Plant Communications, Volume 3

Supplemental information

Structural analysis of receptor-like kinase SOBIR1 reveals mechanisms

that regulate its phosphorylation-dependent activation

Xue Wei, Yulu Wang, Su Zhang, Tianyi Gu, Gabryel Steinmetz, Haiyan Yu, Guoguang Guo, Xin Liu, Shilong Fan, Fengzhong Wang, Yangnan Gu, and Fengjiao Xin



Figure S1. Overall structure of SOBIR1-KD^{D489A}. Related to Figure 2 and Table 1

(A) Schematic representation of two molecules within one asymmetric unit of SOBIR1-KD^{D489A}. Molecule A (right) uses the same color scheme as that in the schematic diagram in Figure 1A, whereas molecule B (left) is presented in gray. The bound AMP-PNPs are presented as black sticks, whereas the Mg²⁺ ions are indicated by red spheres. The disordered β 3- α C loops are presented as red (molecule A) and gray (molecule B) dashed lines.

(B) Molecules A and B superimposed in one asymmetric unit.

(**C**) Comparison between molecule A (right) and molecule B (left), which are colored according to the temperature factors for $C\alpha$ atoms.

				~ "	Glycine-rich	ı loop
						
14000701			++	+-		250
ALSOBIRI ALSOBIRI	AIGGTISGFVFSVLFKLIIQAIRGSEKPP	GPSIFSP		LIKKAEDLA	FLENEEALASLEIIGRGGC	358
SISOBIR1	TFAGGISAVLCSLLFKMVMFFVRRGNNDS	SLTIYSP		LIKKAEDLA	FLEKEDGVASLEMIGKGGC	342
NtSOBIR1	SFAGAISAVLLSVLFKLVMFFVRKGKTDG	TLTIYSP		LIKKAEDLA	FLEKEDGVASLEMIGKGGC	353
GmSOBIR1 RcSOBIR1	ALGGTLSGFVFSLMFKLALALIKGRGRKA AVAGSMSGFVFSVMFKLILATIRGGGKDS	GPDIISP		LIKKAEDLA	FLEKEEGIASLEIIGRGGC	353
AtBAK1	GVAAGAALLFAVPAIALAWWRRKKPQDH	FFDVPAE	EDP	EVHLGQLKRFSLRE	LQVASDNFSNKNILGRGGF	300
AtBRI1	SVAMGLLFSFVCIFGLILVGREMRKRRRKKEA	ELEMYAEGHGNSGD	RTANNTNWKLTGVKEAL	SINLAAFEKPLRKLTFAD	LLQATNGFHNDSLIGSGGF	894
MmPKA Hatraka	AKKGSEQES-VKEFLAKAKEDF	LKKWETP		SQN	-TAQLDQFDRIKTLGTGSF	197
HSIRAR4 HsSrc	TSRTQFNSLQQLVAYYSKHADGLCH	RLTTVCPTS		KPQTQGLAKDA	WEIPRESLRLEVKLGQGCF	281
HsCDK2					MENFQKVEKIGEGTY	15
	β	3-αÇ loop				
	β2 β3		αC	<u> </u>	αD	
	+		+		+	
AtSOBIR1	GEVFKAELPGSNGKIIAVKKVIQPPKDADEI	DEDSKFLNKKMRQI	RSEINTVGHIRHRNLLP	LLAHVSRPECHYLVYEYM	EKGSLQDILTDVQAGNQEL	458
ALSOBIR1	GEVFKAELPGSNGKIIAVKKVIQPPKDADELI	DEDSKFMNKKMRQI	RSEISTVGQIRHRNLLP	LLAHVSRPECHYLVYEYM	EKGSLQDILTDVQAGNQEL	458
SISOBIRI NtSOBIRI	GEVYRAELPGSNGKIIAIKKIIQSPMDAAEIT GEVYRAELPGSNGKIIAIKKIIOPPMDAAEIT	EEDTKALNKKMRQV.	KSEIQIVGQIRHRNLLP KSEIOILGOIRHRNLLP	LLAHMPRPDCHYLVYEYM	KNGSLQDILQQVTEGTREL KNGSLODILOOVTEGTREL	442
GmSOBIR1	GEVYKAELPGSNGKMIAIKKIVQPPKDGAELA	EEDSKVLNKKMRQI	RSEINTVGQIRHRNLLP	LLAHVSRPECHYLVYEFM	KNGSLQDTLSKVERGESEL	453
RcSOBIR1	GEVYKAELPGSNGKMIAIKKIVQPPRDAAEL	EEDTKLLNKKMRQI	RSEINTVGQIRHRNLLP	LLAHVSRPDCHCLVYEFM	KNGSLQDILSDVEQGKREL	450
AtBAK1	GKVYKGRLADGTLVAVKRLKEEFT	QGGELQF	QTEVEMI SMAVHRNLLR	LRGFCMTPTERLLVYPYM	ANGSVASCLRERPESQPPL	385
MmPKA	GRVMLVKHK-ESGNHYAMKILDKOKVV	GQGDIEHT	MAEMETIGKIKHRNLVP LNEKRILOAVNFPFLVK	LEFSFKDNSNLYMVMEYV	AGGEMFSHLRRIGRFSEPH	143
HsIRAK4	GVVYKGYVNNTTVAVKKLAAMVDITTE	ELKQQF	DQEIKVMAKCQHENLVE	LLGFSSDGDDLCLVYVYM	PNGSLLDRLS-CLDGTPPL	283
HsSrc	GEVWMGTWNGTTRVAIKTLKPGTMSP	EAF	LQEAQVMKKLRHEKLVQ	LYAVVSE-EPIYIVTEYM	SKGSLLDFLKGETGKYL	361
HSCDK2	GVVYKARNK-LTGEVVALKKIRLDTET	EGVPSTA	IREISLLKELNHPNIVK	LLDVIHTENKLYLVFEFL	HQD-LKKFMDASALTGIPL	101
	Catalyt	IC IOOP	Activation	S2 AS3 dEF	' αF	
						0
	++++		++	+ +	+	
AtSOBIR1	MWPARHKIALGIAAGLEYLHMDHNPRIIHRDL	KPANVLLDDDMEAR	ISDFGLAKAMPDAVTHI	TISHVAGTVGYIAPEFYQ	THKF-TDKCDIYSFGVILG	557
SISOBIRI	DWLGRHRIAAGVAAGLEYLHINHTQRIIHRDL	KPANILLDDDMEAR	VADFGLAKAVPDAHTHI	TISNVAGIVGFIAPEYYQ	ILKF-TDKCDIYSFGVVLA	541
NtSOBIR1	DWLGRHRIAVGIAAGLEYLHINHSQCIIHRDL	KPANVLLDDDMEAR	IADFGLAKALPDAHTHI	TTSNVAGTVGYIAPEYHQ	TLKF-TDKCDIYSFGVVLA	552
GmSOBIR1	DWLSRHKISLGVAAGLEYLHMNHNPRIIHRDL	KPANILLDDDMEAR	IADFGLAKAMPDYKTHI	TTSNVAGTVGYIAPEYHQ	ILKF-TDKCDIYSYGVILG	552
AtBAK1	DWEARHRIAIGVASGLEYLHTSHSPRIIHRDI	KAANILLDEEFEAV	VGDFGLAKLMDYKDTHV	TTSNVAGTVGYIAPEYHQ TT-AVRGTIGHIAPEYLS	FLKF-TDKCDIYSFGVLLG FGKS-SEKTDVFGYGVMLL	483
AtBRI1	NWSTRRKIAIGSARGLAFLHHNCSPHIIHRDM	KSSNVLLDENLEAR	VSDFGMARLMSAMDTHL	SVSTLAGTPGYVPPEYYQ:	SFRC-STKGDVYSYGVVLL	1077
MmPKA	ARFYAAQIVLTFEYLHSLDLIYRDI	KPENLLIDQQGYIQ	VTDFGFAKRVKGRT <mark>W</mark> TL	CGTPEYLAPEIIL	SKGY-NKAVDWWALGVLIY	230
HsIRAK4	SWHMRCKIAQGAANGINFLHENHHIHRDI	KSANILLDEAFTAK	ISDFGLARASEKFAQTV	MISRIVGITAYMAPEALR	G-EI-TPKSDIYSFGVVLL	378
HSCDK2	PLIKSYLFQLLQGLAFCHSHRVLHRDL	KPQNLLINTEGAIK	LADFGLARAFGVPVRTY	THEVVTLWYRAPEILL	GCKYYSTAVDIWSLGCIFA	194
	αG		αH		αΙ	_
	+++++		++		+	
AtSOBIR1	ILVIGKLPSD-EFFQHTDEMSLIKWMRNIITS	ENPSLAIDPKLMDQ	-GFDEQMLLVLKIACYC	TLDDPKQRPNSKDVRTML	sQIKH	641
AISOBIR1	ILVIGKLPSD-EFFQHTDEMSLIKWMRNIITS	ENPSLAIDPKLMEQ	-GFDEQMLLVLKIACYC	TLDDPKQRPNSKDVRTML	SQIKH	641
NtSOBIRI	VLVIGKLPSD-EFFOHTPEMSLVKWLRNVMTS	EDPKRAIDPKLIGN	-GIDEQMLLVLKIACFC	TLENPKERPNSKDVRCML	IQIKH IOIKH	625
GmSOBIR1	VLVIGKLPSD-DFFQHTEEMSLVKWMRKTLSS	ENPKEAINSKLLGN	-GYEEQMLLVLKIACFC	TMDDPKERPNSKDVRCML	sQIKH	636
RcSOBIR1	VLVMGKLPSD-EFFQTTREMSLVKWMRNIMTS	ENPSQAIDPKMNGN	-GFEEQMLLVLKIACFC	TLDDPRQRPNSKDIRCML	sQIKH	633
AtBAK1	ELITGORAFDLARLANDDDVMLLDWVKGLLKE	KKLEALVDVDLQGN	-YKDEEVEQLIQVALLC	TQSSPMERPKMSEVVRML	EGDGLAERWEEWQKEEMFR	582
MmPKA	EMAAGYPPFFADOPIOIYEKIVSGKVRF	PSHFSSDLKDLLRN	LLQVDLTKRFGNLKNGV	NDIKNHKWFATTDWIAIY	QRKVEAPFI	316
HsIRAK4	EIITGLPAVDEHREPQLLLDIKEEIEDE	EKTIEDYIDKKMND	-ADSTSVEAMYSVASQC	LHEKKNKRPDIKKVQQLL	QEMTAS	460
HsSrc	ELTTKGRVPYPGMVNREVLDQVERGYRMPCPF	ECPESLHDLMCQCW	RKEPEERPTFEYLQAFL	EDYFTSTEPQYQPGENL		536
HSCDK2	EMVTRRALFPGDSEIDQLFRIFRTLGTF	DEVVWPGVTSMPDY	KPSFPKWARQDFS	KVVPPLDEDGRSLLSQML	HYDPNKRISAKAALAHPFF	286
	++++		+	++	+	
AtSOBIR1					641	Q9SKB2
ALSOBIR1					641	D7LEA5
SISOBIR1 NtSORTR1					625	K4C8Q3 081.P72
GmSOBIR1					636	I1JXE0
RcSOBIR1		_			633	B9RAQ8
AtBAK1	QDFNYPTHHPAVSGWIIGDSTSQIENEYPSGP	R T.			615	Q94F62
MmPKA	PKFKGPGDTSNFDDYEEEEIRVSINEKCGKEF	TEF			351	P05132
HsIRAK4					460	Q9NWZ3
HsSrc					536	P12931
HSCDK2	QDVTKPVPHLRL				298	P24941

Figure S2. Amino acid sequence alignment of the kinase domains of SOBIR1 orthologs and selected plant and mammalian kinases. Related to Figure 1, Figure 3, and Figure 4

The Swiss-Prot ID is provided after each protein name. The secondary structure elements of SOBIR1-

KD are indicated above the sequences and colored as in the schematic diagram in Figure 1A. Phosphorylation sites identified in *Arabidopsis thaliana* SOBIR1-KD are highlighted in red. Conserved phosphorylation sites in the kinases are indicated in a red box. The glycine-rich loop, β 3- α C loop, catalytic loop, and activation segment are labeled accordingly. *At*, *Arabidopsis thaliana*; *Al*, *Arabidopsis lyrata*; *Sl*, *Solanum lycopersicum*; *Nt*, *Nicotiana tabacum*; *Gm*, *Glycine max*; *Rc*, *Ricinus communis*; *Mm*, *Mus musculus*; *Hs*, *Homo sapiens*.



Figure S3. Interaction between Ala489 and Asn494 that stabilizes the Src-like inactive conformation of SOBIR1-KD^{D489A}. Related to Figure 2

The residues stabilizing the Src-like inactive conformation are highlighted as lemon sticks. Salt bridges formed by Ala489 and Asn494 (**B**) are presented as black dashed sticks. The Lys512 (**A**) and Lys377 (**C**) side chains cannot be seen due to low resolution, which were supposed to form salt bridges with Glu407 and Asp507, respectively. AMP-PNP is presented as black sticks, and the Mg^{2+} ion is indicated by a red sphere.



Figure S4. Conformations of the nucleotide-binding sites of IRAK4 (PDB code: 6EGF), CDK4 (PDB code: 1HCK), and BRI1 (PDB code: 5LPV). Related to Figure 2

Detailed binding modes of AMP-PNP in the nucleotide-binding pockets of IRAK4 (**A**), CDK4 (**B**), and BRI1 (**C**). The nucleotide-interacting residues of IRAK4, CDK4, and BRI1 are presented as slate, green cyan, and salmon sticks and labeled accordingly. Hydrogen bonds related to the interaction with AMP-PNP and Mg^{2+}/Mn^{2+} are presented as blue or black dashed lines. Mn^{2+} and Mg^{2+} are indicated by purple and red spheres, respectively.



Figure S5. LC-MS/MS spectra revealing the *in vitro* phosphorylation sites in the SOBIR1 activation segment. Related to Figure 3 and Table 2

Mass spectra for the SOBIR1-KD activation segment in different states. The wild-type protein (**A**, **B**) was recombinantly expressed in *E. coli*. The PP2C α -dephosphorylated proteins (**C**) were generated by incubating 10 μ M SOBIR1-KD and 1 μ M GST-PP2C α in the kinase reaction buffer at 25 °C for 1 h. Proteins were purified with GST columns, followed by gel filtration chromatography (Superdex 200

HR 10/30 column). Autophosphorylated SOBIR1-KD (**D**) was generated by incubating 10 μ M dSOBIR1-KD with 1 mM ATP and 10 mM Mg²⁺ at 25 °C for 3 h. BAK1-phosphorylated SOBIR1-KD (**E**) was generated by incubating 10 μ M dSOBIR1-KD and 1 μ M GST-BAK1-CD in the kinase reaction buffer at 25 °C for 1 h. Each LC-MS/MS spectrum presents the ions produced by the collision-induced dissociation of the intact peptides. Predominant b and y product ion peaks are labeled and product ions eliciting neutral mass losses of H₃PO₄ (98 Da) are also indicated. The * and ++ indicate the ions that lost ammonia (-17 Da) and the doubly protonated ions, respectively. The phosphorylated residues in each peptide are highlighted in red.



Figure S6. Overall structure of SOBIR1-KD^{∆389-401}. Related to Figure 4 and Table 1

(A) Determination of the oligomerization states based on size exclusion chromatography. Wild-type SOBIR1-KD and SOBIR1-KD^{GS} mutant eluted as multimers, whereas the deletion of the β 3- α C loop (SOBIR1-KD^{Δ 389-401}) apparently resulted in a monomeric kinase domain. Elution volumes of the protein standards are indicated.

(B) Phosphorylation states of the wild-type SOBIR1-KD as well as the SOBIR1-KD^{Δ 389-401} and SOBIR1-KD^{GS} mutants.

(C) Schematic representation of two molecules within one asymmetric unit of SOBIR1-KD^{Δ 389-401}. N-lobe, C-lobe, α C, and activation segment of molecule A are colored in orange, cyan, blue, and violet, respectively, whereas molecule B (left) is presented in gray. The bound AMP-PNPs are presented as black sticks, and Mg²⁺ ions are indicated by red spheres.

(D) Molecules A and B superimposed in one asymmetric unit.

(E) Comparison between the SOBIR1-KD^{D489A} and SOBIR1-KD^{Δ 389-401} structures. SOBIR1-KD^{D489A} is presented in gray, whereas SOBIR1-KD^{Δ 389-401} is presented in orange (N-lobe) and cyan (C-lobe). The bound AMP-PNP and Mg²⁺ ion of SOBIR1-KD^{Δ 389-401} are presented as a black stick and a red sphere, respectively.





The wild-type (**A**), PP2C α -dephosphorylated (**B**), autophosphorylated (**C**, **D**), and BAK1phosphorylated (**E**) SOBIR1-KDs were generated as described in the Supplemental Figure 5 legend. Each LC-MS/MS spectrum presents the collection of ions produced by the collision-induced dissociation of the intact peptides. Predominant b and y product ion peaks are labeled and product ions eliciting neutral mass losses of H₃PO₄ (98 Da) are also indicated. The * and ++ indicate the ions that lost ammonia (-17 Da) and the doubly protonated ions, respectively. The phosphorylated residues in each peptide are highlighted in red.



Figure S8. LC-MS/MS spectra revealing the *in vitro* phosphorylation sites of SOBIR1phosphorylated BAK1. Related to Figure 5 and Table 2

The reaction was completed by incubating 10 μ M GST-BAK1-CD^{D434N} and 1 μ M SOBIR1-KD at 25 °C for 1 h. Each LC-MS/MS spectrum presents the collection of ions produced by the collision-induced dissociation of the intact peptides. Predominant b and y product ion peaks are labeled and product ions eliciting neutral mass losses of H₃PO₄ (98 Da) are also indicated. The * and ++ indicate ions that lost ammonia (-17 Da) and doubly protonated ions, respectively. Phosphorylated residues in each peptide are highlighted in red.

N. benthamiana



Figure S9. Functional analysis of SOBIR1-KD phosphorylation sites or segment *in planta*. Related to Figure 6 and Figure 7

*At*SOBIR1 and various mutants fused to C-terminally enhanced green fluorescent protein were transiently expressed in *N. benthamiana* by *Agrobacterium*-mediated transient expression (agroinfiltrations). Photographs were taken at 4 days postinfiltration. All assays were performed three times and a representative photograph is shown. Scale bar = 0.5 cm.

	SOBIR1-KDD489A	SOBIR1-KD ^{∆389-401}				
Data collection						
Space group	P 2 ₁	<i>P</i> 2 ₁				
Cell dimensions						
a, b, c (Å)	67.6, 50.9, 105.4	72.4, 50.7, 93.1				
α, β, γ (°)	90°, 95.3°, 90°	90°, 106.6°, 90°				
Popolution(A)	50.00-2.90	50.00-2.90				
Resolution (A)	(3.00 - 2.90) ^a	(3.00-2.90)				
No. of measured reflections	101722	71544				
No. of unique reflections	16161	14273				
R _{sym} or R _{merge}	0.080(0.430)	0.072 (0.411)				
l / σl	19.9 (2.5)	14.6 (2.2)				
Completeness (%)	99.8 (99.1)	98.1 (95.5)				
Redundancy	6.3 (5.2)	5.0 (4.4)				
Refinement statistics						
Resolution (Å)	43.39-2.89	28.63-2.91				
No. reflections	16148	14267				
R _{work} / R _{free} (%)	21.7/23.9	22.0/24.7				
No. atoms						
Protein	4443	4535				
Ligand/ion	64	62				
Water	0	0				
<i>B</i> -factors (Å ²)						
Protein	94.5	58.9				
Ligand/ion	82.1	52.3				
Water	0	0				
R.m.s. deviations						
Bond lengths (Å)	0.010	0.010				
Bond angles (°)	0.74	0.71				
Ramachandran plot statistics						
Most favored (%)	97.59	97.77				
Allowed (%)	2.41	2.23				
Disallowed (%)	0	0				

Supplemental Table 1. Data collection and refinement statistics

^aAll data sets were collected from a single crystal.

^bValues in parentheses are for highest-resolution shell.

Proteins	Identified Sites				
	Thr390, Ser394, Ser406, Thr519, Thr522, Thr523,				
SOBIRT-RD	Ser524, Thr529, Ser566, Thr573, Thr587, Ser592				
PP2Cα-dephosphorylated	no phosphorylation sites				
SOBIR1					
autophosphorylated					
dSOBIR1-KD	Inr390, 3er394, 3er400, Inr529				
BAK1-phosphorylated	Thr390, Ser394, Ser406, Thr410, Thr519, Thr522,				
dSOBIR1-KD	Thr523, Ser524, Thr529, Ser592				
SOBIR1-phosphorylated	Thr324, Thr446, Thr449, Thr450, Thr455, Ser595,				
BAK1-CDD434N	Ser602, Thr603, Ser604				

Supplemental Table 2. Phosphorylated sites identified by LC-MS/MS

Proteins are listed in the first column and the position of the phosphorylated residues corresponding to the proteins are shown in the second column. Phosphorylated sites of proteins identified in the β 3- α C loops or the activation segments are highlighted in black bold.

Mutants	Primers	Sequence (5' to 3')		
Full length	F1	TGCTCTAG AATGGCTGTTCCCACGGGAAGCG		
	R1	CGGGGTAC CCTAGTGCTTGATCTGGGACAACATG		
Kinase domain	F2	CGCCATATG TCCATATTCAGCCCATTGATC		
Rinase domain	R2	CCGCTCGAG CTAGTGCTTGATCTGGGACAAC		
Τ510Δ	F3	CAATGCCAGATGCAGTC <u>GCG</u> CATATTACAACCTCGCATG		
10104	R3	CATGCGAGGTTGTAATATG <u>CGC</u> GACTGCATCTGGCATTG		
Τ522Δ	F4	GATGCAGTCACACATATT <u>GCG</u> ACCTCGCATGTTGCAGGTAC		
IUZZA	R4	GTACCTGCAACATGCGAGGT <u>CGC</u> AATATGTGTGACTGCATC		
Τ523Δ	F5	GCAGTCACACATATTACA <u>GCG</u> TCGCATGTTGCAGGTACTG		
10207	R5	CAGTACCTGCAACATGCGA <u>CGC</u> TGTAATATGTGTGACTGC		
S524A	F6	GTCACACATATTACAACC <u>GCG</u> CATGTTGCAGGTACTG		
0024/1	R6	CAGTACCTGCAACATGCGCGGTTGTAATATGTGTGAC		
Τ529Δ	F7	GCATGTTGCAGGT <u>GCG</u> GTGGGATACATAG		
10207	R7	CTATGTATCCCAC <u>CGC</u> ACCTGCAACATGC		
T529E	F8	GCATGTTGCAGGT <u>GAA</u> GTGGGATACATAG		
1029	R8	CTATGTATCCCAC <u>TTC</u> ACCTGCAACATGC		
A 200 401	F9	GATCCAACCGCCTAAAGACGCCGATGAAAGGCAAATTAGATCCGAGATCA		
△369-401	R9	CTGTGTTGATCTCGGATCTAATTTGCCTTTCATCGGCGTCTTTAGGCGGTT		
00	F10	GCCTAAAGACGCCGATGAA <u>GGUAGCGGUAGCGGUAGCGGUAGCGGUAG</u> <u>CGGU</u> AGGCAAATTAGATCCGAG		
GS	R10	CTCGGATCTAATTTGCCT <u>ACCGCTACCGCTACCGCTACCGCTACCGCTACC</u> TTCATCGGCGTCTTTAGGC		
Τ30 0Δ	F11	GACGCCGATGAACTA <u>GCG</u> GATGAAGATTCCAAG		
13904	R11	CTTGGAATCTTCATC <u>CGC</u> TAGTTCATCGGCGTC		
S30/A	F12	GAACTAACAGATGAAGAT <u>GCG</u> AAGTTTCTGAACAAG		
00047	R12	CTTGTTCAGAAACTT <u>CGC</u> ATCTTCATCTGTTAGTTC		
T300V+6304V	F13	GCCGATGAACTA <u>GCG</u> GATGAAGAT <u>GCG</u> AAGTTTCTG		
1 JJUA 1	R13			

Supplemental Table 3. Primers used in this study

Note: The underline represents the mutagenic position.