Structural dynamics of PRR FLS2, PAMP flg22 and mutated co-receptor BAK1 of *Arabidopsis thaliana*

By

Nazia Nowrid
Student ID:18376004

A thesis submitted to the Department of Mathematics and Natural Sciences in partial fulfillment of the requirements for the degree of
Master of Science in Biotechnology

Department of Mathematics and Natural Sciences
Brac University
March, 2021

Dedicated To

My parents, for their unfaltering guidance, care and support.

Declaration

Declaration

It is hereby declared that

1. The thesis submitted is my/our 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:

_____

Nazia Nowrid
18376004

Approval

The thesis titled "Structural dynamics of PRR FLS2, PAMP flg22 and mutated co-receptor BAK1 of *Arabidopsis thaliana*" submitted by Nazia Nowrid (18376004) has been accepted as satisfactory in partial fulfillment of the requirement for the degree of Master of Science in Biotechnology on March 28, 2021.

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.

Acknowledgement

Firstly, I would like to express my utmost gratitude to my respected supervisor Mr. M H M Mubassir, Lecturer, Biotechnology program, Department of Mathematics and Natural Sciences, BRAC University, for his extensive support, skillful guidance and keen encouragement during the entire timeline of my research work, which significantly aided me 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 Biotechnology laboratory.

Besides, I would like to convey my unreserved gratitude to Professor Dr. Aparna Islam, Department of Mathematics and Natural Sciences, BRAC University for all the inspiration and guidance she provided me as a student. I would like to thank Professor Dr. Iftekhar Bin Naser for his unending encouragement and motivation.

Last but not the least, I am indebted to my parents and my husband, for their unwavering care, encouragement and moral support at every stage during this study.

Nazia Nowrid

Abstract

Pattern-triggered immunity (PTI) acts as the first layer of defense that is deployed by plants in order to avert microbial invasions. It is actually identified by the activities of pattern-recognition receptors (PRRs), that bind with the pathogen-associated molecular patterns (PAMPs) and recruit co-receptor protein(s) to convey a defense signal, thereby initiating the plant's immunity. In this study, the 122nd amino acid in BAK1 from the crystallographic structure of FLS2-BAK1-flg22 complex (PDB ID: 4MN8) was mutated from aspartate to asparagine using *in-silico* method. Molecular dynamics simulations (100 ns) and MM/PBSA calculations were then applied for a thorough and comprehensive analysis of the interactions between PRR FLS2, mutated co-receptor mBAK1 and PAMP flg22. ARG-72 and LYS-36 from FLS2-mBAK1 (in the presence and absence of flg22 respectively), ARG-72 and GLU-28 from mBAK1-flg22 (in the presence and absence of FLS2 respectively) & LYS-77 and ASP-176 from FLS2-flg22 (in the presence and absence of mBAK1 respectively) were revealed to be the most prominent residues, aiding notably in the process of heterodimerization during PTI, subsequently mediated by FLS2. A hypothesis can thereby be established, that mutation at any of these residues will affect the PTI of *Arabidopsis thaliana.* The simulations were also compared using parameters such as root mean square deviation (RMSD), root mean square fluctuation (RMSF), the radius of gyration (Rg), solvent accessible surface areas (SASA) and the number of hydrogen bonds (H-bond) to comprehend the structural integrity of the complex. The obtained results demonstrated that PAMP flg22 interacted much more favorably with the PRR FLS2 in the presence of mutated co-receptor mBAK1 in the complex, which implicates the necessity of the co-receptor in FLS2 mediated PTI to recognize PAMP flg22. Furthermore, since FLS2 has been shown to play a key role in *Arabidopsis thaliana* plant defense mechanism, its assumed binding mechanism with PAMP and co-receptor BAK1 will help paint an improved comprehension of the inceptive stages of PTI.

Table of Contents

List of Figures

List of Tables

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 - Nanometer

ns - Nanosecond

ps - Picosecond

Chapter 1

Introduction:

Plant immune system:

A number of different strategies are adopted by pathogenic organisms attacking plants, in order to infect them and subsequently debilitate the plant growth and maintenance. Lacking an adaptive immune system comprising of moving immune cells, plants are mostly dependent on innate immune system regulations for being able to discern and cease pathogenic infections (Boller and Felix, 2009; Spoel and Dong, 2012; Thomma et al., 2011). The immunological system of plants is generally composed of two interlinked tiers – The first one recognizes the microbe-associated molecular patterns (MAMPs) or the pathogen associated molecular pattern (PAMPs), by utilizing the pattern-recognition receptors present on the cell surfaces. The MAMPs and PAMPs are molecular markers present in varying groups of microorganisms and host-derived damage- associated molecular patterns (DAMPs) (Boller and Felix, 2009). The second one, on the other hand, adequately responds to the effector molecules secreted by pathogens, in order to initiate infections and subdue plant immunity, upon utilization of disease resistance (R) proteins (Upson et al., 2018).

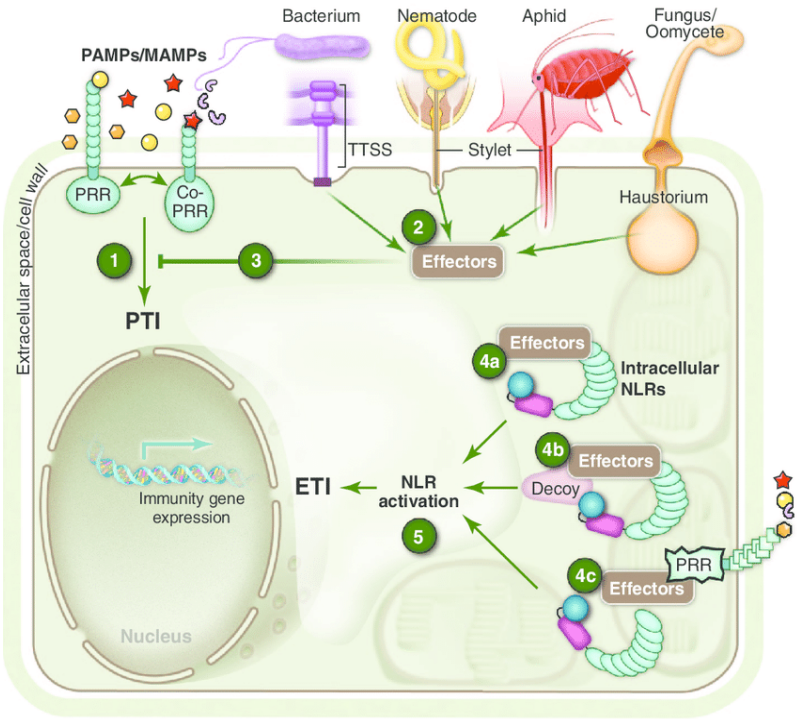


Figure 1: A brief summary of the plant immune system (adapted from (Dangl et al. 2013).

Pattern-triggered immunity (PTI) response:

When different pattern recognition receptors (PRRs) detect PAMPs or DAMPs as a part of plant's innate defense mechanism against attacking and invading pathogens, the pattern-triggered immunity (PTI) simultaneously gets activated (Jones and Dangl, 2006; Zipfel, 2014). PRRs can be categorized into membrane-bound PRRs and cytoplasmic PRRs, depending on their position inside the cell (Gomez and Bowler, 2000; Li and Chory, 1997; Zipfel et al., 2006, Gust and Felix, 2014; Liebrand et al., 2014). Membrane-bound PRRs can be again sub-divided into receptor-like proteins (RLPs) and receptor-like kinases (RLKs) (Boller and Felix, 2009). Some crucial roles played by some members of these protein groups include – growth and development of plants, symbiosis and conferring protection from environmental stresses, which may be abiotic in nature (Tang et al., 2017).

A higher number of both RLPs and RLKs are generally present inside plants, as opposed to animals (Shiu and Bleecker, 2001) and often, there are more RLKs than RLPs. RLKs are known to play a range of crucial roles, and among these, the four most imperative functions include – discerning peptides or ligands as part of extracellular signals, implementing signals that are downstream in nature, predicting the intercommunication sites of proteins and producing ample networks for easily conducting signaling activities (Hohmann et al., 2017).

As PRRs constitute a wide range of extracellular domains, such as - lysine motif (LysM), leucine-rich repeat (LRR), lectin and epidermal growth factor (EGF)-like domains (Yu et al., 2017), they are able to identify different types of ligands. Defense against microbial invasion and the overall counterresponse initiated by PTI can often, get compromised due to existent mutations in PRRs. A relevant example for this would be – In the absence of FLS2 in Arabidopsis, a number of vital functions of the plant can get hindered. These typically include: faulty ROS accumulation upon induction by flg22, stimulation of MAPK and expression of genes involved in defense mechanisms (Asai et al., 2002; Gómez-Gómez et al., 1999). Again, a novel pair of PRR-PAMP interaction has been defined, owing to the discovery of XA21-Ax21. This has duly indicated that the adapted pathogen can be aptly impeded when PTI defenses get activated. Current studies have also proven that PTI creates nutrient deprivation to control colonization of microbes, while simultaneously restricting pathogenic growth (Ranf et al., 2017; Xin et al., 2016).

Effector-triggered immunity (ETI) response:

The second of the two layered defense mechanism of plants, to ward off bacterial and fungal pathogens is known as the "Effector-triggered immunity" (ETI). During infection, pathogens secrete a diverse range of virulence factors known as "effector proteins" – that subvert a number of cellular mechanisms taking place inside the host, including disruption of the cytoskeletal machinery, hindering translation and inhibiting the responses of the immune system (Mattoo et al., 2007; Ribet and Cossart, 2010). Epithelial cells are often faced with substantial exposure to a wide array of micro-organisms, and being a category of non-professional immune cells, are highly dependent on ETI for prompt and adequate response against invasive pathogens. Contrary to this, professional immune cells such as, macrophages typically inhabit inside tissues and are less reliant on ETI. They are also efficient at combatting all types of microbial invasion on the tissues, regardless of the nature of attack: pathogenic or not.

In ETI, the effector protein(s) of the pathogen are recognized by a Resistance (R) gene product, typically present internally in plant cells (Jones and Dangl, 2006; Tsuda and Katagiri, 2010). Most of the R genes carry nucleotide-binding leucine-rich repeat (NB-LRR) proteins (Tsuda and Katagiri, 2010; Collier and Moffett, 2009). ETI is prompted upon the identification of an effector by these NB-LRR proteins. 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). The ETI is therefore, initiated upon the perception of effectors (Spoel and Dong, 2012).

After ETI has started, programmed cell death is prompted at the position/ location of the infection. The subsequent result of this includes the prevention of pathogens to spread (Jones and Dangl, 2006). ETI is also responsible for the production of glycerol-3-phosphate (G3P), methyl salicylic acid (MeSA) and azelaic acid – which cumulatively acts as mobile immune signals in plants, that are consequently transported to other uninfected tissues from the infection site (Fu and Dong, 2013). This immunological process that has been initiated is known as systemic acquired resistance (SAR), that eventually causes the manufacturing of pathogenesis-related (PR) proteins for carrying out activities pertaining to the inhibition of microbes. At the same time, it confers immunity to plants from successive assaults of pathogens (Spoel and Dong, 2012).

RNA interference:

The double-stranded RNA-mediated interference (RNAi) is responsible for promptly silencing the expression of genes in different classes of organisms. RNA is broken down into short RNAs that stimulate ribonucleases for pointing out/ targeting the homologous mRNA – which ultimately causes the silencing of genes. RNAi also confers multifarious benefits, greatly aiding in research. For example, it permits quick categorization of the functions of known genes, alongside the recognition of new genes responsible for causing disease. This has significantly bolstered and fortified functional genomics as a whole (Mocellin and Provenzano, 2004).

Infectious diseases caused by viruses are fended off adequately and efficiently by the process of RNA silencing (Ding, 2010). This typically occurs in two steps: dsRNA are broken down into small interfering RNAs (siRNAs) in the first step, by an activity similar to RNase III. RISC (RNA-induced silencing complex) is formed in the second step, upon the combination of the siRNAs inside an RNase complex; this in turn takes part in the degradation of cognate mRNA by acting on it. The cellular components playing crucial roles in RNAi are generally the Dicer, RNA-dependent RNA polymerase, helicases, and dsRNA endonucleases (Agrawal et al., 2003).

FLS2 mediated pattern triggered immunity:

The leucine-rich repeat (LRR) receptor kinase Flagellin Sensing 2 (FLS2) in Arabidopsis is a plant PRR that has been studied extensively. First discovered in *Arabidopsis Thaliana* (Gómez-Gómez and Boller, 2000), it was later found in other plant species such as - rice (Takai et al., 2008), tobacco (Hann and Rathjen, 2007) and tomato (Robatzek et al., 2007).

A conserved N-terminal 22-amino acid sequence (flg22) of bacterial flagellin is generally recognized by FLS2 (Gómez- Gómez and Boller, 2000), thereby initiating the PTI (Boller and Felix, 2009; Zipfel et al., 2004; Zipfel et al., 2006). Typically, FLS2 consists of three major domains: an extracellular, a helical transmembrane and an intracellular domain. 28 leucine-rich-repeats (LRR) are contained in the extracellular domain where the interaction with flagellin occurs, whereas buried in the phospholipid layer of the cell membrane is the

transmembrane region The serine/threonine kinase domain makes up the intracellular domain of FLS2.

FLS2 forms a complex with an intracellular kinase BIK1 (Botrytis-induced kinase 1), when not faced with an infection (Sun et al., 2013; Lu et al., 2010; Zhang et al., 2010). Another LRR receptor-like kinase (RLK) brassinosteroid insensitive 1-associated kinase 1 (BAK1) is enrolled by FLS2 on the occasion of bacterial infections and upon the perception of flg22 (Chinchilla et al., 2007; Heese et al., 2007). Most recent studies have depicted that BAK1 is a coreceptor of flg22 and intermolecular interaction between the LRR domains of FLS2 and BAK1 stimulate the activities of the PRR complex (Orosa et al., 2018).

BAK1 Co-receptor:

A frequent co-receptor found in plants, that partake in regulating definite cell mechanisms such as growth, development and safeguarding against pathogens – is the membrane-bound Brassinosteroid insensitive 1-associated receptor kinase1 (BAK1). BAK1 attaches itself to ligand-stimulated transmembrane receptors and thereby, stimulates their kinase domains via transphosphorylation (Domínguez-Ferreras et al., 2015). Besides, it also imparts a quintessential role in several pathways of Arabidopsis that are responsible for PTI signaling (Wu et al., 2020).

BAK1 is a leucine-rich repeat receptor-like kinase (LRR-RLK). The extracellular LRR domain of BAK1 consists of five repeats. A serine and proline rich domain follow the LRR domain (Hecht et al., 2001). It also has a membrane-spanning domain, a cytoplasmic kinase domain, and a short C-terminal tail. Being is a member of the SERK protein family, BAK1 possesses 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).

A significant role in plant innate immunity is played by BAK1, upon communicating with the receptor FLS2 (Flagellin Sensing 2) (Chinchilla et al., 2007; Heese et al., 2007). Thereby, the idea was established that BAK1 has an imperative role in the control and management of a number of LRR-RLKs by co-operating with them in a system that is reliant on stimulus (Kemmerling and Nürnberger, 2008; Vert, 2008).

flg22 as an activator of FLS2:

The identification of the highly conserved N-terminal epitope (flg22) by flagellin-sensitive 2 (FLS2) induces Flagellin perception in *Arabidopsis*. When FLS2 binds with flg22, it simultaneously stimulates heteromer formation with Brassinosteroid insensitive 1–associated kinase 1 (BAK1), their reciprocal activation and subsequent induction of plant immunity (Sun et al., 2013).

A receptor like kinase (RLK) with an extracellular leucine-rich repeat (LRR), a single membrane-spanning domain, and a cytoplasmic serine/threonine kinase domain are typically encoded by the Arabidopsis FLS2 (AtFLS2). The consequences of the distinct interaction of FLS2 with AtFLS2 include chemical crosslinking and immunoprecipitation, that subsequently confirm that the specificity of flagellin perception is caused by AtFLS2 (Chinchilla et al. 2006).

flg22 peptides with mutant sequences are present in some bacteria which are pathogenic to plants, subsequently leading to the loss of efficiency in terms of detection by plant defense systems. In few bacteria such as *Agrobacterium tumefaciens* (Albert et al., 2010), *Ralstonia solanacearum* (Pfund et al., 2004), a specific valine/aspartate polymorphism governs the perception of flagellin via FLS2 in *Arabidopsis thaliana* (Sun et al., 2006).

Activation prosses of FLS2 by flg22:

The outward structure of bacteria is known as flagellum and is made of a protein sub-unit termed as ‘flagellin’, which consequently functions as a PAMP in plants (Felix et al., 1999; Che et al., 2000). The Leu-rich repeat transmembrane receptor kinase Flagellin Sensitive 2 (FLS2) is imperative for the identification of flagellin in *Arabidopsis thaliana* (Chinchilla et al., 2006).

In the first step, flagellin binds to the LRR ectodomain and is simultaneously detected by FLS2 (Sun et al., 2013). The FLS2-BAK1 complex is duly formed by the binding of the flg22 to FLS2. Interestingly, the LRR domains of co-receptor BAK1 also communicates with the ligand flg22 (Sun et al., 2013). flg22 then further creates the cross phosphorylation between BAK1 and FLS2 (Wang et al., n.d.). This perception by FLS2 accelerates 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).

When mutations are present inside FLS2, the plants carrying them often demonstrate faulty flagellin perception, reduced binding of flg22 and portray an increased vulnerability to bacterial infections (Bauer et al., 2001; Zipfel et al., 2004). It is thereby assumed that the interaction of flagellin with FLS2 may be secondary, and requires an extra element to create the high-affinity binding site. Though FLS2 mediated signaling can be implemented by a number of protein kinases including BAK1, BKK1, BIK1; however, the distinct FLS2 phosphorylation and kinase activities during receptor activation after ligand binding remains ambiguous.

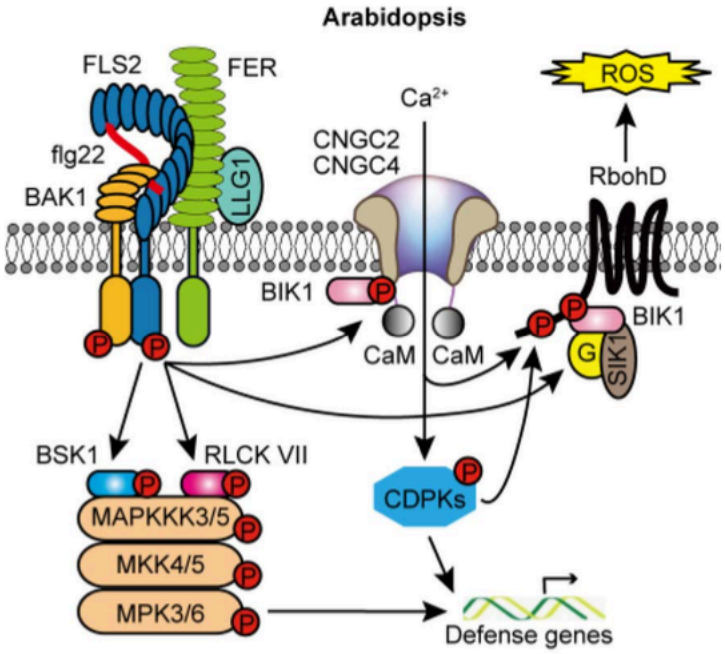


Figure 2: Overview of flg22 perception and molecular response in *Arabidopsis thaliana* (Wang et al., n.d.).

Molecular dynamics simulation:

Molecular dynamics simulation or MD simulation is the study pertaining to protein motions, upon the observation of their conformational changes over time. It is a computational approach and knowledge of the interaction potentials for the particles is imperative for running such a simulation (Karplus and Petsko, 1990). If the coordinate of all the atoms in a biomolecular system is known, for instance that of water and a lipid bilayer surrounding a protein, we can identify the force applied on each atom by all the other atoms. At present, three different types of simulation methods are used in the area of macromolecular research. The first one helps define structures with the data received from experiments on actual or real-life systems, upon working on a mean of sampling configuration space. The second method helps explain the system at equilibrium by utilizing the structural and motional properties, alongside the values of thermodynamic parameters. In the third method, the actual dynamics are examined, owing to the motion and development of specific particles with time (Karplus and McCammon, 2002). A wide variety of significant biomolecular processes can therefore, be captured by these simulations and the examples include - conformational change, protein folding, and ligand binding, which reveals the location of every atoms at femtosecond (fs) time interval resolution. The reaction of these biomolecules at an atomic level to perturbations, such as phosphorylation, mutation, addition/ exclusion of a ligand or protonation can also be predicted using these simulations (Hollingsworth and Dror, 2018).

Program availability and computational power is an essential requirement for carrying out essential studies on biological macromolecules upon utilizing the molecular dynamic (MD) simulation. Numerical stability in such a simulation is ensured by the due selection of short time steps, generally only a few femtoseconds (1015 s) each. Most of the structural changes in proteins, alongside the other biochemical events of interest, are generally known to occur at the timescales of nanoseconds, microseconds, or longer. Therefore, the myriad of time steps in conjunction with the interatomic interactions enumerated at each time step, verily constitute to making the simulations largely demanding, computationally. The most widely used programs in terms of actual simulation software are CHARMM20, AMBER21 and GROMOS22 (Karplus and McCammon, 2002; Becker et al., 2001; Tuckerman and Martyna, 2000; Gunsteren et al., 2013). However, before proceeding to perform the simulation, the molecular system needs to be prepared by adjusting the missing atoms, such as - 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).

MM/PBSA

The free energies generated, owing to the binding of miniscule ligands to larger biological macromolecules is calculated by utilizing the combination of molecular mechanics energies with the Poisson-Boltzmann and surface area continuum solvation (MM/PBSA) method.

This method has been specifically developed and modified according to the category of uses (Genheden and Ryde, 2015; Foloppe and Hubbard, 2006; Homeyer and Gohlke, 2012; Wang et al., 2006). MM/PBSA brings together molecular mechanics and continuum solvent models to estimate the affinities of ligand binding, and being intrinsically dependent on the MD simulations of the receptor-ligand complex, it is extensively used for protein designing, protein-protein interactions, conformer stability and re- scoring (Genheden and Ryde, 2015; Homeyer and Gohlke, 2012; Gohlke and Case, 2004; Hou et al., 2011; Moreira et al., 2007; Réblová et al., 2010; Sirin et al., 2014; Sun et al., 2014). In due accordance to the objectives, three different energy values are typically given as output (Kumari et al., 2014).

Free energy calculation is also used for drug design and the determination of protein structure (de Ruiter and Oostenbrink, 2011; Miller III et al., 2012).

Implications of BAK1 mutation:

A mutation in important amino acids within the BAK1 protein will verily result in a disruption of the developmental and immune signaling pathways in plants, as BAK1 protein is known to impart a valuable role in both these 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, when a ligand is absent

(Jaillais et al., 2011). Mutation in analogues position (D128N) in rice OsSERK2, which is structurally similar to BAK1 and belongs to 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 produces 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) produces 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). Again, in the 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:

This study aims to comprehend the structural dynamics of PRR FLS2, PAMP flg22 and mutated co-receptor mBAK1 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 D122N BAK1 mutation containing 4MN8 complex for 100 ns.
3. Analyzing the molecular interaction among PRR FLS2, PAMP flg22 and Co-receptor BAK1 ectodomain from the mutated complex by using MM/PBSA free energy binding calculation, RMSD, RMSF, Rg, H-bond and SASA.
4. Analyzing the contribution of the mutated Co-Receptor mBAK1 in the PTI mediated by FLS2-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. GROMOS96 54a7 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 water molecules. Next, charge neutralization of the system was achieved by adding positive sodium (NA) ions in place of solvent, as the complex had a net negative charge. In the following step, energy minimization of the system was duly achieved upon uutilizing the steepest descent minimization algorithm. Afterwards, 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 100ns (Nanosecond). The output file was generated as md_0_1.gro. A pdb file from the simulated '. gro' file was also generated using GROMACS.

Analysis:

The simulated protein complex was analyzed for MM/PBSA free energy binding contribution, RMSD (root mean square deviation) of the protein backbone, RMSF (root mean square fluctuation) per residue, R_g (radius of gyration), SASA (solvent accessible surface area) of protein and H-bond (hydrogen bond number) using GROMACS.

Chapter 3

Results:

In this chapter, interactions between PRR FLS2, PAMP flg22 and mutated Co-receptor BAK1 (mBAK1) ectodomain were analyzed, also illustrated with graphs and figures.

For in-depth analysis of interaction, MM/PBSA method was applied and elaborately described with proper figures in this section.

3.1 MM/PBSA

Table 1: The estimated binding free energies and the individual energy components of the studied systems (kJ/mol)

Complex	*ΔEvdw*[a]	*ΔEelec*[b]	*ΔEpolar*[c]	*ΔEsasa*[d]	*ΔEbind*[e]
FLS2+mBAK1	-417.128± 41.034	349.913± 100.612	960.032±174.105	-52.194± 6.877	840.624± 127.064
mBAK1+flg22	-67.914± 14.732	-55.736± 44.535	182.138± 88.961	-10.479± 3.705	48.008± 72.398
FLS2+flg22	-430.275± 30.942	-1397.223± 144.544	1068.419± 138.71	-56.658± 2.926	-815.737± 76.087
FLS2+mBAK1	-481.539± 48.456	344.783± 108.445	1005.547± 186.467	-61.856± 8.774	806.934± 110.225
mBAK1+flg22	-88.460± 0.820	-157.737± 2.964	226.952± 4.291	-12.526± 0.142	-31.830± 2.233
FLS2+flg22	-358.364± 29.562	-1515.022± 114.397	1148.755± 145.438	-50.430± 3.610	-775.061± 85.477

[a] Van der waals energy. [b] Electrostatic energy. [c] Polar solvation energy. [d] Solvent Accessible Surface Area (SASA) energy. [e] Binding Free energy. Each simulation was performed on 100ns and the first three rows are from three protein/ whole complex.

From MM/PBSA calculation (Table 3.1.1), the binding energies between FLS2+mBAK1, mBAK1+flg22 and FLS2+flg22 (from the three protein complex) are 840.624, 48.008 and -815.737 kJ/mol respectively. In contrast to this, the binding energies of FLS2+mBAK1 (in absence of flg22), mBAK1+flg22 (in absence of FLS2) and FLS2+flg22 (in absence of mBAK1) are found to be 806.934, -31.830 and 775.061 kJ/mol respectively.

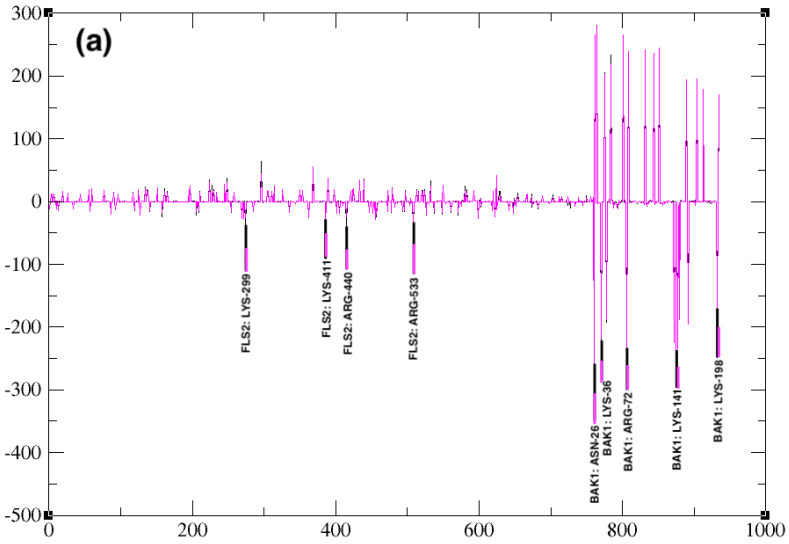
From the data obtained in the table, it is apparent that there is little to no interaction between FLS2+mBAK1 (both in the presence and absence of flg22), due to the high positive binding energy of 840.624 and 806.934 kJ/mol respectively, between them. Therefore, for interaction to happen between FLS2 and mBAK1, there is no significant contribution of PAMP flg22. In case of mBAK1+flg22, the binding energy is 48.008 kJ/mol in the presence of FLS2 and -31.830 kJ/mol in its absence. It can thereby be deduced, that the interaction between mBAK1 and flg22 is slightly more when FLS2 is not present inside the complex. Moving on to FLS2+flg22, there is very high interaction between them (both in the presence and absence of mBAK1). At -815.737 kJ/mol, the binding energy in the presence of mBAK1 increases moderately when compared to that in the absence of mBAK1, at -775.061 kJ/mol. We can therefore conclude, that the presence of mBAK1 plays quite a notable role in terms of increasing the interactions between FLS2 and flg22.

The energy contribution of every single residue was calculated and presented in MM/PBSA graphs, for a thorough and comprehensive analysis of the interaction between these proteins, (Fig. 3). We found that ARG-72, LYS-36, ARG-143, LYS-141 and ARG-138 from mutated co-receptor (coR) mBAK1 demonstrated a valuable role during the interaction, by forming MM energy between FLS2+mBAK1 (both in the presence and absence of flg22). Again, LYS-36, LEU-53, ASN-26 from mutated co-receptor (coR) mBAK1 and GLN-65, ARG-72, LYS-77 from pathogen associated molecular pattern (PAMP) flg22 formed the MM energy between mBAK1+flg22 (in the presence of FLS2), whereas GLU-110, GLU-117, GLU-98, GLU-28, ASP-67 from mutated co-receptor (coR) mBAK1 and GLN-65, ARG-66, ARG-72, LYS-77 from pathogen associated molecular pattern (PAMP) flg22 formed the MM energy between mBAK1+flg22 (in the absence of FLS2). Finally, in case of FLS2+flg22 (in the presence of mBAK1) GLU-81, ASP-102, ASP-150, ASP-176, ASP-220, ASP-222 from pattern recognition receptor (PRR) FLS2 and GLN-65, ARG-66, ARG-72, LYS-77 from pathogen associated molecular pattern (PAMP) flg22, contributed to forming the MM energy. Whereas, ASP-176, ASP-102, ASP-150, ASP-220, GLU-270 from pattern recognition receptor (PRR)

FLS2 and ARG-66, GLN-65, ARG-72, LYS-77 from pathogen associated molecular pattern (PAMP) flg22 produced the MM energy between FLS2+flg22 (in the absence of mBAK1).

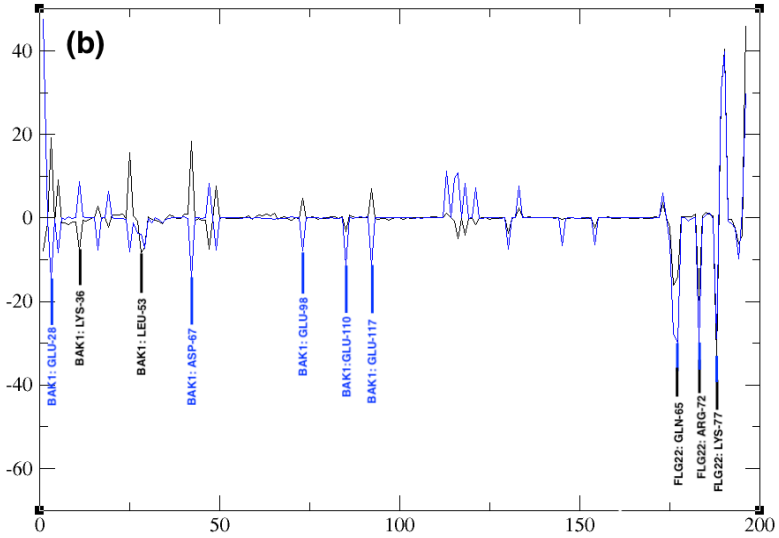
MM/PBSA (Free energy Binding Contribution

Total Energy (kJ/mol)



Residue

Total Energy (kJ/mol)



Residue

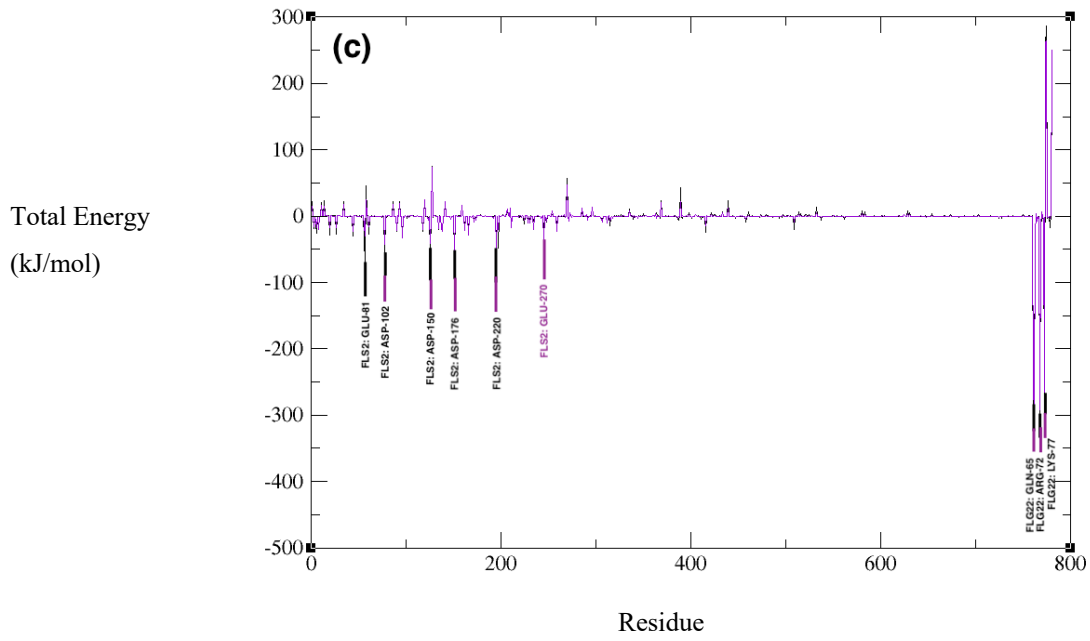


Figure 3: Comparison of the total binding energies obtained from MM/PBSA calculation between **(a)** FLS2-mBAK1 (Black: in the presence of flg22 and Pink: in the absence of flg22) **(b)** mBAK1-flg22 (Black: in the presence of FLS2 and Blue: in the absence of FLS2) **(c)** FLS2-flg22 (Black: in the presence of mBAK1 and Violet: in the absence of mBAK1).

3.2 Determination of the prominent residues of PRR FLS2, PAMP flg22 and mutated CoR mBAK1

The overall review of the MM/PBSA calculations along with the thorough and comprehensive analysis of the energy contribution of every single residue and RMSF values, leads us to discern and recognize some residues that are involved in paramount roles for making the interactions happen between FLS2, mBAK1 and flg22.

For the interaction between FLS2 and mBAK1 (in the presence of flg22), ARG-72 is one such prime residue as it demonstrates – the highest MM energy contribution (-209.7318 kJ/mol) and lowest RMS fluctuation (0.0853 nm). Again, for the interaction between FLS2 and mBAK1 (in the absence of flg22), LYS-36 is the prominent residue as it gives the highest MM energy contribution (-217.5491 kJ/mol) and lowest RMSF fluctuation (0.0993 nm).

When the same analysis of the interactions between mBAK1 and flg22 (in the presence of FLS2) were carried out, ARG-72 is again found to be the most prominent in both the cases (MM/PBSA and RMSF calculation). It gives (-19.7224 kJ/mol) of MM energy and the lowest RMSF fluctuation of (0.0853 nm). Again, for the interaction between mBAK1 and flg22 (in the absence of FLS2), GLU-28 is the prominent residue as it gives the highest MM energy contribution (-14.9059 kJ/mol) and lowest RMSF fluctuation (0.1268 nm).

Lastly, for the interactions between FLS2 and flg22 (in the presence of mBAK1), LYS-77 is the most prominent residue as it gives the highest MM energy contribution (-274.0768 kJ/mol) and lowest RMSF fluctuation (0.1696 nm). Again, for the interaction between FLS2 and flg22 (in the absence of mBAK1), ASP-176 is the prominent residue as it gives the highest MM energy contribution (-51.5362 kJ/mol) and lowest RMSF fluctuation (0.1709 nm).

Therefore, from this overview it can be said that ARG-72 and LYS-36 from FLS2-mBAK1 (in the presence and absence of flg22 respectively), ARG-72 and GLU-28 from mBAK1-flg22 (in the presence and absence of FLS2 respectively) & LYS-77 and ASP-176 (in the presence and absence of mBAK1 respectively) are the most supreme residues for implementing interactions between these three proteins. While other valuable residues may also be present, but in terms of the contribution of van der Waals energy, electrostatic energy and hydrogen bond formation, these six residues are the most imperative in the PTI complex formation. Therefore, we can assume that the plant's ability to trigger PTI can substantially become hindered, if a mutation is carried out at any of these residues.

3.3 RMSD

The stability of the MD simulation was calculated by analyzing Root mean square deviation (RMSD). The time evolution of the RMSDs of FLS2, mBAK1 and flg22 (residues only) are observed as a function of time and the RMSDs of FLS2, mBAK1 and flg22 in the three simulated systems are shown in Fig 4 respectively. The residues of three proteins exhibited varying RMSD in different simulated systems.

In the FLS2, mBAK1 and flg22 interaction when all the three proteins were present in the complex, the graph is equilibrated from time 59.5 ns with RMSD of 0.575 nm. Before 0.575 nm, it gives maximum deviations of 0.8205 nm at 27.7 ns and 0.8026 nm at 28.3 ns. In between

this period, the graph goes slightly down at 67.6 ns after equilibration and then, it moves forward stably with a standard deviation (SD) of 0.457 nm. From 59.5 ns, it gives a stable view and deviated at 67.3 ns, 85.5 ns and 96.4 ns, till the end of simulation, while producing deviations of 0.344, 0.333 and 0.315 nm respectively. Therefore, the deviation was between (0.315 to 0.343) nm.

In case of FLS2 and mBAK1 (when flg22 was not present in the complex), the deviation is unstable at the beginning of the simulation and some very high deviations of 0.764 nm and 0.678 nm are observed at 14.4 ns and 11.6 ns respectively. After this, the graph undergoes a massive decline in RMSD in between 17.25 ns and 34.5 ns. The deviation again rises shortly, and from 51.9 ns (0.705 nm) onwards, the graph is found to be equilibrated and moves forward stably with a standard deviation (SD) of 0.627 nm. However, there is a decline in stability in between 74 ns and 85.5 ns where low RMSD are found, ranging from (0.48 to 0.50) nm

Again, when FLS2 was not present in the complex, deviation of mBAK1 and flg22 complex is not that much stable as before. Rather, at the very beginning of the simulation process, the graph shows some stability. But after 17.5 ns, it continues maximum deviation and at 64.4 ns it deviates at 0.935 nm which is considered as the highest in terms of that condition. Again, instability is observed in this case from 73 ns onwards, as the graph has a massive decline in deviation during this time, and some very low RMSD (0.55 nm to 0.65 nm) are also observed.

The RMSD of FLS2 and flg22 (in the absence of mBAK1) was analyzed similarly. In this case also, the graph is found to be more unstable. Moderate equilibration is achieved from time 44 ns with a RMSD of 0.86 nm. However, at 84 ns there is a sudden rise in RMSD which is considered to be the highest in that condition (1.1 nm). After 84 ns, the graph again moves forward stably with a standard deviation (SD) of 0.892 nm.

In terms of all of the three cases, it is found that RMSD of FLS2, mBAK1 and flg22 gives more stability when all of them are inside the complex and interact with each other. Without mBAK1, RMSD of FLS2 and flg22 deviates more and unstable situation is observed more than another situation. So, for the interaction between FLS2 and flg22 to occur, mBAK1 have to be present inside the complex.

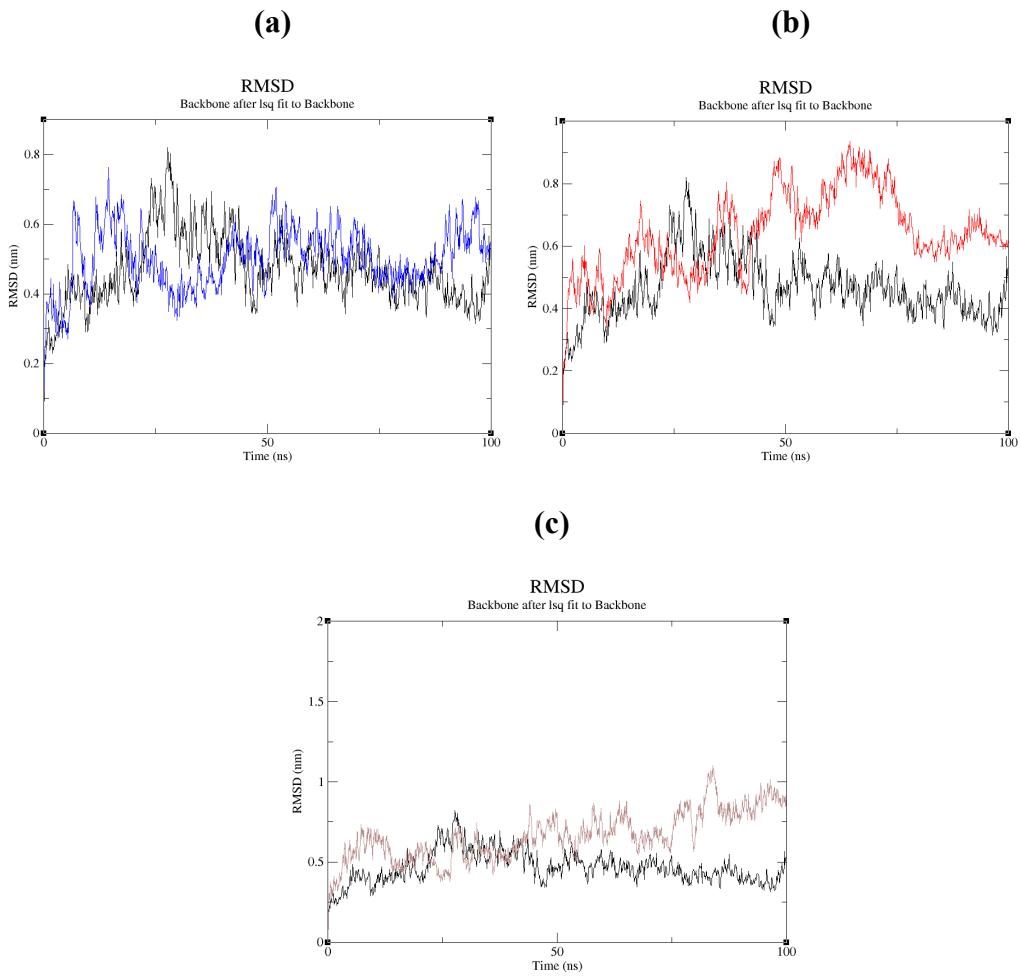


Figure 4: RMSD values of FLS2, mBAK1 and flg22 from 100ns MD trajectories: **(a)** FLS2-mBAK1 (Black: in the presence of flg22 and Blue: in the absence of flg22) **(b)** mBAK1-flg22 (Black: in the presence of FLS2 and Red: in the absence of FLS2) **(c)** FLS2-flg22 (Black: in the presence of mBAK1 and Brown: in the absence of mBAK1).

3.4 Rg

To measure the compactness of all systems, the radius of gyration (Rg) values were measured. When we calculated the Rg values of the FLS2-mBAK1-flg22 three protein complex, we found that the graph reached its highest value of Rg 3.705 nm at 42,200 seconds, after which, it dropped and reached the lowest value of 3.512 nm at 60,500 seconds. After 60,500 seconds, the Rg again continued to rise and stability was found in between 78,000 seconds and 96,700

seconds. However, towards the end of the simulation, Rg increased one more time and formed a value of 3.722 nm at 99,500 seconds.

In case of FLS2 and mBAK1 complex (in the absence of flg22), the graph was found to be unstable and more fluctuations were present. In the inception of the simulation, the Rg value was found to be 3.717 nm at 1500 seconds. However, it quickly fell to 3.433nm at 6700 seconds. Afterwards, the Rg again increased, forming some high values of 3.688 nm, 3.709 nm, 3.736 nm, 3.712 nm and 3.689 nm at 14500, 29100, 32750, 67800 and 84300 seconds respectively. However, during this time, some low Rg values of 3.44 nm at 21,600 seconds, 3.491 nm at 42,600 seconds and 3.499 nm at 95,700 seconds could also be seen.

When the Rg of mBAK1 and flg22 (in the absence of FLS2) was analyzed, the graph was also unstable with more number of fluctuations. Initially, the Rg value was found to be 2.051 nm at 300 seconds, but it soon faced a decline by decreasing to 1.863 nm at 3500 seconds. The graph reached its highest Rg value of 2.023 nm at 7800 seconds. But, it was again found to fall to the lowest Rg value of 1.753 nm at 17,200 seconds. After this, some high Rg values of 1.935 nm and 1.982 nm could also be seen at 40,000 seconds and 52,700 seconds respectively. From 64,600 seconds onwards, the Rg value once again declined and the second lowest value of 1.822 nm was found at 96,300 seconds.

Lastly, in case of FLS2 and flg22 (in the absence of mBAK1), the graph illustrated a continuous increase of Rg values from 0 to 20,250 seconds. After this, there was a quick fall at 23,200 seconds and the lowest Rg value of 3.76 nm was found at 25,000 seconds. The Rg again increased, forming a high value of 4.107 nm at 27,750 seconds. From this point onwards, the graph was found to be equilibrated as the fluctuations were more stable. However, once again, a fall to 3.838 nm was found at 85,550 seconds after which, the Rg again continued to rise by forming the highest value of 4.195 nm at 97,700 seconds.

In conclusion, we can say that more variations in Rg values signify a more changing structure, which is consistent with the higher fluctuations in FLS2-mBAK1, mBAK1-flg22 and FLS2-flg22 complexes; as the proteins are more freely able to uncoil and recoil. On the other hand, when the FLS2 is interacting with both flg22 and mBAK1 in a single complex, it is bound in place and less able to uncoil resulting in a less fluctuating graph.

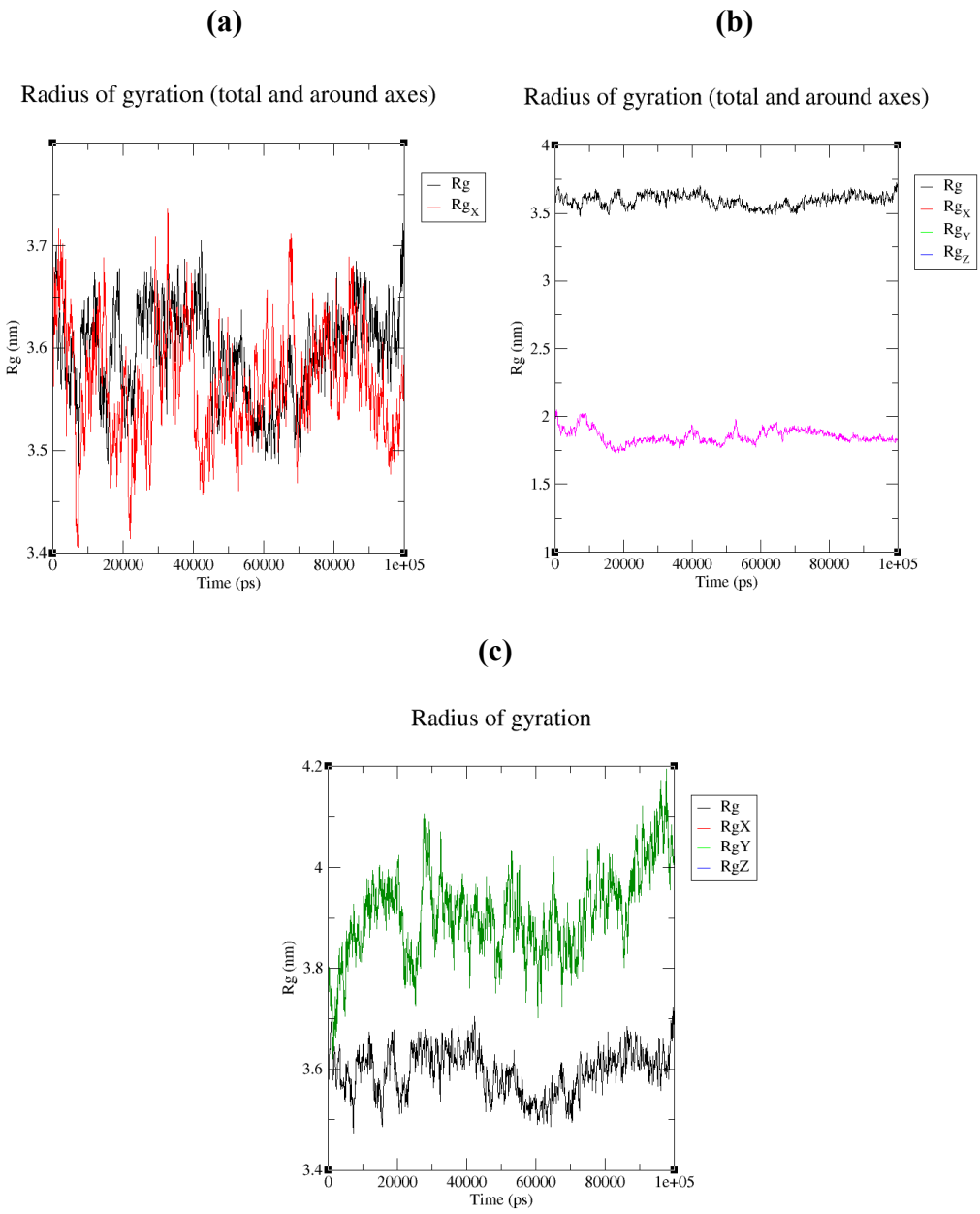


Figure 5: Rg values of FLS2, mBAK1 and flg22 from 100ns MD trajectories: **(a)** FLS2-mBAK1 (Black: in the presence of flg22 and Red: in the absence of flg22) **(b)** mBAK1-flg22 (Black: in the presence of FLS2 and Pink: in the absence of FLS2) **(c)** FLS2-flg22 (Black: in the presence of mBAK1 and Green: in the absence of mBAK1).

3.5 SASA

The solvent accessible surface areas (SASA) were also calculated and it was seen to be moderately higher for the FLS2-mBAK1-flg22 complex than the FLS2-mBAK1 complex, showing a steady mean value of about 400 nm^2 . Opposed to this, the FLS2-mBAK1 complex had a slightly lower mean value of about 395 nm^2. Again, SASA value of mBAK1-flg22 complex is 120 nm^2 and that of FLS2-flg22 complex is 350 nm^2. The low SASA values of mBAK1-flg22 and FLS2-flg22 demonstrate that the interactions of flg22 with both the mutated co-receptor (coR) mBAK1 and the pattern-recognition receptor (PRR) FLS2 regions, somewhat increased the surface area of the complex which water (the solvent in this case) could access.

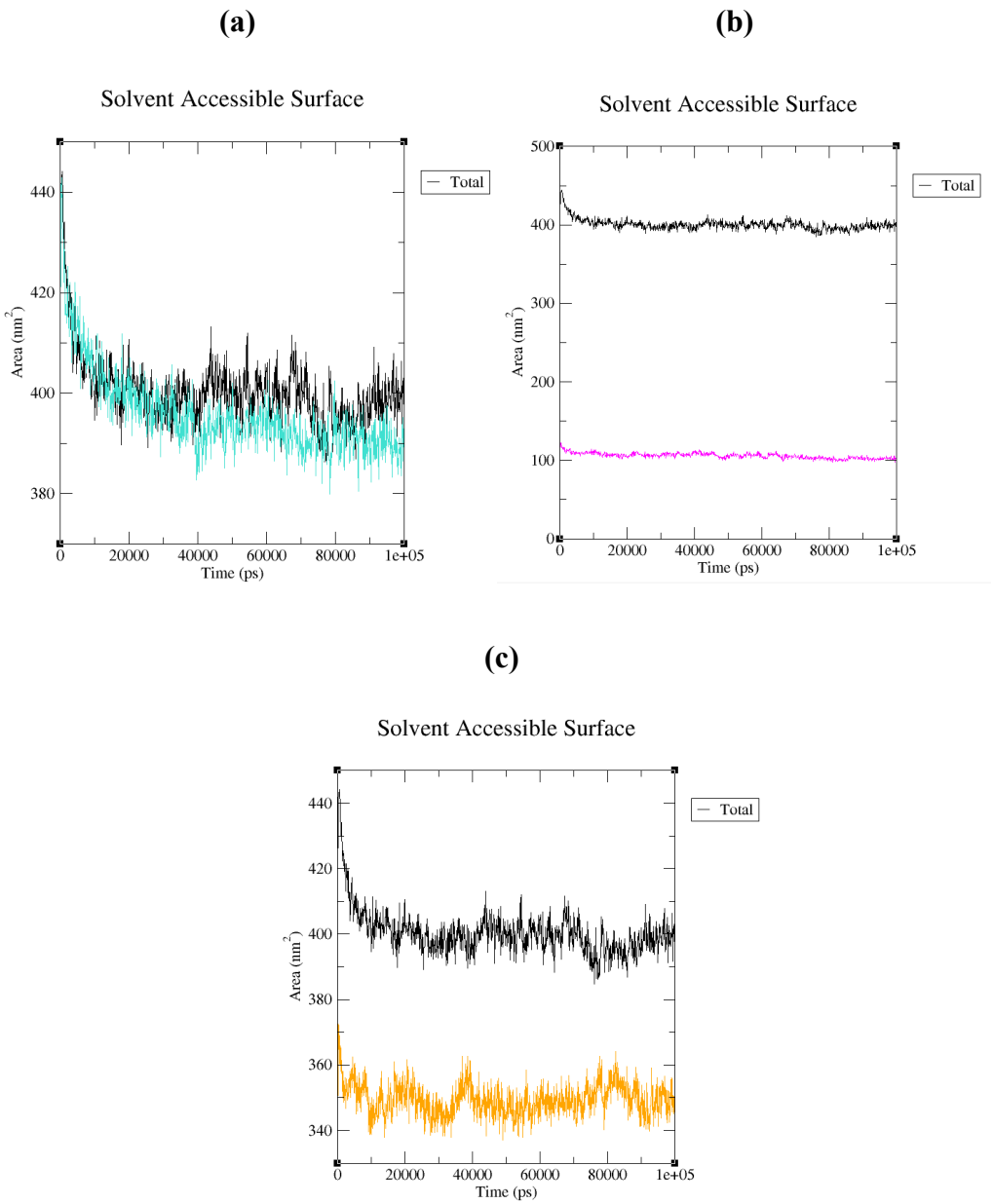


Figure 6: SASA values of FLS2, mBAK1 and flg22 from 100ns MD trajectories: **(a)** FLS2-mBAK1 (Black: in the presence of flg22 and Light Green: in the absence of flg22) **(b)** mBAK1-flg22 (Black: in the presence of FLS2 and Pink: in the absence of FLS2) **(c)** FLS2-flg22 (Black: in the presence of mBAK1 and Yellow: in the absence of mBAK1).

3.6 RMSF

MD trajectories of all the simulated systems were utilized to estimate the RMSF values of the residues of FLS2, mBAK1 and flg22. A simulation period of 100 ns was applied to calculate the RMSFs. The obtained results exhibited that the fluctuations were: less than 0.3 nm for the FLS2-mBAK1-flg22 (three protein complex), less than 0.2 nm for FLS2-mBAK1, less than 0.25 nm for mBAK1-flg22 and less than 0.75 nm for FLS2-flg22 (Fig 7).

From MM/PBSA calculation, some residues show the most favourable condition for interacting between proteins. For FLS2-mBAK1, ARG-72, LYS-36, ARG-143, LYS-141 and ARG-138 give more favourable condition in both cases (when flg22 is present in the complex and absent). From RMSF graphical view, it is obtained that these residues also give less fluctuation in contrast to other residues of FLS2-mBAK1.

When an interaction between mBAK1-flg22 (in the presence of FLS2) were carried out, ARG-72 is again the prime residue in all instances (MM/PBSA and RMSF calculation). It produces (-19.7224 kJ/mol) of MM energy and the lowest RMSF fluctuation of (0.0853 nm). Again, for the interaction between mBAK1 and flg22 (in the absence of FLS2), GLU-28 is the prominent residue as it gives the highest MM energy contribution (-14.9059 kJ/mol) and lowest RMSF fluctuation (0.1268 nm). In contrast to this, the most fluctuating residues were LYS-77 (0.6936 nm) and ARG-72 (0.7031 nm).

When it comes to the interactions between FLS2-flg22 (in the presence of mBAK1), LYS-77 from pathogen associated molecular pattern (PAMP) flg22 range and GLU-81, GLN-65, ARG-66 from pattern-recognition receptor (PRR) FLS2 range showed considerably low fluctuations. It was simultaneously divulged from the MM/PBSA calculation that these same residues produced increased energies for making the due interactions. Besides, LYS-77 is the most prominent residue in this case, as it gives the lowest RMSF fluctuation (0.1696 nm), which further supports the results of MM/PBSA calculation. Again, for the interaction between FLS2-flg22 (in the absence of mBAK1), ASP-176 is the prominent residue as it gives the lowest RMSF fluctuation (0.1709 nm). This residue, along with ARG-66 and GLN-65 from pattern recognition receptor (PRR) FLS2 also produced the MM energy between FLS2+flg22 (in the absence of mBAK1).

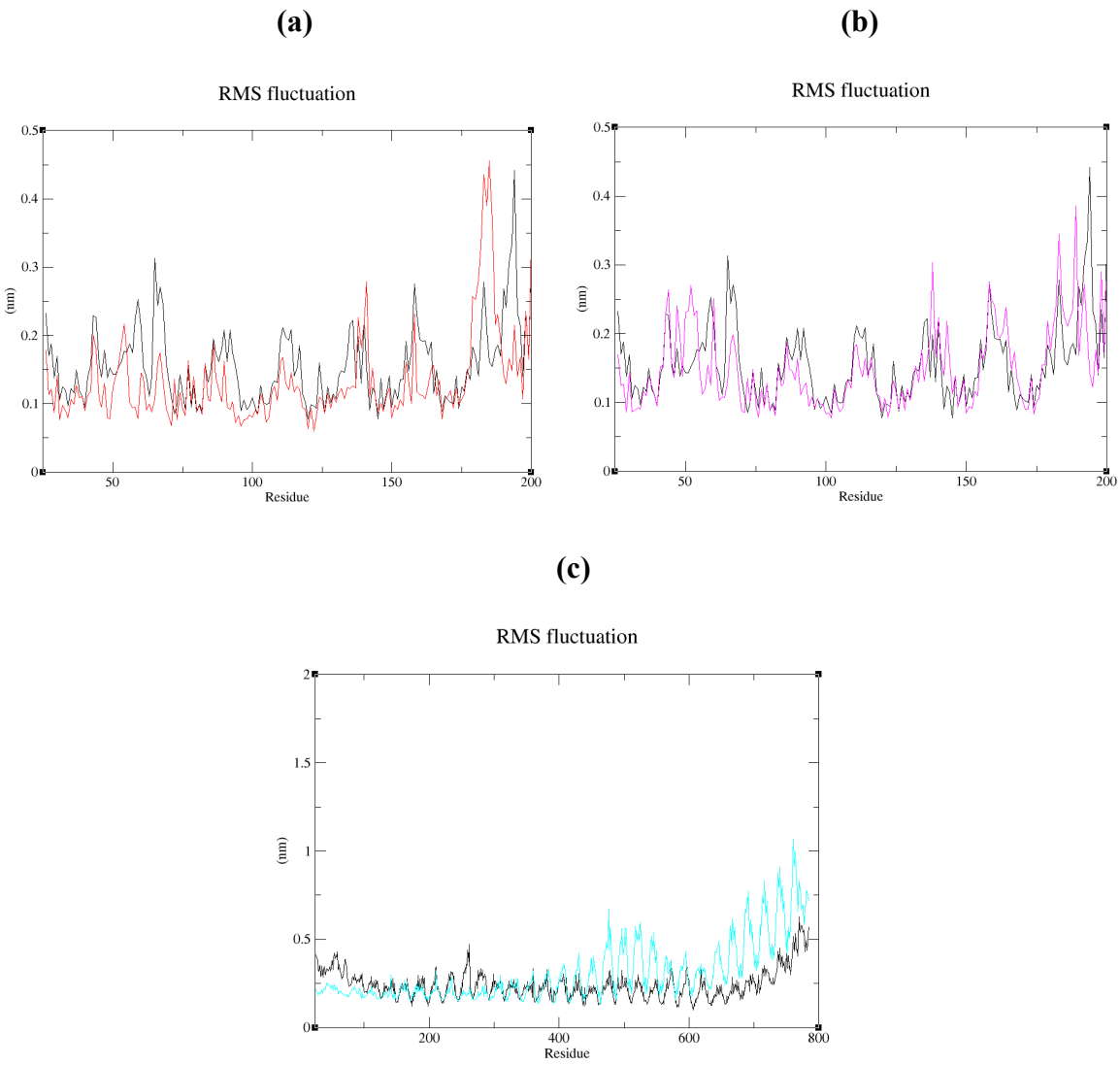


Figure 7: RMSF values of FLS2, mBAK1 and flg22 from 100ns MD trajectories: **(a)** FLS2-mBAK1 (Black: in the presence of flg22 and Red: in the absence of flg22) **(b)** mBAK1-flg22 (Black: in the presence of FLS2 and Pink: in the absence of FLS2) **(c)** FLS2-flg22 (Black: in the presence of mBAK1 and Light Green: in the absence of mBAK1).

3.7 H-bond

Hydrogen bond number was calculated over the 100 ns time of simulation for protein-protein criteria. The protein-protein hydrogen bonds show the overall curve lifting graph throughout the simulation. For a better understanding of forming hydrogen bond within two proteins index file was used which helps to calculate overall hydrogen bond formation at a different period.

From Fig 8, it is visible that maximum 18 hydrogen bond is found between FLS2 and mBAK1, at 69500 ps and 77800 ps respectively, when flg22 is present in the complex. Contrary to this, maximum 21 hydrogen bond is found at 31500 ps, 35500ps and 41100 ps respectively, in the absence of flg22.

Again, a maximum of 6 hydrogen bond is found between mBAK1 and flg22 at 31200 ps, 31600 ps, 41700 ps, 82400 ps, 89900 ps, 92200 ps and 94900 ps respectively, when FLS2 is present and in absence of FLS2, maximum 8 hydrogen bond is formed at 18600ps, 48200 ps and 48300 ps respectively.

Finally, maximum 22 hydrogen bond is found between FLS2 and flg22 at 7800 ps, in the presence of mBAK1 and in its absence, maximum 22 hydrogen bond is also found. This time however, it was at 16650 ps.

It can be thereby concluded that, the total number of hydrogen bonds increased throughout the period of simulation, upon the individual interactions of the three complexes of FLS2-mBAK1, mBAK1-flg22 and FLS2-flg22. Therefore, when all the three proteins are present inside the complex, a slightly lesser number of hydrogen bonds may be received.

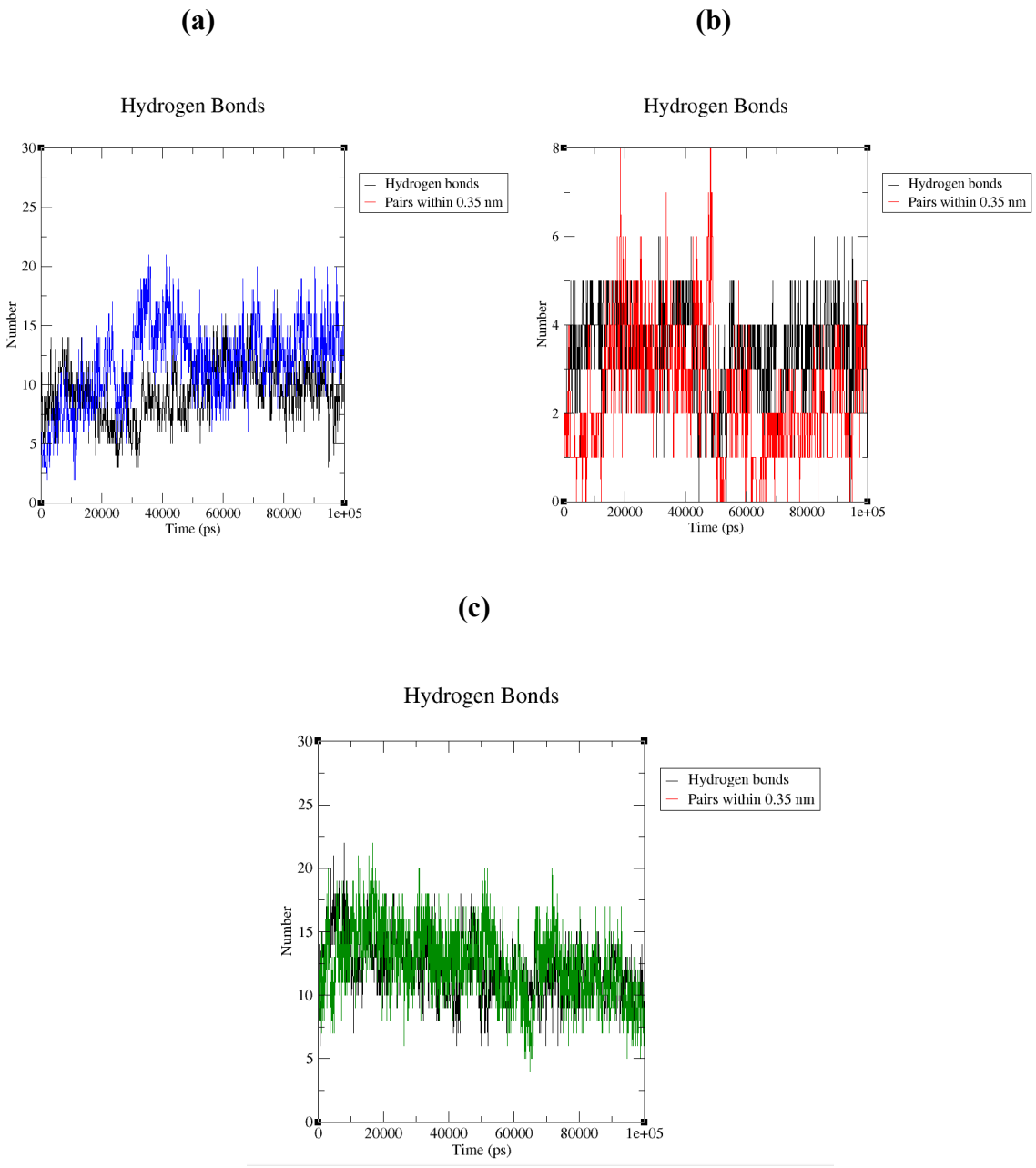


Figure 8: H-bond values of FLS2, mBAK1 and flg22 from 100ns MD trajectories: **(a)** FLS2-mBAK1 (Black: in the presence of flg22 and Blue: in the absence of flg22) **(b)** mBAK1-flg22 (Black: in the presence of FLS2 and Red: in the absence of FLS2) **(c)** FLS2-flg22 (Black: in the presence of mBAK1 and Green: in the absence of mBAK1).

Chapter 4

Discussion:

This study describes the computational analysis of a plant PRR and reports on the interaction of the PRR FLS2 to PAMP flg22 and mutated Co-receptor mBAK1, for understanding Pattern Triggered Immunity (PTI) of *Arabidopsis thaliana*. Different analytical approaches were used for this purpose.

For sustaining the conformation of other amino acids in the original/ native form, the mutation of aspartate to asparagine at 122nd residue of the BAK1 was induced in crystal structure. The mutation was carried out in the raw amino acid sequence and next, the modeling and docking of proteins were carried out to prevent unwanted changes to the conformation of the complex. At the same time, the least steric clash of new amino acid asparagine with its surrounding molecules, was also ensured.

To perform the molecular dynamics simulation, the mutated protein complex was positioned 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. Next, the system was 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. The system was then charge neutralized and energy minimized to relax the protein complex. The system was then equilibrated in two-phases: the first phase brought the temperature of the system around 300 kelvin or room temperature, while the second phase equilibrated the pressure and density of the system. At the end, the system containing the protein complex was stable, equilibrated and ready for molecular dynamics simulation.

Several analytical approaches were utilized for examining and comparing the data of the mutated protein complex, following triumphant execution of the molecular dynamics simulation. First, the MM/PBSA Free Energy Binding Calculation was carried out. The results demonstrated that during the presence of mBAK1 in the complex, flg22 easily interacted with FLS2 where the residue LYS-77 played a crucial role. This residue produced high binding energy which contributed to the making of interaction between FLS2 and flg22 possible in the presence of mBAK1. Contrary to this, in the absence of mBAK1 in the complex, no notable binding energy was found. All in all, ARG-72 and LYS-36 from FLS2-mBAK1 (in the presence and

absence of flg22 respectively), ARG-72 and GLU-28 from mBAK1-flg22 (in the presence and absence of FLS2 respectively) & LYS-77 and ASP-176 (in the presence and absence of mBAK1 respectively) were the paramount residues for instituting interactions between the three proteins.

Next, the Root Mean Square Deviation or RMSD was measured, that analyzed the average deviation of the atoms compared to a reference structure at a moment of the simulation. The results indicated that the RMSD of FLS2, mBAK1 and flg22 produced more stability when all of them were inside the complex and interacted with each other. Without mBAK1, RMSD of FLS2 and flg22 deviated more and unstable situation was observed for more than one situation. Therefore, there is a clear indication that for the interaction between FLS2 and flg22 to occur, mBAK1 have to be present inside the complex.

The Radius of Gyration or Rg is an indicator of protein structure compactness, which is calculated by measuring the distance between the center of mass of the protein to both of its terminal end. As Rg estimates the folding/ unfolding of a protein structure, more variations in observed values indicate a more changing structure. In our study, we found higher fluctuations in the FLS2-mBAK1, mBAK1-flg22 and FLS2-flg22 complexes; as these proteins were more freely able to uncoil and recoil. On the other hand, when FLS2 was interacting with both flg22 and mBAK1 domains in a single complex, it was bound in place and less able to uncoil resulting in lesser number of fluctuations.

The solvent accessible surface areas (SASA) were also calculated and it was seen to be higher for the FLS2-mBAK1-flg22 complex than the FLS2-mBAK1, mBAK1-flg22 and FLS2-flg22 complexes. The low SASA values of mBAK1-flg22 and FLS2-flg22 demonstrated that the interactions of flg22 with both the mutated co-receptor (coR) mBAK1 and the pattern-recognition receptor (PRR) FLS2 regions, somewhat increased the surface area of the complex which water (the solvent in this case) could access.

The Root Mean Square Fluctuation or RMSF is another parameter to determine the deviation of specific atoms or groups of atoms compared to the reference structure average over atoms. Obtained results revealed that during the presence of mBAK1, the interaction between flg22 and FLS2 went to a stable form after a certain period of simulation and also prominent residues were found to be very low fluctuated during this time.

Upon a closer look at the number of hydrogen bonds formed between the three different complexes (FLS2-mBAK1, mBAK1-flg22, FLS2-flg22) over the simulation period, we can get an idea about the trend of intra-protein interaction occurring. From our study, we found that the total number of hydrogen bonds increased throughout the period of simulation, upon the individual interactions of the three complexes of FLS2-mBAK1, mBAK1-flg22 and FLS2-flg22. Therefore, when all the three proteins are present inside the complex, a slightly lesser number of hydrogen bonds may be received.

Taking into account of all the data generated via molecular dynamics simulation, MM/PBSA, RMSD, Rg, SASA, RMSF and Hydrogen bond interaction, we can see the overall effects of the D122N BAK1 mutation on the complex formed by FLS2-flg22-mBAK1.

Conclusion:

D122N mutation in the BAK1 coding protein is known to create elongated (elg) phenotype of *Arabidopsis thaliana,* which in turn causes the plant to have elongated stem and early flowering among other phenotypes. Bolstering plant immunity upon heterodimer formation with plant PRR FLS2 (induced by the bacterial flagellin flg22), alongside crucial roles in various processes of cell development are essentially played the BAK1 protein. It is believed that the findings of this study will aid the scientific community and facilitate more studies to fully perceive the structural dynamics of PRR FLS2, PAMP flg22 and mutated co-receptor mBAK1 of *Arabidopsis thaliana* and better understand PTI, or the first layer of the plant defense as a whole.

References:

Agrawal, N., Dasaradhi, P. V., Mohmmed, A., Malhotra, P., Bhatnagar, R. K., & Mukherjee, S. K. (2003). RNA interference: biology, mechanism, and applications. *Microbiology and molecular biology reviews : MMBR*, *67*(4), 657–685. https://doi.org/10.1128/mmbr.67.4.657-685.2003

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

Alberto P. Macho, Cyril Zipfel, Plant PRRs and the Activation of Innate Immune Signaling, Molecular Cell, Volume 54, Issue 2, 2014, Pages 263-272

Asai, Tsuneaki & Tena, Guillaume & Plotnikova, Joulia & Willmann, Matthew & Chiu, Wan-Ling & Gómez-Gómez, Lourdes & Boller, Thomas & Ausubel, Frederick & Sheen, Jen. (2002). MAP kinase signaling cascade in Arabidopsis innate immunity. Nature. 415. 977-83. 10.1038/415977a.

Bauer, Z., Gómez-Gómez, L., Boller, T., and Felix, G. (2001). Sensitivity of different ecotypes and mutants of *Arabidopsis thaliana* toward the bacterial elicitor flagellin correlates with the presence of receptor-binding sites. J. Biol. Chem. **276** 45669–45676.

Becker, O.M., et al., *Computational biochemistry and biophysics*. 2001: Marcel Dekker New York.

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

Boller, T. and G.J.A.r.o.p.b. Felix, *A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors.* 2009. 60: p. 379-406.

Chinchilla D, Bauer Z, Regenass M, Boller T, Felix G. The Arabidopsis receptor kinase FLS2 binds flg22 and determines the specificity of flagellin perception. *Plant Cell*. 2006;18(2):465-476. doi:10.1105/tpc.105.036574

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/nature0599

Collier SM, Moffett P. NB-LRRs work a "bait and switch" on pathogens. Trends Plant Sci. 2009 Oct;14(10):521-9. doi: 10.1016/j.tplants.2009.08.001. Epub 2009 Aug 31. PMID: 19720556.

Che, F.S. , Nakajima, Y. , Tanaka, N. , Iwano, M. , Yoshida, T. , Takayama, S. , Kadota, I. and Isogai, A. (2000) Flagellin from an incompatible strain of *Pseudomonas avenae* induces a resistance response in cultured rice cells. *J. Biol. Chem*. 275, 32347–32356.

Dangl, J. L., Horvath, D. M., & Staskawicz, B. J. (2013). Pivoting the plant immune system from dissection to deployment. *Science (New York, N.Y.)*, *341*(6147), 746–751. https://doi.org/10.1126/science.1236011

de Ruiter, A. and C.J.C.o.i.c.b. Oostenbrink, *Free energy calculations of protein–ligand interactions.* 2011. **15**(4): p. 547-552.

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

Domínguez-Ferreras A, Kiss-Papp M, Jehle AK, Felix G, Chinchilla D. An Overdose of the Arabidopsis Coreceptor BRASSINOSTEROID INSENSITIVE1-ASSOCIATED RECEPTOR KINASE1 or Its Ectodomain Causes Autoimmunity in a SUPPRESSOR OF BIR1-1-Dependent Manner. Plant Physiol. 2015 Jul;168(3):1106-21. doi: 10.1104/pp.15.00537. Epub 2015 May 5. PMID: 25944825; PMCID: PMC4741324

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 ZQ, Dong X. Systemic acquired resistance: turning local infection into global defense. Annu Rev Plant Biol. 2013;64:839-63. doi: 10.1146/annurev-arplant-042811-105606. Epub 2013 Jan 25. PMID: 23373699.

Foloppe, N. and R.J.C.m.c. Hubbard, *Towards predictive ligand design with free-energy based computational methods?* 2006. **13**(29): p. 3583-3608.

Genheden, S. and U.J.E.o.o.d.d. Ryde, *The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities.* 2015. **10**(5): p. 449-461.

Gohlke, H. and D.A.J.J.o.c.c. Case, *Converging free energy estimates: MM-PB (GB) SA studies on the protein–protein complex Ras–Raf.* 2004. **25**(2): p. 238-250.

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

Gust AA, Felix G. Receptor like proteins associate with SOBIR1-type of adaptors to form bimolecular receptor kinases. Curr Opin Plant Biol. 2014 Oct;21:104-111. doi: 10.1016/j.pbi.2014.07.007. Epub 2014 Jul 26. PMID: 25064074.

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

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

Hohmann, Ulrich and Lau, Kelvin and Hothorn, Michael. (2017). The Structural Basis of Lig-and Perception and Signal Activation by Receptor Kinases, Annual Review of Plant Biology, Volume 68, Number 1, Pages 109-137. 10.1146/annurev-arplant-042916-040957

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

Homeyer, N. and H.J.M.I. Gohlke, *Free energy calculations by the molecular mechanics Poisson− Boltzmann surface area method.* 2012. **31**(2): p. 114-122.

Hou, T., et al., *Assessing the performance of the molecular mechanics/Poisson Boltzmann surface area and molecular mechanics/generalized Born surface area methods. II. The accuracy of ranking poses generated from docking.* 2011. **32**(5): p. 866-877.

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

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

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

Karplus, M. and G.A.J.N. Petsko, *Molecular dynamics simulations in biology.* 1990. **347**(6294): p. 631.

Karplus, M., J.A.J.N.S. McCammon, and M. Biology, *Molecular dynamics simulations of biomolecules.* 2002. **9**(9): p. 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

Kourelis J, van der Hoorn RAL. Defended to the Nines: 25 Years of Resistance Gene Cloning Identifies Nine Mechanisms for R Protein Function. Plant Cell. 2018 Feb;30(2):285-299. doi: 10.1105/tpc.17.00579. Epub 2018 Jan 30. PMID: 29382771; PMCID: PMC5868693.

Li J, Chory J. A putative leucine-rich repeat receptor kinase involved in brassinosteroid signal transduction. Cell. 1997 Sep 5;90(5):929-38. doi: 10.1016/s0092-8674(00)80357-8. PMID: 9298904.

Liebrand, Thomas & Burg, Harrold & Joosten, Matthieu. (2013). Two for all: receptor-associated kinases SOBIR1 and BAK1. Trends in plant science. 19. 10.1016/j.tplants.2013.10.003.

Lourdes Gómez-Gómez, Thomas Boller, FLS2: An LRR Receptor–like Kinase Involved in the Perception of the Bacterial Elicitor Flagellin in Arabidopsis, Molecular Cell, Volume 5, Issue 6, 2000, Pages 1003-1011

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

Mattoo, Seema & Lee, Yvonne & Dixon, Jack. (2007). Interactions of bacterial effector proteins with host proteins. Current opinion in immunology. 19. 392-401. 10.1016/j.coi.2007.06.005.

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

Miller III, B.R., et al., *MMPBSA. py: an efficient program for end-state free energy calculations.* 2012. **8**(9): p. 3314-3321.

Mocellin, S., & Provenzano, M. (2004). RNA interference: learning gene knock-down from cell physiology. *Journal of translational medicine*, *2*(1), 39. https://doi.org/10.1186/1479-5876-2-39

Moreira, I.S., P.A. Fernandes, and M.J.J.T.C.A. Ramos, *Unravelling Hot Spots: a comprehensive computational mutagenesis study.* 2007. **117**(1): p. 99-113.

Orosa, B., Yates, G., Verma, V. *et al.* SUMO conjugation to the pattern recognition receptor FLS2 triggers intracellular signalling in plant innate immunity. *Nat Commun* **9,** 5185 (2018). https://doi.org/10.1038/s41467-018-07696-8

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

Ranf, S.J.C.o.i.p.b., *Sensing of molecular patterns through cell surface immune receptors.* 2017. **38**: p. 68-77.

Réblová, K., et al., *An RNA molecular switch: Intrinsic flexibility of 23S rRNA helices 40 and 68 5′-UAA/5′-GAN internal loops studied by molecular dynamics methods.* 2010. **6**(3): p. 910-929

Ribet D, Cossart P. Pathogen-mediated posttranslational modifications: A re-emerging field. Cell. 2010 Nov 24;143(5):694-702. doi: 10.1016/j.cell.2010.11.019. PMID: 21111231; PMCID: PMC7112265.

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

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

Sirin, S., et al., *A computational approach to enzyme design: predicting ω- aminotransferase catalytic activity using docking and MM-GBSA scoring.* 2014. **54**(8): p. 2334-2346.

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., 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

Sun, H., et al., *Assessing the performance of MM/PBSA and MM/GBSA methods. 5. Improved docking performance using high solute dielectric constant MM/GBSA and MM/PBSA rescoring.* 2014. **16**(40): p. 22035-22045.

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

Tsuda, Kenichi & Katagiri, Fumiaki. (2010). Comparing signaling mechanisms engaged in pattern-triggered and effector-triggered immunity. Current opinion in plant biology. 13. 459-65. 10.1016/j.pbi.2010.04.006.

Tuckerman, M.E. and G.J. Martyna, *Understanding modern molecular dynamics: techniques and applications*. 2000, ACS Publications.

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

van Gunsteren, W.F., P.K. Weiner, and A.J. Wilkinson, *Computer simulation of biomolecular systems: theoretical and experimental applications*. Vol. 3. 2013: Springer Science & Business Media.

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

Wang, J., T. Hou, and X.J.C.C.-A.D.D. Xu, *Recent advances in free energy calculations with a combination of molecular mechanics and continuum models.* 2006. **2**(3): p. 287- 306.

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

Xin, X.-F., et al., *Bacteria establish an aqueous living space in plants crucial for virulence.* 2016. **539**(7630): p. 524.

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

Yujun Wu, Yang Gao, Yanyan Zhan, Hong Kui, Hongyan Liu, Li Yan, Birgit Kemmerling, Ji an-Min Zhou, Kai He, Jia Li. Proceedings of the National Academy of Sciences Oct 2020, 117 (43) 27044-27053; DOI: 10.1073/pnas.1915339117

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, Cyril & Robatzek, Silke & Navarro, Lionel & Oakeley, Edward & Jones, Jonathan & Felix, Georg & Boller, Thomas. (2004). Bacterial disease resistance in Arabidopsis through flagellin perception. Nature. 428. 764-7. 10.1038/nature02485.

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