the plant journal

SEB Society for Experimental Biology

The Plant Journal (2019) **98,** 928–941 doi: 10.1111/tpj.14272

TECHNICAL ADVANCE

Optimized small-molecule pull-downs define MLBP1 as an acyl-lipid-binding protein

Yelena Sterlin[1], Oded Pri-Tal[1], Gil Zimran[1], Sang-Youl Park[2], Julius Ben-Ari[1], Jiorgos Kourelis[2], Inge Verstraeten[2], Maayan Gal[3,4], Sean R. Cutler[2] and Assaf Mosquna[1,*]

[1]*The Robert H. Smith Institute of Plant Sciences and Genetics in Agriculture, the Hebrew University of Jerusalem, Rehovot 7610001, Israel,*

[2]*Department of Botany and Plant Sciences, Center for Plant Cell Biology and Institute for Integrative Genome Biology, University of California, Riverside, CA 92521,USA,*

[3]*Biochemistry Department, MIGAL-Galilee Research Institute, Kiryat-Shmona 11016, Israel, and*

[4]*Faculty of Sciences and Technology, Tel-Hai Academic College, Upper Galilee 1220800, Israel*

Received 3 October 2018; revised 2 January 2019; accepted 29 January 2019; published online 8 February 2019.
*For correspondence (e-mail assaf.mosquna@mail.huji.ac.il).

SUMMARY

Abscisic acid (ABA) receptors belong to the START domain superfamily, which encompasses ligand-binding proteins present in all kingdoms of life. START domain proteins contain a central binding pocket that, depending on the protein, can couple ligand binding to catalytic, transport or signaling functions. In Arabidopsis, the best characterized START domain proteins are the 14 PYR/PYL/RCAR ABA receptors, while the other members of the superfamily do not have assigned ligands. To address this, we used affinity purification of biotinylated proteins expressed transiently in *Nicotiana benthamiana* coupled to untargeted LC-MS to identify candidate binding ligands. We optimized this method using ABA–PYL interactions and show that ABA co-purifies with wild-type PYL5 but not a binding site mutant. The K_d of PYL5 for ABA is 1.1 μM, which suggests that the method has sufficient sensitivity for many ligand–protein interactions. Using this method, we surveyed a set of 37 START domain-related proteins, which resulted in the identification of ligands that co-purified with MLBP1 (At4G01883) or MLP165 (At1G35260). Metabolite identification and the use of authentic standards revealed that MLBP1 binds to monolinolenin, which we confirmed using recombinant MLBP1. Monolinolenin also co-purified with MLBP1 purified from transgenic Arabidopsis, demonstrating that the interaction occurs in a native context. Thus, deployment of this relatively simple method allowed us to define a protein–metabolite interaction and better understand protein–ligand interactions in plants.

Keywords: abscisic acid, ABA signaling, *PYR/PYL/RCAR*, protein–metabolite interactions, *in planta*, START domain, Bet v 1, technical advance, polyketide cyclase-like proteins, technical advance.

INTRODUCTION

In land plants, mechanisms that enable adaptive responses to stresses involve the perception of abscisic acid (ABA) by the pyrabactin resistance1/pyrabactin resistance1-like/regulatory components of ABA receptors (PYR1/PYL/RCAR) (Ma *et al.*, 2009; Park *et al.*, 2009). The ABA receptors belong to the START domain superfamily (Ma *et al.*, 2009; Park *et al.*, 2009). The START domain represents an ancient protein signature originally identified as a lipid-binding domain superfamily present in a broad range of cellular proteins across diverse species (Ponting and Aravind, 1999; Schrick *et al.*, 2004). The name START is an acronym for StAR-related lipid transfer domain, with StAR referring to the steroidogenesis acute regulatory protein involved in the transport of cholesterol to the inner mitochondrial membrane (reviewed in Stocco, 2001). Crystal structures solved for mammalian MLN64 and STARD4 proteins revealed an α-helix/β-grip fold forming a hydrophobic cavity (Tsujishita and Hurley, 2000; Romanowski *et al.*, 2002). This cavity was later predicted to be shared by all the members of the START domain superfamily and was suggested to be involved in ligand binding (Iyer *et al.*, 2001). Evidence for the ability of the START domain to accommodate ligand molecules and its significance derives mainly

from animal proteins, with only a few having been identified in plants. For instance, mammalian StAR binds and transports cholesterol from the outer to the inner mitochondrial membrane for its conversion to biologically active steroid hormones (reviewed in Stocco, 2001). Human phosphatidylcholine transfer protein (PC-TP) specifically interacts with PC to enhance the activity of other enzymes involved in the regulation of cellular fatty acid metabolism (Westerman *et al.*, 1983; Roderick *et al.*, 2002; Ersoy *et al.*, 2013; Kawano *et al.*, 2014) and silkworm carotenoid-binding protein (CBP) manifests binding of carotenoid lutein in a 1:1 ratio and plays a role in the yellow pigmentation of the cocoon (Tabunoki *et al.*, 2002, 2004). Binding of ligands within the START domain cavity has been attributed to the shape and size of the hydrophobic pocket (Iyer *et al.*, 2001). However, it should be pointed out that the residues forming the pocket are not strictly hydrophobic and also accommodate amino acid residues forming polar interactions (Iyer *et al.*, 2001), thus extending the chemical nature of potential ligands.

Genes encoding the START domain and their structural relatives are present in genomes of bacteria, protists, yeast, plants and animals, (Schrick *et al.*, 2004; Gatta *et al.*, 2015), with an over-representation in plant genomes (Schrick *et al.*, 2004). For example, while the genomes of *Drosophila melanogaster* and *Caenorhabditis elegans* encode only a few START domain proteins, the *Arabidopsis thaliana* and rice genomes encode a few dozen (Schrick *et al.*, 2004). Sequence homology suggests birch pollen allergen (Bet v 1) and polyketide cyclase protein subfamilies to be among the closest neighbors of START domain proteins (Iyer *et al.*, 2001). In plants, besides the minimal START domain, many START proteins incorporate additional signaling domains such as a homeodomain (HD), a pleckstrin homology (PH), transmembrane (TM) segments and DUF 1336, a domain with an unknown function (Schrick *et al.*, 2004; Satheesh *et al.*, 2016). Based on the co-occurrence of START with signaling domains, it has been hypothesized that the START proteins may participate broadly in ligand-mediated signaling (Ponting and Aravind, 1999). This hypothesis was reinforced by the discovery that PYR/L proteins serve as ABA receptors (Ma *et al.*, 2009; Park *et al.*, 2009). The PYR/L proteins function through type 2C protein phosphatases (PP2C) and SNF1-related protein kinases (Ma *et al.*, 2009; Park *et al.*, 2009; Umezawa *et al.*, 2009; Vlad *et al.*, 2009). Binding of ABA to PYR/L receptors serves as a molecular switch, altering receptor conformation and increasing their affinity towards PP2Cs. Interaction between PYR/L receptors and PP2Cs inactivates the latter, leading to the activation of SnRK2s, which, in turn, results in regulation of plant development, stress-induced gene expression and physiological changes in stomatal closure (Choi *et al.*, 2000; Uno *et al.*, 2000; Negi *et al.*, 2008; Vahisalu *et al.*, 2008; Geiger *et al.*, 2009; Ma *et al.*, 2009; Melcher *et al.*, 2009; Miyazono *et al.*, 2009; Park *et al.*, 2009; Umezawa *et al.*, 2009; Yin *et al.*, 2009). Because the previously mentioned START domain protein ligands are indispensable for proper functioning of their protein counterparts in both animals and plants, and since plant START domain proteins have been demonstrated to be implicated in essential physiological processes (Rerie *et al.*, 1994; Kubo, 1999; Abe *et al.*, 2003; Prigge *et al.*, 2005; Nakamura *et al.*, 2006; Takada *et al.*, 2013), we decided to optimize a method to allow identification of START domain and START domain-related protein ligands *in planta*.

Tagore *et al.* (2008) described an *in vitro* small-molecule pull-down technique for the identification of protein–metabolite interactions. Previously, Schrick *et al.* (2014) immunoisolated the Arabidopsis HD-STARTs GLABRA2 and protodernal factor 2 from yeast cells and identified a range of metabolites that interact with these proteins, with phospholipids dominating over the other small molecular interactors. Herein, we modified the method of Tagore *et al.* (2008) to screen for endogenous START/Bet v/polyketide cyclase-like protein ligands *in planta* focusing on compounds with an amphiphilic/lipophilic nature. We used PYR/L proteins for optimization of the method, as they are the best-characterized Arabidopsis START domain-related proteins with identified endogenous ligands (Melcher *et al.*, 2009; Santiago *et al.*, 2009a). The overall effort resulted in a simple and time-efficient method for screening of endogenous small molecules found in association with START domain proteins. The technique facilitated the discovery of two START domain-related protein ligands with as yet unknown functions. Although the method was only tested with START domain-related proteins, it was designed as a general tool for identifying protein–small molecule interactions. Thus, by considering chemical nature of a putative ligand, it can be applied to other soluble proteins harboring ligand-binding domains.

RESULTS AND DISCUSSION

Co-purification of ABA with PYR/L receptors *in planta*

Optimization of the small-molecule pull-down was performed with PYL5 as the biochemical properties of the PYR/L-ABA-PP2C complex are well defined (Melcher *et al.*, 2009; Santiago *et al.*, 2009a,b). The dissociation constants of PYL5 (1.1 μM) and PYL5 in the presence of a PP2C (38 nM) provide the range of affinity of the current ligand–protein purification approach (Santiago *et al.*, 2009a). Formation of the PYL5–ABA–PP2C complex offers the experimental means for the verification of protein activity throughout the purification process. In addition, the mechanistic information about the key residues required for ligand binding is well known (Melcher *et al.*, 2009), facilitating a controlled experiment. Moreover, ABA is efficiently

ionized and can thus be detected by mass spectrometry (MS) (López-Carbonell and Jáuregui, 2005; Ma *et al.*, 2008; Pan *et al.*, 2008, 2010; Balcke *et al.*, 2012; Yu *et al.*, 2014). Taken together, we decided to optimize the method using PYL5 co-expressed with HAB1 PP2C.

Complementary DNAs of receptor and phosphatase were co-expressed in *Nicotiana benthamiana* leaves with the receptor fused to an *HPB tag* (Qi and Katagiri, 2009) and the phosphatase fused to GFP. Purification of PYL5 fused to HPB benefits from high biotin–strepavidin binding affinity ($K_d = 10^{-15}$ M) (Green, 1963), which limits the degree of ligand co-purification due to competition with unbound receptors.

To purify the protein–ligand complex, transfected *N. benthamiana* leaves were homogenized and the lysate was further ultra-centrifuged and filtered to get rid of insoluble plant debris to avoid potential clogging of the LC column. The protein of interest was enriched by affinity chromatography using streptavidin Sepharose affinity medium (Figure 1a,b). The next experimental phases were guided by the ligands of interest. The ABA receptor pocket, as well as the putative binding sites of other START, Bet v and polyketide cyclase-like proteins, are mostly hydrophobic (Ponting and Aravind, 1999; Iyer *et al.*, 2001), thus defining the chemical properties of potential ligands and the washing steps necessary to discriminate other molecules. Correspondingly, the elution process should denature the protein while allowing ligand solubilization. On this basis, the washing buffer was supplemented with low levels of detergent and salt, to both remove unbound/non-specifically bound components and to strengthen hydrophobic interactions (Hofmeister, 1988). To prevent interference of the detergent with downstream LC, additional washes were performed with salt only. A number of organic solvents could be used for elution. However, considering that successive chromatographic separation was to be carried out with an acetonitrile (ACN)/water gradient, elution was performed with 60–80% ACN (Figure 1b,c).

The eluate was then subjected to LC-MS. Raw data were analyzed using Thermo Scientific™ SIEVE software which enables semi-quantitative differential analysis of isolated small molecules by identifying components with statistically significant differences in inter-sample abundance (Figure 1c).

As a negative control, a parallel experiment was performed using different START/Bet v/polyketide cyclase-like proteins or GFP. The assay was carried out using three independent biological repeats, each consisting of leaf tissue pooled from four different *N. benthamiana* plants. Expression and purification of the proteins was confirmed by SDS-PAGE and Western blotting (Figure S1a,b in the online Supporting Information). To verify that PYL5 retained its activity throughout the purification process, a small portion of the purified sample was tested for the presence of HAB1 co-receptor (Figure S2), indicating the formation of a biologically active complex.

Statistical analysis of inter-sample differences of ion enrichment in eluates corresponding to PYL5, other START/Bet v/polyketide cyclase-like proteins or GFP revealed PYL5-specific ions matching ABA (*m/z* 263.1284, RT 6.6) $P < 0.05$ (Figure 2a). Recovery of this specific ion was reproducible, as shown by the consistent results of several independent purification events (Figure S3a–c) and by eluate analysis using different MS instruments (see Experimental procedures). The identity of the ABA ions was verified by spiking eluates corresponding to PYL5 samples with an ABA standard (Figure S4) and by MS–MS fragmentation analysis of both the standard and the sample (Figure S5a,b).

The described transient expression in *N. benthamiana* and subsequent protein–ligand co-purification offers a tool for high-throughput identification of protein–ligand complexes. However, established model systems like *A. thaliana* provide genetic and bioinformatics infrastructures that enable/ease subsequent ligand identification (e.g. auxotrophic mutants in a given metabolic pathway may narrow down the range of possible molecule ligands). To test the protein–ligand co-purification protocol in *A. thaliana*, we generated transgenic plants stably expressing PYL5 and two transgenic controls expressing PYL5 K87A, a mutant in one of the key ABA-binding residues (Melcher *et al.*, 2009), or GFP (Pri-Tal *et al.*, 2017). Expression of the proteins was driven by 35S and tagged by *HPB.* Transgenic plants were grown to the bolting stage and protein was isolated from leaf tissue via biotin–streptavidin purification, as described above. The efficiency of expression and purification was verified via SDS-PAGE and Western blotting (Figure S1c) and eluates were analyzed by LC-MS. Statistical analysis of the total ion content indicated that a unique ion corresponding to ABA (*m/z* 263.12, RT 6.5) was present in PYL5 eluates but not in any of the PYL5 K87A or GFP samples (*P*-value 0.00005; Figure 2b). Hence, the same method can be applied for co-purification of the PYL5–ABA complex whether the receptor is transiently expressed in *N. benthamiana* or stably expressed in *A. thaliana*.

Discovering candidate ligands of the *At4G01883* locus encoding MLBP1

In pursuit of ligands for *A. thaliana* START domain-related proteins, 66 HPB-tagged START/Bet v/polyketide cyclase-like proteins for which no ligand had been previously assigned were transiently expressed in *N. benthamiana* (Tables S1 and S2). Of the 66 tested proteins, 37 showed high accumulation and were subsequently purified and co-eluted with small molecules and analyzed via LC-MS (Table S2). Each ligand-screening experiment consisted of four different START domain-related proteins, each represented by three biological repeats. Every protein was

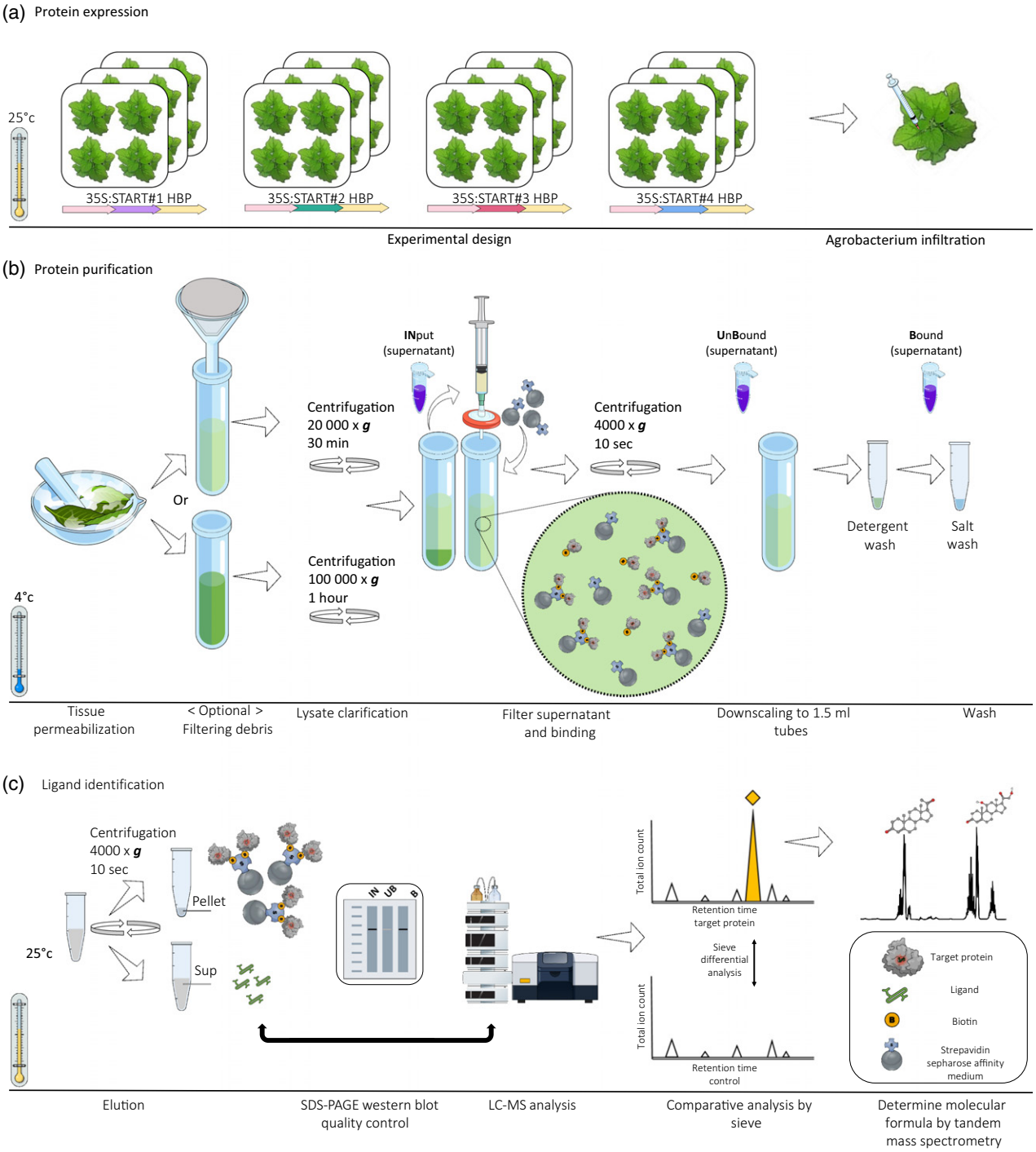
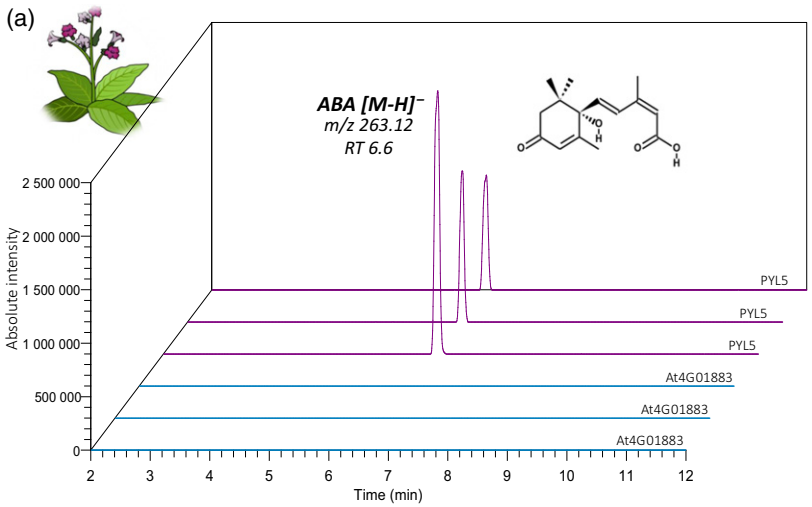


Figure 1. Experimental workflow.
(a) Protein expression: each one of four START proteins was expressed via *Agrobacterium* infiltration into a triplicate of four *Nicotiana benthamiana* plants.
(b) Protein purification: leaves were grand to powder in extraction buffer. The native soluble fraction was obtained by ultracentrifugation or via Miracloth filtering with subsequent centrifugation. An aliquot of supernatant was dissolved in Laemmli sample buffer to test target protein expression (IN). Supernatant containing target protein–metabolite complexes was filtered and incubated with streptavidin Sepharose affinity medium. Affinity medium with bound protein–metabolite complexes was harvested by centrifugation. Affinity medium was washed once with extraction buffer and twice with NaCl. The second and third aliquots were sampled to test the binding efficiency of unbound (UB) and bound (B) target protein. Aliquots were analyzed by Western blot.
(c) Ligand identification: putative ligands were eluted and subjected to LC-MS analysis. SIEVE software was used to detect unique ions by statistical analysis. Tandem mass spectrometry was deployed to unravel molecular structures of putative ligands of interest.



Experiment Screening	Ion in ABA receptor replicates	m/z	RT	P-Value	Control	Fold change	Mass-Analyzer	Figure
I	**$[C_{15}H_{20}O_4 - H]^-$**	**263.12**	**6.6**	**1E-05**	**At4G01883**	**infinite**	**Q Exactive**	**current**
I	$[C_{15}H_{20}O_4 - H]^-$	263.12	6.6	1E-05	At1G35260	infinite	Q Exactive	
I	$[C_{15}H_{20}O_4 - H]^-$	263.12	6.6	1E-05	GFP	infinite	Q Exactive	
II	$[C_{15}H_{20}O_4 - H]^-$	263.12	6.6	0.0272	At4G01883	369.76	Q Exactive	S2a
III	$[C_{15}H_{20}O_4 - H]^-$	263.12	6.9	1E-05	At4G01883	infinite	LTQ Orbitrap	S2b
IV	$[C_{15}H_{20}O_4 - H]^-$	263.12	6.9	0.0147	At4G01883	5,684.05	LTQ Orbitrap	S2c

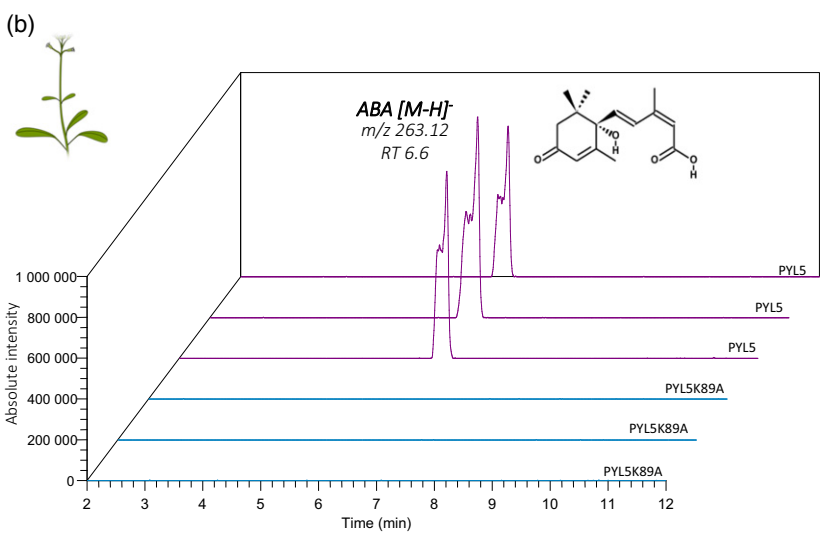


Figure 2. Co-purification of ABA ligand with the PYL5 receptor.
Negative mode LC-MS extracted mass chromatograms (EMCs) (*m/z* range 263.12537–263.13063) of three biological replicates of co-eluted small molecules purified with GFP-PYL5-HPB (AT5G05440), which was (a) transiently co-expressed with HAB1-GFP (AT1G72770) in *Nicotiana benthamiana* or (b) stably expressed in *Arabidopsis thaliana*, are indicated in violet. Blue graphs represent EMCs of eluates corresponding to (a) At4G01883-HPB or (b) PYL5K89A-HPB which served as controls. Molecular ion, *m/z*, retention time (RT) and the structural formula of the detected ABA molecule are indicated above the peaks. Statistical analyses (Student's *t*-test) of current and additional independent experiments are presented in the table (in panel a), summarizing *P*-value and fold change of the identified ion compared with the relevant control. Additionally, the table includes the chemical formula of the ions, *m/z*, RT and the model of mass spectrometer used in the specified experiment. Data displayed in the current figure are highlighted in bold. Corresponding figures of the additional experiments are indicated at the last column.

tested for a unique co-purified ligand with respect to the other three samples. We reasoned that comparison with different START/Bet v/polyketide cyclase-like proteins would be superior to using non-START domain-related protein controls (i.e. GFP), since the former share a hydrophobic cavity, thus presenting a similar background signature (Iyer *et al.*, 2001). Among the different samples analyzed by SIEVE, eluates from two samples corresponding to *At4G01883* and *AT1G35260* protein-encoding loci were associated with unique ions (*m/z* 411.27, RT 12.7, *P*-value 0.01–0.04 and *m/z* 295.22, RT 12.4, *P*-value 0.001–0.04, respectively) (Figures 3a and S6). To further elucidate ligand identity, we focused our efforts on the ion co-eluting with At4G01883, which of the two is phylogenetically closer to ABA receptors. Repeated purifications with At4G01883 and PYL5 or GFP, demonstrated a mean 5–25-fold change in ion enrichment and a *P*-value ranging from 0.03 to 0.04 (Figures 3a and S7a,b). It should be noted that the weak signal corresponding to the unique ion (m/z 411.27 RT 12.7), visible in both the PYL5 and GFP eluates, is a background signal (Figure 3, lower table). Next, co-elution of At4G01883 with the above ion was tested in *A. thaliana*, from which the protein originates. To this end, a transgenic *A. thaliana* plant line stably overexpressing At4G01883 was generated (Figure S1d). Proteins from the plants and from control plants overexpressing PYL5 were

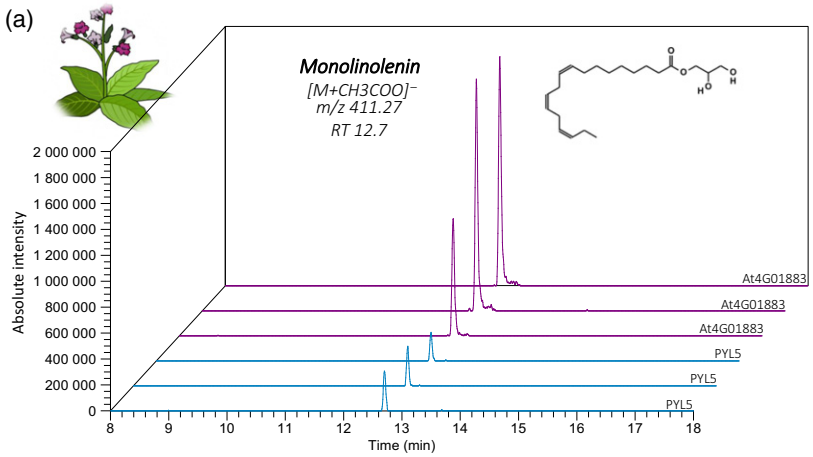
isolated as described earlier. The efficiency of expression and purification was verified via SDS-PAGE and Western blotting, and the eluate was analyzed by LC-MS (Figure S1d). SIEVE data analysis once again detected the unique ion (*m/z* 411.27, RT 12.7) in eluates from three biological replicates of At4G01883 samples (Figure 3b). Detection of the same molecule co-eluting with At4G01883 in both heterologous as well as homologous systems implied that it might be the actual ligand of the target protein. Therefore, we proceeded to characterize the chemical structure of the molecule. The LC-MS data of control (PYL5) and At4G01883 eluates were acquired in positive and negative electrospray ionization (ESI) modes. The same ion unique to At4G01883 was detected in positive mode as *m/z* 353.26818 and in negative mode as *m/z* 411.27520 (Figures 4a and S8a). Atomic composition calculations of the detected ions suggested a formula of $C_{21}H_{37}O_4$ $[C_{21}H_{36}O_4 + H]^+$ for the ion detected in positive mode and $C_{23}H_{39}O_6$ $[C_{21}H_{36}O_4 + CH_3COO]^-$ for that detected in negative mode (Figures 4b and S8b). Next, the unknown compound co-eluting with At4G01883 was analyzed by LC-MS/MS, where the ions (*m/z* 411.27 and 353.27) were subjected to collision-induced dissociation (CID). Negative CID of the parent ion (*m/z* 411.27) gave rise to *m/z* 59.01375, indicating loss of the acetate adduct (CH_3COO^-) from an ion–molecule complex $[C_{21}H_{36}O_4 + CH_3COO]^-$ (Figure S8c). The CID spectrum of the ion measured in positive mode (*m/z* 353.27) was consistent with the structure of the ester of glycerol ($C_3H_8O_3$) with a fatty acid containing three double bonds ($C_{17}H_{29}COOH$) (Figure 4c,d). Further analysis demonstrated that chromatographic retention time, accurate masses and CID spectra of the unknown molecule measured in both positive and negative modes were identical to the characteristics of monolinolenin (AccuStandard UG-024N) (Figure S9). This finding was confirmed by spiking eluates corresponding to At4G01883 (Figure 5) with a monolinolenin standard. This resulted in two overlapping peaks, which is a strong indication that the two compounds are identical. Moreover, spiking recombinant GST-tagged At4G01883 purified from *Escherichia coli* with a monolinolenin standard in an *in vitro* pull-down assay demonstrated a mean 16-fold change in ion enrichment in contrast to the GST control (Figure 6). This result once again supported our proposal that monolinolenin is the At4G01883 protein ligand and it was given the name monolinolenin-binding protein 1 or MLBP1. To assess the specificity of monolinolenin to MLBP1, we spiked HPB-tagged PYL5, GFP and MLBP1-related proteins purified from *N. benthamiana* with monolinolenin and conducted an *in vitro* pull-down assay. Negative-mode LC-MS analysis of the MLBP1 eluates identified an ion corresponding to monolinolenin $[C_{21}H_{36}O_4 + F]^-$ (*m/z* range 371.2598), at an intensity significantly higher than that of the control eluates (Figure S10).

Isothermal calorimetry (ITC) was then applied to study the MLBP1–monolinolenin interaction (Methods S1). Titration of GST-MLBP1 to monolinolenin, but not GST, resulted in an endothermic reaction, which reached saturation with an increase in the protein concentration (Figure S11). Such a heat change was not observed when the buffer or MLBP1 was titrated to buffer free of monolinolenin. A similar response was observed when ITC was carried out by titrating ABA to dimeric PYR1 but not to the mutated $PYR1^{H60P}$ monomer (Dupeux *et al.*, 2011). Dupeux *et al.* (2011) associated the endothermic reaction with ABA-induced dissociation of the PYR1 dimer. While MLBP1 ITC data cannot alone support the hypothesis of MLBP1 dimerization, they do support an interaction between MLBP1 and monolinolenin.

MLBP1 is a protein with an unknown function

At4G01883 is predicted to be a polyketide cyclase/dehydrase and lipid transport protein (TAIR database). In Arabidopsis rosettes and *N. benthamiana* leaves the protein was demonstrated to be localized to chloroplasts [Zybailov *et al.* (2008) and Figure S12a, respectively]. Image analysis of protein localization confirmed the presence of At4G01883-eGFP fusion in chloroplasts and identified the first 18 amino acids as the transit peptide sequence (Figure S12a,c–e). Due to the phylogenetic relationship of the protein to ABA receptors (Figure S13), we hypothesized that MLBP1 and/or monolinolenin might be involved in signaling. To test this hypothesis, 50 μM monolinolenin was applied to Arabidopsis seedlings. While this dose resulted in no visible effect, application of 200 μM monolinolenin had an inhibitory effect on roots. To test whether the inhibitory effect is specific to monolinolenin, several other compounds were applied (Figure S14). The free acid, linolenic acid and the reduced version of monolinolenin, monolinolein ($C_{21}H_{38}O_4$), demonstrated a similar inhibitory effect on roots (Figure S14). Furthermore, application of 200 μM monolinolenin to the *MLBP1* T-DNA mutant (SAIL_504_A06) or transgenic plants overexpressing MLBP1 resulted in similar root inhibition to that seen in the wild type plant (Figures S15a–c and S1d). It is therefore likely that root inhibition is not related to MLBP1.

Neither the T-DNA mutant nor the overexpressing transgenic plants showed any phenotypes to indicate any involvement of MLBP1 in plant development. Although one can argue that the presence of the two MLBP1-related polyketide cyclase/dehydrase-like proteins encoded by At1G02470 and At1G02475 (Figure S13) could obscure the mutant phenotype, it should be noted that neither *in planta* pull-downs nor monolinolenin spiking assays involving these two proteins indicated significant binding of monolinolenin (Figure S10). Moreover, subcellular localization analysis of MLBP1 and its closet homolog, At1G02475, indicated distinct subcellular localization for the two proteins;



Experiment Screening	Ion in At4G01883 replicates	m/z	RT	P-Value	Control	Fold change	Mass-Analyzer	Figure
II	**$[C_{21}H_{36}O_4 + CH_3COO]^-$**	**411.3**	**12.7**	**0.04**	**PYL5**	**6.18**	**Q Exactive**	**current**
III	$[C_{21}H_{36}O_4 + CH_3COO]^-$	411.3	13.2	0.03	PYL5	3.92	LTQ Orbitrap	Sup 4A
	$[C_{21}H_{36}O_4{}^{35}Cl]$	387.2	13.2	0.04	PYL5	4.47		
IV	$[C_{21}H_{36}O_4 + CH_3COO]^-$	411.3	13.2	0.03	PYL5	5.9	LTQ Orbitrap	Sup 4B
	$[C_{21}H_{36}O_4{}^{35}Cl]$	387.2	13.2	0.04	PYL5	6.1		
	$[C_{21}H_{36}O_4{}^{37}Cl]$	389.2	13.2	0.04	PYL5	7.5		
V	$[C_{21}H_{36}O_4{}^{35}Cl]$	387.2	16.9	0.04	GFP	25.6	6224 Mass TOF	

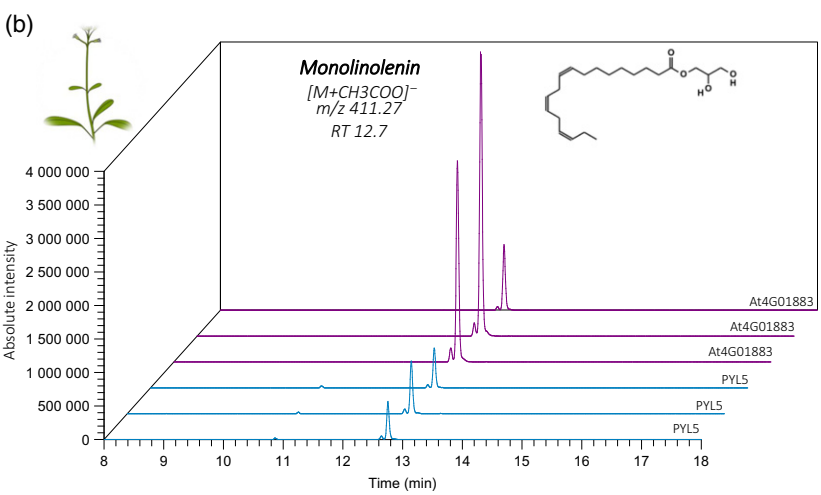


Figure 3. Co-purification of monolinolenin with At4G01883 protein.
Negative mode LC-MS extracted mass chromatograms (EMCs) (*m/z* range 411.27089–411.27911) of three biological replicates of co-eluted small molecules purified with At4G01883-HPB, which was (a) transiently expressed in *Nicotiana benthamiana* or (b) stably expressed in *Arabidopsis thaliana*, are indicated in violet. Blue graphs represent EMCs of eluates corresponding to (a) GFP-PYL5-HPB (AT5G05440) transiently co-expressed with HAB1-GFP (AT1G72770) or (b) PYL5-HPB (AT5G05440), which served as controls. Molecular ion, *m/z*, retention time (RT) and the structural formula of detected monolinolenin molecule are indicated above the peaks. Statistical analyses (Student's *t*-test) of current and additional independent experiments are presented in the table (in panel a), summarizing *P*-value and fold change of the identified adducts in comparison with the relevant control. Additionally, the table includes the chemical formula of adducts, *m/z*, RT and the model of mass spectrometer used in the specified experiment. Data displayed in the current figure are in bold. Corresponding figures of additional experiments are indicated at the last column.

while MLBP1 localized to the chloroplast (Figure S12a), At1G02475 was visualized in the nucleocytoplasm (Figure S12b). We therefore conclude that it is unlikely that either of the two MLBP1-related proteins function in a redundant fashion with respect to monolinolenin binding.

CONCLUSIONS

Protein–metabolite interactions are indisputably essential for the biochemical reactions fundamental to life. Nevertheless, the endogenous small molecular ligands of a large number of proteins with predicted ligand-binding domains have still not been identified (Anantharaman and Aravind, 2001; Schrick *et al.*, 2004; Shinohara *et al.*, 2012; Das and Rahman, 2014; Grotewold *et al.*, 2017; Hohmann *et al.*, 2017). While several methods allowing identification of endogenous protein ligands have been demonstrated in yeast, bacteria and animal systems (Saghatelian *et al.*, 2004; Li and Snyder, 2012; Piazza *et al.*, 2018), only a handful have been applied *in planta*. These include profiling metabolomes in unique genetic backgrounds and the application of affinity co-purification of a target protein in complex with its ligand (Casanal *et al.*, 2013; Luzarowski *et al.*, 2017). Herein, we have presented a robust approach for the discovery of protein ligands *in planta*. Initially, proteins are expressed and co-purified in complex with their putative ligands through an avidin-based handle. Next, small molecule eluates are analyzed by LC-MS. The method was demonstrated to be highly reproducible in

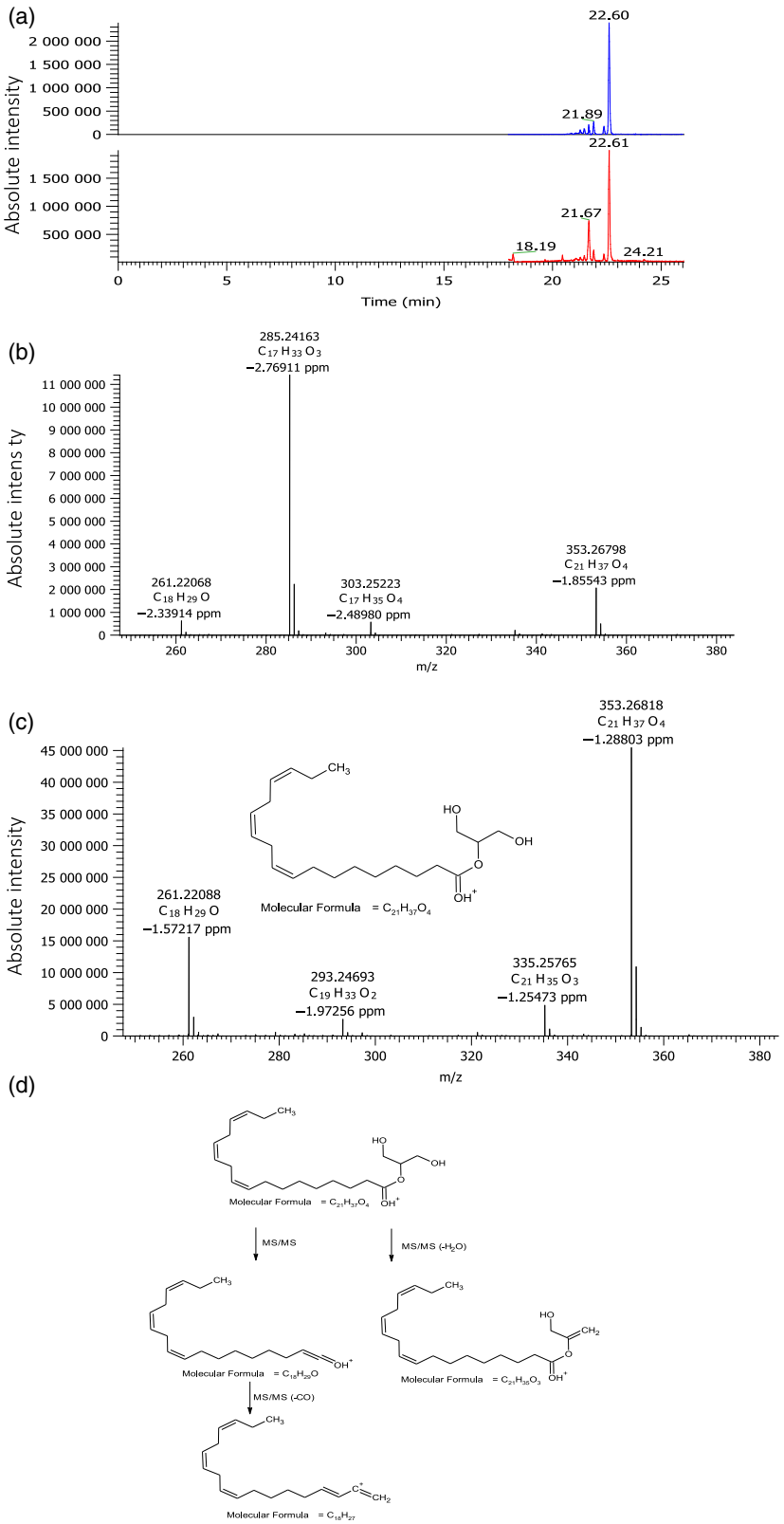


Figure 4. Monolinolenin is co-purified with At4G01883 protein.

(a)–(c) Structural determination of the ligand co-purified with At4G01883-HPB protein from *Arabidopsis thaliana* leaves. Structure was determined by a three-phase analysis. Co-purification of the unique ligand (*m*/*z* 353.2682) was re-confirmed by (a) positive atmospheric pressure chemical ionization extracted-ion chromatogram (APCI-EIC; blue) and tandem mass spectrometry chromatogram (MS/MS; red). The atomic composition of *m*/*z* 353.26798 at retention time (RT) of 22.60 min was obtained by (b) a positive APCI MS spectrum. The fragmentation spectrum of *m*/*z* 353.27 detected at a RT of 22.60 was acquired from (c) the positive APCI MS/MS spectrum.

(d) Proposed chemical structures of the protonated ion and its derivatives detected at positive APCI MS/MS.

numerous independent experiments conducted in various biological systems. Identification of protein–metabolite complexes was equally successful when using different MS apparatuses. An experimental design in which proteins carrying similar structural binding sites served as controls and the application of SIEVE software circumvented an

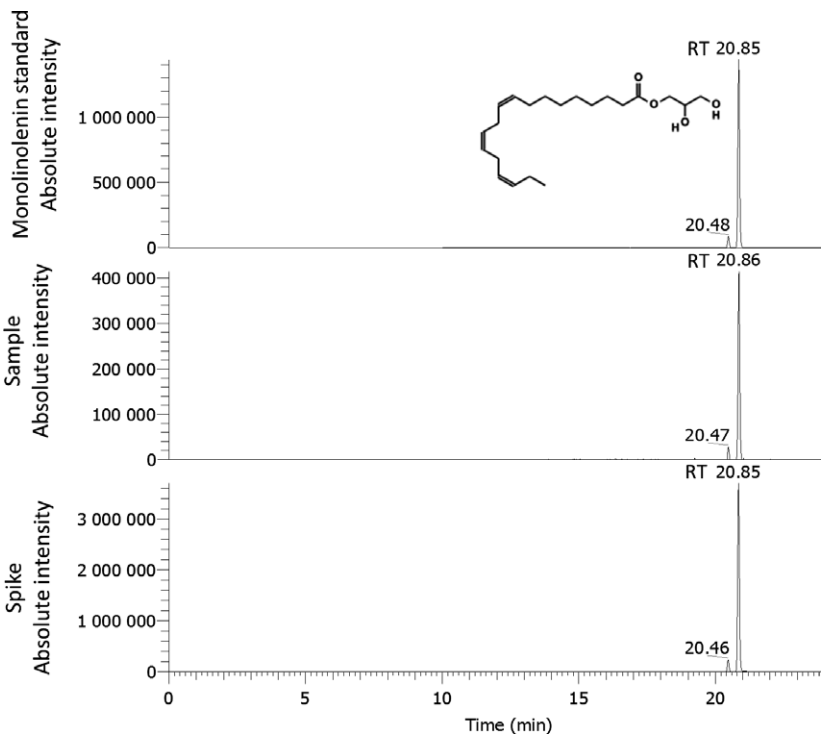
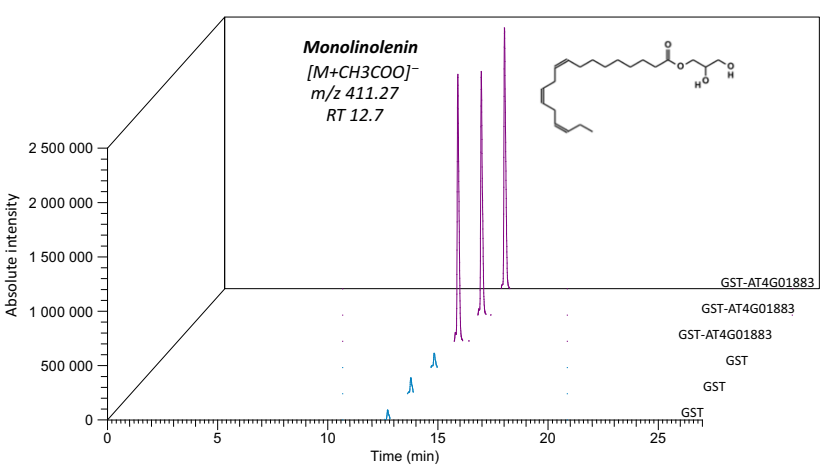


Figure 5. Verification of the identity of the ligand co-purified with At4G01883 protein.
Overlay of negative electrospray ionization LC-MS extracted ion chromatograms (*m/z* 411.27523) of monolinolenin detected in the standard (upper panel), the sample (middle panel) and a sample spiked with 0.1 μg ml^{-1} standard solution of monolinolenin at a ratio of 1:1 (lower panel).



Ion in AT4G01883-GST replicates	m/z	RT	P-Value	Control	Fold change	Mass-Analyzer	Figure
$[C_{21}H_{36}O_4 + CH_3COO]^-$	**411.27**	**12.7**	**0.0007**	**GST**	**16.7**	**Q Exactive**	**current**
$[C_{21}H_{36}O_4{}^{35}Cl]$	387.23	12.7	0.0025	GST	85		
$[C_{21}H_{36}O_4{}^{37}Cl]$	389.22	12.7	0.001	GST	688		

Figure 6. *In vitro* pull-down of monolinolenin with At4G01883-GST protein.
Recombinant GST-At4G01883 or GST alone was purified from *Escherichia coli*. Twenty micrograms of GST-At4G01883 (about 400 nM) or 100 μg of glutathione *S*-transferase (GST; about 400 nM) were bound to glutathione Sepharose beads (20 μl), excess protein was washed and incubated with 4 μM monolinolenin. Bound small molecules were eluted by acetonitrile and subjected to negative-mode LC-MS analysis. Extracted mass chromatograms (EMCs) (*m/z* range 411.27089–411.27911) of three technical replicates eluted from GST-At4G01883 are shown in violet and those for GST alone in blue. Molecular adduct, *m/z*, retention time (RT) and the structural formula of detected ions are indicated above the peaks. Results of non-targeted statistical analyses (Student's *t*-test) are presented in the table, summarizing *P*-value and fold change of the identified adducts compared with GST. Additionally, the table includes the chemical formula of adducts, *m/z*, RT and the model of mass spectrometer used. Data displayed in the current figure are in bold.

uninformed putative ligand elimination process. The technique was demonstrated to be sensitive for detection of low-abundance molecules such as hormones. Finally, the short time frames required for transient *Agrobacterium*-mediated expression together with protein purification made the method favorable for high-throughput screening.

Although the approach was adjusted to discover START ligands it can easily be adapted to ligands of a different chemical nature by adjustment of the analytical conditions. Nevertheless, as with any other, this method is not free of limitations. Although transient protein expression takes place in *N. benthamiana* leaves, these do not represent the physiological environment of all target proteins. However, this might be overcome by spiking lysates with metabolite extracts from the designated tissues of relevant plant species. In addition, proteins in our system are overproduced, violating protein homeostasis within the cells, which may result in false-positive interactions. Moreover, the method is unable for discriminating between direct and indirect protein–metabolite complexes as they might be purified together with associated proteins, which may prove to be binding agents. That said, the current approach successfully identified hitherto unknown ligands of two of the 37 tested START domain-related proteins. While one could argue this is a low success rate, it should be noted that apart from a few known ligands and the structural indications of START domain-related proteins pointing in favor of ligand binding, there is no certainty that a ligand for each and every one of the START/Bet v/polyketide cyclase-like proteins exists. Moreover, expression of target proteins in leaf cells limited the metabolites to which the proteins were exposed, leaving the rest beyond their reach. Considering these limitations, it is impossible to assess the true success rate. Given the alternatives (Casanal *et al.*, 2013; Luzarowski *et al.*, 2017), and the relative ease of application of the proposed approach, this method presents a good technique for initial probing of protein ligands *in planta*.

EXPERIMENTAL PROCEDURES

Binary plasmid construction

All START domain-encoding genes were cloned in a translational fusion to a biotinylation tag sequence and expressed under the 35S constitutive promoter (Pri-Tal *et al.*, 2017). An RNA mixture from *A. thaliana* Col-0 seedlings, seeds, flowers and fruits was isolated using the Concert™ Plant RNA Reagent. Samples were treated with DNase [Ambion™ DNase I (RNase-free), Catalog number AM2222, https://www.thermofisher.com] according to the manufacturer's instructions. Complementary DNA was generated from 1–5 μg of total RNA, using Reverse Transcriptase (SuperScript™ II Reverse Transcriptase, Catalog number 18064014, https://www.thermofisher.com), in a reaction mixture containing oligo-dT_{20}. The START gene coding sequences were amplified from cDNA using the primers listed in Table S3. One of the following restriction site combinations was added to each START domain-coding sequence: *Ngo*MIV and *Eco*RI, *Xma*I or *Xma*I and *Sal*I. The restriction sites generated to clone each START domain-encoding gene are referred to in the primer name (Table S3). Coding sequences were then inserted into the pHeGHPB (Pri-Tal *et al.*, 2017) linearized expression vector in either *Age*I and *Eco*RI, *Xma*I or *Xma*I and *Xho*I restriction sites. ΔN HAB1 was cloned into the pEGAD (Cutler *et al.*, 2000) binary vector using the *Eco*RI and *Xho*I restriction enzymes harboring a GFP translational fusion.

Nicotianan benthamiana growth

Nicotiana benthamiana seeds were densely seeded in a single pot with watered tuff soil mix and kept in the dark at 4°C for 1 week for synchronized germination. After 1 week the seeds were transferred to a PERCIVAL AR-41L2 growth chamber (https://www.percival-scientific.com/) with a light intensity of about 150 μmol photons m^{-2} sec^{-1} and 12-h day/night cycles at 22°C. After 1 week the plants were transferred, one by one, into separate pots with the same soil mix and grown under the same conditions. Plants were infiltrated with *Agrobacterium tumefaciens* at 3–4 weeks following transplant. The described growth conditions result in short internodes and postpone flowering. While the former brings about the necessary rigidity to withstand the load caused by transfection of the leaves, the latter elongates the physiological stage in which the experiment can be performed.

Transient expression in *N. benthamiana*

Agrobacterium tumefaciens strains transformed with each one of the plasmids were grown in 50 ml Luria–Bertani medium, supplemented with 50 mg L^{-1} kanamycin and 50 mg L^{-1} rifampicin, overnight at 28°C while shaking. Bacterial cultures were centrifuged at 3220 ***g***, at room temperature (25°C), for 30 min. The pellets were resuspended in 5 ml of 5% sucrose and the OD_{600} was measured. Next, the cultures were centrifuged at 3220 ***g*** at 28°C for 5 min. The pellets were further resuspended in an induction buffer (10 mM $MgCl_2$, 10 mM MES, pH 5.6, 150 μm acetosyringone) and diluted to an OD_{600} of 1. Bacteria were infiltrated into *N. benthamiana* leaves using a blunt-end 1-ml syringe.

Generation of transgenic Arabidopsis

Arabidopsis thaliana strain Col 0 was grown in 16-h day/8-h night cycles at 20–22°C and 70% humidity in a PERCIVAL AR-41L2 growth chamber with a light intensity of about 150 μmol photons m^{-2} sec^{-1}, and then transformed using the floral-dip method (Clough and Bent, 1998). T_1 transgenic plants were selected with 15 mg L^{-1} glufosinate. Selected transgenic plants bearing a single copy of a transgene were further propagated up to T_3 homozygous plants.

Purification of protein–ligand complexes

For each experiment, approximately 10 g of the leaf tissue was harvested from *N. benthamiana* 2 days post-infiltration or from bolting transgenic *A. thaliana*. Leaves were ground to a powder with a mortar and pestle in liquid nitrogen. The powder was supplemented with two volumes of chilled extraction buffer [50 mM TRIS, pH 7.5, 150 mM NaCl, 10% glycerol, 0.5% NP-40 (Sigma I8896, https://www.sigmaaldrich.com/), supplemented with TRIS (2-carboxyethyl) phosphine) (TCEP; Thermo Scientific 20490, https://www.thermofisher.com/) and protease inhibitor cocktail (Sigma P9599)]. The extract was cleared by ultracentrifugation at 4°C and 100 000 ***g*** for 1 h. Alternatively, sample clearing was performed with a tabletop centrifuge at 4°C and 20 000 ***g***, with pre-cleaning using Miracloth (Calbiochem 475855, https://uk.vwr.com/). To prevent clogging during the liquid chromatography, either ultracentrifuged or Miracloth-treated samples were further filtrated through reinforced 0.45-μm nylon syringe filters (VWR® Syringe Filters, Catalog number 28145–489, https://us.vwr.com/). The clear lysate was loaded onto 50–100 μl streptavidin Sepharose™ affinity chromatography medium (GE Healthcare 17-5113-01, https://www.gehealthcare.com/) and incubated at 4°C, while rotating at a speed of 18 r.p.m. for 30 min. Streptavidin Sepharose™

affinity chromatography medium with bound proteins was separated from the rest of the lysate by a short centrifugation at 4°C at 1810 ***g***. The streptavidin Sepharose™ affinity chromatography medium was then washed once with extraction buffer and twice with 150 mM NaCl. Small molecular ligands were eluted using 60–80% ACN. The lysate and the affinity media were sampled throughout the experiment to ensure protein integrity and check for effective protein purification.

SDS-PAGE and Western blotting

Aliquots sampled before and after addition of the affinity medium were resuspended in sample buffer (final concentration 0.25 M TRIS-HCl pH 6.8, 50% glycerol, 15% SDS, 28% β-mercaptoethanol). Elution of the protein from the streptavidin Sepharose™ affinity chromatography medium was accomplished by resuspension in urea sample buffer (final concentration 0.2 M TRIS-HCl pH 6.8, 10% glycerol, 8% SDS, 20% β-mercaptoethanol, 5 M urea). For Western blot analysis, samples were pre-heated at 80°C for 10 min, separated by SDS-PAGE and then electrophoretically transferred onto nitrocellulose membranes (GE Whatman 10401180, https://www.gelifesciences.com/). Membranes were then treated with blocking buffer [5% non-fat dry milk, 1× TRIS-buffered saline (TBS), 0.1% Tween 20], washed three times with TBS–Tween 20 and incubated with strepavidin–horseradish peroxidase conjugate (GE Healthcare RPN1231). Proteins were detected with the Western LightningR Plus ECL reagent (PerkinElmer NEL103001EA, http://www.perkinelmer.com/) and visualized using the ImageQuant LAS 4000 mini (GE Healthcare).

Purification of recombinant protein from *E. coli*

The coding sequence of At4G01883 was cloned into the pGEX-4T1 vector (GE Healthcare, 28-9545-49), resulting in a glutathione *S*-transferase (GST)-fused protein. At4G01883-GST and GST were expressed in BL21(DE3) pLysS (Promega, https://www.promega.com/) by inoculating 50 ml of the overnight culture into 1 L of TB (1.2% tryptone, 2.4% yeast extract, 0.4% glycerol 17 mM KH_2PO_4, 72 mM K_2HPO_4, pH 7.0) and incubated at 30°C. Protein expression was induced at OD_{600} = 0.8 with 1 mM isopropyl β-D-1-thiogalactopyranoside for 16 h at 15°C. Bacteria were harvested by centrifugation at ≥3220 ***g*** and resuspended in 3 ml g^{-1} TBS [50 mM TRIS, 150 mM NaCl, pH 7.5, and tablet per 50 ml cOmplete™ Protease Inhibitor Cocktail (Sigma)]. The suspension was then frozen and stored at −80°C. Bacteria were thawed, sonicated and centrifuged at 20 000 ***g*** for 10 min. Clear lysate was applied to a 3 ml glutathione Sepharose column (GE Healthcare). The column was then washed with 70 volumes of TBS. Proteins were eluted with TBS supplemented with 20% glycerol, 1 mM TRIS (2-carboxyethyl) phosphine (TCEP; Thermo Scientific 20490), 30 mM reduced glutathione and 300 mM NaCl and then dialyzed in TBS supplemented with 20% glycerol.

In vitro pull-down

For the GST-based pull-down assay, GST (about 400 nM; 100 μg) or GST-At4G01883 (about 400 nM; 20 μg), in triplicate, was dissolved in 1 ml of extraction buffer supplemented with 20 μl of glutathione Sepharose beads and incubated for 1 h at room temperature (25°C). Excess unbound protein was washed with 1 ml of extraction buffer. Bound protein was quantified via SDS-PAGE and compared with BSA standards. The bound protein was incubated in 1 ml of extraction buffer, spiked with 4 μM monolinolenin, for 1 h at room temperature (25°C). Beads were then washed twice with 1 ml of extraction buffer and twice more with extraction buffer without NP-40. Twelve additional samples with GST were similarly treated. These samples were later used to plot a standard curve by spiking with known monolinolenin concentrations (0.125–4 μM in duplicates), while considering the ion suppression resulting from the buffer. Then, ACN (50 μl) was added to all samples and the eluates subjected to LC-MS analysis.

For biotin-based *in vitro* pull-down assay, GFP, PYL5 and eight polyketide cyclase-like proteins fused at the C-terminus to the HPB tag were purified from *N. benthamiana* as described above and stringently washed to remove bound endogenous metabolites [four washes with extraction buffer (150 mM NaCl, 50 mM TRIS-HCl pH 7.5, 10% glycerol, 0.5% NP-40 (Sigma I8896) supplemented with 1 mg ml^{-1} BSA, followed by four washes with extraction buffer supplemented with 1 mg ml^{-1} BSA]. The purified protein was incubated for 1 h in extraction buffer supplemented with 20 nM monolinolenin. Unbound compound was washed away by one rinse with extraction buffer supplemented with 1 mg ml^{-1} BSA and two washes with TBS supplemented with 1 mg ml^{-1} BSA. Metabolites were eluted in ACN and subjected to LC-MS.

Seedling response to chemical treatments

Col-0 wild type seeds were gas sterilized and plated on 0.5 MS media with 0.5% sucrose. Seeds were then stratified for 5 days at 4°C and incubated at 22°C. For the assay, 3-day-old seedlings were transferred to MS medium containing 200 μM of one of the following chemicals: stearic acid (Sigma S4751), 1-stearoyl-*rac*-glycerol (Sigma M2015), linoleic acid (Sigma L1376), linolenic acid (Sigma L2376), monovaccenin (Accustandard UG-018N, https://www.accustandard.com/), monolinolenin (Accustandard UG-024N) or monolinolein (Accustandard UG-021N). All chemicals were dissolved or diluted in methanol to a stock solution of ×1000. Methanol served as a mock control. The plates were oriented vertically and the response was assessed 2 days later.

Q Exactive™ Plus hybrid quadrupole-Orbitrap™ mass spectrometer LC-MS/MS

Samples were analyzed by the LC-MS system which consisted of a Dionex Ultimate 3000 RS HPLC coupled to a Q Exactive Plus hybrid FT mass spectrometer, equipped with a heated electrospray ionization source (Thermo Fisher Scientific). Two different chromatographic separations were utilized. A shorter protocol was used to analyze small molecules in the screening phase and a longer protocol for spiking or MS/MS. In the short protocol, compound separation was carried out using a Kinetex Hexyl-Phenyl column (2.1 × 150 mm, particle size 2.6 μm, Phenomenex, https://www.phenomenex.com/), employing a linear ACN/water (with 0.1% acetic acid) gradient. In the longer protocol, separation was carried out using an Acclaim C18 column (3 × 250 mm, particle size 3 μm, Thermo Scientific) employing a linear ACN/water (with 0.1% acetic acid) gradient.

The negative ionization mode was applied for ionization of small molecules during the screening phase and for ABA MS/MS. The positive ionization mode was only applied for monolinolenin MS/MS analysis. In the negative ionization mode, ion source parameters were as follows: spray voltage 3 kV, capillary temperature 300°C, sheath gas rate (arb.) 40, auxiliary gas rate (arb.) 10. Mass spectra were acquired in scan mode; the resolving power was 140 000. The MS/MS spectra were acquired using an HCD cell in PRM acquisition mode at 20, 30 and 40 normalized collision energy values. For the positive ionization mode, the following ion source parameters were used: corona discharge needle current 5 μA, capillary temperature 300°C, vaporizer temperature 400°C, sheath gas rate (arb.) 40, auxiliary gas rate (arb.) 10.

Mass spectra were acquired in scan mode; resolving power was 70 000. The MS/MS spectra were acquired using an HCD cell in

PRM acquisition mode at 10, 20 and 30 normalized collision energy values.

The LC-MS system was controlled and data were analyzed using Xcalibur software (Thermo Fisher Scientific).

Thermo Scientific LTQ Orbitrap Discovery LC-MS

Samples were analyzed on an LC-MS system comprising a Dionex Ultimate 3000 HPLC coupled to an LTQ Orbitrap Discovery hybrid FT mass spectrometer, equipped with an ESI source (Thermo Fisher Scientific Inc.). The chromatography separation solvent system was A = 0.1% acetic acid in water (JT Baker, 9831-03, https://www.merckgroup.com), B = ACN (Fluka Analytical, 00683, https://www.merckgroup.com) and C = IPA/ACN [acetonitrile (ACN) and isopropanol (IPA)] at a ratio of 7:3. Gradient elution started at 85% A and was followed by 15% B (0–14 min), 5% A and 95% B (14–22 min), 5% A and 95% C (22–27.1 min), 85% A, 15% B (27.1–32 min). The flow rate for each run was 0.35 ml min^{-1} for the duration of the gradient, and the column oven temperature was 40°C. The temperature of the sample tray was 15°C. The injection volumes were 5 μl. The HPLC separations were carried out using a Kinetex Hexyl-Phenyl column (2.1 × 150 mm, particle size 2.6 μm, Phenomenex). The mass spectrometer was operated in negative ionization mode and the ion source parameters were as follows: spray voltage 3.5 kV, capillary temperature 300°C, ion-transfer optics parameters optimized using the automatic tune option, sheath gas rate (arb.) 35, auxiliary gas rate (arb.) 15. Mass spectra were acquired in the *m*/*z* 130–2000 Da range. The LC-MS system was controlled and data were analyzed using Chromeleon and Xcalibur software (Thermo Fisher Scientific).

Agilent 6224 accurate-mass time-of-flight (TOF) LC-MS

Samples were analyzed on a LC-MS system which comprised a HPLC-system (Agilent Technologies 1200 series, G1367B, https://www.agilent.com/) coupled to a 6224 TOF LC/MS (Agilent Technologies), operated in negative ESI mode.

Samples were separated using an Agilent Poroshell, 120 EC-C18 2.7 μm, 50 × 3.0 mm (Agilent, part no. 699975-302). In negative ionization mode the solvent system was A = 1 mM ammonium fluoride (Sigma Aldrich, http://www.sigmaaldrich.com/) in LC-MS-grade water (JT Baker, 9831-03) and B = ACN (Fluka Analytical, 00683). Gradient elution started at 99.5% A (2 min), and was followed by 80% A (4 min), 55% A (11 min), 2% A (18.1–19.6 min) and 99.5% A (20 min). The flow rate for each run was 0.5 ml min^{-1} for the duration of the gradient, and the column oven temperature was 41°C. The injection volumes ranged from 1 to 5 μl.

The ESI conditions were: gas temperature 325°C, drying gas 11 L min^{-1}, nebulizer 30 psig, fragmentor 120 V, skimmer 65 V. Data acquisition was performed in an extended dynamic range (2 GHz mode), with a scan range of 200–1700 *m*/*z* and an acquisition rate of 1.3 spectra sec^{-1}. The MassHunter Workstation (Agilent Technologies, B.04.00) was used to manually quantify extracted ion chromatograms by integrating peak intensities.

Differential ion enrichment analysis using SIEVE®

Differential analysis of full scan MS data was performed using SIEVE® version 2.2. All experiments were specified as *Small molecule* domain type with *Component extraction* algorithm. Frame parameters were adjusted to be automatic for *Frame time width*. The *m*/*z* values ranged between 130.00 and 1500.00 and *m*/*z* window size was adjusted to 10 p.p.m. and retention times ranged between 0.01 and 24.99 min out of 25 min. Component extraction parameters were set to: *Intensity threshold* 3000, *Minimum scans across a peak* 5, *Signal-to-noise* was adjusted to 3 and *m*/*z step size* to 10.

Total ion count chromatograms of the analyzed samples were initially aligned for the peak detection process to provide accurate results. Overlap of the base peaks was visually assessed, and if aggravated by the use of the ChromAlign Algorithm, the unalignment option was preferred. Up to four data files obtained from different samples, each represented by a set of three biological repeats, were processed against each other. A single replicate of one of the samples was assigned as the reference file for chromatographic alignment and a triplicate corresponding to a specific sample served a control. Because statistical SIEVE® analysis performs paired *t*-tests between samples and control, the experiment was processed several times, each time defining a different sample as a control. To overcome the multiple testing problem, a positive result was defined only if it was reproduced in a separate experiment.

Compounds that exhibited $P \leq 0.05$ with an associated *Ratio* value of mean peak intensities ≥5 between a sample and control were selected for further analysis by Xcalibur®. Whenever a *Ratio* value was inaccurate due to the presence of additional non-specific peaks adjacent to the specific peak, the *Ratio* was manually extracted by Xcalibur®. Xcalibur® enabled visual comparison of sample and control chromatograms, provided accurate signal intensities and determined chemical formulae of specific mass-to-charge values. SIEVE® and Xcalibur® software were purchased from Thermo Fisher Scientific.

ACCESSION NUMBERS

PYL5-AT5G05440, MLBP1 – AT4G01883, MLP 165 – AT1G35260, GLABRA 2 – AT1G79840, PROTODERMAL FACTOR 2 – AT4G04890.

AUTHOR CONTRIBUTIONS

YS conducted START/Bet v/polyketide cyclase-like protein expression, purification and data analysis. OP-T preformed the *in vitro* monolinolenin pull-down assays. YS and GZ designed and conducted the computational data analysis. SYP and AM cloned the START coding sequences to binary vectors. JB-A performed all the analytical chemistry procedures. JK helped with the *in vitro* monolinolenin pull-down and conducted protein localization studies. IV helped with data analysis. MG performed the ITC analysis. AM, SRC and YS designed and supervised experiments collaboratively. AM and YS drafted the manuscript with contributions from all co-authors.

ACKNOWLEDGEMENTS

We thank Gil Shoshani for helping with *Agrobacterium* infiltrations and Edgar Renteria for helping with the pull-down and data analysis. This work was supported by the ISF Israel Science Foundation (1069/14), BARD (IS-4919-16 R) and the NSF (1022378 and 1656890).

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Quality control of protein expression and purification.

Figure S2. The SDS-PAGE of AtPYL5-HPB interacting with HAB1-GFP.

Figure S3. Co-purification of ABA ligand with PYL5.

Figure S4. Verification of the identity of the ligand co-purified with PYL5 protein.

Figure S5. The ligand co-purified with PYL5 displays a MS/MS spectrum identical to that of the ABA standard.

Figure S6. Co-purification of $C_{18}H_{32}O_3$ with AT1G35260 protein.

Figure S7. Co-purification of monolinolenin with At4G01883 protein.

Figure S8. The ligand co-purified with At4G01883 protein is an acetate adduct.

Figure S9. The ligand co-purified with At4G01883 displays a MS/MS spectrum identical to that of the monolinolenin standard

Figure S10. At4G01883 polyketide cyclase-like protein binds monolinolenin

Figure S11. Isothermal titration calorimetry with GST-MLBP1 and GST control.

Figure S12. At4G01883 localizes to the chloroplast.

Figure S13. Evolutionary relationships of PYR/PYL ABA receptors and START domain-related polyketide cyclase-like proteins.

Figure S14. Root elongation is inhibited by monolinolenin, monolinolein, linolenic and linoleic acids.

Figure S15. Similar inhibitory effect on root growth of monolinolenin on wild-type and At4G01883 mutant strains.

Table S1. *Arabidopsis thaliana* START-domain and START domain-related proteins.

Table S2. The START domain proteins analyzed in this study.

Table S3. Primers used for cloning START/Bet v/polyketide cyclase-like genes.

Method S1. Supplementary Methods.

REFERENCES

Abe, M., Katsumata, H., Komeda, Y. and Takahashi, T. (2003) Regulation of shoot epidermal cell differentiation by a pair of homeodomain proteins in Arabidopsis. *Development*, **130**, 635–643.

Anantharaman, V. and Aravind, L. (2001) The CHASE domain: a predicted ligand-binding module in plant cytokinin receptors and other eukaryotic and bacterial receptors. *Trends Biochem. Sci.* **26**, 579–582.

Balcke, G.U., Handrick, V., Bergau, N., Fichtner, M., Henning, A., Stellmach, H., Tissier, A., Hause, B. and Frolov, A. (2012) An UPLC-MS/MS method for highly sensitive high-throughput analysis of phytohormones in plant tissues. *Plant Methods*, **8**, 47.

Casanal, A., Zander, U., Munoz, C., Dupeux, F., Luque, I., Botella, M.A., Schwab, W., Valpuesta, V. and Marquez, J.A. (2013) The strawberry pathogenesis-related 10 (PR-10) Fra a proteins control flavonoid biosynthesis by binding to metabolic intermediates. *J. Biol. Chem.* **288**, 35322–35332.

Choi, H.-I., Hong, J.-H., Ha, J.-O., Kang, J.-Y. and Kim, S.Y. (2000) ABFs, a family of ABA-responsive element binding factors. *J. Biol. Chem.* **275**, 1723–1730.

Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J.* **16**, 735–743.

Cutler, S.R., Ehrhardt, D.W., Griffitts, J.S. and Somerville, C.R. (2000) Random GFP:cDNA fusions enable visualization of subcellular structures in cells of Arabidopsis at a high frequency. *Proc. Natl Acad. Sci. USA*, **97**, 3718–3723.

Das, J. and Rahman, G.M. (2014) C1 domains: structure and ligand-binding properties. *Chem. Rev.* **114**, 12108–12131.

Dupeux, F., Santiago, J., Betz, K. *et al.* (2011) A thermodynamic switch modulates abscisic acid receptor sensitivity. *EMBO J.* **30**, 4171–4184.

Ersoy, B.A., Tarun, A., D'Aquino, K., Hancer, N.J., Ukomadu, C., White, M.F., Michel, T., Manning, B.D. and Cohen, D.E. (2013) Phosphatidylcholine transfer protein interacts with thioesterase superfamily member 2 to attenuate insulin signaling. *Sci. Signal.* **6**, ra64–ra64.

Felsenstein, J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, **39**, 783–791.

Gatta, A.T., Wong, L.H., Sere, Y.Y., Calderón-Noreña, D.M., Cockcroft, S., Menon, A.K. and Levine, T.P. (2015) A new family of StART domain proteins at membrane contact sites has a role in ER-PM sterol transport. *Elife*, **4**, 1–46.

Geiger, D., Scherzer, S., Mumm, P. *et al.* (2009) Activity of guard cell anion channel SLAC1 is controlled by drought-stress signaling kinase-phosphatase pair. *Proc. Natl Acad. Sci. USA*, **106**, 21425–21430.

Green, N. (1963) Avidin. *Biochem. J.* **89**, 599–609.

Grotewold, E., Yuan, L. and Feller, A. (2017) The BIF domain in plant bHLH proteins is an ACT-like domain. *Plant Cell*, **29**, 1800–1802.

Hofmeister, F. (1988) On regularities in the albumin precipitation reactions with salts and their relationship to physiological behavior. *Arch. Exp. Pathol. Pharmakol.* **24**, 247–260.

Hohmann, U., Lau, K. and Hothorn, M. (2017) The structural basis of ligand perception and signal activation by receptor kinases. *Annu. Rev. Plant Biol.* **68**, 109–137.

Iyer, L.M., Koonin, E.V. and Aravind, L. (2001) Adaptations of the helix-grip fold for ligand binding and catalysis in the START domain superfamily. *Proteins Struct. Funct. Genet.* **43**, 134–144.

Kawano, Y., Ersoy, B.A., Li, Y., Nishiumi, S., Yoshida, M. and Cohen, D.E. (2014) Thioesterase superfamily member 2 (Them2) and phosphatidylcholine transfer protein (PC-TP) interact to promote fatty acid oxidation and control glucose utilization. *Mol. Cell. Biol.* **34**, 2396–2408.

Kubo, H. (1999) ANTHOCYANINLESS2, a homeobox gene affecting anthocyanin distribution and root development in Arabidopsis. *Plant Cell*, **11**, 1217–1226.

Li, X. and Snyder, M. (2012) Analyzing in vivo metabolite–protein interactions by large-scale systematic analyses. *Curr. Protoc. Chem. Biol.* **3**, 181–196.

López-Carbonell, M. and Jáuregui, O. (2005) A rapid method for analysis of abscisic acid (ABA) in crude extracts of water stressed Arabidopsis thaliana plants by liquid chromatography—mass spectrometry in tandem mode. *Plant Physiol. Biochem.* **43**, 407–411.

Luzarowski, M., Kosmacz, M., Sokolowska, E., Jasinska, W., Willmitzer, L., Veyel, D. and Skirycz, A. (2017) Affinity purification with metabolomic and proteomic analysis unravels diverse roles of nucleoside diphosphate kinases. *J. Exp. Bot.* **68**, 3487–3499.

Ma, Z., Ge, L., Lee, A.S.Y., Yong, J.W.H., Tan, S.N. and Ong, E.S. (2008) Simultaneous analysis of different classes of phytohormones in coconut (Cocos nucifera L.) water using high-performance liquid chromatography and liquid chromatography–tandem mass spectrometry after solid-phase extraction. *Anal. Chim. Acta*, **610**, 274–281.

Ma, Y., Szostkiewicz, I., Korte, A., Moes, D., Yang, Y., Christmann, A. and Grill, E. (2009) Regulators of PP2C phosphatase activity function as abscisic acid sensors. *Science*, **324**, 1064–1068.

Melcher, K., Ng, L.M., Zhou, X.E. *et al.* (2009) A gate-latch-lock mechanism for hormone signalling by abscisic acid receptors. *Nature*, **462**, 602–608.

Miyazono, K., Miyakawa, T., Sawano, Y. *et al.* (2009) Structural basis of abscisic acid signalling. *Nature*, **462**, 609–614.

Nakamura, M., Katsumata, H., Abe, M., Yabe, N., Komeda, Y., Yamamoto, K.T. and Takahashi, T. (2006) Characterization of the class IV homeodomain-leucine zipper gene family in Arabidopsis. *Plant Physiol.* **141**, 1363–1375.

Negi, J., Matsuda, O., Nagasawa, T., Oba, Y., Takahashi, H., Kawai-Yamada, M., Uchimiya, H., Hashimoto, M. and Iba, K. (2008) CO2 regulator SLAC1 and its homologues are essential for anion homeostasis in plant cells. *Nature*, **452**, 483–486.

Pan, X., Welti, R. and Wang, X. (2008) Simultaneous quantification of major phytohormones and related compounds in crude plant extracts by liquid chromatography–electrospray tandem mass spectrometry. *Phytochemistry*, **69**, 1773–1781.

Pan, X., Welti, R. and Wang, X. (2010) Quantitative analysis of major plant hormones in crude plant extracts by high-performance liquid chromatography–mass spectrometry. *Nat. Protoc.* **5**, 986–992.

Park, S.-Y., Fung, P., Nishimura, N. *et al.* (2009) Abscisic acid inhibits type 2C protein phosphatases via the PYR/PYL family of START proteins. *Science*, **324**, 1068–1071.

Piazza, I., Kochanowski, K., Cappelletti, V., Fuhrer, T., Noor, E., Sauer, U. and Picotti, P. (2018) A map of protein-metabolite interactions reveals principles of chemical communication. *Cell*, **172**, 358–372.

Ponting, C.P. and Aravind, L. (1999) START: a lipid-binding domain in StAR, HD-ZIP and signalling proteins. *Trends Biochem. Sci.* **24**, 130–132.

Prigge, M.J., Otsuga, D., Alonso, J.M., Ecker, J.R., Drews, G.N. and Clark, S.E. (2005) Class III homeodomain-leucine zipper gene family members have overlapping, antagonistic, and distinct roles in Arabidopsis development. *Plant Cell*, **17**, 61–76.

Pri-Tal, O., Shaar-Moshe, L., Wiseglass, G., Peleg, Z. and Mosquna, A. (2017) Non-redundant functions of the dimeric ABA receptor BdPYL1 in the grass Brachypodium. *Plant J.* **92**, 774–786.

Qi, Y. and Katagiri, F. (2009) Purification of low-abundance Arabidopsis plasma-membrane protein complexes and identification of candidate components. *Plant J.* **57**, 932–944.

Ratzon, E., Bloch, I., Nicola, M., Cohen, E., Ruimi, N., Dotan, N., Landau, M. and Gal, M. (2017) A small molecule inhibitor of Bruton's Tyrosine kinase involved in B-cell signaling. *ACS Omega*, **2**, 4398–4410.

Rerie, W.G., Feldmann, K.A. and Marks, M.D. (1994) The GLABRA2 gene encodes a homeodomain protein required for normal trichome development in Arabidopsis. *Genes Dev.* **8**, 1388–1399.

Roderick, S.L., Chan, W.W., Agate, D.S., Olsen, L.R., Vetting, M.W., Rajashankar, K.R. and Cohen, D.E. (2002) Structure of human phosphatidylcholine transfer protein in complex with its ligand. *Nat. Struct. Biol.* **9**, 507–511.

Romanowski, M.J., Soccio, R.E., Breslow, J.L. and Burley, S.K. (2002) Crystal structure of the Mus musculus cholesterol-regulated START protein 4 (StarD4) containing a StAR-related lipid transfer domain. *Proc. Natl Acad. Sci. USA*, **99**, 6949–6954.

Saghatelian, A., Trauger, S.A., Want, E.J., Hawkins, E.G., Siuzdak, G. and Cravatt, B.F. (2004) Assignment of endogenous substrates to enzymes by global metabolite profiling. *Biochemistry*, **43**, 14332–14339.

Saitou, N. and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425.

Santiago, J., Rodrigues, A., Saez, A., Rubio, S., Antoni, R., Dupeux, F., Park, S.-Y., Márquez, J.A., Cutler, S.R. and Rodriguez, P.L. (2009a) Modulation of drought resistance by the abscisic acid receptor PYL5 through inhibition of clade A PP2Cs. *Plant J.* **60**, 575–588.

Santiago, J., Dupeux, F., Round, A., Antoni, R., Park, S.-Y., Jamin, M., Cutler, S.R., Rodriguez, P.L. and Márquez, J.A. (2009b) The abscisic acid receptor PYR1 in complex with abscisic acid. *Nature*, **462**, 665–668.

Satheesh, V., Chidambaranathan, P., Jagannadham, P.T., Kumar, V., Jain, P.K., Chinnusamy, V., Bhat, S.R. and Srinivasan, R. (2016) Transmembrane START domain proteins: in silico identification, characterization and expression analysis under stress conditions in chickpea (Cicer arietinum L.). *Plant Signal. Behav.* **11**, e992698.

Schrick, K., Nguyen, D., Karlowski, W.M. and Mayer, K.F.X. (2004) START lipid/sterol-binding domains are amplified in plants and are predominantly associated with homeodomain transcription factors. *Genome Biol.* **5**, R41.

Schrick, K., Bruno, M., Khosla, A. *et al.* (2014) Shared functions of plant and mammalian StAR-related lipid transfer (START) domains in modulating transcription factor activity. *BMC Med.* **12**, 1–20.

Shinohara, H., Moriyama, Y., Ohyama, K. and Matsubayashi, Y. (2012) Biochemical mapping of a ligand-binding domain within Arabidopsis BAM1 reveals diversified ligand recognition mechanisms of plant LRR-RKs. *Plant J.* **70**, 845–854.

Stocco, D.M. (2001) StAR protein and the regulation of steroid hormone biosynthesis. *Annu. Rev. Physiol.* **63**, 193–213.

Tabunoki, H., Sugiyama, H., Tanaka, Y., Fujii, H., Banno, Y., Jouni, Z.E., Kobayashi, M., Sato, R., Maekawa, H. and Tsuchida, K. (2002) Isolation, characterization, and cDNA sequence of a carotenoid binding protein from the silk gland of Bombyx mori larvae. *J. Biol. Chem.* **277**, 32133–32140.

Tabunoki, H., Higurashi, S., Ninagi, O. *et al.* (2004) A carotenoid-binding protein (CBP) plays a crucial role in cocoon pigmentation of silkworm (Bombyx mori) larvae. *FEBS Lett.* **567**, 175–178.

Tagore, R., Thomas, H.R., Homan, E.A., Munawar, A. and Saghatelian, A. (2008) A global metabolite profiling approach to identify protein–metabolite interactions. *J. Am. Chem. Soc.* **130**, 14111–14113.

Takada, S., Takada, N. and Yoshida, A. (2013) ATML1 promotes epidermal cell differentiation in Arabidopsis shoots. *Development*, **140**, 1919–1923.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729.

Todd, J.M. and Gomez, J. (2001) Enzyme kinetics determined using calorimetry: a general assay for enzyme activity? *Anal. Biochem.* **296**, 179–187.

Tsujishita, Y. and Hurley, J.H. (2000) Structure and lipid transport mechanism of a StAR-related domain. *Nat. Struct. Biol.* **7**, 408–414.

Umezawa, T., Sugiyama, N., Mizoguchi, M., Hayashi, S., Myouga, F., Yamaguchi-Shinozaki, K., Ishihama, Y., Hirayama, T. and Shinozaki, K. (2009) Type 2C protein phosphatases directly regulate abscisic acid-activated protein kinases in Arabidopsis. *Proc. Natl Acad. Sci. USA*, **106**, 17588–17593.

Uno, Y., Furihata, T., Abe, H., Yoshida, R., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2000) Arabidopsis basic leucine zipper transcription factors involved in an abscisic acid-dependent signal transduction pathway under drought and high-salinity conditions. *Proc. Natl Acad. Sci. USA*, **97**, 11632–11637.

Vahisalu, T., Kollist, H., Wang, Y.-F. *et al.* (2008) SLAC1 is required for plant guard cell S-type anion channel function in stomatal signalling. *Nature*, **452**, 487–491.

Vlad, F., Rubio, S., Rodrigues, A., Sirichandra, C., Belin, C., Robert, N., Leung, J., Rodriguez, P.L., Lauriere, C. and Merlot, S. (2009) Protein phosphatases 2C regulate the activation of the Snf1-related kinase OST1 by abscisic acid in Arabidopsis. *Plant Cell*, **21**, 3170–3184.

Westerman, J., Wirtz, K.W., Berkhout, T., van Deenen, L.L., Radhakrishnan, R. and Khorana, H.G. (1983) Identification of the lipid-binding site of phosphatidylcholine-transfer protein with phosphatidylcholine analogs containing photoactivable carbene precursors. *Eur. J. Biochem.* **132**, 441–449.

Yin, P., Fan, H., Hao, Q., Yuan, X., Wu, D., Pang, Y., Yan, C., Li, W., Wang, J. and Yan, N. (2009) Structural insights into the mechanism of abscisic acid signaling by PYL proteins. *Nat. Struct. Mol. Biol.* **16**, 1230–1236.

Yu, J.-N., Meng, Q.-Y., Liu, W.-J., Lu, Y.-L. and Ren, X.-L. (2014) Analysis of acidic endogenous phytohormones in grapes by using online solid-phase extraction coupled with LC-MS/MS. *J. Chromatogr. Sci.* **52**, 1145–1149.

Zybailov, B., Rutschow, H., Friso, G., Rudella, A., Emanuelsson, O., Sun, Q. and van Wijk, K.J. (2008) Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. *PLoS One*, **3**, e1994.

