Structural Basis of KAI2 Divergence in Legume

Angelica M. Guercio[1], François-Didier Boyer[2], Catherine Rameau[3], Alexandre de Saint Germain[3†], Nitzan Shabek[1†]

[1] *Department of Plant Biology, University of California – Davis, Davis, CA 95616*

[2] *Université Paris-Saclay, CNRS, Institut de Chimie des Substances Naturelles, UPR 2301, 91198, Gif-sur-Yvette, France*

[3] *Institut Jean-Pierre Bourgin, INRAE, AgroParisTech, Université Paris-Saclay, 78000, Versailles, France*

[†]*Correspondence should be addressed to: nshabek@ucdavis.edu , Alexandre.De-Saint-Germain@inrae.fr*

Abstract

The α/β hydrolase KARRIKIN INSENSITIVE-2 (KAI2) mediates the perception of smoke-derived butenolides (karrikins) and an elusive endogenous hormone (KAI2-ligand, KL) found in all land plants. It has been suggested that *KAI2* gene duplication and sub-functionalization events play an adaptative role for diverse environments by altering the receptor responsiveness to specific KLs. These diversification occurrences are exemplified by the variable number of functional KAI2 receptors among different plant species. Legumes represent one of the largest families of flowering plants and contain many essential agronomic crops. Along the legume lineage the *KAI2* gene underwent a duplication event resulting in *KAI2A* and *KAI2B*. Here we show that the model legume, *Pisum sativum (Ps)*, expresses three distinct *KAI2* homologues, two of which, *KAI2A* and *KAI2B* have uniquely sub-functionalized. We characterize biochemically the distinct ligand sensitivities between these divergent receptors and report the first crystal structure of PsKAI2 in *apo* and butenolide-bound states. Our study provides a comprehensive examination of the specialized ligand binding ability of legume KAI2A and KAI2B and sheds light on the perception and enzymatic mechanism of the KAI2-butenolide complex.

Introduction

Karrikins (KARs) are a family of butenolide small molecules produced from the combustion of vegetation and are a bio-active component of smoke[1–4]. These molecules are capable of inducing germination of numerous species of plants, even those not associated with fire or fire-prone environments such as Arabidopsis[1,5–9]. Through studies in Arabidopsis, KAR sensitivity was shown to be dependent on three key proteins: a KAR receptor, an α/β hydrolase KARRIKIN INSENSITIVE2 (KAI2), an F-box MORE AXILLIARY GROWTH 2 (MAX2) component of the Skp1-Cullin-F-box (SCF) E3 ubiquitin ligase, and the proposed target of ubiquitination and degradation, the transcriptional corepressor SMAX1/SMXL2[4,10–13]. An increasing number of studies have shown that KAI2 and KAR signaling components are involved in the regulation of many plant developmental processes including seedling development, leaf shape, cuticle formation, and root development, as well as play roles in AM fungi symbiosis and abiotic stress response[2–4,14–16].

The striking similarities between KAR and strigolactone (SL) signaling pathways have been the focus of an increasing number of studies. Both SLs and KARs share a similar butenolide ring structure but instead of the KAR pyran moiety, the butenolide is connected via an enol ether bridge to either a tricyclic lactone (ABC rings) in canonical SLs, or to a structural variety in non-canonical SLs[17,18]. The receptor for SL, DWARF14 (D14) shares a similar α/β hydrolase fold as KAI2 and a parallel signaling cascade requiring the function of the MAX2 ubiquitin ligase and downregulation of SMXLs, corepressors which also share some structural elements with SMAX1/SMXL2[4,10,13,19]. Unlike KARs, SLs are plant hormones that act endogenously, but were also found to be exuded by plant roots. SLs regulate diverse physiological responses such promoting hyphal branching of arbuscular mycorrhizal (AM) fungi to enhance the efficiency of

AM symbiosis, stimulating germination of root parasitic plant species, repressing shoot branching, affecting lateral root formation, primary root growth, root hair elongation, secondary growth in the stem, leaf senescence, and adventitious root formation [20–32]. Notably, KAI2 family receptors have undergone numerous duplication events within various land plant lineages. D14 was found to be an ancient duplication in the KAI2 receptor in seed plant lineage followed by sub-functionalization of the receptor, making it uniquely implied in SL signaling[33–36]. The age-old question in receptor diversity has been the evolutionary purpose and functional significance of *KAI2* duplication events. It has been shown that D14 and KAI2 are not able to complement each other functions in planta[37–41]. To this end, within the ligand binding site of KAI2 receptors the substitution of a few amino acids can alter ligand specificity between KAI2 duplicated copies[39,42]. While the role of the D14 receptor in SL signaling is well established, KAI2 receptors and KAR signaling are less understood. Furthermore, given the fact that KAI2 is ancestral to D14 and that KAR signaling controls diverse developmental processes including those unrelated to fire, it has been suggested that KAI2s are able to perceive an endogenous ligand(s), of which is currently unknown and tentatively named KAI2-Ligand (KL)[33,34,41,43]. Thus far, several crystal structures of KAI2/D14 receptors have been reported and have led to a greater understanding of receptor-ligand perception and the hydrolytic activity of the receptor towards certain ligands[11,12,19,31,36,38,44–51]. The divergence between duplications of KAI2 receptors to confer altered ligand specificity has been partially addressed at the physiological and biochemical level for only few plant species, and a structural examination has been limited[38–40,42].

Legumes represent one of the largest families of flowering plants and contain many essential crops. Beyond their agronomic value, most legume species are unique among plants because of

their ability to fix nitrogen by utilizing symbiosis with rhizobia, in addition to AM fungi symbiosis. Because of the potential functional diversification and specialization of KAI2-ligand, we characterized and examined the KAI2 receptor mechanism in legume, using *Pisum sativum (Ps)* as a model.

In this study, we examined the implications of the pre-legume *KAI2* duplication event that resulted in legume *KAI2A* and *KAI2B* clades[42]. We found that *Pisum sativum* expresses three distinct *KAI2* homologues, two of which, *KAI2A* and *KAI2B* have uniquely sub-functionalized. We characterize biochemically the distinct ligand sensitivities between these divergent receptors and further report the first crystal structure of PsKAI2B in *apo* and a unique butenolide-bound state at high resolution (1.6 Å and 2.0 Å, respectively). Altogether our findings provide a comprehensive examination of the specialized ligand binding ability of legume KAI2A and KAI2B and sheds light on the perception and enzymatic mechanism of KAI2 receptors.

Results

Genetic identification and characterization of the legume *Pisum sativum KAI2* genes

To characterize the karrikin pathway sensing mechanisms in legume we examined the evolutionary context of representative legume *KAI2*s. We focused on the *Pisum sativum* genome that encodes distinct *KAI2* gene copies and represents the diversity of legume *KAI2* duplication events (**Figure 1a** and **Figure S1**). Notably, the legume lineage has undergone an independent duplication event resulting in distinct KAI2A and KAI2B protein receptors. We identified three KAI2 homologs in the pea genome[52] that clearly group within the core KAI2 clade by phylogenetic analysis. One (Psat4g083040) renamed *PsKAI2B*, grouped in the same subclade as the legume *KAI2B*s (including *Lotus japonicus, Lj, KAI2B*[42]) and two (Psat2g169960, termed

PsKAI2A and Psat3g014200) in the same subclade as the legume *KAI2A*s (including *LjKAI2A*[42]) (**Figure 1a** and **Figure S1**). Psat3g014200 was very likely a pseudogene as the putative encoded protein is lacking 82 amino acids (aa) in the middle of the protein in comparison to *PsKAI2A* and *PsKAI2B*. By cloning *PsKAI2A* coding sequence (CDS) we identified 2 transcripts for this gene, corresponding to 2 splicing forms (**Figure 1b**). The transcript PsKAI2A.1 comes from intron splicing and produces a protein of 305 aa. Thus, this protein shows a C-terminal extension of 33 aa similar to LjKAI2A (**Figure S2**), missing in other KAI2 proteins. The PsKAI2A.2 transcript arises from the intron retention, which shows a premature STOP codon 2 nucleotides after the end of the first exon. This leads to a 272 aa protein showing a similar size to other KAI2 proteins described (**Figure 1b** and **Figure S2**). From this analysis, it is clear that the KAI2 clade has undergone an independent duplication event in the legume lineage resulting in these KAI2A and KAI2B forms (**Figure S1a-b**). To examine potential functional divergence between the PsKAI2A and PsKAI2B forms, we first analyzed the aa sequences and identified notable alterations in key residues, of which numerous are likely to be functional changes as indicated in later analyses (**Figure S2**). To further characterize divergence of these genes we studied the expression patterns of the two PsKAI2 forms in various tissues of the *Pisum* plant (**Figure 1c-d**). Interestingly, the expression of *PsKAI2s* revealed a ten-fold higher expression of PsKAI2A in comparison to PsKAI2B and distinct patterns between the two forms in the roots, suggesting sub-functionalization between PsKAI2A and PsKAI2B.

***PsKAI2* genes can rescue inhibition of hypocotyl elongation of *kai2-2* Arabidopsis mutant**

To test the function of PsKAI2 proteins *in planta* a cross-species complementation was performed by transforming the Arabidopsis *kai2-2* mutant with the 2 splicing forms of *PsKAI2A*

(*PsKAI2A.1*, *PsKAI2A.2*, and *PsKAI2B*, **Figure 1e**). The proteins were expressed as fusion proteins with mCitrine or HA epitope driven by the native *AtKAI2* promoter (*pAtKAI2*). The widely described hypocotyl elongation assay[37,43] was performed under low light conditions, which causes an elongated hypocotyl phenotype of the *kai2-2* mutant when compared to Ler. All constructs completely restored the phenotype of the *kai2-2* mutant to WT phenotype, except the *pAtKAI2::PsKAI2A.1-6xHA* construct which restored partially the phenotype of the *kai2-2* mutant (**Figure 1e**). Because in Arabidopsis the stereoisomer of the synthetic strigolactone (–)-GR24 may act as KL mimic compound by triggering developmental responses via AtKAI2[11,43,53], we investigated hypocotyl elongation through the PsKAI2 proteins by quantifying hypocotyl length after (–)-GR24 treatment. Only the lines expressing AtKAI2 control protein were able to respond to the treatment whereas all the complemented lines with PsKAI2s did not significantly respond to (–)-GR24 (**Figure 1e**). These results suggest that PsKAI2 proteins are the functional orthologues of the Arabidopsis KAI2, however the differences in the ligand sensitivity between all expressed KAI2s were more elusive compared to the recently reported study in lotus[42] and as suggested by our subsequent biochemical results.

Biochemical data reveal altered ligand specificity and activity between PsKAI2s

To investigate the functional specificity between *Pisum* KAI2 receptors, we have purified PsKAI2 recombinant proteins and investigated various ligand-interaction and ligand-enzymatic activities of the receptors (**Figures 2-3** and **Figures S3-S5**). We first examined KAI2A and KAI2B ligand interactions via the thermal shift assay (DSF) with various KAI2/D14 family ligands including (+) and (–)-GR24 enantiomers (also known as $GR24^{5DS}$ and $GR24^{ent\text{-}5DS}$, respectively[54]) and (+)- and (–)-2'-*epi*-GR24 (also known as $GR24^{ent\text{-}5DO}$ and $GR24^{5DO}$,

respectively) (**Figure 2a-i**). DSF analyses revealed PsKAI2B has an increased change in stability in the presence of (–)-GR24 compared to PsKAI2A which has little to no alteration (**Figure 2c-d**). Thus, PsKAI2B protein differs from its ortholog from lotus, which is not destabilized by (–)-GR24[42] and suggests different ligand specificity among legumes. The other ligands and enantiomers induce no detectable shift in stability for either PsKAI2 proteins. In addition, an extensive interaction screen using intrinsic fluorescence further confirmed that only the (–)-GR24 stereoisomer interacts with PsKAI2 proteins (**Figure 2j** and **Figure S4**). The calculated K_d revealed that PsKAI2B has a better affinity for (–)-GR24 (K_d = 89.43 ± 12.13 μM) than PsKAI2A (115.40 ± 9.87 μM) as also indicated by the DSF assay.

To further examine the catalytic activity of KAI2 enzymes, an enzymatic assay was performed by quantifying the hydrolytic activity of PsKAI2 towards distinct ligands. To that end, KAI2 proteins were incubated with (+)-GR24, (–)-GR24, (+)-2'-epi-GR24 and (–)-2'-*epi*-GR24 in presence of 1-indanol as an internal standard followed by ultraperformance liquid chromatography (UHPLC)/UV DAD analysis (**Figure 3**). The activity PsKAI2A and PsKAI2B was measured in comparison to AtD14, AtKAI2, and RMS3. These results show that PsKAI2A could only cleave (–)-GR24, however PsKAI2B is able to cleave (+)-GR24, (–)-GR24 and (–)-2'-*epi*-GR24 stereoisomers. Unlike RMS3, AtD14 and AtKAI2 have no detectable cleavage for (+)-2'-*epi*-GR24, strongly indicating that PsKAI2s have different stereoselectivity. To further investigate the cleavage kinetics activity of PsKAI2 proteins, we performed an enzymatic assay with the pro-fluorescent probes that were previously designed for detecting SL perception mechanism[55]. Here, (±)-GC240 probe bearing one methyl group on the D-ring was used to measure hydrolysis activity by PsKAI2s, RMS3, AtD14, and AtKAI2 enzymes (**Figure S5a-b**). As expected, PsKAI2A showed no activity, similar to AtKAI2, as previously reported[55].

Surprisingly, PsKAI2B is able to cleave (±)-GC240 probe in a similar manner as AtD14 and RMS3. It has been previously demonstrated that probes without a methyl group, such as dYLG, can serve as the hydrolysis substrate for AtKAI2[53]. To that end, we used the (±)-GC486 probe bearing no methyl on D-ring, and notably, PsKAI2B was able to hydrolyze the probe, whereas PsKAI2A shows little to no activity (**Figure S5c-d**). Furthermore, PsKAI2A and AtKAI2 exhibit biphasic time course of fluorescence, consisting of an initial phase, followed by a plateau phase. By comparing the kinetics profiles, we noticed that with PsKAI2B, RMS3 and AtD14 proteins, the plateau is higher (1 µM versus 0.3 µM of DiFMU), even if it takes PsKAI2B longer to reach this plateau (**Figure S5c-d**). Taken together with the comparative kinetic analysis, PsKAI2B hydrolysis activity is more similar to SL receptors and further highlights the distinct function compared to PsKAI2A.

Structural insights into legume KAI2s divergence

To elucidate the differential ligand selectivity between KAI2A and KAI2B, we first determined the legume crystal structure of *Pisum sativum* KAI2B at 1.6Å resolution (**Figure 4** and **Table 1**). The PsKAI2B structure shares the canonical α/β hydrolase fold and is comprised of base and lid domains (**Figure 4a**). The core domain contains seven-stranded mixed β-sheets (β1–β7), five α-helices (αA, αB, αC, αE and αF) and five 3_{10} helices (ŋ1, ŋ2, ŋ3, ŋ4, and ŋ5). The helical lid domain (residues 124–195, **Figure S2**) is positioned between strands β6 and β7 and forms two parallel layers of V-shaped helices (αD1-4) that create a deep pocket area adjoining the conserved catalytic Ser-His-Asp triad site (**Figure 4a** and **Figure S2**). Despite the sequence variation (77% similarity between PsKAI2B and AtKAI2, **Figure S2**), we did not observe major structural rearrangements between PsKAI2B and the previously determined Arabidopsis KAI2

structure[48] as shown by an Root Mean Squared Deviation (RMSD) of 0.35 Å for superposition of backbone atoms (**Figure 4b**). Nonetheless, further structural comparative analyses have identified two unique residues alterations in positions 129 and 147 within the lid domain. These changes appear to marginally alter the backbone atoms and distinguish legume KAI2s family from other KAI2s species (**Figure S2** and **Figure 4b**). The asparagine residue in position 129 is more variable within legume KAI2s, and alanine or serine in position 147 has diverged from bulky polar residues compared to other plant KAI2s. These amino acids alterations are likely to play role in downstream events rather than directly modulate distinct ligand perception.

To further determine the differential ligand specificity between PsKAI2A and PsKAI2B, we utilized the PsKAI2B crystal structure reported here to generate a 3D model for PsKAI2A. As expected, PsKAI2A structure exhibits a similar backbone atom arrangement (RMSD of 0.34 Å) that parallels the PsKAI2B structure (**Figure 5a**). Nonetheless, we identified eight significant divergent amino acids between the two structures including residues involved in forming the ligand binding pocket as well as solvent-exposed surfaces (**Figure 5b-d** and **Figure S6a-b)**. Because these variants are evolutionarily conserved across legume, the analysis of the underlined residues not only distinguishes between KAI2A and KAI2B in *Pisum* but can be extrapolated to all legume KAI2A/B diverged proteins. Structural comparative analysis within the ligand-binding pocket shows divergent solvent accessibility between PsKAI2A and PsKAI2B (**Figure 5b**). PsKAI2B exhibits a structural arrangement that results in a larger volume of the hydrophobic pocket (125.4 A^3) yet with a smaller entrance circumference (30.3 Å) than PsKAI2A (114.8 A^3 and 33.6 Å, respectively, **Figure 5b**). Further *in silico* docking experiments of (–)-GR24 with PsKAI2B results in a successful docking of the ligand that is totally buried in the pocket and positioned in a pre-hydrolysis orientation nearby the catalytic triad. In contrast,

docking experiments of (–)-GR24 with PsKAI2A results in more restricted interaction where the ligand is partially outside the pocket (**Figure S6c)**. Notably, there are five key residues that are found to directly alter the pocket morphology (**Figure 5c** and **Figure S6a-b**). Among these residues, L160/S190/M218 in PsKAI2A and the corresponding residues, M160/L190/L218 in PsKAI2B are of particular interest because of their functional implications in the pocket volume and solvent accessibility (**Figure 5d**). Residue 160 is positioned at the entrance of the ligand-binding pocket in helix αD2, thus the substitution of leucine (L160 in KAI2A) to methionine (M160 in KAI2B) results in modifying the circumference of PsKAI2B pocket entrance (**Figure 5b-d**). While both L160 and M160 represent aliphatic non-polar residues, the relative low hydrophobicity of methionine as well as its higher plasticity are likely to play major role in modifying the ligand pocket. The conserved legume divergence in residue 190 (S190 in PsKAI2A and L190 in PsKAI2B, **Figure 5d**) is positioned in helix αD4 and represents a major structural arrangement at the back of the ligand envelope. Because leucine has moderate flexibility compared to serine and much higher local hydrophobicity, this variation largely attributes to the changes in the pocket volume as well as fine-tunes available ligand orientations. Further sequence and structural analysis of the variant in position 218 (M218 in KAI2A and L218 in PsKAI2B) placed it in the center of the Asp loop[56] (D-loop, region between β7 and αE, **Figure 5c-d**). In D14, the D-loop has been reported to affect SL perception and cleavage as well as impact protein-protein interactions in SL signaling[51,56]

PsKAI2B forms a complex with the D-OH of (–)-GR24

To further examine the molecular interaction of PsKAI2B with the enantiomeric GR24, we co-crystallized and solved the structure of PsKAI2B-(–)-GR24 at 2.0Å resolution (**Figure 6a** and

Table 1). Electron density map analysis of the ligand-binding pocket revealed the existence of a unique ring-shaped occupancy that is contiguously linked to the catalytic serine (S95) (**Figure 6a-b**). The structural comparison of the backbone atoms between apo-PsKAI2B and PsKAI2B-(–)-GR24 did not reveal significant differences (**Figure S7a**) and is in agreement with previously reported apo and ligand bound D14/KAI2 crystal structures[11,12,19,36,50]. This striking similarity suggests that a major conformational change, if indeed occurs as suggested for D14[51], may happen after the nucleophilic attack of the catalytic serine and the (–)-GR24 cleavage which is likely to be highly unstable state for crystal lattice formation.

Further analysis suggests that 5-hydroxy-3-methylbutenolide (D-OH ring), resulting from the (–)-GR24 cleavage, is trapped in the catalytic site (**Figure S7b-d)**. The lack of a defined electron density fitting with the tricyclic lactone (ABC ring) may exclude the presence of the intact GR24 molecule. Other compounds present in the crystallization condition were tested for their ability to occupy the SER95-contiguous density, and D-OH group of (–)-GR24 demonstrated the highest correlation coefficient calculated score and the best fit in the PsKAI2B co-crystal structure (**Figure S7c**). Additional tests of D-OH binding including *in silico* docking simulations and analyses revealed a high affinity for D-OH in a specific orientation and in agreement with the structure presented here (**Figure S7d**). The most probable orientation of the D-OH positions the methyl group (C4') together with the hydroxyl group of D-OH towards the very bottom/back of the pocket near the catalytic serine, where the O5" atom is coordinated by both N atoms of F26 and V96 (**Figure 6b-c**). The hemiacetal group (C2') of D-OH is oriented towards the access groove of the pocket with angles (between carbon and oxygen atoms) supporting the captured D-OH in an orientation in which cleavage of the intact (–)-GR24 may have taken place.

The C5' of D-OH appears to form a covalent bond with Oγ of S95 (dark gray line in **Figure 6c**) and generates a tetrahedral carbon atom. The overall positioning of this molecule is strictly coordinated by F26, H246, G25, and I193 residues. Remarkably, the electron density around the S95 does not display an open D-OH group (2,4,4,-trihydroxy- 3-methyl-3-butenal as previously described for OsD14[19]) that could directly result from the nucleophilic attack event, but rather correspond to a cyclized D-OH ring linked to the S95. This D-OH ring is likely to be formed by water addition to the carbonyl group at C2' that is generated after cleavage of the enol function and cyclization to re-form the butenolide (**Figure 6d**). The formation of this adduct could also serve as an intermediate before the transfer to the histidine residue. Taken together, our crystal structure highlights a potential new intermediate in the ligand cleavage mechanism by KAI2 proteins.

Discussion

The emerging characterization of karrikin/KL signaling in non-fire ecology plant receptors has been of great interest in the plant signaling field. While there are many missing pieces in the karrikin signaling puzzle, it is clear that KAI2 serves as the key sensor in this pathway. Furthermore, the coevolution between receptors and ligands in diverse contexts throughout plant evolution is of great interest in many biological fields. The limited natural occurrence of karrikin molecules and the evolutionary conservation of KAI2 receptors throughout land plants suggest that the function of KAI2s are preserved to regulate plant development and response to stresses by perceiving an endogenous ligand(s) (KL). Here, we identified and characterized the first KAI2 receptors in pea (*P. sativum*) that serve as representatives of the independent duplication event and subsequent sub-functionalization in legumes. The identification of both *PsKAI2A* and

PsKAI2B genes corroborates the recent finding that the *KAI2* gene duplication event occurred in Papilionoidaea before the diversification of legumes[42]. Interestingly, similarities in expression patterns are found between pea and lotus with global higher expression of the A clade in comparison to the B clade and specific expression in roots of the B clade in comparison to the A clade. Further studies with PsKAI2A/B mutants i.e. for the establishment of the symbioses in *Pisum* roots could explain this differential expression in roots, as no clear root phenotype has been observed in lotus. The occurrences of molecular coevolution of ligands and their specialized receptors have been previously demonstrated for phytohormones such as SL[57], ABA[58], GA[59], and more recently, karrikins[39,42]. Even though the exact identity of KL ligands remains to be revealed, it is likely that the ligands share a common chemical composition to SLs. It has been shown that the synthetic SL analogue, *rac*-GR24, can function by binding KAI2 in Arabidopsis[11,43,53]. In this work we carried out a comprehensive biochemical interrogation and found that PsKAI2B can form stronger interactions with the enantiomeric GR24, (–)-GR24, compared to PsKAI2A. Moreover, we found that while both KAI2s are active hydrolases, they have distinct binding affinity and stereoselectivity towards GR24 stereoisomers. These findings indicate yet again, that sub-functionalization of KAI2s via substitutions in only few amino acids can greatly alter ligand affinity, binding, enzymatic activity, and probably signaling with downstream partners[38,42].

KAI2/D14 crystal structures have greatly impacted our understanding of these receptor ligand-binding pockets and their ability to not only accommodate, but also hydrolyze certain ligands[11,12,19,31,36,38,44–51]. The first crystal structure of legume PsKAI2B together with the PsKAI2A homology model reported here, further substantiates the structural basis of this differential ligand selectivity. We identified conserved key amino acid changes that alter the

shape of the pocket and confer altered ligand specificities. These novel atomic structures of KAI2 enabled us to analyze the distinction between key residues L160/S190/M218 in PsKAI2A and the corresponding residues M160/L190/L218 in PsKAI2B. These findings further support recent in planta and biochemical studies that demonstrate that residues 160 and 190 are required for differential ligand specificity between lotus KAI2A and KAI2B[42]. Furthermore, residue 190 was also identified in the parasitic plant *Striga hermonthica* as being involved in forming differential specificity pockets between the highly variable and functionally distinct ShKAI2s, referred to as HTLs[35,36]. While the changes in positions 160 and 190 directly reshape the pocket morphology, the variant in position 218 is located in the center of the D-loop[56]. The D-loop contains the aspartic acid of the catalytic triad (D217) and has been suggested to play an important role in SL perception and cleavage by D14 as well as downstream protein-protein interactions[51,56]. Therefore, the conserved substitution of KAI2A and KAI2B in M218 to L218 respectively across legumes not only contributes to ligand selectivity and hydrolysis, but may also affect downstream interaction(s).

Based on the analogy with the D14-MAX2 perception mechanism, the KAI2 receptor is likely to adopt different conformational states upon ligand binding and cleavage. As such, the identification of unique residue variations in the lid (between KAI2A and KAI2B, respectively in positions 129 and 147) reported here, infer a sub-functionalization in the receptor regions that are likely to be involved in MAX2 and/or SMAX1 and/or SMXL2 downstream interactions. Therefore, it remains to be further elucidated whether these KAI2A/B distinctive residues play a role in fine tuning the formation of the protein complex with MAX2-SMAX1/SMXL2.

The crystal structure of ligand bound PsKAI2B provides a mechanistic view of perception and cleavage by KAI2s. Based on the crystallization conditions and following a

detailed investigation of the electron density, we were able to overrule common chemicals and place the (–)-GR24 D-OH ring with higher relative fitting values than other components. The absence of positive electron density peaks corresponding to the intact (–)-GR24, and thus the presence of only the D-OH, raise questions of whether the S95-D-OH adduct recapitulates a pre- or post- cleavage intermediate state of (–)-GR24. The possibility that the trapped molecule represents a post cleavage state is intriguing and may provide a new intermediate state where S95 is covalently linked to the cleavage product. As such, the S95-D-OH adduct could explain the single turnover cycle that was observed for KAI2s in this study. Previous studies of the single turnover activity of D14 suggest that a covalent intermediate is formed between the catalytic histidine and serine[51]. The chemical similarity of the D-OH butenolide ring of karrikin and GR24 suggests that the KL signal may share a parallel structure and perhaps will be biochemically processed via multiple steps and intermediate adducts. Therefore, the significance of this study may also reveal a similar mechanism regarding SL perception and cleavage by D14.

Our data *in planta* clearly demonstrate that PsKAI2A and PsKAI2B genes can replace the AtKAI2 ortholog, yet we were unable to conclude KAI2A/B ligand binding specificity by using the KL mimic compound (–)-GR24. The ambiguity in detecting ligand specificity *in vivo* is likely to remain a challenge in the karrikin field until the identification of endogenous KL. Once KL(s) will be revealed, it will be important to test the response of the Arabidopsis complementation lines to KL(s) and further validate the function of the key residues L160/S190/M218 *in planta*. Additionally, future studies with pea mutants will elucidate PsKAI2A and PsKAI2B functional divergence and reveal the distinct physiological functions, and in particularly the symbiotic relationship with AM fungi, that could shed light on the differential expression patterns in the roots.

This study illuminates the complex evolution of KAI2s in plants and particularly in legumes. We provide comprehensive structural and biochemical evidence of the specialization and sub-functionalization of KAI2 receptors and their sensitivity to butenolide compounds. Because of their ability to fix atmospheric nitrogen through plant–rhizobium symbiosis, legume crops such as pea or fava bean are attracting increasing attention for their agroecological potential. Thus, better understanding of KAR/KL perception and signaling in these staple crops may have far-reaching impacts on agro-systems and food security.

Methods

Protein sequence alignment and phylogenetic tree analyses

Representative KAI2 sequences of 41 amino acid sequences were downloaded from Phytozome and specific genome databases as shown in **Figure S1**. Alignment was performed in MEGA X[60] using the MUSCLE multiple sequence alignment algorithm[61]. Sequence alignment graphics were generated using CLC Genomics Workbench v12. The evolutionary history was inferred by using the Maximum Likelihood method and JTT matrix-based model[62]. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the JTT model, and then selecting the topology with superior log likelihood value. The percentage of trees in which the associated taxa clustered together is shown next to the branches[63]. Tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Analysis involved 41 amino acid sequences with a total of 327 positions in the final dataset. Evolutionary analyses were conducted in MEGA X[60].

Constructs and generation of transgenic lines

The expression vectors for transgenic Arabidopsis were constructed by MultiSite Gateway Three-Fragment Vector Construction kit (Invitrogen). *AtKAI2* and *PsKAI2A.2* constructs were tagged with 6xHA epitope tag or mCitrine protein at their C-terminus. Lines were resistant to hygromycin.

The *AtKAI2* native promoter (0.7 kb) was amplified by PCR with the primer AtKAI2_promo_attB4 (5'-ggggacaactttgtatagaaaagttgccTTCACGACCAGTATGGTTTACTCA-3') and AtKAI2_promo_attB1R (5'-ggggactgctttttgtacaaacttgcCTCTCTAAAGAAGATTCTTCTCTGGTT-3') from Col-0 genomic DNA and cloned into the pDONR-P4P1R vector, using Gateway recombination (Invitrogen). The 6XHA with linker and mCitrine tags were cloned into pDONR-P2RP3 (Invitrogen) as described in de Saint Germain et al.[55]. *PsKAI2A.1*, *PsKAI2A.2* and *PsKAI2B* CDS were PCR amplified from *Pisum* cv. Térèse cDNA with the primers PsKAI2A_attB1 (5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTtcATGGGGATAGTGGAAGAAGCA-3'); PsKAI2A.1_attB2_STOP (5'-ggggaccactttgtacaagaaagctgggtcCAAATCTGCCTCAAGTTTCA-3'); PsKAI2A.2_attB2_STOP (5'-ggggaccactttgtacaagaaagctgggtcCCTTATTGGCTCAATATTAA-3'); PsKAI2b_attB1 (5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTtcATGGGAATAGTGGAAGAAGC-3'); PsKAI2B_attB2_STOP (5'-ggggaccactttgtacaagaaagctgggtcAGCTACAATATCATAACGAA-3'); and the *AtKAI2* CDS was PCR amplified from Col-0 cDNA with the primers AtKAI2_attB1 (5'-ggggacaagtttgtacaaaaaagcaggcttcATGGGTGTGGTAGAAGAAGC-3') and AtKAI2_attB2_ΔS (5'-ggggaccactttgtacaagaaagctgggtcCATAGCAATGTCATTACGAAT-3') and then recombined into the pDONR221 vector (Invitrogen). The suitable combination of *AtKAI2* native promoter, *AtKAI2*, *PsKAI2A.1*, *PsKAI2A.2* or *PsKAI2B* and 6XHA or mCitrine

was cloned into the pH7m34GW final destination vectors by using the three fragment recombination system[64] and were thusly named pAtKAI2::AtKAI2-6xHA, pAtKAI2::AtKAI2-mCitrine, pAtKAI2::PsKAI2A.1-6xHA, pAtKAI2::PsKAI2b-6xHA and pAtKAI2::PsKAI2A.2-mCitrine. Transformation of Arabidopsis Atkai2-2 mutant was performed according to the conventional floral dipping method[65], with Agrobacterium strain GV3101. For each construct, only a few independent T1 lines were isolated and all lines were selected in T2. Phenotypic analysis shown in **Figure 1e** was performed on the T3 homozygous lines.

Hypocotyl elongation assays.

Arabidopsis seeds were surface sterilized by consecutive treatments of 5 min 70% (v/v) ethanol with 0.05% (w/v) sodium dodecyl sulfate (SDS) and 5 min 95% (v/v) ethanol. Then seeds were sown on half-strength Murashige and Skoog (½ MS) media (Duchefa Biochemie) containing 1% agar, supplemented with 1 µM (–)-GR24 or with 0.01 % DMSO (control). Seeds were stratified at 4 °C (2 days in dark) then transferred to the growth chamber at 22 °C, under 20-30 µE /m^2/sec of white light in long day conditions (16 hr light/ 8 hr dark). Seedlings were photographed and hypocotyl lengths were quantified using ImageJ[66]. 2 plates of 10-12 seeds were sown for each genotype x treatment. Using Student t-tests, no statistically significantly different means were detected between plates. The data from the 20-24 seedlings were then used for a one-way ANOVA.

Chemicals

Enantiopure GR24 isomers were obtained as described in de Saint Germain et al.[55] or purchased from StrigoLab. Profluorescent probes (GC240, GC486) were obtained as described in de Saint Germain et al.[55].

Protein preparation and purification

PsKAI2A.2 and PsKAI2B were independently cloned and expressed as a 6× His-SUMO fusion proteins from the expression vector pAL (Addgene). These were cloned utilizing primers PsKAI2A_F (5'-aaaacctctacttccaatcgATGGGGATAGTGGAAGAAG-3'), PsKAI2A.1_R (5'-ccacactcatcctccggTTACAAATCTGCCTCAAGTTTC-3'), PsKAI2A.2_R (5'-ccacactcatcctccggTTACCTTATTGGCTCAATATTAAGTTG-3'), PsKAI2B_F (5'-aaaacctctacttccaatcgATGGGAATAGTGGAAGAAGC-3'), and PsKAI2B_R (5'-ccacactcatcctccggTCAAGCTACAATATCATAACGAATG-3').

BL21 (DE3) cells transformed with the expression plasmid were grown in LB broth at 16 °C to an OD_{600} of ~0.8 and induced with 0.2 mM IPTG for 16 h. Cells were harvested, re-suspended and lysed in extract buffer (50 mM Tris, pH 8.0, 200 mM NaCl, 5 mM imidazole, 4% Glycerol). All His-SUMO-PsKAI2s were isolated from soluble cell lysate by Ni-NTA resin. The His-SUMO-PsKAI2 was eluted with 250 mM imidazole and subjected to anion-exchange. The eluted protein was than cleaved with TEV (tobacco etch virus) protease overnight at 4 °C. The cleaved His-SUMO tag was removed by passing through a Nickel Sepharose and PsKAI2 was further purified by chromatography through a Superdex-200 gel filtration column in 20 mM HEPES, pH 7.2, 150 mM NaCl, 5 mM DTT, 1% Glycerol. All proteins were concentrated by ultrafiltration to 3–10 mg/mL^{-1}. RMS3, AtD14, AtKAI2 were expressed in bacteria with TEV cleavable GST tag, purified and used as described in de Saint Germain et al.[55].

Enzymatic degradation of GR24 isomers by purified proteins

Ligands (10 µM) were incubated without and with purified proteins (5 µM) for 150 min at 25 ºC in PBS (0.1 mL, pH 6.8) in presence of (±)-1-indanol (100 µM) as the internal standard. The solutions were acidified to pH 1 with 10% trifluoroacetic acid in CH_3CN (v/v) (2 µL) to quench the reaction and centrifuged (12 min, 12,000 tr/min). Thereafter the samples were subjected to RP-UPLC-MS analyses using Ultra Performance Liquid Chromatography system equipped with a PDA and a Triple Quadrupole mass spectrometer Detector (Acquity UPLC-TQD, Waters, USA). RP-UPLC (HSS C_{18} column, 1.8 µm, 2.1 mm × 50 mm) with 0.1% formic acid in CH_3CN and 0.1% formic acid in water (aq. FA, 0.1%, v/v, pH 2.8) as eluents [10% CH_3CN, followed by linear gradient from 10 to 100% of CH_3CN (4 min)] was carried out at a flow rate of 0.6 mL/min. The detection was performed by PDA using the TQD mass spectrometer operated in Electrospray ionization positive mode at 3.2 kV capillary voltage. The cone voltage and collision energy were optimized to maximize the signal and were respectively 20 V for cone voltage and 12 eV for collision energy and the collision gas used was argon at a pressure maintained near $4.5.10^{-3}$ mBar.

Enzymatic assay with pro-fluorescent probes

Enzymatic assay and analysis have been carried out as described in de Saint Germain et al.[55], using a TriStar LB 941 Multimode Microplate Reader from Berthold Technologies. The experiments were repeated three times.

Protein melting temperatures

Differential Scanning Fluorimetry (DSF) experiments were performed on a CFX96 TouchTM Real-Time PCR Detection System (Bio-Rad Laboratories, Inc., Hercules, California, USA) using excitation and emission wavelengths of 490 and 575 nm, respectively. Sypro Orange ($\lambda_{ex}/\lambda_{em}$: 470/570 nm; Life Technologies Co., Carlsbad, California, USA) was used as the reporter dye. Samples were heat-denatured using a linear 25 to 95 °C gradient at a rate of 1.3 °C per minute after incubation at 25 °C for 30 min in the absence of light. The denaturation curve was obtained using CFX manager™ software. Final reaction mixtures were prepared in triplicate in 96-well white microplates, and each reaction was carried out in 20 µL scale in Phosphate buffer saline (PBS) (100 mM Phosphate, pH 6.8, 150 mM NaCl) containing 6 µg protein (such that final reactions contained 10 µM protein), 0-1000 µM ligand (as shown on the **Figure 2a-h**), 4% (v/v) DMSO, and 0.008 µL Sypro Orange. Plates were incubated in darkness for 30 minutes before analysis. In the control reaction, DMSO was added instead of ligand. The experiments were repeated three times.

Intrinsic tryptophan fluorescence assays and kinetics

Intrinsic tryptophan fluorescence assays and determination of the dissociation constant K_D has been performed as described in de Saint Germain et al.[55], using the Spark® Multimode Microplate Reader from Tecan.

Crystallization, data collection and structure determination

The crystals of PsKAI2B were grown at 25 °C by the hanging-drop vapor diffusion method with 1.0 µL purified protein sample mixed with an equal volume of reservoir solution containing 0.1 M HEPES pH 7.5, 2.75% v/v PEG 4000, 2.75% v/v PEG-ME 5000. The crystals of PsKAI2B in

complex with (–)-GR24 were grown at 25 °C by the hanging-drop vapor diffusion method with 1.0 μL purified protein complex (preincubated with 1 mM (–)-GR24, StrigoLab) and mixed with an equal volume of reservoir solution containing 0.1 M HEPES pH 7.5, 2.75% PEG 2000, 2.75% v/v PEG-ME 5000, 1mM (–)-GR24. Crystals of maximum size were obtained and harvested after 2 weeks from the reservoir solution with additional 20% MPD serving as cryoprotectant. X-ray diffraction data was integrated and scaled with HKL2000 package[67]. PsKAI2s crystal structures were determined by molecular replacement using the AtKAI2 model (PDB: 5Z9H)[68] as the search model. All structural models were manually built, refined, and rebuilt with PHENIX[69] and COOT[70].

Structural biology modelling and analyses

Model structure illustrations were made by PyMOL[71]. PsKAI2A model structure was generated using iTASSER[72–74]. Ligand identification, ligand-binding pocket analyses, and computing solvent accessible surface values analyses were carried out using Phenix LigandFit[69,75,76], CASTp software[77,78], and AutoDock Vina[79], respectively. LigPlot+ program[80] was used for 2-D representation of protein-ligand interactions from standard PDB data format.

Data Availability

The atomic coordinates of apo and ligand-bound forms of PsKAI2 structures has been deposited in the Protein Data Bank with accession codes 7K2Z and 7K38, respectively. All relevant data are available from corresponding authors upon request.

Acknowledgements

We thank the beamline staff at ALS for help with data collection. This work is supported by UC Davis new faculty start-up funds. The Shabek laboratory is supported by National Science Foundation. This work is supported by the Institut Jean-Pierre Bourgin's Plant Observatory technological platforms. F.-D.B. is supported by CHARM3AT Labex program (ANR-11-LABX-39). A.d.S.G. is supported by AgreenSkills from the European Union in the framework of the Marie-Curie FP7 COFUND People Programme and fellowship from Saclay Plant Sciences (ANR-17-EUR-0007).

Author Contributions

AM.G., F.-D.B., C.R., A.dS.G., and N.S. conceived and designed the experiments. N.S., A.dS.G., and AM.G. conducted the protein purification, biochemical and crystallization experiments. N.S. and AM.G. determined and analyzed crystal structures and conducted *in silico* studies. AM.G., A.dS.G., and N.S. wrote the manuscript with the help from all other co-authors.

Author Information

Authors declare no competing interests. Correspondence and requests for materials should be addressed to N.S. (nshabek@ucdavis.edu) and A.dS.G. (Alexandre.De-Saint-Germain@inrae.fr).

References

1. Flematti, G. R., Ghisalberti, E. L., Dixon, K. W. & Trengove, R. D. A compound from smoke that promotes seed germination. *Science (80-.).* **305**, 977 (2004).
2. Nelson, D. C. *et al.* Karrikins enhance light responses during germination and seedling development in Arabidopsis thaliana. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 7095–7100 (2010).
3. Sun, X. D. & Ni, M. HYPOSENSITIVE to LIGHT, an alpha/beta fold protein, acts downstream of ELONGATED HYPOCOTYL 5 to regulate seedling de-etiolation. *Mol. Plant* **4**, 116–126 (2011).
4. Waters, M. T. *et al.* Specialisation within the DWARF14 protein family confers distinct responses to karrikins and strigolactones in Arabidopsis. *Development* **139**, 1285–1295 (2012).
5. Flematti, G. R. *et al.* Preparation of 2H-furo[2,3-c]pyran-2-one derivatives and evaluation of their germination-promoting activity. *J. Agric. Food Chem.* **55**, 2189–2194 (2007).
6. Flematti, G. R., Scaffidi, A., Dixon, K. W., Smith, S. M. & Ghisalberti, E. L. Production of the seed germination stimulant karrikinolide from combustion of simple carbohydrates. *J. Agric. Food Chem.* **59**, 1195–1198 (2011).
7. Dixon, K. W., Merritt, D. J., Flematti, G. R. & Ghisalberti, E. L. Karrikinolide - A phytoreactive compound derived from smoke with applications in horticulture, ecological restoration and agriculture. *Acta Hortic.* **813**, (2009).
8. Stevens, J. C., Merritt, D. J., Flematti, G. R., Ghisalberti, E. L. & Dixon, K. W. Seed germination of agricultural weeds is promoted by the butenolide 3-methyl-2H-furo[2,3-c]pyran-2-one under laboratory and field conditions. *Plant Soil* **298**, 113–124 (2007).
9. Long, R. L. *et al.* Prior hydration of Brassica tournefortii seeds reduces the stimulatory effect of karrikinolide on germination and increases seed sensitivity to abscisic acid. *Ann. Bot.* **105**, 1063–1070 (2010).
10. Nelson, D. C. *et al.* F-box protein MAX2 has dual roles in karrikin and strigolactone signaling in Arabidopsis thaliana. *Proc. Natl. Acad. Sci.* **108**, 8897–8902 (2011).
11. Guo, Y., Zheng, Z., La Clair, J. J., Chory, J. & Noel, J. P. Smoke-derived karrikin perception by the a/B hydrolase KAI2 from Arabidopsis. *Proc. Natl. Acad. Sci.* **110**, 8284–8289 (2013).

12. Kagiyama, M. *et al.* Structures of D14 and D14L in the strigolactone and karrikin signaling pathways. *Genes to Cells* **18**, 147–160 (2013).
13. Stanga, J. P., Smith, S. M., Briggs, W. R. & Nelson, D. C. SUPPRESSOR OF MORE AXILLARY GROWTH2 1 Controls Seed Germination and Seedling Development in Arabidopsis. *Plant Physiol.* **163**, 318–330 (2013).
14. Gutjahr, C. *et al.* Rice perception of symbiotic arbuscular mycorrhizal fungi requires the karrikin receptor complex. *Science (80-.).* **350**, 1521–1524 (2015).
15. Li, W. *et al.* The karrikin receptor KAI2 promotes drought resistance in Arabidopsis thaliana. *PLoS Genet.* **13**, e1007076 (2017).
16. Wang, L., Waters, M. T. & Smith, S. M. Karrikin-KAI2 signalling provides Arabidopsis seeds with tolerance to abiotic stress and inhibits germination under conditions unfavourable to seedling establishment. *New Phytol.* **219**, 605–618 (2018).
17. Scaffidi, A. *et al.* Exploring the molecular mechanism of karrikins and strigolactones. *Bioorganic Med. Chem. Lett.* **22**, 3743–3746 (2012).
18. Yoneyama, K. Recent progress in the chemistry and biochemistry of strigolactones. *J. Pestic. Sci.* **45**, 45–53 (2020).
19. Zhao, L. H. *et al.* Crystal structures of two phytohormone signal-transducing α/β hydrolases: Karrikin-signaling KAI2 and strigolactone-signaling DWARF14. *Cell Res.* **23**, 436–439 (2013).
20. Cook, C. E., Whichard, L. P., Turner, B., Wall, M. E. & Egley, G. H. Germination of witchweed (striga lutea lour.): Isolation and properties of a potent stimulant. *Science (80-.).* **154**, 1189–1190 (1966).
21. Sorefan, K. *et al.* MAX4 and RMS1 are ortholosgous dioxygenase-like genes that regulate shoot branching in Arabidopsis and pea. *Genes Dev.* **17**, 1469–1474 (2003).
22. Kapulnik, Y. *et al.* Strigolactones interact with ethylene and auxin in regulating root-hair elongation in Arabidopsis. *J. Exp. Bot.* (2011) doi:10.1093/jxb/erq464.
23. Rasmussen, A. *et al.* Strigolactones suppress adventitious rooting in arabidopsis and pea. *Plant Physiol.* **158**, 1976–1987 (2012).
24. Lopez-Obando, M., Ligerot, Y., Bonhomme, S., Boyer, F. D. & Rameau, C. Strigolactone biosynthesis and signaling in plant development. *Dev.* (2015) doi:10.1242/dev.120006.
25. Akiyama, K., Matsuzaki, K. I. & Hayashi, H. Plant sesquiterpenes induce hyphal

branching in arbuscular mycorrhizal fungi. *Nature* **435**, 824–827 (2005).

26. Gomez-Roldan, V. *et al.* Strigolactone inhibition of shoot branching. *Nature* **455**, 189–194 (2008).
27. Arite, T. *et al.* D14, a strigolactone-Insensitive mutant of rice, shows an accelerated outgrowth of tillers. *Plant Cell Physiol.* **50**, 1416–1424 (2009).
28. Besserer, A. *et al.* Strigolactones stimulate arbuscular mycorrhizal fungi by activating mitochondria. *PLoS Biol.* **4**, e226 (2006).
29. Li, S. W., Xue, L., Xu, S., Feng, H. & An, L. Mediators, genes and signaling in adventitious rooting. *Bot. Rev.* **75**, 230–247 (2009).
30. Agusti, J. *et al.* Strigolactone signaling is required for auxin-dependent stimulation of secondary growth in plants. *Proc. Natl. Acad. Sci. U. S. A.* **180**, 20242–20247 (2011).
31. Hamiaux, C. *et al.* DAD2 is an α/β hydrolase likely to be involved in the perception of the plant branching hormone, strigolactone. *Curr. Biol.* **22**, 2032–2036 (2012).
32. Kapulnik, Y. *et al.* Strigolactones affect lateral root formation and root-hair elongation in Arabidopsis. *Planta* **233**, 209–216 (2011).
33. Bythell-Douglas, R. *et al.* Evolution of strigolactone receptors by gradual neo-functionalization of KAI2 paralogues. *BMC Biol.* **15**, 1–21 (2017).
34. Swarbreck, S. M., Guerringue, Y., Matthus, E., Jamieson, F. J. C. & Davies, J. M. Impairment in karrikin but not strigolactone sensing enhances root skewing in Arabidopsis thaliana. *Plant J.* **98**, 607–621 (2019).
35. Toh, S. *et al.* Structure-function analysis identifies highly sensitive strigolactone receptors in Striga. *Science (80-.).* **350**, 203–207 (2015).
36. Xu, Y. *et al.* Structural basis of unique ligand specificity of KAI2-like protein from parasitic weed Striga hermonthica. *Sci. Rep.* **6**, 1–9 (2016).
37. Waters, M. T. *et al.* A selaginella moellendorffii ortholog of KARRIKIN INSENSITIVE2 functions in arabidopsis development but cannot mediate responses to karrikins or strigolactones. *Plant Cell* **27**, 1925–1944 (2015).
38. Bürger, M. *et al.* Structural Basis of Karrikin and Non-natural Strigolactone Perception in Physcomitrella patens. *Cell Rep.* **26**, 855–865 (2019).
39. Sun, Y. K. *et al.* Divergent receptor proteins confer responses to different karrikins in two ephemeral weeds. *Nat. Commun.* **11**, (2020).

40. de Saint Germain, Alexandre, Jacobs, A., Brun, G. & Boyer, F.-D. A Phelipanche ramosa KAI2 Protein Perceives enzymatically Strigolactones and 2 Isothiocyanates. *bioRxiv* (2020) doi:10.1101/2020.06.09.136473.
41. Sun, Y. K., Flematti, G. R., Smith, S. M. & Waters, M. T. Reporter gene-facilitated detection of compounds in arabidopsis leaf extracts that activate the karrikin signaling pathway. *Front. Plant Sci.* **7**, 1799 (2016).
42. Carbonnel, S. *et al.* Lotus japonicus karrikin receptors display divergent ligand-binding specificities and organ-dependent redundancy. *bioRxiv* 754937 (2020) doi:10.1101/754937.
43. Conn, C. E. & Nelson, D. C. Evidence that KARRIKIN-INSENSITIVE2 (KAI2) Receptors may Perceive an Unknown Signal that is not Karrikin or Strigolactone. *Front. Plant Sci.* **6**, 1–7 (2016).
44. Shabek, N. *et al.* Structural plasticity of D3–D14 ubiquitin ligase in strigolactone signalling. *Nature* **563**, 652–656 (2018).
45. Xu, Y. *et al.* Structural analysis of HTL and D14 proteins reveals the basis for ligand selectivity in Striga. *Nat. Commun.* **9**, 3947 (2018).
46. Takeuchi, J. *et al.* Rationally designed strigolactone analogs as antagonists of the D14 receptor. *Plant Cell Physiol.* **59**, 1545–1554 (2018).
47. Hamiaux, C. *et al.* Inhibition of strigolactone receptors by N-phenylanthranilic acid derivatives: Structural and functional insights. *J. Biol. Chem.* **293**, 6530–6543 (2018).
48. Bythell-Douglas, R. *et al.* The Structure of the Karrikin-Insensitive Protein (KAI2) in Arabidopsis thaliana. *PLoS One* **8**, e54758 (2013).
49. Nakamura, H. *et al.* Molecular mechanism of strigolactone perception by DWARF14. *Nat. Commun.* **4**, (2013).
50. Zhao, L. H. *et al.* Destabilization of strigolactone receptor DWARF14 by binding of ligand and E3-ligase signaling effector DWARF3. *Cell Res.* **25**, 1219–1236 (2015).
51. Yao, R. *et al.* DWARF14 is a non-canonical hormone receptor for strigolactone. *Nature* **536**, 469–473 (2016).
52. Kreplak, J. *et al.* A reference genome for pea provides insight into legume genome evolution. *Nat. Genet.* **51**, 1411–1422 (2019).
53. Yao, J. *et al.* An allelic series at the KARRIKIN INSENSITIVE 2 locus of Arabidopsis

thaliana decouples ligand hydrolysis and receptor degradation from downstream signalling. *Plant J.* **96**, 75–89 (2018).

54. Scaffidi, A. *et al.* Strigolactone hormones and their stereoisomers signal through two related receptor proteins to induce different physiological responses in arabidopsis. *Plant Physiol.* **165**, 1221–1232 (2014).
55. De Saint Germain, A. *et al.* An histidine covalent receptor and butenolide complex mediates strigolactone perception. *Nat. Chem. Biol.* **12**, 787–794 (2016).
56. Seto, Y. *et al.* Strigolactone perception and deactivation by a hydrolase receptor DWARF14. *Nat. Commun.* **10**, 191 (2019).
57. Conn, C. E. *et al.* Convergent evolution of strigolactone perception enabled host detection in parasitic plants. *Science (80-.).* **349**, 540–543 (2015).
58. Weng, J. K., Ye, M., Li, B. & Noel, J. P. Co-evolution of Hormone Metabolism and Signaling Networks Expands Plant Adaptive Plasticity. *Cell* **166**, 881–893 (2016).
59. Yoshida, H. *et al.* Evolution and diversification of the plant gibberellin receptor GID1. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E7844–E7853 (2018).
60. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).
61. Edgar, R. C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
62. Jones, D. T., Taylor, W. R. & Thornton, J. M. The rapid generation of mutation data matrices from protein sequences. *Bioinformatics* **8**, 275–282 (1992).
63. Felsenstein, J. Confidence Limits on Phylogenies: An Approach Using the Bootstrap. *Evolution (N. Y).* **39**, 783–791 (1985).
64. Karimi, M., Bleys, A., Vanderhaeghen, R. & Hilson, P. Building blocks for plant gene assembly. *Plant Physiol.* **145**, 1183–1191 (2007).
65. Clough, S. J. & Bent, A. F. Floral dip: A simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J.* **16**, 735–743 (1998).
66. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).
67. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation

mode. *Methods Enzymol.* **276**, 307–326 (1997).

68. Lee, I. *et al.* A missense allele of KARRIKIN-INSENSITIVE2 impairs ligand-binding and downstream signaling in Arabidopsis thaliana. *J. Exp. Bot.* **69**, 3609–3623 (2018).
69. Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 213–221 (2010).
70. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 (2010).
71. DeLano, W. L. The PyMOL Molecular Graphics System, Version 2.3. *Schrödinger LLC* (2020).
72. Yang, J. & Zhang, Y. I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Res.* **43**, W174–W181 (2015).
73. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: A unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5**, 725–738 (2010).
74. Yang, J. *et al.* The I-TASSER suite: Protein structure and function prediction. *Nat. Methods* **12**, 7–8 (2014).
75. Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. Electronic ligand builder and optimization workbench (eLBOW): A tool for ligand coordinate and restraint generation. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **65**, 1074–1080 (2009).
76. Terwilliger, T. C., Klei, H., Adams, P. D., Moriarty, N. W. & Cohn, J. D. Automated ligand fitting by core-fragment fitting and extension into density. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **62**, 915–922 (2006).
77. Binkowski, T. A., Naghibzadeh, S. & Liang, J. CASTp: Computed Atlas of Surface Topography of proteins. *Nucleic Acids Res.* **31**, 3352–3355 (2003).
78. Dundas, J. *et al.* CASTp: Computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res.* **34**, W116–W118 (2006).
79. Steffen, C. *et al.* AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **30**, 2785–2791 (2010).
80. Laskowski, R. A. & Swindells, M. B. LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* **51**, 2778–2786 (2011).

Figure Legends

Figure 1. Evolutionary analysis and differential expression of the legume *Pisum sativum* KAI2s. (a) Maximum likelihood phylogeny of 24 representative *KAI2* amino acid sequences. Node values represent percentage of trees in which the associated taxa clustered together. Vertical rectangles highlight distinct *KAI2* family clades. Black circle indicates legume duplication event. Pink and green circles mark the position of *PsKAI2As* and *PsKAI2B* respectively. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. **(b)** *PsKAI2A* and *PsKAI2B* are homologues to *AtKAI2* and encode α-β/hydrolases. Schematic representation of the *PsKAI2A* and *PsKAI2B* genes; Exons are in thick pink and green lines, intron colored in thin gray lines and UTR regions shown as thick gray lines. Bases are numbered from the start codon. *PsKAI2A* shows 2 splicing variants. Spliced introns are shown as bent ("V") lines. Bold lines represent intron retention. Inverted triangle (▼) indicates premature termination codons. **(c-d)** Differential expression pattern of *PsKAI2A* **(c,** pink**)** and *PsKAI2B* **(d,** green**)**. Transcript levels in the different tissues of 21 old wild-type Pisum sativum plants (cv. Terese) were determined by real-time PCR, relative to *PsEF1α*. Data are means ± SE (n = 2 pools of 8 plants). Inset drawing of a node showing the different parts of the pea compound leaf. **(e)** Hypocotyl length of 7-day-old seedlings grown under low light at 21 °C. Data are means ± SE (n = 20-24; 2 plates of 10-12 seedlings per plate). Grey bars: Mock (DMSO), orange bars: (–)-GR24 (1µM). Complementation assays using the *AtKAI2* promoter to express *AtKAI2* (control) or *PsKAI2* genes as noted above the graph. Proteins were tagged with 6xHA epitope or mCitrine protein. Statistical differences were determined using a one-way ANOVA with a Tukey multiple comparison of means post-hoc test, statistical differences of $P<0.05$ are represented by different letters. Means with asterisks indicate significant inhibition

compared to mock-treated seedlings with *** corresponding to $p \leq 0.001$ and * to $p \leq 0.01$, as measured by t- test.

Figure 2. Biochemical analysis of PsKAI2A and PsKAI2B interactions with different GR24 isomers. The melting temperature curves of 10 µM PsKAI2A **(a, c, e, g)** or PsKAI2B **(b, d, f, h)** with (+)-GR24 **(a-b)**, (–)-GR24 **(c-d)**, (+)-*2'-epi*-GR24 **(e-f)**, or (–)-*2'-epi*-GR24 **(g-h)** at varying concentrations are shown as assessed by DSF. Each line represents the average protein melt curve for three technical replicates; the experiment was carried out twice. **(i)** Chemical structure of ligands used in DSF assay **(a-h)**. **(j)** Plots of fluorescence intensity *versus* SL concentrations. The change in intrinsic fluorescence of AtKAI2, PsKAI2A and PsKAI2B was monitored (see **Figure S4**) and used to determine the apparent K_d values. The plots represent the mean of two replicates and the experiments were repeated at least three times. The analysis was performed with GraphPad Prism 7.05 Software.

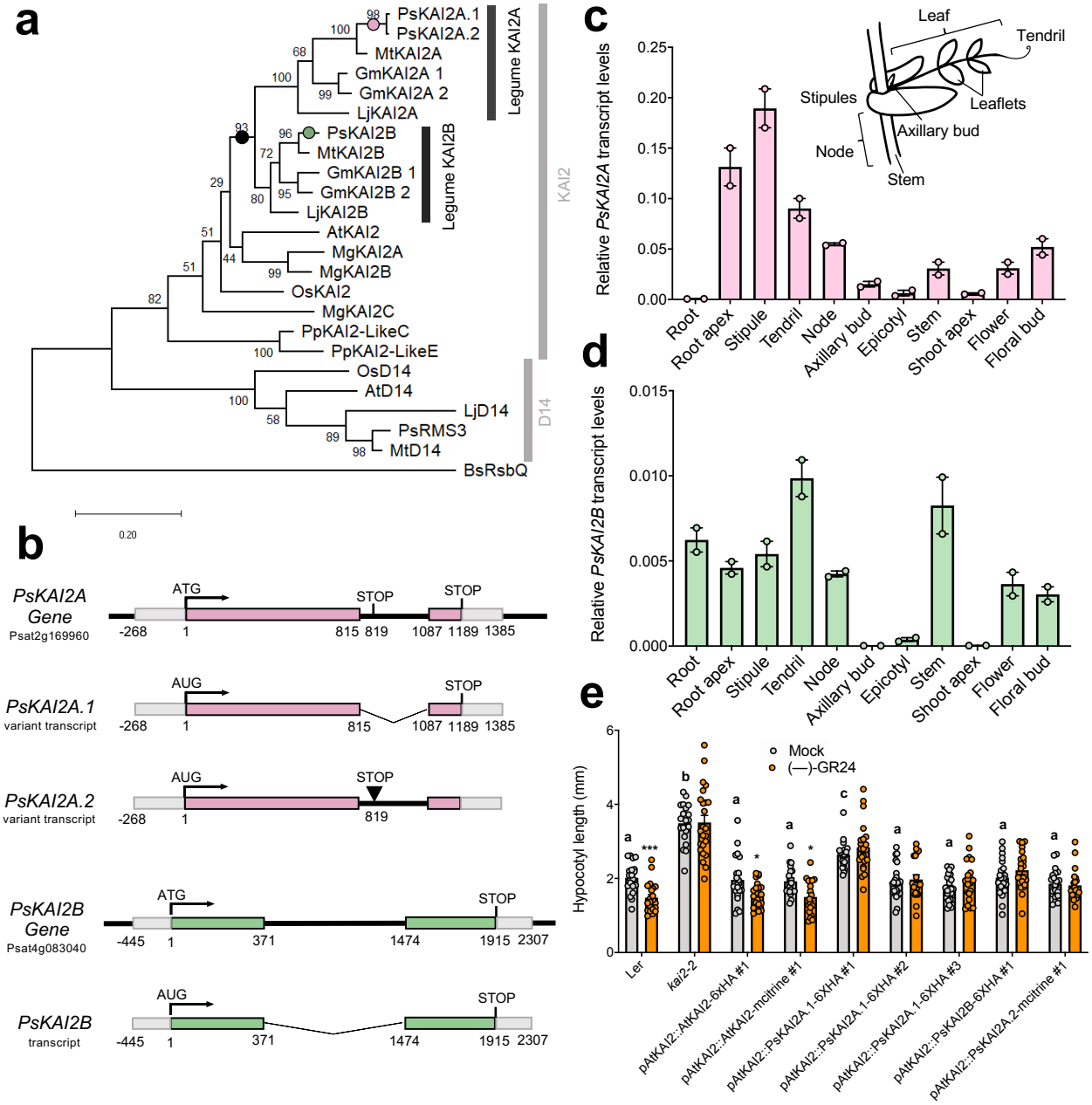
Figure 3. Comparative enzymatic activity of AtD14, AtKAI2, RMS3, PsKAI2A and PsKAI2B proteins with GR24 isomers. UPLC-UV (260 nm) analysis showing the formation of the ABC tricycle from GR24 isomers. The enzymes (10 µM) hydrolysis activity was monitored after incubation with 10 µM (+)-GR24 (yellow), (–)-GR24 (orange), (+)-*2'-epi*-GR24 (blue), or (–)-*2'-epi*-GR24 (purple). The indicated percentage corresponds to the hydrolysis rate calculated from the remaining GR24 isomer, quantified in comparison with indanol as an internal standard. Data are means ± SE (n = 3). nd = no cleavage detected.

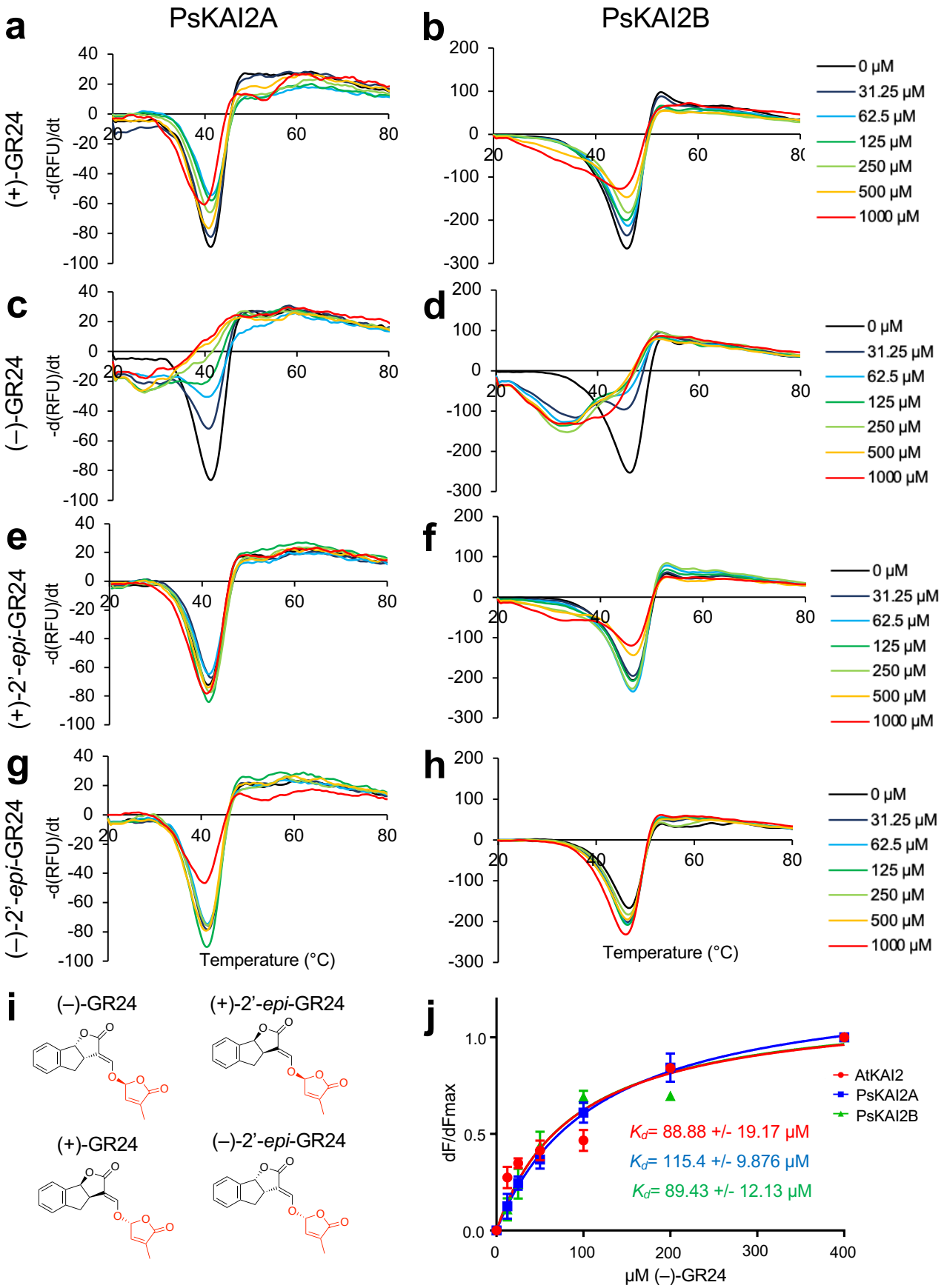
Figure 4. The crystal structure of legume KAI2. (a) Overview of PsKAI2B structure. Lid and base domains are colored in forest and light green respectively with secondary structure elements labeled. **(b)** Structural alignment of PsKAI2B and AtKAI2 (PDB ID: 4HTA) shown in light green and wheat colors respectively. Root-mean-square deviation (RMSD) value of the aligned structures is shown. The location and conservation of legume KAI2 unique residues, alanine in position 147 (A147) and asparagine N129, are highlighted on the structure shown as sticks as well as in reduced Multiple Sequence Alignment from **Figure S1**.

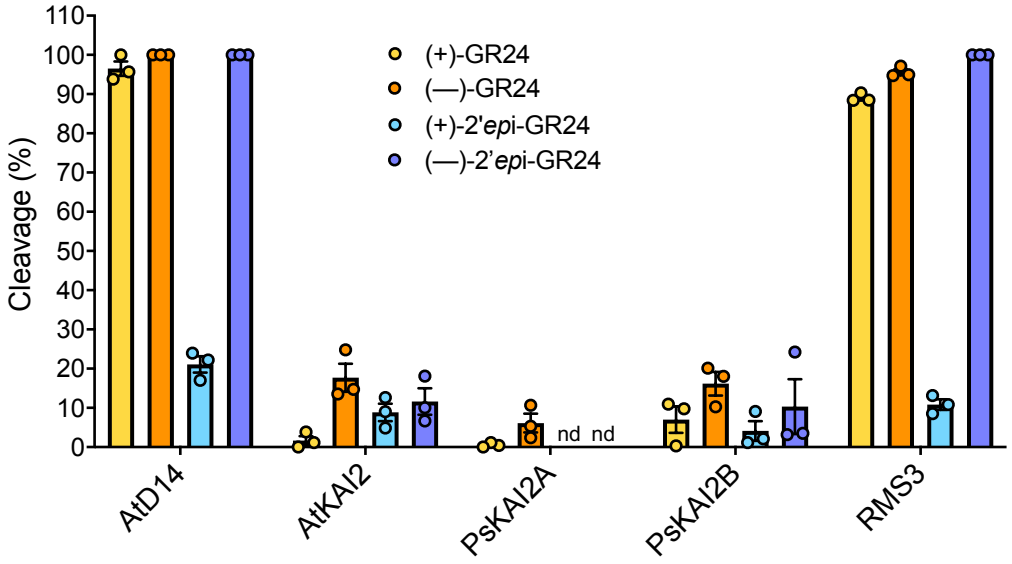
Figure 5. Structural divergence analysis of legume KAI2A and KAI2B. (a) Structural alignment of PsKAI2A and PsKAI2B shown in pink and light green colors respectively. RMSD of aligned structures is shown. **(b)** Analysis of PsKAI2A and PsKAI2B pocket volume, area, and morphology is shown by solvent accessible surface presentation. Pocket size values were calculated via the CASTp server. **(c)** Residues involved in defining ligand-binding pocket are shown on each structure as sticks. Catalytic triad is shown in red. **(d)** Residues L/M160, S/L190, and M/L218 are highlighted as divergent legume KAI2 residues, conserved among all legume KAI2A or KAI2B sequences as shown in reduced Multiple Sequence Alignment from **Figure S1**.

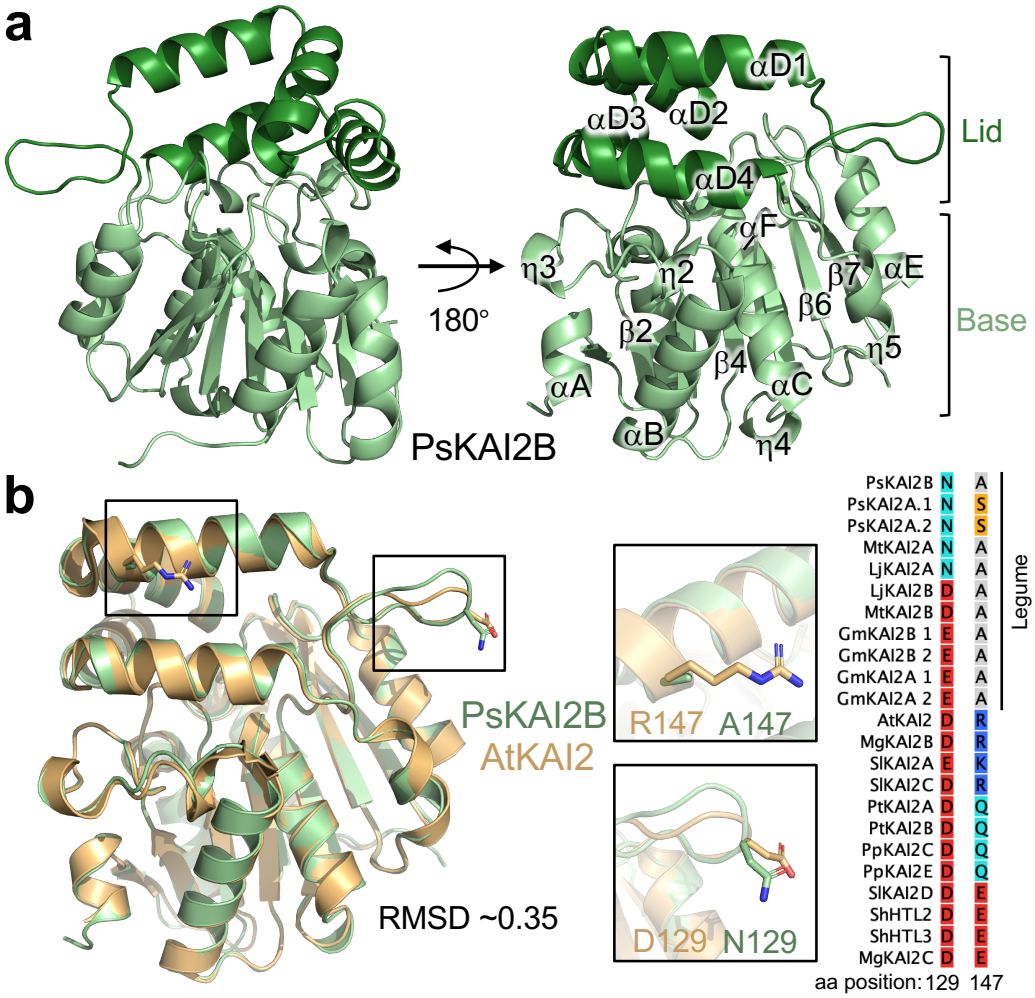
Figure 6. Structural basis of PsKAI2B ligand interaction. (a) Surface (left) and cartoon (right) representations of PsKAI2B crystal structure in complex with (−)-GR24 D-OH ring. Protein structure is shown in blue/gray and ligand in orange. **(b)** Close-up view on ligand interactions and contiguous density with the catalytic serine S95. Electron density for the ligand is shown in navy blue and blue/gray mesh for the labeled catalytic triad. The contiguous density

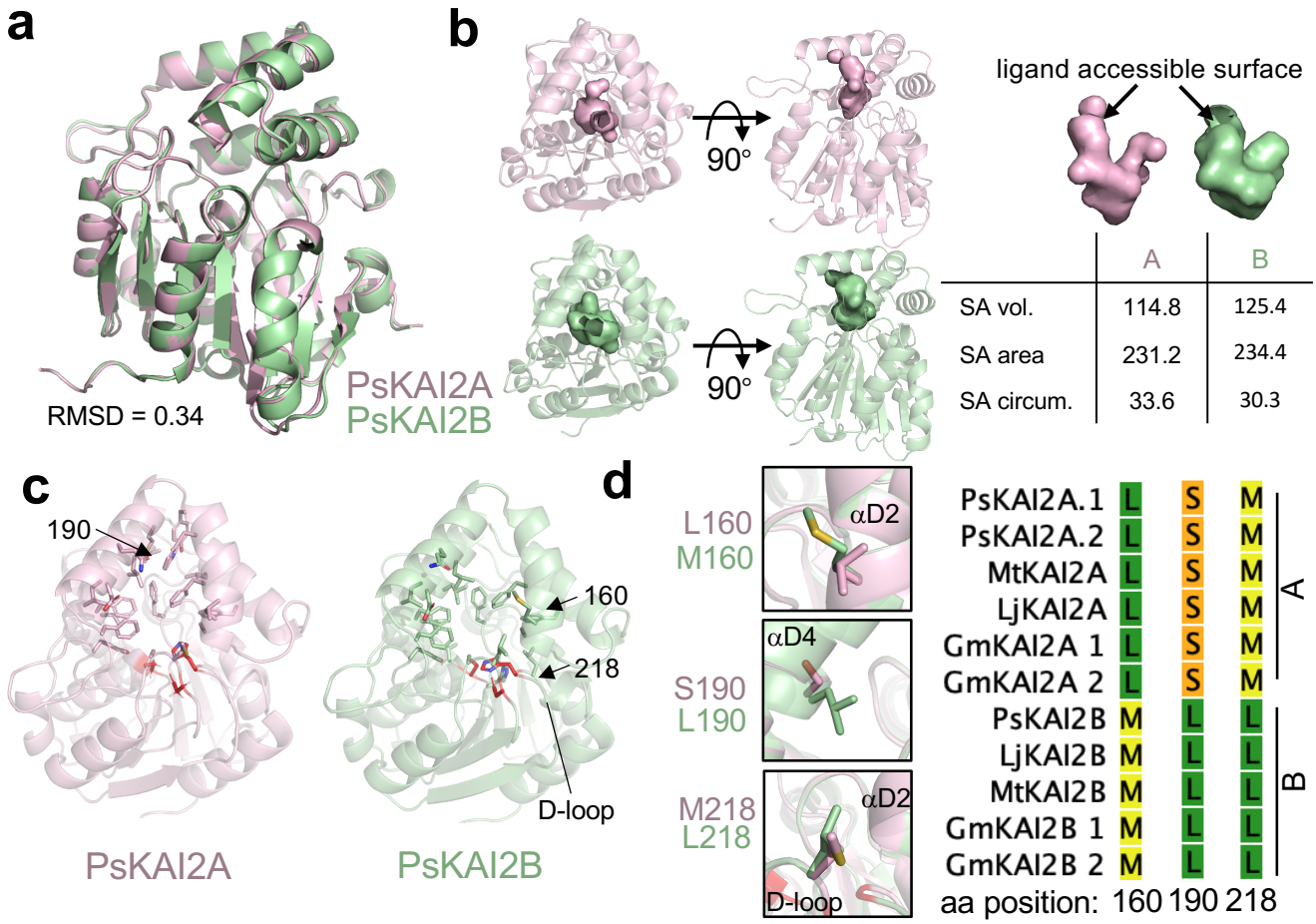
between S95 and the D-OH ring indicates a covalent bond. The electron density is derived from $2mF_oDF_c$ (2fofc) map contoured at 1.0σ. **(c)** Side view of PsKAI2B-D-OH structure shown in cartoon with highlighted (orange) the intact D-OH ring structure. 2-D ligand interaction plot was generated using LigPlot+ software. Dark grey line represents S95-D-OH ring covalent bond. **(d)** Schematic diagram of the proposed mechanism for the formation of the D-ring intermediate covalently bound to S95.


a
PsKAI2A.1
PsKAI2A.2
MtKAI2A
GmKAI2A 1
GmKAI2A 2
LjKAI2A
PsKAI2B
MtKAI2B
GmKAI2B 1
GmKAI2B 2
LjKAI2B
AtKAI2
MgKAI2A
MgKAI2B
OsKAI2
MgKAI2C
PpKAI2-LikeC
PpKAI2-LikeE
OsD14
AtD14
LjD14
PsRMS3
MtD14
BsRsbQ
Legume KAI2A
Legume KAI2B
KAI2
D14
0.20
b
PsKAI2A Gene
Psat2g169960
ATG
STOP
STOP
-268
1
815
819
1087
1189
1385
PsKAI2A.1
variant transcript
AUG
STOP
PsKAI2A.2
variant transcript
AUG
STOP
819
PsKAI2B Gene
Psat4g083040
ATG
STOP
-445
1
371
1474
1915
2307
PsKAI2B
transcript
AUG
STOP
c
Relative PsKAI2A transcript levels
Leaf
Tendril
Stipules
Leaflets
Axillary bud
Node
Stem
Root
Root apex
Stipule
Tendril
Node
Axillary bud
Epicotyl
Stem
Shoot apex
Flower
Floral bud
d
Relative PsKAI2B transcript levels
e
Hypocotyl length (mm)
Mock
(—)-GR24
Ler
kai2-2
pAtKAI2::AtKAI2-6xHA #1
pAtKAI2::AtKAI2-mcitrine #1
pAtKAI2::PsKAI2A.1-6XHA #1
pAtKAI2::PsKAI2A.1-6XHA #2
pAtKAI2::PsKAI2A.1-6XHA #3
pAtKAI2::PsKAI2B-6XHA #1
pAtKAI2::PsKAI2A.2-mcitrine #1


a
PsKAI2A
b
PsKAI2B
(+)-GR24
(–)-GR24
(+)-2'-epi-GR24
(–)-2'-epi-GR24
-d(RFU)/dt
Temperature (°C)
0 µM
31.25 µM
62.5 µM
125 µM
250 µM
500 µM
1000 µM
c
d
e
f
g
h
i
(–)-GR24
(+)-2'-epi-GR24
(+)-GR24
(–)-2'-epi-GR24
j
dF/dFmax
µM (–)-GR24
AtKAI2
PsKAI2A
PsKAI2B
Kd= 88.88 +/- 19.17 µM
Kd= 115.4 +/- 9.876 µM
Kd= 89.43 +/- 12.13 µM


Cleavage (%)
110
100
90
80
70
60
50
40
30
20
10
0
(+)-GR24
(—)-GR24
(+)-2'epi-GR24
(—)-2'epi-GR24
nd nd
AtD14
AtKAI2
PsKAI2A
PsKAI2B
RMS3


a
αD1
αD3
αD2
αD4
Lid
αF
η3
η2
β7
αE
β6
Base
β2
η5
β4
αA
αC
αB
η4
180°
PsKAI2B
b
PsKAI2B
AtKAI2
RMSD ~0.35
R147
A147
D129
N129
PsKAI2B N A
PsKAI2A.1 N S
PsKAI2A.2 N S
MtKAI2A N A
LjKAI2A N A
LjKAI2B D A
MtKAI2B D A
GmKAI2B 1 E A
GmKAI2B 2 E A
GmKAI2A 1 E A
GmKAI2A 2 E A
AtKAI2 D R
MgKAI2B D R
SlKAI2A E K
SlKAI2C D R
PtKAI2A D Q
PtKAI2B D Q
PpKAI2C D Q
PpKAI2E D Q
SlKAI2D D E
ShHTL2 D E
ShHTL3 D E
MgKAI2C D E
aa position: 129 147
Legume


a
RMSD = 0.34
PsKAI2A
PsKAI2B
b
90°
90°
ligand accessible surface
A
B
SA vol. 114.8 125.4
SA area 231.2 234.4
SA circum. 33.6 30.3
c
190
160
218
D-loop
PsKAI2A
PsKAI2B
d
L160
M160
αD2
αD4
S190
L190
M218
L218
αD2
D-loop
PsKAI2A.1 L S M
PsKAI2A.2 L S M
MtKAI2A L S M
LjKAI2A L S M
GmKAI2A 1 L S M
GmKAI2A 2 L S M
PsKAI2B M L L
LjKAI2B M L L
MtKAI2B M L L
GmKAI2B 1 M L L
GmKAI2B 2 M L L
A
B
aa position: 160 190 218

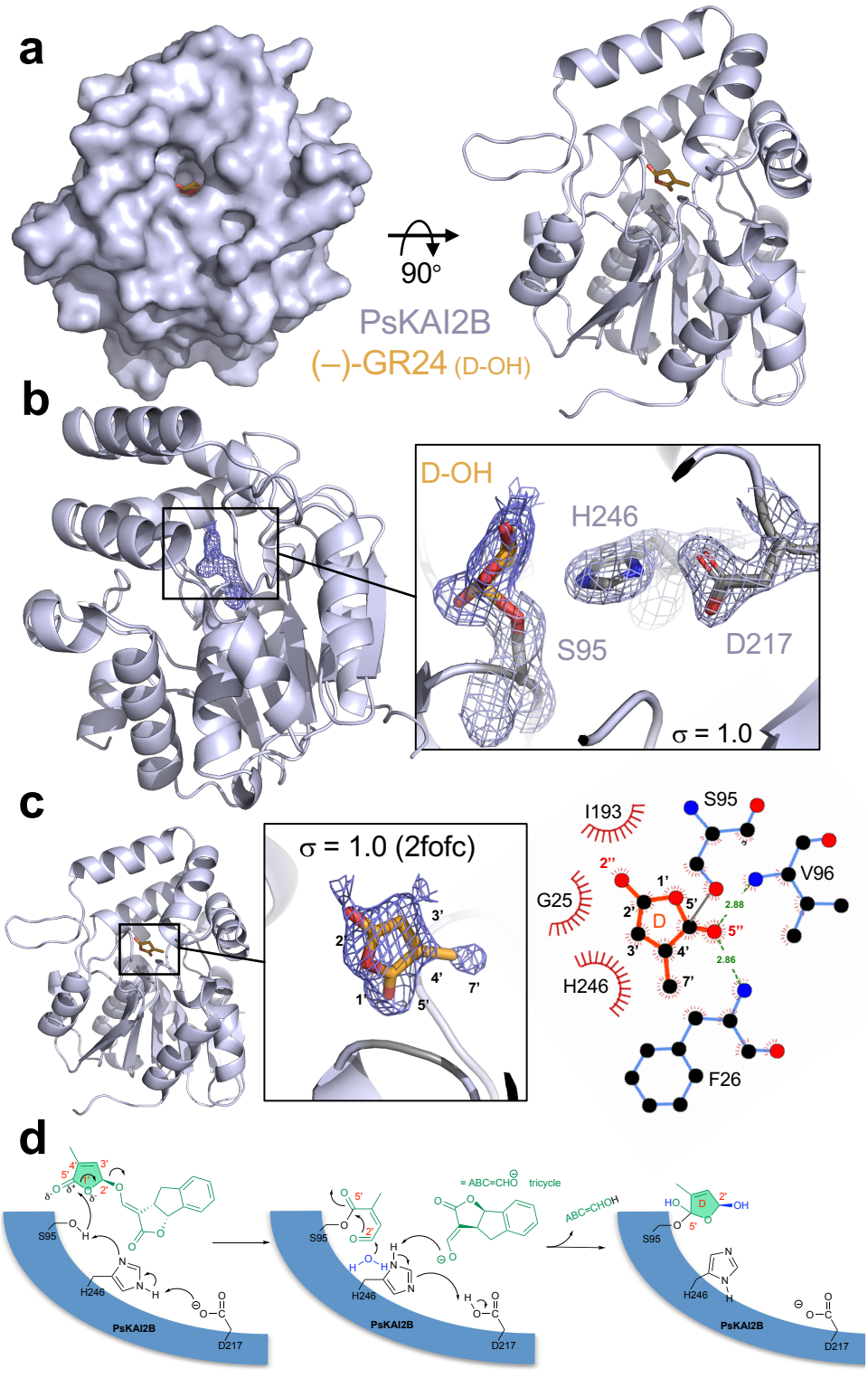

a
90°
PsKAI2B
(–)-GR24 (D-OH)
b
D-OH
H246
S95
D217
σ = 1.0
c
σ = 1.0 (2fofc)
I193
S95
V96
G25
H246
F26
2.88
2.86
d
S95
H246
PsKAI2B
D217
= ABC=CHO⊖ tricycle
ABC=CHOH

Table 1. Data collection, phasing and refinement statistics

	PsKAI2B (apo form, with glycerol)	(−)-GR24 D-OH - bound PsKAI2B
Data collection		
Space group	C2	C2
Cell dimensions		
a, *b*, *c* (Å)	87.59, 71.14, 49.06	87.08, 71.82, 48.79
α, β, γ (°)	90, 117, 90	90, 117.3, 90
Resolution (Å)	43.47-1.61 (1.66-1.61)*	43.36-2.00 (2.07-2.00)
R_{sym}	0.080 (0.589)	0.082 (0.316)
$I / \sigma I$	31.01 (1.52)	35.13 (4.11)
Completeness (%)	99.2 (84.5)	98.73 (87.53)
Redundancy	6.4 (3.2)	6.1 (4.5)
Refinement		
Resolution (Å)	1.61	2.00
No. reflections	34306	17837
R_{work} / R_{free} (%)	15.9/17.7	16.9/21.1
No. atoms	2395	2298
Protein	2110	2110
Ligand/ion	6	8
Water	279	180
B-factors		
Protein	19.92	26.5
Ligand/ion	33.73	24.60
Water	32.07	32.22
R.m.s. deviations		
Bond lengths (Å)	0.009	0.013
Bond angles (°)	0.88	1.03
Ramachandran favored (%)	98.51	98.88
Ramachandran allowed (%)	1.49	1.12
Ramachandran outliers (%)	0	0
PDB ID	7K2Z	7K38

*Statistics for the highest-resolution shell are shown in parentheses.