Structural Basis for DNA Binding Specificity by the Auxin-Dependent ARF Transcription Factors

D. Roeland Boer,1,2,6 Alejandra Freire-Rios,3,6 Willy A.M. van den Berg,3,6 Terrens Saaki,3 Iain W. Manfield,4 Stefan Kepinski,1 Irene López-Vidriero,5 Jose Manuel Franco-Zorrilla,5 Sacco C. de Vries,3 Roberto Solano,5 Dolf Weijers,3,7,* and Miquel Coll1,2,7,*

1Institute for Research in Biomedicine (IRB Barcelona), Baldiri Reixac 10-12, 08028 Barcelona, Spain
2Institut de Biologia Molecular de Barcelona (IBMB-CSIC), Baldiri Reixac 10-12, 08028 Barcelona, Spain
3Laboratory of Biochemistry, Wageningen University, Dreijenlaan 3, 6703 HA Wageningen, the Netherlands
4Astbury Centre for Structural Molecular Biology (IWM) and Centre for Plant Sciences (SK), Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK
5Genomics Unit and Department of Plant Molecular Genetics, Centro Nacional de Biotecnología-Consejo Superior de Investigaciones Científicas, Campus Universidad Autónoma, 28049 Madrid, Spain
6These authors contributed equally to this work
7These authors contributed equally to this work
*Correspondence: dolf.weijers@wur.nl (D.W.), miquel.coll@irbbarcelona.org (M.C.)

http://dx.doi.org/10.1016/j.cell.2013.12.027

SUMMARY

Auxin regulates numerous plant developmental processes by controlling gene expression via a family of functionally distinct DNA-binding auxin response factors (ARFs), yet the mechanistic basis for generating specificity in auxin response is unknown. Here, we address this question by solving high-resolution crystal structures of the pivotal Arabidopsis developmental regulator ARF5/MONOPTEROS (MP), its divergent paralog ARF1, and a complex of ARF1 and a generic auxin response DNA element (AuxRE). We show that ARF DNA-binding domains also homodimerize to generate cooperative DNA binding, which is critical for in vivo ARF5/MP function. Strikingly, DNA-contacting residues are conserved between ARFs, and we discover that monomers have the same intrinsic specificity. ARF1 and ARF5 homodimers, however, differ in spacing tolerated between binding sites. Our data identify the DNA-binding domain as an ARF dimerization domain, suggest that ARF dimers bind complex sites as molecular calipers with ARF-specific spacing preference, and provide an atomic-scale mechanistic model for specificity in auxin response.

INTRODUCTION

The plant hormone auxin controls numerous growth and developmental processes and is a key determinant of plant architecture (Vanneste and Friml, 2009). Physiological approaches in the early 20th century have led to the identification of indole-3-acetic acid as the main natural auxin (Thimann and Koepfli, 1935). In the past decades, genetic studies have revealed mechanisms of hormone biosynthesis (Ljung, 2013), transport (Grunewald and Friml, 2010), and response (Chapman and Estelle, 2009). The cellular response to auxin involves ubiquitin-proteasome-dependent degradation of Auxin/Indole Acetic Acid (Aux/IAA) proteins, transcriptional corepressors (Szemenyi et al., 2008) that act by binding auxin response factors (ARFs) (Tiwari et al., 2003). The latter are DNA-binding transcription factors that control the expression of the large set of auxin-dependent genes that mediate hormone-dependent growth and development (Guilfoyle and Hagen, 2007). A central, yet unanswered question in auxin biology is how the simple tryptophan-like indole-3-acetic acid can trigger a wide variety of cellular responses. As the last step in auxin signaling prior to gene regulation, the ARF transcription factors are likely components to confer specificity to auxin response through selection of target genes. Consistent with a role in response diversification, the ARF family consists of 23 members in Arabidopsis thaliana and contains >10 members even in the moss Physcomitrella patens (Finet et al., 2013). ARF genes are expressed in dynamic and different patterns during development (Radermacher et al., 2011), and genetic studies have shown that individual ARFs control distinct developmental processes (Guilfoyle and Hagen, 2007; Radermacher et al., 2011, 2012; Weijers et al., 2005). For example, ARF5 (also named MONOPTEROS [MP] but for consistency referred to as ARF5 here) is critically required for several developmental auxin responses, including embryonic root and flower formation (Berleth and Jürgens, 1993; Przemek et al., 1996), whereas ARF1 and 2 control senescence and floral organ abscission (Ellis et al., 2005). These differences between ARFs are at least in part due to differences in protein sequence, as misexpression and promoter-swap studies demonstrated that ARF proteins are not equal (Radermacher et al., 2012; Weijers et al., 2005). Instead, double-mutant analysis suggested that ARF1 and ARF5 act
antagonistically (Rademacher et al., 2011). ARFs are modular transcription factors, consisting of several domains that have remained conserved despite hundreds of millions of years of evolution (Finet et al., 2013). At their N terminus, all ARFs have a DNA-binding domain (DBD), followed by a middle region (MR) that determines whether ARFs activate or repress target genes (Tiwari et al., 2003) and a C-terminal interaction domain (domain III/IV). The latter has been shown to mediate interactions between ARFs and their Aux/IAA inhibitors, as well as between ARFs (e.g., Kim et al., 1997). Several lines of evidence suggest that ARF domains are functionally autonomous, i.e., they act in isolation. First, both the DBD and the C-terminal interaction domains are found in other protein families. The DBD harbors a B3 DNA binding motif that is also found in many other plant transcription factors (Swaminathan et al., 2008). Similarly, the C-terminal interaction domain III/IV is also found in the interacting Aux/IAA proteins (Tiwari et al., 2003; Kim et al., 1997).

Second, transient expression assays and domain swaps have demonstrated that each of the three domains can act in isolation (Tiwari et al., 2003).

Here, we address the atomic basis for sequence-specific DNA binding by ARF transcription factors and explore mechanisms by which variation in the ARF DBD selects different target genes. Most ARFs tested have been shown to bind a generic auxin response element (AuxRE; Ulmasov et al., 1999) that was identified based on its occurrence in auxin-dependent promoters (Ulmasov et al., 1995). However, because target sites have not been screened exhaustively, it is unknown whether different ARFs prefer distinct binding sites and, if so, what the molecular basis is for such differences. Here, we have determined high-resolution crystal structures of the DNA-binding domains of two divergent ARFs, as well as an ARF-DNA complex. Structure-function analysis and saturating binding site selection lead to a redefined ARF binding motif, as well as a DNA-binding mechanism in which dimerization of ARF DNA binding domains generates cooperative binding to adjacent sites where spaceing determines ARF binding affinity. Our study provides an atomic-level explanation for DNA-binding specificity in the auxin pathway.

RESULTS

Crystal Structures of ARF DNA-Binding Domains

All ARFs carry a conserved DBD at their N terminus (Figure 1A). This domain is often followed by an MR that directs transcriptional changes and a C-terminal domain (III/IV) that mediates protein-protein interactions (Tiwari et al., 2003). It is well established that domain III/IV is essential for the heterotypic ARF-Aux/IAA interactions that render ARF activity auxin dependent (Figure 1A; Tiwari et al., 2003). The same domain has been proposed to mediate ARF-ARF interactions (Kim et al., 1997), but whether this is biologically meaningful has not been established. The ARF DBD is sufficient for binding auxin-responsive promoters (Tiwari et al., 2003). Its B3 subdomain is found in other transcription factors (Figure 1A; Swaminathan et al., 2008) and was shown to bind DNA in RAV1 (Yamasaki et al., 2004), which suggests domain modularity. Interestingly, phylogenetic trees based on sequence alignments of only the DBD strongly resemble those derived from entire ARF proteins (Figure S1 available online), raising the possibility that variations in this domain contribute to the distinct properties of ARFs.

To gain insight into the mechanism of DNA binding by ARFs, we expressed and purified the DBD of ARF1 and ARF5/MP. These two ARFs are phylogenetically distant (Figure S1), and their divergence occurred early in land plant evolution hundreds of millions of years ago (Finet et al., 2013). Based on the resulting structure, ARF1-DBD is divided into two domains (domain I/II, see below). Thus, the B3 domain appears to be an insertion in the DD. Structure similarity searches identified the Tudor domain of the RAV1 (Yamasaki et al., 2004) and At1g16640 (Waltner et al., 2005). Remarkably, the B3 domain in ARF1-DBD and ARF5-DBD is embedded in a larger fold context. The regions N- and C-terminal to the B3 domain together form a single second domain (Figures 1B and 1C) and constitutes a dimerization domain (DD, see below). Thus, the B3 domain appears to be an insertion in the DD. Structure similarity searches identified the Tudor domain of the human PHD-finger protein 20 (PDB entry 3QII), but the hydrophobic cage that recognizes methylated lysine residues (Adams-Cioaba et al., 2012; Huang et al., 2006) is missing from the ARF-DDBD-AD.

Finally, the last 80 C-terminal residues form a third separate ancillary domain (AD) that tightly interacts with the DD. The AD folds in a small five-stranded β-barrel-like structure. Structural similarity searches identified the Tudor domain of the human PHD-finger protein 20 (PDB entry 3QII), but the hydrophobic cage that recognizes methylated lysine residues (Adams-Cioaba et al., 2012; Huang et al., 2006) is missing from the ARF-DDBD-AD.

ARF Dimerization through the DBD

Strikingly, in all crystal structures, ARF-DBDs homodimerized through their DD (Figures 1D and 1F). The dimer interface contacts include hydrophobic interactions between several highly conserved residues (Figures S3D and S3E), which indicates that this is most probably a physiologically meaningful interaction. It is stabilized by a network of hydrogen bonds, some mediated by water molecules (Figure 1E). Interestingly, α helix 6 (x6) of both monomers is juxtaposed and centered at a
conserved (Figures S3D and S3E) glycine residue (G245 in ARF1-DBD and G279 in ARF5-DBD). Other residues of this helix (A248, T249, and A253; in ARF5: A282, A283, and A287) engage in hydrophobic interactions (Figures 1E and 1F), whereas the P233–S238 loop (ARF5: P267–S272) fits into a groove of the opposite monomer and involves interactions between S235 (ARF5: S269) on one monomer and K265 and E85 (ARF5: N299 and D118) on the other (Figures 1E, S3D, and S3E).

To address whether dimerization is induced by the crystallization conditions or whether this also occurs in solution, small-angle X-ray scattering (SAXS) was used for ARF1-DBD. Although neither monomer nor dimer models explained the scattering data, a monomer:dimer equilibrium improved the fit dramatically (Figure 2A). Hence, homodimerization also occurs in solution, and given that both ARF1 and ARF5 dimerize, this is likely a general property of ARF-DBDs.

To next determine whether ARF-DBD homodimerization is required for biological function, we mutated several amino acids in the dimerization interface of ARFS/MP (Table 1) and tested the ability of mutant proteins to replace the wild-type protein in vivo. The arf5/mp mutant is unable to establish an embryonic root and, as a consequence, forms rootless seedlings (Hardtke and Berleth, 1998). Adventitious roots can, however, be induced postembryonically, and mutant plants have distinctive growth defects, including aberrant flowers or even naked, pin-like inflorescences (Przemeck et al., 1996). Importantly, although S269N, G279A and N299S mutations did not impair ARF5 activity during embryonic root formation (Table 1), G279E, G279I, A282N, and A287N mutations all compromised ARF5 function in vivo (Table 1 and Figure 2B). In some cases, these mutated ARF5/MP proteins even induced dominant-negative defects in wild-type plants (Figure 2C). To ascertain that the failure of these mutant proteins to complement the arf5/mp mutant is due to alterations in dimerization properties, rather than abnormal folding behavior, secondary structures were determined using circular dichroism (CD) spectroscopy. Consistent with the solubility of purified...
mutant proteins, none showed deviations in its CD spectrum (Figure S4), suggesting that all fold normally. Hence, this analysis shows that amino acids at the dimerization interface, in particular in the α-6 helix (G279, A282, and A287), are required for ARF5/MP function in vivo.

To address whether these mutations indeed interfere with homodimerization in the context of a full-length ARF protein that also carries the C-terminal interaction domain (III/IV), we employed a fluorescence resonance energy transfer (FRET)-based interaction assay. Here, interactions between CFP- and YFP-tagged ARF5 are quantified in mesophyll protoplasts (Figure 2D; Russinova et al., 2004), and we have previously used this assay to demonstrate ARF-Aux/IAA interactions in vivo (Radamacher et al., 2012). In this assay, wild-type ARF5 showed clear homodimerization as measured by a decrease in the average lifetime of the ARF5-CFP donor (expressed as FRET efficiency; Figure 2E). As expected from the position of the glycine in the dimerization interface, G279A, G279I, and G279E mutations significantly decreased the FRET efficiency and hence impair dimerization (Figure 2E). To determine the relative contribution of the DBD and domain III/IV in homodimerization, a truncated ARF5 protein was generated in which domain III/IV was deleted (Weijers et al., 2006; Krogan et al., 2012). Homodimerization still occurred, albeit at lower efficiency (Figure 2E). Similarly, the ARF3 protein, which naturally lacks domain III/IV (Ulmasov et al., 1999), was also able to homodimerize (Figure 2E; compare
with ARF3-CFP/free YFP control. Hence, the DBD is sufficient for dimerization in vivo, but interactions through domain III/IV may help to stabilize dimers. Indeed, when the G279I mutation was introduced in the truncated ARF5 protein lacking domains III/IV, FRET efficiency dropped to background levels (Figure 2E).

Collectively, these data show that ARF proteins form dimers through interactions between their DNA-binding domains and that this DBD dimerization is required for ARF function in vivo. Sequence alignments show that the amino acids at the dimerization surface are deeply conserved in the ARF family (Figures S3D and S3E), suggesting that this capacity is both widespread and ancient. The interaction surface is composed of many intermolecular interactions (Figures 1E and 1F), which suggests that the interaction may be robust. Indeed, several mutations in residues at the surface (S269N, G279A, and N299S) are tolerated in vivo (Table 1). To test whether such mutations indeed affect the interaction surface, we purified and crystallized the ARF1-G245A mutant (analogous to ARF5-G279A) and solved its structure to 2.3 Å resolution (Table S1). This showed an overall dimeric structure similar to wild-type ARF1-DBD (Figure S3C). As predicted, the structure showed a disturbed dimerization interface (Figure 2F) in which the 2-fold symmetry is broken and the 233–238 loop of one monomer could not bind the opposing monomer. In summary, the ARF interaction is robust and adaptative, as it tolerates a mutation that significantly affects the interaction surface. Yet, as more drastic mutations (G279E and G279I) in the same residue that measurably interfere with dimerization (Figure 2E) impair biological function (Table 1), we conclude that DBD dimerization is essential for in vivo function of ARF5.

**Mechanism of DNA Binding by ARFs**

Based on promoter analysis of an auxin-responsive gene in soybean, a canonical auxin-response element has been defined as TGTCTC (Ulmasov et al., 1995). ARF1 was first identified in a screen for factors binding this motif (Ulmasov et al., 1997). An inverted repeat of the same element, spaced by seven nucleotides (Figure 3A; ER7), was shown to be efficiently bound by ARF1 (Ulmasov et al., 1997). To determine the structural basis for DNA binding, we cocrystallized ARF1-DBD and a double-stranded ER7 oligonucleotide and solved its structure to 2.9 Å (Table S1 and Figure 3B). The DNA binding interface is located at the tips of the U-shaped dimer. The two B3 domains bind to the inverted AuxRE elements located at both extremes of the oligonucleotide, and the connecting DNA sequence bridges the gap between the B3 domains (Figure 3B). The DNA adopts a B-DNA conformation and is bent by 40°. The structures of apo-ARF1-DBD (without DNA) and DNA-bound ARF1-DBD are very similar, except that the B3 domains are rotated relative to the DDs by 25° (Figures 3C and 3D). As a similar conformational difference is seen between the different apo structures of ARF1 and ARF5 (Figures 1D and S3A), it appears that the B3 domain displays an intrinsic flexibility with respect to the DD and that DNA binding locks the protein into a conformation. Comparing apo- and DNA-bound structures shows that the N-terminal α1 helix functions as a pivot point on which the B3 domain is balanced, and the loops that connect the B3 domain to the DD run down on both sides of the helix (Figure 3C). Interestingly, these loops are mostly disordered in the structures, which indicates flexibility. Given the dimerization of the DBD and the binding of each TGTCTC element to one of the monomers, this structure now explains the efficient binding of ARF1 to an inverted repeat sequence, as well as the constraints of the spacing between repeats (Ulmasov et al., 1997).

The structures of apo-ARF1-DBD (without DNA) and DNA-bound ARF1-DBD are very similar, except that the B3 domains are rotated relative to the DDs by 25°. As a similar conformational difference is seen between the different apo structures of ARF1 and ARF5 (Figures 1D and S3A), it appears that the B3 domain displays an intrinsic flexibility with respect to the DD and that DNA binding locks the protein into a conformation. Comparing apo- and DNA-bound structures shows that the N-terminal α1 helix functions as a pivot point on which the B3 domain is balanced, and the loops that connect the B3 domain to the DD run down on both sides of the helix. Interestingly, these loops are mostly disordered in the structures, which indicates flexibility. Given the dimerization of the DBD and the binding of each TGTCTC element to one of the monomers, this structure now explains the efficient binding of ARF1 to an inverted repeat sequence, as well as the constraints of the spacing between repeats (Ulmasov et al., 1997).

Binding of two AuxRE sites by an ARF dimer suggests that DNA binding may be cooperative. To test whether this is the case, we used surface plasmon resonance (SPR) with immobilized oligonucleotides. Both ARF1-DBD and ARF5-DBD showed
Binding to ER7 (Figures 3E and 3F and Table S2). Binding was sequence specific, as mutating both AuxRE sites completely abrogated binding (Figures 3E and 3F and Table S2). Consistent with cooperative binding, mutating only one of the two AuxRE sites reduced the affinity more than 2-fold (Figures 3E, 3F, S3F, and S3G and Table S2). This cooperative binding behavior of wild-type ARFs predicts that mutations that disturb dimerization should affect DNA binding affinity. We tested this
prediction by performing SPR measurements using S269N, G279E, and A282N mutant proteins. Indeed, all three proteins showed a clear reduction in DNA binding affinity to the ER7 oligonucleotide (Figure 3G and Table S2). These findings show that dimerization of the ARF DBD generates cooperative DNA-binding behavior. As dimerization is important for in vivo function of ARF5, this suggests that cooperative DNA binding is essential for normal ARF function.

**Recognition and Specificity of DNA Binding**

We next addressed the structural basis for specific DNA binding. The B3 domain recognizes the DNA largely at the major groove of both TGTCTC elements (Figures 4A, 4B, 5C, and 5D). The B3 β barrel is positioned laterally to the DNA with the axis of the barrel almost parallel to the axis of the DNA double helix. Two adjacent β strands (β5 and β8) run over the major groove, parallel to the two sugar-phosphate backbones. The loops connecting these strands (R181–R186 and H136–G137; ARF5: R215–R220 and H170–G171), located on either side of the barrel, further penetrate the major groove and make interactions that contribute to specific DNA base recognition (Figure 4A). In addition to these base contacts, DNA binding involves interactions of the DNA backbone with residues S131, S140, T191, and S194 (Figure 4B; ARF5: S165, S174, T227, and S230). In summary, ARF1 binding to the canonical TGTCTC motif involves both base contacts and backbone interactions, and the specific contacts involve only the 5' bases on one strand, which explains why this part of the motif is critical for ARF DNA binding (Ulmasov et al., 1999).

To determine whether the residues that mediate DNA binding in the crystal structure are also required for DNA binding in solution and for ARF function in vivo, several residues were individually mutated to alanines. Even though CD analysis showed that these mutant proteins showed normal overall structure (Figure S4), SPR analysis of ARF5-H170A (ARF1: H136) and ARF5-R215A (ARF1:R181) proteins revealed that, in both cases, ER7 binding was significantly reduced (Figure 4C). We next analyzed DNA-binding specificity of proteins to sequence-specific DNA binding by performing PBM analysis on mutant proteins. Sequence-specific binding was lost when either P218 or R215 residues were mutated (Figure 5), whereas H170A and G171A mutations did not affect binding specificity. This analysis thus helps identify residues within the B3 domain that confer binding specificity (R215 and P218) and distinguish these from residues that contribute to DNA affinity (H170; Figure 4C). We next analyzed DNA-binding specificity of proteins impaired in dimerization. Neither ARF5-S269N nor ARF5-G279E altered the PBM binding profile (Figure 5). This shows that dimerization contributes to DNA binding affinity, but not to the specificity of DNA motif recognition.

**Motif Spacing Constrains Specific ARF Binding**

We found that ARF proteins are extremely conserved at both their dimerization interface and their DNA-contacting residues. As a consequence, both ARF1 and ARF5 dimerize and bind qualitatively similar sequences. A key unanswered question therefore remains how different genes can be selected by different ARFs. We noticed that the largest variation between ARF1 and ARF5 DBDs is in the loops that connect the B3 and DD domains (Figure 6A). Therefore, in addition to quantitative differences in binding of the two ARF DBDs to distinct sequence motifs, one could envisage differences in binding of ARF dimers to complex motifs with varying spacing between AuxRE sites. To test this hypothesis, we performed SPR experiments using ER7 oligonucleotides in which the spacing between the two inverted TGTCTC sites was changed to 5 (ER5), 6 (ER6), 8 (ER8), or 9 (ER9). Both ARF1-DBD and ARF5-DBD bound to ER8 with similar efficiency as ER7 binding (Figures 6B and 6C and Table S2). Affinity of ARF1-DBD to ER5, ER6, and ER9 was strongly reduced (Figure 6B and Table S2) to a level comparable to that of ER7 with one TGTCTC site mutated (Figure 3E and Table S2) or to that of a mutant impaired in dimerization (Figure 3G and Table S2). In contrast, ARF5-DBD retained significant binding to all ER versions, although binding efficiency to ER5, ER6, and ER9 was slightly reduced compared with ER7 and 8 (Figure 6C and Table S2). Hence, in addition to quantitative differences at the level of binding sites, ARF1 and ARF5 markedly differ in their ability to bind complex motifs depending on the spacing of the two binding sites. Such complex sites, with appropriate spacing, are indeed found in the promoters of direct and physiologically relevant ARF5/MP target genes (Figure 6D; Schlereth et al., 2010; Yamaguchi et al., 2013). In the case of LEAFY, mutating this site abrogated MP-dependent gene
Figure 4. Sequence-Specific DNA Binding by ARF-DBDs

(A) Detail of DNA-protein interface of the ARF1-DBD/ER7 complex showing the residues involved in DNA recognition. The two views show the same interaction surface, rotated by 180 degrees. Bases are colored and labeled in italics, and DNA-contacting residues are labeled.

(B) Scheme of intermolecular contacts between ARF1 protein and ER7 DNA bases (A,C,G,T) or backbone phosphates (p). Positively charged amino acids are marked in light blue, polar residues are marked in pink, and Proline is marked in yellow. Dashed lines indicate atomic interactions.

(C) SPR binding profiles of wild-type ARF5-DBD and H170A (ARF1: H136) and R215A (ARF1: R181) mutants on ER7 oligonucleotides. Scales are identical in all three panels.

(D) Conservation of amino acids between ARF1 and ARF5 mapped onto the ARF1-DBD/ER7 protein-DNA interface. Blue, identical; green, conserved; yellow, semi-conserved; red, nonconserved.

(legend continued on next page)
and between ARF and Aux/IAA protein families concluded that a systematic yeast two-hybrid-based interaction analysis among important for ARF dimerization upon DNA binding. Previously, it was demonstrated that low-affinity sites will outnumber high-affinity sites. Strained due to specific spacing requirements, our study predicted that dimer-bound high-affinity sites are more constrained. As dimer-bound high-affinity sites are more constrained due to specific spacing requirements, our study predicts that low-affinity sites will outnumber high-affinity sites.

Our results also show that domains III/IV are not likely to be important for ARF dimerization upon DNA binding. Previously, a systematic yeast two-hybrid-based interaction analysis among and between ARF and Aux/IAA protein families concluded that ARF homodimerization is limited (Vernoux et al., 2011). However, this study used only the previously known C-terminal domain III/IV. Our finding that the DBD represents a critical ARF dimerization domain calls for re-evaluation of this and other studies. Given the high degree of conservation of residues at the ARF-ARF interaction interface (Figure S3E), along with the notion that the phylogenetically diverse ARF1, ARF3, and ARF5 all homodimerize, it is likely that, if not all, ARFs homodimerize through their DBD. Whether or not ARFs can also heterodimerize and whether this would be biologically meaningful is another interesting open question.

The identification of the ARF DBD as a dimerization domain in addition to the previously known domain III/IV also has interesting implications for both DNA recognition and auxin regulation. First, given domain modularity, it is possible that DNA-bound ARF dimers interact through their domains III/IV with other ARF dimers to build higher-order complexes, analogous to what has been suggested for MADS-box transcription factors (Smaczniak et al., 2012). A testable prediction from such interactions would be that ARF complexes can bind more distantly spaced sites and induce DNA looping. Second, the ability of domains III/IV to mediate both homotypic (Aux/IAA or ARF-ARF) and heterotypic (ARF-Aux/IAA) interactions suggested that Aux/IAAs may obstruct ARF dimerization (reviewed in Lokerse and Weijers, 2009). The existence of an additional dimerization domain in the ARFs suggests a different mechanism of Aux/IAA function. Although Aux/IAAs could, in principle, modulate the stability of ARF-ARF dimers formed through their DBDs, an attractive alternative hypothesis is that they act as competitive inhibitors to prevent the formation of domain III/IV-mediated higher-order DNA-bound ARF complexes.

ARFs can have overlapping (Ellis et al., 2005; Nagpal et al., 2005; Okushima et al., 2005), (Rademacher et al., 2011), different (Rademacher et al., 2012; Rademacher et al., 2011; Weijers et al., 2005), or even opposing (Rademacher et al., 2011)
functions, and an important question is how these different activities are encoded in their structures. Often, in the Homeodo-
main family (Berger et al., 2008; Noyes et al., 2008), for example, variation in sequence-specific DNA binding in transcription
factor families is generated by substitutions in the DNA-contact-
ing residues. In contrast, intrinsic DNA-binding specificity among
ARF proteins is highly similar, even between the phylogenetically
diverse family members ARF1 and ARF5, a finding that is consis-
tent with the limited sequence divergence at the DNA-binding
surface. One potential caveat is that, even though ARF domains
can fold and act in isolation (this study and Tiwari et al., 2003), it is
possible that other domains alter DNA binding specificity by
intramolecular interactions with the DBD. Particularly the middle
region is very divergent between ARFs and correlated with the
ability of ARFs to either activate or repress transcription (Tiwari
et al., 2003). Nonetheless, we find that dimerization allows for
variation of ARF DNA recognition at the level of spacing between
two adjacent inverted binding sites. This mechanism, in which
ARFs act as “molecular calipers” to bind uniquely spaced
motifs, can at least account for differences between ARF1 and
ARF5 and is consistent with in vivo binding sites for ARF5
(Schlereth et al., 2010; Yamaguchi et al., 2013). The divergence
in the loops connecting the B3 and DD domains extends beyond
ARF1 and ARF5, and it is therefore conceivable that other ARFs
also have distinct interdomain flexibility that allows unique bind-
ing site spacing. It will be interesting to address what distance
can be accommodated by ARF complexes and if two binding sites
can be separated by larger DNA loops or nucleosomes.

In this context, it is important to note that the distance of seven
bases between AuxRE sites in the ER7 substrate requires little
or no protein conformational change or torsion of the DNA
(Figure 3B). In contrast, increasing or decreasing this distance
will also rotationally displace the two binding sites. Hence, the
different potential in binding between ARF1 and ARF5 depend-
ing on site spacing may either be a consequence of different
flexibility of the dimer or a difference in the capacity of the two
proteins to induce DNA bending or torsion. Given the different
biophysical properties of A:T and G:C pairs, sequence within
the spacer may also contribute to binding affinity.

Transcription factor dimerization is a common element in
transcriptional control. Often, dimerization is required for bind-
ing a single site, such as is the case in basic helix-loop-helix
(bHLH) factors (e.g., MyoD; Ma et al., 1994) or by bZip factors
(e.g., AP-1/CREB; Kim and Struhl, 1995). Unlike many other ex-
amples, however, ARF DNA binding can involve either one or
two binding sites, where the latter case involves cooperativity.
Conceptually, the mechanism underlying sequence-specific
DNA binding in the ARF family is similar to that found in the
animal nuclear receptor (NR) family. Members of this family of
transcription factors, whose nuclear localization and activity is

Figure 6. A Spacing-Based Model for ARF DNA-Binding Specificity
(A) Conservation of amino acids between ARF1 and ARF5 mapped onto the external loop in the ARF1-DBD/ER7 structure. Blue, identical; green, conserved;
yellow, semiconserved; red, nonconserved.
(B and C) SPR binding profiles of ARF1-DBD (B) and ARF5-DBD (C) proteins to ER5, ER6, ER7, ER8, and ER9 (number indicates spacing between two inverted
TGTCTC elements). Values were normalized to the highest value of that same protein on the ER7 oligonucleotide. Scales are identical in all panels.
(D) Complex ARF binding sites in ARF5/MP target genes LFY, TM03, and TM05. Binding sites are in bold, and intervening bases are numbered.
(E) Model for auxin-dependent transcription. ARF proteins bind DNA as dimers, mediated by interactions in the DNA-binding domain. The main determinant of
specificity is the spacing between the two binding sites (1), although quantitative differences in preferences for binding sites may exist (2). Auxin controls ARF
activity by promoting degradation of Aux/IAA proteins that bind to the distant domains III/IV. ARFs may also act through low-affinity DNA-binding as monomers
(3), and heterodimerization (4) may further extend the range of binding preference.
See also Table S2.
modulated by membrane-permeable hormones such as retinoic acid or estrogen (Mangelsdorf et al., 1995), bind DNA either as monomers or dimers (Khorasanizadeh and Rastinejad, 2001). When bound as dimers, the choice of the partners determines the optimal spacing (3, 4, or 5 bases) between two tandem binding sites, a phenomenon that led to the formulation of the 3,4,5 rule (Umesono et al., 1991). The case with ARFs is distinct as symmetric homodimers bind an inverted repeat rather than a tandem repeat as bound by the NR dimers, and in addition, the space between two binding sites is large for ARFs. Nonetheless, both NR (Zechel et al., 1994) and ARFs can bind DNA cooperatively and contribute to generating specific responses to hormonal signals in the animal and plant kingdom.

A key question in auxin biology is how this structurally simple molecule can elicit such a wide range of growth and developmental responses. Our study suggests a model in which diversification of gene expression responses follows from the distinct properties of dimeric complexes formed by the DNA-binding ARF transcription factors. This model, as well as the ARF structures presented here, will now open new avenues to define the mechanistic basis for context-dependent gene regulation in the auxin pathway.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification

Regions corresponding to the DBD of Arabidopsis ARF1 (At1g59750; residues 1–354) and ARF5 (At1g198850; residues 1–390) were amplified from cDNA clones using primers as listed in Table S3 using Phusion Flash polymerase (Finnzymes) and cloned in an expression vector pTWIN1 (New England Biolabs) to generate fusions with chitin-binding domain (CBD) and Intein. ARF-DBD-CBD fusion proteins were expressed in E. coli strain Rosetta DE3 (Novagen). Protein expression was induced by 0.3 mM IPTG for 20 hr at 20 °C, and proteins were purified from cell-free extracts by affinity chromatography on a chitin column followed by size exclusion chromatography on a Superdex 200PG column, both using an Akta Explorer 100 (GE Healthcare). Full details on expression and purification are described in the Extended Experimental Procedures.

X-Ray Crystallography

All crystals were grown at 20 °C using sitting drop vapor diffusion experiments. Initial screens were performed using 80–200 nl droplets on 96-well plates using a Cartesian robot. Additive screens on initial hits showed improved crystal size and longevity with GSH/GSSG. Additional trials using the reducing agents GSH and DTT confirmed the dependence of crystal growth and stability on the reduction potential of the environment. Oligonucleotides used for crystallization were obtained from Biomers (Ulm, Germany). Full details on crystallization conditions, data collection, processing, structure determination, and refinement are given in Extended Experimental Procedures.

Small-Angle X-Ray Scattering Measurements

SAXS data of ARF1-DBD (concentration 3.2 mg/ml) were collected at beamline BM29 (ESRF, Grenoble). BSA references were used for calculating the molecular mass of ARF1-DBD. Measurements were carried out at 293 K, within a momentum transfer range of 0.01 Å⁻¹ < q < 0.45 Å⁻¹. Calculation of the theoretical scattering curves of monomeric and dimeric ARF1 against the scattering data was performed using CRYSOL (Svergun et al., 1995).

Surface Plasmon Resonance and Circular Dichroism Spectroscopy

SPR measurements were performed using eight 2-fold dilution steps (800 nM, 400 nM, 200 nM, 100 nM, 50 nM, 25 nM, 12.5 nM, and 6.25 nM) of purified ARF-DBD proteins on a Biacore 3000 platform using double-stranded biotin-labeled oligonucleotides (5’-biotin: Eurogentec; Table S3) immobilized on SA chips (GE Healthcare). Data were analyzed with Scrubber2-T200 (BioLogic Software Pty Ltd).

Circular dichroism was performed on 0.1 mg/ml dilutions of purified ARF DBD proteins in 0.1 M sodium borate buffer (pH 7.4) using a 1 mm quartz cell in a J-715 CD spectropolarimeter (Jasco). Traces are averages of 20 spectra and smoothed over 3 nm windows.

Site-Directed Mutagenesis and Cloning

Mutations were introduced into cDNA fragments corresponding to the ARF DBDs through PCR using primers as listed in Table S3, and fragments were cloned into pTWIN1. The wild-type and mutated cDNA of ARF5-DBD were amplified and used to replace the genomic DBD in an 8.5 kb MspI genomic fragment (Weijers et al., 2006) using the unique restriction sites Xho1 and BamHI. Wild-type and mutant versions of the full-length ARF5 cDNA or a fragment truncated after T794 (after Krogan et al., 2012) were LIC cloned (De Rybel et al., 2011) using primers as listed in Table S3 into the PMON999 (Monsanto) vector and fused to scFV3A or sYFP2 and transiently expressed in A. thaliana Columbia ecotype mesophyll protoplasts under the 35S promoter for FRET-FLIM assays. The ARF3 plasmids were previously described (Rademacher et al., 2012).

Protein-Binding Microarrays

PBMM1 was performed on ARF1-DBD, ARF5-DBD, and their mutated versions H170A, G171A, R215A, P218A, S269N, and G279E according to Godoy et al. (2011) with modifications detailed in the Extended Experimental Procedures.

Plant Growth and Rescue Experiments

Heterozygous plants of the mp-b4149 strong allele (Weijers et al., 2005) were transformed with a pGREEN vector carrying the construct pMP-mp and its different mutated versions by floral dipping with A. tumefaciens. Seeds carrying the transgene were screened on MS media with 15 mg/ml phosphinothricin (PPT). Segregation of the monopetrous phenotype in the T2 generation was checked to determine the genotype of the T1 plants. The percentage of rootless seedlings observed in the progeny of heterozygous T1 plants was used to determine whether the transgene could rescue the phenotype.

FRET-FLIM

Transfections were performed as described (Rusinova et al., 2004) using Arabidopsis (Columbia wild-type) mesophyll protoplasts. Fluorescence lifetime imaging (FLIM) images were acquired with a Leica TCS SP5 X system equipped with a 63× 1.2 NA water immersion objective lens. scFV fluorophore was excited using a pulse diode laser (40 mW) at a frequency of 40 mHz. Donor fluorescence was recorded via an external fiber output connected to the Leica TCS SP5 X scan head and coupled to a Hamamatsu HPM-100-40 Hybrid detector (Becker & Hickl), which has a time resolution of 120 ps. Donor fluorescence emission was obtained using a 470–500 nm band pass filter. Images of 128 × 128 pixels were acquired with acquisition times of 120 s. FRET-FLIM analysis in Arabidopsis leaf mesophyll protoplasts was performed as described previously (Rademacher et al., 2012).

ACCESSION NUMBERS

The Protein Data Base (PDB) accession numbers for the structures reported in this paper are 4LDV, 4LDW, 4LDU, 4LDX, and 4LDY.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, five figures, and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2013.12.027.

ACKNOWLEDGMENTS

This work was supported by the Netherlands Organization for Scientific Research (NWO; ECHO grant 711.011.002 to D.W.), the Spanish Ministry of
Science and Innovation (grant BFU2011-22588 to M.C.), the Generalitat de Catalunya (grant SGR2009-1309 to M.C.), and the European Commission (FP7 Cooperation Project SILVER - GA 260644 to M.C.). We thank Vincent Rui-grok and Hans Heilig for technical advice; Silke Lochs and Henri Doornbos for experimental support; Bert De Rybel and Lars Østergaard for helpful discus-sion; and the staff of beamlines ID14-1 and ID14-4, in particular Andrew McCarthy at the ESRF (Grenoble, France), and beamline XALOC (ALBA, Cer-danyola del Vallés, Spain). Crystallization was performed at the IBMB Platform for Automated Crystallization (Barcelona, Spain).

Received: August 27, 2013
Revised: November 22, 2013
Accepted: December 11, 2013
Published: January 30, 2014

REFERENCES


