the plant journal

SEB SOCIETY FOR EXPERIMENTAL BIOLOGY

The Plant Journal (2021) **105,** 1710–1726 doi: 10.1111/tpj.15133

TECHNICAL ADVANCE

Computational prediction method to decipher receptor–glycoligand interactions in plant immunity

Irene del Hierro[1,2], Hugo Mélida[1,†], Caroline Broyart[3], Julia Santiago[3] and Antonio Molina[1,2,*]

[1]*Centro de Biotecnología y Genómica de Plantas (CBGP), Universidad Politécnica de Madrid (UPM), Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Campus de Montegancedo-UPM, 28223, Pozuelo de Alarcón, Madrid, Spain,*

[2]*Departamento de Biotecnología-Biología Vegetal, Escuela Técnica Superior de Ingeniería Agronómica, Alimentaria y de Biosistemas, Universidad Politécnica de Madrid (UPM), 28040, Madrid, Spain and*

[3]*Département de Biologie Moléculaire Végétale (DBMV), University of Lausanne (UNIL), Biophore Building, UNIL Sorge, CH-1015, Lausanne, Switzerland*

Received 3 July 2020; revised 30 November 2020; accepted 8 December 2020; published online 14 December 2020.
*For correspondence (e-mail antonio.molina@upm.es).
[†]Present address: Área de Fisiología Vegetal, Departamento de Ingeniería y Ciencias Agrarias, Universidad de León, 24071, León, Spain

SUMMARY

Microbial and plant cell walls have been selected by the plant immune system as a source of microbe- and plant damage-associated molecular patterns (MAMPs/DAMPs) that are perceived by extracellular ectodomains (ECDs) of plant pattern recognition receptors (PRRs) triggering immune responses. From the vast number of ligands that PRRs can bind, those composed of carbohydrate moieties are poorly studied, and only a handful of PRR/glycan pairs have been determined. Here we present a computational screening method, based on the first step of molecular dynamics simulation, that is able to predict putative ECD-PRR/glycan interactions. This method has been developed and optimized with Arabidopsis LysM-PRR members CERK1 and LYK4, which are involved in the perception of fungal MAMPs, chitohexaose (1,4-β-D-$(GlcNAc)_6$) and laminarihexaose (1,3-β-D-$(Glc)_6$). Our *in silico* results predicted CERK1 interactions with 1,4-β-D-$(GlcNAc)_6$ whilst discarding its direct binding by LYK4. In contrast, no direct interaction between CERK1/laminarihexaose was predicted by the model despite CERK1 being required for laminarihexaose immune activation, suggesting that CERK1 may act as a co-receptor for its recognition. These *in silico* results were validated by isothermal titration calorimetry binding assays between these MAMPs and recombinant ECDs-LysM-PRRs. The robustness of the developed computational screening method was further validated by predicting that CERK1 does not bind the DAMP 1,4-β-D-$(Glc)_6$ (cellohexaose), and then probing that immune responses triggered by this DAMP were not impaired in the Arabidopsis *cerk1* mutant. The computational predictive glycan/PRR binding method developed here might accelerate the discovery of protein–glycan interactions and provide information on immune responses activated by glycoligands.

Keywords: *Arabidopsis thaliana*, glycan, immunity, isothermal titration calorimetry, LysM domain, molecular dynamics, pattern recognition receptor, technical advance.

INTRODUCTION

Plant terrestrial colonization and diversification was associated to the evolution of a set of plant protein receptors, called pattern recognition receptors (PRRs), that confer to plants the capacities to perceive environmental and developmental cues. These PRR-based surveillance systems have allowed plants to modulate their adaptive and physiological responses to environmental conditions and to activate defense responses against the diversity of pathogens that can colonize them (Li *et al.*, 2016; Tang *et al.*, 2017). Plant PRRs comprise several groups of extracellular, membrane-anchored proteins that greatly exceed in number their PRR counterparts in animals (Zipfel, 2014). In addition, the plant immune system comprises a set of intracellular protein receptors, mainly resistance (R) proteins, encoded by *R* genes, that perceive pathogen

effectors and activate effector-triggered immunity (ETI), which dampens pathogen colonization (Baltrus *et al.*, 2011; Li *et al.*, 2016; Sonah *et al.*, 2016). Extracellular and membrane-anchored PRRs include three main classes of proteins: (i) receptor-like kinases (RLKs), which contain an extracellular ectodomain (ECD), a transmembrane region (TM) and an intracellular serine/threonine kinase domain (KD); (ii) receptor-like proteins (RLPs), with ECD and TM but lacking the KD; and (iii) receptor proteins (RPs), which either contain an ECD that can be attached to the plasma membrane by a glycosylphosphatidylinositol (GPI)-anchor (RPg) or are extracellular proteins (RPes) not attached to the membrane (Bellande *et al.*, 2017). In *Arabidopsis thaliana*, different subclasses of extracellular RLKs/RLPs/RPs can be considered based on their predicted ECD structures or sequence similarities: lectins (G, L and C-lectins), leucine-rich repeats (LRRs), CRinkly-Like (CR4L), pathogenic-related thaumatin-like (ThaumatinL/PR5), proline-rich extensin-like receptor kinase (PERK), wall-associated kinases (WAKs), malectins, *Catharanthus roseus* receptor-like kinase 1-like (*Cr*RLK1L), lysin motif (LysM) and cysteine-rich kinases (CRK/DUF26, also known as stress-antifungal/salt-response receptors). These RLKs/RLPs/RPs comprise more than 600 members in Arabidopsis, representing approximately 2–3% of Arabidopsis genes (Shiu and Bleecker, 2003; Gish and Clark, 2011; Franck *et al.*, 2018).

RLKs/RLPs/RPs are involved in the perception of developmental cues (e.g., peptide ligands or hormones) (Santiago *et al.*, 2013; Santiago *et al.*, 2016; Stegmann *et al.*, 2017; Tang *et al.*, 2017) and in high-affinity recognition of different conserved structures from microbes called microbe-associated molecular patterns (MAMPs) (Boutrot and Zipfel, 2017). Additionally, these PRRs can recognize other types of patterns, like plant self-molecules that are released or synthesized upon pathogen infection or tissue damage, which are known as damage-associated molecular patterns (DAMPs) (Li *et al.*, 2020). Upon MAMP/DAMP recognition by specific PRRs, formation of protein complexes with other PRRs (co-receptors) takes place and pattern-triggered immunity (PTI) responses are activated (Greeff *et al.*, 2012; Dangl *et al.*, 2013).

Plant RLKs are very similar to animal receptor tyrosine kinases (RTKs), but RLKs have generally serine/threonine kinase specificity instead of the tyrosine specificity of animal RTKs (Shiu *et al.*, 2004; Greeff *et al.*, 2012). The RLK class has an ancient origin after the fungus–metazoan split and is not present in the fungal kingdom. Interestingly, domain fusion resulted in the creation of novel receptors, leading to a high diversity and the appearance of different protein subclasses, and gene duplications contributed to gene expansion, explaining the high number of RLKs in plants (Lehti-Shiu *et al.*, 2009; Li *et al.*, 2016). Most of these RLKs are found in large genomic clusters, supporting the hypothesis of rapid evolution by duplication and gene shuffling under diversifying selection (Shiu and Bleecker, 2003; Fritz-Laylin *et al.*, 2005; Lehti-Shiu *et al.*, 2009; Gish and Clark, 2011; Li *et al.*, 2016). RLPs and RPs have been shown to be involved in the control of defensive and developmental processes, playing roles in RLK signaling as RLK counterparts in co-receptor mechanisms (Jeong *et al.*, 1999; Fritz-Laylin *et al.*, 2005; Gish and Clark, 2011; Bellande *et al.*, 2017; Tang *et al.*, 2017).

There is an extensive diversity of molecules that can be bound by PRRs to trigger plant immune responses, but most of the MAMPs and DAMPs described so far are peptides (Boutrot and Zipfel, 2017; Li *et al.*, 2020). In contrast, the number of glycan structures characterized as MAMPs/DAMPs is very low and accordingly the number of PRR/glycan pairs identified is very restricted (Bacete *et al.*, 2018; Li *et al.*, 2020). However, PRR/glycan interaction is a field in expansion as glycans are cell surface components of major plant pathogens like fungi, oomycete and bacteria (MAMPs) and they are also present in the plant cell walls and can be released as oligosaccharides (DAMPs) (Bacete *et al.*, 2018; Wanke *et al.*, 2020a). On the other hand, one of the reasons explaining the slow progress of this field is the diversity (thus complexity) of glycan ligands in terms of composition: (i) over 20 different monosaccharides can form the backbone and/or ramification building blocks of glycans through a high diversity of glycosidic linkages; (ii) glycans can differ in the degree of polymerization (DP); and (iii) monosaccharides can have different biochemical decorations (e.g., acetylation and methylation) and chemical modifications (e.g., reduction/oxidation) (Carpita and McCann, 2000; Latgé and Calderone, 2006; Mélida *et al.*, 2013; Srivastava *et al.*, 2017).

The interaction of PRRs with carbohydrate-based ligands is well studied in mammals, where several receptor/glycan complexes have been determined such as Dectin-1 (C-lectin ECD)/β-glucans or Galectin-3/β-galactosides (Brown *et al.*, 2003; Chen *et al.*, 2017; Díaz-Alvarez and Ortega, 2017). In contrast, the characterization of PRR–glycoligand interactions in plants has been mainly restricted to PRRs of the LysM subclass that harbor ECDs with lysin motifs, which are promiscuous motifs involved in the recognition of several ligands like chitin, peptidoglycans, β-1,3-glucans and lipopolysaccharides (Miya *et al.*, 2007; Willmann *et al.*, 2011; Desaki *et al.*, 2018; Mélida *et al.*, 2018). In the plasma membrane, CERK1 (LYK1) and LYK5 together with LYK4 are key components in plant immunity acting as co-receptors in the recognition of chitin, a polymer of 1,4-β-D-GlcNAc (Liu *et al.*, 2012; Cao *et al.*, 2014). However, in the plasmodesmata region, LYM2 forms a complex with LYK4 upon chitin perception by a CERK1-independent mechanism (Faulkner *et al.*, 2013; Cheval *et al.*, 2020). In rice (*Oryza sativa*), the molecular mechanism of chitin recognition by the OsCEBiP receptor, a LysM-PRR, has also been

described (Liu *et al.*, 2016), and it has been shown that OsCERK1 cooperates with OsCEBiP module chitin-mediated signaling (Shimizu *et al.*, 2010). Interestingly, LysM domains have also been described in some fungal effector proteins, like *Cf*Avr4F from *Cladosporium fulvum* and Mg1LysM from *Zymoseptoria tritici*, and their crystal structures have been obtained (Hurlburt *et al.*, 2018; Sánchez-Vallet *et al.*, 2020). These fungal proteins bind chitin as a mechanism of fungal virulence aiming to avoid fungal perception by plant LysM-PRRs.

Recent work has demonstrated that LysM-PRRs are involved in the perception by plants of 1,3-β-glucans isolated from fungal cell walls (Mélida *et al.*, 2018; Wanke *et al.*, 2020b). Specifically, the hexasaccharide 1,3-β-D-$(Glc)_6$ (laminarihexaose) is an immune-active structure whose recognition in Arabidopsis is CERK1-dependent. Moreover, molecular docking calculations suggested some interactions between 1,3-β-D-$(Glc)_6$ and CERK1-ECD, though these interactions were not further validated by full molecular dynamics simulations or *in vitro* binding assays (Mélida *et al.*, 2018). CERK1 was also shown to be involved in the perception of bacterial peptidoglycans, with a contribution of LysM members LYM1 and LYM3 (Gust *et al.*, 2007; Willmann *et al.*, 2011). The bases of these additional roles of LysM-PRRs in the perception of these glycans structures are unknown since crystal structures of these ligand/LysM complexes have not been obtained.

Plant cell walls are mainly composed of cellulose, different types of hemicelluloses and pectins (Carpita and McCann, 2000). Upon pathogen infection, some plant cell wall-derived compounds can be released acting as DAMPs recognized by PRRs, activating DAMP-triggered immunity cascades. For example, cellulose-derived oligomers (1,4-β-glucans) have been described as a novel group of plant DAMPs which trigger signaling cascades that share many similarities with the responses activated by the well-characterized plant DAMPs oligogalacturonides (OGs), derived from homogalacturonan pectins (Aziz *et al.*, 2007; de Azevedo Souza *et al.*, 2017; Johnson *et al.*, 2018; Locci *et al.*, 2019). No PRR receptor candidates have been proposed for 1,4-β-glucans oligomers, whereas several PRRs (WAKs, THESEUS and FERONIA) have been suggested to bind pectins, though crystal structures of these ligand/PRR complexes have not been obtained (Hématy *et al.*, 2007; Boisson-Dernier *et al.*, 2011; Kohorn and Kohorn, 2012; Duan *et al.*, 2020). There are many open questions regarding glycan-triggered immunity that could be solved using biochemical and genetic approaches. However, these approaches are time consuming and face the redundant functions of PRR families, as illustrated by the lack of PTI-defective phenotypes of *lyk4* and *lyk5* Arabidopsis single mutants upon chitin and 1,3-β-D-$(Glc)_6$ treatment (Cao *et al.*, 2014; Mélida *et al.*, 2018).

In silico prediction tools (computational modeling) could help in preliminary screening stages to select potential PRRs for further characterization (Das *et al.*, 2018; Fratev *et al.*, 2018). However, structural conformations of glycans are not easy to determine due to their intrinsic mobility and the scarce capability of some structural techniques, like X-ray diffraction, to solve their conformations (Fadda and Woods, 2010; Gimeno *et al.*, 2020). Moreover, protein–glycan interactions are weak, with affinities ranging from the μM to the mM range, due to the formation of transient structural states that result in more dynamic interactions than protein–peptide ones (Otto *et al.*, 2011; Sapay *et al.*, 2013; Isaacson and Díaz-Moreno, 2019; Mende *et al.*, 2019; Gimeno *et al.*, 2020; Haab and Klamer, 2020). *In silico* approaches, particularly molecular dynamics, can aid to solve complex protein–glycan interactions, but simulation of carbohydrate-based structures can also be challenging since the initial protein–glycan conformation and the force field employed during the simulation might be critical steps to obtain reliable results. Moreover, molecular docking parameters in the simulation programs are not specifically designed for carbohydrates, but this limitation can be partially addressed by combining docking and molecular dynamics methods of structural analysis that proved to give excellent results when testing protein–ligand interactions (Woods and Tessier, 2010; Das *et al.*, 2018; Fratev *et al.*, 2018; Kumar *et al.*, 2019).

In this work we introduce an *in silico* pipeline designed to predict PRR–glycan interactions avoiding most of the heavy computing requirements and parameter configuration challenges of molecular dynamics. This method might help to screen the high number of putative interactions between plant PRRs and glycoligands, including MAMPs/DAMPs already identified and those to be discovered. Data obtained with this method would pave the way to design biochemical and genetic approaches to confirm the predicted interactions. Using this pipeline, we have validated the previously described structural interaction of the LysM-PRR member CERK1 with chitohexaose (1,4-β-D-$(GlcNAc)_6$), and we have discarded their direct interaction with laminarihexaose (1,3-β-D-$(Glc)_6$) and cellohexaose (1,4-β-D-$(Glc)_6$), two glycoligands recently described to trigger immunity. These model predictions were confirmed by *in vitro* binding assays with purified ECDs of CERK1 and LYK4 PRR receptors, further validating the feasibility of the modeling method presented here to decipher PRR–glycoligand interactions.

RESULTS

PRRs with glycan-binding ECDs are highly represented in Arabidopsis

Previous studies have predicted different numbers of extracellular PRRs in the Arabidopsis genome (more than 600 members) and considered some of the following subclasses: lectins (C-, L- and G-lectins), LRRs, LysMs,

malectins, *Cr*RLK1Ls, WAKs, CR4Ls, ThaumatinL/PR5, CRKs (also classified as stress-antifungal/salt-response receptors) and PERKs (Gish and Clark, 2011; Bellande *et al.*, 2017; Tang *et al.*, 2017). With this in mind, we decided to perform an additional comprehensive analysis of these proteins in order to establish an updated classification of Arabidopsis extracellular PRRs, comprising RLKs, RLPs and RPs (both RPgs and RPes; Figure 1). The main aim was the identification of PRRs containing putative glycan-binding ECDs among these types of receptor proteins. Classification into RLKs, RLPs, RPgs or RPes was performed depending on the presence of transmembrane regions, serine/threonine kinase domains or GPI-anchor motifs. This characterization allowed the identification of 617 putative PRRs with extracellular ECDs (Table S1), of which 329 (53.3%) have ECDs that putatively could interact with carbohydrates, indicating the high number of plant PRRs that might potentially bind glycoligands (Figure 1a and Table S2). The remaining 288 PRRs (46.7%) were described as putatively not related with carbohydrate binding based on their ECD topology, and comprise PRRs with different ECDs: LRRs (246 members), PERKs (23 members), CR4L (8 members) and non-structural classified ECDs (11 members) (Table S1).

The potential Arabidopsis glycan-binding ECDs were classified based on their similarity with protein domains that have been reported to interact with carbohydrates. As shown in Figure 1(c), the putative glycan-binding PRR members were divided into 10 subclasses based on their different ECDs: C-lectin (1 member), G-lectins (51 members), L-lectins (55 members), WAKs (38 members), LysM (10 members), *Cr*RLK1L (with malectin motif; 20 members), malectins (external malectin motif followed by an LRR; 51 members), LRR-malectins (external LRR domain followed by a malectin motif; 12 members), ThaumatinL/PR5 (23 members) and CRKs (68 members). All these subclasses include RLKs, RLPs and RPs, with the exception of C-lectin and LRR-malectins (which are all RLKs), WAKs (which are all RLKs except for one RPe) and CrRLK1Ls (which are all RLKs except for two RLPs) (Table S2). The LysM-PRR subclass accounts for 10 members in Arabidopsis, with 5 of them considered as RLKs (CERK1/LYK1, LYK2, LYK3, LYK4 and LYK5), 2 as RPes (AT5G62150 and AT4G25433) and 3 as RPgs (LYM1, LYM2 and LYM3). Though LYM1 and LYM2 predictions suggest TM-anchor structures (Table S2), they have been previously classified as GPI-anchored proteins (Faulkner *et al.*, 2013; Bellande *et al.*, 2017).

Computational prediction method for the identification of PRR/glycan complexes

In order to perform *in silico* predictions of PRR/glycan interactions, we developed and optimized a method that consisted on several steps (see Figure 2 and Experimental Procedures). First, the structure of the target glycan (ligand) was generated, either by retrieving it from crystallography or NMR experiments through the Protein Data Bank (PDB; https://www.rcsb.org/) or by generating it from scratch adding carbon rings and additional atoms and bonds since just a few crystallography or NMR data are available for glycans (Figure 2). The structures of 1,4-β-D-$(GlcNac)_6$ and 1,3-β-D-$(Glc)_6$ were obtained from the PDB database, whereas 1,4-β-D-$(Glc)_6$ was both retrieved from crystallized data and built from scratch. In order to provide

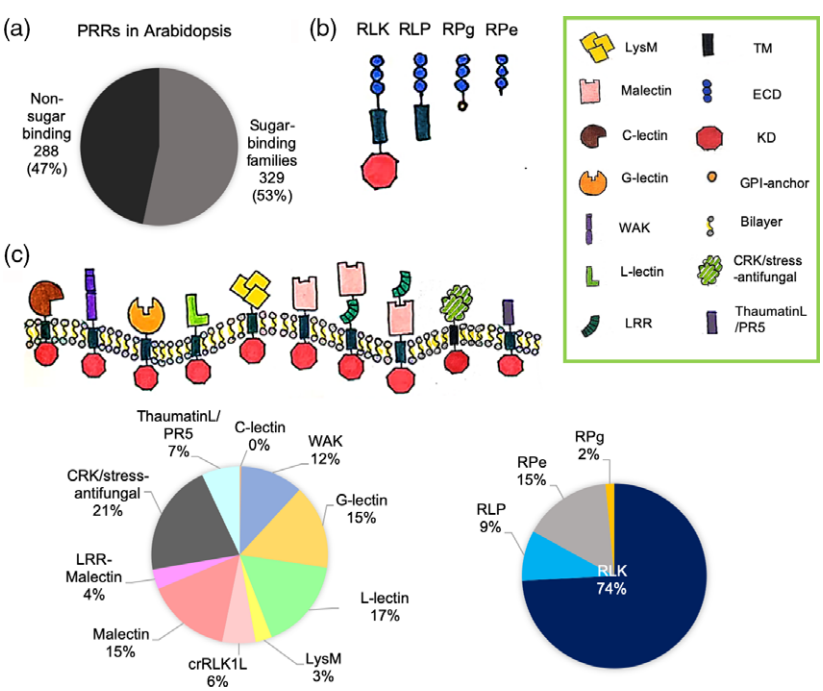

Figure 1. Classification of *Arabidopsis thaliana* PRRs according to the glycan-binding domain of their ECDs. (a) Proportion of Arabidopsis PRRs putatively harboring glycan-binding domains in their ECDs (53.3% of the total 617 PRRs identified). (b) Scheme of the main domain structure of PRRs: receptor-like kinases (RLKs) that contain an extracellular ectodomain (ECD), a transmembrane region (TM) and an intracellular kinase domain (KD); receptor-like proteins (RLPs) with ECD and TM; and receptor proteins (RPs), which either contain an ECD that can be attached to the plasma membrane by a GPI-anchor (RPg) or are extracellular proteins (RPe) not attached to the membrane. (c) Graphical representation of putative glycan-binding PRR-RLK subclasses. Graph depicting the percentages and numbers of each subclass (left) and the proportions of RLKs, RLPs, RPgs and RPes of the putative glycan-binding PRRs (right) are shown.

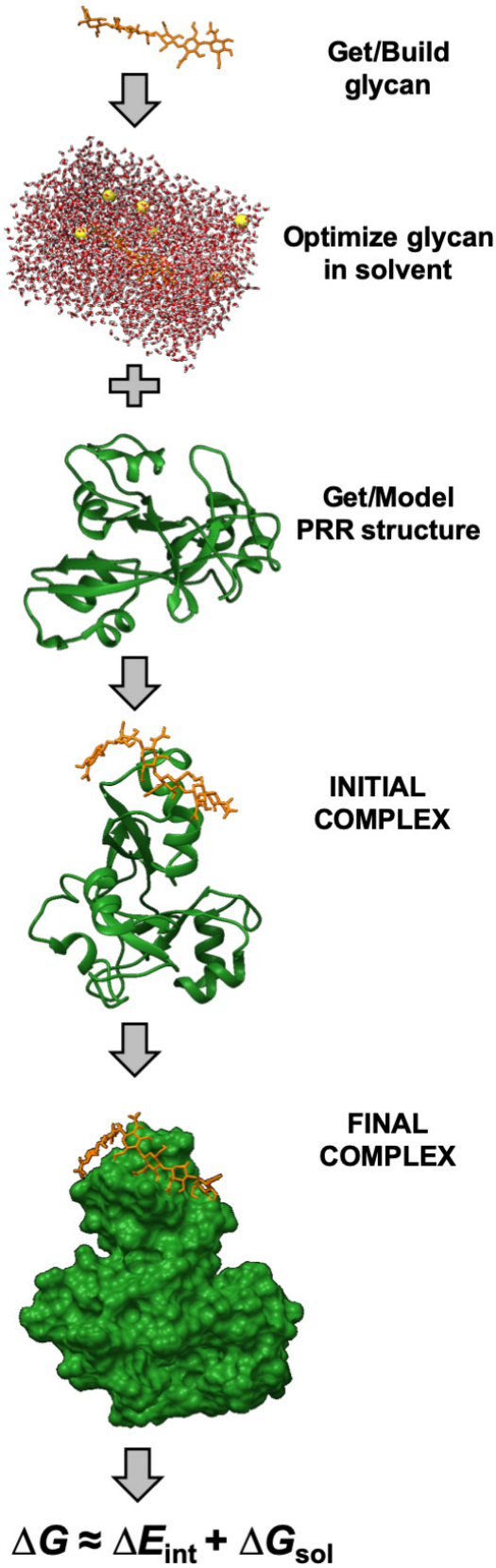

Figure 2. Protocol description of the computational method developed for glycan–PRR binding prediction. Glycan and PRR structures are retrieved through the PDB database, or alternatively glycoligands are built from scratch (with the Chimera *Build Structure* tool) and/or PRR-ECDs are modeled (see Experimental Procedures). Ligands are optimized in vacuum and in solvent using a 12-Å padding box (TIP3 water model, 0.15 M NaCl). Then five minimum different conformations (replicates) of the PRR/glycan complexes are calculated employing AutoDock Vina using a 27-Å padding spatial box. The final complex is then obtained for each replicate applying molecular dynamics with two minimization procedures (SolBox1 and SolBox2) and one full simulation procedure (Full) using the CHARMM force field and TIP3 water with 0.15 M NaCl and a 12-Å padding box. The ΔG energies are obtained from each of these final complexes.

the 'most biologically realistic' glycan conformation, vacuum optimization of glycan structures was performed with Chimera followed by solvent minimization with Visual Molecular Dynamics (VMD)-AutoIMD. In parallel, the PRR-ECD structure of interest was obtained from X-ray or NMR data (e.g., 4EBZ for CERK1-ECD), or it was modeled *in silico* generally through homology modeling (Swiss-Model server; https://swissmodel.expasy.org/). For example, we modeled LYK4-ECD using the rice chitin elicitor-binding protein (OsCEBiP, 5JCE) structure (Liu *et al.*, 2016), while Arabidopsis ANXUR1 (6FIG) (Moussu *et al.*, 2018) was used as a template for modeling THESEUS1 (THE1), which was included as a negative control in our analyses. Next, the initial ECD-PRR/glycan complex was determined by docking the glycan in the putative binding site(s) of the PRR (those described or those that we obtained through sequence homology if they had not been previously identified). We considered the described putative binding sites for chitin in CERK1 (amino acids 109–115 and 137–143, encompassing Q109, E110, E114 and I141 directly involved in chitin binding) and LYK4 (amino acids 125–131 and 154–160) (Iizasa *et al.*, 2010; Wan *et al.*, 2012). THE1 binding sites were obtained by sequence alignment with *Xenopus laevis* malectin complexed with nigerose (2K46), as shown in previous studies (Moussu *et al.*, 2018). Docking was achieved with AutoDock Vina using a 27-Å padding box, retrieving a total of five docking complexes as possible different conformations of the glycan in the PRR-binding site (five initial complexes or replicates). Glycosylation positions and disulfide bonds in CERK1 were obtained from manual assertion inferred from a combination of experimental and computational evidence as shown in UniProt (https://www.uniprot.org/; Liu *et al.*, 2012). LYK4 and THE1 glycosylation patterns and disulfide bonds were identified by sequence analysis prediction and sequence similarity, respectively (UniProt). The identification of glycosylation sites in ECDs is crucial, since they can provoke false positive binding results and mislead predictions if they are close to the glycoligand.

Once the initial complexes were established they were parameterized in the CHARMM force field using CHARMM-GUI and solved using VMD software in order to prepare

them for molecular dynamics with scalable molecular dynamics (NAMD), following three standardized steps: minimization, equilibration and simulation (http://www.iitg.ac.in/tamalb/karp/namd/; https://www.ch.embnet.org/MD_tutorial/). The protocol to complete these three steps was optimized with the CHARMM force field at 298 K, pH 7 for 10 ns with canonical NVT ensemble (at temperature and volume constant allowing for pressure variation; see Table S3(c–e) for an example of configuration files). As full simulations with this standard parameterization tended to crash over time or to produce misleading binding results due to their high computing requirements, we decided to carry out only the minimization steps to test LysM-PRR/glycan interactions. These minimization procedures were SolBox1, with boundary conditions applied to NVT ensemble (constant volume and temperature) (Table S3a), and SolBox2, with boundary conditions applied to NPT ensemble (constant temperature and pressure) having a padding 2 Å bigger than SolBox1 to compensate for allowing volume variation (Table S3b). To properly compare, final energies were retrieved for SolBox1, SolBox2 and a full simulation of 10 ns (Full) as described above.

Once the final complexes were obtained, either from minimization or full simulation, we determined their free energies (ΔG, kcal mol^{-1}), which are used to predict binding events, and that depend on the contribution of internal and solvent energies (ΔE_{int} and ΔG_{sol}; Figure S1). Entropy variations were not calculated since available calculations are still not reliable enough and they cause the majority of the computational cost in the energy calculations (Sadiq *et al.*, 2010; Hou *et al.*, 2011; Genheden and Ryde, 2015). Energy parameters of the final complexes were calculated following the MM/PBSA protocol with the steps summarized in Figure S1 (Hou *et al.*, 2011): (i) internal energies (ΔE_{int}), which are the sum of the electrostatic and Van der Waals interactions; (ii) solvent energies (ΔG_{sol}), which are the sum of polar and non-polar terms, with the polar terms being calculated applying the Poisson–Boltzmann potential to the complex, and then these terms were subtracted separately from the potential of the ECD-PRR alone in the complex and the potential of the glycan in the complex; (iii) ΔG_{sol} non-polar terms, calculated applying a conversion factor into the SASA values (gSASA+*b*) of the complex and the PRR alone (Gilson and Zhou, 2007; Genheden and Ryde, 2015; Das *et al.*, 2018); and then (iv) internal and solvent energies were summed up to obtain the final ΔG energy.

The well-known complex CERK1/1,4-β-D-$(GlcNAc)_6$ was used as positive control for optimizing the methodology (Liu *et al.*, 2012; Cao *et al.*, 2014; Cheval *et al.*, 2020). Three final complexes, two minimization procedures (SolBox1 and SolBox2) and one full simulation procedure (Full) were obtained for the CERK1/1,4-β-D-$(GlcNAc)_6$ complex, which were in accordance with the described binding. Based on the energies obtained with this well-characterized complex, we established some minimal conditions supporting putative glycoligand/PRR interaction assessment (Figure 3): (i) a minimum of five computational replicates were established to assess the binding with reliability; (ii) ΔG energy average values below −6 kcal mol^{-1} (standard deviation of maximum 30%) were considered as evidence of interaction, while higher values may indicate non-specific binding, no binding or even repulsion; (iii) non-polar (ΔG_{np}) values above 0.6 kcal mol^{-1} were considered as reliable and indicative of binding, while conformations with values below 0.6 kcal mol^{-1} were discarded even though the final energy was negative since energy values below 1 kT (where k is the Boltzmann constant and T is the absolute temperature) are noisy and irrelevant because kT or thermic energy is around 0.6–0.7 kcal mol^{-1} at a thermodynamic T of 298 K; and (iv) Van der Waals energies must be more negative than the electrostatic ones with a sufficient difference amongst them of ≥18 kcal mol^{-1} to consider binding conformation as reliable as noted in previous studies (Sadiq *et al.*, 2010; Hou *et al.*, 2011; Kumar *et al.*, 2019; Saravanan *et al.*, 2020; Peng *et al.*, 2020). As additional controls of this last parameter established in the model, we determined the energies of CERK1-, LYK4- and THE1-ECD against an amount of water molecules similar to the size of the ligand (box of 10.0 Å; absence of ligands) and we found that electrostatic terms were more negative than the Van der Waals ones (Table S4), indicating that when ECD-PRRs are bound to the glycan, electrostatic interactions lower their energy values below those of Van der Waals values that become more negative than electrostatic ones (Table 1).

In silico binding prediction of selected LysM-PRR/glycan complexes

The above-described pipeline was applied to predict putative interactions between chitin oligosaccharides of different DP and proteins that have been described to bind chitin oligomers. We first applied the pipeline to determine the binding of chitin oligomers of DP2 and DP4 to

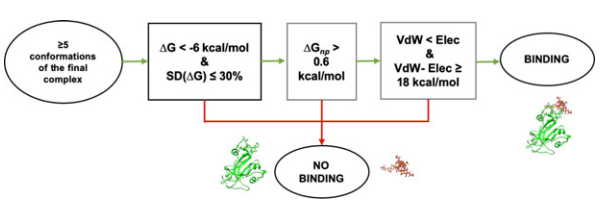

Figure 3. Decision tree for putative binding assessment for ECD-PRR/glycan pairs. Three conditions must be sequentially fulfilled to predict trustable energies and ECD-PRR/glycan putative binding (by using five or more conformations of the complex): (i) negative final energy values (< −6 kcal mol^{-1}) with <30% of variation from average, (ii) non-polar term values higher than 0.6 kcal mol^{-1} and (iii) electrostatic interaction energy values at least 18 kcal mol^{-1} higher than the Van der Waals interaction counterparts.

Table 1 Energy values of CERK1/glycan and LYK4/glycan trials obtained with the *in silico* computational minimization procedure

PRR/glycan	Trial[a]	ΔE_{elec}	ΔE_{vdW}	ΔG_{np}	ΔG_{pol}	ΔG
CERK1/1,4-β-D-$(GlcNAc)_6$	SolBox1	−26.1 ± 2.4	−47.2 ± 3.5	1.8 ± 0.4	49.8 ± 8.9	**−21.6 ± 4.6**
	SolBox2	−26.2 ± 3.2	−46.1 ± 3.3	1.7 ± 0.3	51.3 ± 6.5	**−19.3 ± 4.2**
	Full	−12.6 ± 8.6	−35.1 ± 5.6	2.9 ± 0.4	41.4 ± 14.5	−3.3 ± 16.6
CERK1/1,3-β-D-$(Glc)_6$	SolBox1	−23.3 ± 7.7	−37.4 ± 5.1	1.0 ± 0.3	46.8 ± 10.8	−12.9 ± 10.8
	SolBox2	−24.4 ± 7.6	−37.4 ± 5.3	0.9 ± 0.2	49.8 ± 24.4	−11.1 ± 14.2
	Full	−4.9 ± 9.9	−11.6 ± 14.6	1.6 ± 1.8	16.1 ± 23.0	1.2 ± 4.1
CERK1/1,4-β-D-$(Glc)_6$	SolBox1	−21.2 ± 2.3	−34.8 ± 3.0	1.2 ± 0.2	46.2 ± 6.0	−8.7 ± 7.5
	SolBox2	−21.7 ± 2.3	−34.7 ± 3.5	1.2 ± 0.3	42.9 ± 9.8	−12.2 ± 10.7
	Full	−5.7 ± 6.5	−13.1 ± 12.8	1.5 ± 1.4	14.3 ± 26.2	−3.0 ± 22.4
LYK4/1,4-β-D-$(GlcNAc)_6$	SolBox1	−21.9 ± 4.1	−41.1 ± 11.7	2.3 ± 0.3	39.8 ± 9.7	−20.8 ± 15.5
	SolBox2	−21.8 ± 3.6	−42.6 ± 12.5	2.1 ± 0.3	37.8 ± 5.1	−24.4 ± 11.0
	Full	−6.9 ± 2.5	−22.5 ± 5.5	3.7 ± 0.5	26.2 ± 15.8	0.4 ± 14.5
LYK4/1,3-β-D-$(Glc)_6$	SolBox1	−19.7 ± 9.0	−26.1 ± 5.8	1.8 ± 0.2	37.2 ± 13.4	−8.2 ± 8.9
	SolBox2	−20.1 ± 8.9	−26.5 ± 3.2	1.6 ± 0.3	33.8 ± 13.3	−11.3 ± 11.6
	Full	−3.5 ± 4.5	−12.9 ± 12.4	2.5 ± 1.5	6.3 ± 8.8	−7.5 ± 9.5
LYK4/1,4-β-D-$(Glc)_6$	SolBox1	−25.1 ± 3.4	−29.1 ± 4.0	1.6 ± 0.3	46.7 ± 5.2	−6.0 ± 8.5
	SolBox2	−26.5 ± 2.1	−32.8 ± 3.8	1.2 ± 0.3	47.0 ± 7.1	−11.4 ± 7.8
	Full	−0.9 ± 2.0	−13.4 ± 12.1	1.6 ± 1.4	12.4 ± 10.9	−0.3 ± 3.3

[a]SolBox1, SolBox2 and full trials for every PRR–glycan pair are presented as the average ± SD of five replicates. Positive binding parameters fulfilling the criteria of Figure 3 are highlighted in bold. Only the CERK1/1,4-β-D-$(GlcNAc)_6$ pair was assessed as binding in both SolBox1 and SolBox2 trials.

Arabidopsis CERK1 (CERK1) and we found that CERK1 was predicted to bind DP4, but not the DP2 oligomer (Figure S2), as described previously (lizasa *et al.*, 2010; Wan *et al.*, 2012). Then we applied the pipeline to predict binding of the DP4 oligomer to OsCEBiP, the rice chitin receptor, and its co-receptor OsCERK1, and we found that the pipeline predicted OsCEBiP binding to the DP4 oligomer, but not to OsCERK1 (Figure S2), as described previously (Liu *et al.*, 2016). Next, we applied the pipeline to the fungal effectors *Cf*Avr4F and Mg1LysM, which bind to the DP6 chitin oligomer (Hurlburt *et al.*, 2018; Sánchez-Vallet *et al.*, 2020), and we found that the model positively predicted their binding to the DP6 oligomer (Figure S2), further confirming the robustness of the pipeline developed (Figure 3).

We next expanded the application of the pipeline to obtain predictions of the potential binding of ECDs from Arabidopsis CERK1 and LYK4 to several hexasaccharides that have been described to trigger PTI responses: 1,3-β-D-$(Glc)_6$ and 1,4-β-D-$(Glc)_6$. ECD-THE1 from the *Cr*RLK1L subclass was included in the analysis as a negative control (Table S5). THE1 has been proposed to function as a plant cell wall integrity hub, though it has been recently demonstrated to bind Rapid Alkalinization-Like Factor (RALF) peptides rather than glycans (Hématy *et al.*, 2007; Gonneau *et al.*, 2018). Three final complexes, two minimization procedures (SolBox1 and SolBox2) and one full simulation procedure (Full) were obtained for each PRR–glycan complex. In modeling calculations performed between CERK1 and 1,3-β-D-$(Glc)_6$ we obtained final ΔG energies and nonpolar term values compatible with binding events, but electrostatic and Van der Waals energy ratios and the standard deviation of the final binding ΔG energy did not fulfill the established criteria based on CERK1/1,4-β-D-$(GlcNAc)_6$ interaction determinations, which matched all the established criteria for binding (Table 1). In the case of LYK4, the standard deviation obtained for the calculated values was very high for both ligands, not allowing to predict direct binding of these ligands to LYK4 (Table 1). We next determined the potential binding of the cellulose-derived hexasaccharide DAMP 1,4-β-D-$(Glc)_6$ to ECDs of these PRRs by obtaining the glycan structure from scratch, and the obtained data were incompatible with binding based on the established criteria (Table 1 and Figure 3). Similarly, modeling of ECD-THE1, included as a negative control, and of 1,3-β-D-$(Glc)_6$, 1,4-β-D-$(Glc)_6$ and 1,4-β-D-$(GlcNAc)_6$ yielded parameters that were not compatible with binding of these three glycans (Table S5). Moreover, full trials of the glycan/ECD-PRR analyses gave no clear attraction results, proving that most trajectories were not stable enough, further discarding the Full modeling trials as possible test settings for standard parameter binding prediction.

To further confirm that glycan structures generated from scratch did not affect the modeling calculations, we retrieved the glycan structure of cellohexaose (1,4-β-D-$(Glc)_6$) from the 3D crystal structure of cellobiohydrolase I from *Trichoderma reesei* (PDB: 7CEL; Divne *et al.*, 1998). Notably, we found that this structure is essentially identical to that of *in silico* modeled cellohexaose (Figure S3a), supporting the method developed to optimize glycan structures in solution from scratch (Figure 2). Then, we repeated the modeling calculation with the established

pipeline using this solved structure of cellohexaose and we confirmed the lack of binding of CERK1, LYK4 and THE1 to this glycan (Figure S3b), as predicted with the structure obtained from the *in silico* pipeline (Table 1).

In vitro binding assays confirmed *in silico* predictions

The *in silico* pipeline predicted a clear direct interaction of 1,4-β-D-$(GlcNAc)_6$ with CERK1, but not with the remaining hexasaccharides tested (Table 1). These data were afterwards validated by assessing the capacity of these ligands to trigger an early plant immune event such as Ca^{2+} influxes by using Arabidopsis wild-type and CERK1-defective sensor lines (Col-0^{AEQ} and *cerk1-2*^{AEQ}, respectively; Ranf *et al.*, 2012). Col-0^{AEQ} and *cerk1-2*^{AEQ} seedlings were incubated with the different ligands and Ca^{2+} influxes were monitored through luminescence measurements (Figure 4). All three hexasaccharides triggered Ca^{2+} influxes in Col-0^{AEQ} lines, confirming that they were active glycan ligands in Arabidopsis, as described (Cao *et al.*, 2014; Mélida *et al.*, 2018; Locci *et al.*, 2019): 1,4-β-D-$(GlcNAc)_6$ and 1,3-β-D-$(Glc)_6$ produced a wide peak at about 90 sec after treatment followed by a maintained decrease in luminescence that lasted about 600 sec. Ca^{2+} influx kinetics upon 1,4-β-D-$(Glc)_6$ treatment were different, since cellohexaose triggered a faster response that was followed by a faster luminescence disappearance (Figure 4c). Notably, Ca^{2+} influxes obtained with *cerk1-2*^{AEQ} upon treatment with 1,4-β-D-$(Glc)_6$ were similar to those observed in Col-0^{AEQ} lines, indicating that CERK1 was not required for 1,4-β-D-$(Glc)_6$ perception, whereas 1,4-β-D-$(GlcNAc)_6$ did not active any Ca^{2+} influxes in *cerk1-2*^{AEQ}, as predicted by our *in silico* models and described previously (Table 1; Mélida *et al.*, 2018). In contrast *cerk1-2*^{AEQ} results for 1,3-β-D-$(Glc)_6$ demonstrated a full dependence on CERK1, as described previously (Mélida *et al.*, 2018), which was not consistent with the predictions of our molecular dynamics models (Table 1).

In view of these results and to further clarify this issue, we expressed CERK1 and LYK4 Arabidopsis ECDs in insect cells and purified them by affinity chromatography (Figure S4). Then, isothermal titration calorimetry (ITC) experiments (Sandoval and Santiago, 2020) were carried out to confirm the *in silico* binding predictions of both CERK1-dependent ligands 1,4-β-D-$(GlcNAc)_6$ and 1,3-β-D-$(Glc)_6$. Our ITC results proved direct interactions of CERK1 with 1,4-β-D-$(GlcNac)_6$ (K_d values of 37.5 ± 10.0 μM; Figure 5). The results obtained with 1,3-β-D-$(Glc)_6$ clearly indicated that CERK1 did not bind, at least directly, to this glycan (Figure 5), confirming the *in silico* predictions (Table 1). Similar binding experiments were performed with LYK4 and these two glycans, but no direct binding was detected (Figure 5), which also supported our *in silico* prediction (Table 1). These data suggest that CERK1 takes part of the sensing complex for 1,3-β-D-$(Glc)_6$ in an indirect manner.

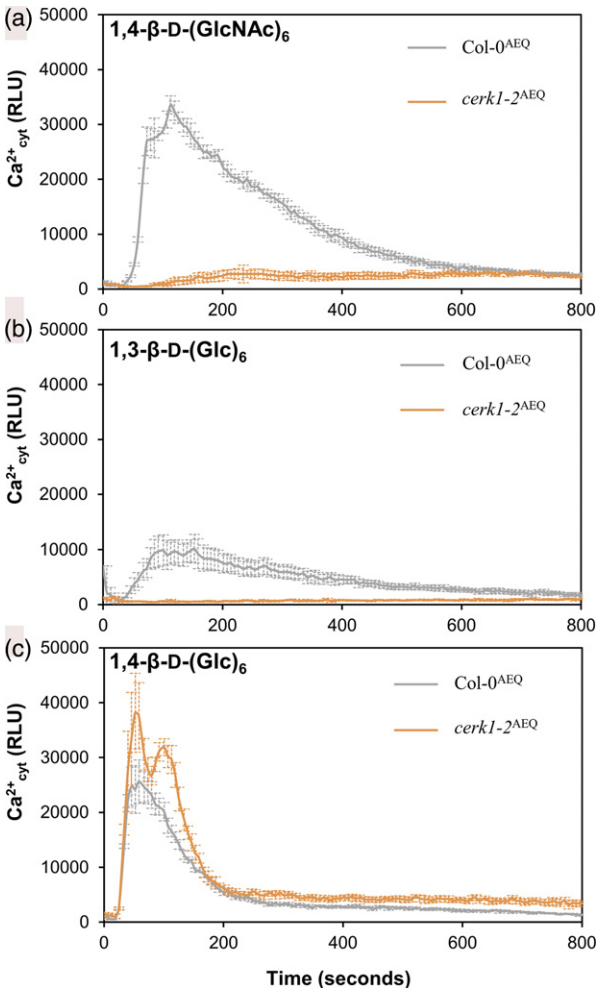

Figure 4. Elevations of cytoplasmic calcium concentrations over time in 8-day-old Arabidopsis Col-0^{AEQ} and *cerk1-2*^{AEQ} seedlings upon treatment with hesaccharides. (a) Chitohexaose (250 μM 1,4-β-D-$(GlcNAc)_6$). (b) Laminarihexaose (250 μM 1,3-β-D-$(Glc)_6$). (c) Cellohexaose (250 μM 1,4-β-D-$(Glc)_6$). Data are presented as the mean ± SD (n = 8). Shown is one of three experiments that gave similar results.

Similarly, LYK4 may be involved as a co-receptor in the recognition of chitin glycans but may not physically bind chitin, as recently suggested (Cheval *et al.*, 2020). As predicted by our model, the THE1 ectodomain did not bind any of the CERK1-dependent ligands tested (Figure S5).

In silico determination of the binding interface residues of CERK1/1,4-β-D-$(GlcNAc)_6$

Since the *in silico* modeling pipeline was found to be robust, we next tested its capacity to predict the amino acids of the CERK1 binding pocket involved in chitin binding. It has been previously shown in the crystal structure of CERK1 (PDB: 4EBZ; Liu *et al.*, 2012) that residues E110, E114, Q109 and I141 seem to be essential for 1,4-β-D-$(GlcNAc)_4$ binding. We carried out an initial comparison of

(a)

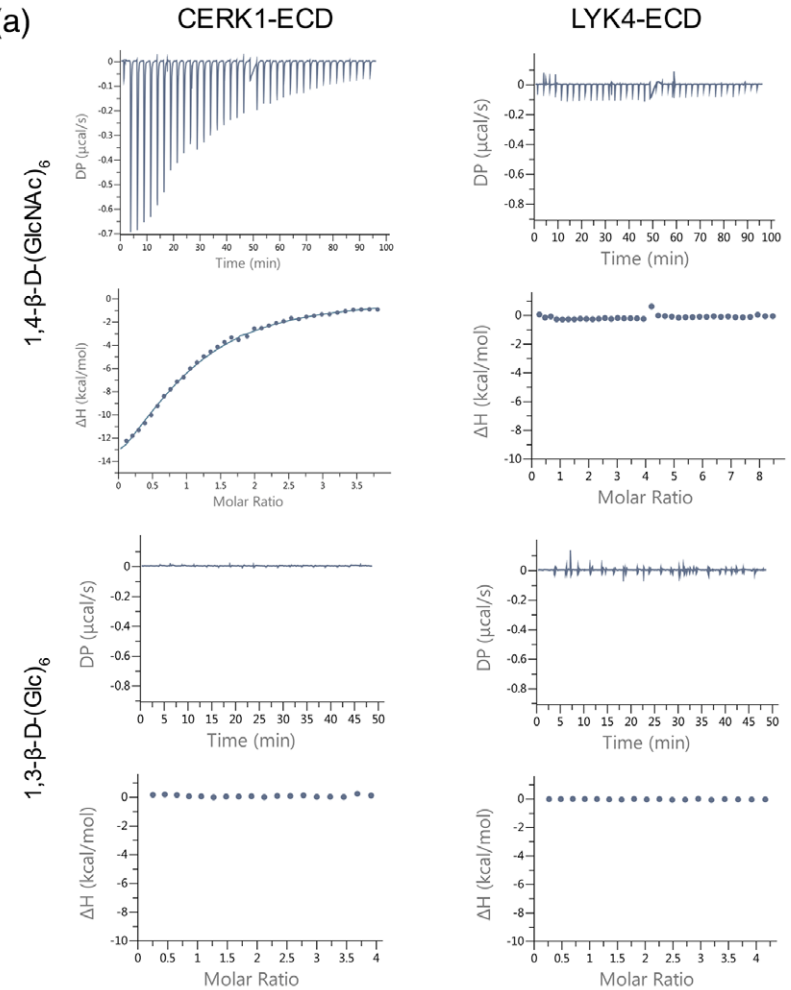

Figure 5. Direct binding assays between the PRR-ECDs and the corresponding glycans. (a) Isothermal titration calorimetry (ITC) experiments of CERK1 and LYK4 ECDs vs 1,4-β-D-$(GlcNAc)_6$ and 1,3-β-D-$(Glc)_6$. (b) ITC table summaries of PRR-ECDs vs 1,4-β-D-$(GlcNAc)_6$ and 1,3-β-D-$(Glc)_6$. The binding affinities between CERK1 and LYK4 and the corresponding glycans are reported as K_d (dissociation constant, in micromoles); N indicates the reaction stoichiometry ($N = 1$ for a 1:1 interaction); ΔH indicates the enthalpy variation. Values indicated in the table are means ± SD of independent experiments ($n = 2$). N.d. indicates no binding detected.

(b)

Ligand (syringe)	Protein (cell)	Kd (uM)	ΔH (kcal/mol)	N
1,4-β-D-$(GlcNAc)_6$	CERK1	37.5 ± 10	-16 ± 1.2	1
1,4-β-D-$(GlcNAc)_6$	LYK4	n.d		
1,3-β-D-$(Glc)_6$	CERK1	n.d		
1,3-β-D-$(Glc)_6$	LYK4	n.d		

the common residues involved in chitin binding of our *in silico* replicates of CERK1-1,4-β-D-$(GlcNAc)_6$ and the CERK1-1,4-β-D-$(GlcNAc)_4$ crystal and we discovered that in our *in silico* models E110, E114 and Q109 formed hydrogen bonds with the ligand in all the replicates and I141 in some models. Next, we performed *in silico* mutations of one or two of these residues by replacing them with alanine in CERK1 (Q109A, E110A, E114A, I141A, Q109A/E110A, E110A/E114A, E110A/I141A, Q109A/I141A and E114A/I141A). These *cerk1 in silico* mutants were subjected to modeling tests against 1,4-β-D-$(GlcNAc)_6$ using the SolBox2 procedure with five replicates. Notably, some individual mutations (e.g., Q109A and I141A) and double mutations (e.g., Q109A/E110A, E110A/E114A and Q109A/I141A) resulted in negative binding predictions of *cerk1* to 1,4-β-D-$(GlcNAc)_6$ (Figure S6), further supporting the relevance of binding of the residues identified in the 4EBZ PDB structure (Liu *et al.*, 2012). These results also indicate that the pipeline described here has the potential to predict the impact on binding of key residues mutations, thus enabling the generation of testable hypotheses for biochemical validation.

DISCUSSION

The plant immune and mammal innate immune systems share a similar conceptual logic, but plants lack the adaptive immunity of mammals (Nürnberger *et al.*, 2004; Haney *et al.*, 2014). Many non-self- (e.g., from microorganisms) and self-macromolecules (e.g., from plants) harbor molecular patterns, MAMPs and DAMPs, respectively, that have been selected through evolution by plants to be perceived by their diverse set of PRRs (617 putative members in Arabidopsis: Figure 1). Modification of these macromolecules, precursors of MAMPs/DAMPs, by plant enzymes might mask, but also unmask, their immunogenic epitopes. In this context, PRRs are essential components for epitope monitoring, allowing plants to recognize and respond to the high diversity of signals that they are exposed to. Moreover, immune signaling upon pattern perception relies not only on the PRRs themselves, but also on immune signaling complexes consisting of membrane-bound and intracellular proteins that are involved in scaffolding or signal transduction (Albert *et al.*, 2020).

The molecular variety of MAMPs/DAMPs requires different ECD structures and properties for recognition (e.g., binding) and immune activation (Saijo *et al.*, 2018). In this work we present an updated classification of the 617 putative Arabidopsis PRRs with extracellular ECDs that comprises at least 14 different subclasses of receptors based on the ECD sequence and conformation (Table S1). LRR-PRRs is the largest group, and shows some similarities with some PRRs in animals, most notably Toll-like receptors (TLRs) that have an LRR-ECD, a transmembrane motif and a cytoplasmic Toll-like domain instead of the kinase domain of plant RLKs (Ronald and Beutler, 2010). Plant LRR-PRRs, which have received most attention from researchers, typically recognize proteinaceous ligands that trigger developmental and immunogenic processes (Smakowska-Luzan *et al.*, 2018). However, 53.3% of Arabidopsis PRRs harbor ECDs putatively able to interact with glycan-based ligands, but these PRRs have not received much attention. These putative glycan-binding PRRs were classified as C-, G- and L-lectins, CRK/stress-antifungal, *Cr*RLK1L, LysM, LRR-malectins, malectins, ThaumatinL/PR5 and WAKs according to their ECD typology (Figure 1 and Table S1). CRKs constitute the most abundant subclass (68 members), followed by L- and G-lectins and malectins (55, 51 and 51 members, respectively; Figure 1). Some members of these groups have already been demonstrated to interact with carbohydrates and others have been associated with the binding of patterns of different nature (Schallus *et al.*, 2008; de Oliveira Figueiroa *et al.*, 2017; Bacete *et al.*, 2018; Moussu *et al.*, 2018; Cheung *et al.*, 2020).

WAK RLKs, with ECDs containing epidermal growth factor motifs, are the proposed receptors for OGs (He *et al.*, 1996; Anderson *et al.*, 2001; Kohorn and Kohorn, 2012). Indeed, a recombinant peptide containing amino acids 67 to 254 of the extracellular domain of Arabidopsis WAK1 bound polygalacturonic acid, OGs, pectins and structurally related alginates (Decreux and Messiaen, 2005). However, a crystal structure of WAK-ECDs is not yet available and detailed structural work is still necessary in order to fully demonstrate OG–WAK interaction. On the other hand, considering that the WAK subclass includes 38 members in Arabidopsis, it would be expected that some of them could recognize other carbohydrate-based patterns, like recently described OG variants (Voxeur *et al.*, 2019), but this hypothesis requires further investigations.

Malectin-like domains in the ECDs of *Cr*RLK1L members are known to bind di-glucose in their animal counterparts (Schallus *et al.*, 2008). Therefore, *Cr*RLK1Ls were suggested to bind carbohydrates, but this hypothesis has not been experimentally demonstrated to date (Lindner *et al.*, 2012; Wolf, 2017). The crystal structures of the C*r*RLK1Ls ANX1 and ANX2 suggest that they are non-canonical malectins or carbohydrate-binding modules (CBMs) as they lack the conserved binding surfaces for carbohydrate ligands (Du *et al.*, 2018; Moussu *et al.*, 2018). Instead, some *Cr*RLK1L members have been demonstrated to bind peptides. For example, FERONIA (FER), one of the best-characterized members of the *Cr*RLK1L subclass, has been shown to bind the peptides RALF1 and RALF23 (Haruta *et al.*, 2014; Stegmann *et al.*, 2017). In spite of this RALF-binding ability, recent results also support FER's role as a sensor of cell wall damage through its interaction with pectins, which could postulate a dual interacting capacity with both peptides and carbohydrates (Feng *et al.*, 2018; Duan *et al.*, 2020). Similarly, lectins are proteins with domains that are well known to bind to carbohydrates and probably have been co-opted for immunity function from their original functions (e.g., catalytic) (Goldstein *et al.*, 1980). Plant lectin-PRRs were defined by homology to their mammalian relatives, and although several non-proteinaceous ligands are known to be perceived by lectin-type RLKs, no ligands of glycan nature have been described for these PRRs. The bulb-type G-lectin LIPOOLIGOSACCHARIDE-SPECIFIC REDUCED ELICITATION (LORE) was initially reported to sense bacterial lipopolysaccharides in Arabidopsis, but later on it was shown that ligands were indeed medium-chain 3-hydroxy fatty acids that copurify with lipopolysaccharides (Ranf *et al.*, 2015; Kutschera *et al.*, 2019). Extracellular ATP (eATP) is one of the best-studied DAMPs in animals. Identification in Arabidopsis of the eATP receptor, the L-lectin DOES NOT RESPOND TO NUCLEOTIDES1 (DORN1), was a major breakthrough in eATP biology and provided a key to address many questions about eATP in plants (Choi *et al.*, 2014). Also, LecRK-I.8 and LecRK-VI.2 L-lectin RLKs have been associated to extracellular NAD (eNAD) and NADP (eNADP) coenzyme perception in Arabidopsis (Singh *et al.*, 2012; Wang *et al.*, 2017), and LecRK-

l.8 has also been shown to be involved in early steps of egg recognition of the white butterfly (Gouhier-Darimont *et al.*, 2019).

LysMs PRRs are the only group of plant PRRs of which all members have been exclusively related to glycan perception, including chitin, peptidoglycans, 1,3-β-glucans and lipopolysaccharides (Miya *et al.*, 2007; Willmann *et al.*, 2011; Mélida *et al.*, 2018). In contrast to WAKs, the availability of crystal structures of several LysM-ECDs (e.g., 4EBZ, 5JCE, 5BUM, 5K2L) made this subclass amenable for the optimization of the computational-based pipeline to study ECD–glycoligand interactions. Until very recently, the main 'Achilles heel' hampering the discovery of PRR–glycan pairs was considered to be the very limited number of identified MAMPs/DAMPs of this nature (Bacete *et al.*, 2018). However, recent discoveries have considerably increased the collection of active glycoligands in plants (de Azevedo Souza *et al.*, 2017; Claverie *et al.*, 2018; Johnson *et al.*, 2018; Mélida *et al.*, 2018, Mélida *et al.*, 2020; Locci *et al.*, 2019; Zang *et al.*, 2019; Wanke *et al.*, 2020b). Moreover, the development of tools such as synthetic glycan collections will soon facilitate growth of the list (Ruprecht *et al.*, 2020). However, none of these MAMPs/DAMPs, including cello-, laminari-, xyloglucan-, mannan- and arabinoxylo-oligosaccharides, have a single candidate PRR for their perception, whose elucidation will require the combination of heavy biochemical and genetic approaches. In particular, biomolecular interaction assays with pure ECDs of putative receptors have been shown to be the golden standard to allow quantification of potential MAMP/DAMP-ECD binding (Sandoval and Santiago, 2020). However, these approaches are time consuming and resource demanding, thus high-throughput screenings based on these technologies are of high risk and not always affordable.

Here we describe a computational minimization procedure that could favor preliminary screenings by narrowing down ECD candidates to be tested in further biomolecular assays. Using the proposed computational minimization procedures, either SolBox1 or SolBox2 determinations would be enough for a rough binding prediction between an ECD and a glycan. Both solvation boxes yield similar results, thus using any of them (or both) would be a user choice. However, 'user-friendly' access to full molecular dynamics simulations with standard parameters proved not to be reliable in the prediction results and is therefore not recommended for the pursued objective of the pipeline described here. The final purpose of the proposed protocol would then be to achieve minimization (the first step of molecular dynamics simulation), which at the end will allow users to obtain results in a computational time frame ranging from minutes to hours, while performing full simulations will increase the time frame from hours to days, depending on computing power and the overall difficulty of the protocol. Additionally, we have developed an *in silico* method to optimize glycan structures in solvent ('structures from scratch'), which has been proved here to generate structures that are similar to glycan crystal structures (Table 1 and Figure S3), thus expanding the possibility to test additional glycan structures without reported crystals, using the computational minimization procedure described here.

In Arabidopsis, chitin is perceived by a complex of LysM-RLKs comprising LYK4, LYK5 and CERK1 (Miya *et al.*, 2007; Cao *et al.*, 2014; Xue *et al.*, 2019). LYK4 and LYK5 interact constitutively (Xue *et al.*, 2019), whereas the LYK5–CERK1 interaction is ligand-dependent (Cao *et al.*, 2014). Our *in silico* computational minimization procedure as well as our *in vitro* analyses confirmed the direct interaction of CERK1 with chitin oligosaccharides of DP4 or DP6 (Figure S2), whereas we discarded the direct LYK4–chitin interaction without the support of other LysM partners (Table 1 and Figure 5). Therefore, it is suggested that LYK4 may help in the formation of a CERK1–LYK5 complex for the recognition of chitin at the plasma membrane. This agrees with recent data that suggested that LYK4 may be in complex with LYK5 at the plasma membrane, and when the chitin concentration rises, LYK4 would dissociate from LYK5 to associate with LYM2 in the plasmodesmata region (Cheval *et al.*, 2020). Our *in silico* computational minimization procedure further supports these recent data. Similarly, the computational calculations performed here with rice OsCEBiP and OsCERK1 and some fungal effectors (*Cf*Avr4 and Mg1LysM) that bind chitin (Figure S2) also confirmed published data (Liu *et al.*, 2016; Hurlburt *et al.*, 2018; Sánchez-Vallet *et al.*, 2020) and supported the robustness of glycoligand/PRR binding predictions by the *in silico* computational minimization procedure described here. Remarkably, this technical advance might also allow the identification of key PRR residues involved in binding and stabilization of glycan structures, as shown here with the *in silico* validation of the relevance of Q109 and I141 from CERK1 in 1,4-β-D-$(GlcNAc)_6$ binding (Figure S6). Moreover, the methodology described might allow to simulate *in silico* point or multiple mutations in PRR binding pockets prior to their biochemical validation through expression of mutated ECD proteins.

β-glucans with 1,3-glycosidic linkages have been shown to be perceived as molecular patterns by different plant species (Klarzynski *et al.*, 2000; Mélida *et al.*, 2013; Wanke *et al.*, 2020b). In mammals, 1,3-β-glucans are recognized by Dectin-1, consisting of an extracellular C-type lectin domain connected to the plasma membrane by a stalk (Brown *et al.*, 2003). The extracellular lectin domain of Dectin-1 binds to 1,3-β-glucans and mixed 1,3-/1,6-β-glucans, with linear β-1,3-$(Glc)_{10}$ or a glucan heptasaccharide with one β-1,6-linked glucose side chain being the minimal structures required for Dectin-1 binding (Brown *et al.*,

2003; Palma *et al.*, 2006; Adams *et al.*, 2008). In Arabidopsis, the PRR required for 1,3-β-glucan perception is yet unknown; however, recent works involved CERK1 in its perception, at least for those with a DP > 5 (Figure 4; Mélida *et al.*, 2018; Wanke *et al.*, 2020b). However, a direct CERK1–1,3-β-D-$(Glc)_6$ interaction was fully discarded based on the molecular dynamics data obtained with the computational minimization method developed here (Figure 4). Notably, this *in silico* prediction of non-direct CERK1–1,3-β-D-$(Glc)_6$ binding was fully confirmed by ITC experiments (Table 1 and Figure 5), further suggesting a co-receptor function of CERK1 in the perception of 1,3-β-D-$(Glc)_6$. This mechanism of 1,3-β-D-$(Glc)_6$ perception by CERK1 has similarities to that described for the perception of chitin by OsCERK1 in rice through its interaction with the OsCEBiP receptor (Liu *et al.*, 2012), which has been also validated *in silico* here by demonstrating that OsCEBiP, but not OsCERK1, binds this ligand (Figure S2). Structural analyses have revealed a common activation mechanism for RLKs, in which the ECD domains of receptors (e.g., LRR and LysM subclasses) and shape complementary co-receptors heterodimerize in the presence of ligand (Albert *et al.*, 2020). Ligands promote dimerization either by binding both proteins directly as 'molecular glue' or allosterically through stabilization of an RLK island (Liu *et al.*, 2012; Sun *et al.*, 2013; Wang *et al.*, 2015; Hohmann *et al.*, 2017). It is known that such activation mechanism exists for LysM-RLKs upon chitin binding (Liu *et al.*, 2012), and our data and previous published data would lead to the hypothesis that CERK1 and maybe also its co-receptor partner LYK5 are involved in the recognition pathway as true receptors, or alternatively as co-receptors together with a *bona fide*, unknown receptor. Cellodextrins, oligosaccharides derived from cellulose, have recently emerged as a group of plant DAMPs showing a great potential for future investigations since they are in high abundance in all plant species and are active at low concentrations, at least in Arabidopsis (Aziz *et al.*, 2007; de Azevedo Souza *et al.*, 2017; Johnson *et al.*, 2018; Locci *et al.*, 2019). In spite of the interest generated by these DAMPs, little is known about their perception by plants and the PRRs involved. Our results clearly show that the main LysM-RLKs CERK1 and LYK4 are not involved in cello-oligosaccharide perception (Table 1).

The study of signal transduction in plants has expanded dramatically from the early efforts to define the basic components of signaling for most known hormones and environmental stresses to the current search for PRR/ligand pairs (Cheung *et al.*, 2020). Different types of ECD-PRRs mediate the perception of distinct MAMPs/DAMPs triggering immunity; however, the identification of receptor PRRs for a given ligand is only the tip of the iceberg. How ligand binding induces complex formation with co-receptors that are required for the activation of downstream immune signaling is really challenging (Wang and Chai, 2020). High-throughput technologies are thus required to accelerate the identification of PRR–ligand pairs. The computational prediction method of glycan/PRR binding presented here might accelerate the discovery of protein–glycan interactions and provide information on immune and/or developmental responses activated by glycoligands. Future work will include (i) the optimization of the method to the specificities of different ECD subclasses and (ii) genetic and biochemical functional validations of wild-type and *in silico* predicted mutant versions of PRRs.

EXPERIMENTAL PROCEDURES

Classification of PRRs with putative glycan-binding ECDs

Bioinformatic classification of PRR-ECDs was performed using public databases at UniProt (https://www.uniprot.org/; Bateman, 2019), ScanProsite (https://prosite.expasy.org/scanprosite/; de Castro *et al.*, 2006) and NCBI (https://www.ncbi.nlm.nih.gov/; Agarwala *et al.*, 2017). Extracellular domains were screened and checked with PFAM (https://pfam.xfam.org/; El-Gebali *et al.*, 2019) and InterPro (http://www.ebi.ac.uk/interpro/search/sequence/; Mitchell *et al.*, 2019) online servers. Transmembrane prediction was done with TMHMM (http://www.cbs.dtu.dk/services/TMHMM/; Krogh *et al.*, 2001), TMPred (https://embnet.vital-it.ch/software/TMPRED_form.html; Hofmann and Stoffel, 1993) and Das-TMFilter (http://www.enzim.hu/DAS/DAS.html; Cserzo *et al.*, 2004). GPI-anchored prediction was carried out with predGPI (http://gpcr.biocomp.unibo.it/predgpi/; Pierleoni *et al.*, 2008) and the presence of signal peptides was determined with SignalP (http://www.cbs.dtu.dk/services/SignalP/; Almagro Armenteros *et al.*, 2019) and PrediSi (http://www.predisi.de/; Hiller *et al.*, 2004). Curated transmembrane and signal peptide prediction was made for all sugar-binding PRRs with Phobius (https://www.ebi.ac.uk/Tools/pfa/phobius/; Käll *et al.*, 2007).

Theoretical protocol development

Glycan conformation was analyzed using the *Build Structure* tool of Chimera (Pettersen *et al.*, 2004), where carbohydrate structures were generated by adding atom by atom. A primary optimization in vacuum was performed with Chimera software and the *Minimize Structure* tool employing several rounds of a combined method of steepest descend (approximately 2000 steps per round) and conjugate gradient algorithms (approximately 300 steps per round). Hydrogen atoms were added considering hydrogen bonds and charges were computed with AMBER ff14SB and the Gasteiger algorithm. PSF and PDB glycan files were obtained from the *CHARMM-GUI PDB-Reader* Tool (Jo *et al.*, 2008; Lee *et al.*, 2016) and they were introduced into VMD 1.9.3 (Humphrey *et al.*, 1996) to create the solvent box with the *Add Solvation Box* tool (the maximum and minimum coordinates of the molecule are obtained with the 'measure minmax [atomselect top all]' command adding a 12-Å padding with TIP3 water). The models were neutralized and an NaCl concentration of 0.15 M was set with the *Add Ions* tool of VMD. Minimization was achieved with the *AutoIMD* tool from VMD treating all atoms in the system as mobile.

PRRs were modeled through Swiss-Model (https://swissmodel.expasy.org/; Waterhouse *et al.*, 2018) and Phyre2 (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index; Kelley *et al.*, 2015). All models and crystal structures were checked with MolProbity (http://molprobity.biochem.duke.edu/; Williams *et al.*,

2018). Binding sites that were not described were obtained with sequence alignment of different templates using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/; Madeira *et al.*, 2019).

AutoDock Vina (Trot and Olsson, 2010) from Chimera was used to create the initial complexes, usually padding with 27 Å in the binding site. Docked PRR–glycan complexes were introduced into CHARMM-GUI *PDB-Reader* to retrieve the coordinate PSF file and its corresponding PDB, and were then introduced into VMD to get the solvation box as stated above. Molecular dynamics analysis was performed using NAMD2 2.13v software (Phillips *et al.*, 2005). For minimization, 5000 steps were used, and 5 000 000 steps were used for the full simulations. Full simulations were computed with a temperature of 298 K, neutral pH and NVT ensemble.

Parameter information and configuration files employed in the analyses are presented in Table S3.

Energy calculations and binding assessment

Energies were calculated following a molecular mechanics approach with Poisson–Boltzmann and a surface area solvation MM/PBSA approach. PSF and corresponding DCD files were introduced into VMD to calculate electrostatic and Van der Waals interactions through the tool *NAMD energy*. Final coordinates were saved in a PDB file for further calculations. The PDB file was introduced in CHARMM-GUI *PBEQ Solver* to calculate the electrostatic potentials of the PRR–glycan complex and of the PRR and glycan alone. SASA calculations were retrieved for the PRR–glycan complex and for the PRR using SPDBV (http://www.expasy.org/spdbv/; Guex and Peitsch, 1997). Afterwards, a conversion was made to obtain $gSASA + b$ (where $g = 0.00526$ kcal mol^{-1} Å^{-2}, $b = 0.918$ kcal mol^{-1}).

In silico carbohydrates and PRRs

The 1,4-β-D-$(GlcNAc)_6$ and 1,3-β-D-$(Glc)_6$ structures were obtained from the 2PI8 and 1W9W PDB structures, respectively. The 1,4-β-D-$(Glc)_6$ structure was built from scratch as described above and was additionally retrieved from the 7CEL crystal. The 1,4-β-D-$(GlcNAc)_2$ and 1,4-β-D-$(GlcNAc)_4$ structures were built *in* silico and retrieved from 5JCE, respectively. The CERK1-ECD structure corresponds to crystal code 4EBZ, while LYK4 was modeled by homology using Swiss-Model and 5JCE as template. THE1 was modeled with Swiss-Model using 6FIG as template. Mg1LysM, *Cf*Avr4 and OsCEBiP correspond to 6Q40, 6BN0 and 5JCE, respectively. OsCERK1 was modeled using 4EBZ as template.

Comparison of ligand structures and study of the binding interface of CERK1

Comparison of the *in silico* 1,4-β-D-$(Glc)_6$ structure and the crystal 1,4-β-D-$(Glc)_6$ structure was made possible with LigPlot+ (Laskowski and Swindells, 2011). LigPlot+ was also used to check CERK1 residues involved in chitin recognition. Each *in silico* mutation study was performed by changing the residues of interest to alanine in the amino acid sequence of CERK1. Then, those sequences were modeled with Swiss-Model against the CERK1 template (4EBZ) to check for binding alterations.

Calcium influx assays

Eight-day-old, liquid-grown transgenic *Arabidopsis thaliana* seedlings of ecotype Col-0 carrying the calcium reporter aequorin (Col-0^{AEQ}; Ranf *et al.*, 2012) were used for cytoplasmic calcium ($Ca^{2+}{}_{cyt}$) measurements using a previously described method (Bacete *et al.*, 2017). The high-purity oligosaccharides used in these assays (1,4-β-D-$(GlcNAc)_6$, 1,3-β-D-$(Glc)_6$ and 1,4-β-D-$(Glc)_6$) were purchased from Megazyme Ltd., Bray, Ireland.

Protein expression and purification

Codon-optimized synthetic genes coding for Arabidopsis CERK1 (residues 1–232), LYK4 (residues 1–275) and THE1 (residues 1–415) ectodomains were cloned into a modified pFastBac (Geneva Biotec) vector, providing a tobacco etch virus protease cleavable C-terminal StrepII-9xHis tag, for expression in *Spodoptera frugiperda* (Invitro GeneArt, Germany). For protein expression, *Trichoplusia ni* Tnao38 cells (Hashimoto *et al.*, 2012) were infected with a multiplicity of infection of 3 and incubated at 28°C for 1 day and at 22°C for another 2 days, at 110 rpm. The secreted ectodomains were purified from the supernatant by sequential nickel affinity chromatography (HisTrap excel; GE Healthcare, Boston, MA, USA); equilibrated in phosphate K buffer, which contains 50 mM K_2HPO_4 and 500 mM NaCl, pH 7.8) and StrepII purification (Strep-Tactin Superflow high capacity; IBA; equilibrated in 25 mM Tris, pH 8.0, 250 mM NaCl, 1 mM EDTA). The proteins were further purified by size-exclusion chromatography on a Superdex 200 increase 10/300 GL column (GE Healthcare) or HiLoad 16/600 superdex 200 (GE Healthcare) equilibrated in 20 mM MES buffer pH 6.5, 150 mM NaCl. Peak fractions were analyzed by SDS-PAGE and then concentrated using Amicon Ultra concentrators (Millipore, MWCO 10 000) to reach a protein concentration of 50 μM.

Isothermal titration calorimetry

Experiments were performed at 25°C using a MicroCal PEAQ-ITC (Malvern Instruments, UK) with a 200-μl standard cell and a 40-μl titration syringe (Moussu *et al.*, 2020). Proteins were gel filtrated into ITC buffer (20 mM MES buffer, pH 6.5, 150 mM NaCl) and carbohydrates (1,4-β-D-$(GlcNAc)_6$ and 1,3-β-D-$(Glc)_6$) were dissolved in the same buffer. Each experiment was performed using an injection pattern of 2 μl of the carbohydrate ligand, at 500 μM or 1 mM, into 50 μM protein in the cell at 150-sec intervals. ITC data were corrected for the heat of dilution by subtracting the mixing enthalpies for titrant solution injections into protein-free ITC buffer. Experiments were done at least in duplicate and data were analyzed using the MicroCal PEAQ-ITC Analysis Software provided by the manufacturer.

Size-exclusion chromatography

Analytical gel filtration experiments were performed using a Superdex 200 increase 10/300 GL column (GE Healthcare) pre-equilibrated in 20 mM MES (pH 6.5) and 150 mM NaCl. Next, 100 μl of the isolated CERK1, LYK4 and THE1 ectodomains was loaded sequentially onto the column and elution at 0.7 ml min^{-1} was monitored by ultraviolet absorbance at 280 nm. The column was calibrated with a mixture of the high-molecular weight (HMW) and low-molecular weight (LMW) kits from GE Healthcare. Peak fractions were analyzed by SDS-PAGE.

ACKNOWLEDGMENTS

This work was supported by grants BIO2015-64077-R of the Spanish Ministry of Economy and Competitiveness (MINECO) and RTI2018-096975-B-I00 of the Spanish Ministry of Science, Innovation and Universities to AM. This work was also financially supported by the 'Severo Ochoa Programme for Centers of Excellence in R&D(2017–2021) from the Agencia Estatal de Investigación of Spain (grant SEV-2016-0672 to CBGP). In the frame of this program HM was supported with a postdoctoral fellow supported by SEV-2016-0672. IdH was the recipient of a PhD FPU

fellow (FPU16/07118) from the Spanish Ministry of Education and from an EMBO Short-Term Fellowship (7985). Research in JS's lab was financially supported by the European Research Council (ERC) grant agreement no. 716358, the Swiss National Science Foundation grants no. 31003A_173101 and the Programme Fondation Philanthropique Famille Sandoz. We thank Luis Fernández-Pacios (CBGP, Spain) for his advice on modeling calculations and the discussions of the results.

AUTHOR CONTRIBUTIONS

AM and HM initiated, conceived and coordinated all the experiments except those related to isothermal titration calorimetry (ITC), which were conceived and initiated by JS. IdH performed the experiments described in Figures 1–3, S1–S3 and S6 and Tables 1 and S1–S5 with help from AM and HM. HM performed the experiments described in Figure 4. Experiments described in Figures 5, S4 and S5 were performed by IdH, CB and JS. IdH and HM prepared the tables and figures. IdH, HM and AM wrote the paper. JS edited the paper.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

All relevant data can be found within the manuscript and its supporting materials.

SUPPORTING INFORMATION

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

Figure S1. Energy calculation procedure.

Figure S2. Energy values obtained in binding trials of 1,4-β-D-$(GlcNAc)_{2\text{-}6}$ and proteins with LysM domains.

Figure S3. Comparison data of 1,4-β-D-$(Glc)_6$ built *in silico* vs 1,4-β-D-$(Glc)_6$ from the 7CEL crystal.

Figure S4. Scheme of the expression constructs and size-exclusion profiles of the expressed proteins.

Figure S5. ITC control experiment of THE1 vs 1,4-β-D-$(GlcNAc)_6$ and 1,3-β-D-$(Glc)_6$.

Figure S6. *In silico* mutagenesis of key CERK1 residues for 1,4-β-D-$(GlcNAc)_6$ stabilization in the binding pocket.

Table S1. Classification of Arabidopsis PRRs with different ECDs.

Table S2. Of 617 PRRs in Arabidopsis, 329 have ECDs that can putatively bind glycans.

Table S3. Example of configuration files for the *in silico* protocol.

Table S4. Energy results obtained in the control trials of CERK1, LYK4 and THE1 against a 10.0-Å box of water molecules.

Table S5. Energy results for the different binding trials performed for THE1.

REFERENCES

Adams, E.L., Rice, P.J., Graves, B. *et al.* (2008) Differential high-affinity interaction of Dectin-1 with natural or synthetic glucans is dependent upon primary structure and is influenced by polymer chain length and side-chain branching. *J. Pharmacol. Exp. Ther.* **325**, 115–123.

Agarwala, R., Barrett, T., Beck, J. *et al.* (2017) Database resources of the national center for biotechnology information. *Nucleic Acids Res.* **45**, D12–D17.

Albert, I., Hua, C., Nürnberger, T., Pruitt, R.N. and Zhang, L. (2020) Surface sensor systems in plant immunity. *Plant Physiol.* **182**, 1582.

Almagro Armenteros, J.J., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O., Brunak, S., von Heijne, G. and Nielsen, H. (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **37**, 420–423.

Anderson, C.M., Wagner, T.A., Perret, M., He, Z.H., He, D. and Kohorn, B.D. (2001) WAKs: cell wall-associated kinases linking the cytoplasm to the extracellular matrix. *Plant Mol. Biol.* 197–206.

de Azevedo Souza, C., Li, S., Lin, A.Z., Boutrot, F., Grossmann, G., Zipfel, C. and Somerville, S.C. (2017) Cellulose-derived oligomers act as damage-associated molecular patterns and trigger defense-like responses. *Plant Physiol.* **173**, 2383–2398.

Aziz, A., Gauthier, A., Bézier, A., Poinssot, B., Joubert, J.M., Pugin, A., Heyraud, A. and Baillieul, F. (2007) Elicitor and resistance-inducing activities of β-1,4 cellodextrins in grapevine, comparison with β-1,3 glucans and α-1,4 oligogalacturonides. *J. Exp. Bot.* **58**, 1463–1472.

Bacete, L., Mélida, H., Pattathil, S., Hahn, M.G., Molina, A. and Miedes, E. (2017) Characterization of plant cell wall damage-associated molecular patterns regulating immune responses. In *Methods in Molecular Biology* (Shan, L., and He, P., eds). New York: Springer, pp. 13–23.

Bacete, L., Mélida, H., Miedes, E. and Molina, A. (2018) Plant cell wall-mediated immunity: cell wall changes trigger disease resistance responses. *Plant J.* **93**, 614–636.

Baltrus, D.A., Nishimura, M.T., Romanchuk, A. *et al.* (2011) Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 pseudomonas syringae isolates. *PLoS Pathog.* **7**, e1002132.

Bateman, A. (2019) UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515.

Bellande, K., Bono, J.J., Savelli, B., Jamet, E. and Canut, H. (2017) Plant lectins and lectin receptor-like kinases: How do they sense the outside? *Int. J. Mol. Sci.* **18**, 1164.

Boisson-Dernier, A., Kessler, S.A. and Grossniklaus, U. (2011) The walls have ears: the role of plant CrRLK1Ls in sensing and transducing extracellular signals. *J. Exp. Bot.* **62**, 1581–1591.

Boutrot, F. and Zipfel, C. (2017) Function, discovery, and exploitation of plant pattern recognition receptors for broad-spectrum disease resistance. *Annu. Rev. Phytopathol.* **55**, 257–286.

Brown, G.D., Herre, J., Williams, D.L., Willment, J.A., Marshall, A.S.J. and Gordon, S. (2003) Dectin-1 mediates the biological effects of β-glucans. *J. Exp. Med.* **197**, 1119–1124.

Cao, Y., Liang, Y., Tanaka, K., Nguyen, C.T., Jedrzejczak, R.P., Joachimiak, A. and Stacey, G. (2014) The kinase LYK5 is a major chitin receptor in Arabidopsis and forms a chitin-induced complex with related kinase CERK1. *Elife*, **3**, e03766.

Carpita, N. and McCann, M. (2000) The cell wall. In *Biochemistry and Molecular Biology of Plants* (Buchanan, B.B., Gruissem, W. and Jones, R.L., eds). Rockville, MD: American Society of Plant Physiologists, pp. 52–108.

de Castro, E., Sigrist, C.J.A., Gattiker, A., Bulliard, V., Langendijk-Genevaux, P.S., Gasteiger, E., Bairoch, A. and Hulo, N. (2006) ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res.* **34**, W362–W365.

Chen, S.M., Shen, H., Zhang, T. *et al.* (2017) Dectin-1 plays an important role in host defense against systemic *Candida glabrata* infection. *Virulence*, **8**, 1643–1656.

Cheung, A.Y., Qu, L.-J., Russinova, E., Zhao, Y. and Zipfel, C. (2020) Update on receptors and signaling. *Plant Physiol.* **182**, 1527.

Cheval, C., Samwald, S., Johnston, M.G. *et al.* (2020) Chitin perception in plasmodesmata characterizes submembrane immune-signaling specificity in plants. *Proc. Natl. Acad. Sci. USA*, **117**, 9621–9629.

Choi, J., Tanaka, K., Cao, Y., Qi, Y., Qiu, J., Liang, Y., Lee, S.Y. and Stacey, G. (2014) Identification of a plant receptor for extracellular ATP. *Science*, **343**, 290–294.

Claverie, J., Balacey, S., Lemaître-Guillier, C., *et al.* (2018) The cell wall-derived xyloglucan is a new DAMP triggering plant immunity in Vitis vinifera and Arabidopsis thaliana. *Frontiers in Plant Science*, **9**, 1725.

Cserzo, M., Eisenhaber, F., Eisenhaber, B. and Simon, I. (2004) TM or not TM: transmembrane protein prediction with low false positive rate using DAS-TMfilter. *Bioinformatics*, **20**, 136–137.

Dangl, J.L., Horvath, D.M. and Staskawicz, B.J. (2013) Pivoting the plant immune system from dissection to deployment. *Science* **341**, 746–751.

Das, S., Khatri, S., Siopsis, G. and Wilde, M.M. (2018) Fundamental limits on quantum dynamics based on entropy change. *J. Math. Phys.* **59**, 012205.

Decreux, A. and Messiaen, J. (2005) Wall-associated kinase WAK1 interacts with cell wall pectins in a calcium-induced conformation. *Plant Cell Physiol.* **46**, 268–278.

Desaki, Y., Miyata, K., Suzuki, M., Shibuya, N. and Kaku, H. (2018) Plant immunity and symbiosis signaling mediated by LysM receptors. *Innate Immun.* **24**, 92–100.

Díaz-Alvarez, L. and Ortega, E. (2017) The many roles of galectin-3, a multifaceted molecule, in innate immune responses against pathogens. *Mediators Inflamm.* **2017**, 1–10. https://doi.org/10.1155/2017/9247574.

Divne, C., Ståhlberg, J., Teeri, T.T. and Jones, T.A. (1998) High-resolution crystal structures reveal how a cellulose chain is bound in the 50 Å long tunnel of cellobiohydrolase I from *Trichoderma reesei*. *J. Mol. Biol.* **275**, 309–325.

Du, S., Qu, L.J. and Xiao, J. (2018) Crystal structures of the extracellular domains of the CrRLK1L receptor-like kinases ANXUR1 and ANXUR2. *Protein Sci.* **27**, 886–892.

Duan, Q., Liu, M.C.J., Kita, D. *et al.* (2020) FERONIA controls pectin- and nitric oxide-mediated male–female interaction. *Nature*, **579**, 561–566.

El-Gebali, S., Mistry, J., Bateman, A. *et al.* (2019) The Pfam protein families database in 2019. *Nucleic Acids Res.* **47**, D427–D432.

Faulkner, C., Petutschnig, E., Benitez-Alfonso, Y., Beck, M., Robatzek, S., Lipka, V. and Maule, A.J. (2013) LYM2-dependent chitin perception limits molecular flux via plasmodesmata. *Proc. Natl. Acad. Sci. USA*, **110**, 9166–9170.

Feng, W., Kita, D., Peaucelle, A. *et al.* (2018) The FERONIA receptor kinase maintains cell-wall integrity during salt stress through Ca2+ signaling. *Curr. Biol.* **28**, 666–675.

Franck, C.M., Westermann, J. and Boisson-Dernier, A. (2018) Plant malectin-like receptor kinases: from cell wall integrity to immunity and beyond. *Annu. Rev. Plant Biol.* **69**, 301–328.

Fratev, F., Steinbrecher, T. and Jónsdóttir, S.Ó. (2018) Prediction of accurate binding modes using combination of classical and accelerated molecular dynamics and free-energy perturbation calculations: an application to toxicity studies. *ACS Omega*, **3**, 4357–4371.

Fritz-Laylin, L.K., Krishnamurthy, N., Tör, M., Sjölander, K.V. and Jones, J.D.G. (2005) Phylogenomic analysis of the receptor-like proteins of rice and Arabidopsis. *Plant Physiol.* **138**, 611–623.

Genheden, S. and Ryde, U. (2015) The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* **10**, 449–461.

Gilson, M.K. and Zhou, H.X. (2007) Calculation of protein-ligand binding affinities. *Annu. Rev. Biophys. Biomol. Str.* **36**(1), 21–42.

Gimeno, A., Valverde, P., Ardá, A. and Jiménez-Barbero, J. (2020) Glycan structures and their interactions with proteins. A NMR view. *Curr. Opin. Struct. Biol.* **62**, 22–30.

Gish, L.A. and Clark, S.E. (2011) The RLK/Pelle family of kinases. *Plant J.* **66**, 117–127.

Goldstein, I.J., Hughes, R.C., Monsigny, M., Osawa, T. and Sharon, N. (1980) What should be called a lectin? [2]. *Nature*, **285**, 66.

Gonneau, M., Desprez, T., Martin, M. *et al.* (2018) Receptor kinase THESEUS1 is a rapid alkalinization factor 34 receptor in Arabidopsis. *Curr. Biol.* **28**, 2452–2458.

Gouhier-Darimont, C., Stahl, E., Glauser, G. and Reymond, P. (2019) The Arabidopsis lectin receptor kinase lecrk-i.8 is involved in insect egg perception. *Front. Plant Sci.* **10**, 623.

Greeff, C., Roux, M., Mundy, J. and Petersen, M. (2012) Receptor-like kinase complexes in plant innate immunity. *Front. Plant Sci.* **3**, 209.

Guex, N. and Peitsch, M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **18**, 2714–2723.

Gust, A.A., Biswas, R., Lenz, H.D. *et al.* (2007) Bacteria-derived peptidoglycans constitute pathogen-associated molecular patterns triggering innate immunity in Arabidopsis. *J. Biol. Chem.* **282**, 32338–32348.

Haab, B.B. and Klamer, Z. (2020) Advances in tools to determine the glycan-binding specificities of lectins and antibodies. *Mol. Cell. Proteomics*, **19**, 224–232.

Haney, C.H., Urbach, J.M. and Ausubel, F.M. (2014) Differences and similarities: Innate immunity in plants and animals. *Biochemistry (Lond)*, **36**, 40–45.

Haruta, M., Sabat, G., Stecker, K., Minkoff, B.B. and Sussman, M.R. (2014) A peptide hormone and its receptor protein kinase regulate plant cell expansion. *Science*, **343**, 408–411.

Hashimoto, E., Yabuta, Y., Watanabe, F., Morimoto, M., Yamaguchi, Y. and Takenaka, H. (2012) Purification and characterization of phycobiliproteins from edible cyanobacterium *Nostochopsis* sp. *Food Sci. Technol. Res.* **18**, 485–490.

He, Z.H., Fujiki, M. and Kohorn, B.D. (1996) A cell wall-associated, receptor-like protein kinase. *J. Biol. Chem.* **271**, 19789–19793.

Hématy, K., Sado, P.E., Tuinen, A.V., Rochange, S., Desnos, T., Balzergue, S., Pelletier, S., Renou, J.P. and Höfte, H. (2007) A receptor-like kinase mediates the response of Arabidopsis cells to the inhibition of cellulose synthesis. *Curr. Biol.* **17**, 922–931.

Hiller, K., Grote, A., Scheer, M., Münch, R. and Jahn, D. (2004) PrediSi: Prediction of signal peptides and their cleavage positions. *Nucleic Acids Res.* **32**, W375–W379.

Hofmann, K. and Stoffel, W. (1993) TMbase: a database of membrane spanning protein segments. *Biol. Chem.* **374**, 166.

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

Hou, T., Wang, J., Li, Y. and Wang, W. (2011) Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J. Chem. Inf. Model.* **51**, 69–82.

Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38.

Hurlburt, N.K., Chen, L.H., Stergiopoulos, I. and Fisher, A.J. (2018) Structure of the *Cladosporium fulvum* Avr4 effector in complex with (GlcNAc)6 reveals the ligand-binding mechanism and uncouples its intrinsic function from recognition by the Cf-4 resistance protein. *PLoS Pathog.* **14**, e1007263.

Iizasa, E., Mitsutomi, M. and Nagano, Y. (2010) Direct binding of a plant LysM receptor-like kinase, LysM RLK1/CERK1, to chitin in vitro. *J. Biol. Chem.* **285**, 2996–3004.

Isaacson, R.L. and Díaz-Moreno, I. (2019) Editorial: Weak interactions in molecular machinery. *Front. Mol. Biosci.* **5**, 117.

Jeong, S., Trotochaud, A.E. and Clark, S.E. (1999) The Arabidopsis CLAVATA2 gene encodes a receptor-like protein required for the stability of the CLAVATA1 receptor-like kinase. *Plant Cell*, **11**, 1925–1933.

Jo, S., Kim, T., Iyer, V.G. and Im, W. (2008) CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865.

Johnson, J.M., Thürich, J., Petutschnig, E.K. *et al.* (2018) A poly(A) ribonuclease controls the cellotriose-based interaction between *Piriformospora indica* and its host Arabidopsis. *Plant Physiol.* **176**, 2496–2514.

Käll, L., Krogh, A. and Sonnhammer, E.L.L. (2007) Advantages of combined transmembrane topology and signal peptide prediction-the Phobius web server. *Nucleic Acids Res.* **35**, W429–W432.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N. and Sternberg, M.J.E. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858.

Klarzynski, O., Plesse, B., Joubert, J.M., Yvin, J.C., Kopp, M., Kloareg, B. and Fritig, B. (2000) Linear β-1,3 glucans are elicitors of defense responses in tobacco. *Plant Physiol.* **124**, 1027–1038.

Kohorn, B.D. and Kohorn, S.L. (2012) The cell wall-associated kinases, WAKs, as pectin receptors. *Front. Plant Sci.* **3**, 88.

Krogh, A., Larsson, B., Heijne, G.V. and Sonnhammer, E.L.L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567–580.

Kumar, A., Srivastava, G., Negi, A.S. and Sharma, A. (2019) Docking, molecular dynamics, binding energy-MM-PBSA studies of naphthofuran derivatives to identify potential dual inhibitors against BACE-1 and GSK-3β. *J. Biomol. Struct. Dyn.* **37**, 275–290.

Kutschera, A., Dawid, C., Gisch, N. *et al.* (2019) Bacterial medium-chain 3-hydroxy fatty acid metabolites trigger immunity in Arabidopsis plants. *Science*, **364**, 178–181.

Laskowski, R.A. and Swindells, M.B. (2011) LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* **51**, 2778–2786.

Latgé, J.P. and Calderone, R. (2006) The fungal cell wall. In *Growth, Differentiation and Sexuality*. Berlin, Heidelberg: Springer, pp. 73–104.

Lee, J., Cheng, X., Swails, J.M. *et al.* (2016) CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the charmm36 additive force field. *J. Chem. Theory Comput.* **12**, 405–413.

Lehti-Shiu, M.D., Zou, C., Hanada, K. and Shiu, S.H. (2009) Evolutionary history and stress regulation of plant receptor-like kinase/pelle genes. *Plant Physiol.* **150**, 12–26.

Li, L., Yu, Y., Zhou, Z. and Zhou, J.M. (2016) Plant pattern-recognition receptors controlling innate immunity. *Sci. China Life Sci.* **59**, 878–888.

Li, Q., Wang, C. and Mou, Z. (2020) Perception of damaged self in plants. *Plant Physiol.* **182**, 1545–1565.

Lindner, H., Müller, L.M., Boisson-Dernier, A. and Grossniklaus, U. (2012) CrRLK1L receptor-like kinases: not just another brick in the wall. *Curr. Opin. Plant Biol.* **15**, 659–669.

Liu, S., Wang, J., Han, Z., Gong, X., Zhang, H. and Chai, J. (2016) Molecular mechanism for fungal cell wall recognition by rice chitin receptor OsCEBiP. *Structure*, **24**, 1192–1200.

Liu, T., Liu, Z., Song, C. *et al.* (2012) Chitin-induced dimerization activates a plant immune receptor. *Science*, **336**, 1160–1164.

Locci, F., Benedetti, M., Pontiggia, D., Citterico, M., Caprari, C., Mattei, B., Cervone, F. and Lorenzo, G.D. (2019) An Arabidopsis berberine bridge enzyme-like protein specifically oxidizes cellulose oligomers and plays a role in immunity. *Plant J.* **98**, 540–554.

Madeira, F., Park, Y.M., Lee, J. *et al.* (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641.

Mélida, H., Bacete, L., Ruprecht, C., Rebaque, D., Del Hierro, I., López, G., Brunner, F., Pfrengle, F. and Molina, A. (2020) Arabinoxylan-Oligosaccharides Act as Damage Associated Molecular Patterns in Plants Regulating Disease Resistance. *Front Plant Sci.*, **11**, 1210. https://doi.org/10.3389/fpls.2020.01210

Mélida, H., Sandoval-Sierra, J.V., Diéguez-Uribeondo, J. and Bulone, V. (2013) Analyses of extracellular carbohydrates in oomycetes unveil the existence of three different cell wall types. *Eukaryot. Cell*, **12**, 194–203.

Mélida, H., Sopeña-Torres, S., Bacete, L., Garrido-Arandia, M., Jordá, L., López, G., Muñoz-Barrios, A., Pacios, L.F. and Molina, A. (2018) Non-branched β-1,3-glucan oligosaccharides trigger immune responses in Arabidopsis. *Plant J.* **93**, 34–49.

Mende, M., Bordoni, V., Tsouka, A., Loeffler, F.F., Delbianco, M. and Seeberger, P.H. (2019) Multivalent glycan arrays. *Faraday Discuss.* **219**, 9–32.

Mitchell, A.L., Attwood, T.K., Babbitt, P.C. *et al.* (2019) InterPro in 2019: Improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res.* **47**, D351–D360.

Miya, A., Albert, P., Shinya, T. *et al.* (2007) CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in Arabidopsis. *Proc. Natl. Acad. Sci. USA*, **104**, 19613–19618.

Moussu, S., Augustin, S., Roman, A.O., Broyart, C. and Santiago, J. (2018) Crystal structures of two tandem malectin-like receptor kinases involved in plant reproduction. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 671–680.

Moussu, S., Broyart, C., Santos-Fernandez, G., Augustin, S., Wehrle, S., Grossniklaus, U. and Santiago, J. (2020) Structural basis for recognition of RALF peptides by LRX proteins during pollen tube growth. *Proc. Natl. Acad. Sci. USA*, **117**, 7494–7503.

Nürnberger, T., Brunner, F., Kemmerling, B. and Piater, L. (2004) Innate immunity in plants and animals: striking similarities and obvious differences. *Immunol. Rev.* **198**, 249–266.

de Oliveira Figueiroa, E., Albuquerque da Cunha, C.R., Albuquerque, P.B.S. *et al.* (2017) Lectin-carbohydrate interactions: implications for the development of new anticancer agents. *Curr. Med. Chem.* **24**, 3667–3680.

Otto, D.M.E., Campanero-Rhodes, M.A., Karamanska, R. *et al.* (2011) An expression system for screening of proteins for glycan and protein interactions. *Anal. Biochem.* **411**, 261–270.

Palma, A.S., Feizi, T., Zhang, Y. *et al.* (2006) Ligands for the β-glucan receptor, dectin-1, assigned using "designer" microarrays of oligosaccharide probes (neoglycolipids) generated from glucan polysaccharides. *J. Biol. Chem.* **281**, 5771–5779.

Peng, C., Zhu, Z., Shi, Y., Wang, X., Mu, K., Yang, Y., Zhang, X., Xu, Z. and Zhu, W. (2020) Exploring the binding mechanism and accessible angle of SARS-CoV-2 Spike and ACE2 by molecular dynamics simulation and free energy calculation. *Chemrxiv. Org*, **11877492**, v1.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.

Phillips, J.C., Braun, R., Wang, W. *et al.* (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802.

Pierleoni, A., Martelli, P. and Casadio, R. (2008) PredGPI: a GPI-anchor predictor. *BMC Bioinformatics*, **9**, 392.

Ranf, S., Gisch, N., Schäffer, M. *et al.* (2015) A lectin S-domain receptor kinase mediates lipopolysaccharide sensing in *Arabidopsis thaliana*. *Nat. Immunol.* **16**, 426–433.

Ranf, S., Grimmer, J., Pöschl, Y., Pecher, P., Chinchilla, D., Scheel, D. and Lee, J. (2012) Defense-related calcium signaling mutants uncovered via a quantitative high-throughput screen in *Arabidopsis thaliana*. *Mol. Plant*, **5**, 115–130.

Ronald, P.C. and Beutler, B. (2010) Plant and animal sensors of conserved microbial signatures. *Science*, **330**, 1061–1064.

Ruprecht, C., Bartetzko, M.P., Senf, D. *et al.* (2020) A glycan array-based assay for the identification and characterization of plant glycosyltransferases. *Angew. Chemie - Int. Ed.* **59**, 12493–12498.

Sadiq, S.K., Wright, D.W., Kenway, O.A. and Coveney, P.V. (2010) Accurate ensemble molecular dynamics binding free energy ranking of multidrug-resistant HIV-1 proteases. *J. Chem. Inf. Model.* **50**, 890–905.

Saijo, Y., Loo, E.-P.-I. and Yasuda, S. (2018) Pattern recognition receptors and signaling in plant–microbe interactions. *Plant J.* **93**, 592–613.

Sánchez-Vallet, A., Tian, H., Rodriguez-Moreno, L. *et al.* (2020) A secreted LysM effector protects fungal hyphae through chitin-dependent homodimer polymerization. *PLoS Pathog.* **16**, e1008652.

Sandoval, P.J. and Santiago, J. (2020) Update on techniques to probe ligand-receptor interactions in vitro analytical approaches to study plant ligand-receptor interactions 1[OPEN]. *Plant Physiol. Ò*, **182**, 1697.

Santiago, J., Brandt, B., Wildhagen, M., Hohmann, U., Hothorn, L.A., Butenko, M.A. and Hothorn, M. (2016) Mechanistic insight into a peptide hormone signaling complex mediating floral organ abscission. *Elife*, **5**, e15075.

Santiago, J., Henzler, C. and Hothorn, M. (2013) Molecular mechanism for plant steroid receptor activation by somatic embryogenesis co-receptor kinases. *Science*, **341**, 889–892.

Sapay, N., Nurisso, A. and Imberty, A. (2013) Simulation of carbohydrates, from molecular docking to dynamics in water. In *Biomolecular Simulations*. Totowa, NJ: Humana Press, pp. 469–483.

Saravanan, K., Hunday, G. and Kumaradhas, P. (2020) Binding and stability of indirubin-3-monoxime in the GSK3β enzyme: a molecular dynamics simulation and binding free energy study. *J. Biomol. Struct. Dyn.* **38**, 957–974.

Schallus, T., Jaeckh, C., Fehér, K. *et al.* (2008) Malectin: a novel carbohydrate-binding protein of the endoplasmic reticulum and a candidate player in the early steps of protein N-glycosylation. *Mol. Biol. Cell*, **19**, 3404–3414.

Shimizu, T., Nakano, T., Takamizawa, D. *et al.* (2010) Two LysM receptor molecules, CEBiP and OsCERK1, cooperatively regulate chitin elicitor signaling in rice. *Plant J.* **64**(2), 204–214. https://doi.org/10.1111/j.1365-313X.2010.04324.x

Shiu, S.H. and Bleecker, A.B. (2003) Expansion of the receptor-like kinase/Pelle gene family and receptor-like proteins in Arabidopsis. *Plant Physiol.* **132**, 530–543.

Shiu, S.H., Karlowski, W.M., Pan, R., Tzeng, Y.H., Mayer, K.F.X. and Li, W.H. (2004) Comparative analysis of the receptor-like kinase family in Arabidopsis and rice. *Plant Cell*, **16**, 1220–1234.

Singh, P., Kuo, Y.C., Mishra, S. *et al.* (2012) The lectin receptor Kinase-VI.2 is required for priming and positively regulates Arabidopsis pattern-triggered immunity. *Plant Cell*, **24**, 1256–1270.

Smakowska-Luzan, E., Mott, G.A., Parys, K. *et al.* (2018) An extracellular network of Arabidopsis leucine-rich repeat receptor kinases. *Nature*, **553**, 342–346.

Sonah, H., Deshmukh, R.K. and Bélanger, R.R. (2016) Computational prediction of effector proteins in fungi: opportunities and challenges. *Front. Plant Sci.* **7**, 126.

Srivastava, V., McKee, L.S. and Bulone, V. (2017) *Plant Cell Walls*. in eLs. Chichester, UK: John Wiley & Sons Ltd. https://doi.org/10.1002/9780470015902.a0001682.pub3

Stegmann, M., Monaghan, J., Smakowska-Luzan, E., Rovenich, H., Lehner, A., Holton, N., Belkhadir, Y. and Zipfel, C. (2017) The receptor kinase FER

is a RALF-regulated scaffold controlling plant immune signaling. *Science*, **355**, 287–289.

Sun, Y., Li, L., Macho, A.P., Han, Z., Hu, Z., Zipfel, C., Zhou, J.M. and Chai, J. (2013) Structural basis for flg22-induced activation of the Arabidopsis FLS2-BAK1 immune complex. *Science*, **342**, 624–628.

Tang, D., Wang, G. and Zhou, J.M. (2017) Receptor kinases in plant-pathogen interactions: More than pattern recognition. *Plant Cell*, **29**, 618–637.

Trott, O. and Olson, A.J. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**(2), 455–461.

Voxeur, A., Habrylo, O., Guénin, S. *et al.* (2019) Oligogalacturonide production upon *Arabidopsis thaliana-Botrytis cinerea* interaction. *Proc. Natl. Acad. Sci. USA*, **116**, 19743–19752.

Wan, J., Tanaka, K., Zhang, X.C., Son, G.H., Brechenmacher, L., Nguyen, T.H.N. and Stacey, G. (2012) LYK4, a lysin motif receptor-like kinase, is important for chitin signaling and plant innate immunity in Arabidopsis. *Plant Physiol.* **160**, 396–406.

Wang, C., Zhou, M., Zhang, X., Yao, J., Zhang, Y. and Mou, Z. (2017) A lectin receptor kinase as a potential sensor for extracellular nicotinamide adenine dinucleotide in *Arabidopsis thaliana*. *Elife*, **6**, e25474.

Wang, J. and Chai, J. (2020) Structural insights into the plant immune receptors PRRs and NLRs. *Plant Physiol.* **182**, 1566–1581.

Wang, J., Li, H., Han, Z., Zhang, H., Wang, T., Lin, G., Chang, J., Yang, W. and Chai, J. (2015) Allosteric receptor activation by the plant peptide hormone phytosulfokine. *Nature*, **525**, 265–268.

Wanke, A., Malisic, M., Wawra, S. and Zuccaro, A. (2020) Unraveling the sugar code: the role of microbial extracellular glycans in plant–microbe interactions. *J. Exp. Bot.*, https://doi.org/10.1093/jxb/eraa414.

Wanke, A., Rovenich, H., Schwanke, F., Velte, S., Becker, S., Hehemann, J.H., Wawra, S. and Zuccaro, A. (2020) Plant species-specific recognition of long and short β-1,3-linked glucans is mediated by different receptor systems. *Plant J.* **102**, 1142–1156.

Waterhouse, A., Bertoni, M., Bienert, S. *et al.* (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303.

Williams, C.J., Headd, J.J., Moriarty, N.W. *et al.* (2018) MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315.

Willmann, R., Lajunen, H.M., Erbs, G. *et al.* (2011) Arabidopsis lysin-motif proteins LYM1 LYM3 CERK1 mediate bacterial peptidoglycan sensing and immunity to bacterial infection. *Proc. Natl. Acad. Sci. USA*, **108**, 19824–19829.

Wolf, S. (2017) Plant cell wall signalling and receptor-like kinases. *Biochem. J.* **474**, 471–492.

Woods, R.J. and Tessier, M.B. (2010) Computational glycoscience: characterizing the spatial and temporal properties of glycans and glycan-protein complexes. *Curr. Opin. Struct. Biol.* **20**, 575–583.

Xue, D.X., Li, C.L., Xie, Z.P., Staehelin, C. and Napier, R. (2019) LYK4 is a component of a tripartite chitin receptor complex in *Arabidopsis thaliana*. *J. Exp. Bot.* **70**, 5507–5516.

Zang, H., Xie, S., Zhu, B., Yang, X., Gu, C., Hu, B., Gao, T., Chen, Y. and Gao, X. (2019) Mannan oligosaccharides trigger multiple defence responses in rice and tobacco as a novel danger-associated molecular pattern. *Mol Plant Pathol.* **20**(8), 1067–1079. https://doi.org/10.1111/mpp.12811

Zipfel, C. (2014) Plant pattern-recognition receptors. *Trends Immunol.* **35**, 345–351.