interacts with the Em1a box within the 76 bp oligonucleotide which was capable of mediating ABA induction in the transient assay but this same protein does not interact with the Em1b box, despite its close sequence

similarity as determined by methylation interference footprinting. In addition, competition experiments demonstrated that the 76bp oligonucleotide containing the Em1a mutation, which was greatly attenuated for ABA induction in the transient assay, was also not an efficient target sequence for this binding activity.

A λgt11 cDNA expression library made from wheat embryo RNA was screened for the presence of fusion proteins which would interact specifically with the ABRE and two positive clones were identified (λgc12 and λgc19). Extracts from lysogens of these recombinant phage contain fusion proteins with higher molecular weights than the λgt11 control extracts and the fusion proteins encoded by these recombinant phage have been shown to exhibit specific binding affinity for the ABRE by electrophoretic mobility shift assay (EMSA) and filter binding assay [10]. Figure 1 shows that the 76bp oligonucleotide probe bound by lysogen extracts encoding the fusion protein is dependent on the presence of a wt Em1a box (compare WT to Em1a single mutant and Em1a/b double mutant probes). In contrast, the Em1b sequence does not appear to contribute to binding either as a mutant or non-mutant site. Methylation interference experiments have established that the specificity of the fusion proteins encoded by the phage is identical to that of the protein found in wheat embryo nuclear extracts [10].

Analysis of the two positive clones revealed almost completely overlapping sequences. The deduced protein sequence contains features which have been found in a number of other transcription factors, in particular a conserved highly basic region juxtaposed with a leucine heptad repeat (leucine zipper). This class of proteins has been given the name basic-leucine zipper or bZIP and members of this class have been described in mammals, yeast, and plants [24]. Interaction of bZIP proteins with the DNA is mediated by the basic region and the leucine zipper sequences are involved in the formation of protein dimers. Chromosomal mapping studies have shown that the plant sequences contained on λgc19 appear to be part of a gene family comprised of 9 or 10 members [25]. Northern hybridization revealed only one transcript length (~1800 bases) indicating that, if these sequences represent a member of a family, all the members produce transcripts of similar length.

The partial cDNA from λgc19 was used to identify and isolate a full length EmBP (Em binding protein) cDNA. Sequence analysis of this full length clone has shown it to be similar to a number of other bZIP proteins in plants. Comparisons of the sequence with the GBF (G-box binding factor) sequences of Cashmore, et al. (1991, personal communication), the HBP (Histone DNA binding protein) sequences of Tabata, et al. (1989, 1991) and the CPRF (common plant regulatory factor) sequences of Weisshaar, et al. (1991) show a very high conservation particularly in the basic region [30]. The amino terminal halves of these proteins, however, share only limited similarity, suggesting that these regions might encode functional differences. Other short conserved regions, which are shared by several of these factors, are not necessarily common to all. As such, these significant similarities/differences may represent a basis on which to begin to group the different factors into subclasses.

The similarities within the basic DNA binding domain of these proteins are not surprising due to the similarity in the target sequences for these proteins. All bind to a sequence which contains a CACGTGG or closely related core [26-29]. Due to these similarities, we investigated the possibility of cross-competition for binding to the EmBP factor. Using 20 bp target sequences, the Em1a sequence was the most efficient competitor for the binding activity present in wheat nuclear extracts and the Em1b sequence was the least effective. While the 20bp target sequences from the ribulose bisphosphate carboxylase and chalcone synthase promoter regions were not as efficient competitors as the Em1a sequence, all were capable of competing for binding (data not shown).

As mentioned earlier, changes in the transcriptional activity of genes are attributed to alterations in the interactions of proteins with specific target sequences on the DNA. However, the mere presence of a particular protein factor may not be sufficient for binding and thereby activation. Data presented by Datta and Cashmore (1989) suggests that protein modifications, e.g. phosphorylation/dephosphorylation, may be involved in the activation of some transcription factors. To begin to address this possibility within the context of EmBP-1, we have analyzed different tissues and developmental stages to determine if the presence of the mRNA for EmBP-1

can be correlated with the presence of Em mRNA. Using specific PCR primers to generate products diagnostic of the presence of these two mRNAs, we have been able to demonstrate that the mRNA for EmBP is present in most tissues of the wheat plant but that this presence does not necessarily correlate with presence of the Em mRNA (data not shown). As expected, both the Em and EmBP mRNAs are detectable in Stage 3 wheat embryos cultured in the presence of ABA. However, the EmBP mRNA is also present in roots and leaves while the Em mRNA is not. This result is consistent with a model in which the EmBP factor must be activated in some way, possibly through protein modification or interaction with accessory proteins whose expression is developmentally or spatially regulated. One must be cautious, however, in the interpretation of these results as they address only expression at the mRNA level and not the presence of the corresponding protein. Extension of this analysis to the protein level will await the generation of antibodies specific for EmBP.

Conclusions

These and similar studies are identifying the molecular components involved in the transcriptional regulation of gene expression in response to a variety of signals. It is only through a thorough knowledge of these most basic cellular processes that it will be possible to address more complex developmental problems. What is beginning to emerge is the realization that general themes in regulation have been highly conserved during the evolution of divergent species of animals and plants. This conservation of structure and function within the plant kingdom is exemplified by the similarity of factors and target sequences involved in responses to such diverse signals as light, hormones and stress. The identification of components of these signal transduction pathways will eventually provide an understanding and appreciation of the mechanisms whereby a whole organism assimilates information into a coordinated physiological response.

The results discussed here provide evidence for the existence of an ABA response element in the 5' flanking region of the Em gene from wheat. The reduction of the required sequences to a

20 bp oligonucleotide and the ability of that sequence to direct ABA-induced expression highlights the versatility of a transient expression system for the rapid analysis of sequence requirements. The identification of a nuclear protein which interacts specifically with those sequences may provide additional insight into the molecular mechanisms involved in ABA signal transduction. Having identified a central point in the ABA signal transduction pathway, it should now be possible to pursue the pathway in two directions, leading ultimately to the identification and characterization of an ABA receptor and also providing some insight into the function of the proteins produced in response to ABA. This work was supported by grants from the USDA Competitive Research Grants Office to W.R.M and R.S.Q.

FIGURE LEGENDS

Table 1. Transient Expression Analysis of Em1a Oligonucleotide

Legend: Transient expression of chimeric constructs was analyzed as described previously [9]. The arrow indicates the presence and orientation of the 20bp oligonucleotide. Protoplasts were incubated for 24 hrs prior to assay. Activities are given in pmoles 4-MU, $\mu g\ protein^{-1}$, hr^{-1}.

Figure 1. Filter Binding Assay

Legend: Immobilized protein extracts were assayed for binding to wildtype (WT) and mutant probe sequences as described previously [10]. Specific nucleotide mutations are shown in lower part of figure. Em1a, Em1b and Em1a/1b designate the box/boxes which contain mutations.

REFERENCES

1. Hetherington, A. & Quatrano, R.S.(1991) New Phytol. in press

2. Broglie, K.E., Biddle, P., Cressman, R. & Broglie, R. (1989) The Plant Cell **1**, 599-607

3. Jacobsen, J.V., & Beach, L. (1985) Nature **316**, 275-277

4. Key, J.L. (1989) Bioessays **11**:52-58

5. Theologis, A., Huynh, T.V., & Davis, R.W. (1985) J. Mol. Biol. **183**:53-68

6. Williamson, J.D. & Quatrano, R.S. (1988) Plant Physiology **86**, 208-215

7. Benfy, P.N., & Chua, N.-H. (1989) Science **244**, 174-181

8. Marcotte, W.R., Bayley, C.C. & Quatrano, R.S. (1988) Nature **335**, 454-457

9. Marcotte, W.R., Russell, S.H. & Quatrano, R.S. (1989) The Plant Cell **1**, 969-979

10. Guiltinan, M.J., Marcotte, W.R., Jr. & Quatrano, R.S. (1990) Science **250**, 267-271

11. Crouch, M.L. & Sussex, I.M. (1981) Planta **153**:64-74

12. Dure, III, L., Greenway, S.C. & Galau, G.A. (1981) Biochemistry **20**:4162-4168

13. Quatrano, R.S., Ballo, B.L., Williamson, J.D., Hamblin, M.T. & Mansfield, M. (1983) in Plant Molecular Biology (R. B. Goldberg, ed.) pp. 343-353, Liss, New York

14. Bartels, D., Singh, M., & Salamini, F. (1988) Planta **175**, 485-492

15. Dure, III, L., Grouch, M., Harada, J., Ho, T.-H.D., Mundy, J., Quatrano, R., Thomas, T. & Sung, Z.R. (1989) Plant Molecular Biology **12**, 475-486

16. Quatrano, R.S. (1987) in Plant hormones and their role in plant growth and development (P.J. Davies, ed.), pp. 494-514, Martinus Nijhoff, Dordrecht

17. Skriver, K. & Mundy, J. (1990) The Plant Cell **2**:503-512

18. Danesch, U., Gloss, B., Schmid, W., Schutz, G., Schule, R. & Renkawitz, R. (1987) EMBO J. **6**:625-630

19. Yamamoto, K.R. (1985) Ann. Rev. Genet. **19**:209-252

20. Chen, Z.L., Schuler, M.A. & Beachy R.N. (1986) Proceedings of the National Academy of Science (USA) **83**, 8560-8564

21. Huttly, A.K & Baulcombe, D.C. (1989) The EMBO J. **8**, 1907-1913

22. Bustos, M.M., Guiltinan, M.J., Jordano, J., Begum, D., Kalkan, F.A., & Hall, T.C. (1989) The Plant Cell **1**, 839-853

23. Pederson, T.J., Arwood, L.J., Spiker, S., Guiltinan, M.J. & Thompson, W.F. (1991) Plant Mol. Biol. **16**: 94-104

24. Vinson, C.R., Sigler, P.B. & McKnight, S.L. (1989) Science **246**, 911-916

25. Devos, K.M., Atkinson, M.D., Chinoy, C.N., Guiltinan, M.J., Quatrano, R.S. & Gale, M.D. (1991) Theoret. Appl. Genet. in press

26. Cashmore, A.R. (1991) Department of Biology, University of Pennsylvania, Philadelphia, PA 19104-6018

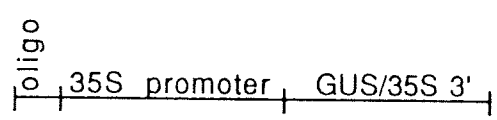
27. Tabata, T., Takase, H., Takayama, S., Mikami, K., Nakatsuka, A., Kawata, T., Nakayama, T. & Iwabuchi, M. (1989) Science **245**, 965-967

28. Tabata, T., Nakayama, T., Mikami, K. & Iwabuchi, M. (1991) EMBO J. **10**, 1459-1467

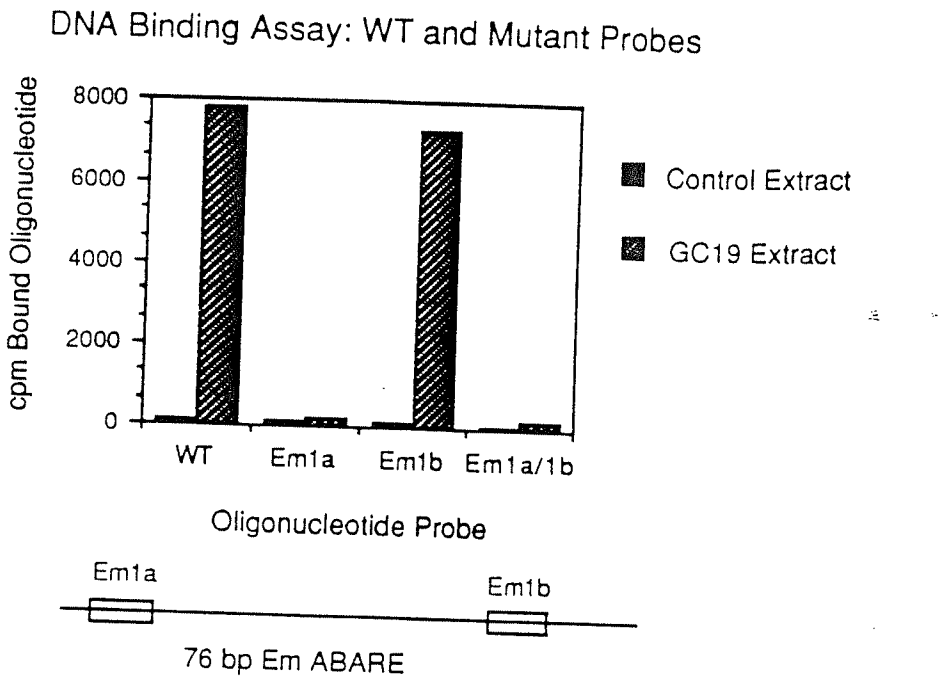
29. Weisshaar, B., Armstrong, G.A., Block, A., da Costa e Silva, O. & Hahlbrock, K. (1991) EMBO J. **10**:1777-1786

30. Quatrano, R. S., Guiltinan, M.J. & Marcotte, Jr., W.R. (1991) in Control of Gene Expression (D-P. S. Verma, ed.) in press, Telford Press, Caldwell, NJ

31. Datta, N. & Cashmore, A.R. (1989) The Plant Cell **1**, 1069-1077



	Activity		
	-ABA	+ABA	RATIO
-90	2.3 ± 0.8	3.8 ± 0.9	1.6
-90	0.5 ± 0.3	8.3 ± 1.0	17.0



			% BOUND
WT Em	:	TGCCGGGACACGTGGC.....CACACGTGCCGCCT	
Em1a	:	C G	100%
Em1b	:	T A	3.5%
Em1a/1b	:	C G T A	99.4%
			4.2%