The Plant Journal (2000) **24**(6), 703–711

Arabidopsis ARR1 and ARR2 response regulators operate as transcriptional activators

Hiroe Sakai, Takashi Aoyama and Atsuhiro Oka*
Laboratory of Molecular Biology, Institute for Chemical Research, Kyoto University, Uji, Kyoto, 611-0011 Japan

Received 24 July 2000; revised 13 September 2000; accepted 14 September 2000.
*For correspondence (fax +81 774 38 3259; e-mail oka@scl.kyoto-u.ac.jp).

Summary

The genes coding for the response regulators ARR1 and ARR2 have previously been identified by *in silico* screening of an expression sequence tag database and subsequent cloning from both *Arabidopsis* cDNA and genomic libraries. Their structures, in which the N-terminal signal receiver domain is followed by the output domain, are characteristic of typical bacterial response regulators of the two-component regulatory systems that control responses to a variety of environmental stimuli. Here we present evidence that these response regulators actually work as transcription factors. ARR1 and ARR2 were localized in the nuclei of plant cells regardless of the presence or absence of their signal receiver domain. Their middle segments, which faintly resemble the mammalian oncogene product Myb, were capable of binding double-stranded DNA in a sequence-specific manner *in vitro*. Their C-terminal halves functioned as transactivation domains in plant cells when combined with the DNA-binding domain of yeast GAL4. They thus possess all the essential components of a transcriptional activator. Both ARR1 and ARR2 promoted expression of a reporter gene in plant cells through their own target sequence. Truncation of their N-terminal signal receiver domain led to an increase in transactivation. An as yet unidentified phospho-relay signal may modulate the capability for transactivation and/or DNA binding through the signal receiver domain.

Keywords: DNA binding, Myb, response regulator, transcription factor, two-component regulatory system.

Introduction

Extracellular stimuli received by living cells are processed through signal transduction pathways and result in orchestrated gene expression. In bacteria, such a response is often controlled by the co-ordinated action of a sensor histidine kinase and a cytoplasmic response regulator. This type of control is called the two-component regulatory system (for reviews see Stock *et al.*, 1989; Stock *et al.*, 1990). Sensor kinases are generally composed of an individual N-terminal periplasmic domain with membrane-anchored regions and a common C-terminal cytoplasmic transmitter domain. The former domain, probably together with the neighbouring region, is thought to be involved in directly or indirectly monitoring an environmental parameter, whereas the latter transmitter domain phosphorylates its own specific His residue and then transfers the phosphoryl group to the cognate response regulator. The response regulators are also, in principle, composed of two functional domains. Their N-terminal halves contain the common signal receiver domain, having three separated hallmark Asp, Asp and Lys residues (hereafter referred to as DDK; Stock *et al.*, 1990), and the central Asp residue acquires the phosphoryl group from the phospho-His of its cognate transmitter. The C-terminal halves of the response regulators are individual output domains, the majority of which have the ability to bind DNA and to activate transcription of the target genes. Asp phosphorylation in the N-terminal domain modulates the output functions of the C-terminal domain. Some sensor kinases have an extra domain at their C-terminal end that resembles the signal receiver domain of the cognate response regulator: these are called hybrid-type sensor kinases (e.g. VirA, BvgC, ArcB, RcsC, BarA). The most common and simplest two-component regulatory systems do not involve additional signal transducers. In some cases, however, the phosphate group is transferred through a bridge component carrying the histidine-

containing phosphotransfer (HPt) domain that exists either alone (e.g. Spo0B) or as a portion of sensor kinases (e.g. ArcB) (Appleby *et al.*, 1996). Another bridge component is a group of polypeptide molecules containing the DDK signal receiver domain without any obvious output domain (e.g. Spo0F) (Appleby *et al.*, 1996). The signal receiver domain present in the hybrid-type sensor kinases might be included in this category. In any case, the typical bacterial response regulators that operate as transcription factors have always been located at the end of the intracellular phospho-relay signal transduction pathway.

The two-component regulatory system was initially thought to be restricted to prokaryotes, but it has since been observed in a wide array of eukaryotic species including plants, slime moulds, fungi and yeast (Wurgler-Murphy and Saito, 1997 and references therein). This suggests that His–Asp phosphotransfer mechanisms may also be involved in a wide variety of sophisticated eukaryotic signal transduction pathways. In higher plants, the *Arabidopsis* ethylene sensors (ETR1 and ERS) and the putative cytokinin sensor (CKI1) have structures that clearly resemble bacterial histidine kinase (Chang *et al.*, 1993; Hua *et al.*, 1995; Kakimoto, 1996). ETR1 and CKI1 are hybrid-type kinases, whereas the ERS members are ordinary kinases. The photoreceptor phytochromes may also be categorized as plant histidine kinase homologues, though they are considerably more divergent than other plant homologues (Schneider-Poetsch *et al.*, 1991). These findings allowed us to search for possible cognate response regulators on the analogy of bacteria. Fourteen *Arabidopsis* response regulator homologues (ARR1 to ARR14) have been identified, including some detected through *in silico* studies (Imamura *et al.*, 1998; Lohrmann *et al.*, 1999; Sakai *et al.*, 1998; Urao *et al.*, 1998). They were structurally classified into two groups (Imamura *et al.*, 1999; Sakai *et al.*, 1998). The members of one group include only the DDK signal receiver domain (Spo0F-type architecture), or at least do not contain additional, established or putative functional domains of a sensible molecular size. Several genes in this group are induced by cytokinin treatment (Brandstatter and Kieber, 1998; Taniguchi *et al.*, 1998). In contrast, the signal receiver domain in the other group is accompanied by additional functional domains. It is followed, in the cases of ARR1 and ARR2, by the potential nuclear localization signal VRK(R/K)R and a large C-terminal region with the following characteristics (Figure 1). A region of about 370 amino acid residues at the C-terminal end is rich in Gln residues (glutamine-rich domain) and the functional domain for transcriptional activation is often glutamine-rich (Triezenberg, 1995). A polypeptide region between the nuclear localization signal and the glutamine-rich domain can be divided into two parts. The first half (acidic domain) has about 70 residues, shows no similarity to any sequence in the protein database, and is a little rich in acidic Asp and Glu residues. The other half (ARRM domain) has 64 residues and faintly resembles the DNA-binding domain of the *c-myb* proto-oncogene product that consists of three repeats. The architecture and amino acid sequence of the ARRM domain are distinct from conventional plant Myb homologues. The majority of conventional plant Myb homologues contain two repeats while the remainder have three (Kranz *et al.*, 2000; Martin and Paz-Ares, 1997; Williams and Grotewold, 1997). The ARRM domain carries only one repeat, and two out of the three landmark Trp residues are substituted by Ala and Ile. Similar single-repeat structures (ARRM-like domains) have recently been found in many plant proteins whose cellular functions are mostly unknown (Sakai *et al.*, 1998), suggesting it is a functional unit that is probably involved in DNA binding. These structural characteristics allow us to presume that ARR1 and ARR2 are transcription factors similar to the majority of bacterial response regulators. In support of this view ARR11 (ARLP1), classified in the same group as ARR1 and ARR2 but considerably diverged from them, has been shown to be localized in the nuclei of parsley protoplasts and to activate transcription in yeast cells (Lohrmann *et al.*, 1999).

In this study we functionally dissected ARR1 and ARR2, and demonstrated their abilities both to bind DNA and to activate transcription. We also show that they work individually as the transcriptional activator in plant cells.

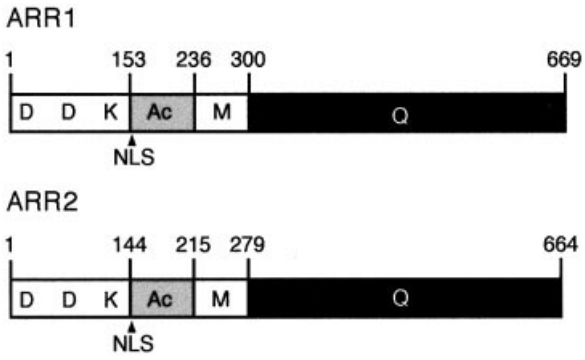

Figure 1. Domain organization of ARR1 and ARR2.
The domain organization of ARR1 and ARR2 has been deduced from their structural characteristics (Sakai *et al.*, 1998). D, D and K in the N-terminal signal receiver domain indicate the approximate positions of three hallmark Asp, Asp and Lys residues, respectively. Ac, M and Q mark the acidic domain, the ARRM domain, and the glutamine-rich domain, respectively. A triangle in the acidic domain shows the position of a potential nuclear localization signal, VRKRR for ARR1 and VRKKR for ARR2. Numerals at the domain boundaries correspond to the number of amino acid residues counted from the N-terminal end.

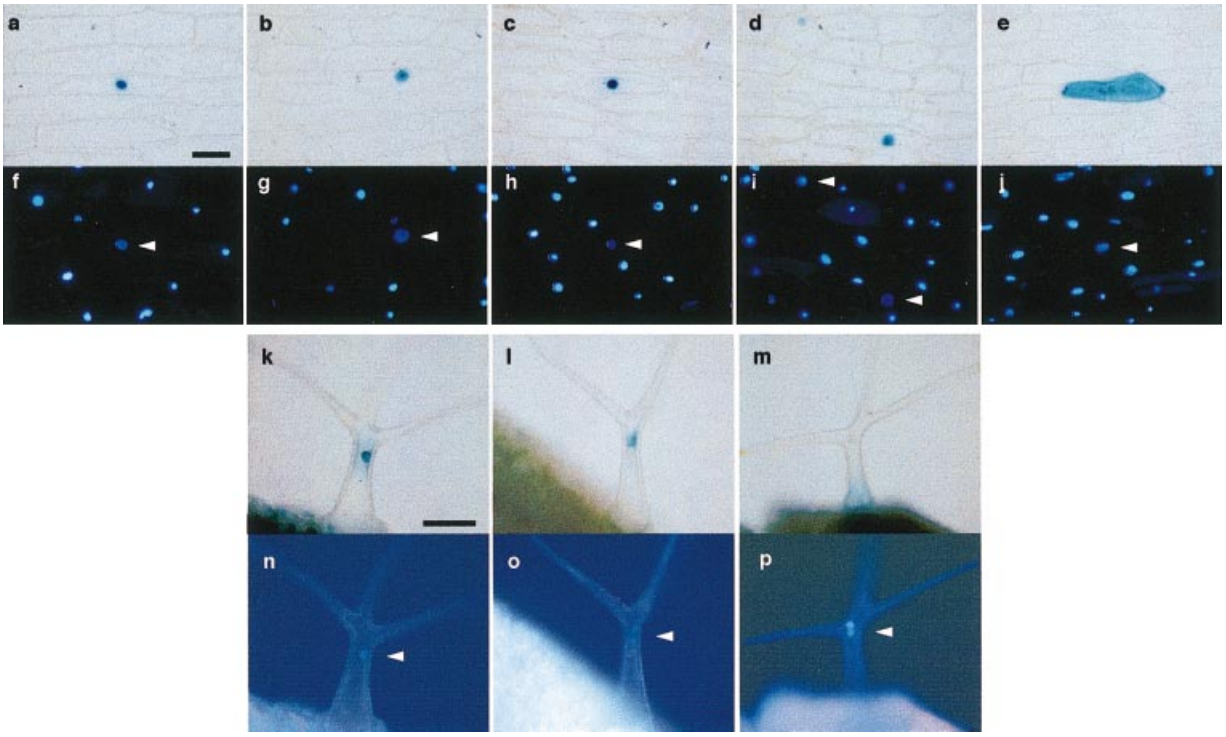

Figure 2. Nuclear localization of ARR1 and ARR2.
Onion epidermal cells to which DNA directing a GUS fusion protein had been delivered were stained for GUS activity: (a) ARR1::GUS; (b) ARR2::GUS; (c) ARR1Δ(DDK)::GUS; (d) ARR2Δ(DDK)::GUS; (e) control GUS without the ARR1 or ARR2 moiety. The same micrographs under UV light for nuclei identification (indicated by an arrowhead) by DAPI staining are presented in (f–j), respectively. *Arabidopsis* trichome cells (2 weeks old) expressing ARR1::GUS (k,n); ARR1Δ(DDK)::GUS (l,o); and GUS (m,p) were similarly treated. Scale bar, 100 μm.

Results

Localization of ARR1 and ARR2 in plant cell nuclei

The cellular locations of ARR1 and ARR2 were examined using β-glucuronidase (GUS) fusion proteins. Both fusion constructs were placed downstream of the cauliflower mosaic virus 35S promoter (hereafter referred to as the 35S promoter) and introduced into onion epidermal cells by the particle-delivery system (Sanford *et al.*, 1993). Histochemical analysis revealed GUS activity predominantly localized in nuclei, whereas activity in cells expressing GUS alone was mainly located in the cytoplasm (Figure 2a,f; b,g; e,j). Removing the DDK signal receiver domain (aa 1–152 for ARR1, aa 1–143 for ARR2; aa = amino acid positions numbered from the N-terminal end) from these fusion proteins did not affect nuclear localization (Figure 2c,h; d,i), suggesting that the transport of ARR1 and ARR2 to nuclei occurs without conditional modulation by the signal receiver domain. Similar results were obtained from histochemical analysis with trichome cells of transgenic *Arabidopsis* plants synthesizing the same fusion proteins (Figure 2k–p). It was thus concluded that ARR1 and ARR2 are normally located in plant cell nuclei.

Transactivation by the glutamine-rich domains of ARR1 and ARR2

Transactivation was examined by transient expression of a reporter gene that had been introduced into tobacco leaves together with an effector plasmid coding for a test polypeptide. The effector plasmids were designed to direct fusions of various portions of ARR1 and ARR2 with the DNA-binding domain of yeast GAL4. The reporter was the firefly luciferase gene (*LUC*) preceded by a promoter containing six repeats of the GAL4 target sequence ($6 \times UAS_G$) and a truncated 35S promoter region. The plasmid constructs and the results are shown in Figure 3. Transactivation occurred only when the effector contained the glutamine-rich domain (aa 300–669 for ARR1, aa 279–664 for ARR2), although the activity varied with the constructs. The results with the ARR1 and ARR2 constructs were comparable, but the activities seen with the ARR2 constructs were always higher than those of the corresponding ARR1 constructs. The highest activity observed was with the construct carrying only the glutamine-rich domain as the ARR2 moiety; this was nearly half the activity seen with the acidic domain of the herpes simplex viral protein VP16 (hereafter referred to as VP16). Attaching the acidic and ARRM domains to this construct appeared

to diminish the activity. Similar experiments with onion epidermal cells gave essentially the same results (data not shown). Therefore the glutamine-rich domains of ARR1 and ARR2 possess the ability to promote transactivation, although other domains may influence that ability.

Sequence-specific DNA binding by the ARRM domains of ARR1 and ARR2

In order to demonstrate the presumed DNA-binding ability of the ARRM domains, gel-retardation assays were carried out. The ARRM regions (aa 236–299 of ARR1-ARRM, aa 215–278 of ARR2-ARRM) were fused with glutathione S-transferase (GST). Crude extracts of *Escherichia coli* cells synthesizing the fusion proteins were prepared, and overproduction of the fusion proteins was verified by SDS–polyacrylamide gel electrophoresis (Figure 4a). These extracts were mixed with synthetic double-stranded oligonucleotides with random sequences of 19 base pairs (bp), and binding site selection experiments with PCR were carried out. Oligonucleotides capable of associating with ARRM were thus enriched, and then sequenced individually upon cloning. The selected oligonucleotides contained the sequence 5′-(G/A)GAT(T/C)-3′ at a frequency that was higher than expected by random occurrence (data not shown). There were no obvious conserved nucleotides flanking this sequence.

In order to prove the significance of this sequence, synthetic 34 bp oligonucleotides containing the 5 bp core sequence 5′-GGATT-3′ or mutant derivatives with a complementary transversion were subjected to gel-retardation assays (Figure 4b, left). Note that the design of the sequences flanking the 5 bp core was such that we rarely found them among the selected 19 bp oligonucleotides. The ARR1-ARRM domain bound 5′-GGATT-3′, and the substitutions within the central 5′-GAT-3′ sequence abolished the DNA–protein interaction, consistent with the results of the binding site selection experiments. Substitutions at the first and fifth positions in the core sequence also affected the gel-retardation pattern.

To further investigate the contributions of these two positions to ARRM binding, a series of substitution mutants were similarly analysed (Figure 4b, right). The first residue appeared to be less important because all substituted sequences were bound by ARRM, though at different degrees dependent on the base (A > G = T > C). Replacement of the fifth T with A or G seriously diminished ARRM binding, whereas replacement with C slightly enhanced binding. Enhancement by replacement with C might be related to the resulting creation of another 5′-GAT-3′ sequence in the complementary strand. Essentially, the same results were obtained with the ARR2-ARRM domain (data not shown). It was thus concluded that the ARRM domains of both ARR1 and ARR2 have the ability to bind double-stranded DNA in a sequence-specific manner, and that the optimum target sequence among the tested oligonucleotide species is 5′-AGATT-3′ under the experimental conditions used.

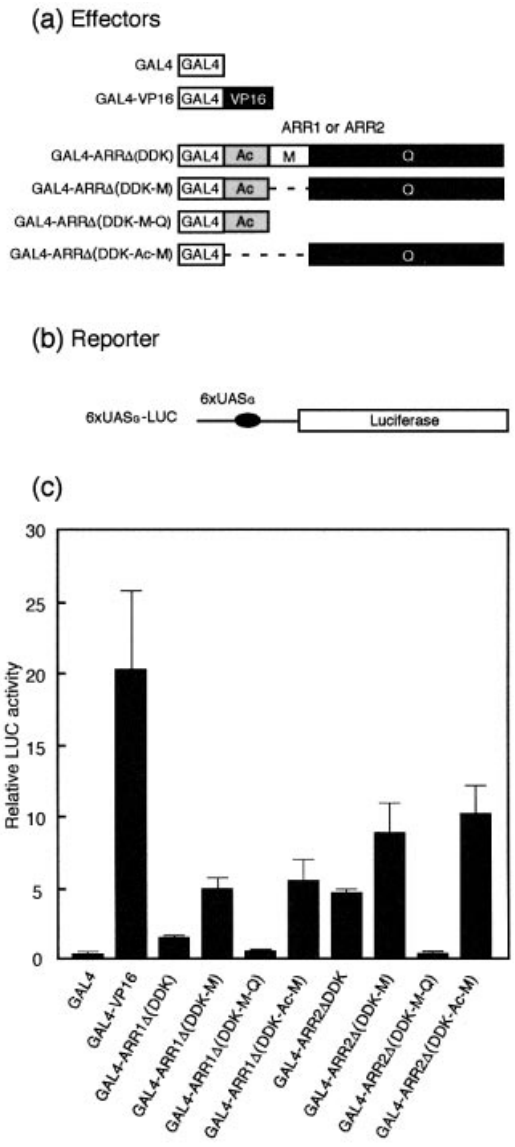

Figure 3. Transactivation by various portions of ARR1 and ARR2.
The chimeric constructs used for the effector and the reporter are schematically illustrated in (a) and (b), respectively. GAL4, VP16, Ac, M and Q indicate the DNA-binding domain of yeast GAL4; the acidic domain of the herpes simplex viral protein VP16; the ARR1/ARR2 acidic domain; the ARR1/ARR2 ARRM domain; and the ARR1/ARR2 glutamine-rich domain, respectively. 6 × UAS_G in (b) is a DNA segment containing six copies of the GAL4 target sequence. The ordinate in (c) shows the ability of each effector to activate transcription, as measured by the dual-luciferase reporter system (10 × LUC/RLUC). Each bar with a standard error of mean represents the average of four independent measurements.

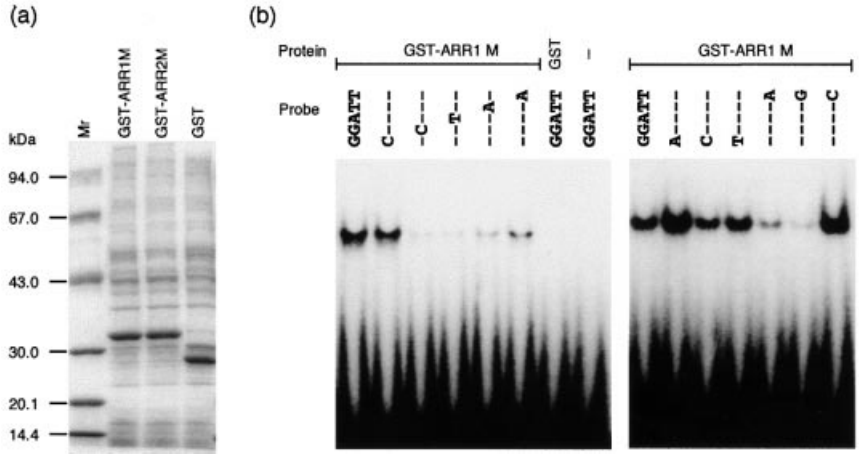

Figure 4. DNA-binding ability of ARR1 revealed by gel-retardation assay.
An electropherogram (a) of SDS–polyacrylamide gel stained with Coomassie brilliant blue R-250 shows overexpression of the GST::ARR1-ARRM (aa 236–299) fusion protein; the GST::ARR2-ARRM (aa 215–278) fusion protein; and GST (the densest band for each lane). Protein size markers were run in parallel. The protein samples used for the gel-retardation assay (b) are indicated at the top of the gel. The core 5 bp sequences of probe DNA are also shown above the gel, where the same residues as the original 5′-GGATT-3′ sequence are indicated by a hyphen. A minus mark means no protein sample added.

Transactivation by ARR1 and ARR2 through their own target sequence

Identification of the ARRM binding site sequence allowed us to examine whether intact ARR1 and ARR2 molecules actually activate transcription of a gene whose promoter contains this sequence. An artificial upstream activating sequence for ARR1 and ARR2 was designed in which the target sequence 5′-GGATT-3′ appears six times at different helical phases (6 × UAS_M). This sequence was connected to a truncated 35S promoter and then placed upstream of the *LUC* gene. The resulting reporter gene was introduced into tobacco leaves together with an effector plasmid, as described above. The effector plasmids carried ARR1 and ARR2 cDNAs driven by the 35S promoter. Their deletion derivatives, in which the DNA regions corresponding to the signal receiver domains were removed, were also used. Full-length ARR1 and ARR2, respectively, gave luminescent transactivation signals about two and three times higher than background levels observed with no effector DNA. Their truncated versions showed much higher activities of transactivation, about 10 and 17 times above background, respectively (Figure 5, black bars). Furthermore, replacement of the target sequence 5′-GGATT-3′ by 5′-GCATT-3′ in the artificial upstream activating sequence (6 × UAS_{Md}) cancelled the increased transactivation signals (Figure 5, grey bars). Since this mutation nearly completely abolished the qualification for the target *in vitro,* as shown in the preceding section, the data in Figure 5 apparently indicate that transactivation by ARR1 and ARR2 *in vivo* occurs through the same, or at least a similar, sequence-specific manner. Similar experimental results were obtained with onion epidermal cells (data not shown). It was thus concluded that both ARR1 and ARR2 potentially possess the ability to activate transcription in plant cells. The presence of a signal receiver domain appeared to partially inhibit transactivation.

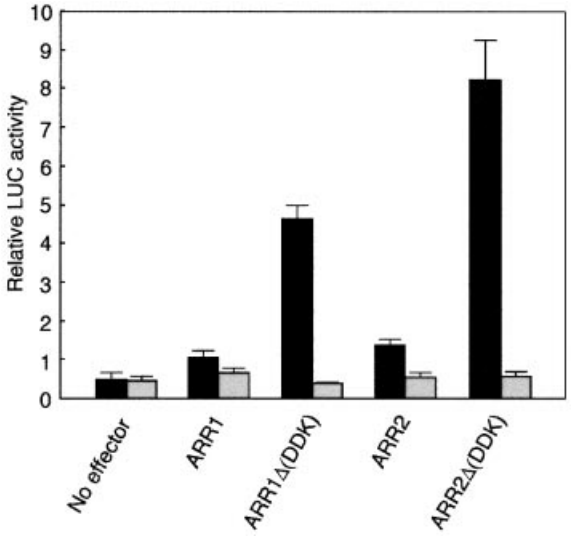

Figure 5. Transactivation by ARR1 and ARR2 through their own target sequence.
The effectors used were ARR1, ARR2 and their deletion derivatives lacking for the signal receiver domain, and the reporter was the same as in Figure 3. The ordinate indicates the transactivation activity of each effector measured by the dual-luciferase reporter system (100 × LUC/RLUC). Each bar with a standard error of mean represents the average of four independent experiments with 6 × UAS_M (black bars) and 6 × UAS_{Md} (grey bars). The 'No effector' sample contained an irrelevant pBI221 DNA to adjust the DNA amount for bombardment.

Discussion

ARR1 and ARR2 were originally identified as *Arabidopsis* homologues of the bacterial response regulators. Their N-terminal signal receiver domain is followed by a long C-terminal region where the potential nuclear localization signal, the short acidic region, the ARRM domain and the glutamine-rich domain are located. This architecture suggests that they are transcription factors. If this is the case, ARR1 and ARR2 should be able to localize to nuclei, and this was in fact observed: the GUS fusion proteins were predominantly seen in the nuclei of onion and *Arabidopsis* cells, regardless of the presence or absence of the signal receiver domain (Figure 2). Several different growth conditions for onion cells, including growth with or without exogenous addition of cytokinin and in the light or dark, had no effect on nuclear localization of ARR1 and ARR2 (unpublished results). These observations strongly suggest that the native ARR1 and ARR2 molecules are constantly located in nuclei. However, the possibility that the signal transduction processes in which ARR1 and ARR2 are involved proceed continuously under the growth conditions used cannot be completely ruled out.

Although the actual signal sequence by which ARR1 and ARR2 localize to nuclei was not determined, our previously argued VRK(R/K)R sequence, present just downstream of the signal receiver domains (aa 153–157 of ARR1, aa 144–148 of ARR2), presumably contributes at least in part, as there are precedents for these sequences actually promoting nuclear localization (Howard *et al.*, 1992; Robbins *et al.*, 1991). Two other probable nuclear localization signals, SRKRK and KKPRV, are also present closely within the acidic and ARRM domains, respectively (aa 213–217 and aa 236–240 of ARR1; aa 193–197 and aa 215–219 of ARR2). These sequences may be functionally redundant in terms of the nuclear targeting of ARR1 and ARR2.

It is well known that glutamine-rich and acidic characteristics are frequently associated with the transactivation domain of eukaryotic transcription factors because they can provide an interface for protein–protein interactions (Triezenberg, 1995). The glutamine-rich domains of ARR1 and ARR2 were no exception. When fused to the DNA-binding domain of yeast GAL4, they activated expression of the *LUC* reporter gene preceded by the GAL4 target sequence (Figure 3). The levels of activation signal given by the ARR1 and ARR2 glutamine-rich domains were about one-third and one-half, respectively, of that seen with the strong transactivation domain VP16, suggesting that their potential for transactivation is high. In contrast, the acidic domains of ARR1 and ARR2 showed almost no transactivation ability; instead, they effected a slight inhibition of transactivation exerted by the glutamine-rich domains. The ARRM domains also lowered transactivation. This reduction might result from titration of the effector by endogenous DNA capable of interacting with the ARRM domain. It was thus concluded that the glutamine-rich domains act as the principal transactivation domains of ARR1 and ARR2. ARR11, belonging to the same group as ARR1 and ARR2, also has the ability to activate transcription, at least in yeast cells, although its C-terminal region has no obvious similarity to those of ARR1 and ARR2 (Lohrmann *et al.*, 1999). The transactivation ability may be commonly associated with this group of plant response regulators.

It was previously argued that the ARRM domain weakly resembles the DNA-binding domain of the mammalian oncogene product Myb, and homology searches using the ARRM domain frequently indicate similarities with sequences in plant DNA databases (Sakai *et al.*, 1998). It is thus reasonable to assume that this domain is utilized as a functional unit for DNA binding in a variety of proteins. The gel-retardation patterns presented in Figure 4 clearly indicate that the ARRM domain has the ability to bind double-stranded DNA in a sequence-specific manner. The best target sequence among the oligonucleotide species tested was identified as 5′-AGATT-3′ in the context used. The actual DNA recognition by the ARRM domain appeared to occur mainly at the last four positions of the 5 bp core sequence. As the specificity determined by a 4 bp sequence does not seem high enough for choosing target genes on the *Arabidopsis* genome, it is possible that an additional protein component(s) interacting with ARR1 and ARR2; phosphorylation of their Asp residues; the context of this 4 bp sequence; or a combination of these factors contributes towards increasing the specificity of interactions between the *cis* and *trans* factors. As we obtained essentially the same gel-retardation patterns with both the crude extracts and the purified GST::ARRM fusion proteins, no other factor is likely to have been involved in the observed specific DNA binding.

The DNA-binding domains of conventional plant Myb homologues are composed of two or three repeated structures, each of which contains three landmark Trp residues (Kranz *et al.*, 2000; Martin and Paz-Ares, 1997; Meshi and Iwabuchi, 1995; Williams and Grotewold, 1997). However, the ARRM domain contains only one repeat, and two out of the three Trp residues are not conserved. The DNA-binding domain of potato MybSt1 transcription factor (Baranowskij *et al.*, 1994) and those of both *Arabidopsis* LHY and CCA1 proteins, whose mutants are defective in circadian rhythm (Schaffer *et al.*, 1998; Wang *et al.*, 1997), appear to belong to the same class as the ARRM domain. They have a similar one-repeat architecture, with amino acid sequences considerably diverged from the mammalian Myb, though LHY and CCA1 retain two out of the three Trp residues. Although these single-repeat Myb homologues were previously thought to be exceptional cases,

the existence of many additional ARRM-like domains in plant species, found by *in silico* screening (Sakai *et al.*, 1998), implies that these Myb homologues constitute a large family of plant DNA-binding proteins that perform a variety of cellular functions. Furthermore, these ARRM-like domains may have similar sequence specificities for DNA binding, because DNA bound to MybSt1 and CCA1 contains 5′-GGATA-3′ and 5′-AGATT-3′ sequences, respectively, which are highly similar or identical to the target sequences of ARR1 and ARR2. It is naturally expected that other ARR members containing the ARRM-like domain, such as ARR11 and ARR10, are also able to bind DNA with similar sequences.

Domain dissection has revealed that ARR1 and ARR2 are comparable in their abilities to localize in nuclei, bind DNA and activate transcription. Therefore both ARR1 and ARR2 are expected to function as transcriptional activators individually, though their intracellular roles may be overlapping. The role of transcriptional activator was confirmed using the *LUC* reporter gene whose promoter contained six copies of the ARRM binding sequence. The full-length versions of ARR1 and ARR2 actually showed this ability in transactivation. In addition to the experiments shown in Figure 5, we repeated similar experiments several times and consistently obtained comparable results (unpublished data). The average levels of transactivation by ARR1 and ARR2 deduced from all these experiments were 2.2 and 2.9 times above background, respectively. Therefore the signal levels exerted by ARR1 and ARR2 appear to be significant, though low. The truncated derivatives without the signal receiver domain gave much higher transactivation signals than those generated by the respective intact version. Therefore the N-terminal signal receiver domain appeared to repress DNA-binding and/or transactivation ability. This repressive role may be modulated by Asp phosphorylation, and the truncated versions might mimic a molecular state of ARR1 and ARR2 that results in transactivation.

This is the first report demonstrating that plant two-component response regulators work independently as transcriptional activators similar to the majority of bacterial response regulators. By way of analogy with the bacterial two-component regulatory system, ARR1 and ARR2 are likely to receive the phospho-relay signal from an as yet unidentified cognate histidine kinase, either directly or through a bridge component(s), triggered by a specific environmental stimulus. On the other hand, the target genes of which ARR1 and ARR2 regulate expression are present downstream of the signal cascade. Thus identification of their target genes will shed light on the kind of signal responses in plant cells that ARR1 and ARR2 are involved in.

Experimental procedures

Recombinant DNA techniques

Standard recombinant DNA techniques have been described previously (Sakai *et al.*, 1998; Sambrook *et al.*, 1989). PCR was carried out according to Saiki *et al.* (1988). In-frame cloning, in which two or more DNA segments were connected to code for a fusion protein, was carried out using DNA fragments prepared by PCR with primers of appropriate sequences.

Plant materials and growth conditions

Arabidopsis thaliana (Columbia ecotype) was grown at 22°C under constant light in soil or on agar plates containing Murashige & Skoog (MS) medium (Murashige and Skoog, 1962; Sigma M-5524, Sigma-Aldrich, Japan) supplemented with 1% sucrose and 0.8% agar. *Nicotiana tabacum* L. cv. SR-1 was grown at 22°C in soil under constant light. Transgenic *Arabidopsis* plants were constructed by vacuum infiltration (Bechtold *et al.*, 1993).

Intracellular localization of proteins

DNA fragments coding for entire or partial regions of ARR1 and ARR2 were inserted between the *Bam*HI and *Sma*I sites of pBI221 (Clontech, Palo Alto, CA, USA) so as to be fused in-frame to the N-terminal end of GUS, downstream of the 35S promoter and upstream of the *Agrobacterium nos* gene terminator. The resulting recombinant DNA (2.5 μg) was used to coat 1.5 mg of 1 μm gold particles, and was introduced into onion epidermal cells with a particle-delivery system (Bio-Rad PDS-1000/He, Biorad Laboratories, Hercules, CA, USA). The onion epidermal cells had previously been incubated on MS agar plates in the light at 22°C for 24 h. The initial pressure of bombardment was 1100 psi, and the travelling distance of the particles to the plant tissues was 6 cm. Bombarded tissues were placed on the same agar plates and incubated at 22°C for 24 h in the light, followed by immersion in a histochemical substrate solution containing 1 mM 5-bromo-4-chloro-3-indolyl-β-D-glucuronid (Sigma B-0522); 50 mM phosphate buffer pH 7.0; 0.5 mM potassium ferricyanide; and 0.5 mM potassium ferrocyanide. Histochemical observations of stained tissues were made using a Carl Zeiss Axiophot 2E microscope. For transgenic *Arabidopsis* plants expressing GUS, leaves were similarly stained and trichome cells were microscopically observed.

Measurement of transcriptional activation

The ability to activate transcription was monitored by the dual-luciferase reporter assay system (Promega Corporation, Madison, WI, USA) according to the instructions provided by the supplier. For transcriptional activation through the GAL4 target sequence as the *cis* factor, the reporter plasmid used was p6GAL4-LUC, a derivative of pGEM-3Z (Promega). It contained the firefly luciferase gene *LUC* (de Wet *et al.*, 1985) preceded by a promoter containing the GAL4 target sequence ($6 \times UAS_G$) and the –46 to +1 region of the cauliflower mosaic virus 35S promoter. $6 \times UAS_G$ was a DNA segment containing six copies of the sequence 5′-CGGGTGACAGCCCTCCG-3′; the truncated 35S promoter (–46 to +1) was the source of the TATA sequence (Aoyama *et al.*, 1995). The control reporter plasmid was pRL carrying the sea pansy luciferase gene *RLUC* directed by the 35S promoter

(Lorenz *et al.*, 1991). Effector plasmids carried a cDNA fragment coding for various portions of ARR1 and ARR2 between the *Xho*I and *Bam*HI sites on the pGAL4 vector (a pUC19 derivative), directing a protein fused to the yeast GAL4 DNA-binding domain (nucleotide positions 1–353) under the control of the 35S promoter. The ARR1 and ARR2 portions inserted in the effectors were (Figures 1 and 3a): ARR1Δ(DDK) (aa 153–669); ARR1Δ(DDK-M) (aa 153–235, aa 300–669); ARR1Δ(DDK-M-Q) (aa 153–235); ARR1Δ(DDK-Ac-M) (aa 300–669); ARR2Δ(DDK) (aa 144–664); ARR2Δ(DDK-M) (aa 144–214, aa 279–664); ARR2Δ(DDK-M-Q) (aa 144–214); and ARR2Δ(DDK-Ac-M) (aa 279–664). All the protein-coding sequences in these plasmids were followed by the poly(A) addition sequence of the pea ribulose-1,5-bisphosphate carboxylase small subunit *rbcS-3A* gene (Fluhr *et al.*, 1986). p6GAL4-LUC, pRL and one of the effector plasmids were simultaneously bombarded into plant cells and incubated as in the previous section. The plant material used was either mature *N. tabacum* leaves (about 15 cm in length) that had been cut from a plant just before use, or onion epidermal cells that had previously been incubated on MS agar plates in the light at 22°C for 24 h. The travelling distance of gold particles was 9 cm for tobacco and 6 cm for onion. After incubation at 22°C for 24 h, bombarded tissues were ground, and the LUC and RLUC luminescent signals were separately measured by a luminometer (model TD-20e Turner). The LUC luminescent signal indicates transactivation by an effector, whereas the RLUC luminescent signal shows the efficiency of the DNA delivery system as an internal control. The normalized transactivation ability was expressed by the ratio of the two signal intensities.

Transactivation through the ARRM target sequence as the *cis* factor was examined in the same way, except for the following changes. The reporter plasmid used was p6ARRM-LUC in which 6 × UAS_G of p6GAL4-LUC was replaced by either 6 × UAS_M or 6 × UAS_{Md}. The 6 × UAS_M and 6 × UAS_{Md} were DNA segments containing six copies of 5′-TCTAGGATTGTCT-3′ and 5′-TCTAGCATTGTCT-3′, respectively. The effector plasmids contained the truncated 35S promoter (–46 to +1)-directed DNA sequence coding for either full-length ARR1 or ARR2, or their deletion versions lacking for the N-terminal signal receiver domains.

Preparation of E. coli *protein extracts*

The DNA segments (Figure 1) corresponding to ARRM (aa 236–299 for ARR1, aa 215–278 for ARR2) were separately inserted between the *Bam*HI and *Eco*RI sites on pGEX-2T (Amersham Pharmacia Biotech). The resulting chimeric genes were expressed in *E. coli* BL21(DE3) cells upon induction by isopropyl-β-D-thiogalactopyranoside, as previously described (Endoh and Oka, 1993). Cells were pelleted and resuspended in binding buffer (20 mM HEPES pH 7.4, 100 mM KCl, 2 mM $MgCl_2$, 0.5 mM EDTA, 1 mM dithiothreitol, 0.2% NP-40, 10% glycerol, 1 mM phenylmethylsulfonyl fluoride, 1 μg ml^{-1} each of aprotinin, leupeptin hemisulfate monohydrate and pepstatin A) of one-tenth the culture volume. Cells were lysed on ice by mild sonication, then subjected to centrifugation at 10 000 *g* for 30 min at 4°C. The supernatant was aliquoted, quickly frozen on liquid nitrogen and stored at –80°C until use. For some experiments we purified fusion proteins by treating the crude extracts with glutathione Sepharose, and samples of more than 90% purity were obtained. Protein concentrations were estimated by staining with Coomassie brilliant blue R-250, after resolving proteins through SDS–polyacrylamide gel electrophoresis (Endoh and Oka, 1993).

Binding site selection by PCR

Binding site selection experiments were carried out as previously described (Sessa *et al.*, 1993). Briefly, an *E. coli* crude extract (about 1 μg protein) containing a GST::ARRM fusion protein was mixed with synthetic double-stranded oligonucleotides with random 19 bp sequences flanked by 18 bp of tag sequences (about 8 ng) in 50 μl binding buffer supplemented with 2 μg poly(dI-dC) (Pharmacia). DNA–protein complexes were trapped on glutathione Sepharose, from which oligonucleotides were isolated and then amplified by PCR. After repeating this cycle four times, oligonucleotides were cloned in pHSG397 (Takara, Kyoto, Japan) and sequenced.

DNA-binding assays

The ability of proteins to bind double-stranded DNA was examined by gel-retardation assays (Sessa *et al.*, 1993). The fragments used were 5′-CGACGTGAATTCTAGGATTGTCTCGCA-TACACTG-3′ and various 1 bp substitution derivatives within the 5 bp core sequence (underlined bases, see text). ^{32}P-labelled DNA fragments (about 2 ng, 1 × 10^5 cpm) were incubated for 30 min at 20°C with an *E. coli* crude extract (about 50 ng proteins) containing a fusion protein or a purified fusion protein sample (about 50 ng) in 20 μl binding buffer supplemented with 2 μg poly(dI-dC). After the incubation step the mixture was immediately loaded on 4% polyacrylamide gels (acrylamide:bis-acrylamide, 29 : 1). Electrophoresis was carried out in 0.25 × TBE (1 × TBE = 89 mM Tris-borate, 2 mM EDTA pH 8.3). Dried gels were then subjected to autoradiography with a Fujix BAS2000 Bio-Image Analyser (Fuji Photo Film).

Acknowledgements

This work was supported in part by the Special Coordination Fund of the Science and Technology Agency of the Japanese Government to A.O., and by a grant from the Bio-oriented Technology Research Advancement Institution to T.A.

References

Aoyama. T., Dong, C.-H., Wu, Y., Carabelli, M., Sessa, G., Ruberti, I., Morelli, G. and Chua, N.-H. (1995) Ectopic expression of the *Arabidopsis* transcriptional activator Athb-1 alters leaf cell fate in tobacco. *Plant Cell*, **7**, 1773–1785.

Appleby, J.L., Parkinson, J.S. and Bourret, R.B. (1996) Signal transduction via the multi-step phosphorelay: not necessarily a road less traveled. *Cell*, **86**, 845–848.

Baranowskij, N., Frohberg, C., Prat, S. and Willmitzer, L. (1994) A novel DNA binding protein with homology to Myb oncoproteins containing only one repeat can function as a transcriptional activator. *EMBO J.* **13**, 5383–5392.

Bechtold, N., Ellis, J. and Pelletier, G. (1993) *In planta Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis* plants. *C.R. Acad. Sci. Paris. Life Sci.* **316**, 1194–1199.

Brandstatter, I. and Kieber, J.J. (1998) Two genes with similarity to bacterial response regulators are rapidly and specifically induced by cytokinin in Arabidopsis. *Plant Cell*, **10**, 1009–1019.

Chang, C., Kwok, S.F., Bleecker, A.B. and Meyerowitz, E.M. (1993) *Arabidopsis* ethylene-response gene *ETR1*: similarity of product to two-component regulators. *Science*, **262**, 539–544.

de Wet, J.R., Wood, K.V., Helinski, D.R. and DeLuca, M. (1985)

Cloning of firefly luciferase cDNA and the expression of active luciferase in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **82**, 7870–7873.

Endoh, H. and Oka, A. (1993) Functional analysis of the VirG-like domain contained in the *Agrobacterium* VirA protein that senses plant factors. *Plant Cell Physiol*. **34**, 227–235.

Fluhr, R., Moses, P., Morelli, G., Coruzzi, G. and Chua, N.-H. (1986) Expression dynamics of the pea *rbcS* multigene family and organ distribution of the transcripts. *EMBO J.* **5**, 2063–2071.

Howard, E.A., Zupan, J.R., Citovsky, V. and Zambryski, P.C. (1992) The VirD2 protein of *A. tumefaciens* contains a C-terminal bipartite nuclear localization signal: implications for nuclear uptake of DNA in plant cells. *Cell*, **68**, 109–118.

Hua, J., Chang, C., Sun, Q. and Meyerowitz, E.M. (1995) Ethylene insensitivity conferred by *Arabidopsis ERS* gene. *Science*, **269**, 1712–1714.

Imamura, A., Hanaki, N., Umeda, H., Nakamura, A., Suzuki, T., Ueguchi, C. and Mizuno, T. (1998) Response regulators implicated in His-to-Asp phosphotransfer signaling in *Arabidopsis*. *Proc. Natl Acad. Sci. USA*, **95**, 2691–2696.

Imamura, A., Hanaki, N., Nakamura, A., Suzuki, T., Taniguchi, M., Kiba, T., Ueguchi, C., Sugiyama, T. and Mizuno, T. (1999) Compilation and characterization of *Arabidopsis thaliana* response regulators implicated in His–Asp phosphorelay signal transduction. *Plant Cell Physiol.* **40**, 733–742.

Kakimoto, T. (1996) *CKI1*, a histidine kinase homolog implicated in cytokinin signal transduction. *Science*, **274**, 982–985.

Kranz, H., Scholz, K. and Weisshaar, B. (2000) *c-Myb* oncogene-like genes encoding three MYB repeats occur in all major plant lineages. *Plant J.* **21**, 231–235.

Lohrmann, J., Buchholz, G., Keitel, C., Sweere, U., Kircher, S., Bäurle, I., Kudla, J., Schäfer, E. and Harter, K. (1999) Differential expression and nuclear localization of response regulator-like proteins from *Arabidopsis thaliana*. *Plant Biol.* **1**, 495–505.

Lorenz, W.W., McCann, R.O., Longiaru, M. and Cormier, M.J. (1991) Isolation and expression of a cDNA encoding *Renilla reniformis* luciferase. *Proc. Natl Acad. Sci. USA*, **88**, 4438–4442.

Martin, C. and Paz-Ares, J. (1997) MYB transcription factors in plants. *Trends Genet.* **13**, 67–73.

Meshi, T. and Iwabuchi, M. (1995) Plant transcription factors. *Plant Cell Physiol.* **36**, 1405–1420.

Murashige, T. and Skoog, F. (1962) A revised medium for rapid growth and bioassays with tobacco tissue. *Physiol. Plant*, **15**, 493–497.

Robbins, J., Dilworth, S.M., Laskey, R.A. and Dingwall, C. (1991) Two independent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. *Cell*, **64**, 615–623.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, **239**, 487–491.

Sakai, H., Aoyama, T., Bono, H. and Oka, A. (1998) Two-component response regulators from *Arabidopsis thaliana* contain a putative DNA-binding motif. *Plant Cell Physiol.* **39**, 1232–1239.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Sanford, J.C., Smith, F.D. and Russell, J.A. (1993) Optimizing the biolistic process for different biological applications. *Meth. Enzymol.* **217**, 483–509.

Schaffer, R., Ramsay, N., Samach, A., Corden, S., Putterill, J., Carre, I.A. and Coupland, G. (1998) The *late elongated hypocotyl* mutation of *Arabidopsis* disrupts circadian rhythms and the photoperiodic control of flowering. *Cell*, **93**, 1219–1229.

Schneider-Poetsch, H.A., Braun, B., Marx, S. and Schaumburg, A. (1991) Phytochromes and bacterial sensor proteins are related by structural and functional homologies. Hypothesis on phytochrome-mediated signal-transduction. *FEBS Lett.* **281**, 245–249.

Sessa, G., Morelli, G. and Ruberti, I. (1993) The Athb-1 and Athb-2 HD-Zip domains homodimerize forming complexes of different DNA binding specificities. *EMBO J.* **12**, 3507–3517.

Stock, J.B., Ninfa, A.D. and Stock, A.M. (1989) Protein phosphorylation and regulation of adaptive response in bacteria. *Microbiol. Rev.* **53**, 450–490.

Stock, J.B., Stock, A.M. and Mottonen, J.M. (1990) Signal transduction in bacteria. *Nature*, **344**, 395–400.

Taniguchi, M., Kiba, T., Sakakibara, H., Ueguchi, C., Mizuno, T. and Sugiyama, T. (1998) Expression of *Arabidopsis* response regulator homologs is induced by cytokinins and nitrate. *FEBS Lett.* **429**, 259–262.

Triezenberg, S.J. (1995) Structure and function of transcriptional activation domains. *Curr. Opin. Genet. Dev.* **5**, 190–196.

Urao, T., Yakubov, B., Yamaguchi-Shinozaki, K. and Shinozaki, K. (1998) Stress-responsive expression of genes for two-component response regulator-like proteins in *Arabidopsis thaliana*. *FEBS Lett.* **427**, 175–178.

Wang, Z.-Y., Kenigsbuch, D., Sun, L., Harel, E., Ong, M.S. and Tobin, E.M. (1997) A Myb-related transcription factor is involved in the phytochrome regulation of an *Arabidopsis Lhcb* gene. *Plant Cell*, **9**, 491–507.

Williams, C.E. and Grotewold, E. (1997) Differences between plant and animal Myb domains are fundamental for DNA binding activity, and chimeric Myb domains have novel DNA binding specificities. *J. Biol. Chem.* **272**, 562–571.

Wurgler-Murphy, S.M. and Saito, H. (1997) Two-component signal transducers and MAPK cascades. *Trends Biochem. Sci.* **22**, 172–176.