IN SILICO ANALYSIS OF STRUCTURAL COMPLEX OF HAESA RECEPTOR LIKE KINASE, PAMP IDA AND SERk1 CO-RECEPTOR

By

Raghib Ishraq Alvy
15236001

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

Mathematics and Natural Sciences
Brac University
December 2019

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.

Student's Full Name & Signature:

Raghib Ishraq Alvy

15236001

Approval

The thesis titled “In silico analysis of structural complex of HAESA receptor like kinase with PAMP IDA and SERk1 co-receptor” submitted by Raghib Ishraq Alvy (15236001) has been accepted as satisfactory in partial fulfillment of the requirement for the degree of Bachelor of Science on 5 December, 2019.

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 animal or living things were used in this study.

ABSTRACT

Pattern Triggered Immunity (PTI) is actually identified by the activities of pattern recognition receptors (PRRs), which play the most important role in first layer of plant defense mechanism. During aggression of microbes, PRRs immediately bind with the pathogen-associated molecular patterns (PAMPs) and recruit co-receptor protein(s) to activate the defense signal and begin plant's immunity. Although several plant PRRs have been discovered, very few of them have been fully characterized and their functional parameters studied. In this study, the crystallographic structure of HAESA-IDA-SERk1 (PDB ID: 5IYX) complex was simulated for 30 ns in different stages like there were five different combinations of the main crystallographic structure. For every combination 30ns simulated trajectories were analyzed for getting an overview of the interaction and immune response of HAESA towards IDA with the help of co-receptor SERk1. Analyzing the interaction revealed its remarkable resemblance to the other crystallographic structures, which are also a member of the same LRR-RK subfamily (leucine-rich repeat – receptor kinase subfamily XII). Furthermore, it was observed that Tyr56 and Arg67 of IDA contributed significantly to the formation of the HAESA-IDA complex. It can thus be predicted that any change to the PAMP at these positions can be greatly detrimental to the plant, leading to the PRR losing its ability to recognize the PAMP. Moreover, from MM/PBSA analysis it is found that binding energy between HAESA and IDA is more favourable than HAESA and SERk1. -636.824 kJ/mol binding energy is found when co-receptor SERk1 is present inside the complex and during its absence the binding energy between HAESA and IDA goes higher. Again, a notable binding energy between IDA and SERk1 is obtained from this analysis. There are several prominent residues found from HAESA and SERk1, mutation at any of these residues will affect on PTI of Arabidopsis thaliana. Since HAESA has been shown to play a key role in Arabidopsis thaliana plant defense mechanism, its hypothesized binding mechanism with IDA and co-receptor SERk1 will help paint a better understanding of the inceptive stages of PTI.

Dedication

I want to dedicate this work to

My Mother

who is the very first teacher and friend of my life.

Thank you for being beside me every time

ACKOWLEDGEMENT

I proclaim my submissive gratitude to the Creator of the universe, by who's grace I adapted the strength, patience and understanding to complete this thesis work.

I am countlessly indebted to my family members for their unconditional support, prayer and faith on me since my childhood.

I convey my sincere gratitude to Professor A.F.M Yusuf Haider, the Chairperson of the Department of Mathematics and Natural Sciences, late Professor A. A. Ziauddin Ahmed, former Chairperson of the Department of Mathematics and Natural Sciences and Professor Naiyyum Choudhury, former coordinator of the Biotechnology and Microbiology Program, for giving me the opportunity to continue my research work in my desired subject and to grow up with outstanding guidance during my study life in Brac University.

I would like reveal my gratitude to my supervisor, M H M Mubassir, Lecturer, Biotechnology Program, Department of Mathematics and Natural Sciences for his gentle and humble guidance, supervision and cooperation. His intellectual vision and constant support lead me to complete my research work through hardship. I will always grateful to him for giving me the worthy opportunity.

I would like to thank specially Tasfia Tawhid for her constant psychological support during the research work and her belief on my abilities make me motivated to do work harder. Moreover, I want to thank my seniors as well as my mates for their supports.

Sincerely,

Raghib Ishraq Alvy

Department of Mathematics and Natural Sciences

Brac University, Dhaka

Table of Contents

List of Figures

List of Tables

List of Acronyms

AA - Amino Acid

ETI - Effector Triggered Immunity

GROMACS - Groningen Machine for Chemical Simulations

H-bonds - Hydrogen Bonds

LRR - Leucine Rich Repeat Domain

MD - Molecular Dynamics

PAMP - Pathogen Associated Molecular Pattern

PDB - Protein Data Bank

PRR - Pattern Recognition Receptor

PRR-RKs - Pattern Recognition Receptor like kinases

PTI - Pattern Triggered Immunity

Rg - Radius of Gyration

RMSD - Root Means Square Deviation

RMSF - Root Means Square Fluctuations

List of Symbols

Aº - Angstrom

K - Kelvin

nm - Nano meter

ns - Nano second

ps - Pico second

Chapter 1

Introduction:

1.1 Background Study

Plants are surrounded innumerable pathogenic microbes but they remain green. To detect and eliminate these perilous pathogens, plants use their innate immune system.[1-4]

Plant pathogens utilize assorted life techniques. Bacteria, virus, fungi these types of pathogens multiply in intercellular spaces (the apoplast) subsequent to entering through gas or water pores (stomata and hydathodes, individually), or get access by means of wounds. Plants, in contrast to well-evolved creatures (like a human being), have no versatile protector cells and a physical versatile resistant framework. Rather, they depend on the natural resistance of each cell and on foundational signals exuding from contamination destinations. [5]

The staggeringly sensitive impression of pathogen or microbe-associated molecular patterns (PAMPs or MAMPs) through pattern recognition receptors (PRRs) at the plant's cell surface is the first one of this system. For an example, for detecting bacterial flagellin, plants get help from a pattern recognition receptors (PRR) 'FLS2 (flagellin sensitive 2)' a leucine-rich repeat receptor kinase (LRR-RK) situated in the plasma membrane. Here comes the term 'PRR triggered immunity (PTI)' which defines reactions to different MAMPs and flagellin. [2]

Alongside PTI, Effector Triggered Immunity (ETI) is another immune response system for plants where Intracellular Immune receptors are utilized and most of them are nucleotide-binding site leucine-rich repeat (NBS-LRR) proteins, which are used to differentiate directly or indirectly discharged harmfulness effectors. [6]

Thus, the first one of the plant immune system is transmembrane pattern recognition receptors (PRRs) that react step by step creating microbial or pathogen-associated molecular patterns (MAMPS or PAMPs), like flagellin, and the second one works inside the cell by considering all things, using the polymorphic NB-LRR protein things encoded by most R genes. [5, 7]

1.2 Research aim and objectives

The actual intention of this research paper is to understand pattern triggered immunity (PTI) in Arabidopsis thaliana, a model plant which mediates recognition with HAESA receptors using Bioinformatics perspective. To be more specific:

- Analyzing the interaction between PRR HAESA, PAMP IDA and Co-receptor SERK1 by using Molecular Dynamic (MD) Simulation and free energy binding calculation
- Contribution of SERK1 Co-Receptor in PTI mediated by HAESA-IDA complex
- Analyzing structural dynamics of HAESA-IDA-SERK1 complex

1.3 Literature review

In this section, simple and general overview of pattern triggered immunity, the targeted model plant and its' morphology, some terminologies according to pattern triggered immunity is discussed. Computational approaches and analyzing terminologies are also overviewed at the end of this section.

1.3.1 Arabidopsis Thaliana

Arabidopsis Thaliana, a flowering plant recently becomes an essential model system to analyze genes and their functional system.[8] In-plant science, for doing wide range studies A. Thaliana becomes a popular choice in these days.[9] Recent research of Arabidopsis shows that to do basic analysis of the structural and functional area of biology for most of the eukaryotes this angiosperm is the best choice. As most of the eukaryotic organisms have a common genetic ancestry and this was decided while genome project.[10]

Having a very small nuclear (sn) genome make Arabidopsis more desirable model for plant genomics.[11] For the small genome, this angiosperm becomes highly resistant to ionizing radiation, the relation between radiation sensitivity and DNA contents in plants was proved around sixty years ago.[12, 13]

From the above, it can be understood easily the reasons behind make Arabidopsis Thaliana a convenient model for genome analysis. Beside small nuclear genome, it has a short generation time, a large number of offspring and obviously small size.[8, 14-16]

1.3.2 Morphology of Arabidopsis Thaliana

It needs almost six weeks to complete the entire life cycle of Arabidopsis thaliana including seed germination, rosette formation, main stem bolting, flowering and first seed maturation. Most of the things are very tiny after coming to a size such as 2mm long flower, self-pollination system because of open bud, pollen can be applied to the stigma surface for crossing.[10]

At maturity seed length is 0.5mm, into the rosette plant seedlings are developed which ranges from 2 to 10 cm in diameter according to growth conditions. Trichomes, small unicellular hairs cover the leaves which are used for studying cellular differentiation and morphogenesis.[10]

This mustard family (Cruciferae or Brassicaceae) member takes three weeks for bolting after planting. Silique forming a central gynoecium, pollen-bearing six stamens, four white petals containing inner whorls and four green sepals containing outer whorl compose flowers of A. thaliana. Almost 5000 total seeds with several hundred siliques are produced when the plants become 15-20 cm in height, which is a mature plant also.[10]

There is no evidence of symbiotic relationship with nitrogen-fixing microbes of the roots and very structure to study in culture.[10]

1.3.3 History of Arabidopsis Thaliana

In 1873, Alexander Braun described a mutant plant near Berlin in the literature of botany, which was the very first Arabidopsis's non-taxonomic appearance.[17] In the AGAMOUS gene, the mutation happened which is known as floral ABC regulators in these days.[17, 18]

By publishing chromosome number of various plants Friedrich Laibach (1885-1967) student of Strasburger's laboratory made Arabidopsis features clear in 1907. Though the chromosomes of this plant are very small.[17, 19]

During Russian expedition in 1935 appearance of Arabidopsis was found again to use in genetics and cytogenetics instead of Drosophila because of small chromosome number and flowering time

is so rapid than others but small size and tiny structure made problems to differentiate chromosome pairs at that time.[17, 20, 21]

In 1943 Arabidopsis again cropped up because Laibach showed that it has a short generation time, ease to crosses, mutagenesis possibility, fecundity and so on features which made it a genetic model organism.[17, 22]

Revolution during the 1980s in plant genetics, molecular genetics and physiology make extensive acceptance of Arabidopsis as a model plant. There were several suggestions at that time for model plant choosing like petunia for its' ease of transformation and haploid lines availability, tomato for the mutant availability but in 1975 George Re´dei proposed Arabidopsis as a model plant for auxotrophic mutation finding, an article of Annual Review of Genetics described it takes attention of molecular cloners and young geneticists. [17, 23]

The 11th International Conference on Arabidopsis research was held in Summer 2001 wherein the presence of around 1000 people extensive acceptance of this plant for laboratory model system was led. A fusion of molecular and classical genetics in plant biology has been brought by this model genetic system with development, physiology and pathology of plants. Information transfer and cellular process mechanisms in plant life became understandable for the first time.[17]

1.3.4 Pattern Triggered Immunity (PTI)

According to the sensation of the microbe-associated or pathogen-associated molecular pattern (MAMPs or PAMPs) by pattern recognition receptor (PRRs), the very first action of plant immune response is triggered, which is Pattern triggered immunity (PTI).[1, 2, 5, 24, 25] To inhibit PTI, effector proteins have been attained by adapted pathogens through evolution which is transferred into the plant cell.[24, 26-28]

At the Period of infection MAMPs (Microbe-Associated Molecular Patterns), DAMPs (damage-associated Molecular Patterns) evolved from the host and so many types of pathogens are released which are recognized by cell surface-localized Pattern Recognition Receptors (PRRs) and these PRRs activate PTI for immune response.[29, 30]

Because of the conserved nature of PAMPs (e.g., bacterial flagellin, fungal chitin), multiple microbes are fended off by PTI. Thus, it contributes to basal immunity during infection.[6]

For most of the non-host pathogens, PTI is effective by providing a strong and abstinent immune response. As a contributor to basal resistance PTI works against adapted pathogens where effectors proteins are employed to knock out PTI for host metabolism manipulation is called Effector-Triggered Susceptibility (ETS). [30]

Current study proved that beside restricting pathogenic growth PTI creates nutrient deprivation to control colonization of microbes.[30, 31]

1.3.5 Effector-triggered Immunity (ETI)

Another immunity system of the plant is specialized in effectors identification by additional receptors, known as effector-triggered immunity (ETI). These receptors are obtained during co-evolution. [5, 24, 32]

Pathogen's effector protein(s) of the pathogen are identified by a Resistance (R) gene product inside the plant cell, in ETI.[5, 24] Most of the R genes carry nucleotide-binding leucine-rich repeat (NB-LRR) proteins.[24, 33] When these NB-LRR proteins identify an effector, ETI starts, Effector identification process of NB-LRR depends on the results of their actions. [5, 24]

Basically, these effectors are the result of adapted pathogens and their intention is to redesign host's physiology for making it contagious compatible, but NB-LRR protein's identification ability creates a war of nerves between plants and their pathogens.[6]

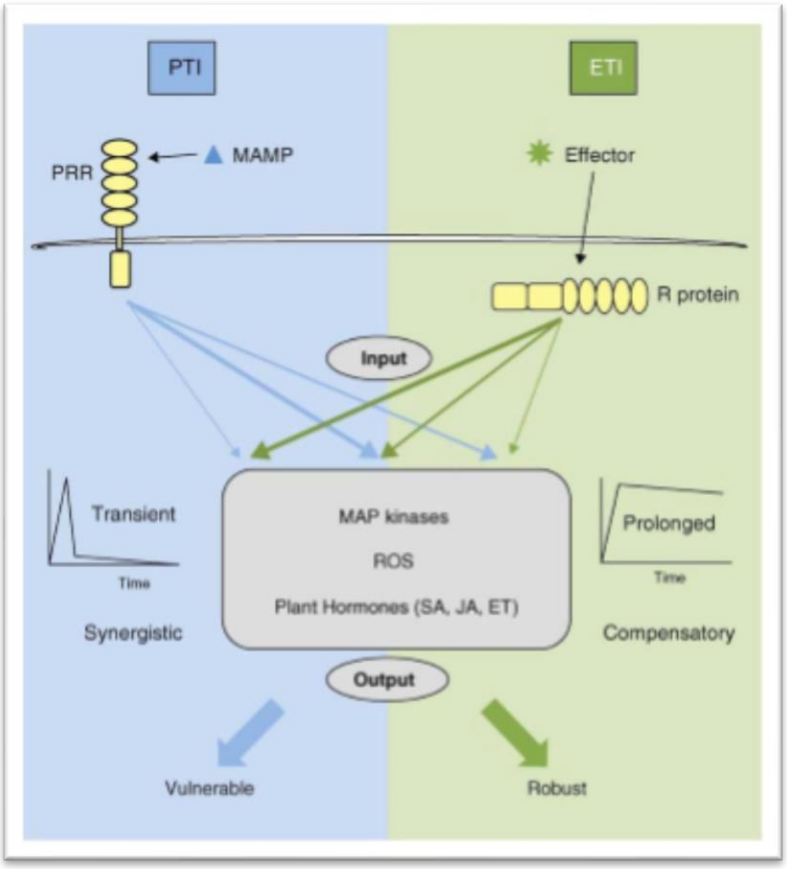

Fig 1.1: Differences in signaling between PTI and ETI. [25]

1.3.6 Pattern Recognition Receptor (PRR)

In Arabidopsis thaliana, more than 200 genes encoding leucine-rich repeat receptor-like kinases and more than 1000 genes encoding putative secreted peptides have been identified, which shows that in plants for a cell to cell communication peptide ligand-receptor interactions are important.[34-36]

LRR-RLK flagellin-sensitive 2 (FLS2) binds to flg22 which is a 22 residues epitope situated at the N terminus of flagellin from gram-negative bacteria.[37]

Inflorescence Deficient in Abscission (IDA)-HAESA / HAESA Like2 (HLS2) mainly control the floral organ abscission. At the base of organs to be shed both the IDA mutant and the HAESA/HLS2 double mutant hold their floral organs indefinitely due to lack of breakdown of the middle lamella between cell layers of the abscission zone.[38]

Formerly named RLK5, Arabidopsis LRR-RLK HAESA is plasma membrane-associated containing serine/threonine-protein kinase activity. Beside abscission zones of the floral organs, HAESA is expressed at the base of the petioles and pedicels.[39]

A structure of HAESA with a PKGV-IDA peptide at 1.94 A resolution has no extra electron density at the IDA N-terminus rather it identifies a dodecamer peptide comprising residues 58-69^{IDA}.[38]

IDL1, an IDA peptide family member can easily be sensed by HAESA, though it has low affinity.[38, 40] Though most of the PRRs are identified and illustrated clearly their mechanism of interaction with PAMPs is not specified yet. Interaction between PAMP IDA with PRR HAESA is studied here to identify PTI mechanism among them.

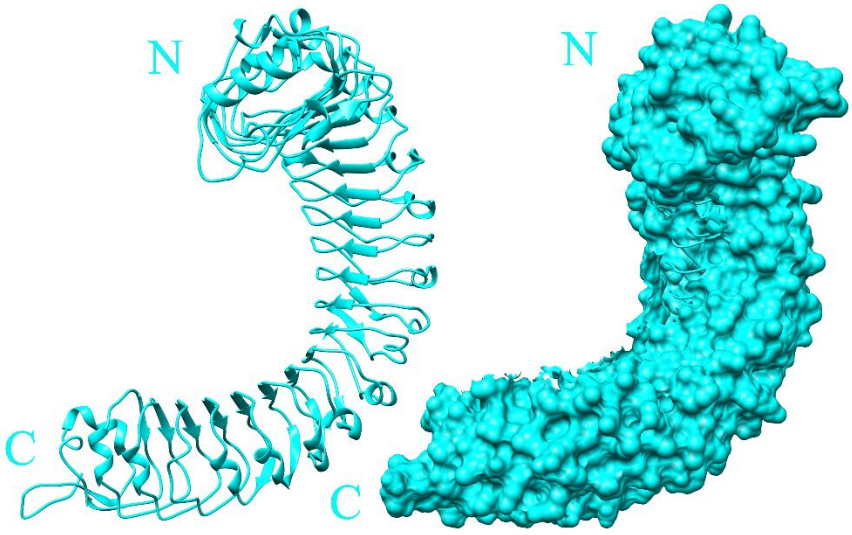

Fig 1.2: Illustration of HAESA LRR in cartoon and surface structural view.

1.3.7 Leucine-Rich Repeat Receptor-like Kinase

Specific receptors are working for conducting signals and sensing outside interaction in the cell surface of living organisms, but in the case of plants, this procedure is induced by receptor-like kinases (RLKs). Leucine-rich repeat RLK family is one of them and also the largest family having two types of domains: Extracellular domain (ECD) and Kinase domain (KD). The differences between ECD and KD are: ECD contains a numerous number of LRR repeats which help to sense small molecules, entire protein or peptides but KD contains 12 subdomains which are basically conserved and fold into three-dimensional catalytic core consisting two-lobed structure and these subdomains are playing an important role in enzyme function. [41, 42]

1.3.8 Pathogen Associated Molecular Pattern (PAMP)

Pathogen Associated Molecular Patterns (PAMPs) are specific molecular pattern of microbes playing elicitor key role of innate immunity in the plant system.[43-46] Most probably specificities of recognizing immensely conserved molecules for different kingdoms gained independently by convergent evolution. [45, 47, 48]In the First stage, PAMPs sensing by the host leads activation of defense system rapidly by reinforcing cell wall by callose deposition, reactive oxygen species (ROS) production etc. These PAMP elicited fundamental defense system can be inhibited by virulence factors of successful pathogens just after the first stage. Finally, specialized resistance (R) proteins have been evolved to identify these virulence factors derived from pathogen and their effectiveness on the host. [45, 49-51]

Result of R protein-dependent perception, a hypersensitive response (HR) occurs that consists of localized cell death and capture of PAMPs expansion.[45, 52, 53]

1.3.9 Inflorescence Deficient in Abscission (IDA)

No longer required, damaged or senescent organs are detached by plants at the period of their growth, development and reproduction known as a cell separation process.[38, 54] Floral abscission is excessively controlled by LRR-RKs HAESA and HAESA LIKE 2 (HLS2).[38-40, 55] Inflorescence Deficient in Abscission (IDA) makes abscised floral organs to stay linked while

overexpression of IDA leads to ill-timed shedding.[38, 56, 57] the Proteolytically full length of IDA is processed and the IDA loss-of-function phenotype can be rescued by 20 amino acids conserved stretch (termed as EPIP).[38, 40] It is shown that a dodecamer peptide

within EPIP has the ability to activate HAESA and HLS2 in temporary research in tobacco cells.[38, 58] This sequence motif presumed to be post-translationally modified to hydroxyproline, contains a central Pro residue and highly conserved among IDA family members (IDA-LIKE PROTEINS, IDLs).[38, 56, 58]

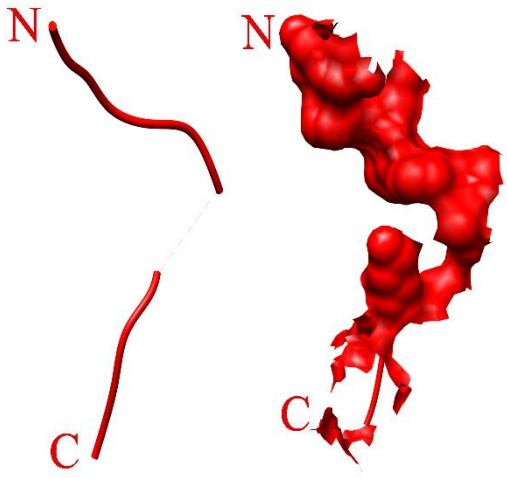

Fig 1.3: Cartoon and surface structural view of IDA

1.3.10 Somatic Embryogenesis Receptor-Like kinase1 (SERk1)

Somatic Embryogenesis Receptor-Like kinase1 (SERk1), an LRR-RLK which works for marking the formation of embryogenic cells in culture and it's expression in ovule primordia both in male and female gametophytes.[59] In vascular tissue of all organ, its' expression is highest but in sporophytic tissues, it expresses a complex pattern.[59-62]

In Arabidopsis plants, ectopic expression of SERK1 has no result in an exact phenotype but in culture helps to increase somatic embryos. It is a member of a small family of five related RLKs, all of which have typical Ser-Pro-rich juxta membrane region.[59, 61]

SERK1 interacts with the KINASE-ASSOCIATED PROTEIN PHOSPHATASE (KAPP) which plays a role in receptor internalization. [59, 63]

Knockout alleles of SERK1 (serk1-1 and serk1-2) have no morphological phenotype but the combination with serk2 null mutant resulted in complete male-sterile plants.[59, 60, 64]

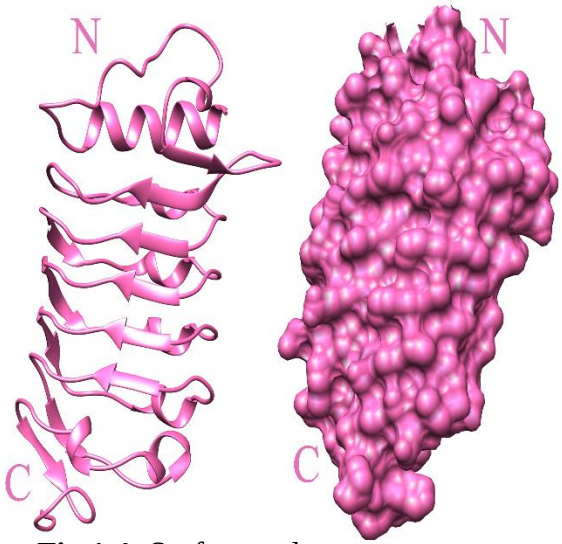

Fig 1.4: Surface and cartoon structural illustration of Co-receptor SERk1

1.3.11 Activation of SERK1 with PRRs

Somatic Embryogenesis Receptor Kinases (SERKs) regulates positively floral abscission and have the ability to interact with HAESA and HLS2 in an IDA-dependent manner.[38, 65] In the Arabidopsis abscission zone all five SERK family members are appeared for expression, by using petal break strength assay relative contribution of SERKs can be quantified.[38, 66, 67] Required force to remove petals of serk1-1 mutants is much higher than that required for wild type plants which was observed for haesa/hls2 at the very first time.[38, 40] In plant development SERKs play an extra role like in pollen formation and brassinosteroid signaling. [38, 60, 64]

The residue of LRR ectodomain SERK1 is from 24-213 forms stable, in vitro. In the presence of SERK1, HAESA gets the ability to identify IDA with a ~60 folds higher affinity, which indicates that SERK1 is working for specific recognition of the peptide hormone. [38]

There was no notable interaction found between HAESA and SERK1 in the absence of IDA, but in the presence of IDA, SERK1 binds with HAESA very tightly.[38, 68-70]

An important fact is hydroyprolination of IDA is critical for HAESA-IDA-SERK1 complex forming and HAESA and other receptor kinases may be rendered by SERKs for their agnate ligand sensing.[38] Though cell surface is the IDA binding site, In the cytoplasm, HAESA and SERK1 may have interacted to active each other.[38, 39, 71, 72]

1.3.12 HAESA-IDA complex

The protein complex structure was collected from “Protein Data Bank” (PDB code: 5IYX). From baculovirus-infected insect cells HAESA ectodomain (residues 20-620) was purified. LRR ectodomain HAESA is bind with the peptide hormone IDA. Crystal structures of apo HAESA and HAESA-IDA complex are respectively at 1.74 and 1.86 A resolution. Along the inner surface of the HAESA ectodomain, IDA binds in a complete extended conformation by covering LRRs 2-14. A specific pocket formed by HAESA LRRs 8-10 interred the central Hyp64 with its hydroxyl group for setting up hydrogen bonds with the strictly conserved Glu266 with a water molecule which severally is coordinated by the primary chain oxygens of Phe289 and Ser311. HAESA LRRs 11-14 with the C-terminal Arg-His-Asn motif in IDA maps to a cavity formation.[38, 73]

COO group of Arg69 contacts directly with Arg407 and Arg409, HAESA have no capability to bind with C-terminally prolonged IDA-SFVN peptide. From there, it can be believed that the conserved Asn69 most probably construct the very C-terminus of the mature IDA peptide inside the plant and this activated IDA is brought forth by proteolytic processing from a longer pre-protein.[38, 40] Loss of function phenotype in HLS2 is occurred by mutation of Arg417 (similar to Arg409) which is the indication of peptide binding pockets in various HAESA receptors carry similar structural and sequence properties.[38, 66]

1.3.13 Computation approach for Molecular Dynamic (MD) simulation

Simulating the motions of a system of specific particles scientifically known as Molecular dynamics. The system is generally tiny like an atom and a diatomic molecule undergoes a chemical reaction larger like a galaxy.[74, 75] For running a molecular dynamics simulation, knowledge of the interaction potentials for the particles is necessary. [75]

For getting conclusive details of specific particle motions as an act of time and answering about the properties of a model system simulations are used. Three different types of simulation method exist in the area of macromolecular research. The first one is working on a mean of sampling configuration space, which helps to define structures with the data obtaining from experiments on actual or real-life systems. The second method is used for describing the system at equilibrium with the help of structural and motional properties and the values of thermodynamic parameters. The actual dynamics are examined by the third method where the motion and development of specific particles with time are the main concern. [76]

For doing meaningful studies on biological macromolecules by using molecular dynamic (MD) simulation, programs availability and computational power must be required. The simulation was doing less than 10 ps in length in old days but the current scenario shows that 1000 times longer simulations of same sizes are often done within half of the time than previous. Besides time, another important fact in simulation is computational power by which multiple numbers of simulation can be performed. CHARMM20, AMBER21 and GROMOS22 are the most widely used programs.[76-79] Sometimes it becomes impossible to get the desired result in the laboratory for not providing extreme and controlled temperature and pressure where computational simulations solve this problem.[80]

1.3.14 Protein Interaction Calculator (PIC)

For a better understanding of sequence-structure relationships, the stability of protein on the basis of structure and evolution of protein atomic interaction analysis in their tertiary structures is needed. Strategies for remote homology identification, recognition of protein folding, comparisons between protein structures and modelling a protein knowledge about sidechain interaction is used.[81-87]This knowledge also helps to do site-directed mutagenesis research and understand the residue conservation in homologous proteins.[87-89]Moreover, interaction between subunits of multimeric proteins and also between interacting protein modules are the areas of acute knowledge. [87, 90-94]Characteristics of sidechain-sidechain interactions between protein modules give us an indication of evolutionary conservation of protein-protein interactions.[87, 95-97]Protein Interaction Calculator (PIC) is a web-based server which is working for analyzing different types of interactions in tertiary structures of proteins and protein-protein complexes. It has also solvent accessibility calculator to recognize of interacting motifs those are exposed. Basically, PIC server supports the atomic coordinate set of protein structure in the standard format which is PDB or Protein Data Bank format.[87]

1.3.15 MM/PBSA

Calculation of the free energy of the binding of tiny ligands to biological macromolecules is done by molecular mechanics energies combined with the Poisson-Boltzmann and surface area continuum solvation (MM/PBSA) method. This method is basically depending on molecular dynamics simulations of the receptor-ligand complex. [98]Free energy calculation is used for drug design and determining protein structure. [99, 100]Particularly MM-PBSA combines molecular mechanics and continuum solvent models to calculate affinities of ligand binding. [98]The MM-PBSA method has been developed and modified according to the type of uses. [98, 101-103]The method is used for protein designing, protein-protein interactions, conformer stability and re-scoring. [98, 102, 104-109]Here MM stands for molecular mechanics, PB for Poisson-Boltzmann and SA for surface area. In this method, three different energy values as output are given according to their objectives.[110]

1.3.16 SASA, RMSD, RMSF, Rg

Solvent accessible surface area previously designed in a way that traced out by the centre of a probe sphere representing a solvent molecule which rolled over the surface of the molecule of interest. This invention used as a tool for attacking the protein folding problem. [111, 112]Only measuring a quantity area is not sufficient for further studies so another method was developed. On that method without displacing van der Waals surface of the atoms which are accessible to the contact surface, make a connection by a network of concave and saddle-shaped surfaces.[111-113]

Latterly the method is developed again by which a minor problem is solved. The problem is if the probe sphere is excluded and not to experience van der Waals overlap then how to calculate SASA of that molecule? By improving algorithms for doing calculation of contact and reentrant surfaces and also for the solvent accessible area which helps to calculate SASA of any molecules.[111, 114-116]

Root mean square deviation (RMSD) is used to analyze the stability of protein during interaction for a several period of time which helps to understand the proteins stability character. Root mean square fluctuation (RMSF) is also analyzed like RMSD but in this case fluctuation of individual residues are studied. Low fluctuation indicates the possibility of more interaction. Radius of gyration (Rg) is used to determine the compactness of the protein. Coiling and uncoiling condition in the mean of the time is analyzed by Rg value.

1.3.17 Hydrogen Bond

The most important interatomic interaction in protein folding is the hydrogen bond.[117, 118] Comparing to covalent bond average intermolecular hydrogen bond energy is very low but their massive number of existences gives a major influence on protein folding in another sense it dominates the folding procedure but their role is mostly reserved for hydrophobic interaction. [117-120]

Most of the hydrogen bonds in proteins are main chains NH to CO bonds and bonds between main-chains and side-chains make a cluster around the caps of helices. Failure of main-chain NH or CO groups to form hydrogen bonds is very rare.[118]

Chapter 2

Materials and Methods

In this part procedures which were followed for analyzing the interaction between HAESA-IDA-SERK1 is described. For analyzing different types of open-source computational tool and servers were used. Those are GROMACS 5.0, MM-PBSA, RCSB Protein Data Bank, Protein Interaction Calculator (PIC), XmGrace, Chimera and so on. All of the softwares are developed for the open-source platform and the analyzing procedures are performed on Linux based Ubuntu operating system.

2.1 Molecular dynamic simulation of HAESA LRR, PAMP IDA and Co-Receptor SERk1:

To begin with, the first protein containing the receptor ectodomain HAESA incorporation with the PAMP IDA and Co-Receptor SERk1 (PDB code: 5IYX) was collected from Protein Data Bank web server. The file was downloaded as '.pdb' format. The PDB file was opened in a text editor and it was edited by keeping all the residues of proteins inside. After saving the file, PIC data were collected in accordance with protein-protein interaction search category. Then the PDB file was subjected to molecular dynamic (MD) simulation with GROMACS software suite.[121] Like a force field, GROMOS 54a7 united force field was selected for this simulation. The system was solvated, neutralized, energy minimized and equilibrated accordingly. During solvation, protein complex was inserted into a cubic box with minimum distance 1Å ranging from protein surface to the edges. The box newly created with the protein complex inside was solvated with SPC water model.[122] Before proceeding to energy minimization, the system was neutralized with genion tool of GROMACS. During equilibration, 1 ns NPT coordinates followed by 1 ns NVT coordinates were equilibrated while keeping constant 1 atm pressure and 300 K temperature. The output file generated as 'md_0_1.gro'. Given gro file was converted from 'md_0_1.gro' to 'md_0_1.pdb' using GROMACS tool. Again, solutions were removed apart from residues through text editor and updated the save file. The pdb file was used for PIC in case of protein-protein interaction data collection.

2.2 Analysis of binding mode of HAESA LRR with PAMP IDA and Co-receptor SERk1:

Using Chimera, a molecular visualization system, the intermolecular interactions among HAESA LRR, IDA and SERk1 co-receptor were visualized in the complex. H-bonds, hydrophobic interactions, ionic interactions, aromatic interactions and cation-Pi interactions were measured using the tool protein interaction calculator (PIC). All these analyses were done for the structural complex before and after the simulation. Binding free energy calculations were done using the g_mmpbsa tool.

2.3 Comparative study between HAESA LRR-IDA-SERk1 complex FLS2 LRR-flg22-BAK1 and TDR-TDIF complex

To compare the binding mechanism between HAESA LRR-IDA-SERk1 complex with FLS2 LRR-flg22-BAK1 crystal complex[123] and TDR LRR-TDIF complex[124], the crystal structure of FLS2 LRR-flg22-Bak1 (PDB ID: 4mn8) and TDR LRR-TDIF (PDB ID: 5GIJ) were obtained from the protein data bank. Then the structure which most resembled the binding mechanism of 4mn8 was selected using Chimera of the complex obtained from the protein data bank of HAESA LRR-IDA-SERk1; then a comparative study of binding conformation between HAESA LRR-IDA-SERk1 complex and FLS2 LRR-flg22 complex was conducted alongside TDR LRR-TDIF crystal complex

Chapter 3

Result and Discussion

In this chapter, interactions between HAESA LRR with PAMP IDA and co-receptor SERk1 were analyzed, also illustrated with graphs and figures. Finally, molecular interactions between the three proteins complex before and after MD simulation, molecular interactions between HAESA-IDA before and after MD simulation and molecular interaction between HAESA-SERk1 was calculated and described at the end of this chapter. 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 3.1.1: The predicted binding free energies and the individual energy components for the studied systems (kJ/mol)

Complex	ΔE_{vdw} [a]	ΔE_{elec} [b]	ΔE_{polar} [c]	ΔE_{sasa} [d]	ΔE_{bind} [e]
HAESA+IDA	-248.577± 22.518	-899.515± 69.213	542.589± 94.864	-31.321± 2.992	-636.824± 41.829
HAESA+SERk1	-391.376± 31.304	128.769± 107.256	1235.945± 48.734	-53.971± 7.029	919.368± 112.720
IDA+SERk1	-72.693± 12.840	-557.404± 52.415	317.766± 86.629	-11.585± 3.616	-323.915± 62.562
HAESA+SERk1	-204.673± 20.065	521.165± 103.766	473.147± 144.563	-24.226± 5.885	765.412± 125.846
HAESA+IDA	-181.989± 26.706	-891.846± 149.439	447.872± 202.042	-24.749± 5.077	-650.712± 57.583
IDA+SERk1	-54.883± 40.691	-422.722± 191.277	215.321± 158.997	-8.071± 6.514	-270.355± 118.330

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

From MM/PBSA calculation (Table 3.1.1), obtained binding energies between HAESA, IDA and SERk1 are -636.824 $kJmol^{-1}$ (when HAESA interacts with IDA) and -323.915 $kJmol^{-1}$(when IDA interacts with SERk1). Besides these interactions, HAESA interacts with IDA in the absence of SERk1 and their binding energy is -650.712 $kJmol^{-1}$, which is another notable interaction. But in case of absence and presence of IDA, there is no interaction happened between HAESA and SERk1, because in that situation the binding energies are 919.368 $kJmol^{-1}$ and 765.412 $kJmol^{-1}$

respectively. From this overview, it can be easily said that for happening interaction between HAESA and IDA there is no negligible contribution of co-receptor SERk1.

Again, in the case of intermolecular Van Der Waals energy, Electrostatic energy highly favourable value can be observed for interacting between HAESA, IDA and SERk1. HAESA interacts easily with IDA when SERk1 is present in the complex and when SERk1 is absent polar solvation energy of HAESA and IDA goes lower which makes it more interacting complex. There is another point of view that indicates different situation of interacting HAESA with SERk1 in the absence of IDA, because from intermolecular polar solvation energy it is found that when IDA is present in the complex polar solvation is not favourable for HAESA and SERk1 complex but in case of absence of IDA it gives much lower polar solvation energy, which indicates polar solvation energy also takes part in binding affinity of HAESA and SERk1 in the absence of IDA. But for happening interaction between HAESA and IDA, Co-Receptor SERk1 needs not to be present in the complex rather the presence of SERk1 causes expansion the polar solvation energy of HAESA and IDA, which makes it lower binding energy between these two proteins than another phase where SERk1 is absent.

For in-depth analysis different energy contribution for every single residue is calculated. From this calculation it is found that Asp120, Glu123, Glu145, Glu191, Asp242, Glu263, Glu266, Glu310, Glu316, Glu335, Glu382 from HAESA and Tyr56, Lys66, Arg67 from IDA play an important role by forming MM energy for interacting between themselves. Where MM energy is a combination of Van Der Waals energy, Electrostatic energy and covalent bond. When SERk1 is present in the complex Glu123, Asp242, Glu335 of HAESA and Tyr56, Arg67 of IDA give comparatively more non-polar energy than others which indicates these residues are responsible for van der Waals interaction between HAESA and IDA. Though maximum energy comes from electrostatic at this condition. But in the absence of SERk1, there is no negligible non-polar energy for Glu335 of HAESA, other mentioned residues of HAESA and IDA show higher value as before. Moreover, when HAESA interacts with SERk1 prominent residues of HAESA are Lys337, Arg407, Arg503, Lys523, Lys571 and for SERk1 Arg37, Lys142, Arg144. Lys571 from HAESA only gives negligible non-polar energy from SERk1 there is no residue which gives minimum non-polar value. So only from HAESA beside electrostatic energy van der Waals energy also works to interact with SERk1. From (Table 3.1.2 – 3.1.7) Glu123 of HAESA shows maximum binding free

energy for interacting with IDA which is -53.2099 ± 0.2607 $kJmol^{-1}$ in the presence of SERk1 and -40.5769 ± 0.6238 $kJmol^{-1}$ in the absence of SERk1. As opposed to Arg144 from SERk1 shows maximum binding free energy when it interacts with HAESA which is -100.4988 ± 0.4003 $kJmol^{-1}$ in the presence of IDA and -102.6319 ± 0.4762 $kJmol^{-1}$ in the absence of IDA. In both cases, electrostatic interaction is more prominent than van der Waals interaction.

In Graphical view, the mentioned residues from HAESA and IDA play similarly an important role, whereas, Asp361 from HAESA interacts with IDA much favourably when SERk1 is absent in the complex, but other residues show favourable interaction in the presence of SERk1. From IDA interaction of Lys66 is more favourable in the presence of SERk1 but Tyr56 gives favourable interaction in the absence of SERk1. For HAESA and SERk1 interaction, all of the residues give more favourable condition when IDA is present in the complex but from SERk1 Lys142 and Arg144 are more favourable when IDA is not present in the complex.

From Protein Interaction Calculation (PIC) (Table 3.3.1 – 3.3.21) for HAESA and IDA interaction, those residues which are found favourable from MM/PBSA calculation show participation in different interactions before and after the simulation process. Tyr56 and Arg67 from IDA participate in the protein-protein main chain – side chain hydrogen bond interaction and only Tyr56 participates in hydrophobic interaction for both cases (presence and absence of SERk1 in the complex). Similarly, for HAESA only Glu123 plays a vital role and participate in different interaction before and after simulation, basically for hydrogen bond. For HAESA and SERk1 complex, Lys571 of HAESA participates in protein-protein main chain-side chain hydrogen bond interaction with the residues of SERk1 before and after simulation, besides this very little amount of ionic interaction is happened with the residues of SERk1 by Lys571 of HAESA when there is IDA present inside the complex but in absence of IDA besides these interactions Lys571 of HAESA participates in protein-protein cation-pi interaction with residues of SERk1. From SERk1 only Arg144 participates in ionic interaction with the residues of HAESA after the simulation process when IDA is present in the complex but in the absence of IDA, it participates in protein-protein side chain-side chain hydrogen bond interaction after the simulation alongside ionic interaction.

From the overview of MM/PBSA calculation alongside an in-depth analysis of energy contribution of every single residue and with the help of Protein interaction calculation (PIC) data there are

some specific residues from these three proteins are found which are responsible for causing interaction between HAESA, IDA and SERk1. For HAESA, Glu123 is the most responsible residue for making interaction with IDA. From MM/PBSA it gives maximum MM energy than any other residues of HAESA alongside negligible non-polar energy. Again, from PIC this residue also participates in different hydrogen bond interaction beside hydrophobic interaction. So Glu123 from HAESA gives electrostatic energy for rising binding affinity with IDA with some other interactions. For IDA, though Arg67 gives more MM energy for interacting with HAESA Tyr56 participates in more interactions. These two residues from IDA are responsible for making interaction with HAESA.

For another case, Lys571 and Lys523 from HAESA are responsible for interacting with SERk1. But between these residues, Lys571 is the most responsible in every portion of calculations. In the presence of IDA, it gives -138.4027 $kJmol^{-1}$ MM energy which comes from electrostatic energy and few contributions of van der Waals energy and in the absence of IDA, it shows -157.1811 $kJmol^{-1}$ MM energy. These two values are higher in contrast of contribution of other residues of HAESA. Moreover, Lys571 also participates in hydrogen bond interaction. Arg144 from SERk1 is another residue responsible for making interaction with HAESA. -159.6342 $kJmol^{-1}$ MM energy makes it more vital as the contribution is higher than the other residues of SERk1 in the presence of IDA when IDA is absent in the complex the MM energy is -154.3252 $kJmol^{-1}$, which is also higher on that condition. Besides this energy contribution, Arg144 also participates in ionic interaction with HAESA alongside hydrogen bond.

So, from HAESA Glu123 and Lys571 from IDA Tyr56 and Arg67 and from SERk1 only Arg144 are the most responsible and vital residues for making the interaction between these three proteins. Though there are other important residues in case of a contribution of Van Der Waals energy and Electrostatic energy and different hydrogen bond and hydrophobic interaction as well as cation-pi interaction and ionic interaction these five residues play the most important role which makes a favourable condition for interacting between HAESA, IDA and SERk1.

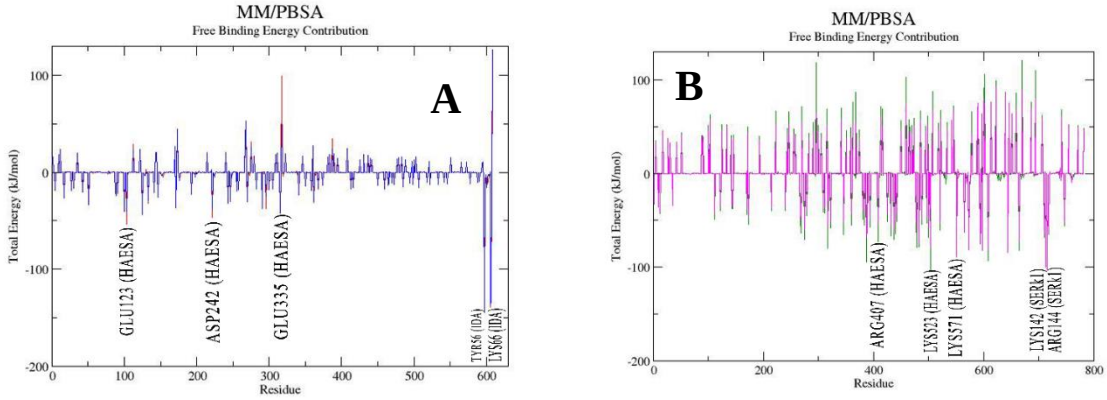

Fig 3.1.1: (A) MM/PBSA total energy value from 30ns MD trajectories. HAESA and IDA complex (Red) presence of Co-receptor SERk1 inside the complex. HAESA and IDA complex (Blue) absence of Co-receptor SERk1. (B) MM/PBSA total energy value from 30ns MD trajectories. HAESA and SERk1 complex (Green) presence of PAMP IDA inside the complex. HAESA and SERk1 complex (Purple) absence of PAMP IDA.

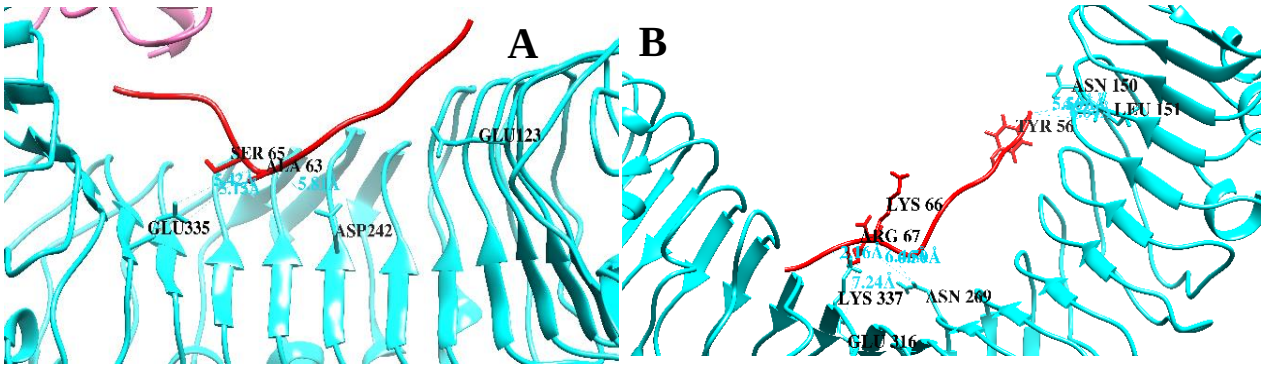

Fig 3.1.2: Cartoon structural view of prominent residues for interacting between HAESA and IDA during the presence of Co-receptor SERk1 where the interaction distance is calculated for H-bond. (A) Prominent residues of HAESA LRR. (B) Prominent residues of PAMP IDA.

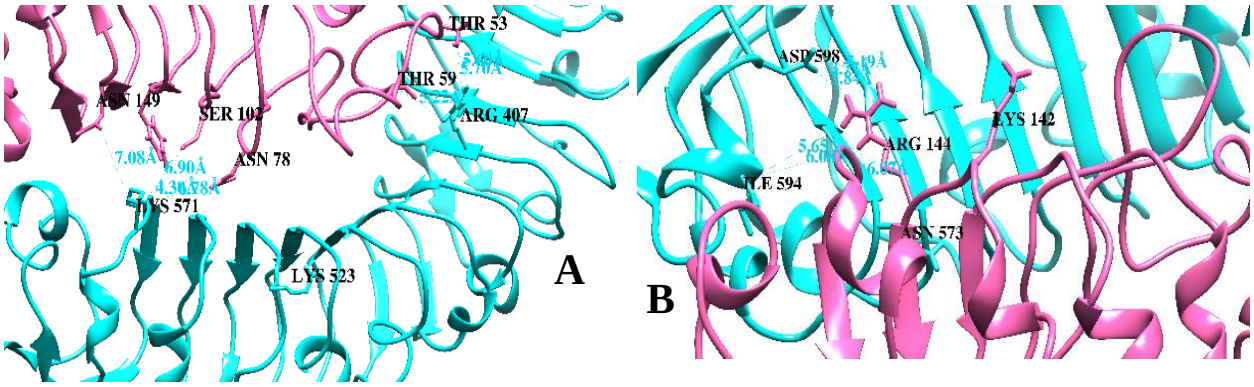

Fig 3.1.3: Cartoon structural view of prominent residues for interacting between HAESA and SERk1 during the presence of PAMP IDA where the interaction distance is calculated for H-bond. (A) Prominent residues of HAESA LRR. (B) Prominent residues of Co-Receptor SERk1.

Table 3.1.2. Binding free energy contribution of the key binding-site residues calculated from the binding energy decomposition for HAESA ($kJmol^{-1}$) from HAESA-IDA interaction. Marked residues are from three protein complex.

Residues	MM Energy	Polar Energy	Apolar Energy	Total Energy
ASP-24	-17.7658 ± 0.0317	0.0767 ± 0.0029	0.0000 ± 0.0000	-17.6893 ± 0.0287
ASP-24	-19.2659 ± 0.0614	0.1783 ± 0.0046	0.0000 ± 0.0000	-19.0865 ± 0.0595
ASP-37	-24.5349 ± 0.0865	0.1606 ± 0.0084	0.0000 ± 0.0000	-24.3775 ± 0.0839
ASP-37	-26.6638 ± 0.1366	0.4815 ± 0.0161	0.0000 ± 0.0000	-26.1901 ± 0.1241
ASP-47	-16.9813 ± 0.0374	0.0187 ± 0.0009	0.0000 ± 0.0000	-16.9623 ± 0.0371
ASP-47	-18.9686 ± 0.0716	0.0682 ± 0.0021	0.0000 ± 0.0000	-18.9008 ± 0.0690
ASP-50	-14.6373 ± 0.0290	0.0047 ± 0.0002	0.0000 ± 0.0000	-14.6322 ± 0.0294
ASP-50	-16.4523 ± 0.0509	0.0167 ± 0.0005	0.0000 ± 0.0000	-16.4343 ± 0.0516
ASP-62	-16.6779 ± 0.0265	0.0196 ± 0.0007	0.0000 ± 0.0000	-16.6577 ± 0.0268
ASP-62	-20.2631 ± 0.0575	0.0717 ± 0.0020	0.0000 ± 0.0000	-20.1927 ± 0.0571
ASP-71	-26.1143 ± 0.0742	0.2689 ± 0.0121	0.0000 ± 0.0000	-25.8461 ± 0.0667
ASP-71	-34.8983 ± 0.2418	1.2048 ± 0.0430	0.0000 ± 0.0000	-33.6960 ± 0.2020
ASP-108	-22.4073 ± 0.0569	0.1330 ± 0.0056	0.0000 ± 0.0000	-22.2734 ± 0.0534
ASP-108	-21.2874 ± 0.0526	0.2325 ± 0.0070	0.0000 ± 0.0000	-21.0537 ± 0.0478
ASP-109	-23.0774 ± 0.0629	0.2595 ± 0.0092	0.0000 ± 0.0000	-22.8174 ± 0.0594
ASP-109	-24.6091 ± 0.0613	0.6515 ± 0.0138	0.0000 ± 0.0000	-23.9542 ± 0.0534
ASP-111	-21.7888 ± 0.0389	0.1860 ± 0.0064	0.0000 ± 0.0000	-21.6023 ± 0.0343
ASP-111	-20.2643 ± 0.0440	0.2458 ± 0.0063	0.0000 ± 0.0000	-20.0192 ± 0.0391
ASP-120	-36.9090 ± 0.1277	1.0336 ± 0.0487	0.0000 ± 0.0000	-35.8811 ± 0.1173
ASP-120	-44.8829 ± 0.2634	4.5455 ± 0.1172	0.0000 ± 0.0000	-40.3241 ± 0.1866
GLU-123	**-58.2884 ± 0.5745**	**5.4351 ± 0.6328**	**-0.3673 ± 0.0238**	**-53.2099 ± 0.2607**
GLU-123	**-99.7720 ± 1.4265**	**60.0211 ± 1.9570**	**-0.9540 ± 0.0164**	**-40.5769 ± 0.6238**
GLU-145	-36.9753 ± 0.1226	0.4398 ± 0.0342	-0.0002 ± 0.0001	-36.5363 ± 0.1286
GLU-145	-45.0630 ± 0.2661	1.7588 ± 0.0854	-0.0050 ± 0.0014	-43.3069 ± 0.2431
ASP-153	-33.4400 ± 0.0894	1.2326 ± 0.0442	0.0000 ± 0.0000	-32.2073 ± 0.0548
ASP-153	-30.3659 ± 0.1404	1.3300 ± 0.0547	0.0000 ± 0.0000	-29.0399 ± 0.0944
GLU-161	-21.3922 ± 0.0292	0.1220 ± 0.0037	0.0000 ± 0.0000	-21.2718 ± 0.0270
GLU-161	-20.4571 ± 0.0474	0.1703 ± 0.0079	0.0000 ± 0.0000	-20.2879 ± 0.0417
GLU-166	-24.5358 ± 0.0645	0.1917 ± 0.0061	0.0000 ± 0.0000	-24.3450 ± 0.0629
GLU-166	-25.9645 ± 0.0812	0.1992 ± 0.0139	0.0000 ± 0.0000	-25.7641 ± 0.0810
GLU-191	-33.7732 ± 0.0940	0.3541 ± 0.0184	0.0000 ± 0.0000	-33.4198 ± 0.0936
GLU-191	-36.9471 ± 0.1477	0.1812 ± 0.0449	0.0000 ± 0.0000	-36.7669 ± 0.1579
GLU-213	-22.0366 ± 0.0362	0.2478 ± 0.0060	0.0000 ± 0.0000	-21.7892 ± 0.0318
GLU-213	-22.9510 ± 0.0563	0.0966 ± 0.0096	0.0000 ± 0.0000	-22.8558 ± 0.0561
ASP-242	**-59.1789 ± 0.2378**	**12.4402 ± 0.3999**	**-0.0570 ± 0.0040**	**-46.8052 ± 0.2529**
ASP-242	**-59.7577 ± 0.6153**	**22.2094 ± 0.8417**	**-0.2991 ± 0.0117**	**-37.8606 ± 0.3396**

GLU-263	-29.3044 ± 0.0951	1.2464 ± 0.0234	0.0000 ± 0.0000	-28.0577 ± 0.0806
GLU-263	-31.3711 ± 0.1619	-0.2027 ± 0.0390	-0.0002 ± 0.0001	-31.5772 ± 0.1723
GLU-266	-70.4221 ± 0.3560	40.2983 ± 0.7170	-0.3076 ± 0.0065	-30.4004 ± 0.4527
GLU-266	-61.2544 ± 0.9119	33.2274 ± 1.4418	-0.3651 ± 0.0148	-28.3875 ± 0.6188
GLU-275	-19.8343 ± 0.0425	0.4032 ± 0.0148	0.0000 ± 0.0000	-19.4304 ± 0.0318
GLU-275	-19.1931 ± 0.0739	0.2247 ± 0.0116	0.0000 ± 0.0000	-18.9676 ± 0.0639
GLU-278	-18.3506 ± 0.0356	0.2847 ± 0.0083	0.0000 ± 0.0000	-18.0656 ± 0.0294
GLU-278	-18.4403 ± 0.0706	0.1515 ± 0.0089	0.0000 ± 0.0000	-18.2869 ± 0.0654
ASP-290	-65.0326 ± 0.2794	46.4904 ± 0.5476	-0.2530 ± 0.0038	-18.8044 ± 0.3712
ASP-290	-52.6715 ± 0.5377	22.8081 ± 0.8383	-0.2086 ± 0.0101	-30.0661 ± 0.5421
ASP-302	-16.6940 ± 0.0276	0.1412 ± 0.0074	0.0000 ± 0.0000	-16.5530 ± 0.0232
ASP-302	-17.4680 ± 0.0524	0.0895 ± 0.0069	0.0000 ± 0.0000	-17.3798 ± 0.0479
GLU-310	-31.7668 ± 0.1257	0.5941 ± 0.0303	0.0000 ± 0.0000	-31.1699 ± 0.1125
GLU-310	-37.1414 ± 0.3145	-0.2016 ± 0.0779	0.0000 ± 0.0000	-37.3364 ± 0.2975
GLU-316	-41.8801 ± 0.2000	4.7216 ± 0.1762	-0.0033 ± 0.0008	-37.1572 ± 0.1852
GLU-316	-24.4655 ± 0.3190	-0.5458 ± 0.1181	-0.0043 ± 0.0015	-25.0146 ± 0.2682
GLU-320	-18.2401 ± 0.0369	0.0962 ± 0.0152	0.0000 ± 0.0000	-18.1444 ± 0.0266
GLU-320	-17.8162 ± 0.0826	0.1092 ± 0.0105	0.0000 ± 0.0000	-17.7079 ± 0.0736
GLU-325	-15.2162 ± 0.0299	-0.0550 ± 0.0060	0.0000 ± 0.0000	-15.2714 ± 0.0269
GLU-325	-17.6311 ± 0.0687	0.0256 ± 0.0132	0.0000 ± 0.0000	-17.6082 ± 0.0618
GLU-335	**-45.8056 ± 0.2872**	**5.6822 ± 0.3978**	**-0.1210 ± 0.0050**	**-40.2675 ± 0.2599**
GLU-335	**-44.4852 ± 0.6641**	**2.8126 ± 0.7537**	**-0.0555 ± 0.0063**	**-41.7417 ± 0.4125**
ASP-361	-37.1579 ± 0.2828	38.2959 ± 0.4109	-0.5370 ± 0.0067	0.5993 ± 0.3283
ASP-361	-24.6077 ± 0.2773	1.2345 ± 0.2153	-0.0481 ± 0.0050	-23.4266 ± 0.2807
GLU-370	-13.3487 ± 0.0271	-0.3334 ± 0.0059	0.0000 ± 0.0000	-13.6823 ± 0.0271
GLU-370	-14.5145 ± 0.0463	-0.0093 ± 0.0036	0.0000 ± 0.0000	-14.5229 ± 0.0445
GLU-378	-17.7845 ± 0.0700	-0.2342 ± 0.0048	0.0000 ± 0.0000	-18.0185 ± 0.0712
GLU-378	-18.7079 ± 0.0914	-0.0286 ± 0.0050	0.0000 ± 0.0000	-18.7392 ± 0.0886
GLU-382	-25.5684 ± 0.2289	-5.9103 ± 0.1328	-0.0082 ± 0.0016	-31.4784 ± 0.2160
GLU-382	-29.7281 ± 0.1937	-0.2552 ± 0.0458	0.0000 ± 0.0000	-29.9793 ± 0.1841
ASP-388	-9.2420 ± 0.0903	-3.9487 ± 0.0690	0.0000 ± 0.0000	-13.1887 ± 0.0464
ASP-388	-17.3564 ± 0.1398	-0.1045 ± 0.0186	0.0000 ± 0.0000	-17.4565 ± 0.1270
GLU-394	-11.9943 ± 0.0253	-0.3420 ± 0.0049	0.0000 ± 0.0000	-12.3366 ± 0.0242
GLU-394	-13.6290 ± 0.0414	-0.0104 ± 0.0024	0.0000 ± 0.0000	-13.6378 ± 0.0402
GLU-433	-7.0215 ± 0.0732	-5.7321 ± 0.0624	0.0000 ± 0.0000	-12.7498 ± 0.0504
GLU-433	-17.4499 ± 0.1660	-0.1940 ± 0.0261	0.0000 ± 0.0000	-17.6411 ± 0.1564
ASP-436	-10.5517 ± 0.0351	-1.2709 ± 0.0164	0.0000 ± 0.0000	-11.8239 ± 0.0277
ASP-436	-15.3434 ± 0.0983	-0.0476 ± 0.0073	0.0000 ± 0.0000	-15.3911 ± 0.0948
GLU-470	-12.2791 ± 0.0258	-0.3267 ± 0.0036	0.0000 ± 0.0000	-12.6058 ± 0.0259
GLU-470	-13.7488 ± 0.0343	-0.0116 ± 0.0014	0.0000 ± 0.0000	-13.7594 ± 0.0334
GLU-479	-11.7331 ± 0.0561	-1.0638 ± 0.0142	0.0000 ± 0.0000	-12.7968 ± 0.0484
GLU-479	-17.2304 ± 0.1377	0.0001 ± 0.0046	0.0000 ± 0.0000	-17.2291 ± 0.1331
GLU-484	-10.7726 ± 0.0217	-0.2394 ± 0.0031	0.0000 ± 0.0000	-11.0126 ± 0.0214

GLU-484	-13.2940 ± 0.0641	-0.0046 ± 0.0010	0.0000 ± 0.0000	-13.2981 ± 0.0643
ASP-486	-10.3414 ± 0.0165	-0.2125 ± 0.0028	0.0000 ± 0.0000	-10.5547 ± 0.0163
ASP-486	-11.9482 ± 0.0326	-0.0075 ± 0.0009	0.0000 ± 0.0000	-11.9550 ± 0.0324
GLU-490	-10.3485 ± 0.0145	-0.0788 ± 0.0013	0.0000 ± 0.0000	-10.4278 ± 0.0136
GLU-490	-11.2745 ± 0.0268	-0.0021 ± 0.0003	0.0000 ± 0.0000	-11.2757 ± 0.0261
GLU-493	-11.5130 ± 0.0185	-0.0716 ± 0.0010	0.0000 ± 0.0000	-11.5850 ± 0.0189
GLU-493	-12.5043 ± 0.0281	-0.0037 ± 0.0004	0.0000 ± 0.0000	-12.5090 ± 0.0275
ASP-505	-11.4446 ± 0.0290	-0.4942 ± 0.0062	0.0000 ± 0.0000	-11.9396 ± 0.0260
ASP-505	-14.5814 ± 0.0756	-0.0040 ± 0.0019	0.0000 ± 0.0000	-14.5851 ± 0.0740
GLU-514	-9.8568 ± 0.0129	-0.0262 ± 0.0004	0.0000 ± 0.0000	-9.8828 ± 0.0130
GLU-514	-10.8598 ± 0.0224	-0.0012 ± 0.0001	0.0000 ± 0.0000	-10.8611 ± 0.0228
GLU-518	-11.4571 ± 0.0214	-0.0306 ± 0.0006	0.0000 ± 0.0000	-11.4873 ± 0.0211
GLU-518	-12.9586 ± 0.0351	-0.0020 ± 0.0002	0.0000 ± 0.0000	-12.9600 ± 0.0339
GLU-527	-12.6749 ± 0.0352	-0.2487 ± 0.0036	0.0000 ± 0.0000	-12.9248 ± 0.0350
GLU-527	-16.0263 ± 0.0961	0.0077 ± 0.0015	0.0000 ± 0.0000	-16.0232 ± 0.0970
GLU-538	-9.8532 ± 0.0132	-0.0151 ± 0.0002	0.0000 ± 0.0000	-9.8684 ± 0.0133
GLU-538	-10.6860 ± 0.0224	-0.0006 ± 0.0001	0.0000 ± 0.0000	-10.6868 ± 0.0228
GLU-542	-11.6502 ± 0.0188	-0.0474 ± 0.0006	0.0000 ± 0.0000	-11.6980 ± 0.0189
GLU-542	-12.7199 ± 0.0313	-0.0028 ± 0.0003	0.0000 ± 0.0000	-12.7225 ± 0.0306
ASP-553	-11.4292 ± 0.0220	-0.0843 ± 0.0011	0.0000 ± 0.0000	-11.5123 ± 0.0218
ASP-553	-12.9313 ± 0.0510	-0.0004 ± 0.0005	0.0000 ± 0.0000	-12.9318 ± 0.0499
GLU-562	-9.7980 ± 0.0142	-0.0088 ± 0.0001	0.0000 ± 0.0000	-9.8067 ± 0.0139
GLU-562	-10.4353 ± 0.0248	-0.0003 ± 0.0000	0.0000 ± 0.0000	-10.4357 ± 0.0244
GLU-566	-11.5225 ± 0.0196	-0.0170 ± 0.0002	0.0000 ± 0.0000	-11.5391 ± 0.0202
GLU-566	-12.1002 ± 0.0339	-0.0010 ± 0.0001	0.0000 ± 0.0000	-12.1013 ± 0.0342
ASP-598	-11.5091 ± 0.0216	-0.0039 ± 0.0002	0.0000 ± 0.0000	-11.5128 ± 0.0220
ASP-598	-11.9515 ± 0.0454	0.0010 ± 0.0001	0.0000 ± 0.0000	-11.9502 ± 0.0462
ASP-608	-10.0750 ± 0.0178	-0.0042 ± 0.0001	0.0000 ± 0.0000	-10.0788 ± 0.0180
ASP-608	-10.3103 ± 0.0273	-0.0001 ± 0.0000	0.0000 ± 0.0000	-10.3109 ± 0.0281
ILE-616	-9.5002 ± 0.0143	-0.0017 ± 0.0000	0.0000 ± 0.0000	-9.5021 ± 0.0146
ILE-616	-10.0265 ± 0.0287	-0.0000 ± 0.0000	0.0000 ± 0.0000	-10.0272 ± 0.0285

Table 3.1.3. Binding free energy contribution of the key binding-site residues calculated from the binding energy decomposition for HAESA ($kJmol^{-1}$) from HAESA-SERk1 interaction. Marked residues are from three protein complex.

Residues	MM Energy	Polar Energy	Apolar Energy	Total Energy
LEU-21	-32.5687 ± 0.0617	0.0810 ± 0.0726	0.0065 ± 0.0061	-32.4813 ± 0.0900
LEU-21	-31.7099 ± 0.0772	0.0941 ± 0.0936	-0.0125 ± 0.0062	-31.6279 ± 0.1209
ARG-29	-35.9141 ± 0.0756	-0.0270 ± 0.0283	0.0022 ± 0.005	-35.9403 ± 0.0800
ARG-29	-35.7214 ± 0.1130	-0.0033 ± 0.0264	0.0011 ± 0.0068	-35.7237 ± 0.1143
LYS-32	-42.6849 ± 0.1208	0.1389 ± 0.0608	0.0017 ± 0.0070	-42.5424 ± 0.1364

LYS-32	-42.1665 ± 0.1512	-0.0152 ± 0.0668	0.0139 ± 0.0067	-42.1612 ± 0.1623
LYS-55	-34.1970 ± 0.0908	-0.0871 ± 0.0737	-0.0065 ± 0.0051	-34.2907 ± 0.1188
LYS-55	-33.9545 ± 0.1117	-0.0313 ± 0.0956	-0.0036 ± 0.0046	-33.9952 ± 0.1396
LYS-132	-48.9207 ± 0.0715	0.0685 ± 0.0583	0.0023 ± 0.0062	-48.8511 ± 0.0910
LYS-132	-41.8256 ± 0.1438	-0.0119 ± 0.0897	-0.0153 ± 0.0054	-41.8528 ± 0.1748
LYS-142	-39.8055 ± 0.0642	-0.0596 ± 0.0653	0.0033 ± 0.0058	-39.8585 ± 0.0938
LYS-142	-35.2700 ± 0.0855	0.0141 ± 0.0909	-0.0050 ± 0.0050	-35.2609 ± 0.1226
ARG-163	-37.5808 ± 0.0485	0.0251 ± 0.0366	-0.0010 ± 0.0053	-37.5589 ± 0.0582
ARG-163	-34.7953 ± 0.0656	0.0313 ± 0.0406	-0.0080 ± 0.0055	-34.7701 ± 0.0781
LYS-164	-37.0673 ± 0.0462	-0.0019 ± 0.0679	0.0030 ± 0.0061	-37.0663 ± 0.0799
LYS-164	-33.9760 ± 0.0638	0.1974 ± 0.0871	0.0036 ± 0.0052	-33.7757 ± 0.1021
LYS-190	-43.2099 ± 0.0970	-0.1944 ± 0.0676	-0.0082 ± 0.0055	-43.4057 ± 0.1256
LYS-190	-37.4229 ± 0.0889	-0.0023 ± 0.0764	0.0008 ± 0.0058	-37.4256 ± 0.1199
LYS-193	-53.9578 ± 0.0925	-0.2446 ± 0.0531	0.0079 ± 0.0064	-54.1947 ± 0.1104
LYS-193	-45.7108 ± 0.1384	0.0920 ± 0.0793	0.0013 ± 0.0074	-45.6138 ± 0.1568
ARG-234	-39.8943 ± 0.0530	-0.0521 ± 0.0342	-0.0076 ± 0.0054	-39.9568 ± 0.0670
ARG-234	-35.8171 ± 0.0797	0.0071 ± 0.0398	0.0021 ± 0.0067	-35.8079 ± 0.0897
LYS-260	-39.1204 ± 0.0546	-0.1273 ± 0.0709	-0.0011 ± 0.0046	-39.2522 ± 0.0933
LYS-260	-34.7367 ± 0.0690	0.0771 ± 0.0826	0.0059 ± 0.0060	-34.6552 ± 0.1013
LYS-287	-49.0873 ± 0.0968	-0.1920 ± 0.0723	-0.0020 ± 0.0047	-49.2729 ± 0.1228
LYS-287	-41.7686 ± 0.1316	0.0995 ± 0.0851	-0.0007 ± 0.0073	-41.6722 ± 0.1563
ARG-288	-62.2906 ± 0.1046	-0.9683 ± 0.0318	-0.0059 ± 0.0066	-63.2634 ± 0.1039
ARG-288	-50.3135 ± 0.1792	-0.0449 ± 0.0327	-0.0029 ± 0.0099	-50.3659 ± 0.1859
LYS-295	-68.6384 ± 0.2379	-1.5897 ± 0.0748	-0.0136 ± 0.0050	-70.2411 ± 0.2620
LYS-295	-59.5407 ± 0.2458	-0.0164 ± 0.0740	-0.0138 ± 0.0061	-59.5641 ± 0.2598
LYS-299	-44.6641 ± 0.0910	-0.2155 ± 0.0717	-0.0008 ± 0.0051	-44.8780 ± 0.1140
LYS-299	-39.0469 ± 0.0878	0.0941 ± 0.0851	0.0069 ± 0.0064	-38.9571 ± 0.1269
ARG-329	-41.2856 ± 0.0505	-0.1129 ± 0.0307	-0.0056 ± 0.0049	-41.4054 ± 0.0564
ARG-329	-36.0621 ± 0.0716	-0.0337 ± 0.0424	-0.0068 ± 0.0069	-36.0992 ± 0.0814
LYS-331	-42.1537 ± 0.0576	-0.0709 ± 0.0686	-0.0005 ± 0.0045	-42.2298 ± 0.0898
LYS-331	-36.2754 ± 0.0932	-0.0168 ± 0.0928	-0.0011 ± 0.0083	-36.2946 ± 0.1336
LYS-337	-77.6730 ± 0.1871	-2.1014 ± 0.0785	-0.0009 ± 0.0036	-79.7780 ± 0.1763
LYS-337	-56.2900 ± 0.2433	-0.2693 ± 0.0792	0.0054 ± 0.0057	-56.5604 ± 0.2570
ARG-342	-55.9431 ± 0.2115	-2.0447 ± 0.0837	-0.0006 ± 0.0056	-57.9882 ± 0.2207
ARG-342	-50.9481 ± 0.1810	-0.0353 ± 0.0398	-0.0020 ± 0.0066	-50.9823 ± 0.1756

ARG-366	-67.9929 ± 0.4287	3.1197 ± 0.2704	-0.0180 ± 0.0046	-64.8759 ± 0.2575
ARG-366	-51.2452 ± 0.1673	0.0737 ± 0.0430	0.0011 ± 0.0077	-51.1772 ± 0.1799
LYS-380	-54.7762 ± 0.0987	0.1096 ± 0.0626	0.0007 ± 0.0065	-54.6674 ± 0.1192
LYS-380	-41.7206 ± 0.1759	-0.0203 ± 0.0754	0.0004 ± 0.0081	-41.7469 ± 0.1908
LYS-401	-49.2065 ± 0.0908	0.2926 ± 0.0713	0.0053 ± 0.0062	-48.9015 ± 0.1097
LYS-401	-38.2859 ± 0.1419	0.1619 ± 0.0736	-0.0053 ± 0.0070	-38.1297 ± 0.1595
LYS-403	-54.8503 ± 0.1316	0.2102 ± 0.0724	-0.0053 ± 0.0076	-54.6517 ± 0.1465
LYS-403	-42.8225 ± 0.1768	-0.0197 ± 0.0822	0.0021 ± 0.0077	-42.8452 ± 0.1964
ARG-407	**-113.4999 ± 0.4190**	**19.9251 ± 0.4347**	**-1.0823 ± 0.0177**	**-94.6407 ± 0.2506**
ARG-407	**-67.4324 ± 0.4222**	**-0.2368 ± 0.0382**	**0.0017 ± 0.0067**	**-67.6695 ± 0.4340**
ARG-409	-122.6583 ± 0.7268	62.9510 ± 0.8774	-1.7440 ± 0.0132	-61.4415 ± 0.3362
ARG-409	-63.2655 ± 0.2986	-0.2820 ± 0.0362	0.0075 ± 0.0075	-63.5480 ± 0.3106
LYS-414	-59.9443 ± 0.2630	2.8078 ± 0.1203	-0.0027 ± 0.0050	-57.1419 ± 0.1981
LYS-414	-47.0357 ± 0.1393	-0.0173 ± 0.0746	0.0040 ± 0.0070	-47.0535 ± 0.1591
ARG-428	-73.0511 ± 0.2243	0.6136 ± 0.0450	-0.0047 ± 0.0065	-72.4497 ± 0.2214
ARG-428	-53.0131 ± 0.3205	-0.0775 ± 0.0510	0.0022 ± 0.0075	-53.1060 ± 0.3280
LYS-445	-51.3617 ± 0.0925	0.3947 ± 0.0733	0.0134 ± 0.0082	-50.9526 ± 0.1153
LYS-445	-41.9214 ± 0.1493	0.0355 ± 0.0836	-0.0004 ± 0.0079	-41.8864 ± 0.1770
LYS-451	-67.0293 ± 0.1795	0.5255 ± 0.0754	-0.0044 ± 0.0056	-66.5041 ± 0.2001
LYS-451	-55.1876 ± 0.3377	0.0184 ± 0.0963	0.0069 ± 0.0087	-55.1739 ± 0.3367
ARG-457	-129.9870 ± 0.5177	61.1992 ± 0.7698	-1.4292 ± 0.0165	-70.2436 ± 0.4185
ARG-457	-61.0519 ± 0.3268	-0.3668 ± 0.0366	-0.0184 ± 0.0075	-61.4519 ± 0.3314
LYS-460	-69.0115 ± 0.4348	4.5280 ± 0.3472	-0.0923 ± 0.0085	-64.5615 ± 0.2673
LYS-460	-49.9862 ± 0.1987	-0.0347 ± 0.0758	-0.0229 ± 0.0080	-50.0268 ± 0.2059
ARG-462	-52.3749 ± 0.1254	1.1966 ± 0.0417	-0.0038 ± 0.0049	-51.1763 ± 0.1114
ARG-462	-41.6525 ± 0.1116	-0.0438 ± 0.0436	-0.0018 ± 0.0066	-41.6945 ± 0.1153
LYS-497	-63.2035 ± 0.1370	0.7054 ± 0.0656	-0.0013 ± 0.0060	-62.4914 ± 0.1454
LYS-497	-51.4362 ± 0.1987	-0.1081 ± 0.0929	0.0038 ± 0.0062	-51.5301 ± 0.2162
ARG-503	-94.2044 ± 0.5125	16.5274 ± 0.5668	-0.3921 ± 0.0224	-78.0726 ± 0.2272
ARG-503	-71.5822 ± 0.6615	-2.4374 ± 0.1731	-0.0319 ± 0.0094	-74.0738 ± 0.6997
LYS-508	-54.6900 ± 0.1700	0.2375 ± 0.0816	-0.0005 ± 0.0052	-54.4477 ± 0.1940

LYS-508	-46.6851 ± 0.1814	-0.3467 ± 0.0867	0.0016 ± 0.0106	-47.0219 ± 0.1898
ARG-517	-52.5233 ± 0.1122	0.1501 ± 0.0408	0.0011 ± 0.0055	-52.3720 ± 0.1161
ARG-517	-46.2090 ± 0.1128	-0.1693 ± 0.0457	-0.0002 ± 0.0068	-46.3828 ± 0.1206
ARG-520	-68.2821 ± 0.2724	0.5538 ± 0.0348	-0.0054 ± 0.0073	-67.7416 ± 0.2696
ARG-520	-56.8929 ± 0.2019	-0.3786 ± 0.0423	0.0028 ± 0.0096	-57.2594 ± 0.2025
LYS-523	**-104.7475 ± 0.5165**	**0.9850 ± 0.0893**	**0.0052 ± 0.0068**	**-103.7482 ± 0.5120**
LYS-523	**-77.8643 ± 0.4893**	**-0.3788 ± 0.0927**	**-0.0019 ± 0.0104**	**-78.2325 ± 0.5036**
LYS-541	-60.6786 ± 0.1665	0.1065 ± 0.0711	0.0050 ± 0.0067	-60.5599 ± 0.1830
LYS-541	-52.9379 ± 0.1410	-0.1438 ± 0.0993	0.0153 ± 0.0065	-53.0567 ± 0.1587
LYS-571	**-138.4027 ± 0.7386**	**51.9142 ± 1.1539**	**-2.2878 ± 0.0269**	**-88.7892 ± 0.6939**
LYS-571	**-157.1811 ± 1.5041**	**69.5700 ± 2.6548**	**-1.4160 ± 0.0385**	**-89.0764 ± 1.3326**
LYS-585	-59.3027 ± 0.1288	-0.1287 ± 0.0634	0.0186 ± 0.0072	-59.4079 ± 0.1416
LYS-585	-62.2674 ± 0.1469	-0.1997 ± 0.0815	-0.0136 ± 0.0084	-62.4831 ± 0.1502
LYS-593	-87.2303 ± 0.9032	11.9046 ± 0.7692	-0.4393 ± 0.0189	-75.7603 ± 0.3840
LYS-593	-85.4179 ± 0.5481	4.3375 ± 0.2954	-0.1442 ± 0.0148	-81.2111 ± 0.4151
ARG-614	-62.6623 ± 0.2327	-0.0926 ± 0.0314	-0.0067 ± 0.0049	-62.7515 ± 0.2281
ARG-614	-74.8271 ± 0.4158	0.0978 ± 0.0603	-0.0031 ± 0.0070	-74.7367 ± 0.4145
LYS-615	-51.0710 ± 0.0956	-0.1255 ± 0.0772	0.0041 ± 0.0058	-51.1912 ± 0.1248
LYS-615	-53.3135 ± 0.1488	-0.0733 ± 0.0941	-0.0118 ± 0.0056	-53.4034 ± 0.1780

Table 3.1.4. Binding free energy contribution of the key binding-site residues calculated from the binding energy decomposition for IDA ($kJmol^{-1}$) from HAESA-IDA interaction. Marked residues are from three protein complex.

Residues	MM Energy	Polar Energy	Apolar Energy	Total Energy
TYR-56	**-156.0128 ± 0.8144**	**32.8366 ± 1.1105**	**-1.1374 ± 0.0321**	**-124.2754 ± 0.6482**
TYR-56	**-229.6908 ± 1.2987**	**87.4998 ± 1.2997**	**-2.1683 ± 0.0353**	**-144.3737 ± 0.6748**
ILE-59	-8.6530 ± 0.2258	3.5453 ± 0.1530	-0.4761 ± 0.0168	-5.5880 ± 0.1520
ILE-59	-16.9705 ± 0.1307	3.0770 ± 0.0665	-1.8666 ± 0.0204	-15.7639 ± 0.1151
LYS-66	**-151.0040 ± 0.4502**	**12.0672 ± 0.2357**	**-0.0786 ± 0.0093**	**-139.0063 ± 0.3722**
LYS-66	**-149.9025 ± 0.8650**	**13.2792 ± 0.7895**	**-0.1645 ± 0.0163**	**-136.7618 ± 0.7747**
ARG-67	**-159.6867 ± 0.7076**	**31.0118 ± 0.9721**	**-1.7052 ± 0.0331**	**-130.3741 ± 0.4156**
ARG-67	**-164.5130 ± 1.0677**	**32.6473 ± 0.8672**	**-1.3839 ± 0.0312**	**-133.2729 ± 0.4412**

Table 3.1.5. Binding free energy contribution of the key binding-site residues calculated from the binding energy decomposition for IDA ($kJmol^{-1}$) from IDA-SERk1 interaction. Marked residues are from three protein complex.

Residues	MM Energy	Polar Energy	Apolar Energy	Total Energy
TYR-56	-74.9683 ± 0.2387	-1.5498 ± 0.1327	0.0146 ± 0.0067	-76.5029 ± 0.2755
TYR-56	-56.1934 ± 0.5788	2.2192 ± 0.1790	0.0051 ± 0.0073	-53.9602 ± 0.6237
LYS-66	**-95.5235 ± 0.4025**	**4.2566 ± 0.2393**	**-0.4049 ± 0.0099**	**-91.6671 ± 0.3075**
LYS-66	**-105.2887 ± 2.6585**	**30.7005 ± 1.6892**	**-0.5351 ± 0.0252**	**-75.1556 ± 1.4666**
ARG-67	-159.8879 ± 1.1256	69.7261 ± 1.4027	-2.3518 ± 0.0298	-92.5308 ± 0.4236
ARG-67	-119.3229 ± 3.0573	56.1458 ± 2.4535	-1.9194 ± 0.0724	-65.1654 ± 1.1806

Table 3.1.6. Binding free energy contribution of the key binding-site residues calculated from the binding energy decomposition for SERk1 ($kJmol^{-1}$) from HAESA-SERk1 interaction. Marked residues are from three protein complex.

Residues	MM Energy	Polar Energy	Apolar Energy	Total Energy
ALA-26	-151.0284 ± 0.4921	74.1103 ± 0.7335	-1.0410 ± 0.0103	-77.9653 ± 0.5511
ALA-26	-81.1722 ± 0.2440	-0.4613 ± 0.0761	-0.0093 ± 0.0079	-81.6388 ± 0.2447
ARG-37	-92.5341 ± 0.3105	-0.3315 ± 0.0599	-0.0019 ± 0.0101	-92.8791 ± 0.3077
ARG-37	-81.1542 ± 0.4048	-0.4240 ± 0.0377	-0.0227 ± 0.0099	-81.5856 ± 0.3985
ARG-73	-150.6381 ± 0.8656	85.9147 ± 0.8472	-1.5858 ± 0.0169	-66.2532 ± 0.4482
ARG-73	-87.3424 ± 1.0307	3.3779 ± 0.8124	-0.4109 ± 0.0243	-84.3558 ± 0.5223
LYS-93	-83.6902 ± 0.3236	1.7314 ± 0.1338	-0.0051 ± 0.0092	-81.9557 ± 0.3201
LYS-93	-71.5407 ± 0.1866	0.2480 ± 0.1014	0.0010 ± 0.0133	-71.2927 ± 0.2121
LYS-139	-69.3564 ± 0.2062	1.1857 ± 0.0987	0.0069 ± 0.0084	-68.1721 ± 0.2255
LYS-139	-64.0857 ± 0.1755	0.0680 ± 0.1210	-0.0116 ± 0.0119	-64.0300 ± 0.2059
LYS-142	**-91.6536 ± 0.3608**	**2.2227 ± 0.1035**	**0.0174 ± 0.0100**	**-89.4065 ± 0.3684**
LYS-142	**-99.9321 ± 0.4770**	**0.1071 ± 0.1241**	**-0.0123 ± 0.0108**	**-99.8596 ± 0.4426**
ARG-144	**-159.6342 ± 0.9370**	**59.6461 ± 1.1085**	**-0.5233 ± 0.0190**	**-100.4988 ± 0.4003**
ARG-144	**-154.3252 ± 0.5627**	**52.9457 ± 0.7387**	**-1.2386 ± 0.0230**	**-102.6319 ± 0.4762**
ARG-147	-131.8008 ± 0.5643	68.3170 ± 0.6651	-1.5497 ± 0.0160	-65.0661 ± 0.4658
ARG-147	-73.1901 ± 1.1568	10.1373 ± 1.1504	-0.9028 ± 0.0288	-63.8895 ± 0.3714
ARG-176	-55.2927 ± 0.1427	-0.2523 ± 0.0546	0.0005 ± 0.0110	-55.5416 ± 0.1524
ARG-176	-51.9715 ± 0.1567	-0.1959 ± 0.0684	0.0063 ± 0.0127	-52.1553 ± 0.1694

Table 3.1.7. Binding free energy contribution of the key binding-site residues calculated from the binding energy decomposition for SERk1 ($kJmol^{-1}$) from IDA-SERk1 interaction. Marked residues are from three protein complex.

Residues	MM Energy	Polar Energy	Apolar Energy	Total Energy
GLU-29	-22.5547 ± 0.1430	-4.4439 ± 0.1477	0.0028 ± 0.0103	-27.0031 ± 0.0901
GLU-29	-28.1672 ± 0.6408	4.5924 ± 0.2057	-0.0062 ± 0.0076	-23.5762 ± 0.4671
ASP-31	-64.4057 ± 0.7235	18.7544 ± 1.0199	-0.1795 ± 0.0108	-45.7926 ± 0.3792
ASP-31	-68.8336 ± 2.0106	40.1661 ± 1.6610	-0.3523 ± 0.0153	-29.0351 ± 0.8138
ASP-42	-31.0453 ± 0.1140	-0.1855 ± 0.0562	0.0046 ± 0.0041	-31.2251 ± 0.1258
ASP-42	-24.3308 ± 0.4931	1.3188 ± 0.0841	-0.0112 ± 0.0058	-23.0333 ± 0.4548
ASP-51	-55.9795 ± 0.3939	4.0347 ± 0.2979	-0.0020 ± 0.0043	-51.9412 ± 0.2125
ASP-51	-40.6566 ± 1.0793	10.1105 ± 0.5080	-0.0184 ± 0.0059	-30.5340 ± 0.7157
GLU-68	-10.5841 ± 0.0606	-1.6192 ± 0.1165	-0.0082 ± 0.0062	-12.2069 ± 0.1176
GLU-68	-18.4338 ± 0.3321	0.3713 ± 0.0978	-0.0023 ± 0.0041	-18.0681 ± 0.3374
ASP-75	-16.2847 ± 0.0545	-1.6429 ± 0.0692	-0.0005 ± 0.0067	-17.9328 ± 0.0754
ASP-75	-19.1131 ± 0.2623	0.8509 ± 0.0735	-0.0073 ± 0.0068	-18.2806 ± 0.2541
GLU-80	-20.6783 ± 0.0585	-0.1806 ± 0.1076	-0.0133 ± 0.0074	-20.8761 ± 0.1244
GLU-80	-19.1115 ± 0.3001	0.4586 ± 0.0803	-0.0126 ± 0.0069	-18.6648 ± 0.2953
GLU-88	-31.0709 ± 0.1222	0.0453 ± 0.0846	0.0002 ± 0.0093	-31.0252 ± 0.1297
GLU-88	-25.0544 ± 0.4876	0.5583 ± 0.0864	-0.0004 ± 0.0060	-24.4947 ± 0.4988
GLU-99	-15.6928 ± 0.0412	-0.6078 ± 0.0851	0.0005 ± 0.0064	-16.2985 ± 0.0896
GLU-99	-18.2847 ± 0.2210	0.5744 ± 0.0770	-0.0045 ± 0.0057	-17.7112 ± 0.2129
ASP-115	-19.2698 ± 0.0600	-0.2037 ± 0.0973	-0.0019 ± 0.0074	-19.4759 ± 0.1109
ASP-115	-20.3251 ± 0.3439	0.4314 ± 0.0799	0.0021 ± 0.0054	-19.9085 ± 0.3438
ASP-123	-15.5287 ± 0.0369	-0.2390 ± 0.0613	0.0039 ± 0.0055	-15.7641 ± 0.0714
ASP-123	-17.7293 ± 0.1799	0.6657 ± 0.0544	0.0000 ± 0.0046	-17.0718 ± 0.1751
GLU-135	-14.5635 ± 0.0370	-0.0050 ± 0.0963	-0.0148 ± 0.0079	-14.5834 ± 0.1001
GLU-135	-15.5755 ± 0.1257	0.0636 ± 0.0505	-0.0012 ± 0.0044	-15.5067 ± 0.1359
ASP-171	-13.9175 ± 0.0269	0.0809 ± 0.0636	-0.0089 ± 0.0064	-13.8477 ± 0.0685
ASP-171	-15.4665 ± 0.1073	0.1636 ± 0.0515	-0.0016 ± 0.0047	-15.3066 ± 0.1183
ASP-183	-10.6774 ± 0.0197	0.3192 ± 0.1075	0.0031 ± 0.0065	-10.3542 ± 0.1095
ASP-183	-12.7810 ± 0.0932	0.1493 ± 0.0850	0.0065 ± 0.0038	-12.6282 ± 0.1149
ASP-200	-11.9128 ± 0.0222	0.1836 ± 0.0921	0.0005 ± 0.0085	-11.7263 ± 0.0946
ASP-200	-13.2592 ± 0.0623	0.0445 ± 0.0767	-0.0025 ± 0.0063	-13.2195 ± 0.0941
PRO-211	-10.5971 ± 0.0220	0.3086 ± 0.1073	0.0013 ± 0.0069	-10.2934 ± 0.1102
PRO-211	-12.0858 ± 0.0643	0.0884 ± 0.0886	-0.0099 ± 0.0047	-12.0067 ± 0.1066

3.2 Root Mean Square Fluctuation (RMSF)

The RMSFs of the residues of HAESA, IDA and SERk1 in all the simulated systems were calculated from the MD trajectories. 30ns were used to calculate the RMSF values. The results revealed that most of the residues fluctuated by less than 0.40 nm for HAESA, 0.25nm for IDA, 0.30 nm for SERk1 (Fig 3.2.1). Though some residues exceed 0.50 nm for HAESA, 0.30 nm for IDA and SERk1. From MM/PBSA calculation, some residues show the most favourable condition for interacting between proteins. For HAESA, Asp120, Glu123, Glu145, Glu191, Asp242, Glu263, Glu266, Glu310, Glu316, Glu335, Glu382 give more favourable condition in both cases (when SERk1 is present in the complex and absent in the complex). From RMSF graphical view, it is obtained that these residues also give less fluctuation in contrast to other residues of HAESA. When it is an interaction between HAESA and IDA, from HAESA Glu123 is the most prominent residue, in RMSF, it shows the lowest fluctuation (0.24 nm) when SERk1 is present inside the complex and in the absence of SERk1 it gives the second-lowest fluctuation (0.22 nm). The most fluctuated residues of HAESA is Arg50 (0.76 nm) and Lys593 (0.42 nm) at the presence of SERk1 and when there is no SERk1 in the complex these two residues give similarly maximum fluctuation on that condition. In the case of IDA, Lys66 shows the lowest fluctuation in RMSF, as it is shown another prominent residue in MM/PBSA calculation. During interaction with HAESA Lys66 from IDA remains at 0.13 nm when there is SERk1 is present in the complex and it rises to 0.27 nm when SERk1 is removed from the complex. Another residue of IDA is Arg67, which is also low fluctuating residue in the complex where SERk1 exists, but after removing SERk1 from the complex Arg67 exceeds highly and gives comparatively higher fluctuation.

On the other hand, in case of HAESA and SERk1 interaction Lys337, Arg407, Arg503, Lys523, Lys571 of HAESA show the lowest fluctuation rate in both conditions where IDA is present in the complex and also removed from the complex. From MM/PBSA calculation these residues of HAESA also contribute more energy for interacting with SERk1. Lys571 of HAESA, which gives the consistent RMSF value (0.27nm) when IDA is present in the complex and absent in the complex which indicates that it remains the most prominent residue as it is shown in MM/PBSA calculation before. The most fluctuated residue of HAESA in this interaction is Lys523. For SERk1 Arg37, Lys142, Arg144 are called the most prominent residues by MM/PBSA calculation. In RMSF graphical view these residues also give less fluctuation value in contrast of other residues

of SERk1. Between these residues, Arg144 gives the lowest fluctuation value (0.17 nm) when IDA is present but in the absence of IDA, it slightly fluctuated down to 0.14 nm.

Though there are some other lowest fluctuated residues are obtained from RMSF analysis and also some highly fluctuated residues are found but in terms of making the interaction between HAESA, IDA and SERk1 the mentioned residues are responsible mostly. Again, for happening interaction between HAESA and IDA, Co-Receptor SERk1 is not that much mandatory in the complex. Because after removing SERk1 the fluctuation rate of the residues of HAESA and IDA go lower which increases the possibilities of interaction between them.

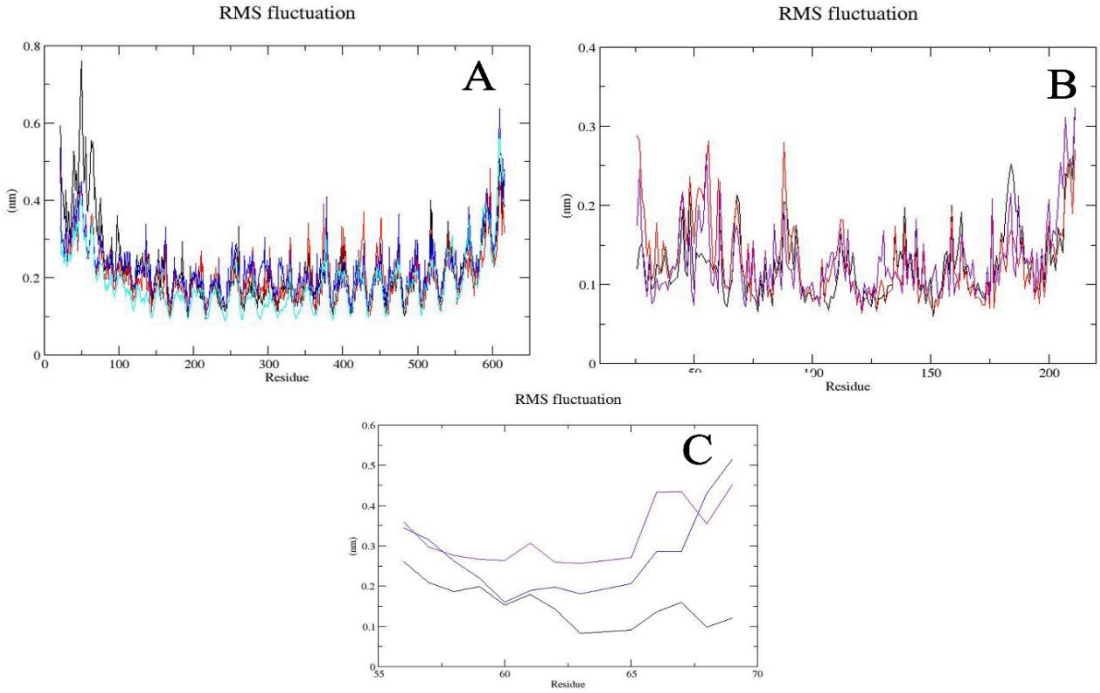

Fig 3.2.1: (A) RMSF value of HAESA from 30ns MD trajectories. HAESA, IDA complex (Black) presence of SERk1 in the complex, HAESA and IDA complex (Blue) absence of SERk1, HAESA and SERk1 complex (Red) absence of IDA, only HAESA (Sky Blue). (B) RMSF value of SERk1 from 30ns MD trajectories. HAESA, SERk1 complex (Black) in the presence of IDA, HAESA and SERk1 complex (Red) absence of IDA, IDA and SERk1 complex (Purple) absence HAESA. (C) RMSF value of IDA from 30ns MD trajectories. HAESA, IDA complex (Black) presence of SERk1, HAESA and IDA complex (Blue) absence of SERk1, IDA and SERk1 complex (Purple) absence of HAESA.

3.3 H-Bond

Hydrogen bond number was calculated over the 30 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 3.3.1(A), it is visible that maximum 10 hydrogen bond is found between HAESA and IDA at 850ps and 1850ps when SERk1 is present and in absence of SERk1 maximum 10 hydrogen bond is formed at 10850ps. Protein Interaction Calculation (PIC) (Table 3.3.1 – 3.3.21) data also said that there is very few sustainable hydrogen bond interactions exist between HAESA and IDA, rather sustainable hydrogen bond interactions are existing between HAESA and SERk1 with a very high number. In details, Arg73 of SERk1 donates atoms to Asn526 and Glu527 of HAESA and remain same after simulation where it is a side chain-side chain interaction again Asp123 and Arg147 of SERk1 also create side chain-side chain interaction with Asn573 of HAESA. IDA makes a very negligible amount of sustainable main chain-side chain hydrogen bond with HAESA when SERk1 is absent in the complex. Again, there are very few hydrogen bonds found in between IDA and SERk1 when HAESA is present in the complex but 8-9 hydrogen bonds are found in absence of HAESA protein after simulation.

From PIC data, the number of every possible interaction are summarized (Table 3.3.22). Where it is found that at the very initial stage or before simulation there are 36 hydrogen bonds, 5 hydrophobic interactions and 1 ionic interaction have happened between HAESA and IDA when SERk1 is also present with them. Though the number of hydrogen bonds reduces between HAESA and IDA in this situation remain the same after the 30ns simulation period some hydrophobic interactions are increased and the only ionic interaction is reduced. Moreover, when SERk1 is not present in the complex hydrogen bonds of HAESA and IDA are also reduced. On the other hand, the number of hydrogen bonds of HAESA and SERk1 is 35 before the simulation. When there is IDA is present hydrogen bonds are drastically increased between HAESA and SERk1. But in the absence of IDA, numbers of hydrogen bond of HAESA and SERk1 are reduced. Hydrophobic and ionic interaction reduction is normal in both cases. In another case, though before simulation there were 4 hydrogen bonds found between IDA and SERk1 after simulation there is only 1 hydrogen bond reduced. This is happened when HAESA is present in the complex, during the absence of

HAESA numbers of hydrogen bonds of IDA and SERk1 are increased. There are no negligible other interactions found in between IDA and SERk1.

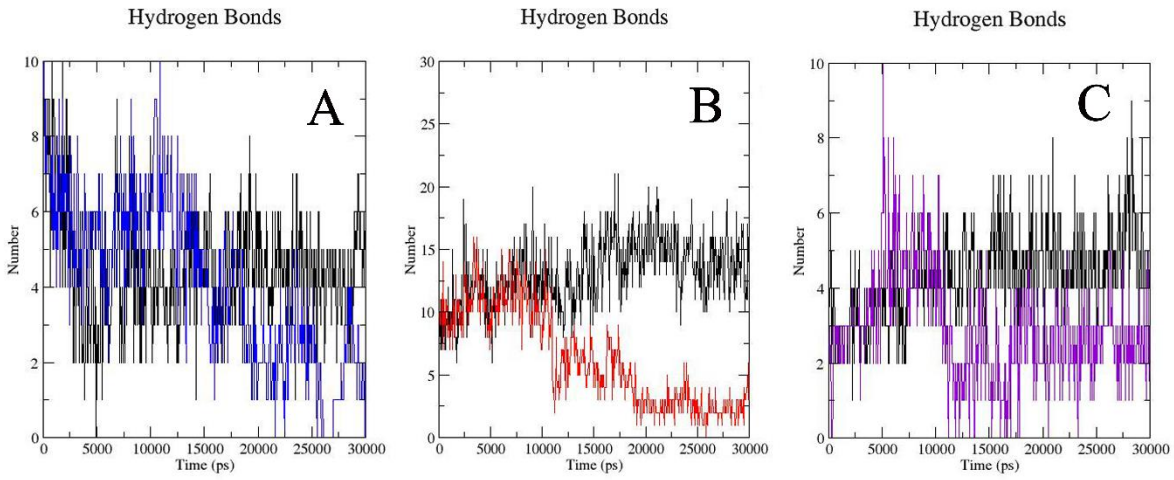

Fig 3.3.1: (A) H-bond value of HAESA and IDA from 30ns MD trajectories. HAESA, IDA complex (Black) presence of SERk1 in the complex, HAESA and IDA complex (Blue) absence of SERk1. (B) H-bond value of HAESA and SERk1 from 30ns MD trajectories. HAESA, SERk1 complex (Black) in the presence of IDA, HAESA and SERk1 complex (Red) absence of IDA. (C) H-bond value of IDA and SERk1 from 30ns MD trajectories. IDA and SERk1 complex (Black) presence of HAESA, IDA and SERk1 complex (Purple) absence of HAESA.

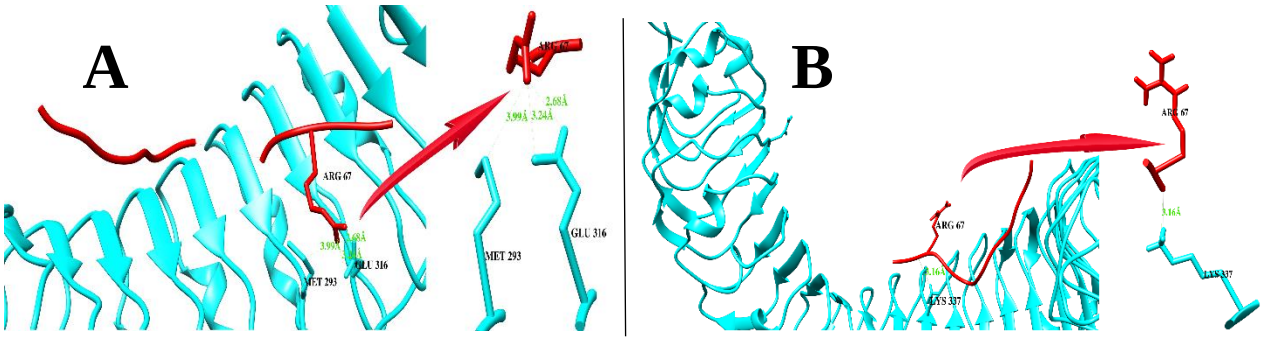

Fig 3.3.2: (A) H-bond of Arg67 from IDA before simulation in cartoon structure. (B) H-bond of Arg67 from IDA After simulation in cartoon structure.

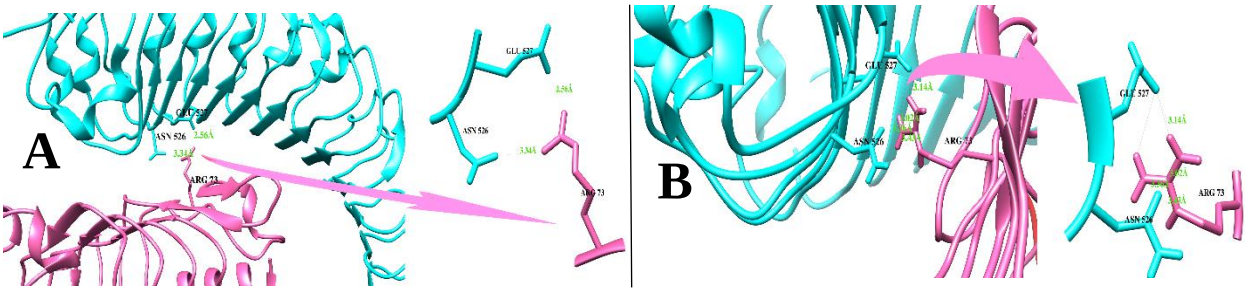

Fig 3.3.3: (A) H-bond of Arg73 from SERk1 before simulation in cartoon structure. (B) H-bond of Arg73 from SERk1 After simulation in cartoon structure.

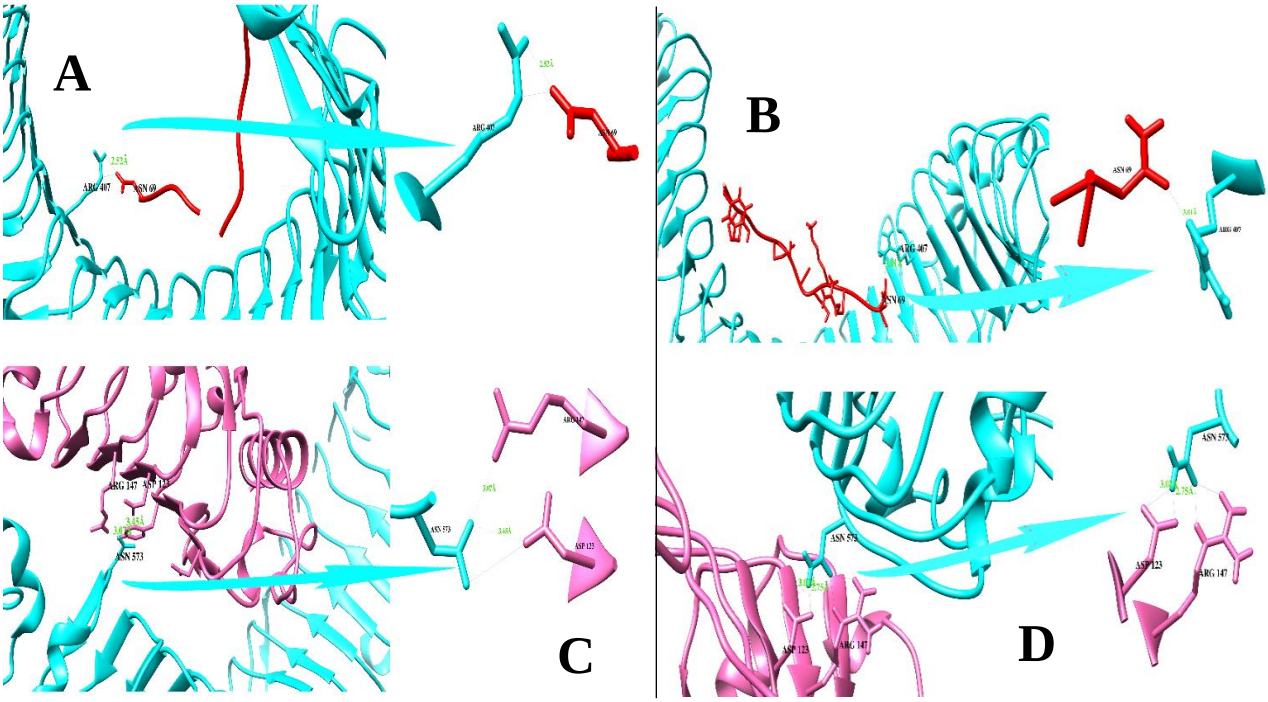

Fig 3.3.4: (A) H-bond of Arg407 from HAESA before simulation in cartoon structure. (B) H-bond of Arg407 from HAESA After simulation in cartoon structure. (C) H-bond of Asn573 from HAESA before simulation in cartoon structure. (D) H-bond of Asn573 from HAESA After simulation in cartoon structure.

Table 3.3.1: Protein-Protein Hydrophobic Interaction of HAESA-IDA-SERk1 complex before and after simulation.

Protein-Protein Hydrophobic Interactions					
Before Simulation					
Position	**Residue**	**Chain**	**Position**	**Residue**	**Chain**
97	TYR	A	56	TYR	B
171	ALA	A	59	ILE	B
196	TYR	A	59	ILE	B
196	TYR	A	60	PRO	B
218	TRP	A	59	ILE	B
339	PHE	A	55	VAL	C
478	ILE	A	61	PHE	C
548	VAL	A	101	TYR	C
551	TYR	A	97	TYR	C
574	VAL	A	97	TYR	C
594	ILE	A	169	VAL	C
594	ILE	A	189	LEU	C
595	TYR	A	145	PHE	C
595	TYR	A	169	VAL	C
After Simulation					
Position	**Residue**	**Chain**	**Position**	**Residue**	**Chain**
174	PHE	A	56	TYR	B
196	TYR	A	58	PRO	B
196	TYR	A	60	PRO	B
196	TYR	A	61	PRO	B
198	LEU	A	56	TYR	B
198	LEU	A	58	PRO	B
218	TRP	A	60	PRO	B
339	PHE	A	55	VAL	C
551	TYR	A	145	PHE	C
551	TYR	A	97	TYR	C
574	VAL	A	145	PHE	C
594	ILE	A	189	LEU	C
594	ILE	A	190	PHE	C
595	TYR	A	145	PHE	C
595	TYR	A	169	VAL	C

Table 3.3.2: Hydrogen Bond (main chain-main chain) of HAESA-IDA-SERk1 before and after simulation.

Protein-Protein Main Chain-Main Chain Hydrogen Bonds								
Before Simulation								
Donor				**Acceptor**				
POS	**CHAIN**	**RES**	**ATOM**	**POS**	**CHAIN**	**RES**	**ATOM**	**Dd-a**
68	B	HIS	N	53	C	THR	O	3.19
55	C	VAL	N	68	B	HIS	O	3.14
After Simulation								
Donor				**Acceptor**				
POS	**CHAIN**	**RES**	**ATOM**	**POS**	**CHAIN**	**RES**	**ATOM**	**Dd-a**
68	B	HIS	N	53	C	THR	O	3.33
55	C	VAL	N	68	B	HIS	O	3.26

Table 3.3.3: Hydrogen bond (main chain-side chain) of HAESA-IDA-SERk1 before and after simulation.

Protein-Protein Main Chain-Side Chain Hydrogen Bonds								
Before Simulation								
Donor				**Acceptor**				
POS	**CHAIN**	**RES**	**ATOM**	**POS**	**CHAIN**	**RES**	**ATOM**	**Dd-a**
123	A	GLU	OE1	57	B	VAL	O	3.45
123	A	GLU	OE1	57	B	VAL	O	3.45
123	A	GLU	OE2	57	B	VAL	O	2.65
123	A	GLU	OE2	57	B	VAL	O	2.65
264	A	GLN	NE2	62	B	SER	O	3.15
264	A	GLN	NE2	62	B	SER	O	3.15
266	A	GLU	OE1	63	B	ALA	O	3.34
266	A	GLU	OE1	63	B	ALA	O	3.34
313	A	ASN	ND2	65	B	SER	O	3.31
313	A	ASN	ND2	65	B	SER	O	3.31
337	A	LYS	NZ	65	B	SER	O	2.88
337	A	LYS	NZ	67	B	ARG	O	2.66
361	A	ASP	OD2	69	B	ASN	OXT	2.98
361	A	ASP	OD2	69	B	ASN	OXT	2.98
457	**A**	**ARG**	**NH2**	58	C	CYS	O	3.11
457	**A**	**ARG**	**NH2**	58	C	CYS	O	3.11
57	B	VAL	N	123	A	GLU	OE1	2.7
65	B	SER	N	290	A	ASP	OD2	2.97
67	B	ARG	NH2	55	C	VAL	O	3.39
67	B	ARG	NH2	55	C	VAL	O	3.39

69	B	ASN	N	361	A	ASP	OD2	2.48
101	C	TYR	OH	548	A	VAL	O	3.42
144	C	ARG	NH2	**594**	**A**	**ILE**	**O**	3.04
144	C	ARG	NH2	**594**	**A**	**ILE**	**O**	3.04
147	C	ARG	NH2	572	A	LEU	O	3.38
147	C	ARG	NH2	572	A	LEU	O	3.38
After Simulation								
Donor				**Acceptor**				
POS	**CHAIN**	**RES**	**ATOM**	**POS**	**CHAIN**	**RES**	**ATOM**	**Dd-a**
150	A	ASN	OD1	56	B	TYR	O	3.09
150	A	ASN	OD1	56	B	TYR	O	3.09
337	A	LYS	NZ	65	B	SER	O	2.91
337	A	LYS	NZ	67	B	ARG	O	3.16
409	A	ARG	NH1	68	C	GLU	O	3.43
409	A	ARG	NH1	68	C	GLU	O	3.43
409	A	ARG	NH2	68	C	GLU	O	2.82
409	A	ARG	NH2	68	C	GLU	O	2.82
457	A	ARG	NH2	65	C	CYS	O	3.09
457	**A**	**ARG**	**NH2**	65	C	CYS	O	3.09
457	**A**	**ARG**	**NH1**	66	C	ASN	O	2.82
457	**A**	**ARG**	**NH1**	66	C	ASN	O	2.82
457	**A**	**ARG**	**NH2**	66	C	ASN	O	3.12
457	**A**	**ARG**	**NH2**	66	C	ASN	O	3.12
592	A	ASN	ND2	189	C	LEU	O	3.28
592	A	ASN	ND2	189	C	LEU	O	3.28
56	B	TYR	N	150	A	ASN	OD1	3
69	B	ASN	N	361	A	ASP	OD2	3.06
27	C	ASN	N	316	A	GLU	OE1	2.95
27	C	ASN	N	316	A	GLU	OE2	3.26
147	C	ARG	NH2	571	A	LYS	O	2.97
147	C	ARG	NH2	571	A	LYS	O	2.97
147	C	ARG	NH2	572	A	LEU	O	3.34
147	C	ARG	NH2	572	A	LEU	O	3.34
168	C	GLN	NE2	**594**	**A**	**ILE**	**O**	2.68
168	C	GLN	NE2	**594**	**A**	**ILE**	**O**	2.68

Table 3.3.4: Hydrogen bond (side chain-side chain) of HAESA-IDA-SERk1 before and after simulation.

Protein-Protein Side Chain-Side Chain Hydrogen Bonds								
Before Simulation								
Donor				Acceptor				
POS	CHAIN	RES	ATOM	POS	CHAIN	RES	ATOM	Dd-a
240	A	ASN	ND2	62	B	SER	OG	2.98
240	A	ASN	ND2	62	B	SER	OG	2.98
264	A	GLN	NE2	62	B	SER	OG	3.17
264	A	GLN	NE2	62	B	SER	OG	3.17
407	A	ARG	NE	69	B	ASN	OD1	3.19
407	A	ARG	NE	69	B	ASN	ND2	3.14
407	**A**	**ARG**	**NH2**	**69**	**B**	**ASN**	**OD1**	2.52
407	**A**	**ARG**	**NH2**	**69**	**B**	**ASN**	**OD1**	2.52
409	A	ARG	NH1	69	B	ASN	OD1	2.58
409	A	ARG	NH1	69	B	ASN	OD1	2.58
526	A	ASN	ND2	75	C	ASP	OD2	3.49
526	A	ASN	ND2	75	C	ASP	OD2	3.49
550	A	ASN	OD1	101	C	TYR	OH	2.7
550	A	ASN	OD1	101	C	TYR	OH	2.7
550	A	ASN	ND2	101	C	TYR	OH	3.38
550	A	ASN	ND2	101	C	TYR	OH	3.38
551	A	TYR	OH	75	C	ASP	OD1	2.73
573	A	ASN	ND2	99	C	GLU	OE2	2.85
573	A	ASN	ND2	99	C	GLU	OE2	2.85
573	A	ASN	OD1	121	C	SER	OG	2.72
573	A	ASN	OD1	121	C	SER	OG	2.72
573	**A**	**ASN**	**OD1**	**123**	**C**	**ASP**	**OD2**	3.45
573	**A**	**ASN**	**OD1**	**123**	**C**	**ASP**	**OD2**	3.45
62	B	SER	OG	240	A	ASN	ND2	2.98
62	B	SER	OG	242	A	ASP	OD2	2.72
62	B	SER	OG	264	A	GLN	NE2	3.17
67	B	ARG	NH1	293	A	MET	SD	3.99
67	B	ARG	NH1	293	A	MET	SD	3.99
67	B	ARG	NH1	316	A	GLU	OE1	2.68
67	B	ARG	NH1	316	A	GLU	OE1	2.68
67	B	ARG	NH1	316	A	GLU	OE2	3.24
67	B	ARG	NH1	316	A	GLU	OE2	3.24
73	**C**	**ARG**	**NH1**	**526**	**A**	**ASN**	**ND2**	3.34
73	**C**	**ARG**	**NH1**	**526**	**A**	**ASN**	**ND2**	3.34
73	**C**	**ARG**	**NH2**	**527**	**A**	**GLU**	**OE1**	2.56

73	**C**	**ARG**	**NH2**	**527**	**A**	**GLU**	**OE1**	2.56
75	**C**	**ASP**	**OD2**	**526**	**A**	**ASN**	**ND2**	3.49
75	C	ASP	OD2	526	A	ASN	ND2	3.49
99	**C**	**GLU**	**OE2**	**573**	A	ASN	ND2	2.85
99	**C**	**GLU**	**OE2**	**573**	A	ASN	ND2	2.85
101	C	TYR	OH	550	A	ASN	OD1	2.7
101	C	TYR	OH	550	A	ASN	ND2	3.38
121	C	SER	OG	573	A	ASN	OD1	2.72
123	**C**	**ASP**	**OD2**	**573**	**A**	**ASN**	**OD1**	3.45
123	**C**	**ASP**	**OD2**	**573**	**A**	**ASN**	**OD1**	3.45
147	**C**	**ARG**	**NH2**	**573**	**A**	**ASN**	**OD1**	3.07
147	**C**	**ARG**	**NH2**	**573**	**A**	**ASN**	**OD1**	3.07
After Simulation								
Donor				**Acceptor**				
POS	**CHAIN**	**RES**	**ATOM**	**POS**	**CHAIN**	**RES**	**ATOM**	**Dd-a**
316	A	GLU	OE1	27	C	ASN	ND2	2.93
316	A	GLU	OE1	27	C	ASN	ND2	2.93
361	A	ASP	OD2	69	B	ASN	ND2	2.84
361	A	ASP	OD2	69	B	ASN	ND2	2.84
388	A	ASP	OD1	69	C	ASN	ND2	3.24
388	A	ASP	OD1	69	C	ASN	ND2	3.24
388	A	ASP	OD2	69	C	ASN	ND2	2.91
388	A	ASP	OD2	69	C	ASN	ND2	2.91
407	A	ARG	NH1	68	B	HIS	NE2	3.15
407	A	ARG	NH1	68	B	HIS	NE2	3.15
407	**A**	**ARG**	**NH1**	**69**	**B**	**ASN**	**OD1**	3.01
407	**A**	**ARG**	**NH1**	**69**	**B**	**ASN**	**OD1**	3.01
409	A	ARG	NH1	56	C	ASN	ND2	3.48
409	A	ARG	NH1	56	C	ASN	ND2	3.48
409	A	ARG	NH1	58	C	CYS	SG	3.89
409	A	ARG	NH1	58	C	CYS	SG	3.89
409	A	ARG	NH1	69	C	ASN	OD1	3.3
409	A	ARG	NH1	69	C	ASN	OD1	3.3
526	A	ASN	ND2	99	C	GLU	OE1	3.07
526	A	ASN	ND2	99	C	GLU	OE1	3.07
526	A	ASN	ND2	99	C	GLU	OE2	2.9
526	A	ASN	ND2	99	C	GLU	OE2	2.9
550	A	ASN	ND2	99	C	GLU	OE1	2.89
550	A	ASN	ND2	99	C	GLU	OE1	2.89
550	A	ASN	ND2	123	C	ASP	OD2	3.4
550	A	ASN	ND2	123	C	ASP	OD2	3.4

573	A	ASN	ND2	123	C	ASP	OD1	3.02
573	A	ASN	ND2	123	C	ASP	OD1	3.02
573	A	ASN	ND2	123	C	ASP	OD2	2.75
573	A	ASN	ND2	123	C	ASP	OD2	2.75
67	B	ARG	NE	27	C	ASN	OD1	2.86
69	B	ASN	ND2	361	A	ASP	OD2	2.84
69	B	ASN	ND2	361	A	ASP	OD2	2.84
27	C	ASN	ND2	293	A	MET	SD	3.12
27	C	ASN	ND2	293	A	MET	SD	3.12
27	C	ASN	ND2	316	A	GLU	OE1	2.93
27	C	ASN	ND2	316	A	GLU	OE1	2.93
69	C	ASN	ND2	388	A	ASP	OD1	3.24
69	C	ASN	ND2	388	A	ASP	OD1	3.24
69	C	ASN	ND2	388	A	ASP	OD2	2.91
69	C	ASN	ND2	388	A	ASP	OD2	2.91
73	C	ARG	NE	526	A	ASN	OD1	3.43
73	C	ARG	NH1	526	A	ASN	OD1	3.02
73	C	ARG	NH1	526	A	ASN	OD1	3.02
73	C	ARG	NH2	526	A	ASN	OD1	3.36
73	C	ARG	NH2	526	A	ASN	OD1	3.36
73	C	ARG	NH1	527	A	GLU	OE2	3.14
73	C	ARG	NH1	527	A	GLU	OE2	3.14
99	C	GLU	OE1	526	A	ASN	ND2	3.07
99	C	GLU	OE1	526	A	ASN	ND2	3.07
99	C	GLU	OE2	526	A	ASN	ND2	2.9
99	C	GLU	OE2	526	A	ASN	ND2	2.9
99	C	GLU	OE1	550	A	ASN	ND2	2.89
99	C	GLU	OE1	550	A	ASN	ND2	2.89
123	C	ASP	OD2	550	A	ASN	ND2	3.4
123	C	ASP	OD2	550	A	ASN	ND2	3.4
123	C	ASP	OD1	573	A	ASN	ND2	3.02
123	C	ASP	OD1	573	A	ASN	ND2	3.02
123	C	ASP	OD2	573	A	ASN	ND2	2.75
123	C	ASP	OD2	573	A	ASN	ND2	2.75
147	C	ARG	NE	573	A	ASN	OD1	2.93
147	C	ARG	NH2	573	A	ASN	OD1	3.01
147	C	ARG	NH2	573	A	ASN	OD1	3.01

Table 3.3.5: Protein-Protein ionic interaction of HAESA-IDA-SERk1 before and after simulation.

Protein-Protein Ionic Interactions					
Before Simulation					
Position	**Residue**	**Chain**	**Position**	**Residue**	**Chain**
68	HIS	B	51	ASP	C
316	GLU	A	67	ARG	B
503	ARG	A	75	ASP	C
527	GLU	A	73	ARG	C
571	LYS	A	171	ASP	C
598	ASP	A	144	ARG	C
After Simulation					
Position	**Residue**	**Chain**	**Position**	**Residue**	**Chain**
527	GLU	A	73	ARG	C
598	ASP	A	144	ARG	C

Table 3.3.6: Protein-Protein Aromatic interaction of HAESA-IDA-SERk1 complex before and after simulation.

Protein-Protein Aromatic-Aromatic Interactions							
Before Simulation							
Position	**Residue**	**Chain**	**Position**	**Residue**	**Chain**	**D(centroid-centroid)**	**Dihedral Angle**
551	TYR	A	97	TYR	C	5.48	100.22
After Simulation							
Position	**Residue**	**Chain**	**Position**	**Residue**	**Chain**	**D(centroid-centroid)**	**Dihedral Angle**
174	PHE	A	56	TYR	B	6.1	98.16

Table 3.3.7: Protein-Protein Cation-Pi interaction of HAESA-IDA-SERk1 complex before and after simulation.

Protein-Protein Cation-Pi Interactions							
Before Simulation							
Position	**Residue**	**Chain**	**Position**	**Residue**	**Chain**	**D(cation-Pi)**	**Angle**
551	TYR	A	73	ARG	C	4.2	19.98
After Simulation							
Position	**Residue**	**Chain**	**Position**	**Residue**	**Chain**	**D(cation-Pi)**	**Angle**
551	TYR	A	73	ARG	C	5.99	66.31

Table 3.3.8: Protein-Protein Hydrophobic Interaction of HAESA-SERk1 complex (IDA absent) before and after simulation.

Protein-Protein Hydrophobic Interactions					
Before Simulation					
Position	**Residue**	**Chain**	**Position**	**Residue**	**Chain**
339	PHE	A	55	VAL	C
478	ILE	A	61	PHE	C
548	VAL	A	101	TYR	C
551	TYR	A	97	TYR	C
574	VAL	A	97	TYR	C
594	ILE	A	169	VAL	C
594	ILE	A	189	LEU	C
595	TYR	A	145	PHE	C
595	TYR	A	169	VAL	C
After Simulation					
Position	**Residue**	**Chain**	**Position**	**Residue**	**Chain**
589	LEU	A	189	LEU	C
589	LEU	A	190	PHE	C
590	TYR	A	145	PHE	C
590	TYR	A	169	VAL	C
594	ILE	A	189	LEU	C

Table 3.3.9: Hydrogen bond (main chain-side chain) of HAESA -SERk1 (IDA absent) before and after simulation.

Protein-Protein Main Chain-Side Chain Hydrogen Bonds								
Before Simulation								
Donor				**Acceptor**				
POS	**CHAIN**	**RES**	**ATOM**	**POS**	**CHAIN**	**RES**	**ATOM**	**Dd-a**
457	A	ARG	NH2	58	C	CYS	O	3.11
457	A	ARG	NH2	58	C	CYS	O	3.11
101	C	TYR	OH	548	A	VAL	O	3.42
144	C	ARG	NH2	594	A	ILE	O	3.04
144	C	ARG	NH2	594	A	ILE	O	3.04
147	C	ARG	NH2	572	A	LEU	O	3.38
147	C	ARG	NH2	572	A	LEU	O	3.38
After Simulation								
Donor				**Acceptor**				
POS	**CHAIN**	**RES**	**ATOM**	**POS**	**CHAIN**	**RES**	**ATOM**	**Dd-a**
592	A	ASN	ND2	167	C	LEU	O	2.92

592	A	ASN	ND2	167	C	LEU	O	2.92
595	A	TYR	OH	166	C	THR	O	2.69
144	C	ARG	NE	573	A	ASN	O	3.19
168	C	GLN	NE2	590	A	TYR	O	2.78
168	C	GLN	NE2	590	A	TYR	O	2.78

Table 3.3.10: Hydrogen bond (side chain-side chain) of HAESA -SERk1 (IDA absent) before and after simulation.

Protein-Protein Side Chain-Side Chain Hydrogen Bonds								
Before Simulation								
Donor				**Acceptor**				
POS	**CHAIN**	**RES**	**ATOM**	**POS**	**CHAIN**	**RES**	**ATOM**	**Dd-a**
526	A	ASN	ND2	75	C	ASP	OD2	3.49
526	A	ASN	ND2	75	C	ASP	OD2	3.49
550	A	ASN	OD1	101	C	TYR	OH	2.7
550	A	ASN	OD1	101	C	TYR	OH	2.7
550	A	ASN	ND2	101	C	TYR	OH	3.38
550	A	ASN	ND2	101	C	TYR	OH	3.38
551	A	TYR	OH	75	C	ASP	OD1	2.73
573	A	ASN	ND2	99	C	GLU	OE2	2.85
573	A	ASN	ND2	99	C	GLU	OE2	2.85
573	A	ASN	OD1	121	C	SER	OG	2.72
573	A	ASN	OD1	121	C	SER	OG	2.72
573	A	ASN	OD1	123	C	ASP	OD2	3.45
573	A	ASN	OD1	123	C	ASP	OD2	3.45
73	C	ARG	NH1	526	A	ASN	ND2	3.34
73	C	ARG	NH1	526	A	ASN	ND2	3.34
73	C	ARG	NH2	527	A	GLU	OE1	2.56
73	C	ARG	NH2	527	A	GLU	OE1	2.56
75	C	ASP	OD2	526	A	ASN	ND2	3.49
75	C	ASP	OD2	526	A	ASN	ND2	3.49
99	C	GLU	OE2	573	A	ASN	ND2	2.85
99	C	GLU	OE2	573	A	ASN	ND2	2.85
101	C	TYR	OH	550	A	ASN	OD1	2.7
101	C	TYR	OH	550	A	ASN	ND2	3.38
121	C	SER	OG	573	A	ASN	OD1	2.72
123	C	ASP	OD2	573	A	ASN	OD1	3.45
123	C	ASP	OD2	573	A	ASN	OD1	3.45
147	C	ARG	NH2	573	A	ASN	OD1	3.07
147	C	ARG	NH2	573	A	ASN	OD1	3.07

After Simulation								
Donor				Acceptor				
POS	CHAIN	RES	ATOM	POS	CHAIN	RES	ATOM	Dd-a
573	A	ASN	OD1	96	C	GLN	OE1	3.5
573	A	ASN	OD1	96	C	GLN	OE1	3.5
573	A	ASN	ND2	96	C	GLN	OE1	2.94
573	A	ASN	ND2	96	C	GLN	OE1	2.94
96	C	GLN	OE1	573	A	ASN	OD1	3.5
96	C	GLN	OE1	573	A	ASN	OD1	3.5
96	C	GLN	OE1	573	A	ASN	ND2	2.94
96	C	GLN	OE1	573	A	ASN	ND2	2.94

Table 3.3.11: Protein-Protein ionic interaction of HAESA- SERk1 (IDA absent) before and after simulation.

Protein-Protein Ionic Interactions					
Before Simulation					
Position	Residue	Chain	Position	Residue	Chain
503	ARG	A	75	ASP	C
527	GLU	A	73	ARG	C
571	LYS	A	171	ASP	C
598	ASP	A	144	ARG	C
After Simulation					
Position	Residue	Chain	Position	Residue	Chain
571	LYS	A	99	GLU	C

Table 3.3.12: Protein-Protein aromatic interaction of HAESA- SERk1 (IDA absent) before and after simulation.

Protein-Protein Aromatic-Aromatic Interactions							
Before Simulation							
Position	Residue	Chain	Position	Residue	Chain	D(centroid-centroid)	Dihedral Angle
551	TYR	A	97	TYR	C	5.48	100.22
After Simulation							
Position	Residue	Chain	Position	Residue	Chain	D(centroid-centroid)	Dihedral Angle
590	TYR	A	145	PHE	C	6.29	112.3

Table 3.3.13: Protein-Protein cation-pi interaction of HAESA- SERk1 (IDA absent) before and after simulation.

Protein-Protein Cation-Pi Interactions							
Before Simulation							
Position	**Residue**	**Chain**	**Position**	**Residue**	**Chain**	**D(cation-Pi)**	**Angle**
97	TYR	C	571	LYS	A	5.14	136.74
After Simulation							
Position	**Residue**	**Chain**	**Position**	**Residue**	**Chain**	**D(cation-Pi)**	**Angle**
97	TYR	C	571	LYS	A	5.14	136.74
145	PHE	C	571	LYS	A	3.09	179.73

Table 3.3.14: Protein-Protein Hydrophobic Interaction of HAESA-IDA complex (SERk1 absent) before and after simulation.

Protein-Protein Hydrophobic Interactions					
Before Simulation					
Position	**Residue**	**Chain**	**Position**	**Residue**	**Chain**
97	**TYR**	**A**	**56**	**TYR**	**B**
171	**ALA**	**A**	**59**	**ILE**	**B**
196	TYR	A	59	ILE	B
196	**TYR**	**A**	**60**	**PRO**	B
218	**TRP**	**A**	**59**	**ILE**	B
After Simulation					
Position	**Residue**	**Chain**	**Position**	**Residue**	**Chain**
97	**TYR**	**A**	**56**	**TYR**	**B**
171	**ALA**	**A**	**59**	**ILE**	**B**
196	**TYR**	**A**	**60**	**PRO**	**B**
218	**TRP**	**A**	**59**	**ILE**	B
218	TRP	A	60	PRO	B

Table 3.3.15: Hydrogen bond (main chain-side chain) of HAESA -IDA (SERk1 absent) before and after simulation.

Protein-Protein Main Chain-Side Chain Hydrogen Bonds								
Before Simulation								
Donor				Acceptor				
POS	**CHAIN**	**RES**	**ATOM**	**POS**	**CHAIN**	**RES**	**ATOM**	**Dd-a**
123	A	GLU	OE1	57	B	VAL	O	3.45
123	A	GLU	OE1	57	B	VAL	O	3.45
123	A	GLU	OE2	57	B	VAL	O	2.65
123	A	GLU	OE2	57	B	VAL	O	2.65

264	**A**	**GLN**	**NE2**	**62**	**B**	**SER**	**O**	3.15
264	**A**	**GLN**	**NE2**	**62**	**B**	**SER**	**O**	3.15
266	A	GLU	OE1	63	B	ALA	O	3.34
266	A	GLU	OE1	63	B	ALA	O	3.34
313	A	ASN	ND2	65	B	SER	O	3.31
313	A	ASN	ND2	65	B	SER	O	3.31
337	**A**	**LYS**	**NZ**	**65**	**B**	**SER**	**O**	2.88
337	A	LYS	NZ	67	B	ARG	O	2.66
361	A	ASP	OD2	69	B	ASN	OXT	2.98
361	A	ASP	OD2	69	B	ASN	OXT	2.98
57	**B**	**VAL**	**N**	**123**	**A**	**GLU**	**OE1**	2.7
65	**B**	**SER**	**N**	290	A	ASP	OD2	2.97
69	B	ASN	N	361	A	ASP	OD2	2.48
After Simulation								
Donor				**Acceptor**				
POS	**CHAIN**	**RES**	**ATOM**	**POS**	**CHAIN**	**RES**	**ATOM**	**Dd-a**
264	**A**	**GLN**	**NE2**	**62**	**B**	**SER**	**O**	3.07
264	**A**	**GLN**	**NE2**	**62**	**B**	**SER**	**O**	3.07
337	**A**	**LYS**	**NZ**	**65**	**B**	**SER**	**O**	3.38
56	B	TYR	N	98	A	ASN	ND2	3.47
57	**B**	**VAL**	**N**	**123**	**A**	**GLU**	**OE2**	2.92
65	**B**	**SER**	N	266	A	GLU	OE2	3.41

Table 3.3.16: Hydrogen bond (side chain-side chain) of HAESA -IDA (SERk1 absent) before and after simulation.

Protein-Protein Side Chain-Side Chain Hydrogen Bonds								
Before Simulation								
Donor				**Acceptor**				
POS	**CHAIN**	**RES**	**ATOM**	**POS**	**CHAIN**	**RES**	**ATOM**	**Dd-a**
240	A	ASN	ND2	62	B	SER	OG	2.98
240	A	ASN	ND2	62	B	SER	OG	2.98
264	A	GLN	NE2	62	B	SER	OG	3.17
264	A	GLN	NE2	62	B	SER	OG	3.17
407	A	ARG	NE	**69**	**B**	**ASN**	**OD1**	3.19
407	A	ARG	NE	**69**	**B**	**ASN**	**ND2**	3.14
407	A	ARG	NH2	**69**	**B**	**ASN**	**OD1**	2.52
407	A	ARG	NH2	**69**	**B**	**ASN**	**OD1**	2.52
409	A	ARG	NH1	**69**	**B**	**ASN**	**OD1**	2.58
409	A	ARG	NH1	**69**	**B**	**ASN**	**OD1**	2.58
62	B	SER	OG	240	A	ASN	ND2	2.98

62	**B**	**SER**	**OG**	**242**	**A**	**ASP**	**OD2**	2.72
62	B	SER	OG	264	A	GLN	NE2	3.17
67	B	ARG	NH1	293	A	MET	SD	3.99
67	B	ARG	NH1	293	A	MET	SD	3.99
67	B	ARG	NH1	316	A	GLU	OE1	2.68
67	B	ARG	NH1	316	A	GLU	OE1	2.68
67	B	ARG	NH1	316	A	GLU	OE2	3.24
67	B	ARG	NH1	316	A	GLU	OE2	3.24
After Simulation								
Donor				**Acceptor**				
POS	**CHAIN**	**RES**	**ATOM**	**POS**	**CHAIN**	**RES**	**ATOM**	**Dd-a**
147	A	SER	OG	56	B	TYR	OH	2.93
361	A	ASP	OD1	**69**	**B**	**ASN**	**ND2**	3.15
361	A	ASP	OD1	**69**	**B**	**ASN**	**ND2**	3.15
62	**B**	**SER**	**OG**	**242**	**A**	**ASP**	**OD2**	2.9
69	B	ASN	ND2	361	A	ASP	OD1	3.15
69	B	ASN	ND2	361	A	ASP	OD1	3.15

Table 3.3.17: Protein-Protein ionic interaction of HAESA- IDA (SERk1 absent) before and after simulation.

Protein-Protein Ionic Interactions					
Before Simulation					
Position	**Residue**	**Chain**	**Position**	**Residue**	**Chain**
316	GLU	A	67	ARG	B
After Simulation					
NO PROTEIN-PROTEIN IONIC INTERACTIONS FOUND					

Table 3.3.18: Hydrogen bond (main chain-main chain) of IDA-SERk1 complex (HAESA absent) before and after simulation.

Protein-Protein Main Chain-Main Chain Hydrogen Bonds								
Before Simulation								
Donor				**Acceptor**				
POS	**CHAIN**	**RES**	**ATOM**	**POS**	**CHAIN**	**RES**	**ATOM**	**Dd-a**
68	**B**	**HIS**	**N**	**53**	**C**	**THR**	**O**	3.19
55	**C**	**VAL**	**N**	**68**	**B**	**HIS**	**O**	3.14

After Simulation								
Donor				Acceptor				
POS	CHAIN	RES	ATOM	POS	CHAIN	RES	ATOM	Dd-a
68	B	HIS	N	53	C	THR	O	3.42
55	C	VAL	N	68	B	HIS	O	2.91

Table 3.3.19: Hydrogen bond (main chain-side chain) of IDA-SERk1 complex (HAESA absent) before and after simulation.

Protein-Protein Main Chain-Side Chain Hydrogen Bonds								
Before Simulation								
Donor				Acceptor				
POS	CHAIN	RES	ATOM	POS	CHAIN	RES	ATOM	Dd-a
67	B	ARG	NH2	55	C	VAL	O	3.39
67	B	ARG	NH2	55	C	VAL	O	3.39
After Simulation								
Donor				Acceptor				
POS	CHAIN	RES	ATOM	POS	CHAIN	RES	ATOM	Dd-a
67	B	ARG	NE	52	C	PRO	O	2.92
67	B	ARG	NH2	52	C	PRO	O	2.92
67	B	ARG	NH2	52	C	PRO	O	2.92
67	B	ARG	NH2	54	C	LEU	O	3.37
67	B	ARG	NH2	54	C	LEU	O	3.37

Table 3.3.20: Hydrogen bond (side chain-side chain) of IDA-SERk1 complex (HAESA absent) before and after simulation.

Protein-Protein Side Chain-Side Chain Hydrogen Bonds								
Before Simulation								
NO PROTEIN-PROTEIN SIDE CHAIN-SIDE CHAIN HYDROGEN BONDS FOUND								
After Simulation								
Donor				Acceptor				
POS	CHAIN	RES	ATOM	POS	CHAIN	RES	ATOM	Dd-a
67	B	ARG	NH1	27	C	ASN	OD1	2.86
67	B	ARG	NH1	27	C	ASN	OD1	2.86

Table 3.3.21: Protein-Protein ionic interaction of IDA-SERk1 (HAESA absent) before and after simulation.

Protein-Protein Ionic Interactions					
Before Simulation					
Position	**Residue**	**Chain**	**Position**	**Residue**	**Chain**
68	HIS	B	51	ASP	C
After Simulation					
NO PROTEIN-PROTEIN IONIC INTERACTIONS FOUND					

Table 3.3.22: Summary of interactions among HAESA, IDA and SERk1 before and after simulation

Complex	**Inter. Bet.**	**H-Bond**		**Hydrophobic Interaction**		**Ionic Interaction**		**Cation - Pi Interaction**		**Aromatic - Aromatic Interaction**		**Aromatic - Sulphur Interaction**	
		B. MD	*A. MD*	*B. MD*	*A. MD*	*B. MD*	*A. MD*	*B. MD*	*A. MD*	*B. MD*	*A. MD*	*B. MD*	*A. MD*
HAESA+ IDA+ SERk1	*HAESA + IDA*	36	14	5	7	1	0	0	0	0	0	0	0
	HAESA + SERk1	35	74	9	8	4	2	1	1	1	0	0	0
	IDA+ SERk1	4	3	0	0	1	0	0	0	0	0	0	0
HAESA+ SERk1	*HAESA + SERk1*	35	14	9	5	4	1	1	2	1	1	0	0
HAESA+ IDA	*HAESA + IDA*	36	12	5	5	1	0	0	0	0	0	0	0
IDA+ SERk1	*IDA+ SERk1*	4	9	0	0	1	0	0	0	0	0	0	0

3.4 Root Mean Square Deviation (RMSD)

The stability of the MD simulation was measured in terms of deviations by analyzing Root mean square deviation (RMSD). The time evolution of the RMSDs of HAESA, IDA and SERk1 (residues only) are monitored as a function of time. The RMSDs of HAESA, IDA and SERk1 in the four simulated systems are shown in Fig 3.4.1 respectively. The residues of three proteins showed variable RMSD in different simulated systems. In HAESA and IDA interaction when SERk1 is present in the complex, it takes 5ns to equilibrate. Before 5ns of the period, it gives maximum deviation at 1.3ns and 2.1ns which are 0.50nm and 0.55nm. Though in between this period graph slightly goes down at 3ns after equilibration it moves forward stably with a standard deviation (SD) of 0.20nm. From 5 ns it gives a stable view and deviated in between 0.41nm to 0.46nm till the end of the simulation. On the other hand, when SERk1 is not present in the complex, deviation of HAESA and IDA complex is not that much stable as before. Rather, at the very beginning of the simulation process graph shows stability. But after 10ns it continues maximum deviation, at 25ns it deviates at 1.125nm which is considered as second highest in terms of that condition. Again, instability is observed in this case from 20ns to 21ns because between this period it shows some high RMSD like 0.85nm to 1.15nm and some very low RMSD (0.25nm to 0.35nm) are also observed.

After analyzing the RMSD of HAESA and IDA in both situations (presence of SERk1 and absence of SERk1), RMSD of HAESA and SERk1 are analyzed similarly. In the presence of IDA inside the complex, HAESA and SERk1 give a stable deviated RMSD. Like before, after 5ns the complex is equilibrated and shows stability. At 10ns it deviates to 1.22nm and after 10ns it dramatically goes down to 1.15nm, but at 25ns of the simulation, the RMSD rises to 1.19nm and with a standard deviation of 0.39nm it continues till the end of the simulation. From 5ns till the end the RMSD deviates from 1.10nm to 1.49nm. As observed before in the case of HAESA and IDA RMSD without SERk1, an unstable deviation is found similarly HAESA and SERk1 RMSD without IDA also shows instability. As it is observed that when IDA is present in the complex HAESA and SERk1 RMSD shows stability from 5ns but in this case, it unstable from 5ns with a standard deviation of 0.98nm till the end of the simulation period. From 5ns till the end of the period RMSD of HAESA and SERk1 has deviated from 0.94nm to 1.95nm.

Moreover, on another simulated system where IDA and SERk1 complex gives RMSD in both situation where HAESA is present and absent is analyzed. In case of HAESA's presence, RMSD of IDA and SERk1 gives 1.05nm deviation at 10ns and slows down to 0.90nm at 20ns. After 25ns the RMSD tries to be stable with a standard deviation of 0.6nm till the end. From 25ns the RMSD of IDA and SERk1 deviates from 1.20nm to 1.40nm. But in another case, where HAESA is absent inside the complex, the RMSD of IDA and SERk1 is very unstable. At 15ns it deviates to 4.5nm but from 25ns to 30ns it has deviated from 5.0nm to 3.5nm. Though after 25ns it tries to be stable with a 2.0nm standard deviation and deviated in between 3.0nm to 5.0nm till the end of the simulation period.

In terms of all of the three cases, it is found that RMSD of HAESA, IDA and SERk1 gives more stability when all of them are inside the complex and interact with each other. Without SERk1, RMSD of HAESA and IDA deviates more and unstable situation is observed more than another situation. So, for interacting IDA with HAESA, SERk1 have to be present inside the complex.

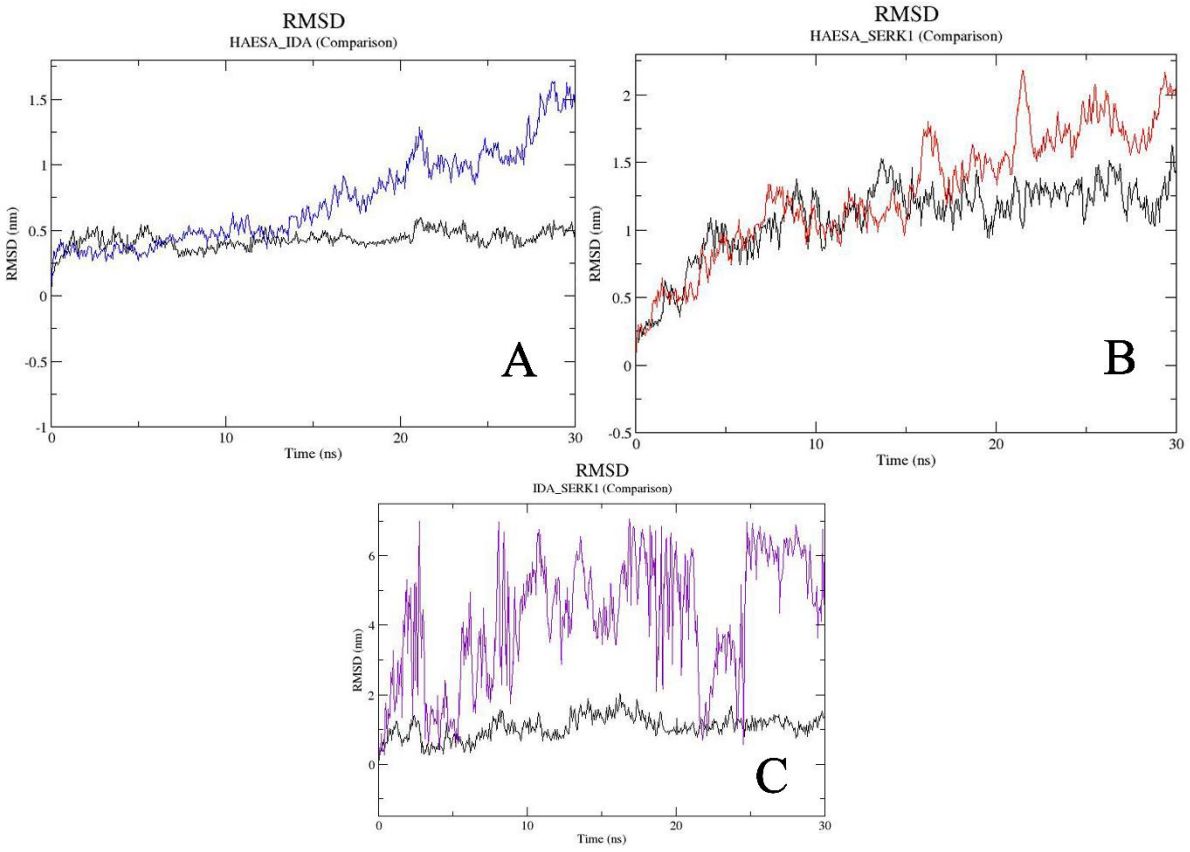

Fig 3.4.1: (A) RMSD value of HAESA and IDA from 30ns MD trajectories. RMSD of HAESA and IDA when SERk1 is present in the complex (Black). RMSD of HAESA and IDA when SERk1 is absent in the complex (Blue). (B) RMSD value of HAESA and SERk1 from 30ns MD trajectories. RMSD of HAESA and SERk1 when IDA is present in the complex (Black). RMSD of HAESA and SERk1 in the absence of IDA (Red). (C) RMSD value of IDA and SERk1 from 30ns MD trajectories. RMSD of IDA and SERk1 at the presence of HAESA (Black). RMSD value of IDA and SERk1 in the absence of HAESA inside the complex (purple).

3.5 Radius of Gyration

To measure the compactness of all systems, the radius of gyration (Rg) values were measured. Overall, it could be seen that the Rg of the HAESA-SERk1 complex fluctuated more compared to that of the HAESA-IDA-SERk1 complex and HAESA-IDA complex. This trend is most pronounced in the 21 ns of the simulation; as the HAESA-SERk1 complex reaches the highest value of about 3.68 nm in the first 16 ns and then drops and reaches the lowest value of about 3.40 nm near 20 ns mark. Whereas, the graph remains relatively steady for the HAESA-IDA-SERk1 complex in this time frame; though it also reaches its highest (around 3.4 nm) and lowest (around 3.25 nm) points at around 16.6 ns and 3 ns, respectively. Though in HAESA-IDA complex at the very beginning of the simulation (3ns) it gives the lowest value of 3.32nm after 16.6ns it rises to 3.6nm and continues the stability. In IDA-SERk1 complex, many instabilities occur. Though at 21ns it rises to 1.95nm at 22.4ns it drops to 1.81nm. More variations in the Rg values signify a more changing structure, which is consistent with higher fluctuations in the HAESA-SERk1 complex and HAESA-IDA complex; as the proteins are more freely able to uncoil and recoil. On the other hand, when the HAESA is interacting with both IDA and SERk1 in a single complex, it is bound in place and less able to uncoil resulting in a less fluctuating graph.

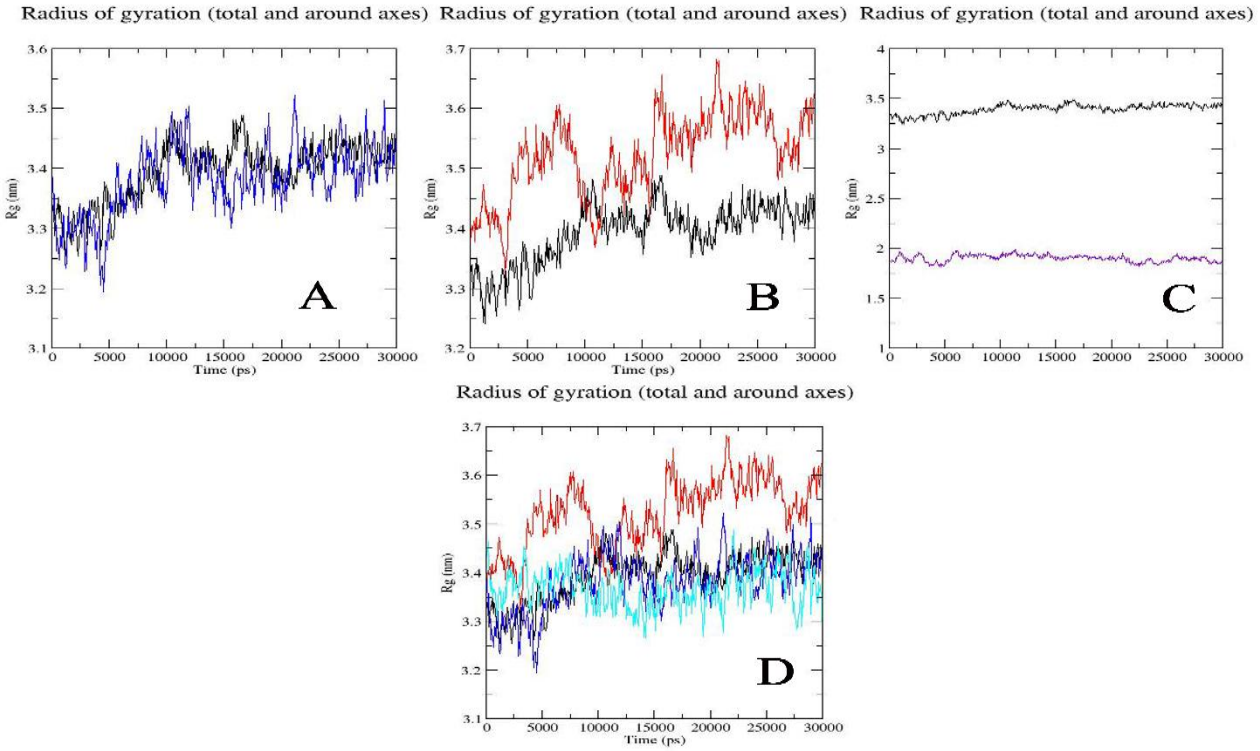

Fig 3.5.1: (A) Rg value of HAESA and IDA from 30ns MD trajectories. Rg of HAESA and IDA when SERk1 is present in the complex (Black). Rg of HAESA and IDA when SERk1 is absent in the complex (Blue). (B) Rg value of HAESA and SERk1 from 30ns MD trajectories. Rg of HAESA and SERk1 when IDA is present in the complex (Black). Rg of HAESA and SERk1 in the absence of IDA (Red). (C) Rg value of IDA and SERk1 from 30ns MD trajectories. Rg of IDA and SERk1 at the presence of HAESA (Black). Rg value of IDA and SERk1 in the absence of HAESA inside the complex (purple). (D) Rg value of all complex with the addition of HAESA only (Sky Blue).

3.6 Solvent Accessible Surface Area (SASA)

The solvent assessable surface areas (SASA) were also calculated and it was seen to be considerably lower for the HAESA-IDA-SERk1 complex than HAESA-SERk1 complex, showing a steady mean value of about 381 nm^2. Opposed to this, the HAESA-SERk1 complex had a much higher mean value of about 386 nm^2. Again, SASA value of HAESA-IDA complex is 305 nm^2 where IDA-SERk1 value is very low.

This meant that the interaction of IDA with the HAESA LRR region and also interaction of HAESA, IDA and SERk1 greatly increased the surface area of the complex which water (the solvent in this case) could access.

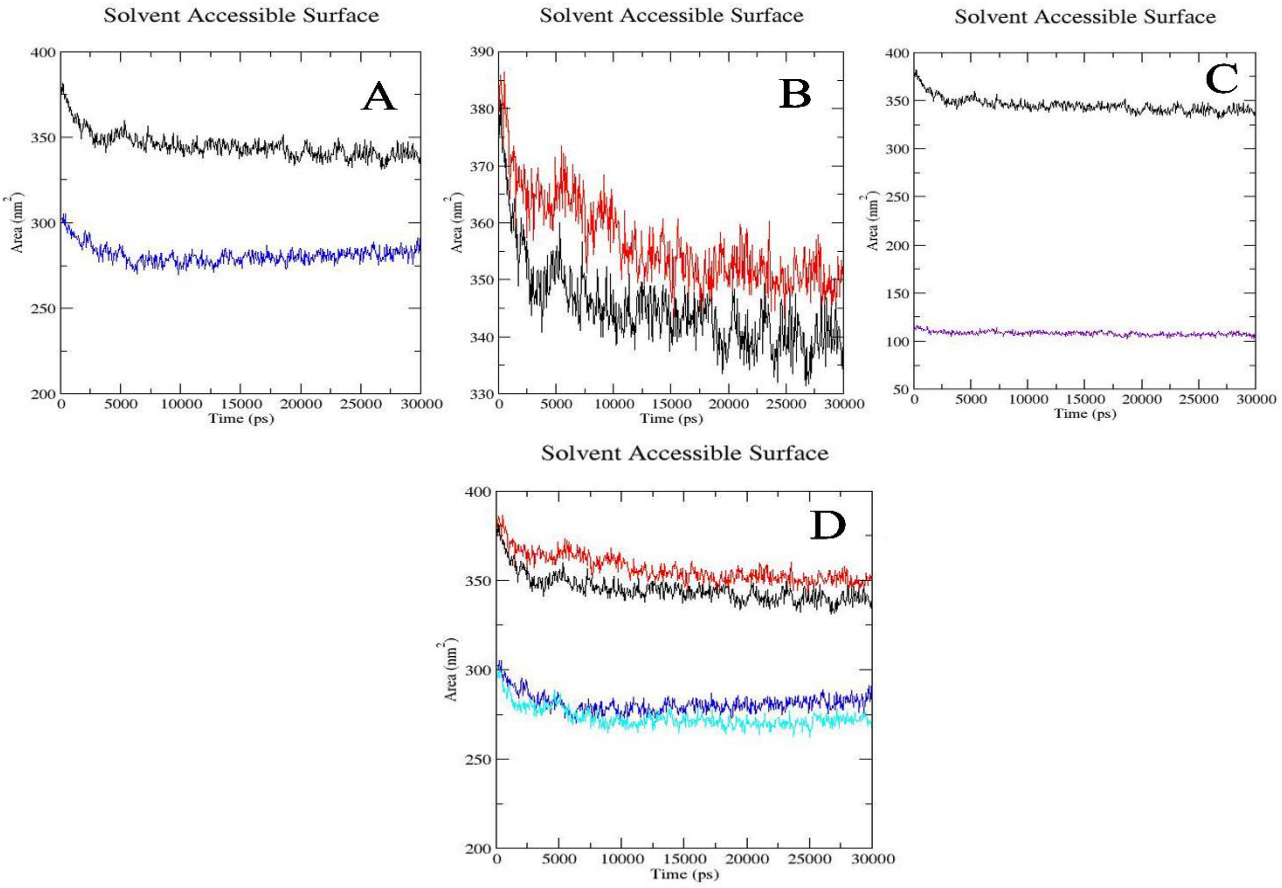

Fig 3.6.1: (A) SASA value of HAESA and IDA from 30ns MD trajectories. SASA of HAESA and IDA when SERk1 is present in the complex (Black). SASA of HAESA and IDA when SERk1 is absent in the complex (Blue). (B) SASA value of HAESA and SERk1 from 30ns MD trajectories. SASA of HAESA and SERk1 when IDA is present in the complex (Black). SASA of HAESA and SERk1 in the absence of IDA (Red). (C) SASA value of IDA and SERk1 from 30ns MD trajectories. SASA of IDA and SERk1 at the presence of HAESA (Black). SASA value of IDA and SERk1 in the absence of HAESA inside the complex (purple). (D) SASA value of all complex with the addition of HAESA only (Sky Blue).

3.7 Discussion

This study describes the computational analysis of a plant PRR and reports on the interaction of PRR-RLK HAESA to PAMP IDA and Co-receptor SERk1. Different analytical approaches used in this study for understanding Pattern Triggered Immunity (PTI) of Arabidopsis thaliana towards IDA by using HAESA LRR domain.

Besides, the in-depth analysis to the interactions between IDA and HAESA LRR domain, reveals that the 56^{th} and 67^{th} (tyrosine and arginine, respectively) AA residues of IDA play an important role in the interaction with HAESA LRR. Moreover, in MM/PBSA calculation it is found that during the presence of SERk1 in the complex IDA easily interacts with HAESA where C-terminus residue Arg67 plays an important role for binding with Lys337 of HAESA, this residue gives higher binding energy which makes the interaction between IDA and HAESA has happened in the presence of SERk1. Where in the absence of SERk1 there is no notable binding energy found.

From RMSF and RMSD studies, during the presence of SERk1, the interaction between IDA and HAESA goes to a stable form after a certain period of simulation and also prominent residues are found very low fluctuated during this time. From H-bond analysis and protein interaction calculation (PIC) data, it is observed, previously described favourable residues give a various type of interactions alongside hydrogen bond.

For another point of view, a comparison between previously studied crystallographic structures and HAESA-IDA-SERk1 complex is followed. For interacting between LRR-RKs and peptide hormones some steps are followed naturally. Firstly, residues of peptidyl hormones interact with the inner portion of LRR in a full conformation.[125, 126] Here, IDA binds into the inner superhelix portion of HAESA. In the case of flg22, it also binds to the inner portion of FLS2 and TDIF binds to the inner portion of TDR in Arabidopsis thaliana. Secondly, the direction of peptide hormones or ligands and ectodomains is same or vice versa.[125, 126] Though, in this case, IDA binds inversely with HAESA. Arg67 (C-terminus residue) from IDA binds with Lys337 of HAESA. Finally, LRR-RK ectodomains play the role of rulers which measure the appropriate

length of a bioactive peptide by interacting with very specific mature C terminus of peptide hormones.[125, 126] In terms of IDA and HAESA interaction Lys56 which is a mature C-terminus residue binds with the receptor of HAESA. In the case of flg22 and TDIF, their mature C-terminus residues are also bound with FLS2 and TDF respectively.

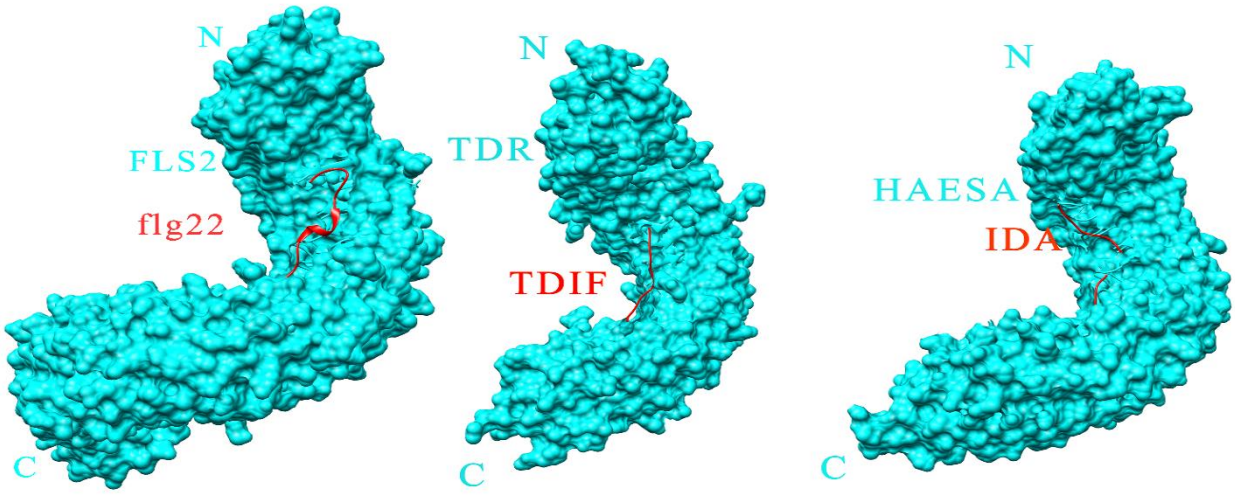

Fig 3.7.1: Surface structural comparison of HAESA LRR with FLS2 and TDR LRR during interaction of PAMP.

In the case of LRR-RK and Co-Receptor interaction, a large number of residues from the inner portion of the ectodomains of SERks bind with different types of receptors. Without interacting with small molecules or peptidyl hormones SERks highly overlap on LRR-RK ectodomains.[4] When HAESA and SERk1 are interacting within themselves binding energy is getting lower and in the presence of IDA, the binding affinity is higher without happening negligible interaction with IDA.

past studies show that various peptidyl hormones or ligands can tightly bind with LRR-RKs with their favourable co-receptors. The interacting heterodimeric complex brings co-receptor close to the transmembrane helices

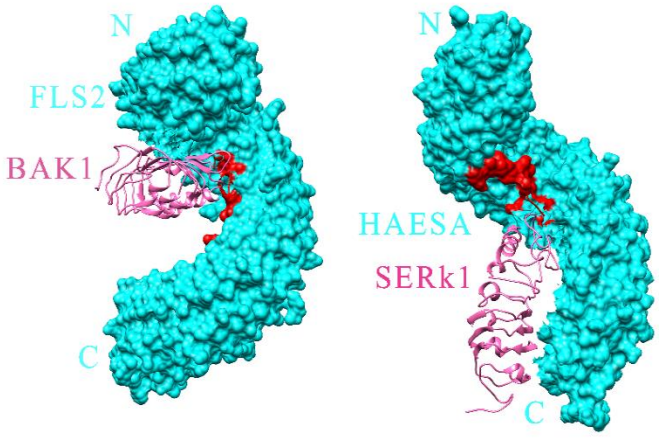

Fig 3.7.2: Surface structural comparison of HAESA LRR with FLS2 during interaction of Co-Receptor.

which help to do interaction between co-receptor and the cytoplasmic kinase domains.[4] In this study, it is also found that in the presence of SERk1 co-receptor IDA can bind favourably with HAESA and the position of SERk1 is nearly one end of HAESA.

And it can be hypothesized that a mutation at these points of LRR-RK of HAESA can greatly affect the plants' ability to recognize the PAMP. Again, co-receptor helps to recognize PAMP, so mutation at co-receptor also causes conformational changes in binding with LRR-RK which also affect the recognition of PAMP.

Chapter 4

Conclusions and Recommendations

4.1 Conclusion:

From the analysis of 30ns trajectories of HAESA, IDA and SERk1 complex by using RMSD, RMSF, H-bond, PIC and MM/PBSA it can be said that for immune response against IDA there should be co-receptor SERk1 present inside the complex. Though from 30ns trajectories RMSD it is found that after 30ns the whole complex tends to be stable but more stability can be found if the simulation period will be extended to 50ns or 100ns. Moreover, the prominent residues found through MM/PBSA calculation is calculated from 30ns trajectories. If the simulation period will be extended in future more prominent residues can be obtained again the observed prominent residues obtained from 30ns trajectories can be confirmed if they will remain prominent like before or not.

From H-bond and PIC result after 30ns simulation there are notable differences found in different types of interaction alongside H-bond than before the simulation. Extended simulation period will provide that what types of interactions are more responsible and which interaction is going less important.

But from 30ns trajectories, as it is found for immune response from HAESA co-receptor plays a notable role, so it can be assumed that after the extension of the simulation period it will remain the same.

4.2 Recommendations for Future Works:

This research can be further developed by adopting some measure such as:

1. This study can be improved by running the molecular dynamic simulation for much longer (50 ns or 100 ns), this would allow more selective and conclusive results from the study. More understanding of protein nature can be built.

2. This study might be useful for the interaction of HAESA LRR with other mutated PAMPs in plant Arabidopsis thaliana. Interaction of PRR-PAMP complex and the involvement of mutated

co-receptor can show how mutant molecule can make changes in certain residues and effects in their interactions which leads to pattern triggered immunity (PTI).

3. MM/PBSA is a post-processing method whereby the free energy of a state is determined from the interior energy (MM) of the residues and its connection with an understandable portrayal of solvent (PBSA). FEP, like other free energy assessments (TI, BAR, and so forth) assesses free energy contrasts of a given design under an alternate Hamiltonian (for example one whose cooperations are scaled by a lambda factor). Estimations that utilize FEP are substantially costlier than MM/PBSA, which uses a single trajectory, run utilizing typical MD.

While other free energy techniques, similar to TI and BAR, are more precise than MM/PBSA, they are restrictively costly for bigger solutes like macromolecules. For assessing restricting free energies where the ligand to be decoupled is little. For solvation free energies of proteins, DNA, and so forth, MM/PBSA is adequately exact and maintains a strategic distance from the immense cost that would emerge from utilizing decoupling strategies.

References

1. 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.
2. Boller, T. and S.Y.J.S. He, *Innate immunity in plants: an arms race between pattern recognition receptors in plants and effectors in microbial pathogens.* 2009. **324**(5928): p. 742-744.
3. Chisholm, S.T., et al., *Host-microbe interactions: shaping the evolution of the plant immune response.* 2006. **124**(4): p. 803-814.
4. Hohmann, U., K. Lau, and M.J.A.r.o.p.b. Hothorn, *The structural basis of ligand perception and signal activation by receptor kinases.* 2017. **68**: p. 109-137.
5. Jones, J.D. and J.L.J.n. Dangl, *The plant immune system.* 2006. **444**(7117): p. 323.
6. Zipfel, C.J.T.i.i., *Plant pattern-recognition receptors.* 2014. **35**(7): p. 345-351.
7. Ting, J.P. and B.K.J.A.R.I. Davis, *CATERPILLER: a novel gene family important in immunity, cell death, and diseases.* 2005. **23**: p. 387-414.
8. nature, A.G.I.J., *Analysis of the genome sequence of the flowering plant Arabidopsis thaliana.* 2000. **408**(6814): p. 796.
9. Meyerowitz, E.M. and C.R. Somerville, *Arabidopsis*. 1994: Cold Spring Harbor Laboratory Press.
10. Meinke, D.W., et al., *Arabidopsis thaliana: a model plant for genome analysis.* 1998. **282**(5389): p. 662-682.
11. Feldmann, K.A. and S.A. Goff, *The first plant genome sequence—Arabidopsis thaliana*, in *Advances in Botanical Research*. 2014, Elsevier. p. 91-117.
12. Sparrow, A., H. Price, and A. Underbrink. *A survey of DNA content per cell and per chromosome of prokaryotic and eukaryotic organisms: some evolutionary considerations*. in *Brookhaven symposia in biology*. 1972.
13. Sparrow, A.H. and J.P.J.S. Miksche, *Correlation of nuclear volume and DNA content with higher plant tolerance to chronic radiation.* 1961. **134**(3474): p. 282-283.
14. Lin, X., et al., *Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.* 1999. **402**(6763): p. 761.

15. Mayer, K., et al., *Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.* 1999. **402**(6763): p. 769.
16. Theologis, A., et al., *Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana.* 2000. **408**(6814): p. 816.
17. Meyerowitz, E.M.J.P.p., *Prehistory and history of Arabidopsis research.* 2001. **125**(1): p. 15-19.
18. Yanofsky, M.F., et al., *The protein encoded by the Arabidopsis homeotic gene agamous resembles transcription factors.* 1990. **346**(6279): p. 35.
19. Laibach, F., *Zur Frage nach der Individualität der Chromosomen im Pflanzenreich*. 1907.
20. Ezhova, T.J.G., *Arabidopsis thaliana (L.) Heynh. as a model object for studying genetic control of morphogenesis.* 1999. **35**(11): p. 1522-1537.
21. Titova, N.J.S.B., *Poiski rastitePnoy drozophily.* 1935. **2**: p. 61-67.
22. Laibach, F.J.B.A., *Arabidopsis thaliana (L.) Heynh. als Objekt für genetische und entwicklungsphysiologische Untersuchungen.* 1943. **44**: p. 439-455.
23. Rédei, G.P.J.A.r.o.g., *Arabidopsis as a genetic tool.* 1975. **9**(1): p. 111-127.
24. Tsuda, K. and F.J.C.o.i.p.b. Katagiri, *Comparing signaling mechanisms engaged in pattern-triggered and effector-triggered immunity.* 2010. **13**(4): p. 459-465.
25. Tsuda, K., et al., *Network properties of robust immunity in plants.* 2009. **5**(12): p. e1000772.
26. Block, A., et al., *Phytopathogen type III effector weaponry and their plant targets.* 2008. **11**(4): p. 396-403.
27. Cunnac, S., M. Lindeberg, and A.J.C.o.i.m. Collmer, *Pseudomonas syringae type III secretion system effectors: repertoires in search of functions.* 2009. **12**(1): p. 53-60.
28. Göhre, V. and S.J.A.R.P. Robatzek, *Breaking the barriers: microbial effector molecules subvert plant immunity.* 2008. **46**: p. 189-215.
29. Böhm, H., et al., *Immune receptor complexes at the plant cell surface.* 2014. **20**: p. 47-54.
30. Ranf, S.J.C.o.i.p.b., *Sensing of molecular patterns through cell surface immune receptors.* 2017. **38**: p. 68-77.
31. Xin, X.-F., et al., *Bacteria establish an aqueous living space in plants crucial for virulence.* 2016. **539**(7630): p. 524.

32. Rosebrock, T.R., et al., *A bacterial E3 ubiquitin ligase targets a host protein kinase to disrupt plant immunity.* 2007. **448**(7151): p. 370.
33. Collier, S.M. and P.J.T.i.p.s. Moffett, *NB-LRRs work a "bait and switch" on pathogens.* 2009. **14**(10): p. 521-529.
34. Kumpf, R.P., et al., *Floral organ abscission peptide IDA and its HAE/HSL2 receptors control cell separation during lateral root emergence.* 2013. **110**(13): p. 5235-5240.
35. Lease, K.A. and J.C.J.P.p. Walker, *The Arabidopsis unannotated secreted peptide database, a resource for plant peptidomics.* 2006. **142**(3): p. 831-838.
36. Shiu, S.-H. and A.B.J.P.o.t.N.A.o.S. Bleecker, *Receptor-like kinases from Arabidopsis form a monophyletic gene family related to animal receptor kinases.* 2001. **98**(19): p. 10763-10768.
37. Dunning, F.M., et al., *Identification and mutational analysis of Arabidopsis FLS2 leucine-rich repeat domain residues that contribute to flagellin perception.* 2007. **19**(10): p. 3297-3313.
38. Santiago, J., et al., *Mechanistic insight into a peptide hormone signaling complex mediating floral organ abscission.* 2016. **5**: p. e15075.
39. Jinn, T.-L., et al., *HAESA, an Arabidopsis leucine-rich repeat receptor kinase, controls floral organ abscission.* 2000. **14**(1): p. 108-117.
40. Stenvik, G.-E., et al., *The EPIP peptide of INFLORESCENCE DEFICIENT IN ABSCISSION is sufficient to induce abscission in Arabidopsis through the receptor-like kinases HAESA and HAESA-LIKE2.* 2008. **20**(7): p. 1805-1817.
41. Hanks, S.K. and T.J.T.F.j. Hunter, *Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification.* 1995. **9**(8): p. 576-596.
42. Hanks, S.K., A.M. Quinn, and T.J.S. Hunter, *The protein kinase family: conserved features and deduced phylogeny of the catalytic domains.* 1988. **241**(4861): p. 42-52.
43. Akira, S. and K.J.N.r.i. Takeda, *Toll-like receptor signalling.* 2004. **4**(7): p. 499.
44. Medzhitov, R. and C.A.J.S. Janeway, *Decoding the patterns of self and nonself by the innate immune system.* 2002. **296**(5566): p. 298-300.
45. Zipfel, C., et al., *Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts Agrobacterium-mediated transformation.* 2006. **125**(4): p. 749-760.

46. Nürnberger, T., et al., *Innate immunity in plants and animals: striking similarities and obvious differences.* 2004. **198**(1): p. 249-266.
47. Ausubel, F.M.J.N.i., *Are innate immune signaling pathways in plants and animals conserved?* 2005. **6**(10): p. 973.
48. Zipfel, C. and G.J.C.o.i.p.b. Felix, *Plants and animals: a different taste for microbes?* 2005. **8**(4): p. 353-360.
49. Espinosa, A. and J.R.J.C.m. Alfano, *Disabling surveillance: bacterial type III secretion system effectors that suppress innate immunity.* 2004. **6**(11): p. 1027-1040.
50. Kim, M.G., et al., *Two Pseudomonas syringae type III effectors inhibit RIN4-regulated basal defense in Arabidopsis.* 2005. **121**(5): p. 749-759.
51. Nomura, K., M. Melotto, and S.-Y.J.C.o.i.p.b. He, *Suppression of host defense in compatible plant–Pseudomonas syringae interactions.* 2005. **8**(4): p. 361-368.
52. Jones, D.A. and D.J.C.o.i.i. Takemoto, *Plant innate immunity–direct and indirect recognition of general and specific pathogen-associated molecules.* 2004. **16**(1): p. 48-62.
53. Nimchuk, Z., et al., *Recognition and response in the plant immune system.* 2003. **37**(1): p. 579-609.
54. Aalen, R.B., et al., *IDA: a peptide ligand regulating cell separation processes in Arabidopsis.* 2013. **64**(17): p. 5253-5261.
55. Cho, S.K., et al., *Regulation of floral organ abscission in Arabidopsis thaliana.* 2008. **105**(40): p. 15629-15634.
56. Butenko, M.A., et al., *Inflorescence deficient in abscission controls floral organ abscission in Arabidopsis and identifies a novel family of putative ligands in plants.* 2003. **15**(10): p. 2296-2307.
57. Stenvik, G.-E., et al., *Overexpression of INFLORESCENCE DEFICIENT IN ABSCISSION activates cell separation in vestigial abscission zones in Arabidopsis.* 2006. **18**(6): p. 1467-1476.
58. Butenko, M.A., et al., *Tools and strategies to match peptide-ligand receptor pairs.* 2014. **26**(5): p. 1838-1847.
59. Karlova, R., et al., *The Arabidopsis somatic embryogenesis receptor-like kinase1 protein complex includes brassinosteroid-insensitive1*. 2006. **18**(3): p. 626-638.

60. Albrecht, C., et al., *The Arabidopsis thaliana SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASES1 and 2 control male sporogenesis.* 2005. **17**(12): p. 3337-3349.
61. Hecht, V., et al., *The Arabidopsis SOMATIC EMBRYOGENESIS RECEPTOR KINASE 1 gene is expressed in developing ovules and embryos and enhances embryogenic competence in culture.* 2001. **127**(3): p. 803-816.
62. Kwaaitaal, M.A.C.J., S. De Vries, and E.J.P. Russinova, *Arabidopsis thaliana Somatic Embryogenesis Receptor Kinase 1 protein is present in sporophytic and gametophytic cells and undergoes endocytosis.* 2005. **226**(1-2): p. 55-65.
63. Shah, K., et al., *The Arabidopsis kinase-associated protein phosphatase controls internalization of the somatic embryogenesis receptor kinase 1.* 2002. **16**(13): p. 1707-1720.
64. Colcombet, J., et al., *Arabidopsis SOMATIC EMBRYOGENESIS RECEPTOR KINASES1 and 2 are essential for tapetum development and microspore maturation.* 2005. **17**(12): p. 3350-3361.
65. Meng, X., et al., *Ligand-induced receptor-like kinase complex regulates floral organ abscission in Arabidopsis.* 2016. **14**(6): p. 1330-1338.
66. Niederhuth, C.E., O.R. Patharkar, and J.C.J.B.g. Walker, *Transcriptional profiling of the Arabidopsis abscission mutant hae hsl2 by RNA-Seq.* 2013. **14**(1): p. 37.
67. Lease, K.A., S.K. Cho, and J.C.J.P.M. Walker, *A petal breakstrength meter for Arabidopsis abscission studies.* 2006. **2**(1): p. 2.
68. Santiago, J., C. Henzler, and M.J.S. Hothorn, *Molecular mechanism for plant steroid receptor activation by somatic embryogenesis co-receptor kinases.* 2013. **341**(6148): p. 889-892.
69. Sun, Y., et al., *Structure reveals that BAK1 as a co-receptor recognizes the BRI1-bound brassinolide.* 2013. **23**(11): p. 1326.
70. Wang, J., et al., *Allosteric receptor activation by the plant peptide hormone phytosulfokine.* 2015. **525**(7568): p. 265.
71. Shah, K., J. Vervoort, and S.C.J.J.o.B.C. de Vries, *Role of threonines in the Arabidopsis thaliana somatic embryogenesis receptor kinase 1 activation loop in phosphorylation.* 2001. **276**(44): p. 41263-41269.

72. Taylor, I., et al., *Analysis of phosphorylation of the receptor-like protein kinase HAESA during Arabidopsis floral abscission.* 2016. **11**(1): p. e0147203.
73. Ohyama, K., et al., *A glycopeptide regulating stem cell fate in Arabidopsis thaliana.* 2009. **5**(8): p. 578.
74. Barnes, J.E.J.N., *Evolution of compact groups and the formation of elliptical galaxies.* 1989. **338**(6211): p. 123.
75. Karplus, M. and G.A.J.N. Petsko, *Molecular dynamics simulations in biology*. 1990. **347**(6294): p. 631.
76. Karplus, M., J.A.J.N.S. McCammon, and M. Biology, *Molecular dynamics simulations of biomolecules.* 2002. **9**(9): p. 646.
77. Becker, O.M., et al., *Computational biochemistry and biophysics*. 2001: Marcel Dekker New York.
78. Tuckerman, M.E. and G.J. Martyna, *Understanding modern molecular dynamics: techniques and applications*. 2000, ACS Publications.
79. 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.
80. Allen, M.P.J.C.s.m.f.s.p.t.p., *Introduction to molecular dynamics simulation.* 2004. **23**: p. 1-28.
81. Bhaduri, A., et al., *Conserved spatially interacting motifs of protein superfamilies: application to fold recognition and function annotation of genome data.* 2004. **54**(4): p. 657-670.
82. Gromiha, M.M., S.J.P.i.b. Selvaraj, and m. biology, *Inter-residue interactions in protein folding and stability.* 2004. **86**(2): p. 235-277.
83. Reva, B.A., et al., *Recognition of protein structure on coarse lattices with residue-residue energy functions.* 1997. **10**(10): p. 1123-1130.
84. Russell, R.B. and G.J.J.J.o.m.b. Barton, *Structural features can be unconserved in proteins with similar folds: An analysis of side-chain to side-chain contacts secondary structure and accessibility.* 1994. **244**(3): p. 332-350.
85. Šali, A. and T.L.J.J.o.m.b. Blundell, *Comparative protein modelling by satisfaction of spatial restraints.* 1993. **234**(3): p. 779-815.

86. Shao, X. and N.V.J.N.A.R. Grishin, *Common fold in helix–hairpin–helix proteins.* 2000. **28**(14): p. 2643-2650.
87. Tina, K., R. Bhadra, and N.J.N.a.r. Srinivasan, *PIC: protein interactions calculator.* 2007. **35**(suppl_2): p. W473-W476.
88. Krupa, A., G. Preethi, and N.J.J.o.m.b. Srinivasan, *Structural modes of stabilization of permissive phosphorylation sites in protein kinases: distinct strategies in Ser/Thr and Tyr kinases.* 2004. **339**(5): p. 1025-1039.
89. Van Roey, P., et al., *Crystallographic and mutational studies of Mycobacterium tuberculosis recA mini-inteins suggest a pivotal role for a highly conserved aspartate residue.* 2007. **367**(1): p. 162-173.
90. Jones, S., J.M.J.P.i.b. Thornton, and m. biology, *Protein-protein interactions: a review of protein dimer structures.* 1995. **63**(1): p. 31-65.
91. Nooren, I.M. and J.M.J.T.E.j. Thornton, *Diversity of protein–protein interactions.* 2003. **22**(14): p. 3486-3492.
92. Ofran, Y. and B.J.J.o.m.b. Rost, *Analysing six types of protein–protein interfaces.* 2003. **325**(2): p. 377-387.
93. Rekha, N., et al., *Interaction interfaces of protein domains are not topologically equivalent across families within superfamilies: Implications for metabolic and signaling pathways.* 2005. **58**(2): p. 339-353.
94. Saha, R.P., et al., *Interaction geometry involving planar groups in protein–protein interfaces.* 2007. **67**(1): p. 84-97.
95. Bahadur, R.P., et al., *A dissection of specific and non-specific protein–protein interfaces.* 2004. **336**(4): p. 943-955.
96. De, S., et al., *Interaction preferences across protein-protein interfaces of obligatory and non-obligatory components are different.* 2005. **5**(1): p. 15.
97. Valdar, W.S. and J.M.J.J.o.m.b. Thornton, *Conservation helps to identify biologically relevant crystal contacts.* 2001. **313**(2): p. 399-416.
98. 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.
99. 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.

100. Miller III, B.R., et al., *MMPBSA. py: an efficient program for end-state free energy calculations.* 2012. **8**(9): p. 3314-3321.
101. 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.
102. 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.
103. 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.
104. 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.
105. 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.
106. 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.
107. 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.
108. 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.
109. 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.
110. Kumari, R., et al., *g_mmpbsa□ A GROMACS tool for high-throughput MM-PBSA calculations.* 2014. **54**(7): p. 1951-1962.
111. Connolly, M.L.J.S., *Solvent-accessible surfaces of proteins and nucleic acids.* 1983. **221**(4612): p. 709-713.
112. Lee, B. and F.M.J.J.o.m.b. Richards, *The interpretation of protein structures: estimation of static accessibility.* 1971. **55**(3): p. 379-IN4.

113. Richards, F., *Annu. Rev. Biophys. Bioeng.* 1977.
114. Greer, J. and B.L.J.P.o.t.N.a.o.S. Bush, *Macromolecular shape and surface maps by solvent exclusion.* 1978. **75**(1): p. 303-307.
115. Jiang, F. and S.-H.J.J.o.m.b. Kim, *"Soft docking": matching of molecular surface cubes.* 1991. **219**(1): p. 79-102.
116. Shrake, A. and J.J.J.o.m.b. Rupley, *Environment and exposure to solvent of protein atoms. Lysozyme and insulin.* 1973. **79**(2): p. 351-371.
117. Huggins, M.L.J.A.C.I.E.i.E., *50 years of hydrogen bond theory.* 1971. **10**(3): p. 147-152.
118. McDonald, I.K. and J.M.J.J.o.m.b. Thornton, *Satisfying hydrogen bonding potential in proteins.* 1994. **238**(5): p. 777-793.
119. Creighton, T.E.J.C.o.i.s.b., *Stability of folded conformations: Current opinion in structural biology 1991, 1: 5–16.* 1991. **1**(1): p. 5-16.
120. Dill, K.A.J.B., *Dominant forces in protein folding.* 1990. **29**(31): p. 7133-7155.
121. Van Der Spoel, D., et al., *GROMACS: fast, flexible, and free.* 2005. **26**(16): p. 1701-1718.
122. van der Spoel, D., P.J. van Maaren, and H.J.J.T.J.o.c.p. Berendsen, *A systematic study of water models for molecular simulation: derivation of water models optimized for use with a reaction field.* 1998. **108**(24): p. 10220-10230.
123. Sun, Y., et al., *Structural basis for flg22-induced activation of the Arabidopsis FLS2-BAK1 immune complex.* 2013. **342**(6158): p. 624-628.
124. Morita, J., et al., *Crystal structure of the plant receptor-like kinase TDR in complex with the TDIF peptide.* 2016. **7**: p. 12383.
125. Shinohara, H., et al., *Identification of three LRR-RKs involved in perception of root meristem growth factor in Arabidopsis.* 2016. **113**(14): p. 3897-3902.
126. Somssich, M., et al., *Real-time dynamics of peptide ligand–dependent receptor complex formation in planta.* 2015. **8**(388): p. ra76-ra76.

Appendixes

Different bioinformatics tools used in this study

Serial No.	Tool	Used for
1	Gromacs	MD simulation
2	Protein interaction calculator: PIC (online)	Residual bond identification
3	Chimera	Molecular visualization
4	g_mmpbsa	Binding free energy calculation
5	xmgrace	Graph generation and analysis