Investigation of triterpene synthesis and regulation in oats reveals a role for β-amyrin in determining root epidermal cell patterning

Ariane C. Kemen,a,1 Suvi Honkanen,a,b,2 Rachel E. Meltona, Kim C. Findlayc, Sam T. Mugforda, Keiko Hayashia,d and Anne Osbournea,5

Departments of aMetabolic Biology and cCell and Developmental Biology, John Innes Centre, Norwich, Norfolk NR4 7UH, United Kingdom; bInstitute of Molecular Cell and Systems Biology, University of Glasgow, Glasgow G12 8QQ, United Kingdom; and dDepartment of Biology, University of Athens, 15784 Athens, Greece

Edited by Joe Chappell, University of Kentucky, Lexington, KY, and accepted by the Editorial Board April 28, 2014 (received for review January 28, 2014)

Sterols have important functions in membranes and signaling. Plant sterols are synthesized via the isoprenoid pathway by cyclization of 2,3-oxidosqualene to cycloartenol. Plants also convert 2,3-oxidosqualene to other sterol-like cyclization products, including the simple triterpene β-amyrin. The function of β-amyrin per se is unknown, but this molecule can serve as an intermediate in the synthesis of more complex triterpene glycosides associated with plant defense. β-Amrin is present at low levels in the roots of diploid oat (Avena strigosa). Oat roots also synthesize the β-amyrin-derived triterpene glycoside avenacin A-1, which provides protection against soil-borne diseases. The genes for the early steps in avenacin A-1 synthesis [saponin-deficient 1 and 2 (Sad1 and Sad2)] have been recruited from the sterol pathway by gene duplication and neofunctionalization. Here we show that Sad1 and Sad2 are regulated by an ancient root developmental process that is conserved across diverse species. Sad1 promoter activity is dependent on an L1 box motif, implicating sterol/lipid-binding class IV homeodomain-leucine zipper transcription factors as potential regulators. The metabolism of β-amyrin is blocked in sad2 mutants, which therefore accumulate abnormally high levels of this triterpene. The accumulation of elevated levels of β-amyrin in these mutants triggers a “superhairy” root phenotype. Importantly, this effect is manifested very early in the establishment of the root epidermis, causing a greater proportion of epidermal cells to be specified as root hair cells rather than nonhair cells. Together these findings suggest that simple triterpenes may have widespread and as yet largely unrecognized functions in plant growth and development.

Significance

Sterols and triterpenes are complex molecules that are synthesized from the isoprenoid pathway. The functions of sterols in plants have been studied extensively, but the role of triterpenes is less well understood. Here we investigate triterpene synthesis and regulation in diploid oat. We show that the genes for triterpene synthesis are regulated by an ancient root development process that is conserved across diverse plants. We further show that mutants in which the metabolism of the most common plant triterpene, β-amyrin, is blocked undergo a change early in the development of the root epidendermis that leads to a “superhairy” root phenotype. Our findings shed light on triterpene synthesis and provide evidence for a role for the simple triterpene β-amyrin in plant development.
Sad1 promoter function in cots and eudicots. We use promoter are under the control of an ancient regulatory mechanism as- in oats. Importantly, this effect impacts early in the establish- vation of avenacin biosynthesis at any level or indeed about the physical clustering may facilitate regulation at the level of biosynthesis (7, 9, 11)
a metabolic gene cluster together with other genes for avenacin expressed specifically in the epidermal cells of oat root tips, the (Fig. 1B) (22). Previously, we have shown by mRNA in situ hybridization that the genes for avenacin bio- synthesis are expressed specifically in this cell layer (8, 9, 11–13, 15). The physical clustering of the avenacin-pathway genes in oat is suggestive of regulation at the level of chromatin, a pre- diction supported by DNA FISH analysis of chromatin decon- densation in this region (16). However, when the Sad1 promoter was fused to the β-glucuronidase (GUS) reporter gene and in- troduced into A. thaliana, the resulting GUS-expression pattern mirrored that of avenacin A-1 fluorescence in oats; i.e., GUS staining was detected in the root tips and lateral root initials (Fig. 1C). Similar results were obtained with rice and M. truncatula (Fig. 1C). This promoter therefore contains the information necessary for appropriate expression in the root tips and lateral root initials of diverse plant species, including both monocots and eudicots. Similar results were obtained with the Sad2 promoter in A. thaliana (Fig. S1A).

Further analysis using GUS and GFP reporters revealed that the Sad1 and Sad2 promoters drive reporter gene expression predominantly in the epidermal cells of the root tips in both A. thaliana and rice (Fig. 1D). Expression was also observed in the lateral root primordia (Fig. S1 A and B). Differences observed with the GUS and GFP reporters are likely to be caused by the diffusion of the GUS product into nonexpressing neighboring tissue. The activity of the Sad2 promoter (pSad2) in the root tip was more restricted than that of the Sad1 promoter (pSad1) and was localized primarily in the cells of the quiescent center and the epidermal stem cells (Fig. 1D). Interestingly in M. truncatula roots pSad1 activity was observed not only in primary roots and lateral roots but also in the meristematic regions of root nodules, suggesting similarities in the regulatory mech- anisms within the root meristem and nodules (Fig. S1C).

These experiments led us to conclude that the Sad1 and Sad2 promoters are under the control of a highly conserved root de- velopment process that is present in both monocots and eudicots.

Analysis of the Sad1 Promoter. We next carried out deletion ex- periments using the 3-kb Sad1 promoter and A. thaliana as our expression system to identify promoter elements implicated in the regulation of gene expression. Preliminary experiments re- vealed that deletion of ~500 bp from the 5’ end of the Sad1 promoter region resulted in the loss of GUS expression (Fig. S2A) and defined a region between −2598 and −2488 as critical for promoter function (Fig. S2B). Insertion of the 5′-most kilobase of the Sad1 promoter (the region from −2992 to −1874) directly in front of the GFP/GUS reporter genes (construct pHGWFSAS2992–1847) resulted in GUS expression similar to that observed with the full 3-kb pSad1 region (Fig. S2B). Further fine deletion nar- rowed the critical region down to only 22 nucleotides (Fig. 2A). This region contains a pseudopalindromic sequence CATTATA- TGTGA with an embedded L1 box motif [5′-TAAATTC(C/ T)A-3′] (Fig. 2B) (23). Pseudopalindromic sequences that overlap L1 boxes are binding sites for HD-ZIP IV proteins (24). Members of this family of transcription factors contain a steroidogenic acute regulatory protein-related lipid transfer (START) domain suggestive of sterol/lipid binding and are implicated in epidermis-
Sad2 threshold (CT) method (42). Relative transcript levels are shown. **

expression of transcript levels in the terminal 1 seedlings is shown. (B) The L1 box and associated pseudopalindromic sequence identified in the Sad1 promoter sequence. The presence or absence of full L1 boxes with pseudopalindromic sequences (red) and core L1 box sequences (yellow) in the promoter regions of Sad1, Sad2, CAS1, and CYP51G1. (D) The genes for sterol and avenacin biosynthesis are inversely regulated at the level of transcription. qRT-PCR analysis of transcript levels in the terminal 1 seedlings is shown. (**p < 0.001; *p < 0.05) related developmental processes (17). Seedlings transformed with a −2541 nucleotide construct disrupted in the L1 box motif showed very faint GUS staining in occasional individual cells, and the typical pSad1-expression pattern disappeared (Fig. 2A). Bioinformatics analysis revealed that a pseudopalindromic L1 box sequence is also present in the Sad2 promoter. The Sad1 and Sad2 promoters also have multiple core L1 box motifs. The promoters of the oot sterol biosynthesis genes CAS1 and CYP51G1 from which Sad1 and Sad2 have been recruited (8, 9, 15) do not contain pseudopalindromic L1 box sequences, although the CAS1 promoter has a single core L1 box motif (Fig. 2C).

Expression of Sad1, Sad2, and the other avenacin biosynthesis genes is restricted to the epidermal cells of root tips and lateral root primordia (8–13, 15). We compared the relative transcript levels of Sad1, Sad2, and the corresponding sterol biosynthesis genes CAS1 and CYP51G1 in seedlings of diploid oat (Avena strigosa). CAS1 and CYP51G1 transcript levels were higher in the maturation zone than in the root tips, whereas the reverse was true for Sad1 and Sad2 transcript levels (Fig. 2D). These results are consistent with our earlier mRNA in situ experiments, which suggest that the sterol and avenacin pathways are spatially separated and inversely regulated at the level of transcription (16).

Experiments in which pSad1-GUS reporter lines of A. thaliana were crossed with mutants defective in auxin-related transport (25, 26) or signaling (27) gave normal GUS-expression patterns (Fig. S3). Homozygous mutant lines of the auxin response mutant MONOPTEROS, mpg12 (28), do not form primary root meristems during embryogenesis, and GUS expression was not seen in this mutant background. Occasionally mpg12 mutant seedlings produce adventitious roots (28), and GUS staining was detected in these rare adventitious roots (Fig. S3). Thus, Sad1 promoter activity is auxin independent but depends on the formation of the primary root meristem.

To investigate likely regulators of the avenacin-pathway genes, we generated a root tip transcriptome database for A. strigosa. Representation of root tip-expressed genes within this database was validated by searching for matches to each of the five cloned avenacin genes (SI Materials and Methods and Table S1). BLAST searches using all 16 A. thaliana HD-ZIP IV family members as probe sequences (24) identified multiple reads corresponding to a single class I HD-ZIP sequence, AsHDZ1. AsHDZ1 shares close amino acid sequence similarity (ε-values = 0) with HD-ZIP IV proteins from other cereals, including Triticum aestivum (AK335308), Hordeum vulgare (AK362919), ROC5 from rice (Orzya sativa; AB101648), a ROC5-like protein from Brachypodium distachyon (XM_003570037), and OCL1 from Zea mays (NM_00112023) (Fig. 3A). These sequences together form a monocot-specific subclade within the α-clade of the plant HD-ZIP IV superfamily (29). OCL1 and ROC5 are expressed in the epidermal cell layers of organs in monocots and are involved in regulation of processes such as the determination of epidermal cell fate, anthocyanin accumulation, and root development (30, 31). The wider α-clade includes the A. thaliana HD-ZIP IV proteins HDG1, HDG6, HDG7, and ANL2 (24, 32). GLABRA2 (2G), a well-characterized negative regulator of root hair development in A. thaliana (17, 33), belongs to a different clade—the β-clade (Fig. 3A).

AsHDZ1 has an expression pattern similar to that of Sad1 and Sad2 in roots (Fig. 3B). Unlike Sad1 and Sad2, however, this gene is also expressed in young leaves (as are OCL1 and ROC5 from rice) (25, 26). The expression pattern of the A. thaliana HD-ZIP IV genes belonging to the α-clade overlap, suggesting functional redundancy (24). HDG7 is a basal member of this clade and is expressed in roots. Experiments evaluating pSad1-GUS reporter activity in a homozygous hdg7 mutant background indicated that Hdg7 is dispensable for GUS expression (Fig. S4). Future experiments to identify regulators of pSad1 and pSad2 expression in A. thaliana are likely to require the generation of multiple combinatorial mutants for the α-clade HD-ZIP IV genes. Although it is clear that the Sad1, Sad2, and HDG7 promoters are under the control of an ancient development process that is conserved across diverse plant species, it does not necessarily follow that the transcription factors involved are orthologous. The establishment of a TILLING platform (34) for diploid oat will open future opportunities to carry out functional analysis of AsHDZ1 in A. strigosa using reverse genetics approaches.

Accumulation of β-Amyrin Leads to a Change in Cell Fate in the Root Epidermis. In most plant species the cells that will give rise to root hairs are predetermined (20). These predetermined trichoblasts are smaller than those that are destined to become atrichoblasts and differentiate into root hair cells in the differentiation zone of the root. Trichoblasts develop in one of two distinct patterns. They may alternate with the atrichoblasts along longitudinal cell files, as is the case in oat; alternatively, trichoblasts and atrichoblasts may develop in separate longitudinal files as in A. thaliana (Fig. 4A) (18, 19). Fluorescence microscopy indicates that the fluorescence attributable to avenacin A-1 is present predominantly in the atrichoblasts of oat root epidermal cells (Fig. 4A). Interestingly, investigation of pSad1 reporter lines revealed that GUS activity was detectable primarily in the atrichoblasts of A. thaliana (Fig. 4B). Thus, Sad1 is expressed in atrichoblasts, and this cell type-specific expression is conserved in A. thaliana.

The SAD1 product β-amyрин is widespread in plants (5). It is present at low levels in the roots of wild-type A. strigosa seedlings (~1.5–3 μg/g dry weight) and at even lower but detectable levels
in sad1 mutants (−0.7–0.8 μg/dry weight) (9, 35). sad2 mutants are unable to convert β-amyrin toavenacin (Fig. 1A) and thus accumulate considerably elevated levels of this triterpene (−40–50 μg/dry weight) (9, 35). sad2 mutants have a short root phenotype, whereas the wild type and other avenacin biosynthesis mutants (11, 12) do not (Fig. 4 C and D). The short root phenotype therefore is associated with the accumulation of abnormal levels of β-amyrin and not with the loss of the entire pathway (sad1 mutants) or the accumulation of later pathway intermediates (sad7 and other downstream mutants). A partial sad2 mutant (no. 791) that accumulates intermediate levels of avenacin A-1 and β-amyrin (Fig. 55) (7, 9, 35) has an intermediate root-length phenotype (Fig. 56).

Cryo-scanning electron microscopy (cryo-SEM) revealed that the roots of sad2 mutants have a superhairy phenotype (Fig. 4E). The lengths of the trichoblast and atrichoblast cells in the various mutant lines were unaltered as compared with the wild type (Fig. 4F). However, the sad2 mutants had significantly more root hair cells per unit area than the wild-type or sad1 lines. Interestingly, these cells have the shorter dimensions typical of trichoblast cells, i.e., they were predetermined to become hair cells (Figs. 4G and 5). Thus, the accumulation of elevated levels of β-amyrin triggers a change in cell specification in roots. The intermediate sad2 mutant #791 had an intermediate phenotype (Fig. 4 E and G). Thus, the accumulation of elevated levels of β-amyrin in oat roots results in a greater proportion of epidermal cells becoming trichoblasts. This effect is distinct from the actions of most genetically characterized root hair determinants, including GL2, which act after cell specification to regulate root hair growth (17–21).

Interestingly, our earlier mRNA FISH experiments using highly sensitive fluorophore detection methods detected low levels of Sad1 transcript in 50% of the cells of the subepidermal cell layer, but the Sad2 transcript was not detectable in this cell layer (16). The subepidermal cell layer is known to influence cell specification in the root epidermis through as yet unknown signaling processes (18–21). Therefore it is tempting to speculate that β-amyrin may play a role in such processes, perhaps by influencing asymmetric cell division and hence specification of trichoblasts and atrichoblasts as the epidermal cells leave the meristematic zone. Application of exogenous β-amyrin to wild-type and sad1 mutant A. strigosa lines did not result in an increase in root hair cells (Fig. 57). The failure to phenocopy the sad2 root phenotype was not caused by the conversion of β-amyrin toavenacin, as evidenced by LC-MS analysis (Fig. 58). Likewise transgenic rice lines expressing Sad1 under the control of the maize ubiquitin promoter did not show a short root or root hair phenotype (Fig. 58). This finding is not surprising, however, given that the effect of β-amyrin on cell fate is likely to depend on local concentrations of the triterpene in particular cell types at a specific stage in development.

Discussion

Previously we reported the discovery and characterization of a gene cluster for avenacin synthesis in oat and showed that Sad1 and Sad2, the first two genes in the pathway, have been recruited from sterol biosynthesis (8, 9, 15). These genes have arisen by gene duplication and neofunctionalization, coupled with a change in expression pattern (from being widely expressed to being highly tissue specific). Here we show that the promoters of these genes retain their characteristic primary and lateral root meristem expression patterns when introduced into other plant species, indicating that the promoters are appropriately regulated in diverse heterologous plant backgrounds, including both monocots and dicots. Other genes within the avenacin cluster have expression patterns very similar to those of Sad1 and Sad2, with expression for EF1-α. Values are means ± SD (three plants). Data were analyzed using the comparative C(T) method (42). Relative transcript levels are shown.*P < 0.05 (unpaired two-tailed t test).
the function of AsHDZ1 in the regulation of triterpene synthesis and to establish the roles of lipid and sterol ligands in pathway regulation.

2,3-Oxidosqualene can be converted to a diverse array of triterpene products in plants. Of these, the most common cyclization products β-amyrin, α-amyrin, and lupeol are widespread in plants and have been reported in unconjugated form in numerous species, including *A. thaliana* (1–5). Little is known about the physiological functions of these compounds in *plants*. However, an increasing body of evidence suggests that they may have important roles. For example, they accumulate in the epidermal layers of stem and leaf surfaces and have been implicated in protection against dehydration and potentially also against herbivores (reviewed in ref. 5). They are also involved in signaling. Previously we reported that lupeol is synthesized in the nodules of the legume *Lotus japonicus* and is involved in nodule development (36). Interestingly, accumulation of the simple triterpene thalianol in *A. thaliana* leads to enhanced root length (37). Our discovery that the accumulation of elevated levels of β-amyrin in oat roots causes a change in epidermal cell patterning provides further intriguing evidence suggesting that simple triterpenes may have widespread and as yet largely unrecognized functions in plant growth and development. At present we cannot distinguish between direct effects (e.g., specific binding to a receptor) and indirect effects (e.g., perturbation of membranes). Nevertheless, elucidation of the mechanistic basis of these effects is likely to provide new insights into the factors that govern plant growth and development.

**Materials and Methods**

**Plant Material.** Wild-type and mutant lines of *A. strigosa* are as described (7, 9). The methods used to measure root length and metabolite analysis after β-amyrin feeding are described in *SI Materials and Methods*. *Col-O* was used as the wild-type *A. thaliana* accession for this work. Mutant *A. thaliana* lines used were *tr1-1* (N3798) (27) and *pin1-7* (SALK_047613) (26), *mp12G* (28), *pil9* (25), and *hdg7* (SALK_132114), all from Nottingham Arabidopsis Stock Center. Crosses with the *pSad1-GUS* reporter line were genotyped as described being tightly restricted to the epidermal cells of the root meristems (11–13). Thus, the genes for avenacin synthesis have acquired a common and specific expression pattern during pathway evolution. Our deletion analysis of the *Sad1* promoter identified a classical L1-box motif with a pseudopalindromic border sequence that is essential for promoter function, implicating the HD-ZIP IV family of transcription factors in the regulation of triterpene synthesis. A future challenge will be to determine

---

**Fig. 4.** Accumulation of β-amyrin triggers changes in root development. (A) (Left) Cryo-SEM image of wild-type *A. strigosa* root epidermis showing the root hair pattern. Hair cells are highlighted in pink in the cryo-SEM image at the far left. The patterning of trichoblast (pink) and atrichoblast (gray) cells before root hair emergence (modified from ref. 20) is illustrated in the adjacent diagram. (Right) The blue fluorescence in the maturation zone of oat roots is stronger in atrichoblasts than in trichoblasts. The white box indicates the magnified region shown in the panel at the far right. (Scale bars: 200 μm and 20 μm, respectively.) (B) cryo-SEM image of *A. thaliana* Col0 root epidermis (Left) and schematic view (Center) showing the epidermal cell patterning before root hair emergence; trichoblasts are indicated in blue and atrichoblasts in gray (modified from ref. 20). (Right) *Sad1*-driven GUS expression in *A. thaliana* roots is stronger in atrichoblasts than in trichoblasts, creating a striped pattern in the elongation zone. (Scale bar: 50 μm.) (C and D) Growth of wild-type, *sad1*, and *sad2* lines of *A. strigosa* after 15 d on water agar. Results in D are represented as mean ± SE; *P < 0.0001 (unpaired two-tailed t test); n indicates number of seedlings analyzed. (E) Representative cryo-SEM images of the maturation zones of wild-type *A. strigosa* and the *sad1* mutant #109 and #610, and *sad2* mutant #791, #1027, and #500. *A. strigosa* lines. (Scale bars: 200 μm.) (F and G) Lengths of trichoblasts (one-way ANOVA: *F*<sub>4,294</sub> = 2.01; *P = 0.08) and atrichoblasts (one-way ANOVA: *F*<sub>4,294</sub> = 1.66; *P = 0.14) (F) and numbers of trichoblasts and atrichoblasts per square millimeter (G) in the maturation zones for the lines shown in E. ***P < 0.001; **P < 0.01; *P < 0.05 (unpaired two-tailed t test). Mutant #791 is a partial *sad2* mutant (9).

---

**Fig. 5.** Increased β-amyrin content alters epidermal cell patterning in oat roots. The levels of β-amyrin in root extracts of wild-type and *sad2* mutant lines are shown (9, 35). *sad2* mutants are unable to metabolize β-amyrin to avenacin A-1 and therefore accumulate considerably elevated levels of this simple triterpene. The effects of the accumulation of elevated levels of β-amyrin are seen early in the establishment of the root epidermis, causing a greater proportion of epidermal cells to be specified as trichoblasts (specialized shorter cells that give rise to root hairs) (indicated in pink).
Accession Numbers. Sequence data for the genes and promoters referred to in this article can be found in the GenBank/European Molecular Biology Laboratory database libraries under accession nos. DQ680849 (Sad1 and Sad2, including promoter regions); KF857483 (CAS1 promoter sequence); KF857482 (AsCYP51G1 promoter sequence); AY16893 (CAS1 coding sequence); DQ680850 (AsCYP51G1 coding sequence); and KF857484 (AsHD21 coding sequence). The oat root tip transcriptome data have been deposited in the European Bioinformatics Institute database, accession no. ERA148431.

Transgenic Plants. The promoter regions of the Sad1 and Sad2 genes were amplified from BAC clone 460D15 (9) using the primers listed in Table S2. For construct pK1848aGUS a BamHI/EcoRI fragment harboring 1,848 bp of the Sad1 promoter, the GUS gene, and the NOS-terminator were cloned into the vector pBI101, and amplified from BAC clone 460D15 (9) using the primers listed in Table S2. For Transgenic Plants.

Generation of Oat Root Transcriptome Database, Phylogenetic Analysis, and Microscopy. Details are provided in SI Materials and Methods.

Quantitative RT-PCR. Total RNA was extracted from roots of 3-day-old oat seedlings. Details of RNA extraction are described in SI Materials and Methods. Quantitative RT-PCR (qRT-PCR) and relative quantification were performed as previously described (41). Gene-specific primers for Sad1, CAS1, Sad2, CYP51G1, and AshDZ1 were designed around introns within the coding sequences (Table S3). Oat elongation factor-1α (EF-1α) was used as an internal standard. The method described by Livak and Schmittgen (42) was used to quantify gene expression.

ACKNOWLEDGMENTS. We thank Lionel Hill for metabolite analysis, the John Innes Centre horticulture staff for plant care, Andrew Davis for assistance with photography of plants, the Genome Analysis Centre for generation of the oat root transcriptome resource, and Brenda Blacklock for useful discussions. This work was supported by Deutsche Forschungsgemeinschaft Fellowship KE 1519/1-1 (to A.C.K.); a Joint Engineering and Physical Sciences Research Council/National Science Foundation (NSF/EPSCoR) award (to A.O. and S.J.R.) as part of the Syntegron consortium NSF/EPSCoR Grant EPIK334951; the University of Glasgow Work Placement Programme (S.H.); Plant Bioscience Limited; the Biotechnology and Biological Sciences Research Council Institute Strategic Programme Grant BB/J000461/1 “Understanding and Exploiting Plant and Microbial Secondary Metabolism”; and the John Innes Foundation.