A conserved amino acid residue critical for product and substrate specificity in plant triterpene synthases

Melissa Salmon1, Ramesha B. Thimmappa1, Robert E. Minto6, Rachel E. Melton9, Richard K. Hughes2, Paul E. O’Maille5, Andrew M. Hemmings5,8, and Anne Osbourn3,2

1John Innes Centre, Norwich Research Park, Norwich NR4 7UH, United Kingdom; 2Department of Chemistry and Chemical Biology, Indiana University–Purdue University, Indianapolis, IN 46202; 3Food & Health Programme, Institute of Food Research, Norwich Research Park, Norwich NR4 7UH, United Kingdom; 4School of Chemistry, University of East Anglia, Norwich NR4 7TJ, United Kingdom; and 5School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, United Kingdom

Edited by Rodney B. Croteau, Washington State University, Pullman, WA, and approved June 9, 2016 (received for review April 5, 2016)

Triterpenes are structurally complex plant natural products with numerous medicinal applications. They are synthesized through an origami-like process that involves cyclization of the linear 30 carbon precursor 2,3-oxidosqualene into different triterpene scaffolds. Here, through a forward genetic screen in planta, we identify a conserved amino acid residue that determines product specificity in triterpene synthases from diverse plant species. Mutation of this residue results in a major change in triterpene cyclization, with production of tetracyclic rather than pentacyclic products. The mutated enzymes also use the more highly oxygenated substrate dioxidosqualene in preference to 2,3-oxidosqualene when expressed in yeast. Our discoveries provide new insights into triterpene cyclization, revealing hidden functional diversity within triterpene synthases. They further open up opportunities to engineer novel oxygenated triterpene scaffolds by manipulating the precursor supply.

Significance

The triterpenes are a large and highly diverse group of plant natural products. They are synthesized by cyclization of the linear isoprenoid 2,3-oxidosqualene into different triterpene scaffolds by enzymes known as triterpene synthases. This cyclization process is one of the most complex enzymatic reactions known and is only poorly understood. Here, we identify a conserved amino acid residue that is critical for both product and substrate specificity in triterpene synthases from diverse plant species. Our results shed new light on mechanisms of triterpene cyclization in plants and open up the possibility of manipulating both the nature of the precursor and product specificity, findings that can be exploited for the production of diverse and novel triterpenes.

Author contributions: M.S., R.B.T., R. E. Minto, P.E.O., A.M.H., and A.O. designed research; M.S., R.B.T., and R. E. Minto contributed new reagents/analytic tools; M.S., R.B.T., R. E. Minto, R. E. Melton, A.M.H., and A.O. analyzed data; and M.S., R.B.T., R. E. Minto, A.M.H., and A.O. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. Freely available online through the PNAS open access option.

1 M.S. and R.B.T. contributed equally to this work.
2 To whom correspondence should be addressed. Email: anne.osbourn@jic.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1605509113/-/DCSupplemental.
triterpene scaffolds by manipulation of the precursor supply. Our results further illustrate the power of using a forward genetics approach to identify residues that are critical for the stability and functional diversification of triterpene synthases.

Results

Identification of Mutant SAD1 Protein Variants. The avenacin-deficient A. strigosa mutants were generated by using the chemical mutagen sodium azide (13), which causes single-base substitutions, usually from guanine to adenine (22, 23). DNA sequence analysis of the Sad1 gene in each of the 16 new candidate sad1 mutants (21) revealed single-point mutations in each case, the majority of which involved guanine-to-adenine transitions as expected. The mutants could be divided into three categories (Table 1)—those with predicted premature termination of translation (as for the two original sad1 mutants 109 and 610) (14); those with mutations at intron-exon boundaries that may give rise to splicing errors; and those with predicted amino acid substitutions.

We then assessed the sad1 transcript levels in RNA from the root tips of these mutants by RT-PCR. The four new mutants with predicted premature termination of translation codons (Table 1), like 109 and 610 (14), had substantially reduced transcript levels (Fig. 1B), most likely due to nonsense-mediated mRNA decay (24). As expected, Western blot analysis using polyclonal antisera specific for SAD1 (16) failed to detect cross-reacting protein in protein preparations from the roots of these mutants (SI Appendix, Fig. S1B). These deletions may result in the formation of misfolded proteins that are targeted for degradation (25).

The transcript levels for the seven mutants with predicted amino acid substitutions were unaltered (Fig. 1D). Western blot analysis revealed a protein of the same molecular mass as SAD1 in root

---

**Fig. 1.** Characterization of sad1 mutants. (A) Biosynthesis of phytosterols and avenacin A-1 in oat. (B–D) RT-PCR analysis of mutant sad1 transcript levels in mRNA extracted from the roots of wild-type (WT) oats and predicted premature termination of translation (B), splicing error (C), and amino acid substitution (D) mutants (Table 1). The oat glyceraldehyde-3-phosphate dehydrogenase gene (GAPDH) was used as a control. (E) Analysis of protein extracts from root tips of WT and sad1 mutants probed with antisera raised against SAD1. A single band of ∼86 kDa corresponding to full-length SAD1 protein is present in the WT and mutants 358, 384, and 1023. (F) Locations of point mutations in sad1 mutants. A schematic of the WT Sad1 gene is shown at the top in red. Exons are represented by boxes and introns by lines. The location of each mutation within the gene is indicated by a vertical line.
extracts from three of these mutants (358, 384, and 1023) (Fig. 1E).
The mutations in the remaining four mutants are located in regions that are likely to be important for protein structure and presumably lead to unstable proteins that are degraded (SI Appendix, Fig. S4). A schematic summarizing the nature and locations of all of the sad1 mutations is shown in Fig. 1F.

Conversion of S728 to F Results in the Formation of Tetracyclic Instead of Pentacyclic Triterpenes in Planta. We next examined the triterpene content of extracts from the root tips of seedlings of A. strigosa mutants 358, 384, and 1023. We expected to see loss of the SAD1 cyclization product β-amyrin with associated accumulation of the precursor OS. This result is indeed what we observed for the previously characterized sad1 mutant 109, a predicted premature termination of a translation mutant that does not produce SAD1 protein; also for mutant 358, suggesting that this mutant SAD1 variant is inactive (Fig. 2A and SI Appendix, Fig. S5). Surprisingly, however, a new compound was observed in root extracts of mutants 384 and 1023 that was not present in extracts from the wild-type or sad1 mutants 109 and 358 (Fig. 2A and SI Appendix, Fig. S5). The new compound had an elution profile and mass spectrum identical to dammaranediol-II (DM) (Fig. 2A and SI Appendix, Figs. S6 and S7). DM was not detectable in wild-type root extracts by GC-MS, although a more polar minor peak with an elution profile and mass spectrum consistent with that of epoxydammarane (epDM) was observed. Accumulation of DM in mutants of 384 and 1023 implies specificity of the downstream avenacin pathway enzymes for the β-amyrin scaffold.

Ginseng and other medicinal plants accumulate biologically active epoxydammarane saponins at levels as high as 5%, but the biosynthetic origin of the oxacyclic triterpenoid scaffold is not known (26). Stereoisomers of 2-epoxydammarane have been identified as cyclization products generated by the A. thaliana mixed product triterpene synthase AtLUP1 in yeast when fed with exogenous DOS (26). Although small peaks with similar elution times to epDM were observed in the mutant extracts, we were unable to detect epDM in these lines by GC-MS. Collectively, these data indicate that the S728F mutation results in a change in product specificity, converting SAD1 into an enzyme that yields primarily tetracyclic rather than pentacyclic cyclization products.

### Table 1. Sequence analysis of sad1 mutants

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Mutation event</th>
<th>Predicted amino acid change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Premature termination of translation:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1</td>
<td>G1912A</td>
<td>Tyr-165 Stop*</td>
</tr>
<tr>
<td>B1</td>
<td>G1912A</td>
<td>Tyr-165 Stop*</td>
</tr>
<tr>
<td>109</td>
<td>G3417A</td>
<td>Tyr-380 Stop</td>
</tr>
<tr>
<td>610</td>
<td>G1912A</td>
<td>Tyr-165 Stop*</td>
</tr>
<tr>
<td>1146</td>
<td>G4169A</td>
<td>Tyr-471 Stop</td>
</tr>
<tr>
<td>1293</td>
<td>G39A</td>
<td>Tyr-13 Stop</td>
</tr>
<tr>
<td>Predicted splicing errors:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>110</td>
<td>G6689A</td>
<td>—</td>
</tr>
<tr>
<td>225</td>
<td>G3302A</td>
<td>—</td>
</tr>
<tr>
<td>589</td>
<td>G3914A</td>
<td>—</td>
</tr>
<tr>
<td>1001</td>
<td>G4365A</td>
<td>—</td>
</tr>
<tr>
<td>Predicted amino acid substitutions:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>297</td>
<td>G3939A</td>
<td>Glu-419 Lys</td>
</tr>
<tr>
<td>358</td>
<td>G5234A</td>
<td>Cys-563 Tyr</td>
</tr>
<tr>
<td>384</td>
<td>C7249T</td>
<td>Ser-728 Phe*</td>
</tr>
<tr>
<td>532</td>
<td>G549A</td>
<td>Gly-121 Glu</td>
</tr>
<tr>
<td>599</td>
<td>G2809A</td>
<td>Gly-277 Glu</td>
</tr>
<tr>
<td>1023</td>
<td>C7249T</td>
<td>Ser-728 Phe*</td>
</tr>
<tr>
<td>1217</td>
<td>G2025A</td>
<td>Gly-203 Glu</td>
</tr>
</tbody>
</table>

*Identical mutation (G1912 → A). Although mutants A1, B1, and 109 all have a mutation at G1912, these mutants were isolated from different M2 families and so represent independent mutation events.

†Identical mutation (C7249 → T). Although mutants 384 and 1023 have both undergone a cytidine to thymidine change at C7249, these mutants were isolated from different M2 families and so represent independent mutation events.
Heterologous Expression of the S728F SAD1 Mutant Variant in Yeast. We then expressed the S728F SAD1 variant in yeast. cDNAs encoding the wild-type SAD1 protein and the mutant SAD1 variant were cloned into the yeast expression vector pYES2 under the control of a galactose-inducible promoter, expressed in the yeast strain GIL77 (gal2 hem3-6 erg7 ura3-167) (27) (SI Appendix, Fig. S8), and yeast extracts were analyzed by GC-MS (Fig. 2B). β-Amyrin was the major triterpene product detected when the wild-type SAD1 protein was expressed in yeast. epDM was also detected as a minor product (Fig. 2B). The S728F SAD1 mutant variant produced small amounts of β-amyrin by comparison. Unexpectedly, however, this variant generated a major peak that appeared to correspond to epDM, with only trace amounts of DM (Fig. 2B and SI Appendix, Fig. S9). DOS was also clearly detectable in extracts from yeast expressing the SAD1 mutant variant but not in those from the empty vector control or expressing the wild-type SAD1 protein. Thus, in yeast, the SAD1 mutant enzyme appears to preferentially cyclize DOS rather than OS, yielding predominantly epDM rather than DM (Fig. 3B). Accumulation of DOS in yeast expressing the SAD1 mutant variant may be an equilibrium effect due to pull through by DOS cyclization and could be suggestive of metabolome formation.

To confirm the identity of the putative epDM cyclization product, we grew a large-scale (1 L) culture of the yeast strain expressing the S728F SAD1 mutant variant and purified ~2 mg of this compound (Methods). The purified triterpene was examined by 1H-NMR.

![Diagram of triterpene synthesis](image-url)

**Fig. 3.** Cyclization reactions carried out by the WT and mutant triterpene synthases from oat and A. thaliana. (A) The WT SAD1 protein first converts OS to the tetracyclic dammarenyl cation and then traverses through a series of cations to give the oleanyl cation. The final step is the deprotonation of the oleanyl cation and the release of β-amyrin (BA). In contrast, the S728F SAD1 variant catalyzes an alternative cyclization reaction along the path indicated to give tetracyclic products. This mutant cyclase can accept both OS and DOS as substrates, cyclizing them to DM and epDM, respectively. Accumulation of OS can lead to reacceptance of OS as a substrate by endogenous squalene epoxidase, resulting in transformation of OS into DOS. The WT SAD1 enzyme is also able to accept DOS and cyclize it to epDM, but to a much lesser extent than the S728F SAD1 variant (~6% of total cyclic products in oat and ~3% of total cyclic products in yeast). (B) The cyclization reactions catalyzed by WT ATLUP1 and the ATLUP1-T729F variant are shown. WT ATLUP1 is also able to accept DOS, cyclizing it to 2-epDM epimers as opposed to the single epimer observed for SAD1. For ATLUP1-T729F, cyclization is predominantly DOS-mediated.
spectroscopy at 400 MHz in CDCl₃ solution (SI Appendix, Figs. S10 and S11). C-24$S$ or C-24$R$ epimers of the epoxidammaranes can be distinguished by comparing the $^1$H-NMR chemical shifts of H-24, Me-26, and Me-27 positions. Between C-20 and C-24, four different combinations of configurations are possible. Molecules with 20$R$, 24$R$ configuration are not known in nature. Three possible configuration pairs (20$S$, 24$S$; 20$R$, 24$S$; 20$S$, 24$R$) are known to occur in natural products; chemical shifts of assignable resonances are listed in SI Appendix, Table S1. Because epDM, which is made together with DM, has 20$S$ configuration, it was anticipated to maintain the $S$ configuration at C-20. Chemical shifts at the diagnostic protons (H-24, Me-26, Me-27) confirmed epDM has 20$S$, 24$S$ configuration. Chemical shifts and coupling constants at H-24 vary dramatically for molecules with locally diastereomeric configurations at C-20 and C-24. For example, epDM with 20$R$, 20$S$, at H-24 $\delta$ 3.73 (dd, $J$ = 7.7, 6.9) was similar to the skeleton with 20$S$, 24$R$ at H-24 $\delta$ 3.73 (dd, $J$ = 7.5, 7.5). For an arrangement with both stereocenters $S$, H-24 appeared at $\delta$ 3.639 (dd, $J$ = 10.1, 5.3), a signal similar to that of the yeast-derived epDM [H-24 $\delta$ 3.64 (dd, $J$ = 9.9, 5.4)]. Based on this shift and coupling data, together with other spectral attributes, the epDM product generated by the S728F mutant variant of SAD1 was assigned the configuration 20$S$, 24$S$ and designated as (3$S$, 20$S$, 24$S$)-20,24-epoxydammarane-3,25-diol (Fig. 3A).

Homology Modeling. To further investigate the likely impact of the amino acid substitutions that we had observed on SAD1 stability and function, we generated a homology model of SAD1 by using the human lanosterol synthase crystal structure (28) as a template. The locations of the seven predicted SAD1 amino acid substitutions are shown in Fig. 4B. The C563Y mutation in line 358 results in stable but inactive protein (Figs. 1 and 2). This mutation affects
a residue close to the catalytic aspartate D484 of the conserved DCTAED motif (Fig. 4A). This conserved aspartate has been implicated in oxidosqualene cyclase function in *Euphorbia tirucalli* β-amyrin synthase (7). C485 (Fig. 4C) is known to be required for the function of the oxidosqualene cyclase lanosterol synthase in *Saccharomyces cerevisiae* (29). C563 has been proposed to have a role in initiation of cyclization based on its proximity to D484, following the determination of the human lanosterol synthase crystal structure (28), but its function has not been tested. The C563Y substitution introduces a bulky tyrosine residue that is likely to interfere with the hydrogen bonding interaction critical for lowering the pKb of D484 to facilitate protonation of the epoxide group (Fig. 3C) (30). We predict that this substitution will inhibit reaction initiation, which would inactivate cyclization.

The wild-type SAD1 protein converts OS to the pentacyclic cyclization product β-amyrin through a series of cationic intermediates (Fig. 3A). Protein modeling and docking analysis predict that in the S728F variant, the aromatic side chain of the newly introduced phenylalanine together with that of F725 sandwich the carboxcational center of the dammarenyl cation (Fig. 4D). As a result, ring expansion of the dammarenyl cation is likely to be compromised, giving rise to formation of a truncated tetraacyclic OS cyclization product (Fig. 3A). This product is presumably converted to the triterpene derived from C-20 cation by a water molecule. The wild-type SAD1 protein is able to generate low levels of epDM in oat and in yeast (Fig. 2 A and B), suggesting that although OS is its preferred substrate, it is capable of using DOS. In contrast, the S728F mutant SAD1 variant preferentially accepts DOS as a substrate when expressed in yeast and cyclizes this to epDM (Fig. 3B). As speculated in Shan et al. (26), side-chain rotation of the distal epoxide to the carboxcation at C-20 may enable movement of a positive charge from carbon to oxygen and attack of a nearby water molecule to give an alcohol at C-25. The epoxide-containing side chain of the dammarenyl cation is therefore free to rotate within the active site to participate in cyclization. S728 is located close to a loop containing IS54, T260, and Y264. This loop is believed to block the opening to the substrate access channel and undergo a conformational change to allow substrate entry to the active site (Fig. 4D). The S728F amino acid substitution may alter the hydrogen-bonding network of this loop and control access of the novel DOS (OS and DOS) into the active site. However, further experiments are required to investigate this mechanism in detail. The addition of water in triterpene cyclization is known for the triterpene synthases PgPNA, AtLUP1, and AtPEN1 (ARAB) (31–35). Unlike AtLUP1, which makes two stereoisomers of epDM (20R, 24S and 20S, 24S), S728F accumulates only 20S, 24S-epDM in a highly stereospecific manner. Cyclization of the side chain of DOS occurs in a manner consistent with cation quenching by a positioned water molecule, resulting in a reaction that is more stereospecific in S728F than in the AtLUP1 enzyme.

We next investigated the effects of amino acid substitution at the corresponding position in the *A. italicana* triterpene synthase AtLUP1 (Thr729; Fig. 4A and SI Appendix, Figs. S11). Wild-type AtLUP1 and a mutated version in which Thr729 had been converted to phenylalanine (T729F) were expressed in yeast. Wild-type AtLUP1 produced the OS-derived pentacyclic triterpenes lupeol and lupanediol as the major cyclization products, and also the epoxydammarenadiols epDM (20R, 24S) and 17,24-exopoybaccharane diol from DOS as minor products (Figs. 3B and 5). Interestingly, the AtLUP1-T729F mutant yielded primarily epoxydammarendiol products (Figs. 3B and 5 and SI Appendix, Figs. S13 and S14). Thus, the T729F mutation in AtLUP1 leads to a change in product specificity from pentacyclic to tetraacyclic triterpenes, and also to preferential cyclization of DOS instead of OS in yeast as seen for the SAD1 S728F mutant variant. As seen for SAD1 S728F, yeast strains expressing the mutant form of AtLUP1 also accumulated elevated levels of DOS (SI Appendix, Fig. S15). The S728F SAD1 mutant enzyme shows high stereoselective cyclization of DOS to the epDM-20S, 24S epimer, whereas AtLUP1 is less selective, generating two epimers and other additional minor products. Elucidation of the crystal structures of the wild-type and mutant forms of these enzymes may in the future enable the exact nature of this control to be understood in more detail.

**Discussion**

In triterpenoid biosynthesis, cyclization of OS is the key step that determines the nature of the triterpene scaffold. The enzymes that catalyze this process—triterpene synthases—belong to multigenic families in plant genomes. Here, we have shown that a single-point mutation that causes an amino acid substitution close to the active site dramatically alters product specificity in both SAD1 and LUP1, so uncovering hidden functional diversity in these triterpene synthases. It is conceivable that nature explores alternate modes of cyclization through such single mutational steps, as has been suggested for diterpene synthases (36–43). A dedicated cyclase that makes epDM as its major product has not been reported before our work to our knowledge. Triterpene glycosides based on the DM and epDM skeletons have important pharmaceutical and antibacterial properties (44–46). Our results open up the possibility of manipulating both the nature of the precursor and the product specificity of the cyclization process for the production of diverse and novel triterpenes. They further demonstrate the power of forward genetics screens in plants for elucidating the enzymatic mechanisms of this versatile and fascinating group of enzymes.

**Methods**

**Plant Material.** Wild-type and mutant *A. strigosa* lines were grown as described (13). Transcript, protein, and triterpene analysis were carried out by using root tips (terminal 0.5 cm) from 3-d-old seedlings, and genomic DNA was extracted from 6-d-old seedlings.

**DNA, RNA, and Protein Analysis.** Genomic DNA was isolated from 6-d-old seedlings of *A. strigosa* by using the DNeasy Plant Mini kit (Qiagen). The wild-type and mutant forms of the Sad1 gene were amplified in four segments. Purified DNA segments from each of the sad1 mutants were sequenced by using sets of primers along the length of the Sad1 gene, and each mutant was sequenced to at least twofold coverage. Primers used for *Sad1* amplification and sequencing are shown in SI Appendix, Table S2. Genomic DNA samples from each mutant were sent for Diversity Array Technology (DArT) analysis (Diversity Arrays Technology Pty Ltd) to confirm that each mutant line was independent (47).

For analysis of transcripts in oats, total RNA was extracted from 0.5-cm root tips of 3-d-old oat seedlings by using TRI-REAGENT (Sigma catalog no. T9424) and the extract was treated with DNase I (Roche). For RT-PCR, cDNA was synthesized by using 1 μg of DNase-treated RNA. First-strand cDNA synthesis was carried out by using SuperScript II Reverse Transcriptase (Invitrogen) according to the manufacturer’s instructions, and cDNA was amplified by PCR. For Northern blot analysis, 10 μg of total RNA was used. RNA was separated on a 1.2% (wt/vol) agarose/0.25 M formaldehyde gel and transferred to a Hybond-N+ nylon membrane (Amersham) overnight. cDNA probes were labeled with 32P-CTP by using the Rediprime II Random Prime Labeling Kit (Amersham). Hybridizations were carried out overnight at 65 °C in 10 mL of Church Buffer containing 0.1 mg/mL salmon sperm DNA (Sigma) and 50 μL of 103P-CTP labeled probe. The membrane was exposed to a BAS-1000 Imaging plate (Fuji) overnight and imaged by using a Typhoon 9200 Variable Mode Imager (Amersham).

For protein and immunoblot analysis of oat root protein, total protein was extracted from 0.5-cm root tips of 3-d-old oat seedlings. Root tips were ground in protein extraction buffer (50 mM Tris HCl pH 7.5, 150 mM NaCl, 5 mM EDTA, 10% [vol/vol] glycerol, 1% [wt/vol] PVPP, 1% [vol/vol] Triton X-100 (Boehringer Mannheim), 1× Complete protease inhibitor Cocktail (Roche)) for 1 min with a plastic pestle followed by incubation at 4 °C for 2 h. Proteins were denatured, separated on NuPAGE gels (4–12% acrylamide gel) (Invitrogen), and blotted onto nitrocellulose membranes (Bio-Rad) by using the manufacturer’s protocol. Membranes were probed with anti-SAD1 antisera (1:10,000 dilution)
were run on either a HP-5MS column (30 m × 0.25 mm i.d., 0.25–μm film) (Agilent) or a 2B-SHT column (35 m × 0.25 mm i.d., 0.10-μm film) ( Phenomenex) by using an Agilent 7890B GC machine. The injector port, source, and transfer line temperatures were set at 250 °C; an oven temperature program from 80 °C (2 min) to 290 °C (30 min) at 20 °C/min was used. The carrier gas was helium; the flow rate was 1.2 mL/min. Samples were injected in splitless mode with either a 1-μL or a 3-μL sample volume. The output was used to search the NIST8 library to assign identity to peaks in the GC-MS traces. Product abundance was calculated as the percentage of total cyclic products using integrated peak areas. All experiments were repeated to confirm reproducibility of the triterpene profiles of the wild-type and mutant samples.

**Triterpene Standards.** Dammarenediol-ii (catalog no. CFN94976, 98% HPLC pure) was purchased from Wuhan ChemFases Biochemical Co. Ltd, China, and β-Amyrin and cycloartenol from Extrasynthese. The standards were dissolved and diluted to 0.5 mg/mL in hexane before derivatization and GC-MS.

**ATLUP1 Cloning and Site-Directed Mutagenesis.** ATLUP1 (AT1G78970)-ORF was amplified from total cDNA of young A. thaliana (Col-0) seedlings by PCR using Gateway primers (SI Appendix, Table S2) and cloned into pDONOR207. The T729F mutation was created by site-directed mutagenesis using pDONOR207-ATLUP1 as the template. The oligonucleotide strategy and conditions for PCR amplification followed those described (48). The oligonucleotides used for site-directed mutagenesis are listed in SI Appendix, Table S2.

**Yeast Cloning and Expression.** All cloning and expression analysis was carried out in the yeast strain GIL77 (gal2 hem3-6 erg7 ura3-167) (32). Expression vectors were constructed by using in vivo homologous recombination in yeast. The ORFs of the wild-type Sad1 gene, ATLUP1 (AT1G78970), and mutant variants were amplified from pDONOR207 entry vectors by using the oligonucleotides for yeast cloning shown in SI Appendix, Table S2. Each primer contained a region that overlapped with the pYES2 vector sequence (the 5’ end of the forward primer overlapped with the GAL1 promoter sequence, and the 5’ end of the reverse primer with the CYC1 terminator sequence). The 3’ ends of the primers matched the beginning and end of the Sad1 ORFs. The ORF of the wild-type and mutant lines were amplified by using these primers, and the PCR fragments obtained were cotransformed into GIL77 strain along with XbaI/Hindlll-linearized pYES2 vector. Yeast transformation was performed by using standard protocols (Yeastmaker Yeast transformation system 2, Clontech Laboratories). This resulted in in vivo recombination between the pYES2 vector and the Sad1 ORFs. Plasmids were recovered from yeast, transformed into E. coli, and checked by sequencing.

For expression analysis, yeast strains were grown at 28 °C in 5-mL cultures in selective medium [SD-URA+ 2% (wt/vol) glucose + supplements] until saturation (~2 d). The supplements used were as follows: ergosterol (Fluka), 20 μg/mL; hemin (Sigma-Aldrich), 13 μg/mL; and Tween-80 (Sigma-Aldrich), 5 mg/mL. Cells were then pelleted, washed in ddH2O, transferred to incubation medium [SD-URA+ + 2% (wt/vol) galactose], and incubated at 30 °C for further 2 d to allow accumulation of triterpenes. They were then pelleted and washed once with ddH2O before triterpene extraction as described for oat roots.

For protein analysis, yeast cells were resuspended in protein extraction buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM EDTA, 10% (vol/vol) glycerol, 1% (wt/vol) PVPP, 1% (vol/vol) Triton X-100 (Boehringer Mannheim), 1x Complete protease inhibitor (Roche)) and lysed by using a French press (with two passes at 1125 p.s.i., 4 °C). The preparations were then incubated on ice for 2 h and then centrifuged at 21,130 × g for 20 min. The supernatants were used for protein and Western blot analysis as described previously (16).

**Purification and Structural Elucidation of** (35, 20S, 24S)-20,24-epoxydammarane-3, 25-diol. epDM was extracted from a 1-L culture of a yeast transformant expressing the SAD1 S728F variant. Cells were pelleted, rounded, and then extracted by using the methods described above for oat roots. The organic residue was loaded onto a silica gel column 10 cm long and 0.5 cm in diameter (LC60A35-70 μm; Fluorochem) in a Pasteur pipette that had been prequillibrated with an ethyl acetate/hexane (1:9) solvent system. The column was washed with 5–6 column volumes of 1:9 ethyl acetate/hexane to remove oxidosqualene, dioxidosqualene, and other nonpolar yeast components. Next, the solvent was switched in a step gradient to 1:6 and then 1:4 ethyl acetate/hexane, and 0.5-mL fractions were collected and analyzed by TLC. Fractions containing epDM were combined and dried in a rotary evaporator. The purity of the compound was assayed by using

---

**Extraction and Analysis of Triterpenes from Oat Roots.** The triterpene content of oat root tips was analyzed by TLC and GC-MS. Root tips (~50 per line) were ground, mixed with 0.5 mL of saponification reagent [20% (wt/vol) KOH in 50% (vol/vol) ethanol], and incubated at 65 °C for 2 h before extraction with an equal volume of hexane. The extraction step was repeated twice more to maximize triterpene recovery. The extract was then dried down, and the residue dissolved in 500 μL of hexane. For rapid qualitative analysis, extracts were run on TLC plates (Silica gel on Al foil, 10 cm × 5 cm, FLU.K.A, catalog no. 70644) by using a hexane:ethyl acetate (6:1) solvent system. Compounds were visualized by spraying the plates with acetic acid: H2SO4: p-anisaldehyde-hyde (48:1:1 vol/vol) and heating to 120 °C for 5 min on a TLC plate heater. For GC-MS analysis, 100-μL aliquots of hexane extract were dried down and the residues were resuspended in 100 μL of Tri-Sil Z reagent (Sigma, catalog no. 92718) before incubating at 65 °C for 30 min in a dry heat bath. Samples

---

**Fig. 5.** Total ion chromatograms of yeast extracts expressing the WT and mutant triterpene synthases. In ATLUP1, lupeol and lupanediol are derived from OS cyclization, whereas epDM isomers (20S, 24R and 20S, 24S) and 17, 24-epoxybaccharane diol are derived from DOS cyclization. For the SAD1 mutant variant S728F (384), the epDM isomer-20S, 24S is derived from DOS. Peaks with red arrows indicate OS-derived cyclization products, and ones with green indicate DOS-derived cyclization products.

(16) followed by detection with a goat anti-rat IgG horseradish peroxidase-labeled secondary antibody (Sigma-Aldrich) according to the manufacturer’s protocol.
The orientation and position of SAD1 relative to a virtual membrane were calculated using the program MACPHERSON. The models obtained were visualized by using PyMOL (53). Protein sequences were aligned by using Clustal W, and sequence features were viewed and annotated manually using functional information available for human lanosterol synthase (28).

**Homology Modeling and Sequence Alignments.** For homology modeling of SAD1, human lanosterol synthase was used as a template (PDB ID code: 1W6K) to generate a model using Modeller (49). The models obtained were subjected to stereochemical validation by using Prosa II (50), Prove (51), and Procheck (52). Models were visualized by using PyMOL (53). Protein sequences were aligned by using Clustal W, and sequence features were viewed and annotated manually using functional information available for human lanosterol synthase (28).


**Acknowledgments.** This work was supported by European Union Grant KBBE-2013-7 (TriforC), the Biotechnology and Biological Sciences Research Council (BBSRC) Phytogenic Grant Understanding and Exploring Plant and Microbial Metabolism BB/J004561/1, the John Innes Foundation (A.O., R. E. Melton, R.K.H., and P.E.O.), and a Norwich Research Park studentship award (to M.S.). E. R. Minto is grateful for sabbatical leave provided by Indiana University-Purdue University, Indianapolis.

8 of 8


Salmon et al.