

# Computational investigation of drought-resistance in plants

## Blue Waters Progress Report, 2016

Diwakar Shukla

Blue Waters Assistant Professor, Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign. Email: diwakar@illinois.edu

### Executive Summary

With the 240,000 hours allotted to our group starting in Feb. 2015 and the extension of the allocation (additional 240,000 hours) in Sep. 2015, we have worked on problems that range from applications of high performance computing on Blue Waters to investigate conformational dynamics of plant receptor-like kinases involved in Brassinosteroid signaling to a methodological study utilizing evolutionary information to guide protein simulations. This allocation has led to two publications and several manuscript submissions are planned in near future. Given our productive usage of the computing time originally allocated to us, we believe that time allotted to our group will be put to good use and will allow us to continue to do the science uniquely enabled by the computational power of Blue Waters. We request an allocation of 290,000 node hours for the current allocation period (Mar 2016-Feb 2017) in order to continue our research work. Below, we have provided project updates and publications directly resulting from our work on Blue Waters.

### Challenge.

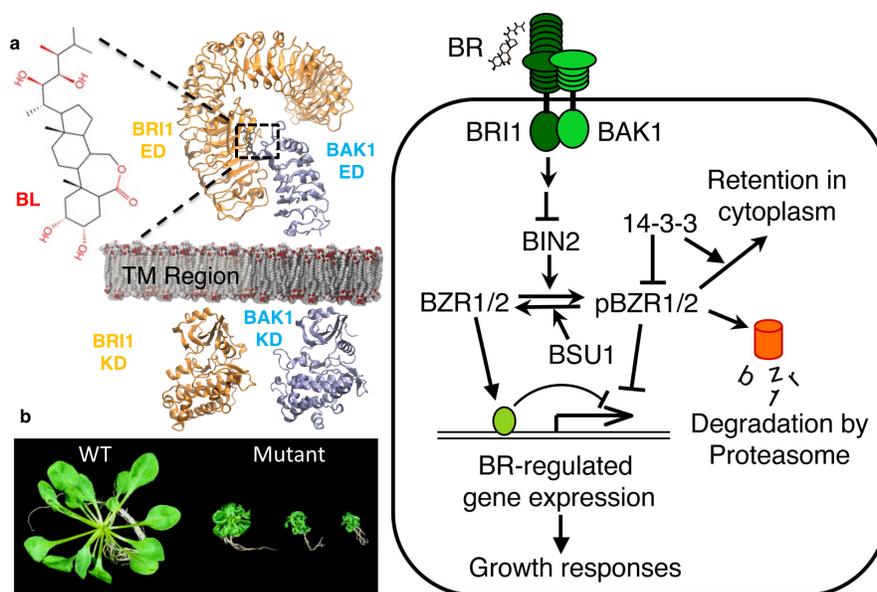


Figure 1: Mechanistic understanding of Brassinosteroid signaling process could lead to enhanced plant productivity under stress. a) Crystal structures of the extracellular domains (ED) bound to brassinolide (BL) (RCSB ID: 4M7E)[1] and the cytoplasmic kinase domains (KD) of *AtBRI1* (RCSB ID: 4OA2)[2] and *AtBAK1* (RCSB ID: 3TL8)[3] receptor-like kinases. b) *Arabidopsis thaliana* wild-type (WT) and dwarf plants from the brassinosteroid deficient mutants.[4] c) Role of BRI1-BAK1 kinases in the plant growth and development signaling.

Increase in demand for our primary foodstuffs is outstripping increase in yields, an expanding gap that indicates large potential food shortages by mid-century.[5, 6] This comes at a time when yield improvements are slowing or stagnating as the approaches of the Green Revolution reach their biological limits.[7] **With the threat of global climate change and frequent occurrence of extreme weather events such as droughts, the task of producing sufficient food and biofuels is expected to become even more challenging in future.**[8] Plant respond to changing environmental conditions by translating extracellular signals (typically in the form of small molecules such as hormones) into appropriate intracellular responses.

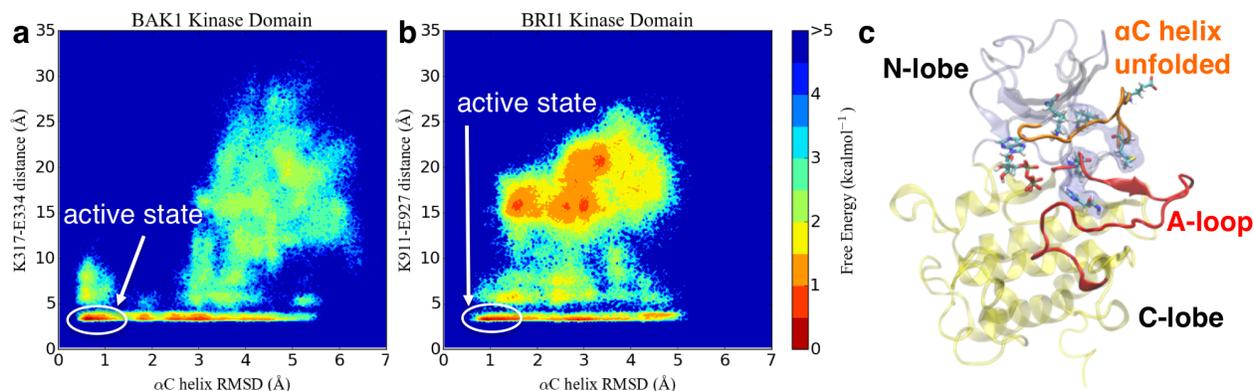


Figure 2: Free energy landscape of fully phosphorylated, ATP-bound kinase domain as a function of  $\alpha$ C helix RMSD from the crystal structures and K317 side-chain amine nitrogen to E334 side-chain carboxyl carbon distance for a) BAK1 b) BRI1 kinase. c) Structure of BRI1 kinase domain with a  $\alpha$ C-helix in unfolded state. The active state (available crystal structure pose) is also encircled. (Unpublished results)

Cell-surface receptor-like kinases (RLKs) play a major role in extracellular sensing and transmitting the information into the cytoplasm for downstream signaling.[9, 10] In plants, the receptor like kinases play key roles in regulating growth and development, protection against pathogens, and reproductive success in generating seeds and fruits and hindering premature abscission.[11, 12, 13] **Therefore, a quantitative molecular-level understanding of these plant signaling processes is fundamental to the future food and energy security.** In this proposed work, we focus on one of the best-characterized Leucine-rich repeat (LRR) RLK in plants, BRASSINOSTEROID INSENSITIVE 1 (BRI1), the receptor for plant steroid hormones called Brassinosteroids, which are crucial for the cell growth and plant morphogenesis.[14, 15]

## Approach.

Molecular simulations and theoretical models have been used to obtain atomistic insights into underlying dynamics and structural ensembles of key signaling proteins (e.g., G-Protein Coupled Receptors and Kinases) related to human health.[16, 17, 18] **However, these methods have not been widely used for investigating plant signaling proteins.** In brief, Molecular simulations would be used to obtain thermodynamic (stability of a particular conformation of protein or a protein-protein complex) and kinetic information (rate of interconversion among different conformations) about the protein. Such detailed structural and dynamic information could be then used to obtain mechanistic insights and complement experimental observations. Finally, the concept of energy landscapes could be used as a framework to connect protein conformational dynamics and experimental information on BRI1 activation. In this project, we plan to perform computational investigations of Brassinosteroid signaling from hormone reception to regulation of the BR activity. Specifically, we plan to build kinetic models of the conformational dynamics of BRI1 and BAK1 kinases under different conditions such as presence or absence of ligand, different post-translational modifications. **Such large transitions are difficult to observe by running simulations on small computers. Powerful resources like Blue Waters are required to study such complex biological process. Blue Waters provides thousands of Graphical Processing Units (GPU) that are utilized for parallel molecular dynamics simulations to perform Markov state model based adaptive sampling of conformational energy landscape of proteins. Use of Blue Waters increases the overall compute performance by several orders of magnitude (in terms of the real time required for simulation) and is therefore suitable for our study.**

## Accomplishments.

**Conformational dynamics of fully phosphorylated BRI1 and BAK1 kinase domain.** We have recently performed extensive MD simulations of BAK1 and BRI1 kinase domain starting from the available crystal structures using Blue Waters Supercomputer. The simulation data was used to build Markov state models (MSMs) of kinase dynamics by clustering the structurally similar conformations into states and

obtaining the interconversion rates between these states from the simulated trajectories.[16, 19] These preliminary kinetic models have provided new insights into the regulation of BAK1 and BRI1 kinase function. The crystal structures of the fully phosphorylated BAK1 and BRI1 kinase domain show a folded  $\alpha$ C-helix conformation. However, simulations reveal completely unfolded conformations of the  $\alpha$ C-helix. The folded conformation of the  $\alpha$ C-helix is a critical feature of the active state as the catalytically important K317-E334 h-bond is disrupted in the unfolded state.[20, 21] The free energy diagram (Fig. 2a-b; low value of free energy (red color regions) indicates more stable states) shows regions with an unfolded  $\alpha$ C-helix (high  $\alpha$ C root mean square deviation (RMSD) from the folded structure) and a broken K-E h-bond. A simulation snapshot of BRI1 kinase domain with an unfolded  $\alpha$ C-helix is also shown (Fig. 2c). **These results indicate the presence of catalytically incompetent conformations in the ensemble of kinase domain in its fully phosphorylated state.** Similar observations about the regulation of kinase activation mechanism via  $\alpha$ C-helix have been reported for a variety of human kinases such as EGFR, PKA, RET, Aurora/TP2X, PKB/Akt, Rho and Fes kinase.[20] However, this behavior is not universal and several well-studied human kinases such as c-SRC, c-ABL maintain a folded  $\alpha$ C-helix conformation during activation.[22] **Our investigations (using bioinformatics methods) reveal that SERK-family kinases show large differences in their  $\alpha$ C-helix unfolding propensity, which could provide new avenues for conformational control of kinase activity of individual SERK-family kinases.** The discussed above provide information that is not readily accessible from the available crystal structures and experimental data. Therefore, synergistic combination of simulations and available experimental data could answer challenging mechanistic questions beyond the capability of the individual techniques. **We have also performed experiments in collaboration with Prof. S.C. Huber (Plant Biology, UIUC) to validate these computational predictions.**

## List of Publications

1. Conformational Heterogeneity of the Calmodulin Binding Interface. *Shukla et al., Nature Communications, In press, 2016*
2. Application of Hidden Markov Models in Biomolecular Simulations. *Shukla et al., Methods in Molecular Biology, In press, 2016*
3. Guiding functional dynamics and dimerization of proteins using evolutionary couplings. *Moffett et al., Submitted, Journal of Chemical Theory and Computation, 2015.*
4. Elucidating the mechanism and regulatory role of sodium binding in different classes of GPCRs. *In preparation, 2016.*
5. Conformational engineering of  $\alpha$ C helix for regulating the activity of plant kinases. *In preparation, 2016.*
6. Molecular mechanism for inhibitory glutathionylation of Arabidopsis BRI1-associated receptor-like kinase 1 (BAK1). *In preparation, 2016.*

## Research Activities for Next Year

In the period March 2016-Feb 2017, we intend to extend our current project in new directions. Example of our planned work on Blue Waters also includes starting a new collaborative project with Prof. Huimin Zhao, Chemical & Biomolecular Engineering department, UIUC to investigate the specificity of gene-editing proteins such as CRISPR/Cas9 and TALENs.

**Aim 1: Constructing full-length receptor complex (BRI1-BAK1) model using evolutionary-coupling guided dynamics and experimental mutagenesis data.** The missing pieces in the full-length receptor complex model of BRI1-BAK1 receptor are the complexes formed by transmembrane helices and the cytoplasmic kinase domain. We plan to employ a novel method developed by our group called evolutionary coupling-guided dynamics to accelerate the simulations of protein-protein association. The

ECGD method uses a combination of bioinformatic analysis and molecular simulation, along with any experimental structural information that becomes available through literature and collaborators, to construct an atomistic, functional model of the assembled BRI1-BAK1 complex.

**Aim 2: Enabling conformational control of the Brassinosteroid activity of RLKs.** We hypothesize that understanding the mechanism by which post-translational modifications such as residue phosphorylation, glutathionylation, protein-protein interaction etc. modulate the conformational landscape of RLKs could help us in designing targeted interventions for controlling the RLK activity in plants. We plan to build kinetic models of BRI1 and BAK1 kinase domain to investigate the individual effects of multiple phosphorylated or glutathionylated residues on BRI1-BAK1 kinase domain. **It is challenging to estimate the contribution of an individual phosphorylation site to the overall protein activity experimentally without introducing mutations at other phosphorylation sites. This information can be obtained using *in silico* approaches.**

## Resource Requirement for Next Year

1. Storage: Default storage quota in project home and Nearline are sufficient for this project. Our current usage on Nearline is 14.5TB/50 TB. We anticipate the usage for next year to be similar.
2. Estimated Blue Waters Usage Schedule: Q1: 25%, Q2: 25%, Q3: 25%, Q4: 25%.
3. Data Transfer: Statistics for the data transfer from or to Blue Waters are not directly available. We plan to transfer simulation data to local UIUC clusters for analysis or archiving. We anticipate the total transfers to be of same magnitude as the amount of the data generated per year(15-30 TB in 1 year).
4. Memory Requirements: The planned simulations do not need more memory than currently available on the XE and XK nodes.
5. Node hour Requirement:

Table 1: Simulations planned on Blue Waters for the allocation period March, 2016-February, 2017.

System	Simulation time ( $\mu$ s)	Node-hours
Post-translational modifications of BAK1 kinases (6 modifications for 20 $\mu$ s each )	120 $\mu$ s <sup>a</sup>	127,500
BRI1 receptor with Brassinosteroid bound	50 $\mu$ s <sup>b</sup>	80,000
BRI1 receptor with Brassinosteroid & co-receptor bound	50 $\mu$ s <sup>b</sup>	80,000
Total		~290,000 hours

Note: a) The total time for conformational dynamics of post-translationally modified proteins is based on the the time estimates obtained from the simulations of the native BAK1 and BRI1 kinase domains. b) The total time requirement for these simulations are estimated using the experimental ligand binding kinetic measurements.

## References

- [1] Z. Han, Y. Sun, and J. Chai, "Structural insight into the activation of plant receptor kinases," *Current Opinion in Plant Biology*, vol. 20, pp. 55–63, 2014.
- [2] Y. Sun, Z. Han, J. Tang, Z. Hu, C. Chai, B. Zhou, and J. Chai, "Structure reveals that bak1 as a co-receptor recognizes the bri1-bound brassinolide," *Cell Research*, vol. 23, no. 11, p. 1326, 2013.
- [3] W. Cheng, K. R. Munkvold, H. Gao, J. Mathieu, S. Schwizer, S. Wang, Y.-b. Yan, J. Wang, G. B. Martin, and J. Chai, "Structural analysis of pseudomonas syringae avrptob bound to host bak1 reveals two similar kinase-interacting domains in a type iii effector," *Cell Host & Microbe*, vol. 10, no. 6, pp. 616–626, 2011.

- [4] C. Müssig and T. Altmann, “Brassinosteroid signaling in plants,” *Trends in Endocrinology & Metabolism*, vol. 12, no. 9, pp. 398–402, 2001.
- [5] D. K. Ray, N. Ramankutty, N. D. Mueller, P. C. West, and J. A. Foley, “Recent patterns of crop yield growth and stagnation,” *Nature Communications*, vol. 3, no. 1293, 2012.
- [6] D. Tilman, C. Balzer, J. Hill, and B. L. Befort, “Global food demand and the sustainable intensification of agriculture,” *Proceedings of the National Academy of Sciences, USA*, vol. 108, no. 50, pp. 20260–20264, 2011.
- [7] S. P. Long, A. Marshall-Colon, and X.-G. Zhu, “Meeting the global food demand of the future by engineering crop photosynthesis and yield potential,” *Cell*, vol. 161, no. 1, pp. 56–66, 2015.
- [8] T. Wheeler and J. von Braun, “Climate change impacts on global food security,” *Science*, vol. 341, no. 6145, pp. 508–513, 2013.
- [9] I. De Smet, U. Voss, G. Jürgens, and T. Beeckman, “Receptor-like kinases shape the plant,” *Nature Cell Biology*, vol. 11, no. 10, pp. 1166–1173, 2009.
- [10] J.-Y. Zhu, J. Sae-Seaw, and Z.-Y. Wang, “Brassinosteroid signalling,” *Development*, vol. 140, no. 8, pp. 1615–1620, 2013.
- [11] N. Sanabria, D. Goring, T. Nürnberger, and I. Dubery, “Self/nonself perception and recognition mechanisms in plants: a comparison of self-incompatibility and innate immunity,” *New Phytologist*, vol. 178, no. 3, pp. 503–514, 2008.
- [12] M. Tör, M. T. Lotze, and N. Holton, “Receptor-mediated signalling in plants: molecular patterns and programmes,” *Journal of Experimental Botany*, vol. 60, no. 13, pp. 3645–3654, 2009.
- [13] L. A. Gish and S. E. Clark, “The rlk/pelle family of kinases,” *The Plant Journal*, vol. 66, no. 1, pp. 117–127, 2011.
- [14] Z. He, Z.-Y. Wang, J. Li, Q. Zhu, C. Lamb, P. Ronald, and J. Chory, “Perception of brassinosteroids by the extracellular domain of the receptor kinase bri1,” *Science*, vol. 288, no. 5475, pp. 2360–2363, 2000.
- [15] J. Li and J. Chory, “A putative leucine-rich repeat receptor kinase involved in brassinosteroid signal transduction,” *Cell*, vol. 90, no. 5, pp. 929–938, 1997.
- [16] D. Shukla, C. X. Hernandez, J. K. Weber, and V. S. Pande, “Markov state models provide insights into dynamic modulation of protein function,” *Accounts of Chemical Research*, vol. 48, no. 2, pp. 414–422, 2015.
- [17] D. Shukla, Y. Meng, B. Roux, and V. S. Pande, “Activation pathway of src kinase reveals intermediate states as targets for drug design,” *Nature Communications*, vol. 5, no. 3397, 2014.
- [18] K. Kohlhoff, D. Shukla, M. Lawrenz, G. Bowman, D. E. Konerding, D. Belov, A. R. B., and V. S. Pande, “Cloud-based simulations on google exacycle reveal ligand modulation of gpcr activation pathways,” *Nature Chemistry*, vol. 6, pp. 15–21, 2014.
- [19] D. Shukla, M. Lawrenz, and V. S. Pande, “Elucidating ligand-modulated conformational landscape of gpcrs using cloud-computing approaches,” *Methods in Enzymology*, vol. 557, pp. 551–572, 2015.
- [20] N. Jura, X. Zhang, N. F. Endres, M. A. Seeliger, T. Schindler, and J. Kuriyan, “Catalytic control in the egf receptor and its connection to general kinase regulatory mechanisms,” *Molecular Cell*, vol. 42, no. 1, pp. 9–22, 2011.
- [21] S. S. Taylor, R. Ilouz, P. Zhang, and A. P. Kornev, “Assembly of allosteric macromolecular switches: lessons from pka,” *Nature Reviews Molecular Cell Biology*, vol. 13, no. 10, pp. 646–658, 2012.
- [22] Y. Shan, M. P. Eastwood, X. Zhang, E. T. Kim, A. Arkhipov, R. O. Dror, J. Jumper, J. Kuriyan, and D. E. Shaw, “Oncogenic mutations counteract intrinsic disorder in the egfr kinase and promote receptor dimerization,” *Cell*, vol. 149, no. 4, pp. 860–870, 2012.