

# Role of substrate recognition in modulating strigolactone receptor selectivity in witchweed

Jiming Chen<sup>1</sup>, Alexandra White<sup>2</sup>, David C. Nelson<sup>2</sup>, and Diwakar Shukla<sup>\*1,3,4,5,6</sup>

<sup>1</sup>Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801

<sup>2</sup>Department of Botany and Plant Sciences, University of California, Riverside, Riverside, CA 92521

<sup>3</sup>Center for Biophysics and Quantitative Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801

<sup>4</sup>National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Urbana, IL 61801

<sup>5</sup>Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801

<sup>6</sup>NIH Center for Macromolecular Modeling and Bioinformatics, University of Illinois at Urbana-Champaign, Urbana, IL 61801

\*Email: diwakar@illinois.edu

Witchweed, or *Striga hermonthica*, is a parasitic weed that destroys billions of dollars worth of crops globally every year. Its germination is stimulated by strigolactones exuded by its host plants. Despite high sequence, structure, and ligand binding site conservation across different plant species, one strigolactone receptor in witchweed (*ShHTL7*) uniquely exhibits a picomolar EC<sub>50</sub> for downstream signaling. Previous biochemical and structural analyses have hypothesized that this unique ligand sensitivity can be attributed to a large binding pocket volume in *ShHTL7* resulting in enhanced ability to bind substrates. Additional structural details of the substrate binding process can help explain its role in modulating the ligand selectivity. Using long-timescale molecular dynamics simulations, we demonstrate that mutations at the entrance of the binding pocket facilitate a more direct ligand binding pathway to *ShHTL7*, whereas hydrophobicity at the binding pocket entrance results in a stable “anchored” state. We also demonstrate that several residues on the D-loop of *AtD14* stabilize catalytically inactive conformations. Finally, we show that strigolactone selectivity is not modulated by binding pocket volume. Our results indicate that while ligand binding is not the sole modulator of strigolactone receptor selectivity, it is a significant contributing factor. These results can be used to inform the design of selective antagonists for strigolactone receptors in witchweed.

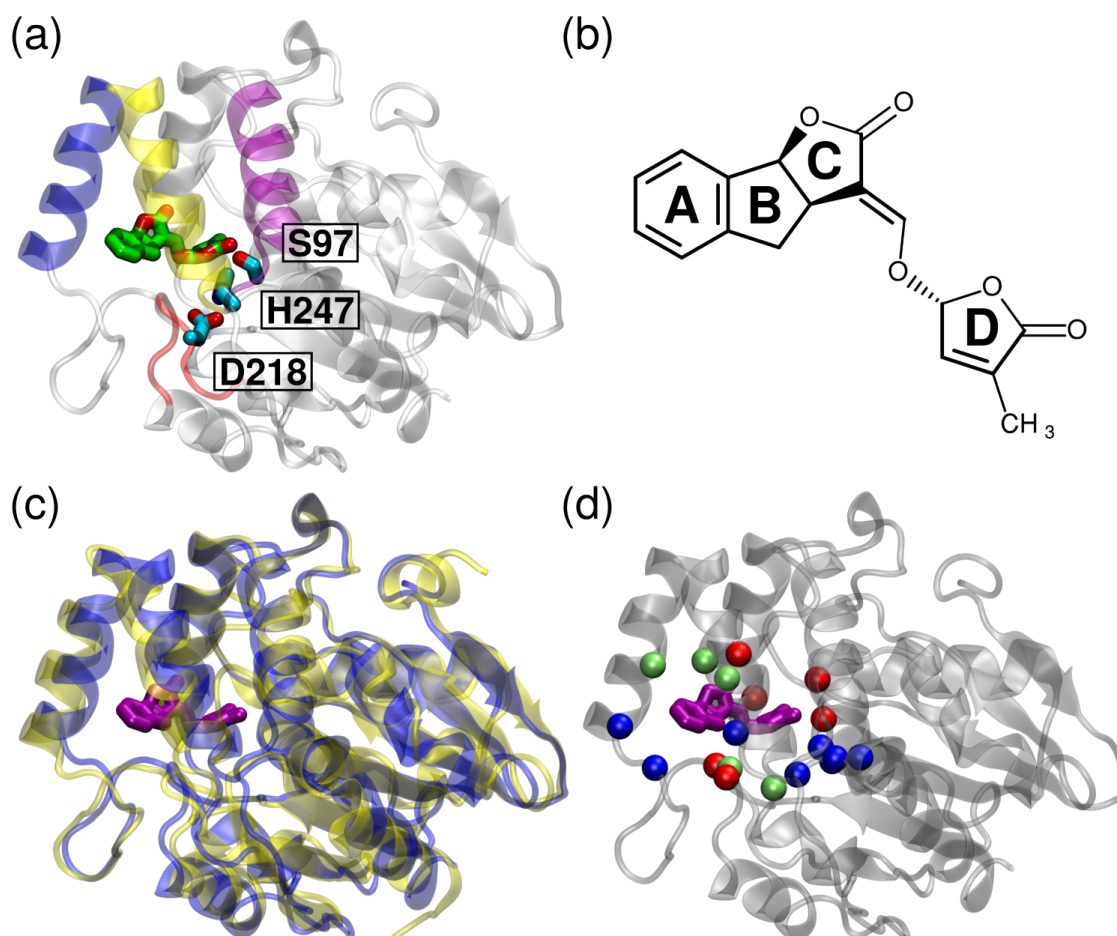
## 1 Introduction

Strigolactones are a class of plant hormones responsible for regulating shoot branching and root architecture in plants<sup>1-4</sup>. They have also been found to induce seed germination in the parasitic *Striga* genus<sup>5</sup>. Estimates of global crop losses due to *Striga* parasites are in excess of \$10 billion per year, warranting a need for effective *Striga* control<sup>6</sup>. Strigolactone perception is controlled by a family of proteins called DWARF14, which possess a conserved  $\alpha$ - $\beta$  hydrolase fold with a hydrophobic cavity in which the substrate binds. D14 and its closely related homolog KAI2, contain a strictly conserved Ser-His-Asp catalytic triad (Fig. 1a). Strigolactone signalling responses are believed to be dependent on enzymatic hydrolysis of the substrate and subsequent covalent modification of the enzyme by a hydrolysis product<sup>7-10</sup>. Following hydrolysis, the receptor undergoes a large conformational change that enables it to associate with MAX2 and SMXL proteins, which

are then ubiquitinated and degraded by the proteasome<sup>11</sup>. Recent evidence has also suggested that signal can be transduced by intact strigolactone molecules<sup>12</sup>, and that MAX2 proteins may act as a repressor of strigolactone hydrolysis<sup>13</sup>.

Despite 44% sequence identity between *Arabidopsis thaliana* and *Striga hermonthica* strigolactone receptors, 78% sequence similarity, and a highly conserved structure between different species (Fig. 1c), one receptor in *Striga hermonthica*, *ShHTL7*, uniquely exhibits a picomolar range EC50 for downstream signaling for inducing a germination response, compared to micromolar ranges for other strigolactone receptors<sup>14</sup>. An evolutionary analysis by Conn *et al.* revealed that *ShHTL* proteins evolved the ability to perceive strigolactone via convergent evolution<sup>15</sup>. *ShHTL* proteins are paralogs of KAI2 proteins, which perceive seed germination stimulants in plants and evolved strigolactone sensitivity independently of D14 proteins<sup>15</sup>. KAI2 proteins are generally grouped into three clades, the KAI2c (conserved) clade which is the most KAI2-like and has sensitivity to karrikins but not strigolactones, the KAI2i intermediate clade, and the divergent KAI2d clade that is strigolactone-sensitive but not karrikin-sensitive. Further studies have hypothesized that the high strigolactone sensitivity found in several members of the divergent clade of *ShHTL* proteins, which includes *ShHTL7*, can be attributed to their larger binding pocket volume compared to other members of the D14/KAI2 superfamily of proteins<sup>14,16</sup>. Additionally, an isothermal titration calorimetry and crystallography study by Burger *et al.* suggested that the T2-T3 loop of KAI2 proteins is able to modulate pocket size which in turn is able to influence binding affinity<sup>17</sup>. However, this hypothesis relies on pocket volumes computed from crystal structures, which can only provide static “snapshots” of the protein. In an aqueous environment, the pocket volume is likely to fluctuate due to conformational flexibility of the protein.

Alternatively, differences in the substrate binding mechanism can contribute to enhanced signaling ability in one protein over the other. Differences in the substrate binding process can enhance signaling in two ways: (i) A higher binding affinity for the ligand can increase the residence time of the ligand in the pocket, leading to increased probability of enzymatic hydrolysis occurring, or (ii) a lower free energy barrier of binding can enhance the rate of binding, thus enhancing the apparent rates of subsequent steps. Characterizing the role of these effects in producing the uniquely high sensitivity of *ShHTL7* requires a detailed structural and dynamical characterization of the binding process. While structures of the protein-ligand complexes could provide insights into differences in binding affinity, mechanistic details of the binding process can additionally determine the effects of sequence differences in residues outside the binding pocket on ligand binding. The only currently available crystal structure of a strigolactone-bound D14 protein is a structure of *OsD14*, the *Oryza sativa* ortholog of *AtD14* (~74% sequence identity), bound to GR24, a synthetic strigolactone analog<sup>18</sup>. There is uncertainty surrounding the accuracy of this as well as other strigolactone receptor crystal structures bound to various ligands due to low electron density of ligands within the binding pocket<sup>19</sup>, and there is an inherent lack of dynamical information contained in crystal structures. Making direct biophysical measurements on the binding process is also particularly challenging for strigolactone receptors since it is known to hydrolyze its ligand. This coupling of binding and hydrolysis makes it difficult to elucidate the effects of substrate binding on signaling independently of subsequent steps. A powerful method that can be used to characterize the strigolactone binding process is molecular dynamics (MD) simulations<sup>20-26</sup>. When used with Markov state models (MSMs), simulations can provide detailed kinetic and thermodynamic information about ligand binding processes at atomic-level resolution<sup>27-32</sup>. Furthermore, MSMs allow us to perform analysis on a large number of short simulations rather than a single long simulation<sup>33,34</sup>, which greatly decreases the time required to obtain sufficient data. MD simulations have previously been used to characterize other conformational dynamics and substrate binding in other plant



**Fig. 1.** (a) Structure of *Arabidopsis thaliana* strigolactone receptor *AtD14* in complex with strigolactone analog GR24 (green). The T1, T2, and T3 helices are shown in blue, yellow, and purple, respectively, and the D-loop is shown in red. The serine-histidine-aspartate catalytic triad is shown in cyan. (b) Structure of synthetic strigolactone analog GR24<sup>5DS</sup>, which possesses the stereochemistry of naturally occurring strigolactones. (c) Structural alignment of *AtD14* (blue) and *ShHTL7* (yellow). Bound GR24 is shown in purple. (d) Similarity of binding pocket residues. Blue spheres indicate conserved residues between *AtD14* and *ShHTL7*, red spheres indicate similar residues, and green spheres indicate different residues. Full sequence and secondary structure alignments are shown in Fig. S1.

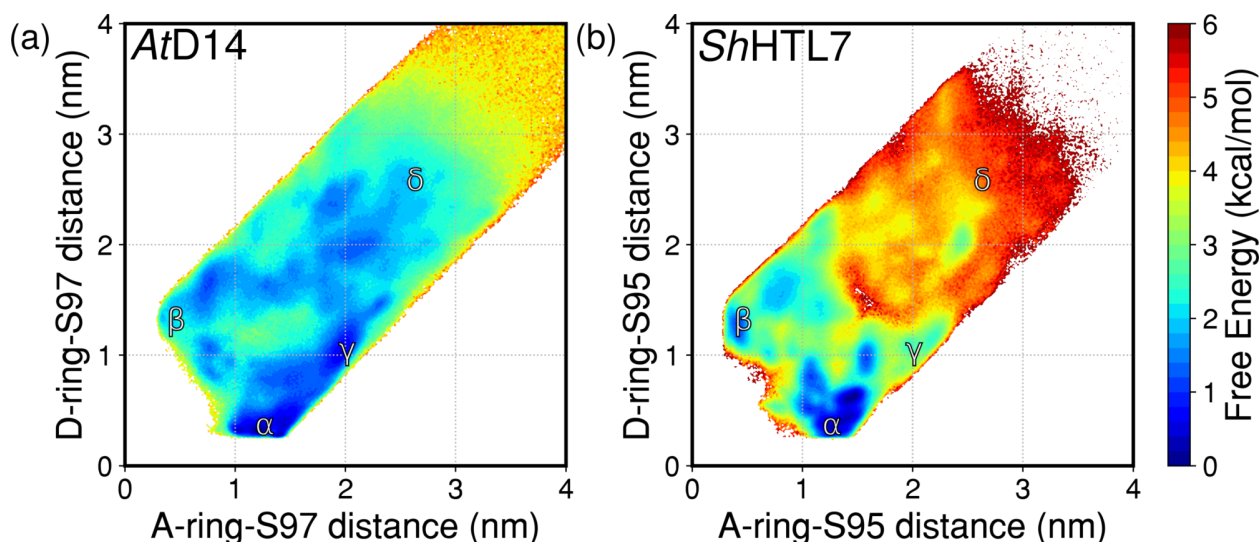
proteins<sup>20, 21, 24–26</sup>.

Recently, Hu *et al.* employed biased MD simulations to characterize the mechanism of the smoke-derived compound KAR<sub>1</sub> to *AtKAI2*<sup>26</sup>, a homolog of *AtD14* (~50% sequence identity). However, a limitation to this study is the biasing methods that were used have an inherent assumption that the ligand can only bind in a single binding pose and via a single pathway. Here, we employ long timescale (~400  $\mu$ s aggregate) unbiased MD simulations, allowing for a high resolution, dynamical view of the substrate recognition mechanisms in *AtD14* and *ShHTL7*. We demonstrate that *ShHTL7* is more efficient at binding GR24 and is also more effective at positioning GR24 for hydrolysis than *AtD14*. Additionally, we show that while differences in the ligand binding process do contribute to the high ligand sensitivity in *ShHTL7*, these differences are not caused by the

difference in the crystal structure pocket volume.

## 2 Results

**Free energy profile of the binding process.** Using  $\sim 200 \mu s$  aggregate of MD simulations each on GR24 binding to *AtD14* and *ShHTL7*, we computed the free energy landscapes of the complete ligand binding processes (Fig. 2). These landscapes were projected onto A-ring-catalytic serine distance and D-ring-catalytic serine distance for the purpose of distinguishing different binding modes of the ligand. Free energy minima discernable from these landscapes are the bound state ( $\alpha$ ), consistent with the crystal structure of GR24-bound *OsD14* (PDB 5DJ5)<sup>18</sup>; an inverse bound state ( $\beta$ ), and an “anchored” intermediate state ( $\gamma$ ). The most stable minimum for both proteins corresponds to the bound state with the butenolide ring of the ligand oriented into the binding pocket and close to S97/95 of the catalytic triad. Both *AtD14* and *ShHTL7* are also capable of binding GR24 in an inverse pose, in which the A-ring is oriented into the pocket and the butenolide ring (D-ring) is oriented toward the pocket entrance. The canonical model of strigolactone signaling involves a catalytic mechanism in which S97/95 nucleophilically attacks the ligand upon the D-ring<sup>7,8</sup>, indicating that this inverse-bound pose is likely catalytically inactive, and thus signaling incompetent.



**Fig. 2.** Free energy landscapes of GR24 binding to (a) *AtD14* and (b) *ShHTL7*. Labeled states are  $\alpha$ : Bound state,  $\beta$ : Inverse bound state,  $\gamma$ : Anchored state,  $\delta$ : Unbound states

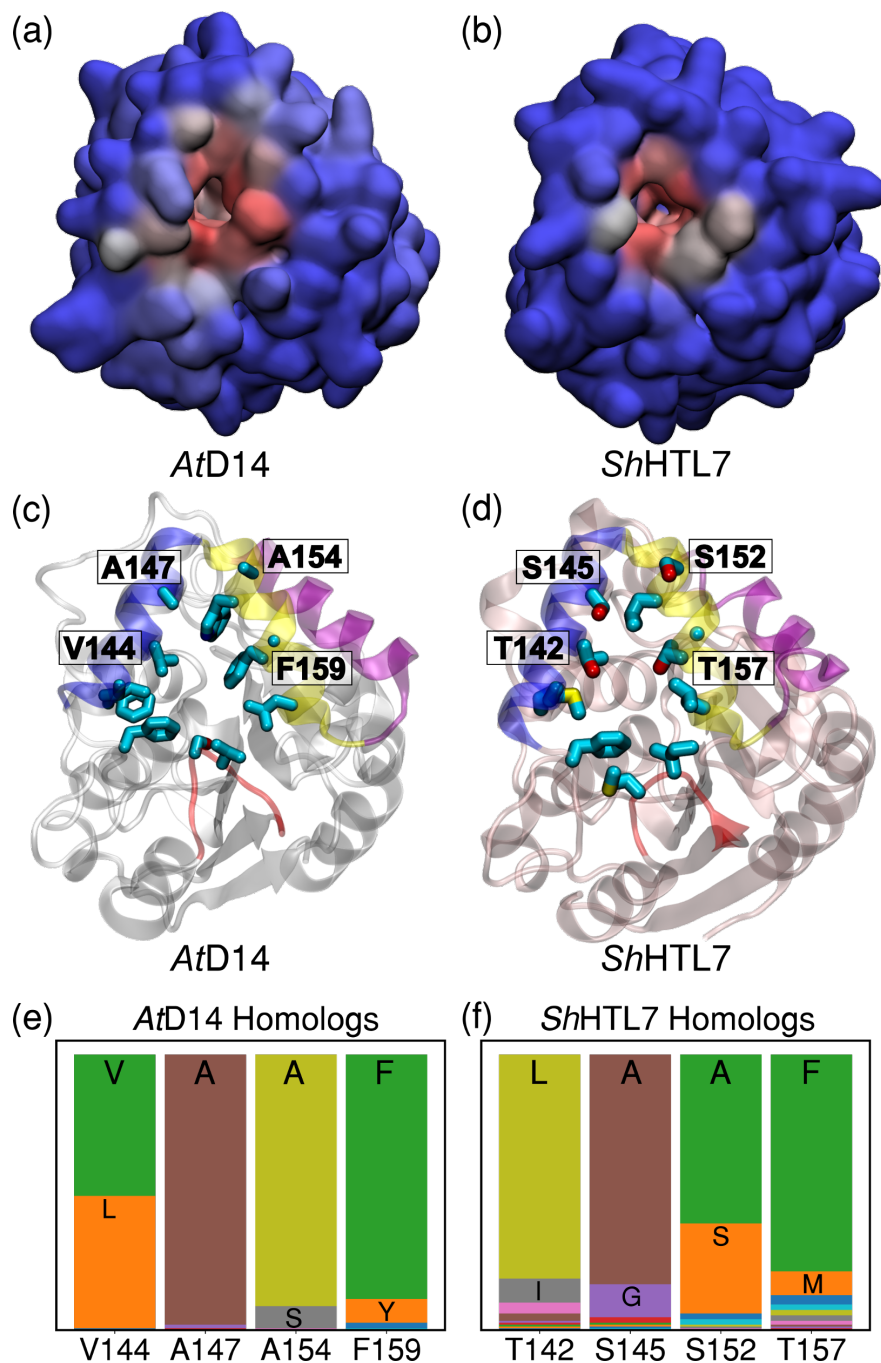
Using the method in Buch *et al.*<sup>27</sup>, we calculated the free energy for GR24 binding to be  $-5.5$  kcal/mol in *AtD14* and  $-5.7$  kcal/mol in *ShHTL7*. Free energy landscapes with respect to the slowest motions in the binding process are shown in Fig. S2. A previously reported dissociation constant ( $K_d$ ) for GR24 binding to *AtD14* based on an isothermal titration calorimetry measurement is  $0.30 \pm 0.02 \mu M$ , which corresponds to a free energy of  $-8.7$  kcal/mol at the experimental temperature of  $293$  K<sup>35</sup>. These free energy values were computed using the equation  $\Delta G = -RT \ln K_d$ , where  $T$  is the temperature at experimental conditions and  $K_d$  is the reported dissociation constant. However, this value is likely the free energy associated with both binding and hydrolysis because a significant positive entropy change ( $+19.5$  cal/mol\*K) upon binding is reported. Ligand binding is expected to have a negative entropy change associated with loss of configurational entropy of the ligand, but a hydrolysis reaction in which GR24 is split into its ABC-ring and D-ring would



more likely yield a positive entropy change. The  $K_d$  for GR24 binding to *ShHTL7* is estimated to be  $0.92 \pm 0.01 \mu\text{M}$  based on microscale thermophoresis assay and  $0.39 \pm 0.05 \mu\text{M}$  based on a tryptophan fluorescence assay, which correspond to a free energies of  $-8.7 \text{ kcal/mol}$  and  $-8.2 \text{ kcal/mol}$ , respectively, at  $298\text{K}$ <sup>36</sup>.

**Hydrophobic to polar mutations at pocket entrance destabilize anchored intermediate state in *ShHTL7*.** A notable difference in the binding pathways is the stability of the “anchored” intermediate state ( $\gamma$ ). Based on the free energy landscapes, the anchored intermediate state is  $\sim 1.5$  to  $2 \text{ kcal/mol}$  more stable in relation to the bound minimum in *AtD14* than in *ShHTL7*. This indicates that the ligand is more likely to interact with the pocket entrance during the binding process in *AtD14* than in *ShHTL7*. To further investigate the pocket entrance-anchoring observed in *AtD14*, we computed per-residue ligand contact probabilities for both *AtD14* and *ShHTL7* (Fig. 3a,b). In agreement with the free energy landscapes of the binding process, the regions of highest ligand contact probabilities in both proteins were the interior of the binding pocket. Additionally, the region directly outside the binding pocket shows considerably higher ligand contact probability in *AtD14* than in *ShHTL7*. This further indicates the presence of a stable interaction between the ligand and a patch of T1 and T2 helix residues outside the binding pocket. A comparison of residues in contact with the ligand in the *AtD14* anchored state and their corresponding residues in *ShHTL7* is shown in Fig. 3c,d. Four residues on the T1 and T2 helices are mutated from hydrophobic to polar residues between *AtD14* and *ShHTL7*: V144 (T142), A147 (S145), A154 (S152), and F159 (T157). Increased polarity at the pocket entrance prevents stable hydrophobic interactions from forming with the ABC rings of GR24. Since interactions between GR24 and the pocket entrance are largely hydrophobic, these mutations are likely to destabilize the anchored state in *ShHTL7*, thus leading to enhanced binding kinetics. This is consistent with the observation that a G158E mutant of *AtD14* displayed increased hydrolytic activity toward GR24 despite becoming signalling inactive<sup>7</sup>. Additionally, in a recent study introducing a femtomolar-range suicide germination compound for *Striga*, several mutations at the binding pocket entrance resulted in an increase in IC<sub>50</sub> for competitive binding of the compound to SPL7<sup>37</sup>. While this is not directly comparable since the measurements were done with a different ligand, it nonetheless supports the hypothesis that residues at the pocket entrance play an important role in ligand binding. Using the ConSurf server<sup>38</sup>, we also computed the site conservation of these four residues among homologs of *AtD14* and *ShHTL7*. Most frequent residues occupying the four pocket entrance sites among *AtD14* and *ShHTL7* are shown in Fig. 3e and f. The four sites show high conservation among both sets of homologs. However, while the pocket entrance residues in *AtD14* all match the most frequent residues of the given sites, the pocket entrance residues in *ShHTL7* are all less common substitutions. Notably, the most common residues at pocket entrance sites in *ShHTL7* homologs are hydrophobic, as in *AtD14*, indicating that polarity at the pocket entrance is not a common feature even among close homologs of *ShHTL7*. Parameters for the ConSurf calculation can be found in Table S1, and conservation scores and residue lists can be found in Table S2.

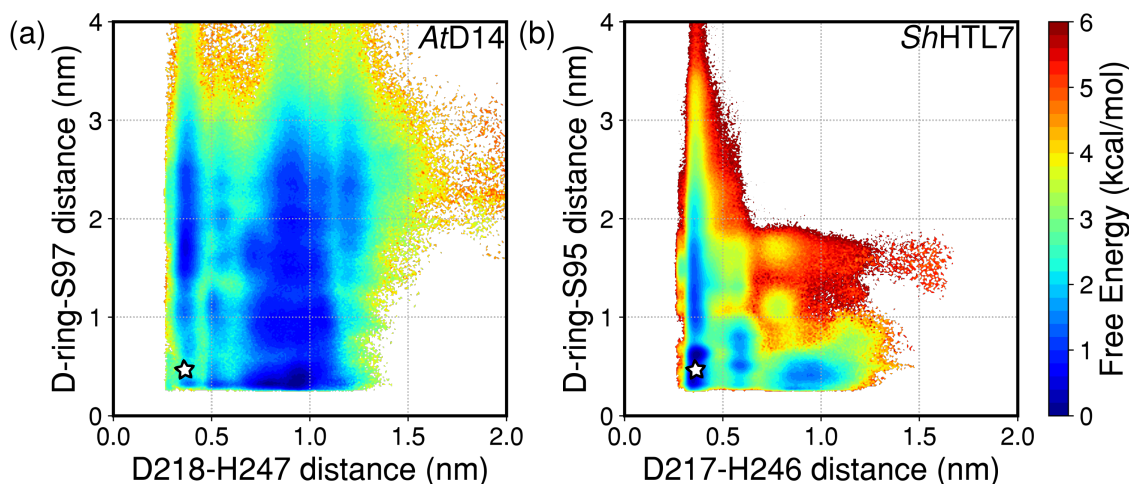
**Catalytically competent D-loop conformation more stable in *ShHTL7* than in *AtD14*.** Enzymatic hydrolysis of the substrate requires the D-loop of the protein to be in a D-in conformation, in which there is interaction between the aspartate (D218/217) and histidine (H247/246) of the catalytic triad. This is known due to previous mutagenesis experiments that have shown elimination of enzymatic activity upon mutation of any of the catalytic triad residues<sup>7,12,39,40</sup>. We computed free energy profiles of the binding process projected onto catalytic D-catalytic H distance and ligand-pocket distance (Fig. 4). The catalytically active state in which D218/217 is oriented into the binding pocket (D-in), substrate bound state is most stable in *ShHTL7*, however, *AtD14* exhibits highly stable conformations in which the D218 is oriented away from H247. The



**Fig. 3.** Residue ligand contact probabilities for (a) *AtD14* and (b) *ShHTL7*. Red regions indicate high ligand contact probability, and blue regions indicate low ligand contact probability. (c) *AtD14* pocket entrance residues in contact with the ligand in the anchored state and (d) corresponding residues in *ShHTL7*. Hydrophobic to polar substitutions are labeled. (e) Most frequently occurring residues in pocket entrance sites for *AtD14* homologs and (f) *ShHTL7* homologs. Residue lists for each site can be found in Table S2.

D-out conformations in *AtD14* are ~3-5 kcal/mol more stable than in *ShHTL7*, indicating the presence of stabilizing interactions that facilitate the formation of these catalytically incompetent

conformations. Upon comparison of the D-loop sequences in *AtD14* (AKDVSVPA) and *ShHTL7* (SNDIMVPV), we identified three mutations to differing residue types (i.e. hydrophobic to hydrophilic, charged to neutral): A216S, K217N, and S220M. Based on free energy profiles of key contacts involving these residues, we determine that each of these three mutations contribute to stabilization of the D-out conformation in *AtD14* (Fig. 5).



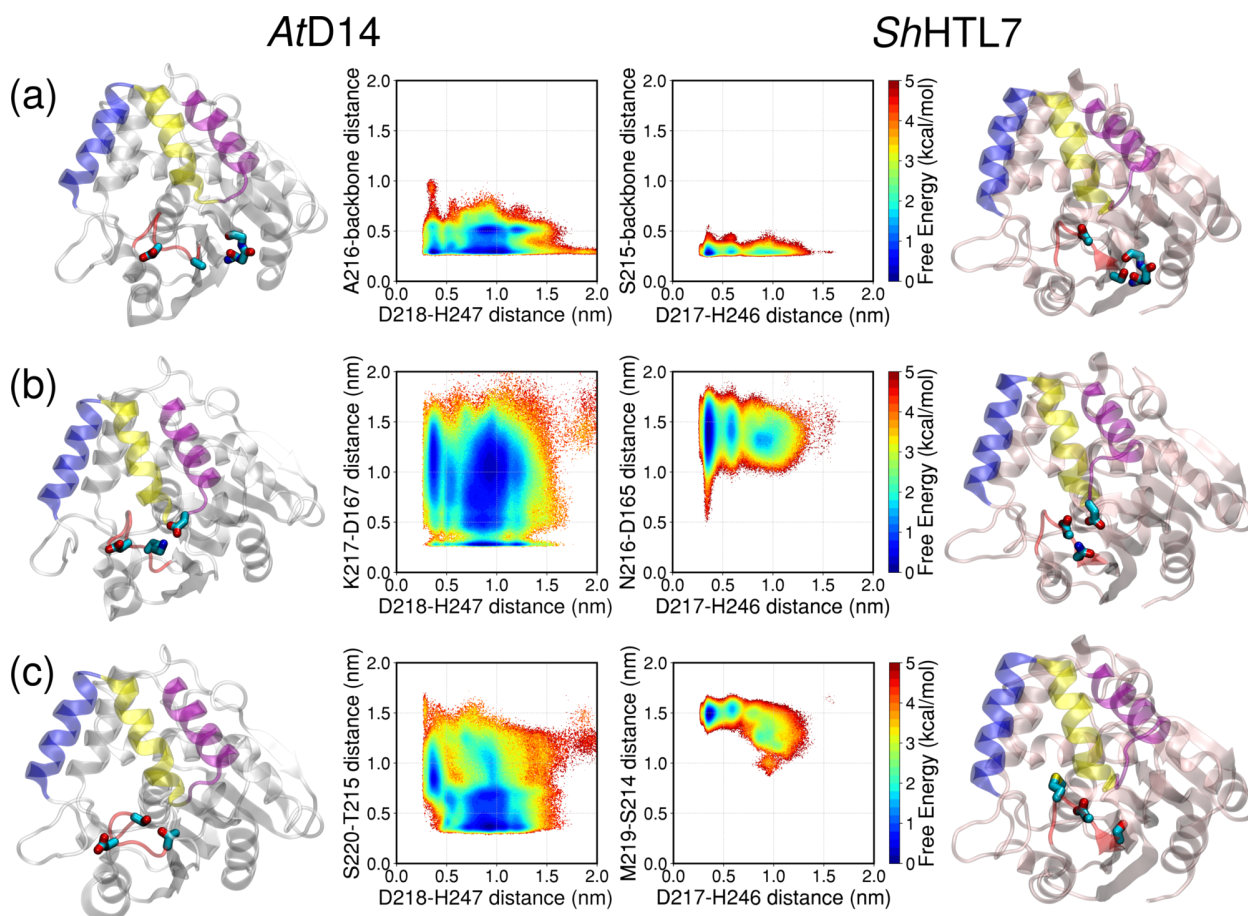
**Fig. 4.** Free energy landscapes of GR24 binding to (a) *AtD14* and (b) *ShHTL7* projected onto D218/217-H247/246 distance and ligand-S97/95 distance. The star demarcates the catalytically active state in which the ligand is bound and the aspartate and histidine of the catalytic triad are in contact.

The A216S mutation is located on the end of the D-loop closest to the T2 helix. The corresponding residue in *ShHTL7*, is S215, which is able to form hydrogen bonding interactions with the adjoining  $\beta$ -strand. This limits the range of motion of the D-loop. The A216 residue in *AtD14* is unable to form a hydrogen bond with the adjacent  $\beta$ -strand, allowing for increased D-loop motion. Additionally, the free energy landscape indicates that less stable D218-H247 interaction is observed in the absence of A216-backbone interaction (Fig. 5a). This implies that the interaction between S215 in *ShHTL7* and the adjacent backbone helps to stabilize the D-loop in the catalytically active D-in conformation.

The K217 residue in *AtD14* can form salt bridges with nearby negatively charged residues (D167, E244). In *ShHTL7*, the corresponding residue is N216, which eliminates a positive charge and prevents the formation of stable salt bridges. In particular, formation of the K217-E244 salt bridge in *AtD14* destabilizes the D218-H247 interaction, whereas absence of this salt bridge in *ShHTL7* allows for a stable D217-H246 interaction (Fig. 5b). In addition to D167, K217 can also form a salt bridge with E244. However, equally stable D218-H247 interactions are observed both in the presence and absence of the K217-E244 salt bridge, which indicates that this interaction does not destabilize the D218-H247 interaction (Fig. S3). In *ShHTL7*, N216 is able to form a hydrogen bond with E243. As observed in *AtD14*, the D217-H246 interaction remains intact both in the presence and absence of the N216-E243 hydrogen bond (Fig. S3).

Finally, S220 in *AtD14* can form a hydrogen bond with T215, which locks D218 in an outward-oriented position. The free energy landscape with respect to D218-H247 distance and S220-T215 distance indicates that D218-H247 contact is nearly eliminated in the presence of the S220-T215 hydrogen bond (Fig 5c). The corresponding residue to S220 in *ShHTL7* is M219, which is hy-

drophobic and thus unable to form a hydrogen bond with S214, the corresponding *Sh*HTL7 residue to T215 in *At*D14. In the absence of a M219-S214 hydrogen bond, a stable D217-H246 interaction is observed.

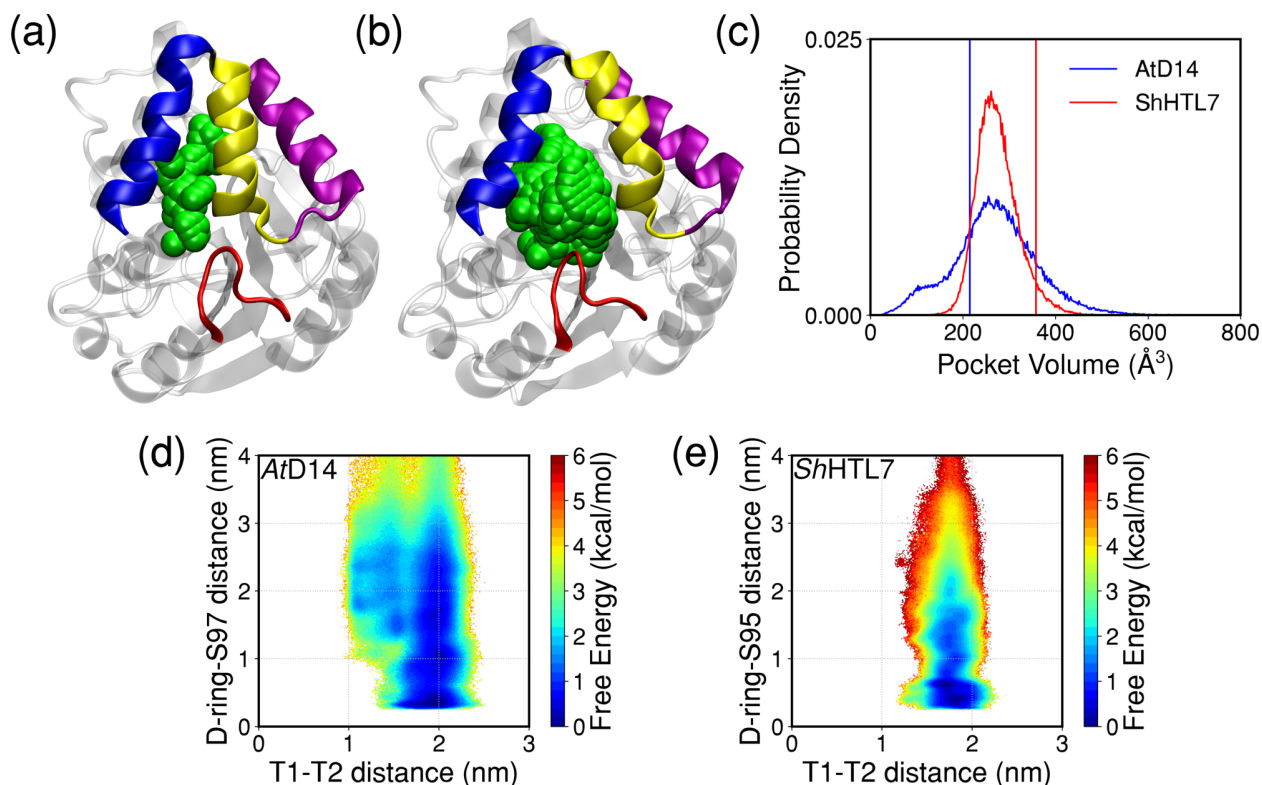


**Fig. 5.** Contacts stabilizing the D-out conformation relative to the D-in conformation in *At*D14. The D-in conformation is defined as having an interaction between D218/217 and H247/246, which means that the D-H distance is within  $\sim 0.5$  nm. (a) The S215-backbone interaction (S215-backbone distance  $< 0.5$  nm) in *Sh*HTL7 is lost in *At*D14. (b) The K217-D167 salt bridge in *At*D14 (K217-D167 distance  $< 0.5$  nm) is lost in *Sh*HTL7. (c) The S220-T215 hydrogen bond in *At*D14 (S220-T215 distance  $< 0.5$  nm) is lost in *Sh*HTL7.

**Large fluctuations in *At*D14 pocket volume facilitate binding-incapable states and non-productive binding.** Previous structural studies have hypothesized that a large binding pocket volume in *Sh*HTL7 is responsible for its uniquely high sensitivity to strigolactones<sup>14,16,41</sup>. To evaluate this hypothesis, we computed the probability distributions of pocket volumes in *At*D14 and *Sh*HTL7 over the course of our simulations (Fig. 6). The average pocket volumes of the two proteins are in close agreement with each other ( $\mu = 268 \text{ \AA}^3$  and  $274 \text{ \AA}^3$  for *At*D14 and *Sh*HTL7, respectively). However, *At*D14 displays a significantly broader distribution of binding pocket volumes ( $\sigma = 90 \text{ \AA}^3$  and  $45 \text{ \AA}^3$  for *At*D14 and *Sh*HTL7, respectively). Using the same pocket volume calculation metrics, we computed the pocket volumes of the *apo* crystal structures of *At*D14 and *Sh*HTL7 to be  $215 \text{ \AA}^3$  and  $358 \text{ \AA}^3$ , respectively. While the *Sh*HTL7 crystal structure does have a larger binding pocket volume than the *At*D14 crystal structure, this difference in pocket volume decreases significantly in



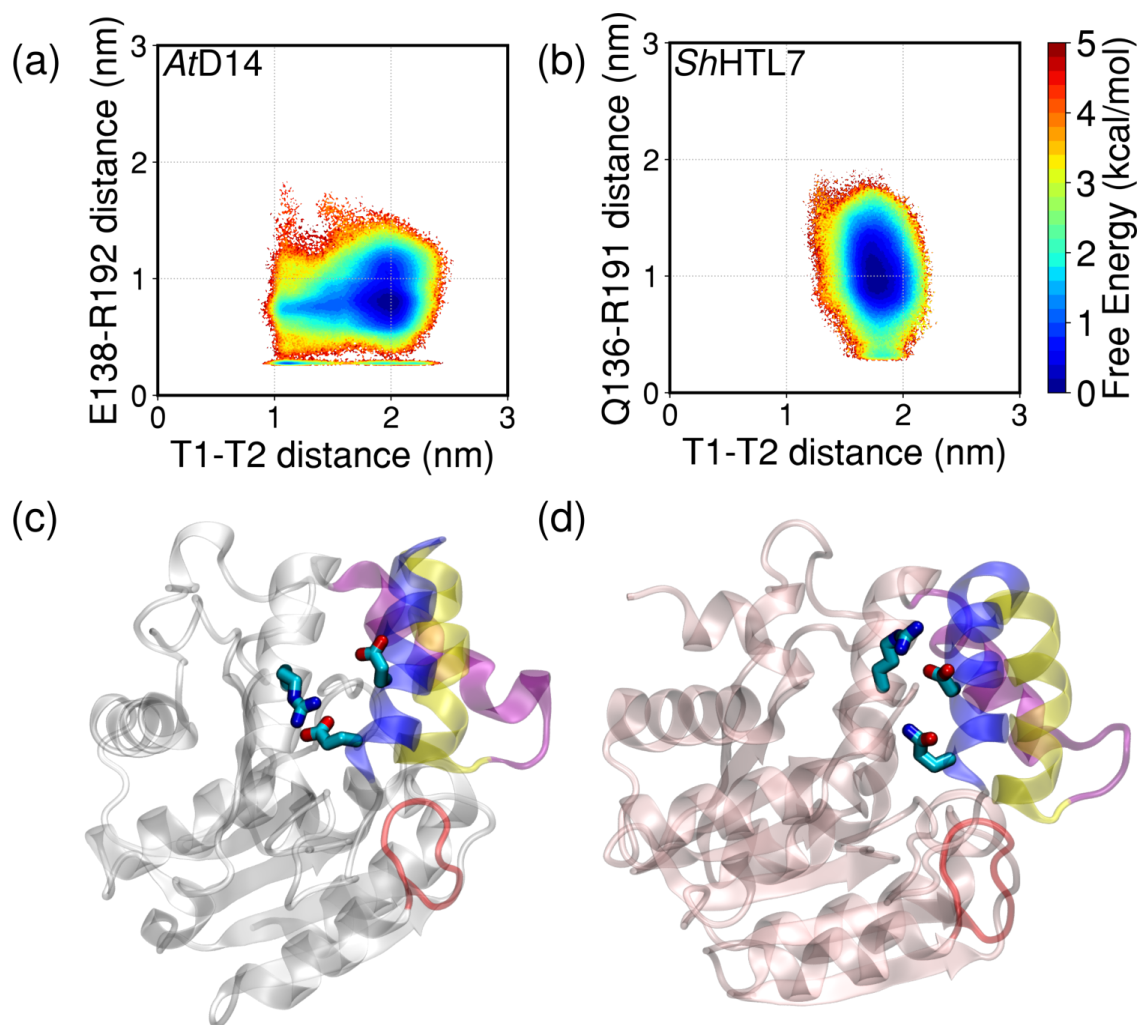
an aqueous environment. In both proteins, the primary modulator of binding pocket volume is a hinging motion between the T1 and T2 helices (Fig. S4). In the lowest-volume states, the binding pocket becomes solvent-inaccessible, rendering the protein unable to bind ligand. The highest-volume states allow for a large ensemble of ligand binding poses to form including many non-productive binding states in which the ligand is inside the pocket but not positioned for hydrolysis. These results indicate that the decreased tendency of *ShHTL7* to change its pocket volume may serve to increase its catalytic efficiency by retaining the pocket in a solvent-accessible conformation while also decreasing the stability of non-productive binding poses.



**Fig. 6.** (a) Representative low-volume structure of *AtD14* observed in simulations (b) Representative high-volume structure of *AtD14* observed in simulations. (c) Probability distributions of pocket volume for *AtD14* and *ShHTL7*. Vertical lines indicate the crystal structure pocket volumes. (d) Free energy landscape of ligand binding and T1-T2 distance for *AtD14*. (e) Free energy landscape of ligand binding and T1-T2 distance for *ShHTL7*.

The low-volume states accessible by *AtD14* are stabilized by hydrophobic interactions between the T1 and T2 helices, indicating that polarity in this region may play a dual role in modulating ligand selectivity. As previously stated, these residues enable a stable intermediate state to form, which acts as a barrier to binding. The low-volume states in which the pocket is nearly solvent inaccessible also show hydrophobic interactions between T1 and T2 helix residues, which indicates that these interactions could play a role in stabilizing low-volume states as well. In addition to the hydrophobic contacts between the T1 and T2 helices, a non-conserved salt bridge between the T1 and T4 helices in *AtD14* provides stabilization to the low-volume states as well (Fig. 7). Both *AtD14* and *ShHTL7* have a conserved arginine on the T4 helix (R192/191) that are able to form a salt bridge with E142/E140 on the T1 helix. However, *AtD14* also has a second negative residue, E138, on T1 that can form a salt bridge with R192. The free energy landscape of this interaction

and the T1-T2 distance indicates that in the presence of this E138-R192 salt bridge, low volume states are stabilized. This residue is mutated to Q136 in *ShHTL7*, so a salt bridge cannot form and stabilize low-volume states of the protein.



**Fig. 7.** Additional salt bridge in *AtD14* which stabilizes the low-volume states. (a) Free energy landscape showing stabilization of low-volume states in the presence of the E138-R192 salt bridge in *AtD14*, (b) *ShHTL7* free energy landscape showing little Q136-R191 interaction as well as lower population of low-volume states than in *AtD14*. (c) E138-R192 salt bridge in *AtD14*. E142, which can also form a salt bridge with R192, is also shown. (d) *ShHTL7* salt bridge between R191 and D140, the corresponding residue to E138 in *AtD14*. Q136, the corresponding residue to E138 in *AtD14*, is also shown.

### 3 Discussion

Using extensive MD simulations, we have characterized the mechanism of substrate binding to strigolactone receptors in full atomistic detail and identified several key differences in the binding mechanisms that contribute to the ligand selectivity between strigolactone receptors. Based on our simulations, GR24 binds to both *AtD14* and *ShHTL7* in the same binding pose as the reported crystal structure of *OsD14*-GR24 complex<sup>18</sup>. Additionally, since our simulations were performed

in an unbiased manner, we were also able to identify several non-productive bound states in which the ligand is bound but improperly positioned for hydrolysis.

In addition to characterizing the possible ligand poses within the receptor binding pockets, we identified a key anchored intermediate along the binding pathway that is  $\sim 1.5$  kcal/mol more stable in *AtD14* than in *ShHTL7*. This difference in stability indicates that the anchored state population relative to the bound state is  $\sim 12$  times higher in *AtD14* than in *ShHTL7*. This likely results in faster binding kinetics in *ShHTL7* than in *AtD14*, which in turn would lead to a higher observed catalytic turnover rate. Since hydrolysis is driver of receptor activation and downstream signaling, faster binding kinetics can contribute to enhanced signaling in *ShHTL7* compared to *AtD14*.

We also identified several key interactions involving residues on the D-loop of *AtD14* that stabilize the D-loop in a catalytically inactive, D218-out conformation. While these interactions do not preclude binding, they likely hinder the catalytic process in *AtD14* compared to in *ShHTL7* by stabilizing catalytically inactive states of the protein. Assuming that the hydrolysis is an inducer of receptor activation and downstream signaling, the stabilization of catalytically inactive conformations of the protein can lead to decreased signaling as well.

Finally, we evaluated the hypothesis that a larger binding pocket in *ShHTL7* enables its unique sensitivity to strigolactone. The average binding pocket volumes of *AtD14* and *ShHTL7* are nearly identical, however, *AtD14* is able to access more low-volume states which preclude ligand binding as well as more high-volume states which allow more non-productive binding to occur. These effects play a dual role in modulating ligand selectivity: *ShHTL7* is less likely to adopt low-volume states which are unable to bind ligand as well as high-volume states that are likely to bind ligand in non-productive, signaling-inactive poses.

A 10000-fold difference in EC<sub>50</sub> for downstream signaling implies a difference of  $\sim 5.5$ - $6$  kcal/mol along the strigolactone signaling process. In total, combined effects of enhanced binding and more stable catalytically active state contribute  $\sim 2$ - $3$  kcal/mol. This indicates that while the differences in the ligand binding process are contributors to the unique sensitivity of *ShHTL7*, subsequent steps in the signaling process, such as hydrolysis, activation, and association with signaling partners, play important roles in modulating this selectivity as well.

To our knowledge, this is the first in-depth characterization of the ligand binding process in strigolactone receptors. This study demonstrates the utility of molecular simulations approaches in providing mechanistic insights into fundamental questions in the field of plant biology<sup>42,43</sup>. Due to the importance of strigolactone signaling in crop productivity and parasitic weed germination, there is great interest in developing strigolactone signalling antagonists<sup>44-47</sup>. The factors we have identified that modulate ligand selectivity in strigolactone receptors can be used to inform the design of selective signalling agonists to enhance shoot branching in crops, induce suicidal germination in parasitic weeds, or prevent parasitic weed germination.

## 4 Methods

**Molecular dynamics simulations.** MD simulations were prepared using AmberTools 14/18 and run using Amber 14/18<sup>48</sup>. The protein was described using the ff14SB force field and water was described using the TIP3P model<sup>49</sup>. The GR24 ligand was described using the generalized AMBER force field (GAFF)<sup>50</sup>. Force field parameters for GR24 were generated using Antechamber. Initial structures for *AtD14* and *ShHTL7* were obtained from Protein Data Bank entries 4IH4<sup>51</sup> and 5Z7Y<sup>16</sup>, respectively. The GR24 substrate was superimposed into the binding pocket by structural

alignment of the bound structure of *OsD14* bound to GR24 (PDB 5DJ5)<sup>18</sup>. For *ShHTL7*, an additional system was prepared with the ligand randomly placed in solution using Packmol<sup>52</sup>. Each protein-ligand complex was solvated in a TIP3P water box of size  $\sim 70 \times 70 \times 70$  Å. NaCl was added at a concentration of 0.15M to neutralize the system. Each structure was minimized for 10000 steps using the conjugate gradient descent method and equilibrated for 10 ns. Production runs were performed for an aggregate of 207  $\mu$ s for *ShHTL7* and 198  $\mu$ s for *AtD14*. Temperatures were held constant at 300 K using the Berendsen thermostat, and pressures were held constant at 1.0 bar using the Berendsen barostat. Full electrostatics were computed using the Particle Mesh Ewald algorithm with a cutoff distance of 10 Å<sup>53</sup>. Bonds to hydrogen were constrained using the SHAKE algorithm<sup>54</sup>.

**Markov state model construction.** Markov state models (MSMs) were constructed using the PyEmma<sup>55</sup> package. Thirty-one input distance features were computed using MDTraj 1.9.0<sup>56</sup> (Table S3). The input distances were projected onto a reduced set of coordinates using time-lagged independent component analysis (TICA)<sup>57,58</sup>. The dimensionality-reduced coordinates were then clustered into states using the Mini-Batch K-Means algorithm prior to MSM estimation. The hyperparameters (number of TICA dimensions and number of clusters) were chosen via maximization of cross-validation scores (Fig. S6)<sup>59</sup>. Lag time was chosen by convergence implied timescales with respect to lag time (Fig. S5). Final parameters for MSM construction are listed in Table S4. Markovianity of the model was further validated using the Chapman-Kolmogorov test (Fig. S8). Free energy landscapes were calculated by computing the probability distribution along chosen sets of order parameters ( $x, y$ ) and weighting each point by the equilibrium probability of its associated MSM state (Eq. 1).

$$F(x, y) = -RT \ln[P_{\text{raw}}(x, y) * \pi_i(x, y)] \quad (1)$$

**Standard binding free energy calculation.** Standard free energies of binding were calculated using the volume correction method as detailed in Buch *et al.*<sup>27</sup>. Briefly, this method corrects for non-standard ligand concentration in the simulation by introducing a correction term (Eq. 2) that corresponds to the free energy of moving the ligand from a 1M solution to simulation conditions. For calculation of bound state volume, the bound state was defined as points within 1.0 nm and 4.0 kcal/mol of the minimum free energy point on the 3-dimensional MSM-weighted free energy landscape projected onto ligand position. Convergence of  $\Delta G_0$  with respect to bound state definition is shown in Fig. S9.

**Binding pocket volume calculation.** Binding pocket volumes were calculated using the POVME 2.0 package<sup>60</sup>. For each protein, a “maximum englobing region” was defined as a sphere centered at the midpoint between the geometric center of the T1 and T2 helix C- $\alpha$  atoms (residues 138-165 for *AtD14*, residues 136 to 163 for *ShHTL7*) and the C- $\alpha$  atom of the catalytic serine (Fig. S10). The radius of the maximum englobing sphere was set as the distance between the center and the C- $\alpha$  of the catalytic serine. The probability distribution of binding pocket volumes was weighted by Markov state model equilibrium probability. The average and standard deviation of the pocket volumes were calculated using the MSM-weighted probability distributions.

$$\Delta G_0 = -RT \ln \frac{V_b}{V_0} - \Delta W \quad (2)$$

**Residue-ligand contact probability calculation.** Residue-ligand distances for each residue were computed using MDTraj 1.9.0<sup>56</sup>. Contacts were defined as residue-ligand distances within a cutoff distance of 4.0 Å. The equilibrium contact probability was calculated as the product of raw



contact probability within each MSM state multiplied by the equilibrium probability of the MSM state as shown in Eq. 3:

$$P_{\text{contact, eq}} = \sum_i^{N_{\text{states}}} P_{\text{contact, raw}}|_i * \pi_i \quad (3)$$

## 5 Acknowledgements

This research was supported by the Blue Waters sustained-petascale computing project, which is funded by the National Science Foundation (OCI-0725070 and ACI-1238993) and the state of Illinois. J.C. acknowledges support from the Samuel W. Parr Graduate Fellowship (Department of Chemical and Biomolecular Engineering, University of Illinois) and the National Institutes of Health Chemistry-Biology Interface Training Grant (T32-GM070421). D.S. acknowledges support from the Center for Advance Study at University of Illinois at Urbana-Champaign.

## References

1. Umehara, M. *et al.* Inhibition of shoot branching by new terpenoid plant hormones. *Nature* **455**, 195–200 (2008).
2. Gomez-Roldan, V. *et al.* Strigolactone inhibition of shoot branching. *Nature* **455**, 189–194 (2008).
3. Waters, M. T., Gutjahr, C., Bennett, T. & Nelson, D. C. Strigolactone signaling and evolution. *Annual Review of Plant Biology* **68**, 291–322 (2017).
4. Lumba, S., Holbrook-Smith, D. & McCourt, P. The perception of strigolactones in vascular plants. *Nature Chemical Biology* **13**, 599–606 (2017).
5. Cook, C. E., Whichard, L. P., Turner, B., Wall, M. E. & Egle, G. H. Germination of witchweed (*striga lutea* Lour.): Isolation and properties of a potent stimulant. *Science* **154**, 1189–1190 (1966).
6. Ejeta, G. & Gressel, J. *Integrating New Technologies for Striga Control* (WORLD SCIENTIFIC, 2007).
7. Yao, R. *et al.* Dwarf14 is a non-canonical hormone receptor for strigolactone. *Nature* **536**, 469473 (2016).
8. de Saint Germain, A. *et al.* An histidine covalent receptor and butenolide complex mediates strigolactone perception. *Nature Chemical Biology* **12**, 787–794 (2016).
9. Yao, R. *et al.* ShHTL7 is a non-canonical receptor for strigolactones in root parasitic weeds. *Cell Research* **27**, 838–841 (2017).
10. Yao, R. *et al.* Rice DWARF14 acts as an unconventional hormone receptor for strigolactone. *Journal of Experimental Botany* **69**, 2355–2365 (2018).
11. Jiang, L. *et al.* DWARF 53 acts as a repressor of strigolactone signalling in rice. *Nature* **504**, 401–405 (2013).
12. Seto, Y. *et al.* Strigolactone perception and deactivation by a hydrolase receptor DWARF14. *Nature Communications* **10** (2019).

13. Shabek, N. *et al.* Structural plasticity of d3–d14 ubiquitin ligase in strigolactone signalling. *Nature* **563**, 652–656 (2018).
14. Toh, S. *et al.* Structure-function analysis identifies highly sensitive strigolactone receptors in striga. *Science* **350**, 203–207 (2015).
15. Conn, C. E. *et al.* Convergent evolution of strigolactone perception enabled host detection in parasitic plants. *Science* **349**, 540–543 (2015).
16. Xu, Y. *et al.* Structural analysis of HTL and d14 proteins reveals the basis for ligand selectivity in striga. *Nature Communications* **9** (2018).
17. Bürger, M. *et al.* Structural basis of karrikin and non-natural strigolactone perception in physcomitrella patens. *Cell Reports* **26**, 855–865.e5 (2019). URL <https://doi.org/10.1016/j.celrep.2019.01.003>.
18. Zhao, L.-H. *et al.* Destabilization of strigolactone receptor DWARF14 by binding of ligand and e3-ligase signaling effector DWARF3. *Cell Research* **25**, 1219–1236 (2015).
19. Carlsson, G. H., Hasse, D., Cardinale, F., Prandi, C. & Andersson, I. The elusive ligand complexes of the DWARF14 strigolactone receptor. *Journal of Experimental Botany* **69**, 2345–2354 (2018).
20. Moffett, A. S., Bender, K. W., Huber, S. C. & Shukla, D. Molecular dynamics simulations reveal the conformational dynamics of Arabidopsis thaliana BRI1 and BAK1 receptor-like kinases. *Journal of Biological Chemistry* **292**, 12643–12652 (2017).
21. Moffett, A. S., Bender, K. W., Huber, S. C. & Shukla, D. Allosteric control of a plant receptor kinase through s-glutathionylation. *Biophysical Journal* **113**, 2354–2363 (2017).
22. Aldukhi, F., Deb, A., Zhao, C., Moffett, A. S. & Shukla, D. Molecular mechanism of brassinosteroid perception by the plant growth receptor BRI1. *The Journal of Physical Chemistry B* **124**, 355–365 (2019). URL <https://doi.org/10.1021/acs.jpcc.9b09377>.
23. Lawrenz, M., Shukla, D. & Pande, V. S. Cloud computing approaches for prediction of ligand binding poses and pathways. *Scientific reports* **5**, 7918 (2015).
24. Shukla, S., Zhao, C. & Shukla, D. Dewetting controls plant hormone perception and initiation of drought resistance signaling. *Structure* **27**, 692–702.e3 (2019).
25. Selvam, B., Yu, Y.-C., Chen, L.-Q. & Shukla, D. Molecular basis of the glucose transport mechanism in plants. *ACS Central Science* (2019).
26. Hu, F. *et al.* MD simulation investigation on the binding process of smoke-derived germination stimulants to its receptor. *Journal of Chemical Information and Modeling* **59**, 1554–1562 (2019).
27. Buch, I., Giorgino, T. & Fabritiis, G. D. Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations. *Proceedings of the National Academy of Sciences* **108**, 10184–10189 (2011).
28. Shan, Y. *et al.* How does a drug molecule find its target binding site? *Journal of the American Chemical Society* **133**, 9181–9183 (2011).

29. Selvam, B., Shamsi, Z. & Shukla, D. Universality of the sodium ion binding mechanism in class a g-protein-coupled receptors. *Angewandte Chemie International Edition* **57**, 3048–3053 (2018).
30. Shukla, D., Hernández, C. X., Weber, J. K. & Pande, V. S. Markov state models provide insights into dynamic modulation of protein function. *Accounts of chemical research* **48**, 414–422 (2015).
31. Shamsi, Z., Cheng, K. J. & Shukla, D. Reinforcement learning based adaptive sampling: Reaping rewards by exploring protein conformational landscapes. *The Journal of Physical Chemistry B* **122**, 8386–8395 (2018).
32. Shamsi, Z., Moffett, A. S. & Shukla, D. Enhanced unbiased sampling of protein dynamics using evolutionary coupling information. *Scientific reports* **7**, 1–13 (2017).
33. Prinz, J.-H. *et al.* Markov models of molecular kinetics: Generation and validation. *The Journal of Chemical Physics* **134**, 174105 (2011).
34. Husic, B. E. & Pande, V. S. Markov state models: From an art to a science. *Journal of the American Chemical Society* **140**, 2386–2396 (2018).
35. Kagiya, M. *et al.* Structures of d14 and d14l in the strigolactone and karrikin signaling pathways. *Genes to Cells* **18**, 147–160 (2013).
36. Hameed, U. S. *et al.* Structural basis for specific inhibition of the highly sensitive ShHTL7 receptor. *EMBO reports* e45619 (2018).
37. Uraguchi, D. *et al.* A femtomolar-range suicide germination stimulant for the parasitic plant striga hermonthica. *Science* **362**, 1301–1305 (2018).
38. Ashkenazy, H. *et al.* ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Research* **44**, W344–W350 (2016). URL <https://doi.org/10.1093/nar/gkw408>.
39. Nakamura, H. *et al.* Molecular mechanism of strigolactone perception by DWARF14. *Nature Communications* **4** (2013).
40. Abe, S. *et al.* Carlactone is converted to carlactonoic acid by MAX1 in Arabidopsis and its methyl ester can directly interact with AtD14 in vitro. *Proceedings of the National Academy of Sciences* **111**, 18084–18089 (2014).
41. Xu, Y. *et al.* Structural basis of unique ligand specificity of KAI2-like protein from parasitic weed striga hermonthica. *Scientific Reports* **6** (2016).
42. Feng, J., Chen, J., Selvam, B. & Shukla, D. Computational microscopy: Revealing molecular mechanisms in plants using molecular dynamics simulations. *The Plant Cell* **31**, tpc.119.tt1219 (2019).
43. Moffett, A. S. & Shukla, D. Using molecular simulation to explore the nanoscale dynamics of the plant kinome. *Biochemical Journal* **475**, 905–921 (2018).
44. Holbrook-Smith, D., Toh, S., Tsuchiya, Y. & McCourt, P. Small-molecule antagonists of germination of the parasitic plant striga hermonthica. *Nature Chemical Biology* **12**, 724–729 (2016).

45. Xiang, H. *et al.* Simple  $\beta$ -lactones are potent irreversible antagonists for strigolactone receptors. *Cell Research* **27**, 1525–1528 (2017).
46. Takeuchi, J. *et al.* Rationally designed strigolactone analogs as antagonists of the d14 receptor. *Plant and Cell Physiology* (2018).
47. Nakamura, H. *et al.* Triazole ureas covalently bind to strigolactone receptor and antagonize strigolactone responses. *Molecular Plant* **12**, 44–58 (2019).
48. Case, D. *et al.* AMBER2018. University of California, San Francisco (2018).
49. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **79**, 926–935 (1983).
50. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *Journal of Computational Chemistry* **25**, 1157–1174 (2004).
51. Zhao, L.-H. *et al.* Crystal structures of two phytohormone signal-transducing  $\alpha/\beta$  hydrolases: karrikin-signaling KAI2 and strigolactone-signaling DWARF14. *Cell Research* **23**, 436–439 (2013).
52. Martínez, L., Andrade, R., Birgin, E. G. & Martínez, J. M. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *Journal of Computational Chemistry* **30**, 2157–2164 (2009).
53. Darden, T., York, D. & Pedersen, L. Particle mesh ewald: An  $n \cdot \log(n)$  method for ewald sums in large systems. *The Journal of Chemical Physics* **98**, 10089–10092 (1993).
54. Ryckaert, J.-P., Ciccotti, G. & Berendsen, H. J. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics* **23**, 327–341 (1977).
55. Scherer, M. K. *et al.* PyEMMA 2: A software package for estimation, validation, and analysis of markov models. *Journal of Chemical Theory and Computation* **11**, 5525–5542 (2015).
56. McGibbon, R. T. *et al.* MDTraj: A modern open library for the analysis of molecular dynamics trajectories. *Biophysical Journal* **109**, 1528–1532 (2015).
57. Schwantes, C. R. & Pande, V. S. Improvements in markov state model construction reveal many non-native interactions in the folding of ntl9. *Journal of chemical theory and computation* **9**, 2000–2009 (2013).
58. Schwantes, C. R., Shukla, D. & Pande, V. S. Markov state models and tica reveal a nonnative folding nucleus in simulations of nug2. *Biophysical journal* **110**, 1716–1719 (2016).
59. McGibbon, R. T. & Pande, V. S. Variational cross-validation of slow dynamical modes in molecular kinetics. *The Journal of Chemical Physics* **142**, 124105 (2015).
60. Durrant, J. D., Votapka, L., Sørensen, J. & Amaro, R. E. POVME 2.0: An enhanced tool for determining pocket shape and volume characteristics. *Journal of Chemical Theory and Computation* **10**, 5047–5056 (2014).