EFFECT OF D122N MUTATION IN BAK1 ON STRUCTURAL INTEGRITY OF PROTEIN COMPLEX CONSISTING FLS2, FLG22 AND BAK1 ECTODOMAIN- A MOLECULAR DYNAMICS STUDY

by
Arka Roy
18276002

A dissertation submitted to the Department of Mathematics and Natural Sciences, BRAC university in partial fulfillment of the requirement for the degree of Master of Science in Biotechnology

Department of Mathematics and Natural Sciences
BRAC University
January, 2020

Declaration

Declaration

It is hereby declared that

1. The thesis submitted is my own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. I have acknowledged all main sources of help.

Candidate

_____

Arka Roy

18276002

Approval

The thesis titled “Effect of D122N mutation in BAK1 on structural integrity of protein complex consisting FLS2, flg22 and BAK1 ectodomain- A molecular dynamics study” submitted by Arka Roy (18276002) has been accepted as satisfactory in partial fulfillment of the requirement for the degree of Master of Science on January 30, 2020.

Examining Committee:

Supervisor:

_____

M H M Mubassir
Lecturer, Biotechnology Program
Department of Mathematics and Natural Sciences

Program Coordinator:

_____

Iftekhar Bin Naser, PhD
Assistant Professor
Department of Mathematics and Natural Sciences

Department Head:

_____

Prof. Dr. A F M Yusuf Haider
Professor and Chairperson
Department of Mathematics and Natural Sciences

Ethics Statement

No living organism was used in this study.

Abstract

Elongated (elg) phenotype of *Arabidopsis thaliana* is caused by D122N mutation in the third leucine repeat of BRI1-associated kinase 1 (BAK1) protein that changes the 122nd amino acid aspartate to asparagine. The mutation is mainly associated with photomorphogenesis since plants with this mutation are characterized to be early flowering and have elongated stem and petioles. However, the BAK1 protein also plays a crucial role in plant immunity by forming a heterodimer with pattern triggered receptor Flagellin sensing 2 (FLS2) upon the perception of flg22, an elicitor protein from bacterial flagellin, that cause transphosphorylation of BAK1 and FLS2 and subsequent initiation of signal transduction pathway leading to activation of the immune response. Here we investigated the impact of D122N mutation in BAK1 on the structural integrity of the FLS2-flg22-BAK1 complex through molecular dynamics simulation. We induced the D122N mutation in BAK1 protein of native FLS2-flg22-BAK1 crystallographic structure using the in-silico method. Subsequently, we simulated both native and mutated complex using molecular dynamics methods for 30ns. The simulations were compared using parameters such as root mean square deviation (RMSF), root mean square fluctuation (RMSF), the radius of gyration (Rg) and the number of hydrogen bonds to compared the structural integrity of two complexes. We found that the mutation resulted in often 2 Armstrong higher deviation of atoms from reference structure for the whole complex, FLS2, BAK1, and flg22 compared to counterpart from the non-mutated complex. RMSF analysis revealed that mutation caused higher fluctuation of amino acid in Both the N terminal and C terminal of FLS2 that hinder interaction with both flg22 and BAK1. Except for flg22, both FLS2 and BAK1 had a higher radius of gyration in the complex containing mutation compared to the non-mutated complex. The mutation hindered the formation of hydrogen bonds among all the peptides, often terminating all the hydrogen bonds between peptide. Even though ASP122 does not form a direct connection with FLS2 atoms, the mutation results in the termination of intra-protein interaction of 122nd amino acid with ARG146 in BAK1, which undermines the interaction of BAK1 peptide with FLS2.

Dedicated To

My Mother and Father without whom I am nothing

Table of Contents

Acknowledgement

First of all, I would like to express my deepest gratitude to Almighty God for everything he has blessed me with in my whole life and for giving me patience and strength to make this project a success.

I would like to thank Professor A F M Yusuf Haider, Ph.D., Professor and Chairperson of Department of Mathematics and Natural Sciences, BRAC University for allowing me to continue my work at the BRAC University microbiology laboratory. I express my heartiest gratitude to late Professor A. A. Ziauddin Ahmad**,** former Chairperson of the Department of Mathematics and Natural Sciences, BRAC University for his valuable advice, enthusiastic encouragement and the support he gave me throughout my undergraduate study. He has been my role model.

I am obliged to my supervisor Mr. M H M Mubassir, Lecturer, Biotechnology program, Department of Mathematics and Natural Sciences, BRAC University for believing in me and giving me support whenever I needed. I would like to again thank him for his constant encouragement and guidance throughout this project.

I would like to especially thank Professor Dr. Aparna Islam, Department of Mathematics and Natural Sciences, BRAC University for the inspiration and guidance she provided me as a student. I would like to thank Professor Dr. Mahboob Hossain for his unending encouragement and motivation. I would also like to show my thanks to late Professor Naiyyum Choudhury, former Coordinator of the Biotechnology and Microbiology program, Department of Mathematics and Natural Sciences, BRAC University for his exemplary guidance and encouragement during his short stay with us.

I especially thank my friends and classmates Fahim Sarwar, D.M. Asiquzzaman and all other classmates for giving me company and helping me throughout my post-graduate study.

Arka Roy

List of Tables

List of Figures

List of Acronyms

PRR: Pattern Recognition Receptor

LRR - Leucine Rich Repeat Domain

GROMACS - Groningen Machine for Chemical Simulations

PAMP - Pathogen Associated Molecular Pattern

MD - Molecular Dynamics

PTI - Pattern Triggered Immunity

ETI - Effector Triggered Immunity

Rg - Radius of Gyration

RMSD - Root Means Square Deviation

RMSF - Root Means Square Fluctuations

List of Symbols

K - Kelvin

nm - Nano meter

ns - Nanosecond

ps - Picosecond

Chapter 1

Introduction:

Plant immune system:

The pathogenic organisms that attack plants use varied strategies to infect plants and weaken plant growth and maintenance. Unlike some higher-order aminol plants lack an adaptive immune system made up of mobile immune cells. Hence, Plants mainly rely on the innate immune system to detect and halt pathogen infections (Boller and Felix, 2009; Spoel and Dong, 2012; Thomma et al., 2011). It has two interconnected tiers. One uses pattern-recognition receptors (PRRs) present on the surface of the cell to recognize microbe-associated molecular patterns (MAMPs) or often called pathogen associated molecular pattern (PAMPs). These are molecular markers present in a diverse group of microorganisms and host-derived damage-associated molecular patterns (DAMPs) (Boller and Felix, 2009). On the other hand, the second innate immune system responds to effector molecules secreted by pathogens, which helps them to establish infections and suppress plant immunity, using disease resistance (R) proteins (Upson et al., 2018).

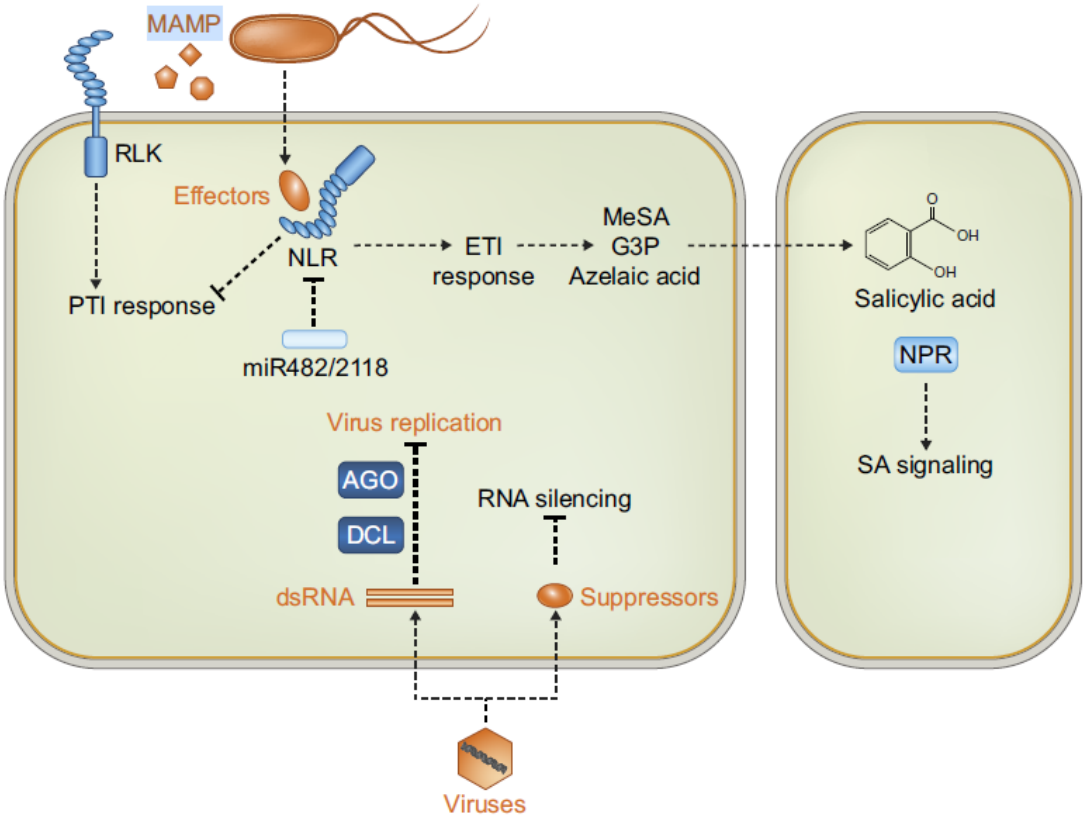

Figure 1: A brief overview of the plant immune system (adapted from (Han, 2019).

Pattern-triggered immunity (PTI) response:

The detection of PAMPs or DAMPs by different pattern recognition receptors (PRRs) activates the defense mechanism against infecting pathogens, termed as pattern triggered immunity (PTI) (Jones and Dangl, 2006; Zipfel, 2014). Till now, receptor- like proteins (RLPs)and Receptor like kinases (RLKs) have been discovered to perform as PRRs (Boller and Felix, 2009). Some members of these protein groups play essential roles in plant growth, symbiosis and resistance to abiotic stresses (Tang et al., 2017). A transmembrane receptor is characterized as RLK when it has an extracellular domain taking part in ligand perception, a transmembrane domain and an cytoplasmic kinase domain (Couto and Zipfel, 2016; Zipfel and Oldroyd, 2017). RLPs are almost similar to RLKs except that they don't have the kinase domain like RLKs have (Zipfel, 2014). Pattern recognition receptors contains a range of extracellular domains, such as lysine motif (LysM), leucine-rich repeat (LRR), lectin and epidermal growth factor (EGF)-like domains. Therefore, these receptors can recognize a varied types of ligands (Yu et al., 2017). In early steps of ligand perception, PRRs form heterodimer with specific co-receptors (Ranf, 2017). Following the ligand binding activation of receptor like cytoplasmic kinases (RLCKs) by PRR complexes occur to transduce the PTI signal (Yu et al., 2017). These RCLKs then initiate the activation of calcium-dependent protein kinases (CDPKs) as well as mitogen-activated protein kinase (MAPK) cascades, which in turn activate the transcription of defense genes involved in PTI response (Couto and Zipfel, 2016).

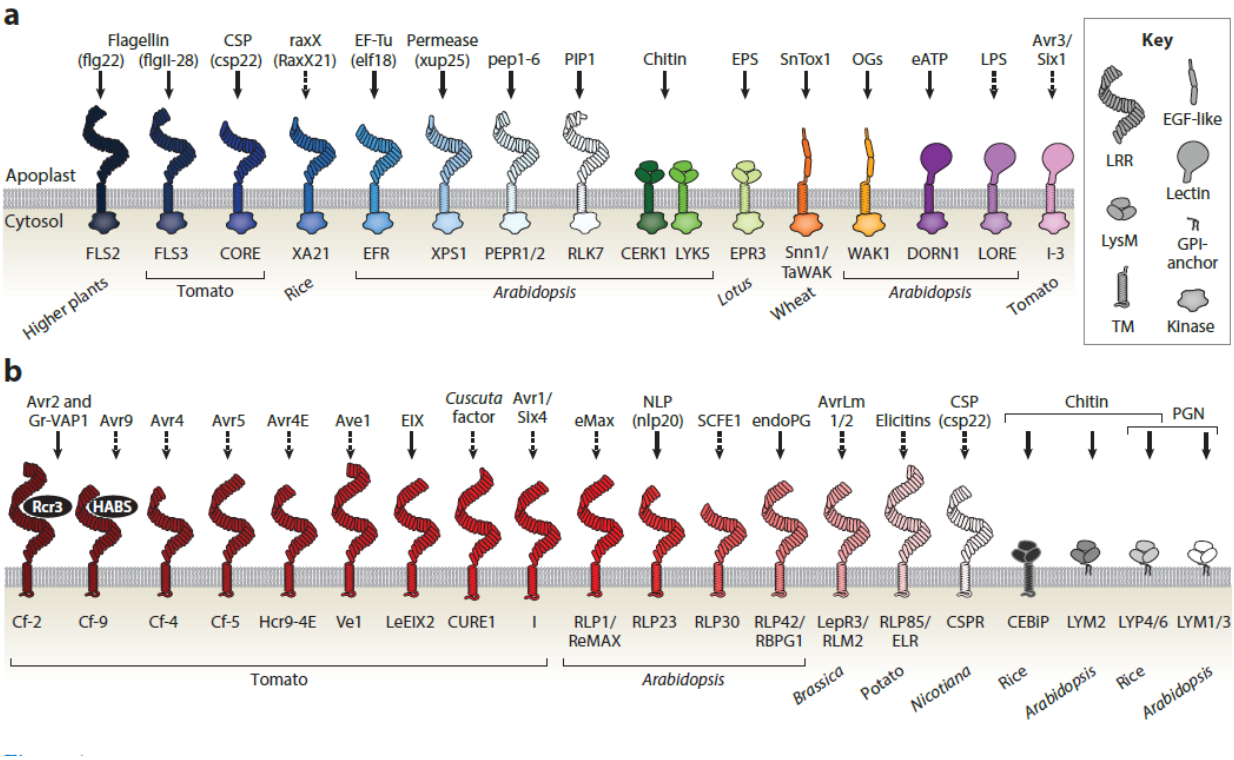

Figure 2: Different pattern-recognition receptors (PRRs) of plant and their respective recognized bacterial PAMPs. (a) Receptor kinases (RKs). (b) Receptor-like proteins (RLPs) (Boutrot and Zipfel, 2017).

Effector-triggered immunity (ETI) response:

To circumvent bacterial PTI immune mechanism pathogens have established a diverse collection of effector molecules that can bypass PTI and insert into plant cells to interfere with PTI (Jones and Dangl, 2006). These effectors are generally pathogen-specific and are usually conserved only within the family level of microbial groups (Boller and Felix, 2009). To counter this infection strategy, R proteins are used by plants to respond to effectors. Plants use two strategies in this mechanism. One is through direct binding and the other is through perceiving perturbations of host molecules by effectors (Kourelis and Hoorn, 2018). Perception of effectors initiates the effector triggered immunity (ETI) (Spoel and Dong, 2012). Most of the genes encoding plant R proteins contain nucleotide binding-site leucine-rich repeat (NLR), which is also known as NBS-LRR, proteins. Initiation of ETI usually results in the induction of programmed cell death at the site of infection, called the hypersensitive response (HR). As a result, it prevents pathogen to spread (Jones and Dangl, 2006). ETI initiates the production of glycerol-3-phosphate (G3P), methyl salicylic acid (MeSA) and azelaic acid. These compounds act as mobile immune signals in plants that are transported to other uninfected tissues from the infection site (Fu and Dong, 2013). When these signals are perceived by

uninfected tissues, they induce the production of salicylic acid and facilitates massive transcription of necessary genes (Fu and Dong, 2013). This initiated immune mechanism is called systemic acquired resistance (SAR). Eventually, it causes the production of pathogenesis-related (PR) proteins that convey antimicrobial activity. Consequently, it protects plants from successive pathogen attacks (Spoel and Dong, 2012). In *Arabidopsis thaliana*, both the NPR3/4 proteins and non-expresser of PR genes 1 (NPR1) protein bind salicylic acid and serve as salicylic acid receptors (Ding et al., 2018). NPR1 protein act as a transcriptional co-activator and NRP3/4 proteins act as transcriptional co-repressors, and they play opposite roles in transcriptional control of defense gene expression (Ding et al., 2018).

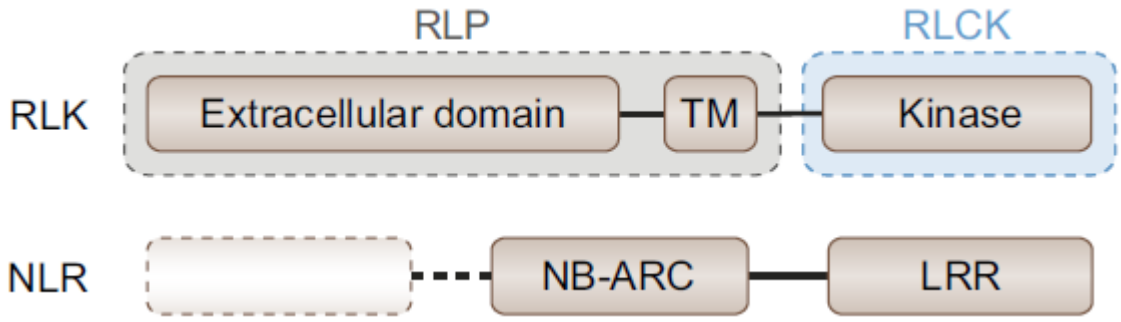

Figure 3: Domain architecture of RLKs and NLRs (Han, 2019).

RNA interference:

The RNA interference (RNAi), also known as RNA silencing, works as defense machinery against viral infections in plants (Ding, 2010). Dicer like (DCL) proteins process Double-stranded RNA (dsRNA) molecules, which are generally replicative intermediates of RNA viruses, and produce viral small interfering RNAs (siRNAs). RNA-based antiviral immunity induced when loading viral siRNAs into Argonaute (AGO) proteins that guide viral mRNA cleavage (Ding, 2010). Moreover, small RNA molecules have recently been indicated to play important roles PTI and ETI signaling pathways as well as regulating the expression of R genes in plants (Pumplin and Voinnet, 2013).

Fls2 mediated pattern triggered immunity:

FLAGELLIN-SENSING 2 (FLS2) is a receptor like kinase protein that contains leucine-rich repeat (LRR-RLK). It is a plant receptor that is activated by bacterial elicitor flagellin or the flagellin-derived flg22 peptide. This receptor was first found in *Arabidopsis Thaliana* (Gómez-

Gómez and Boller, 2000). Later it was discovered in other plant species such as rice (Takai et al., 2008), tobacco (Hann and Rathjen, 2007) and tomato (Robatzek et al., 2007).

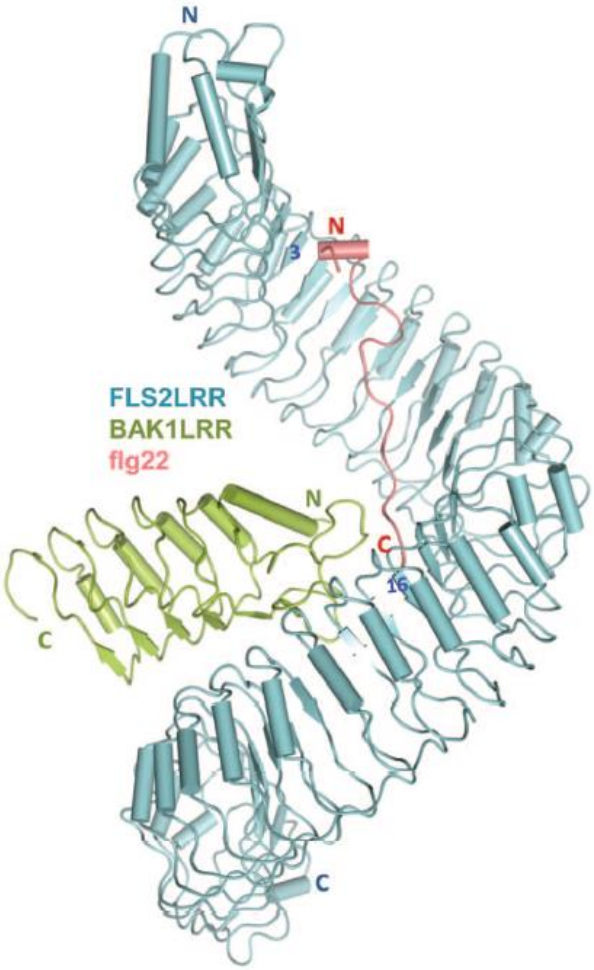

Figure 4: overall structure and binding pattern of FLS2, BAK1 and flg22 (Sun et al., 2013).

Fls2 consists of three major domains. An extracellular, a helical transmembrane and an intracellular domain. The extracellular domain consists of 28 leucine-rich-repeat(LRR) where the interaction with flagellin occurs. The transmembrane region is buried in the phospholipid layer of the cell membrane. The serine/threonine kinase domain makes up the intracellular domain of fls2. here, phosphorylation catalyzes a protein kinase cascade which in turn leads to signal transduction (Shiu and Bleecker, 2001).

Prior to ligand exposure, FLS2 is present in FLS2-FLS2 and FLS2- BIK1 complexes(Sun et al., 2012; Zhang et al., 2010). FLS2 also associates with close homologs of BIK1 such as PBS1, PBL1 and, PBL2, all of which are predicted cytoplasmic protein kinases (Zhang et al., 2010).

BAK1 Co-receptor:

BAK1 (BRI1-associated kinase 1) is a leucine-rich repeat receptor-like kinase (LRR-RLK). It consists of an extracellular LRR domain made of five repeats. A serine and proline rich domain follow the LRR domain which actually defines the SERK protein family (Hecht et al., 2001). It also has a membrane-spanning domain, a cytoplasmic kinase domain, and a short C-terminal tail. BAK1 is a member of the SERK protein family and has four homologs. The somatic embryogenesis receptor kinase 1 (SERK1) is the first identified member of this family and that is why BAK1 is also called SERK3 (Hecht et al., 2001). SERK3 was renamed as BAK1 after researchers found out in 2002 that SERK3 acts as a signaling partner of another LRR-receptor kinase (LRR-RK), BRI1 (Brassinosteroid Insensitive1) (Li et al., 2002; Nam and Li, 2002). interestingly, BAK1 was also found out to take part in light signaling (Whippo and Hangarter, 2005) and in the containment of cell death (He et al., 2007; Kemmerling et al., 2007). additionally, BAK1 was discovered to interact with receptor FLS2 (Flagellin Sensing 2) (Chinchilla et al., 2007; Heese et al., 2007) which has a function in plant innate immunity. These researches resulted in the formulation of the idea that BAK1 has a crucial role in the regulation of several LRR-RLKs by interacting with them in a stimulus-dependent manner (Kemmerling and Nürnberger, 2008; Vert, 2008).

flg22 as an activator of FLS2:

flg22 is a 22 amino acid long peptide sequence derived from the conserved part of the N-terminal region of flagellin (Felix et al., 1999). The sequence of flg22 is based on that of flagellin from *Pseudomonas aeruginosa* (Felix et al., 1999). Bacterial flagellum Filaments form tubular structures made of protofilaments, which are composed of arrays consisting of up to hundreds of flagellin molecules (O'Brien and Bennett, 1972). However, the flg22 epitope is not in the exterior of flagellum making accessibility difficult for the FLS2 receptor (Samatey et al., 2001; Yonekura et al., 2003). surprisingly, Flagellin peptide might defuse into the environment during the creation of the flagella (Komoriya et al., 1999) or could be detected after the flagella break down.

Bacterial flagellin flg22 that is perceived by FLS2, is highly conserved since it is a must requirement for the function of flagella and bacterial mobility (Haefele and Lindow, 1987). However, some bacteria that are pathogenic to pant have flg22 peptides with mutant sequences that lead to loss of efficiency in terms of detection by the plant defense systems. Few Bacteria such as *Agrobacterium tumefaciens* (Albert et al., 2010), *Ralstonia solanacearum* (Pfund et al., 2004), where a specific valine/aspartate polymorphism governs the perception of flagellin via FLS2 in *Arabidopsis thaliana* or not (Sun et al., 2006).

Activation prosses of FLS2 by flg22:

In the first step, the flagellin signal is detected by FLS2 independently of BAK1 by binding to its LRR ectodomain (Sun et al., 2013). This binding of the flg22 to FLS2 stimulates the formation of the FLS-BAK1 complex. Interestingly, the LRR domains of co-receptor BAK1 also interacts with the ligand flg22 (Sun et al., 2013). Flg22 promotes the cross phosphorylation between BAK1 and FLS (Wang et al., n.d.). this perception by FLS2 leads to phosphorylation and dissociation of RCLKs, such as BIK1 and BSK1 to initiate an immune response (Lu et al., 2010; Schwessinger et al., 2011; Zhang et al., 2010). BSK1 along with RLCK group VII members directly phosphorylate MAPKKK5 (Mitogen-activated protein kinase kinase kinase 5) and stimulate MAPK cascades which results in the up-regulation of defense genes expression(Yamada et al., 2016). Moreover, BIK1 promotes the activation of RbohD (Respiratory burst oxidase homolog protein D) and heterotrimeric G proteins to trigger ROS (Reactive Oxygen Species) burst (Liang et al., 2016). Besides, BIK1 activates the CNGC2–CNGC4 calcium channel by phosphorylating CNGC4 which causes in an increase of cytoplasmic Ca^{2+} concentration and activation of CDPKs (Couto and Zipfel, 2016; Tang et al., 2017; Wu and Zhou, 2013), which can further up-regulate RbohD activation and defense gene expression.

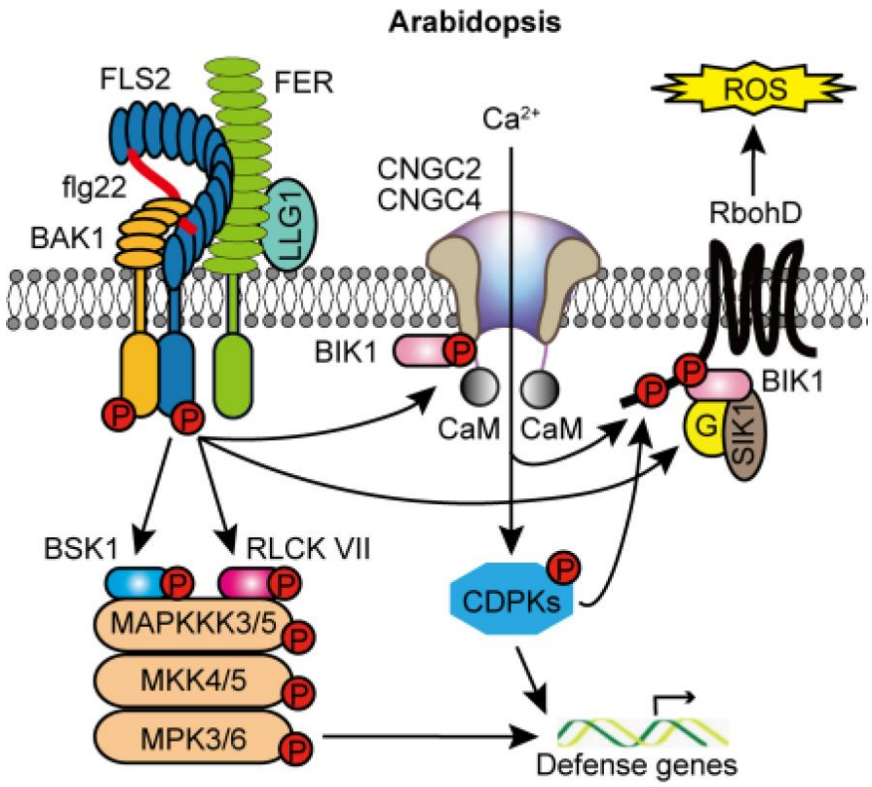

Figure 5: Overview of flg22 perception and molecular response in *Arabidopsis thaliana* (Wang et al., n.d.).

Molecular dynamics simulation:

Molecular dynamics simulation or MD simulation is a computational approach where it allows protein motion to be studied, by observing their conformational changes over time. The principle behind an MD simulation is that Given the coordinate of all of the atoms in a biomolecular system, such as water and perhaps a lipid bilayer surrounding a protein, it is possible to estimate the force applied on each atom by all of the other atoms. Thus, Newton's laws of motion can be used to predict the spatial location of each atom as a function of time. Especially, one steps over time, recurrently enumerating the forces on each atom and then using those forces to update the location and velocity of each atom. The trajectory that we get at the end is a three-dimensional motion picture that describes the atomic-level arrangement of the system at every point throughout the simulated time interval (Karplus and McCammon, 2002). Therefore, a wide variety of important biomolecular processes can be captured by these simulations, such as conformational change, protein folding, and ligand binding, which reveals the location of all of the atoms at femtosecond (fs) time interval resolution. Prominently, these simulations can also forecast how biomolecules will react at an atomic level to perturbations

such as phosphorylation, mutation, addition or removal of a ligand or protonation (Hollingsworth and Dror, 2018).

These simulations are a powerful and great addon to any experiment for some reason. First, they detect the location and motion of each atom at every point in time, which is highly difficult with the current experimental technique. Second, the simulation parameters are precisely known and can be controlled. These include the initial conformation of a protein, specific ligands that are bound to it, presence of any post-translational modifications or mutations, presence of other molecules that are present in the simulation environment the temperature, the voltage passing across a membrane, etc. By juxtaposing simulations performed under dissimilar conditions, we can detect the effects of a diverse range of molecular perturbations (Hollingsworth and Dror, 2018).

A model called molecular mechanics force field is used to calculate the forces in an MD simulation, which is fit to the results of quantum mechanical calculations and, generally, to few experimental measurements. Numerical stability in an MD simulation is ensured by selecting short time steps, generally only a few femtoseconds (1015 s) each. Most of the biochemical events of interest, such as structural changes in proteins take place on timescales of nanoseconds, microseconds, or longer. Thus, millions or billions of time steps are required in a typical simulation. This combined with the millions of interatomic interactions typically enumerated at each time step makes simulations very computationally demanding. There are several choices when it comes to choosing a force field. The most common force fields are different versions of AMBER, CHARMM, and OPLS (Harder et al., 2016; Huang et al., 2017; Robustelli et al., 2018). All of these force fields depend on similar functional forms, although, each has certain strengths and weaknesses. In terms of actual simulation software GROMACS, NAMD, AMBER, CHARMM, Desmond, and OpenMM are generally used (Abraham et al., 2015; Bowers et al., 2006; Brooks et al., 2009; Case et al., 2005; Eastman et al., 2017; Phillips et al., 2005). Before proceeding to perform the simulation, molecular system needs to prepared by building in missing atoms, which includes hydrogen atoms, adding in solvent particles such as water, salt ions, lipids in case of membrane protein, and selecting force field parameters (Betz, 2017; Jo et al., 2008; Sastry et al., 2013).

Implications of BAK1 mutation:

As BAK1 protein plays a crucial role in both developmental and immune signaling pathways in plants, a mutation in important amino acids within the protein will result in a disruption in both the pathways (McAndrew et al., 2014). *Arabidopsis thaliana* mutant line with the elongated phenotype (elg) (characterized by elongated hypocotyls and petioles, narrow leaves and early flowering) is caused by a point mutation (GGA to GAA) 1,370 base pair downstream from the transcriptional start site. The mutation (D122N) changes an Aspartate (D-122) to an Asparagine (N-122) in the third LRR (Whippo and Hangarter, 2005). Laboratory experiments predicted that the mutation obstructs the interaction of BAK1 with FLS2 but allows aberrant dimerization with the BRI1 receptor in the absence of ligand (Jaillais et al., 2011). Mutation in analogues position (D128N) in rice OsSERK2, which is structurally similar to BAK1 and a member of the SERK protein family, predicts the underlying molecular mechanism of the effect caused by D122N mutation in BAK1 during flg22 induced fls2 activation. Wild type osSERK2 aspartate128 forms a salt bridge to arginine152 on the fourth LRR. The D128N mutation not only disrupts the interaction between 128th and 152nd amino acids but also causes the Asn128 to form a salt bridge with Glu174 (McAndrew et al., 2014). The analogues arginine in BAK1 (Arg146) forms a salt bridge with Glu749 from BRI1 and BAK1 Asp122 plays a crucial role in this interaction by forming a salt bridge with BAK1 Arg146 that hold the amino acid in position to allow interaction with Glu749 (McAndrew et al., 2014). In the case of FLS2-BAK1 complex BAK1 Arg146 also directly contacts FLS2 LRR (Sun et al., 2013). However, the alteration in molecular interaction due to mutation of BAK1 Asp122 to Asn122 on flg22 induced FLS2 activation is still unknown.

Research Aim:

The goal of this study is to understand the molecular mechanism of FLS2-BAK1 mediated pattern triggered immunity (PTI) of *Arabidopsis thaliana*.

Specific Objectives:

1. Mutate 122nd amino acid in BAK1 from aspartate to asparagine in the crystal structure (PDB ID: 4MN8) using in silico method.
2. Performing molecular dynamics simulation of both native 4MN8 and D122N BAK1 mutation containing 4MN8 complex for 30 ns.
3. Analyzing the molecular interaction among PRR FLS2, PAMP flg22 and Co-receptor BAK1 ectodomain from the both native and mutated complex during simulation using parameters RMSD, RMSF, Rg, and H-bond.
4. Identification of change in molecular interaction within the mutated ectodomain complex compared to native ectodomain complex.
5. Analyzing structural dynamics of FlS2-D122N BAK1-flg22 complex.

Chapter 2

Methods and materials:

Mutation of BAK1:

First, the crystal protein structure of FLS2-BAK1-flg22 ectodomain complex of *Arabidopsis thaliana* was retrieved from the protein data bank (PDB ID: 4MN8) in '.pdb' format. The pdb file was processed in a text editor by removing water and ions. The mutation of BAK1 122nd amino acid aspartate to asparagine was done using the processed pdb file in pymol software. among all probable orientation of mutated amino acid asparagine, the final one was selected that had the least steric clash with surrounding molecules.

MD simulation:

The mutated pdb file was used to perform molecular dynamics simulation using GROMACS software. OPLS-AA/L all-atom force field was used to generate a topology file, position restraint file, and a post-processed structure file. Then the simulation system was defined by placing the protein complex at the center of a cubic box keeping 1nm minimum distance between the protein complex and the box edges to satisfy the minimum image convention of periodic boundary condition. The system was solvated using SPC/E water molecule type from spc216.gro configuration file, which filled the system with 104173 water molecules. Then the system was charge neutralized by adding 28 positive sodium (NA) ions in place of solvent as the complex had a net negative charge. Then the system was energy minimized using steepest descent minimization algorithm. Following energy minimization, the solvent and ions around the protein complex were equilibrated in two-phase. First, it was equilibrated under NVT (constant Number of particles, Volume, and Temperature) for 100ps (Picosecond). Second, the system was equilibrated under the NPT ensemble for 100ps to stabilize pressure and thus also density of the system. Finally, the system was simulated for 30ns (Nanosecond). The output file was generated as md_0_1.gro. A pdb file from the simulated '. gro' file was also generated using GROMACS. The pdb file was used to evaluate protein-protein interaction using PIC (protein interaction calculator).

Analysis:

The simulated protein complex was analyzed for RMSD (root mean square deviation) of the protein backbone, RMSF (root mean square fluctuation) per residue, R_g (radius of gyration) of

protein, using GROMACS. Additionally, intra-protein interaction consisting Hydrogen bonds, hydrophobic interaction and ionic interaction was calculated using PIC server.

Chapter 3

Results:

Pre Molecular Dynamics Simulation:

The protein complex was first energy-minimized after adding water solvent and neutralizing the charge of the system by adding ion. The energy minimization was completed at 1320 steps with the potential energy of -5.8×10^6 kJ/mol. Following energy minimization, the system was equilibrated first under NVT ensemble for 100 ps where the whole system stabilizes at an average temperature of 299.797 K with an RMSD of 3.06128 K. the second phase of equilibration was conducted under NPT ensemble where the system stabilizes at an average pressure of 1.38255 bar with an RMSD of 58.7928 bar. The density of the system stabilizes at an average density of 1012.31 kg/m^3 with an RMSD of 1.67151 kg/m^3. Then the molecular dynamic simulations were performed for 30 ns in a couple of Desktop computers consisting of intel i5 7500 processor and 16 gigabytes of RAM. The simulations took around 12 days to complete using all 4 threads with an average of 2.841 ns per day.

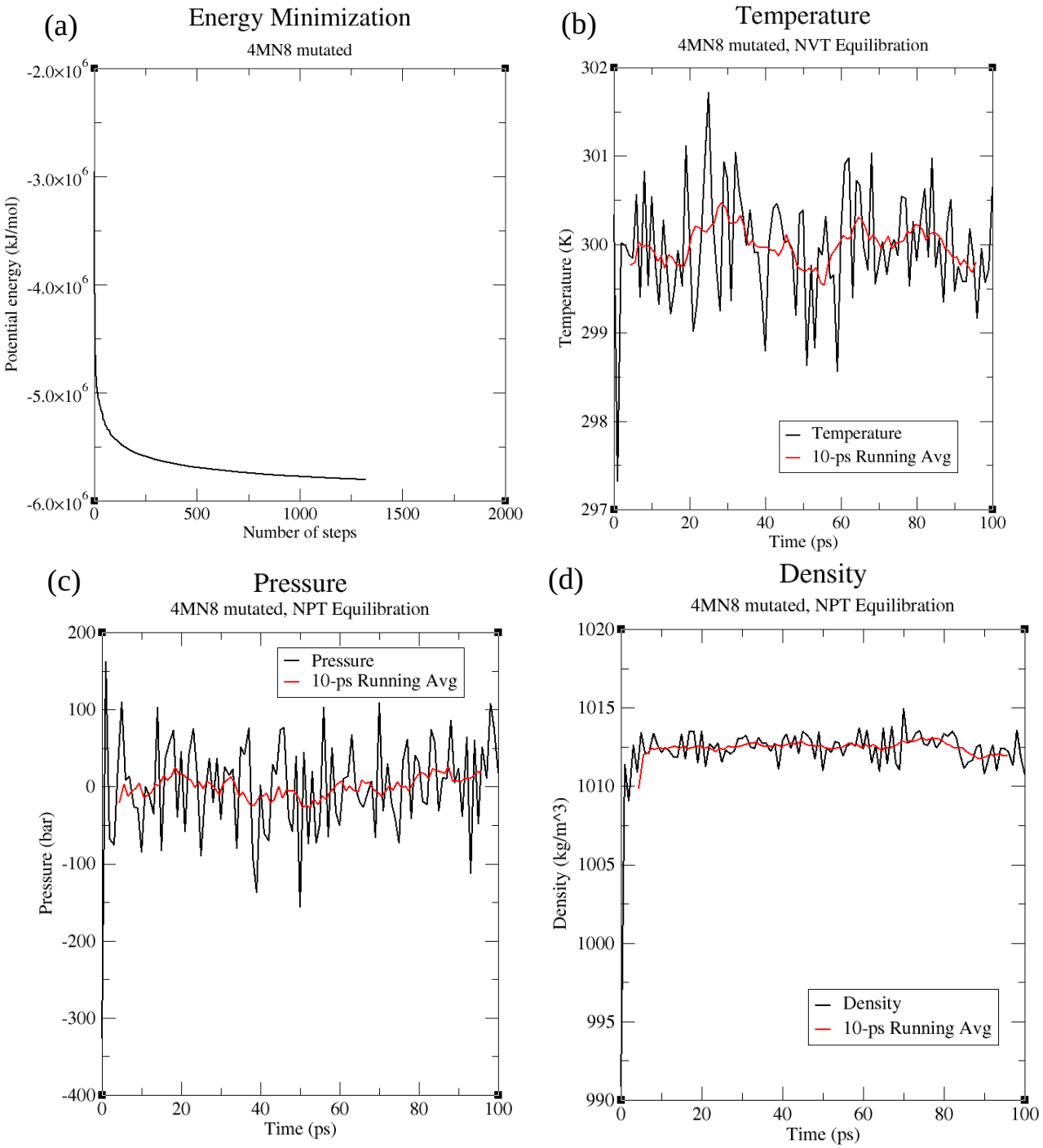

Figure 6: Different parameter of pre-molecular dynamics processing of protein complex. (a) energy minimization steps of 4MN8 mutated complex. (b) Temperature equilibration of 4MN8 mutated complex. (c) Pressure equilibration of 4MN8 mutated complex (d) Density equilibration of 4MN8 mutated complex.

Post Molecular Dynamics Simulation:

As molecular dynamic simulations were complete, several key parameters were analyzed to distinguish the difference between native and mutated protein complex. These analyses include RMSD over time of the whole complex, FLS2, BAK1, flg22; RMSF analysis of FLS2, BAK1, flg22; R_g analysis of the whole complex, FLS2, BAK1, flg22; Number of Hydrogen bonds over time between FLS2-BAK1, BAK1-flg22, FLS-flg22. Moreover, intra-protein interactions between native and mutated protein complex were also analyzed.

RMSD:

In Root mean square deviation or RMSD comparison of the whole complex between native and mutation, the RMSD of the mutated complex stayed lower than the native complex in the first 7ns. And after that time the RMSD of mutated complex stayed higher than the native complex RMSD most of the time. In the case of just FLS2 peptide, the mutated RMSD stayed lower than the native counterpart till 7.5ns and after that deviated higher than native FLS2 RMSD. RMSD values for mutated BAK1 also stayed lower than native one till 7.5ns after which it jumped higher than native one a few times till 11.7ns. from 11.7ns to 19.5ns the mutated data had lower value but 19.5ns onward mutated BAK1 RMSD got higher than native BAK1 RMSD. The flg22 PAMP RMSD was less than native flg22 RMSD till 10.2ns but onwards the RMSD value for both was quite similar.

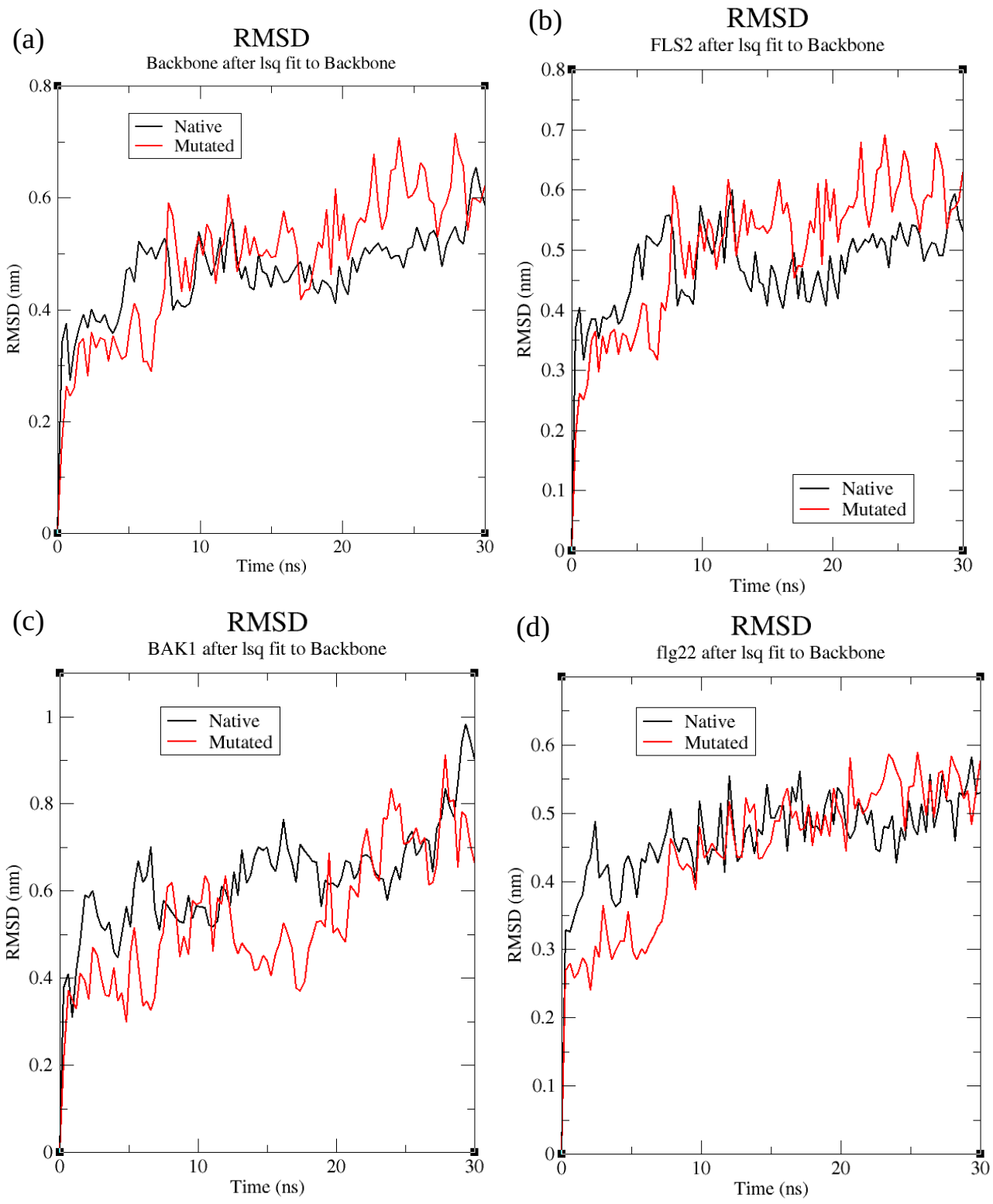

Figure 7: RMSD calculation of all protein atoms of both native and mutated complex. (a) RMSD values of mutated(red) and native(black) whole complex. (b) RMSD values of FLS2 from mutated(red) and native(black) complex (c) RMSD values of BAK1 from mutated(red) and native(black) complex (d) RMSD values of flg22 from mutated(red) and native(black) complex.

RMSF:

Root mean square fluctuation or RMSF value for mutated FLS2 was higher for most of the residues except for a region consisting of residue number 400 to 550 compared to the native FLS2. For BAK1, the mutated version had similar or higher RMSF than the native version with residue number 43, 83, 158, 192 having significantly higher fluctuation. RMSF value for flg22 from the mutated complex was either similar or lower for most residue compared with the native counterpart except residue 75,76,78,79,80.

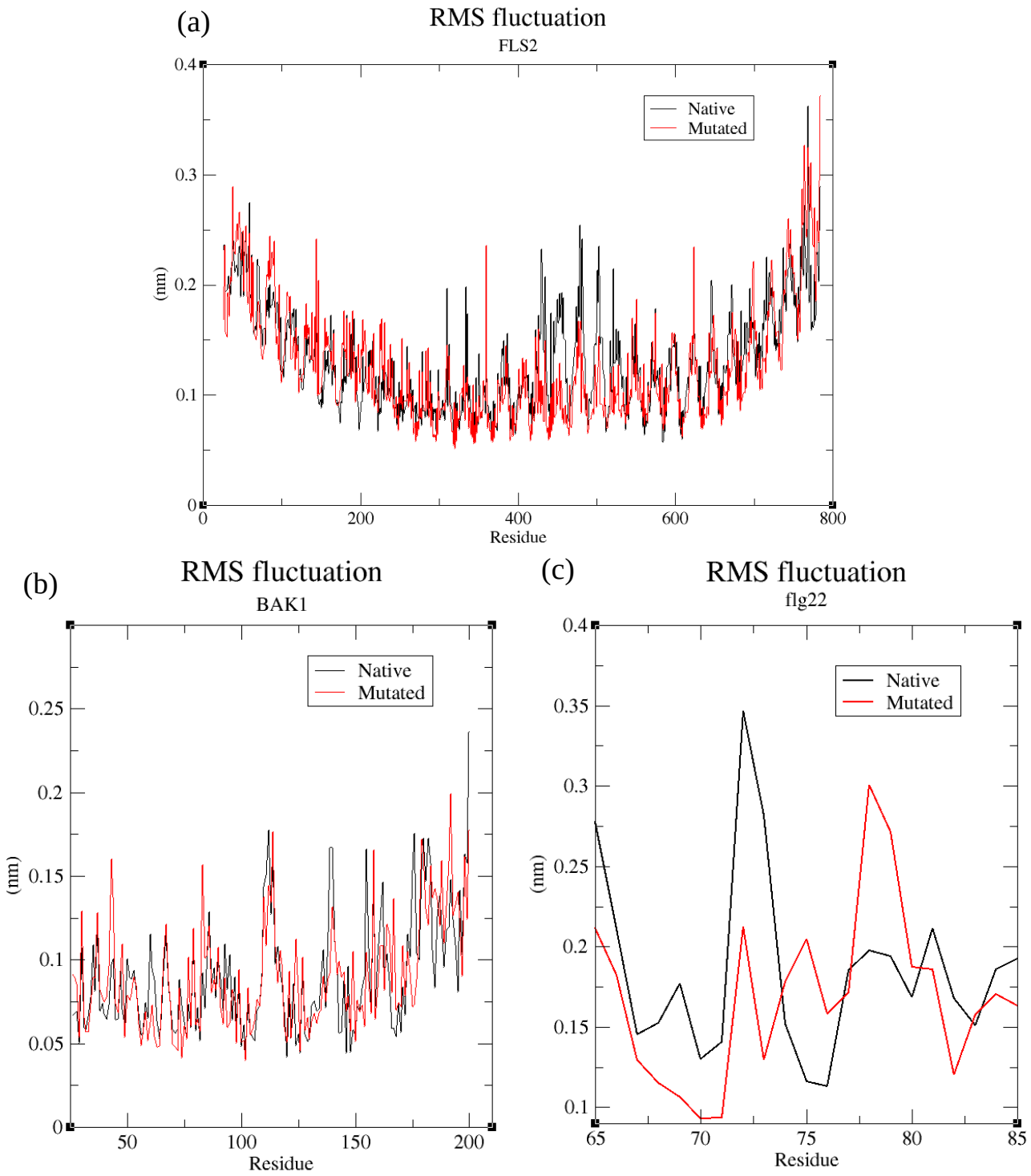

Figure 8: RMSF calculation of residues from both native and mutated complex. (a) RMSF values of FLS2 from mutated(red) and native(black) complex (b) RMSF values of BAK1 from mutated(red) and native(black) complex (c) RMSF values of flg22 from mutated(red) and native(black) complex.

Rg:

The radius of gyration or Rg for mutated whole complex stayed less than the native complex until 12ns. After that, the Rg for mutated complex maintained a higher value than the native complex. For FLS2 peptide only the Rg was slightly higher for mutated complex than native complex after 14ns. For BAK1 the radius of gyration of mutated complex deviated from the native complex after 2 ns and maintained a much higher Rg. The radius of gyration for flg22 PAMP of the mutated complex was lower than the native complex for most of the time.

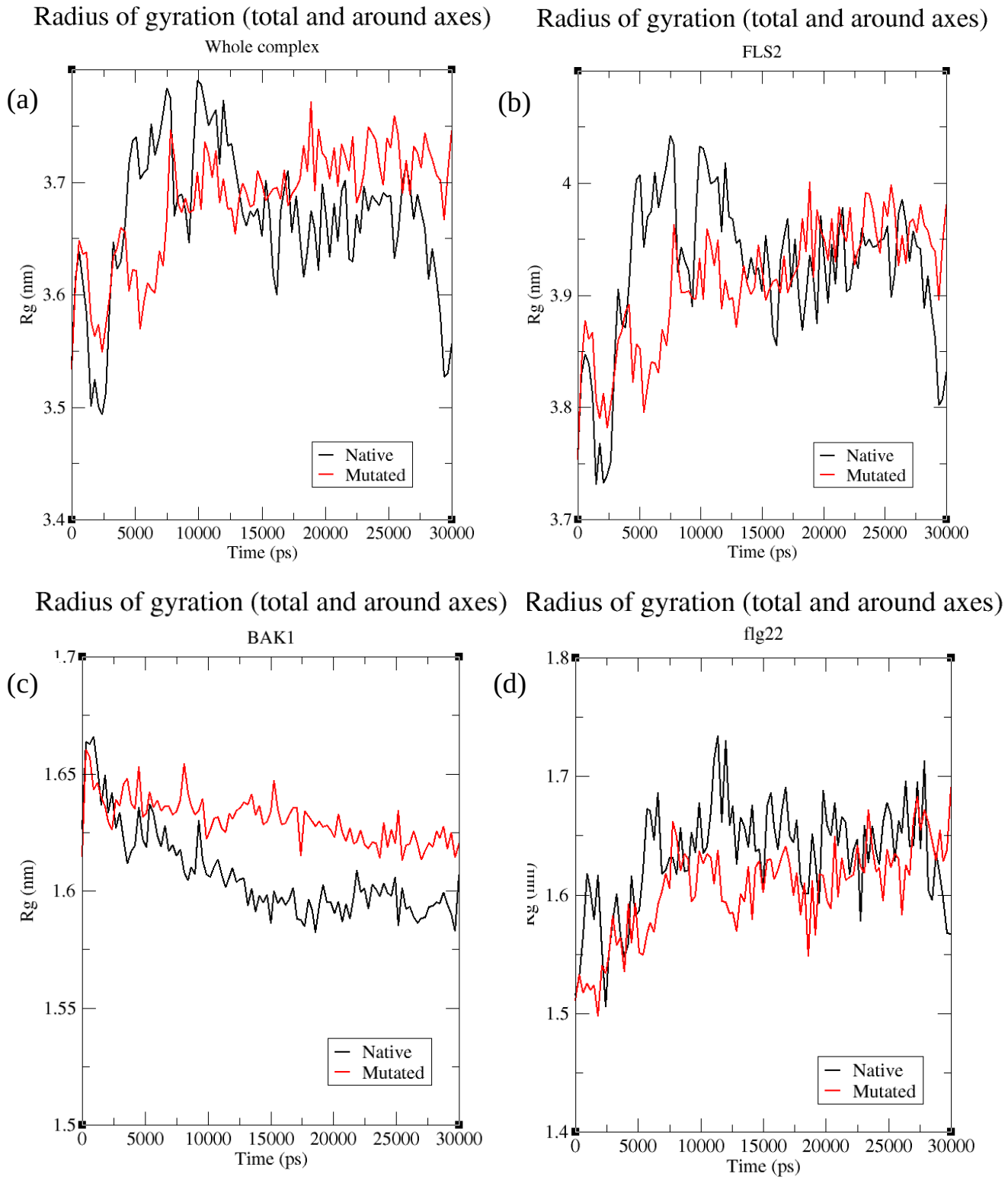

Figure 9: Rg calculation of all protein atoms of both native and mutated complex. (a) Rg values of mutated(red) and native(black) whole complex. (b) Rg values of FLS2 from mutated(red) and native(black) complex (c) Rg values of BAK1 from mutated(red) and native(black) complex (d) Rg values of flg22 from mutated(red) and native(black) complex.

Hydrogen bond:

Total Number of hydrogen bonds between FLS2 and BAK1 peptide reveals that the mutated BAK1 formed a lower number of hydrogen bonds thought the simulation compared to non-mutated BAK1-FLS2 interaction. The lowest number of the hydrogen bond between the muted version was 3 compared to 4 hydrogen bonds by native version in case of BAK1-flg22 interaction we see a similar trend. However, around 20ns all the hydrogen bonds were broken down between BAK1-flg22, whereas the native version had at least 1 hydrogen bond between those two peptides. For FLS -flg22 the mutated version also interacted with a lower number of hydrogen bonds between themselves. The native version had 9 hydrogen bonds during simulation compared to 7 hydrogen bonds in the mutated version.

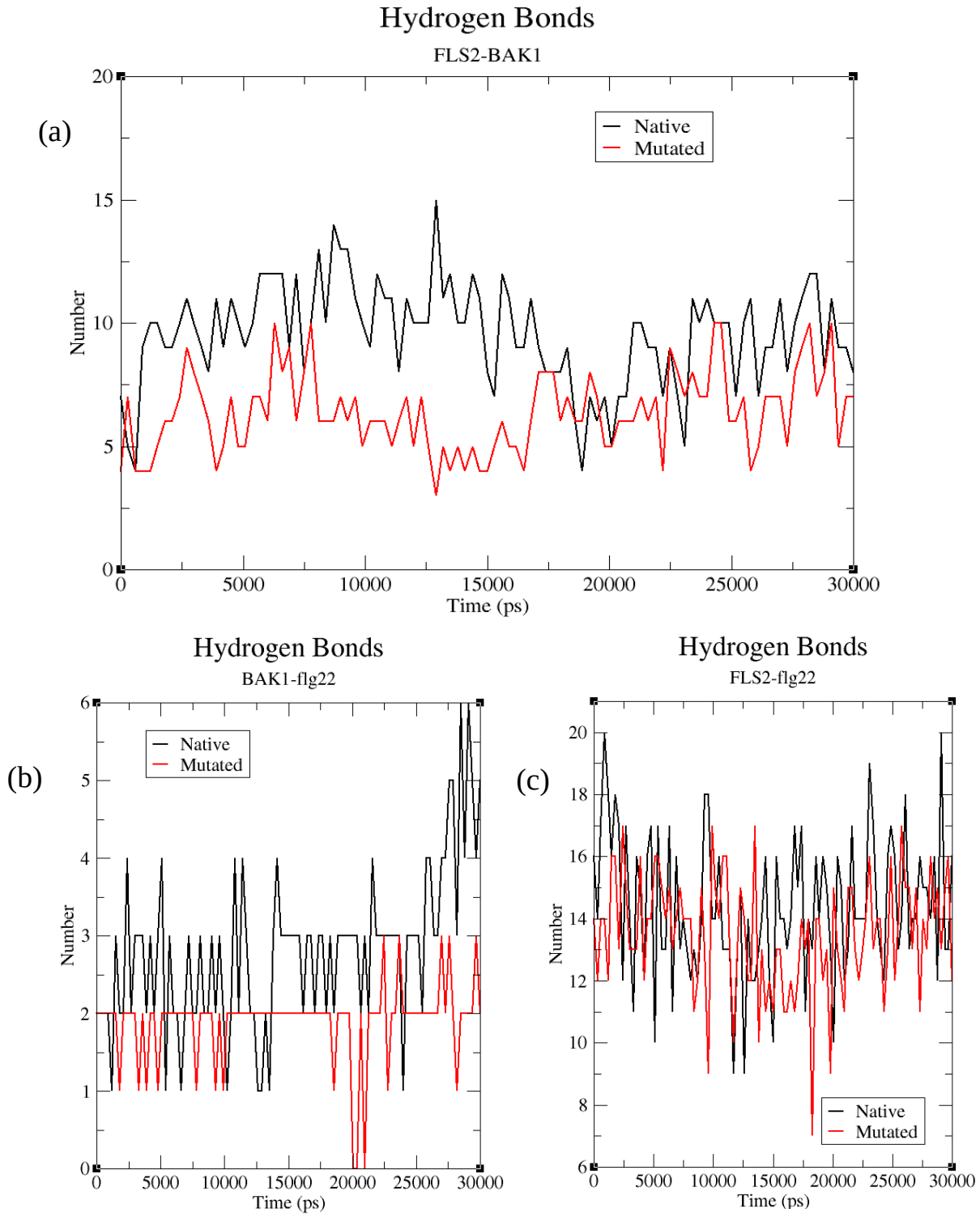

Figure 10: Hydrogen bond calculation between two peptides from both native and mutated complex. (a) Number of Hydrogen bond between FLS2-BAK1 from mutated(red) and native(black) complex (b Number of Hydrogen bond between BAK1-flg22 from mutated(red) and native(black) complex (c) Number of Hydrogen bond between FLS2-flg22 from mutated(red) and native(black) complex.

Molecular interaction of BAK1 122nd amino acid with other atoms:

The data obtained from the protein interaction calculator server shows that change in amino acid from aspartate to asparagine in the 122nd position in BAK1 causes some change in molecular interaction within the BAK1 peptide. In the native protein complex, the aspartate maintains the main chain-main chain Hydrogen bond with residue 97 and 99, Side Chain-Side Chain Hydrogen Bonds with residue 120 and 146, and Ionic Interaction with residue 146. On the other hand, mutation of ASP122 to ASN122 in BAK1 causes this amino acid to lose interaction with residue 97 and 146, rather it makes new connections with residue 98 and 100.

Table 1: Intra-protein Interaction of 122nd amino acid in BAK1 after molecular dynamics simulation

		Donor				Acceptor								
		POS	**CHAIN**	**RES**	**ATOM**	**POS**	**CHAIN**	**RES**	**ATOM**	**MO**	**Dd-a**	**Dh-a**	**A(d-H-N)**	**A(a-O=C)**
Native	Main chain main chain H-bond	99	B	LEU	N	122	B	ASP	O	-	3.08	2.18	150.07	170.02
		122	B	ASP	N	97	B	LEU	O	-	3.37	2.84	114.26	146.44
	Side Chain-Side Chain Hydrogen Bonds	120	B	SER	OG	122	B	ASP	OD1	-	-	2.63	9.99	999.99
		146	B	ARG	NH1	122	B	ASP	OD2	-	1	3.12	2.41	123.56
		146	B	ARG	NH1	122	B	ASP	OD2	-	2	3.12	3.66	49.89
	Ionic Interactions	122	B	ASP	-	146	B	ARG	-	-	-	-	-	-
Mutated	Main chain main chain H bond	99	B	LEU	N	122	B	ASN	O	-	2.81	1.91	152.94	157.99
		100	B	TYR	N	122	B	ASN	O	-	3.38	3.32	85.20	142.98
	Main Chain-Side Chain Hydrogen Bonds	122	B	ASN	N	98	B	GLU	OE1	-	-	3.31	2.52	134.14
	Side Chain-Side Chain Hydrogen Bonds	98	B	GLU	OE1	122	B	ASN	ND2	1	2.71	1.69	156.22	999.99
		98	B	GLU	OE1	122	B	ASN	ND2	2	2.71	3.43	39.38	999.99
		98	B	GLU	OE2	122	B	ASN	ND2	1	3.18	2.37	130.20	999.99
		98	B	GLU	OE2	122	B	ASN	ND2	2	3.18	4.09	24.00	999.99
		120	B	SER	OG	122	B	ASN	ND2	-	3.26	9.99	999.99	999.99
		122	B	ASN	ND2	98	B	GLU	OE1	1	2.71	3.34	46.25	999.99
		122	B	ASN	ND2	98	B	GLU	OE1	2	2.71	1.73	163.39	999.99
		122	B	ASN	ND2	98	B	GLU	OE2	1	3.18	3.66	55.58	999.99
		122	B	ASN	ND2	98	B	GLU	OE2	2	3.18	2.41	132.49	999.99

		122	B	ASN	ND2	120	B	SER	OG	1	3.26	3.14	87.29	999.99
		122	B	ASN	ND2	120	B	SER	OG	2	3.26	2.83	106.36	999.99

Dd-a = Distance Between Donor and Acceptor

Dh-a = Distance Between Hydrogen and Acceptor

A(d-H-N) = Angle Between Donor-H-N

A(a-O=C) = Angle Between Acceptor-O=C

MO = Multiple Occupancy

ASP= Aspartic acid

ASN= Asparagine

SER= Serine

TYR= Tyrosine

LEU= Leucine

ARG= Arginine

GLU= Glutamic acid

POS= Amino acid position

CHAIN= Chain identifier

RES= Residue name

Note that angles that are undefined are written as 999.99

Chapter 4

Discussion:

The mutation of aspartate to asparagine at 122nd residue in the BAK was induced in crystal structure to maintain the conformation of other amino acids to the native form. As mutating in raw amino acid sequence and then modeling and docking the proteins will induce unwanted changes to the conformation of the complex. The mutation was done in a way that ensured the least steric clash of new amino acid asparagine with surrounding molecules.

To perform the molecular dynamics simulation both, the native and mutated protein complex was placed at the center of a three-dimensional cubic box with at least 1.0 nm distance from the edge of the protein. The distance was set to satisfy the minimum image convention since we used periodic boundary condition (PBC) in the simulation. The system was then solvated with water model SPC/E which is a generic three-point water model since no biological reaction occurs without the presence of a solvent. Then the system was charge neutralized with 28 positive sodium ions as the protein complex had a net negative charge of 28 and since biomolecules do not exist at a net charge. The system was energy minimized afterward to relax the protein complex as the crystallized form of protein does not fully represent the actual state in nature due to the complex Bing treated with the crystallization process. The system was then equilibrated in two-phase with the first phase bringing the temperature of the system around 300 kelvin or room temperature. The pressure and density of the system were then equilibrated in the second phase bringing them to a running average pressure of 1.38255 ± 58.7928 bar and a running average density of 1012.31 ± 1.67151 kg/m$^{3.}$ These values were close to the reference value of 1 bar and expected density of SPC/E model 1008 kg/m^3 respectively. All in all, the system containing the protein complex was stable, equilibrated and ready for molecular dynamics simulation.

Following the completion of molecular dynamics simulation, various analysis was conducted on both mutated and native protein complex and their data were compared. First, the Root Mean Square Deviation or RMSD was calculated for both complexes. The RMSD data calculates the average deviation of the atoms compared to a reference structure at a moment of the simulation, generally, the reference structure is the first frame of the simulation or the crystallographic structure. RMSD is useful in determining the motions of the structure in the simulation trajectory in a time-dependent manner (Martínez, 2015). Here, we used the first

frame of simulation as the reference structure as the equilibrated structure was what we used to initiate the simulation rather than the unchanged crystallographic structure. The RMSD value comparison between native and mutated versions showed that initially, atoms of mutated complex deviated less than the native complex. However, after around 7.5ns both whole complex and FLS2 RMSD got higher than native counterpart. which means they deviated more from the reference structure. For BAK1 the atoms deviated less initially for mutated structure but hen after 12ns the atoms deviated less than the native then again after 23ns the mutated version deviated more than native. These fluctuations in data could be due to a small simulation period of 30ns. a simulation of 100ns would smoothen out the fluctuations. For flg22 the RMSD value indicates that the atoms of this peptide fluctuated more than native after some time. All these RMSD values indicate that the mutation changes the way some amino acids interact with each other which resulted in the interaction pattern change, which caused the atoms to deviate more than the native structure.

The Root Mean Square Fluctuation or RMSF is another parameter to determine the deviation of particular atoms or groups of atoms compared to the reference structure average over atoms. RMSF data shows averaged over each residue shows that almost all the residues of FLS2 and majority of residues of BAK1 form mutated complex fluctuate more than the native counterpart. For FLS2 the fluctuations in the amino acid region 120-240 mean it loses rigid interaction with flg22 PAMP in the N terminal as residue 148,152,272,296 of FLS2 directly contacts with flg22. Lower deviation in the residue region 400 to 550 in mutated complex indicate that these amino acids do not lose interaction with the C terminal of flg22. Higher fluctuation in the residue 627-721 or LRR23-26 in mutated complex means that these residues lost interaction with BAK1 as they directly contact BAK1 residue Arg72, Tyr96, Tyr100, Arg143, and Phe144. In the case of BAK1, the RMSF values were higher for mutate complex over native complex for the majority of residues but most importantly the fluctuations in the C terminal indicate a lower level interaction with flg22 and the fluctuations in N terminal indicates a lower level interaction with FLS2 complex. Interestingly, the flg22 had a lower fluctuation in the N terminal but had a higher fluctuation in the C terminal indicating flg22 C terminal that acts as a molecular glue between FLS2 and BAK1 was moving more than native structure probably due to loss of connection among residue in this region.

The Radius of Gyration or Rg is a measure of the compactness of protein, which is calculated by measuring the distance between the center of mass of the protein to both of its terminal end.

It measures how much a protein structure folds or unfolds. Rg values of the whole complex demonstrate that the mutated complex unfolds more than the native complex which indicates that rigid interactions losses over time and mutated protein stretches more than the native counterpart. For FLS2 the radius of gyration of the mutated complex takes some time to match the native values but after that, it increases over time and stays above the native values. In the case of BAK1 peptide, the mutated version attains higher Rg values than the native version very early and always stays over native for the rest of the simulation. It means the BAK1 mutated peptide tries to get compact but due to low atomic interaction with flg22 and FLS2, it can't get as compact as the native form. The radius of gyration for flg22 form mutated version gets an upward trend like native version but it hardly reaches the native values let alone surpasses it, which indicates that flg22 in mutated complex gets more compact as it terminal gets more close to its center of mass.

If we take a close look at the number of hydrogen bonds formed between two different complexes over the simulation period, we can get an idea about the trend of intra-protein interaction occurring. The number of hydrogen bonds formed between FLS2 and BAK demonstrates that the mutation-induced in BAK1 had a clear impact on the formation of hydrogen bonds between these two peptides. The mutated complex allowed fewer hydrogen bonds between FLS2 and BAK1 compared to the native complex. Taking into consideration of BAK1-flg22 interaction via hydrogen bond also shows the trend we saw for BAK1-FLS2. The BAK1-flg22 had no hydrogen bond interaction for a spilt time. Looking at FLS2-flg22 interaction, we also see the detrimental effect of the mutation on the interaction between these two peptides by forming hydrogen bonds.

The intra-protein interaction data of BAK1 before and after mutation from the protein interaction calculator server reveals that change in aspartate to asparagine results into some major change in the interaction of 122nd amino acid with other residues. The previous study has found that ASP122 of BAK1 does not make any direct contact with FLS2 and flg22 (Sun et al., 2013). However, it interacts with other important residues within BAK1 such as ARG146 that directly interact with FLS2 amino acid 674. The mutation causes the interaction between 122nd and 146th to dissolve which hampers the interaction of BAK1 with FLS2. As a result, the overall interaction between BAK1 and FLS2 ectodomain lowers which in turn hampers the induction of immune activation by flg22.

Taking into account of all the data generated via molecular dynamics simulation, RMSD, RMSF, Rg, hydrogen bond interaction, we can see that the mutation had a harmful effect on the overall complex formed by FLS2-flg22-BAK1. The mutation lowers the integrity of the complex, which in turn make the formation of heterodimer of FLS2-BAK1 via flg22 induction less likely. As a result, the mutation undermines the immune system of *Arabidopsis thaliana*.

Conclusion:

Elongated (elg) phenotype of *Arabidopsis thaliana* is caused by D122N mutation in the BAK1 coding protein which causes the plant to have elongated stem and early flowering among other phenotypes. BAK1 protein not only play role in cell development but also takes part in plant immunity by forming heterodimer with FLS2 plant PRR that is induced by bacterial flagellin flg22. Here we observed that the elg phenotype causing mutation has a negative effect on the plant immunity as the mutation disrupts key atomic interaction between receptor and co-receptor. Consequently, the likelihood of heterodimer formation between FLS2 and BAK1 reduces and subsequent activation of signal transduction leading immune response might get halted.

References:

Abraham, M.J., Murtola, T., Schulz, R., Páll, S., Smith, J.C., Hess, B., Lindahl, E., 2015. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 1–2, 19–25. https://doi.org/10.1016/j.softx.2015.06.001

Albert, M., Jehle, A.K., Lipschis, M., Mueller, K., Zeng, Y., Felix, G., 2010. Regulation of cell behaviour by plant receptor kinases: Pattern recognition receptors as prototypical models. Eur. J. Cell Biol. 89, 200–207. https://doi.org/10.1016/j.ejcb.2009.11.015

Betz, R., 2017. Dabble. https://doi.org/10.5281/zenodo.836914

Boller, T., Felix, G., 2009. A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. Annu Rev Plant Biol 60, 379–406. https://doi.org/10.1146/annurev.arplant.57.032905.105346

Boutrot, F., Zipfel, C., 2017. Function, Discovery, and Exploitation of Plant Pattern Recognition Receptors for Broad-Spectrum Disease Resistance. Annual Review of Phytopathology 55, 257–286. https://doi.org/10.1146/annurev-phyto-080614-120106

Bowers, K.J., Chow, D.E., Xu, H., Dror, R.O., Eastwood, M.P., Gregersen, B.A., Klepeis, J.L., Kolossvary, I., Moraes, M.A., Sacerdoti, F.D., Salmon, J.K., Shan, Y., Shaw, D.E., 2006. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters, in: ACM/IEEE SC 2006 Conference (SC'06). Presented at the SC 2006 Proceedings Supercomputing 2006, IEEE, Tampa, FL, pp. 43–43. https://doi.org/10.1109/SC.2006.54

Brooks, B.R., Brooks, C.L., Mackerell, A.D., Nilsson, L., Petrella, R.J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A.R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R.W., Post, C.B., Pu, J.Z., Schaefer, M., Tidor, B., Venable, R.M., Woodcock, H.L., Wu, X., Yang, W., York, D.M., Karplus, M., 2009. CHARMM: the biomolecular simulation program. J Comput Chem 30, 1545–1614. https://doi.org/10.1002/jcc.21287

Case, D.A., Cheatham, T.E., Darden, T., Gohlke, H., Luo, R., Merz, K.M., Onufriev, A., Simmerling, C., Wang, B., Woods, R.J., 2005. The Amber biomolecular simulation programs. J Comput Chem 26, 1668–1688. https://doi.org/10.1002/jcc.20290

Chinchilla, D., Zipfel, C., Robatzek, S., Kemmerling, B., Nürnberger, T., Jones, J.D.G., Felix, G., Boller, T., 2007. A flagellin-induced complex of the receptor FLS2 and

BAK1 initiates plant defence. Nature 448, 497–500. https://doi.org/10.1038/nature05999

Couto, D., Zipfel, C., 2016. Regulation of pattern recognition receptor signalling in plants. Nat. Rev. Immunol. 16, 537–552. https://doi.org/10.1038/nri.2016.77

Ding, S.-W., 2010. RNA-based antiviral immunity. Nat. Rev. Immunol. 10, 632–644. https://doi.org/10.1038/nri2824

Ding, Y., Sun, T., Ao, K., Peng, Y., Zhang, Yaxi, Li, X., Zhang, Yuelin, 2018. Opposite Roles of Salicylic Acid Receptors NPR1 and NPR3/NPR4 in Transcriptional Regulation of Plant Immunity. Cell 173, 1454-1467.e15. https://doi.org/10.1016/j.cell.2018.03.044

Eastman, P., Swails, J., Chodera, J.D., McGibbon, R.T., Zhao, Y., Beauchamp, K.A., Wang, L.-P., Simmonett, A.C., Harrigan, M.P., Stern, C.D., Wiewiora, R.P., Brooks, B.R., Pande, V.S., 2017. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. PLoS Comput. Biol. 13, e1005659. https://doi.org/10.1371/journal.pcbi.1005659

Felix, G., Duran, J.D., Volko, S., Boller, T., 1999. Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. Plant J. 18, 265–276. https://doi.org/10.1046/j.1365-313x.1999.00265.x

Fu, Z.Q., Dong, X., 2013. Systemic acquired resistance: turning local infection into global defense. Annu Rev Plant Biol 64, 839–863. https://doi.org/10.1146/annurev-arplant-042811-105606

Gómez-Gómez, L., Boller, T., 2000. FLS2: An LRR Receptor–like Kinase Involved in the Perception of the Bacterial Elicitor Flagellin in Arabidopsis. Molecular Cell 5, 1003–1011. https://doi.org/10.1016/S1097-2765(00)80265-8

Haefele, D.M., Lindow, S.E., 1987. Flagellar Motility Confers Epiphytic Fitness Advantages upon Pseudomonas syringae. Appl. Environ. Microbiol. 53, 2528–2533.

Han, G.-Z., 2019. Origin and evolution of the plant immune system. New Phytologist 222, 70–83. https://doi.org/10.1111/nph.15596

Hann, D.R., Rathjen, J.P., 2007. Early events in the pathogenicity of Pseudomonas syringae on Nicotiana benthamiana. Plant J. 49, 607–618. https://doi.org/10.1111/j.1365-313X.2006.02981.x

Harder, E., Damm, W., Maple, J., Wu, C., Reboul, M., Xiang, J.Y., Wang, L., Lupyan, D., Dahlgren, M.K., Knight, J.L., Kaus, J.W., Cerutti, D.S., Krilov, G., Jorgensen, W.L., Abel, R., Friesner, R.A., 2016. OPLS3: A Force Field Providing Broad Coverage of

Drug-like Small Molecules and Proteins. J. Chem. Theory Comput. 12, 281–296. https://doi.org/10.1021/acs.jctc.5b00864

He, K., Gou, X., Yuan, T., Lin, H., Asami, T., Yoshida, S., Russell, S.D., Li, J., 2007. BAK1 and BKK1 regulate brassinosteroid-dependent growth and brassinosteroid-independent cell-death pathways. Curr. Biol. 17, 1109–1115. https://doi.org/10.1016/j.cub.2007.05.036

Hecht, V., Vielle-Calzada, J.P., Hartog, M.V., Schmidt, E.D., Boutilier, K., Grossniklaus, U., de Vries, S.C., 2001. The Arabidopsis SOMATIC EMBRYOGENESIS RECEPTOR KINASE 1 gene is expressed in developing ovules and embryos and enhances embryogenic competence in culture. Plant Physiol. 127, 803–816.

Heese, A., Hann, D.R., Gimenez-Ibanez, S., Jones, A.M.E., He, K., Li, J., Schroeder, J.I., Peck, S.C., Rathjen, J.P., 2007. The receptor-like kinase SERK3/BAK1 is a central regulator of innate immunity in plants. Proc. Natl. Acad. Sci. U.S.A. 104, 12217–12222. https://doi.org/10.1073/pnas.0705306104

Hollingsworth, S.A., Dror, R.O., 2018. Molecular Dynamics Simulation for All. Neuron 99, 1129–1143. https://doi.org/10.1016/j.neuron.2018.08.011

Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., de Groot, B.L., Grubmüller, H., MacKerell, A.D., 2017. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. Nat. Methods 14, 71–73. https://doi.org/10.1038/nmeth.4067

Jaillais, Y., Belkhadir, Y., Balsemão-Pires, E., Dangl, J.L., Chory, J., 2011. Extracellular leucine-rich repeats as a platform for receptor/coreceptor complex formation. Proc Natl Acad Sci U S A 108, 8503–8507. https://doi.org/10.1073/pnas.1103556108

Jo, S., Kim, T., Iyer, V.G., Im, W., 2008. CHARMM-GUI: a web-based graphical user interface for CHARMM. J Comput Chem 29, 1859–1865. https://doi.org/10.1002/jcc.20945

Jones, J.D.G., Dangl, J.L., 2006. The plant immune system. Nature 444, 323–329. https://doi.org/10.1038/nature05286

Karplus, M., McCammon, J.A., 2002. Molecular dynamics simulations of biomolecules. Nat. Struct. Biol. 9, 646–652. https://doi.org/10.1038/nsb0902-646

Kemmerling, B., Nürnberger, T., 2008. Brassinosteroid-independent functions of the BRI1-associated kinase BAK1/SERK3. Plant Signal Behav 3, 116–118. https://doi.org/10.4161/psb.3.2.4981

Kemmerling, B., Schwedt, A., Rodriguez, P., Mazzotta, S., Frank, M., Qamar, S.A., Mengiste, T., Betsuyaku, S., Parker, J.E., Müssig, C., Thomma, B.P.H.J., Albrecht, C., de Vries, S.C., Hirt, H., Nürnberger, T., 2007. The BRI1-associated kinase 1, BAK1, has a brassinolide-independent role in plant cell-death control. Curr. Biol. 17, 1116–1122. https://doi.org/10.1016/j.cub.2007.05.046

Komoriya, K., Shibano, N., Higano, T., Azuma, N., Yamaguchi, S., Aizawa, S.I., 1999. Flagellar proteins and type III-exported virulence factors are the predominant proteins secreted into the culture media of Salmonella typhimurium. Mol. Microbiol. 34, 767–779. https://doi.org/10.1046/j.1365-2958.1999.01639.x

Kourelis, J., Hoorn, R.A.L. van der, 2018. Defended to the Nines: 25 Years of Resistance Gene Cloning Identifies Nine Mechanisms for R Protein Function. The Plant Cell 30, 285–299. https://doi.org/10.1105/tpc.17.00579

Li, J., Wen, J., Lease, K.A., Doke, J.T., Tax, F.E., Walker, J.C., 2002. BAK1, an Arabidopsis LRR receptor-like protein kinase, interacts with BRI1 and modulates brassinosteroid signaling. Cell 110, 213–222. https://doi.org/10.1016/s0092-8674(02)00812-7

Liang, X., Ding, P., Lian, K., Wang, J., Ma, M., Li, Lin, Li, Lei, Li, M., Zhang, X., Chen, S., Zhang, Y., Zhou, J.-M., 2016. Arabidopsis heterotrimeric G proteins regulate immunity by directly coupling to the FLS2 receptor. Elife 5, e13568. https://doi.org/10.7554/eLife.13568

Lu, D., Wu, S., Gao, X., Zhang, Y., Shan, L., He, P., 2010. A receptor-like cytoplasmic kinase, BIK1, associates with a flagellin receptor complex to initiate plant innate immunity. Proc. Natl. Acad. Sci. U.S.A. 107, 496–501. https://doi.org/10.1073/pnas.0909705107

Martínez, L., 2015. Automatic Identification of Mobile and Rigid Substructures in Molecular Dynamics Simulations and Fractional Structural Fluctuation Analysis. PLoS One 10. https://doi.org/10.1371/journal.pone.0119264

McAndrew, R., Pruitt, R.N., Kamita, S.G., Pereira, J.H., Majumdar, D., Hammock, B.D., Adams, P.D., Ronald, P.C., 2014. Structure of the OsSERK2 leucine-rich repeat extracellular domain. Acta Crystallogr. D Biol. Crystallogr. 70, 3080–3086. https://doi.org/10.1107/S1399004714021178

Nam, K.H., Li, J., 2002. BRI1/BAK1, a receptor kinase pair mediating brassinosteroid signaling. Cell 110, 203–212. https://doi.org/10.1016/s0092-8674(02)00814-0

O'Brien, E.J., Bennett, P.M., 1972. Structure of straight flagella from a mutant Salmonella. J. Mol. Biol. 70, 133–152. https://doi.org/10.1016/0022-2836(72)90168-4

Pfund, C., Tans-Kersten, J., Dunning, F.M., Alonso, J.M., Ecker, J.R., Allen, C., Bent, A.F., 2004. Flagellin is not a major defense elicitor in Ralstonia solanacearum cells or extracts applied to Arabidopsis thaliana. Mol. Plant Microbe Interact. 17, 696–706. https://doi.org/10.1094/MPMI.2004.17.6.696

Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kalé, L., Schulten, K., 2005. Scalable molecular dynamics with NAMD. J Comput Chem 26, 1781–1802. https://doi.org/10.1002/jcc.20289

Pumplin, N., Voinnet, O., 2013. RNA silencing suppression by plant pathogens: defence, counter-defence and counter-counter-defence. Nat. Rev. Microbiol. 11, 745–760. https://doi.org/10.1038/nrmicro3120

Ranf, S., 2017. Sensing of molecular patterns through cell surface immune receptors. Curr. Opin. Plant Biol. 38, 68–77. https://doi.org/10.1016/j.pbi.2017.04.011

Robatzek, S., Bittel, P., Chinchilla, D., Köchner, P., Felix, G., Shiu, S.-H., Boller, T., 2007. Molecular identification and characterization of the tomato flagellin receptor LeFLS2, an orthologue of Arabidopsis FLS2 exhibiting characteristically different perception specificities. Plant Molecular Biology 64, 539–547. https://doi.org/10.1007/s11103-007-9173-8

Robustelli, P., Piana, S., Shaw, D.E., 2018. Developing a molecular dynamics force field for both folded and disordered protein states. PNAS 115, E4758–E4766. https://doi.org/10.1073/pnas.1800690115

Samatey, F.A., Imada, K., Nagashima, S., Vonderviszt, F., Kumasaka, T., Yamamoto, M., Namba, K., 2001. Structure of the bacterial flagellar protofilament and implications for a switch for supercoiling. Nature 410, 331–337. https://doi.org/10.1038/35066504

Sastry, G.M., Adzhigirey, M., Day, T., Annabhimoju, R., Sherman, W., 2013. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. J. Comput. Aided Mol. Des. 27, 221–234. https://doi.org/10.1007/s10822-013-9644-8

Schwessinger, B., Roux, M., Kadota, Y., Ntoukakis, V., Sklenar, J., Jones, A., Zipfel, C., 2011. Phosphorylation-Dependent Differential Regulation of Plant Growth, Cell Death, and Innate Immunity by the Regulatory Receptor-Like Kinase BAK1. PLOS Genetics 7, e1002046. https://doi.org/10.1371/journal.pgen.1002046

Shiu, S.H., Bleecker, A.B., 2001. Plant receptor-like kinase gene family: diversity, function, and signaling. Sci. STKE 2001, re22. https://doi.org/10.1126/stke.2001.113.re22

Spoel, S.H., Dong, X., 2012. How do plants achieve immunity? Defence without specialized immune cells. Nat. Rev. Immunol. 12, 89–100. https://doi.org/10.1038/nri3141

Sun, W., Cao, Y., Jansen Labby, K., Bittel, P., Boller, T., Bent, A.F., 2012. Probing the Arabidopsis flagellin receptor: FLS2-FLS2 association and the contributions of specific domains to signaling function. Plant Cell 24, 1096–1113. https://doi.org/10.1105/tpc.112.095919

Sun, W., Dunning, F.M., Pfund, C., Weingarten, R., Bent, A.F., 2006. Within-species flagellin polymorphism in Xanthomonas campestris pv campestris and its impact on elicitation of Arabidopsis FLAGELLIN SENSING2-dependent defenses. Plant Cell 18, 764–779. https://doi.org/10.1105/tpc.105.037648

Sun, Y., Li, L., Macho, A.P., Han, Z., Hu, Z., Zipfel, C., Zhou, J.-M., Chai, J., 2013. Structural Basis for flg22-Induced Activation of the Arabidopsis FLS2-BAK1 Immune Complex. Science 342, 624–628. https://doi.org/10.1126/science.1243825

Takai, R., Isogai, A., Takayama, S., Che, F.-S., 2008. Analysis of flagellin perception mediated by flg22 receptor OsFLS2 in rice. Mol. Plant Microbe Interact. 21, 1635–1642. https://doi.org/10.1094/MPMI-21-12-1635

Tang, D., Wang, G., Zhou, J.-M., 2017. Receptor Kinases in Plant-Pathogen Interactions: More Than Pattern Recognition. Plant Cell 29, 618–637. https://doi.org/10.1105/tpc.16.00891

Thomma, B.P.H.J., Nürnberger, T., Joosten, M.H.A.J., 2011. Of PAMPs and effectors: the blurred PTI-ETI dichotomy. Plant Cell 23, 4–15. https://doi.org/10.1105/tpc.110.082602

Upson, J.L., Zess, E.K., Białas, A., Wu, C.-H., Kamoun, S., 2018. The coming of age of EvoMPMI: evolutionary molecular plant-microbe interactions across multiple timescales. Curr. Opin. Plant Biol. 44, 108–116. https://doi.org/10.1016/j.pbi.2018.03.003

Vert, G., 2008. Plant signaling: brassinosteroids, immunity and effectors are BAK ! Curr. Biol. 18, R963-965. https://doi.org/10.1016/j.cub.2008.09.006

Wang, W., Feng, B., Zhou, J.-M., Tang, D., n.d. Plant immune signaling: Advancing on two frontiers. Journal of Integrative Plant Biology n/a. https://doi.org/10.1111/jipb.12898

Whippo, C.W., Hangarter, R.P., 2005. A brassinosteroid-hypersensitive mutant of BAK1 indicates that a convergence of photomorphogenic and hormonal signaling modulates phototropism. Plant Physiol. 139, 448–457. https://doi.org/10.1104/pp.105.064444

Wu, Y., Zhou, J.-M., 2013. Receptor-like kinases in plant innate immunity. J Integr Plant Biol 55, 1271–1286. https://doi.org/10.1111/jipb.12123

Yamada, K., Yamaguchi, K., Shirakawa, T., Nakagami, H., Mine, A., Ishikawa, K., Fujiwara, M., Narusaka, M., Narusaka, Y., Ichimura, K., Kobayashi, Y., Matsui, H., Nomura, Y., Nomoto, M., Tada, Y., Fukao, Y., Fukamizo, T., Tsuda, K., Shirasu, K., Shibuya, N., Kawasaki, T., 2016. The Arabidopsis CERK1-associated kinase PBL27 connects chitin perception to MAPK activation. EMBO J. 35, 2468–2483. https://doi.org/10.15252/embj.201694248

Yonekura, K., Maki-Yonekura, S., Namba, K., 2003. Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. Nature 424, 643–650. https://doi.org/10.1038/nature01830

Yu, X., Feng, B., He, P., Shan, L., 2017. From Chaos to Harmony: Responses and Signaling upon Microbial Pattern Recognition. Annu Rev Phytopathol 55, 109–137. https://doi.org/10.1146/annurev-phyto-080516-035649

Zhang, J., Li, W., Xiang, T., Liu, Z., Laluk, K., Ding, X., Zou, Y., Gao, M., Zhang, X., Chen, S., Mengiste, T., Zhang, Y., Zhou, J.-M., 2010. Receptor-like cytoplasmic kinases integrate signaling from multiple plant immune receptors and are targeted by a Pseudomonas syringae effector. Cell Host Microbe 7, 290–301. https://doi.org/10.1016/j.chom.2010.03.007

Zipfel, C., 2014. Plant pattern-recognition receptors. Trends in Immunology 35, 345–351. https://doi.org/10.1016/j.it.2014.05.004

Zipfel, C., Oldroyd, G.E.D., 2017. Plant signalling in symbiosis and immunity. Nature 543, 328–336. https://doi.org/10.1038/nature22009