1. Computational Ancestral Gene Resurrection for Investigating Selectivity Mechanisms of Kinases and GPCRs

2. Principal Investigator

Diwakar Shukla,
Assistant Professor, Department of Chemical and Biomolecular Engineering,
Affiliate faculty, Plant Biology, Center for Biophysics & Quantitative Biology and NCSA
University of Illinois at Urbana-Champaign,
Urbana, IL 61801

3. Executive Summary

Kinases and G-protein coupled receptors (GPCRs) are cell signaling proteins involved in various physiological functions. Small molecules and allosteric modulators such as sodium (Na^+) ion binds to these proteins and modulates their function. The drug binding sites of Kinases and GPCRs share a high degree of similarity and shows various degree of effect. Designing selective molecules and elucidating functional mechanism is a key challenge in drug discovery pipeline. Since the pharmacology of these protein rely on structural changes it is important to clearly understand the mechanistic basis of transition between the states. Such processes are slow and occurs at long time scales, powerful computational resources are required to study these complex systems. We used Blue Waters Supercomputers to understand the fundamental biology behind such rare events at millisecond timescale time scale to design better drugs.

4. Key Challenge

Selectivity Mechanisms in Kinases and GPCRs. Kinases and GPCRs are key proteins involved in a large number of human diseases. These proteins are targets for approximately 60% of all drugs on market. Kinases are cell-signaling proteins involved in cell division. Mutations in protein kinases can cause cancer, which makes kinases one of the major cancer drug targets [1][2]. G protein-coupled receptors (GPCR) belong to the largest class of integral membrane proteins. GPCRs are activated by ion, peptides, hormone etc., and transduce signal to downstream effectors responsible for regulating key physiological functions in human body. These proteins adopt multiple structural conformations and their function depends on high degree of plasticity. Therefore, their conformational equilibrium is modulated using ions, drugs, and binding proteins to regulate their function. However, the mechanism by which their conformational equilibrium is perturbed has remained elusive. For example, Imatinib is one of the clinically successful cancer drugs, called to be the wonder drug of the century, mainly because of its least amount of side effects [3]. This cancer drug inhibits protein kinase, Abl, and strongly binds to it, while does not inhibit kinase Src from the same family with ~46% sequence identity and high structure similarity [4]. However, the **decades of computational and experimental research has not revealed the molecular origin of Imatinib's selectivity towards Abl kinase.** Similalry, GPCRs are flexible proteins and their conformational equilibrium is perturbed by sodium ions. High resolution crystal structures are required to

identify the ion binding sites on GPCRs. It is very challenging to obtain crystal structure due to their low expression, detergent instability and lack of stable crystal contacts [5].

5. Why it matters

Kinases. After a decade of research the reason behind high selectivity of Imatinib stayed unclear [4,6]. In this project, we studied the behavior of five common ancestors of Src and Abl to find the evolutionary pathway, which makes them behave differently in Imatinib binding process. Understanding the exact atomistic processes can help us in design of more powerful drugs with the least side effects.

G-Protein Coupled Receptors. GPCRs are called as allosteric machines as many ligands modulate its function and signaling mechanisms. Ligands can bind to the sites other than the primary ligand binding site (orthosteric) and alters the receptor function [7]. The recent experimental and biochemical results show that the increase in ion concentration increases the affinity towards the antagonist and decreases the agonist binding. The mutation of $Asp^{2.50}$ abolishes the Na^{+} ion binding and increases the affinity for the agonist [8]. Understanding of functional interactions of ions at the allosteric site at the atomic level is important to modulate the function of GPCR.

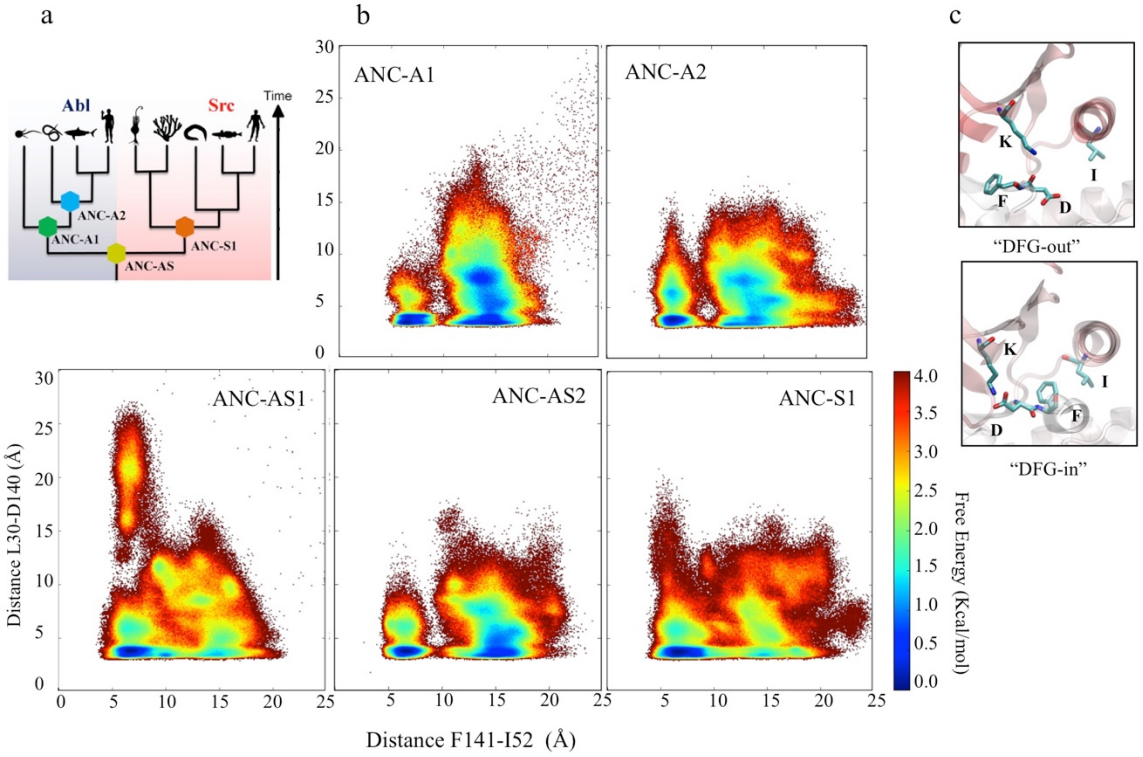

Figure 1. (a) Phylogenetic tree between Abl and Src. (b) DFG-flip related distance distributions for different ancestors. ANC denotes ancestor, S denotes that ancestor is closer to Src in the phylogenetic tree, A denotes that ancestor is closer to Abl and AS denotes that ancestor lies in between the Src and Abl. (c) DFG-out and DFG-in conformations

6. Why Blue Waters?

A total of 600 starting structures were subjected to molecular dynamics (MD) simulations and performed ~500μs of unbiased multiple short MD simulation guided with adaptive sampling method. Similarly, the 17 GPCRs with available crystal structures were subjected to MD simulation and multiple copies of each GPCR were simulated to produce statistically significant results. A total ~250μs simulation data were generated on Blue Waters to visualize the ion binding to the allosteric site. We also identified residues that act as a major barrier for the ion permeation for the receptors $A_{2A}R$, M_3, $5HT_{1B}$, $5HT_{2B}$, LA_1R and high barrier is noticed for β2AR, S_1P and H_1 (Figure. 2). Such extensive parallel MD simulation was only feasible using powerful supercomputers like Blue Waters.

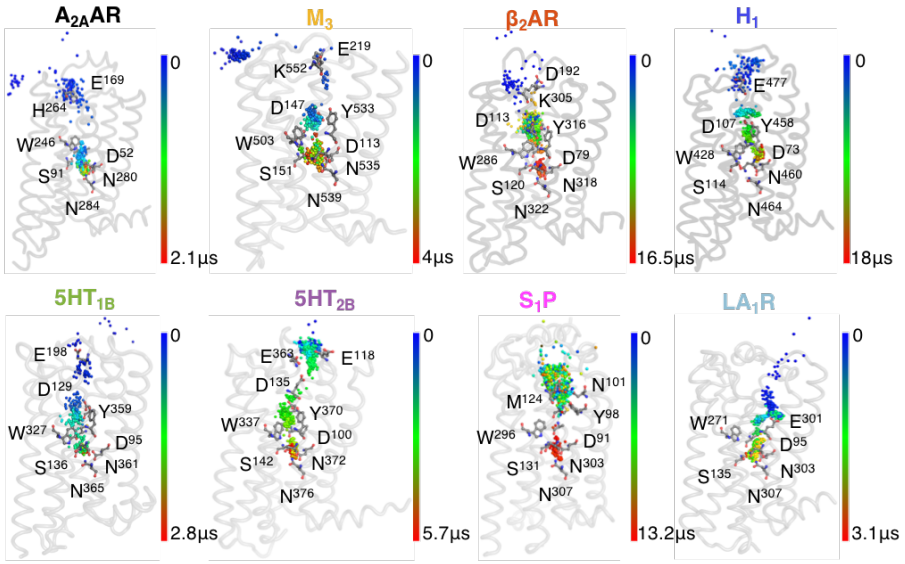

Figure 2. Ion binding mechanism in different GPCRs. The color bar shows the time required for the ion to reach the allosteric site.

7. Accomplishments

Kinases. Several plausible mechanisms of drug binding selectivity have been reported in the literature. The DFG-motif (Asp-Phe-Gly) is a highly conserved segment of the activation loop in kinase domains that is proposed to play a major role in the selection mechanism. Several groups have argued that the kinetic basis of Gleevec selectivity for Abl, compared to c-Src kinase is rooted in a pre-existing equilibrium between two conformations of the DFG motif, the inactive "D-out/F-in" or "DFG-out" and the active "D-in/F-out" or "DFG-in" [9]. Gleevec binds the "DFG-out" conformation. We investigated this hypothesis by measuring distances between two pairs of residues and weight them using equilibrium probably derived from Markov State Model calculations as shown in Figure 1. Our results indicate that all ancestors can adopt both the DFG-in and DFG-out conformations. The stability differences among the ancestors for the DFG-out conformation are not large; MD simulations results indicates the difference to be ~1 kcal/mol as compared to the overall binding free energy difference of ~5 kcal/mol. **We have successfully completed this work and a manuscript is under preparation for submission to a top**

interdisciplinary science journal. We also constructed markov state model (MSM) using a machine learning approach to analyze and obtained kinetics of ion binding to various GPCRs. **We have successfully completed this work and an article is submitted to the journal *Angewandte Chemie International Edition*.**

8. Next Generation Work

Using evolutionary information, we studied the changes of protein kinases in the past and understood the mechanism of a clinically successful present drug. In the next step, we want to use evolutionary information to study the changes of protein kinases in the future by predicting drug resistant mutations in cancer patients. We are going to use predicted drug resistant mutations to design the next generation of cancer drugs before they will be needed. Similarly, our results on ion-binding to GPCRs reveal the molecular origin of different effect of ion on GPCRs. For example, the close homologues like adrenergic receptors the nature of ion effect varies from millimolar to micromolar inhibition concentration. Our future work, involves the understanding the evolution of ion effect in different GPCRs.

9. References

[1] Newton et al., vol. 1, no. 27, pp.28495-28498, 1995.
[2] Sawyers et al., Cancer Cell , vol. 1, no. 1, pp. 13–15, 2002.
[3] Capdeville et al., Nat. Rev. Drug Discov. , vol. 1, no. 7, pp. 493–502, 2002.
[4] Seeliger et al., Structure , vol. 15, no. 3, pp. 299–311, 2007.
[5] 6. Carpenter et al., Curr. Opin. Struct. Biol, 2008, 18(5), 581-586
[6] Lin et al., Proc. Natl. Acad. Sci. U. S. A. , vol. 110, no. 5, pp. 1664–9, 2013.
[7] Katritch etal., Trends Biochem Sci, 39(5), 2014, 233-233
[8] Teran et al., Structure. 2013;21:2175–2185.
[9] Wilson et al., Nature, vol. 5287, no. 1995, pp. 201–206, 2014.