## **Structural and Functional Analyses Explain Pea KAI2 Receptor Diversity and Reveal Stereoselective Catalysis During Signal Perception**

Angelica M. Guercio<sup>1</sup>, Salar Torabi<sup>2</sup>, David Cornu<sup>3</sup>, Marion Dalmais<sup>4</sup>, Abdelhafid Bendahmane<sup>4</sup>, Christine Le Signor<sup>5</sup>, Jean-Paul Pillot<sup>6</sup>, Philippe Le Bris<sup>6</sup>, François-Didier Boyer<sup>7</sup>, Catherine Rameau<sup>6</sup>, Caroline Gutjahr<sup>2</sup>, Alexandre de Saint Germain<sup>6\*</sup>, and Nitzan Shabek<sup>1\*</sup>

## **Supplementary Materials**

**Supplementary Fig. 1.** Proteins Sequence analysis of KAI2/D14 family.

**Supplementary Fig. 2.** Alternative splicing of the *PsKAI2A* transcript and expression analysis of *PsKAI2C* pseudogene.

**Supplementary Fig. 3.** Multiple sequence alignment and conservation analysis of representative legume and non-legume KAI2s.

**Supplementary Fig. 4.** TILLING mutant residues position on KAI2 structures.

**Supplementary Fig. 5.** Branching and root hair phenotypes *of Pskai2* mutants.

**Supplementary Fig. 6.** Hypocotyl elongation in Ler *kai2-2* mutant and KAI2s protein expression in complementation assay.

**Supplementary Fig. 7**. Purification of PsKAI2 Proteins.

**Supplementary Fig. 8.** Biochemical analysis of the interaction between PsKAI2 proteins and the (+)-2' epi-GR24 and (-)-2'-epi-GR24 isomers ligands by DSF.

**Supplementary Fig. 9.** Intrinsic tryptophan fluorescence of PsKAI2s, AtKAI2 and RMS3 proteins in the presence of SL analogs.

**Supplementary Fig. 10.** Structural divergence analysis of legume KAI2A and KAI2B and effect on ligand docking.

**Supplementary Fig. 11.** Biochemical and structural analysis of the interaction between wildtype and residue 160 and 190 swap mutant PsKAI2 proteins and (–)-GR24 by DSF and effect on ligand docking.

**Supplementary Fig. 12.** Structural interrogation of the ligand bound PsKAI2B crystal structure.

**Supplementary Fig. 13.** Mass spectrometry characterization of covalent PsKAI2-ligand complexes.

**Supplementary Table 1.** List of the mutations identified during TILLING and mutant alleles used in the study.

**Supplementary Table 2.** Primer sequences used in study.



**Supplementary Fig. 1. Proteins Sequence analysis of KAI2/D14 family.** (**a**) Maximum likelihood phylogeny of 41 KAI2/D14-family proteins with bacterial RBSQ\* as outgroup. Node values represent percentage of trees in which the associated taxa clustered together. Vertical rectangles highlight distinct *KAI2/D14* 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**) Sequence identifiers and sources for each. \*RBSQ is a bacterial hydrolase with the most similar a/b fold.





Primer : *PsKAI2A\_*F100/*PsKAI2A\_*R1233

Primer : *PsKAI2A\_*F898/*PsKAI2A\_*R1233



**Supplementary Fig. 2. Alternative splicing of the** *PsKAI2A* **transcript and expression analysis of** *PsKAI2C* **pseudogene.** (**a**) Schematic representation of the splicing events leading to *PsKAI2A.1* and *PsKAI2A.2* forms of mature transcripts. Exons are indicated as pink boxes. (**b-c**) Electrophoresis gel of PCR products obtained after amplification of the PsKAI2A coding sequence with primers specific to *PsKAI2A.1 f*orm (**b**) and primers specific to *PsKAI2A.2 f*orm (**c**) using genomic DNA (gDNA) or first-strand cDNA (cDNA) as template; H2O: negative control; Lane L: DNA ladder. (**d**) Electrophoresis gel of PCR products obtained after amplification of the predicted *PsKAI2C* coding sequence and the reference gene *PsEF1α* using genomic DNA (gDNA) or first-strand cDNA (cDNA) from different tissues as template; H2O: negative control; Lane L: DNA ladder.











**Supplementary Fig. 3. Multiple sequence alignment and conservation analysis of representative legume and non-legume KAI2s.** Multiple sequence alignment and conservation analysis of selected KAI2s. Amino acid alignment of 7 plant KAI2 proteins. Intensity of red behind residues shows degree of divergence. Numbers on residues refer to position in PsKAI2B sequence. PsKAI2B lid (forest green) and base (light green) domains are indicated above alignment. Secondary structure of PsKAI2B sequence is shown above sequence in blue as alpha helices (αA- αF, base; and αD1- αD4, lid), beta sheets (β1- β7), 310 helices (ŋ1- ŋ5) and non-secondary structurecontaining loops. Red arrows indicate catalytic triad residues, green arrows indicate legume KAI2-unique residues as shown in **Fig. 6**, dark grey arrows indicate residues diverged between legume KAI2A and KAI2B as described in **Fig. 7** and **Supplementary Fig. 10**.



**Supplementary Fig. 4. TILLING mutant residues position on KAI2 structures.** Superposition of PsKAI2A, PsKAI2B, and example KAI2s AtKAI2 (PDB ID: 4HTA), and PpKAI2E (PDB ID: 6AZB) structures are shown in grey. The identified TILLING mutant residues from **Fig. 2a** are highlighted and shown as sticks and labelled in rainbow colors. PsKAI2B was used as the reference for amino acid sequence position.



of (**b-d**) plant height and (**c-e**) branch length at node 2 (mm) of 20-day-old (**b-c**) or 10-day-old (**d-e**) *Pisum sativum* plant. Bar, 8.5 cm. Data are means ± SE (n=11-12). Statistical differences were determined using the Kruskal-Wallis rank sum test (Asterisks indicate significant differences between WT and the mutant \*P < 0.05 , \*\*P< 0.01 and \*\*\*P < 0.001). (**f**) Root hair length phenotype and quantification (**g**) of 10 day old WT and *Pskai2a-6 Pskai2b-3* seedlings. Bars, 2 mm. Data are means of measurements (n=8) of individual root (n=8-9), Statistical analysis Welch t-test, pvalue  $< 0.05$ 



**Supplementary Fig. 6. Hypocotyl elongation in Ler** *kai2-2* **mutant and KAI2s protein expression in complementation assay. (a)** 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). Light blue bars: Mock (DMSO), lavender bars: (-)-GR24 (1µM). Complementation assays using the *AtKAI2* promoter to express *AtKAI2* (control) or *PsKAI2* genes in the null *kai2-2* mutant background (Ler ecotype) as noted above the graph. Proteins were tagged with 6xHA epitope or mCitrine protein. For DMSO controls, 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  $p \le 0.001$  and  $*$  to  $p \le 0.01$ , as measured by t- test. Picture are representative of the mock traiteent. Bar, 2 mm **(b-c)** AtKAI2 and PsKAI2 level analyzed by immunoblot using α-HA antibody **(b)** or α-GFP antibody **(c)** in wild type, *kai2-2* and transformed *kai2-2* plants, expressed under the control of *AtKAI2* promoter. Protein extracts from 10-d-old seedlings were separated by 10% SDS-PAGE. Ponceau staining is included for loading reference.



**Supplementary Fig. 7. Purification of PsKAI2 Proteins.** Size exclusion peaks eluates for PsKAI2A (**a**) and PsKAI2B (**b**). Proteins were resolved on SDS-PAGE gels and visualized via Coomassie stain. Molecular weight (MW) markers are labeled to show the relative size of proteins.



**Supplementary Fig. 8. Biochemical analysis of the interaction between PsKAI2 proteins and the (+)-2'** *epi***-GR24 and (-)-2'-***epi***-GR24 isomers ligands by DSF.** (**a**) Chemical structure of ligands used in DSF assay. The melting temperature curves of 10 µM PsKAI2A (**b**, **d**) or PsKAI2B (**c**, **e**) with (+)-*2'-epi*-GR24 (b-c), or (–)-*2' epi*-GR24 (**d-e**) at varying concentrations are shown as assessed by DSF. Each line represents the average protein melt curve for three technical replicates and the experiment was carried out twice.



**Supplementary Fig. 9. Intrinsic tryptophan fluorescence of PsKAI2s, AtKAI2 and RMS3 proteins in the presence of SL analogs.** Intrinsic tryptophan fluorescence of PsKAI2A (**a-d**), PsKAI2B (**e-h**), AtKAI2 **(i-l)** and RMS3 (**m-p**) proteins in the presence of SL analogs. Changes in intrinsic fluorescence emission spectra in the presence of various concentrations of (+)-GR24 (**a;e;i;m;q**), (–)-GR24 (**b;f;j;n;r**), (+)-2'-*epi-*GR24 (**c;g;k;o;s**), (–)- 2'-*epi*-GR24 (**d;h;l;p;t**). Proteins (10 μM) were incubated with increasing amounts of ligand (0–400 μM, top line to bottom line, respectively). The observed relative changes in intrinsic fluorescence were plotted as a function of SL analog concentration and transformed to degree of saturation and used to determine the apparent *K*D values relevant to **Fig. 4b**. The plots represent the mean of two replicates and the experiments were repeated at least three times. The analysis was performed with GraphPad Prism 8.0 Software.



**Supplementary Fig. 10. Structural divergence analysis of legume KAI2A and KAI2B and effect on ligand docking. (a)** Structural alignment of PsKAI2A and PsKAI2B shown in pink and light green respectively. Calculated RMSD of aligned structures is shown. Residues differentiating all legume KAI2A from KAI2B are shown on each structure as sticks and labeled with residue number. Catalytic triad is shown in red. **(b)** Residues 102, 140, 160, 190, 191, 214, 218, and 225, L190, and L218 are highlighted as divergent legume KAI2 residues, conserved among all legume KAI2A or KAI2B sequences (with the exception of D140 in GmKAI2A 2) as shown in reduced Multiple Sequence Alignment from **Supplementary Fig. 1**. **(c)** *In silico* docking analysis of intact (–)-GR24 with solventaccessible pocket shown for each structure and corresponding docking scores reported in kcal/mol.



**Supplementary Fig. 11. Biochemical and structural analysis of the interaction between wildtype and residue 160 and 190 swap mutant PsKAI2 proteins and (–)-GR24 by DSF and effect on ligand docking.** The melting temperature curves of 10 µM PsKAI2A (a), PsKAI2A L160M, S190L swap (b), PsKAI2B (c), and PsKAI2B M160L, L190S swap **(d)**, with (+)-GR24 or (–)-GR24 at the effective concentration of 62.5 µM are shown as assessed by DSF**.** Each line represents the average protein melt curve for three technical replicates and the experiment was carried out twice. **(e)** pocket surface representation of PsKAI2A model, PsKAI2B apo structure, and modelled PsKAI2A L160M, S190L and PsKAI2B M160L, L190S swap mutants. **(f)** Analysis of solvent accessible pocket volumes and areas calculated via the CASTp server. **(g)** In silico docking analysis of intact (–)-GR24 with solvent-accessible pocket shown for each structure and corresponding docking scores reported in kcal/mol.



**Supplementary Fig. 12. Structural interrogation of the ligand bound PsKAI2B crystal structure. (a)** Structural alignment of PsKAI2B apo (light green) and PsKAI2B in complex with (–)-GR24 D- OH (gray/blue). Calculated RMSD value is shown. Similar orientation of the superposition is shown in surface (right) and cartoon (left) representations. **(b)** Chemical structure of intact (–)- GR24 molecule with numbered carbons. Electron density mesh fit with D-OH ring of (–)-GR24. Protein structure is shown in blue/gray and ligand in orange. The electron density of the ligand is derived from 2mFoDFc (2fofc) map contoured at 1.0s. **(c)** LigandFit examination of crystallization and purification conditions reported here. Glycerol denoted: GOL, (+/-)-2-Methyl- 2,4 pentanediol denoted: MPD. Polyethylene glycol, PEG, denoted: PE4. Ligand access pocket is shown together with the ligand placed via LigandFit fitting software (top), and the obtained electron density maps following 3 cycles of Phenix refine. The electron densities shown are derived from 2mFoDFc map (2fofc, blue mesh) contoured at 1.5s and mFoDFc map (fofc, green mesh) contoured at 3s. Correlation Coefficient (CC) scores were calculated via LigandFit. **(d)** *In silico* analysis of D-OH ligand docking is shown in orange (found in the structure) and predicted orientations in navy and magenta with corresponding docking scores reported in kcal/mol.



**Supplementary Fig. 13**. **Mass spectrometry characterization of covalent PsKAI2-ligand complexes.** A Deconvoluted electrospray mass spectra of PsKAI2A and PsKAI2B before (**a-b**) and after (**c-d**). adding of ligand (–)- GR24 are shown respectively on upper and lower panels. Peaks with an asterisk correspond to PsKAI2 covalently bound to a (–)-GR24 ligand (PsKAI2-ligand). A mass increment of 96 Da is measured for two PsKAI2-ligand complexes. (**e-h**) Ligand-modified H246 amino-acids were identified by nano LC-MSMS analyses after chymotrypsin proteolysis. Fragmentation spectra of unmodified and ligand-modified peptides are shown. Labeled peaks correspond to b and y fragments of the triple charged precursor ion. Histidine 246 residue modified by ligand is marked with GR on the sequence displayed on the top.

**Supplementary Table 1. List of the mutations identified during TILLING and mutant alleles used in the study.**





