# DESIGN OF ION TRANSPORTER-BASED ON POLYPEPTIDE MARKERS FOR DETECTION OF SALT TOLERANCE IN PLANTS

Thesis

Submitted by

**Sambit Majumdar**

Roll No.: 602204016

In partial fulfilment of the requirements for the award of the degree

Master of Technology

Biotechnology

Under the supervision of

**Dr. Debajyoti Dutta**

Assistant Professor

Department of Biotechnology

THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

**DEPARTMENT OF BIOTECHNOLOGY ENGINEERING**
**THAPAR INSTITUTE OF ENGINEERING AND TECHNOLOGY,**
**PATIALA, PUNJAB - 147004**
**(Deemed-to-be-University u/s 3 of the UGC Act,1956)**
**(July, 2024)**

# DECLARATION

I Sambit Majumdar, hereby declare that the thesis entitled "**Design of ion transporter-based on polypeptide markers for detection of salt tolerance in plants**" submitted by me, for the degree of Master of Technology in Biotechnology of Thapar Institute of Engineering and Technology, Patiala is a record of honest work carried by me under the supervision and guidance of supervision of Dr. Debajyoti Dutta, Assistant Professor at the Department of Biotechnology, Thapar Institute of Engineering and Technology, Patiala. I have not submitted the matter embodied in this thesis for the award of any other degree.

Sambit Majumdar

Date: **29/07/2024**

Place: **TIET, Patiala**

**Sambit Majumdar**
**M.tech Biotechnology**
**602204016**

THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

# CERTIFICATE

It is to be certified that the thesis entitled **“Design of ion transporter-based on polypeptide markers for detection of salt tolerance in plants”** submitted by Sambit Majumdar (602204016) to the Department of Biotechnology, Thapar Institute of Engineering and Technology Patiala, in partial fulfilment for the award of the degree Master of Technology in Biotechnology, is a record of original work carried out by him from 1st August 2023 to 29th July 2024 under the supervision and guidance of supervision of Dr. Debajyoti Dutta, Assistant Professor at the Department of Biotechnology, Thapar Institute of Engineering and Technology, Patiala. The matter embodied in the thesis work has not been submitted for the award of any degree.

Debajyoti Dutta

**Dr. Debajyoti Dutta**
Assistant Professor
(Supervisor)
Department of Biotechnology
Thapar Institute of Engineering and Technology
Patiala- 147004.

# ACKNOWLEDGEMENT

I would like to express my sincere thanks to all those who have supported and guided me throughout the journey of accomplishing this thesis. It is a great honour and privilege to complete my dissertation at the Thapar Institute of Engineering and Technology, Patiala.

I would also like to express my profound, deep gratitude and deep regards **to Dr. Debajyoti Dutta**, Assistant Professor, Department of Biotechnology, Thapar Institute of Engineering and Technology, Patiala, for giving me this opportunity. His knowledge and guidance were essential to accomplishing this thesis. I am extremely thankful for the opportunity and for having faith on my potential.

I express my gratitude to **Dr. M.S.Reddy, Professor, Head of Department of Biotechnology**, for consenting me for this thesis.

I would also like to thank PhD scholars **Ms Sanhita Banik** and **Ms Tanya Bhatnagar,** laboratory superintendent **Mr. Ram Newal Yadav** and lab technician **Mr. Mohinder**, for providing their assistance in laboratory work and other lab requirements. Also, I would like to thank my parents, lab mates and friends for their constant support and encouragement.

**Sambit Majumdar**

**M.tech Biotechnology**

**Roll No.- 602204016**

# Tables of Contents

# LIST OF FIGURES

# LIST OF TABLES

# ANNEXURE-I

# LIST OF ABBREVIATION USED

| | |
|---|---|
| At | *Arabidopsis thaliana* |
| ATP | Adenosine Triphosphate |
| BLAST | Basic Local Alignment Search |
| DNA | Deoxyribose Nucleotide Phosphate |
| $H_2O_2$ | Hydrogen Peroxide |
| SOS1 | Salt Overall Sensitivity 1 |
| HKT1 | High Affinity $K^+$ Transporter 1 |
| NHX1 | Sodium/ Hydrogen Exchanger 1 |
| MAFFT | Multiple Alignment Using Fast Fourier Transform |
| MSA | Multiple Sequence Alignment |
| NaCl | Sodium Chloride |
| NCBI | National Center for Biotechnology Information |
| PDB | Protein Data Bank |
| PM | Plasma Membrane |
| ROS | Reactive Oxygen Species |
| SS | Salinity Stress |
| AtSOS1 | *Arabidopsis thaliana* Salt Overall Sensitivity |
| TtSOS1 | *Triticum turgidum* Salt Overall Sensitivity |
| OsSOS1 | *Oryza sativa* Salt Overall Sensitivity |
| NtSOS1 | *Nicotiana tabacum* Salt Overall Sensitivity |
| StSOS1 | *Solanum tuberosum* Salt Overall Sensitivity |
| AtHKT1 | *Arabidopsis thaliana* High Affinity K+ Transporter |
| GmHKT1 | *Glycine max* High Affinity K+ Transporter |

| | |
|---|---|
| GaHKT1 | *Gossypium arboretum* High Affinity K+ Transporter |
| SlHKT1 | *Solanum lycopersicum* High Affinity K+ Transporter |
| StHKT1 | *Solanum tuberosum* High Affinity K+ Transporter |
| RMSD | Root Mean Square Deviation |
| RMSF | Root Mean Square fluctuation |
| Rubisco | Ribulose 1,5 biphosphate carboxylase/ oxygenase |
| PEPCase | Phosphoenolpyruvate carboxylase |
| NADP-ME | NADP- malic enzyme |

# Abstract

Salinity is one of the environmental stresses that impact growth and development of plants. Sodium chloride is one of the predominant salt responsible for salinity. Beyond the allowable limit in the soil sodium chloride exerts osmotic and ionic stresses. Upon perception of the salinity stress plants undergo different mechanisms such as ion homeostasis, signaling, redox homeostasis and transport of the saline solute in the cell to overcome the salinity. It is managed by the different membrane proteins present in the plant cell that help with sodium distribution, sodium extrusion, sodium compartmentalization, osmotic balance and cellular damage repair. Some of the plant's conserved membrane proteins that help to maintain the plants in salinity stress are Aquaporins, HKT1 (High-Affinity $K^+$ Transporter 1), NHX1 ($Na^+/H^+$ exchangers 1), and SOS1 (Salt Overly Sensitive 1). In this thesis, we have studied two membrane proteins SOS1 and HKT1 protein that help to maintain salinity stress in plants. For the above studies, we have selected a few plants where the role of SOS1 and HKT1 are shown in salt tolerance including the model plant *Arabidopsis thaliana*. We have used different bioinformatics tools and servers to identify the conserved amino acid sequence and design the most appropriate polypeptide markers. The designed polypeptide markers can be used to determine the salinity stress in multiple crops.

# Introduction

Plant stress can be defined as biotic or abiotic stresses that affects by disrupting of proper functioning of the plant’s membrane (Zhang *et al.*,2022; Dresselhaus *et al.*, 2018). Soil salinity, drought and temperature are some examples of abiotic stress whereas pathogens, parasites, allelopathy and weeds are some examples of biotic stress. (Zhu *et al.*, 2016; Roeber *et al.*, 2021).

Salinity stress in plants is a significant issue for crop productivity around the world. These stresses impact the plant productivity in the same way as the abiotic stress. Other than the climate change, degradation of the land and droughts are responsible for increase in soil salinity (Zhu *et al*., 2016). Sometimes, these stresses affect the plants so severely that they get over damaged or die due to stresses. It also creates conductivity and porosity in soils, which causes low water potential in plants. Salinity stress impacts the plants by creating osmotic stress, ion toxicity, oxidative stress thereby affecting photosynthesis and carbon assimilation, stomatal regulation and transpiration, secondary metabolism and hence leading to loss in crop productivity.

As data reported in 2016 shows approx. 60% of the world’s total land area is suitable for cultivation, of which 80% are used for crops and 6% are used for rangelands (Mythili *et al.*, 2016). According to estimates, salinity will impact 50% of the world’s arable land by 2050 (Bartels *et al.*,2005; Kumar *et al.*, 2020). According to data reported in 2017, approximately 2.1% of the India’s total land area or 7 million hectares is impacted by salt of these 3 million hectares are salty and the remaining 4 million hectares are sodic (Arora *et al.*, 2016;). In all developing countries, intensive agricultural practices are responsible for deteriorating agricultural lands. One of the crucial factors for the deduction of plant growth and production is the salt stress (Szabolcs *et al.*, 1974).

Soil salinity has various types depending on the types of salt, such as sodic soils, saline soils, and saline-sodic soils. When the concentration of sodium ions increases as compared to other cations in the soil, it is considered sodic soil. The decrease of water flow in the soil initiates the accumulation of salts (Qadir *et al.*,2007). When water-soluble salts are present in the soil, in amounts that they affect normal crop production, then it is considered saline soil. This type of soil pH is always less than 8.5 (Negacz *et al.*, 2022). If the concentration of both saline and sodic levels is high in soils, it is considered saline-sodic soils. This type of soil is poor in soil structures, has an imbalance in nutrient content, moderately alkalinity in pH and many other characteristics (Qadir *et al.*,2007).

Salinity stress started in crops due to the solubilization of salts in soil. There are some positively and negatively charged ions such as Ammonium ($NH^{4+}$), Phosphate ($PO_4^{3-}$), Potassium($K^+$), Calcium ($Ca^{2+}$), Magnesium ($Mg^{2+}$), Phosphate ($PO_4^{3-}$), Sulfate ($SO_4^{2-}$), Nitrate ($NO^{3-}$), Sodium ($Na^+$) and Chlorine ($Cl^-$). Out of these ion's potassium, sodium, chloride and sulfate are the main ions present in the soil that salinity stress. NaCl containing soil causes various types of stresses in plants which leads to osmotic stress, ion toxicity, root damage, nutrient imbalance, physiological disorders, reduces growth and yields.

***Figure 1:*** Land affected by soil salinity

Salinity effects the plants by osmotic stress and then by phytotoxic ions. Continuous intake of salt from roots creates a decrease in water conductivity in cells. Long exposure of plant roots to this process creates toxicity and nutrient imbalance in plants.

There are two different categories of plants that reveal different tolerability to salinity and also to adaptability, they are glycophytes and halophytes. Halophytes are the ones that can grow in saline soils. These types of plants can be grown in coastal areas. Some of the halophyte plants are mangroves, glasswort, seagrasses and many others. Glycophytes are simply the opposite, meaning they grow in saline soils; they are less tolerant as compared to halophytes. Some of the glycophyte plants are wheat, rice, corn, soybeans and many others.

There are various physiological responses to salinity stress such as ionic imbalance and osmotic stress, reactive oxygen species accumulation in cells, changes in photosynthesis and carbonic metabolism, variations in water uptake and transport. Inside the cells plants maintain the low concentration of $Na^+$ and the high concentration of $K^+$. High concentration of $K^+$ helps the plants for proper growth and development. During the salt stress $Na^+$ levels increase in the cells and create ion imbalance in plants, which disrupts the normal function of plants (Li, J. *et al.*, 2023). This imbalance can be maintained by segregating sodium, compartmentalization of sodium or by distribution of sodium from cells. There are various transmembrane protein that helps to maintain the concentration.

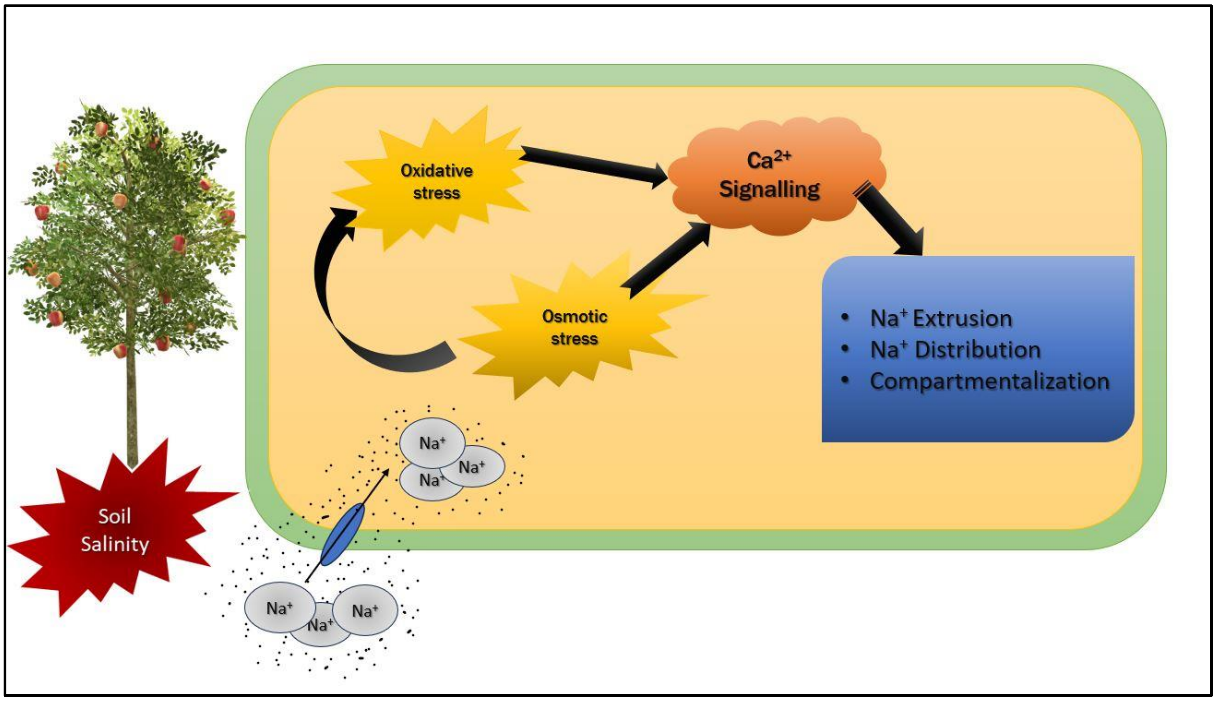


***Figure 2:*** Mechanism to salinity stress in plants

Salinity stress is managed by $Na^+$ distribution, extrusion and compartmentalization, osmotic balance and cellular damage repair by the membrane proteins (Dutta *et al.*, 2023). Some critical plant membrane protein that helps tin the management of salinity stress are HKT1 (High Affinity $K^+$ Transporter 1), NHX1 (Sodium/ Hydrogen Exchanger 1), SOS1 (Salt Overall Sensitivity 1), Aquaporins and various ion channels. Different membrane protein has different function in various membrane parts in plants. During salinity stress, certain membrane protein in plants recognize osmotic stress and activate calcium ions which signal to manage salinity stress and produce reactive oxygen species (ROS) by activating the plasma membrane enzyme RBOH, aquaporins membrane transporter transport reactive oxygen species (ROS) into the cytosol and causing oxidative stress which signals to manage salinity stress (Banik and Dutta, 2023).

During salinity stress in soil, the water uptake in plants decreases. Referring to figure 2, we can understand that concentration of salt creates the salinity in soil. This pressure impacts plants by water absorption and by plant development. Osmotic stress develops due to the variation of water potential in the soil. Reactive oxygen species (ROS) is another physiological response to salinity stress. It is a common byproduct of cells produced at low rates. During salinity stress, it is produced at high rate in the cell (Moller *et al.*, 2007; Miller *et al.*, 2010; Hossain *et al.*, 2016). They are the free radicals like superoxide and hydroxyl radicals as well as non-radical molecules. Excess amounts of reactive oxygen species (ROS) can damage the cells by damaging and by metabolic disorders. It helps to maintain the biotic and abiotic stresses in the plants (Mittler *et al.*, 2004; Hossain *et al.*, 2016).

SOS1 (Salt Overly Sensitive 1) is one of the major membrane proteins associated with the salinity tolerance.  It is localized in the plasma membrane to extrudes the sodium ion from the cell. The SOS1 membrane protein is composed of 13 transmembrane domains. It helps to maintain the cell viability and it is expressed in the parenchyma cells present in the boundary between the xylem and symplast and also in the epidermal cells at the tips of the roots (Zhang *et al.*, 2023). It is responsible for the excretion of sodium ions out of the cell across the plasma membrane. The SOS1 pathway involves the binding and exchange of the $Na^+$/ $H^+$ antiporter. SOS1 is a large integral membrane protein it has multiple transmembrane domains in various locations, different domains have different functions. It has been seen that on overexpression of SOS1 protein can also show the potential activity for salt tolerance in plants. Various SOS membrane protein is also available which shows a similar function to other membrane protein.

HKT1 (High affinity $K^+$ Transporter 1) is also a membrane protein associated with the salinity tolerance. It is also primarily localized to the plasma membrane and acts as a $Na^+$ transporter in the membrane (Horie *et al.*, 2001). It is composed of 4 transmembrane domain-pore loop- transmembrane domain patterns (Hamamoto *et al.*, 2015). They are expressed not only present in xylem parenchyma but also present in roots, shoots, leaves and flowers (Riedelsberger *et al.*, 2021). It functions as a uniporter or symporter for the transport of potassium ions across the plasma membrane. When the concentration of the sodium ions is high and potassium ions is less inside the cell, it helps to uptake the potassium ions from the soils to maintain the physiological characteristics inside the cells.

In this project, we have used the SOS1 and HKT1 membrane protein. We attempt to use the conserved membrane protein transporter sequences, involved in plant salt tolerance, to design polypeptide markers for the identification of salinity stress.

# Review of literature

Salt stress is a major issue in the environment that affects the growth and development of plants. Presence of the sodium concentration affects the plant's growth and developments. Like animals, plants also do not require excess Sodium ions inside the cells. If present in excess amount, it affects the plant in various physiological ways such as oxidative stress, nutrient imbalance, delayed growth and development of death of the plant's cells (Assaha *et al.*, 2017; Cabot *et al.*, 2014).

Restriction of accumulation of the sodium ions and ion homeostasis are the key function of plant tolerance. Roots of the plant gets exposed to saline soil first. Sodium is then absorbed from soil to plants through roots by certain non-selective cation channel (Assaha *et al.*, 2017). Sodium ions are then transported to the shoot through the xylem with the support of different transporters and channels and then to the leaves where they impact the most (Munns *et al.*, 2007). It decreases the agricultural lands for production and increases the demand for foods with the increase in the population (Kumar *et al.*, 2020). Some of the ways by which plant manages the salinity are distribution, exclusion and compartmentalization of the sodium ions (Munns *et al.*, 2008).

Ion homeostasis in plants is defined as how plants respond to the different ions in their cells and tissues, such as potassium, sodium, calcium, magnesium, and chloride. Some of the characteristics of ion homeostasis are the uptake of ions from the soils through different transport mechanisms such as ion channels and transporters and ion partitioning by different cellular compartments to maintain the ion concentration and toxicity inside the cells. Salinity stress affects plants by disrupting the ionic balance inside the cells. When the imbalance occurs in plants due to an excess of $Na^+$ and $Cl^-$ ions, they uptake the other essential ions from the soils to recover the plants from the imbalance. On accumulation of the $Na^+$ and $Cl^-$ ions inside the cells it disrupts the plant cells and creates the osmotic stress. Finally, the plant's metabolism gets affected by hinderance of enzyme activities, protein synthesis and membrane integrity.

For detecting stress some markers are used such as markers, such as physiological markers, morphological markers, oxidative stress markers and molecular stress markers (Soltabayeva *et al.*,2021). Morphological and physiological changes are followed by the orders of visualization of the stress in the plants and then by the physiologically. On the other hand, ROS molecules play a role in signaling stress and the oxidative stress markers changes are detected with the changes in the molecular markers (Ellouzi *et al.*, 2011). Salt stress can be more easily detected in the older plants leaves than in the other parts of the plants by morphologically, physiologically and molecular changes.

During salinity stress, certain ion's transporters in plant's membrane gets activated. These transporters help to control the uptake, transport or compartmentalization of ions in the plant's cells. Some of the ion transporters that plays the critical role in regulation of the salinity stress are SOS1 protein which is known as the salt overly sensitive 1 and localized in the plasma membrane and helps in extrusion of sodium ions from cytoplasm in the exchange of the proton ($H^+$); NHX protein which is known as the $Na^+/ H^+$ exchanger, and localized in the intracellular membrane in the sequestration of sodium ions inside the vacuoles; HKT1 which is also known as the High affinity $K^+$ transporter 1 and localized in the plasma membrane and regulates the sodium ions toxicity; AKT1 which is known as the Arabidopsis potassium ion transporter 1 localized in plasma membrane and maintain the osmotic stress in salinity condition; CNGC which is known as Cyclic Nucleotide Gated Ion Channel and maintain the calcium signaling pathways in cells during salinity stress condition; GORK which is known as the Guard Cell Outward Rectifying potassium channels localized in the plasma membrane in guard cells and helps during water loss in salt stress condition by mediating the potassium ions the cells; CAX protein which is known as $Ca^{2+}/ H^+$ exchanger and transport the calcium ions during the salinity stress conditions.

In this project, we will mainly focus on SOS1 and HKT1 membrane transporter protein which on expression causes drastic changes in the salinity stress in plants.

The SOS (salt overly sensitive 1) system is an important mechanism that can be used by plants to resist salinity. There are three SOS proteins that are generally involved in the SOS pathway SOS1, SOS2 and SOS3 (Figure 3). All three types of protein have different functions in the SOS pathway. SOS1 is activated by the collaborative reaction of the SOS2 and SOS3. SOS3, which is also known as a calcium-binding protein, binds to the stress activated released calcium and activates the SOS2 kinase (Ali *et al.*, 2023). SOS2, which is known as the serine/threonine kinase, creates the complex kinase of the SOS2-SOS3, which combinedly interacts with the SOS1 and activates it to release $Na^+$ from the cells, promoting salt tolerance (Ali *et al.*, 2023).

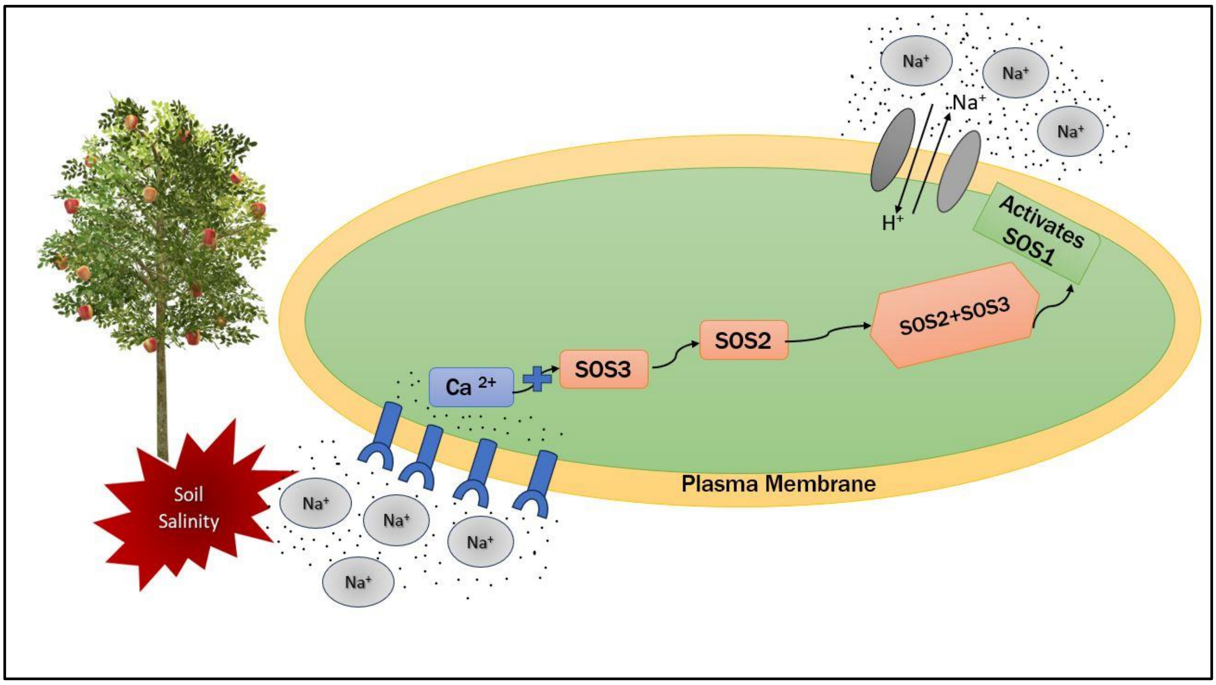


***Figure 3:*** SOS pathway

SOS1 pathways can also be activated by the SCaBP8-SOS2 complex in shoots. SCaBP is known as the SOS3 Like Calcium Binding Protein, also known as the CBL (Calcineurin B-like), which has been identified in the Arabidopsis genome (Laun *et al.*, 2002; Gong *et al.*, 2004; Lin *et al.*, 2009). It has been observed that CBL protein can also control ion homeostasis. The interaction of the CBL1 and CBL9 activates the CIPK 23 (CBL- Interacting Protein Kinase 23). Activation of CIPK23 helps to control the $K^+$ homeostasis by activating the activity of the AKT1, plasma membrane $K^+$ channel.

***Table 1:*** Different SOS1 genes studied in plants

| **Sl.No.** | **Protein Name** | **Organism Name** | **Comments** | **Reference** |
|---|---|---|---|---|
| 1 | *At*SOS1 | *Arabidopsis thaliana* | They have identified the SOS1 gene, which functions by removing sodium from the cell in the exchange of hydrogen ions. They show that a mutation of the SOS1 gene makes the plants more sensitive to salinity stress. | Shi *et al.*,2000 |

| | | | | |
|---|---|---|---|---|
| 2 | *Td*SOS1 | *Triticum turgidum* | On overexpression of *Arabidopsis thaliana* protein complex (SOS2/ SOS3) with *Td*SOS1 it shows the salt tolerance in the drum wheat plant. | Feki *et al.*,2011 |
| 3 | *Os*SOS1 | *Oryza sativa* | They suggested that the efflux of sodium ions in cytosolic was greater in the wild type as compared to the mutant type. | Shahzad *et al.*,2022 |
| 4 | *Gm*SOS1 | *Glycine max* | The SOS1 protein, as an antiporter, has an important role in sodium ion homeostasis by improving salinity in crops. | Zhang *et al.*, 2022 |
| 5 | *Nt*SOS1 | *Nicotiana tabacum* | They have used the overexpressed *At*SOS1 gene in Nicotiana tabacum and compared it with non-transgenic plants and found that transgenic crops can effectively grow in salinity stress. | Yue *et al.*,2011 |
| 6 | *Gh*SOS1 | *Gossypium hirsutum* | They observed that the efflux of sodium ions shows the tolerance to salinity stress. | Cushman *et al.*, 2020 |
| 7 | *So*SOS1 | *Spinacia oleracea* | It demonstrates that the highly expressed of SOS1 in plants shows elution of the sodium ions and its tolerability to salts in plants. | Katschnig *et al.*,2015 |
| 8 | *Sp*SOS1 | *Sesuvium portulacastrum* | They identified *Sp*SOS1 tolerability to salinity in transgenic plants. It was also observed that expression of Arabidopsis thaliana plants in *Sp*SOS1 shows better tolerance to salinity. | Fan *et al.*,2019 |

| | | | | |
|---|---|---|---|---|
| 9 | *St*SOS1 | *Solanum tuberosum* | The position of *St*SOS1 was identified and recognised as a cis-element that helps in the various types of stress. | Liang and Guo, 2023 |
| 10 | *Kc*SOS1 | *Karelinia caspia* | They state that *Kc*SOS1 regulates sodium transport and also regulates the potassium transporter in plants. | Guo *et al.*,2020 |
| 11 | *Pe*SOS1 | *Populus euphratica* | They state that *Pe*SOS1 helps in the control of the $Na^+/H^+$ antiporter in plants. | Wu *et al.*,2016 |
| 12 | *Cq*SOS1 | *Chenopodium quinoa* | They show that *Cq*SOS1 present in plant membranes can help tolerate quinoa in saline environments. | Adolf *et al.*, 2012 |
| 13 | *Mc*SOS1 | *Mesembryanthemum crystallinum* | Expression of $Na^+/H^+$ antiporters response to salinity in plants. It shows the function of antiporter in chloroplasts in plants. | Cosentino *et al.*,2009 |
| 14 | *Nt*SOS1 | *Nitraria tangutorum* | Overexpression of *Nt*CIPK9 from Nitraria tangutorum in Arabidopsis increased seed germination rate under salt stress. | Zhu *et al.*, 2011 |

HKT1 (High affinity $K^+$ Transporter) is a non-specific cation transporter protein primarily located in xylem parenchyma plasma membrane in plants (Horie *et al.*, 2005). It belongs to superfamily of proteins such as HKT/Trk/Ktr, which are present in plants, bacteria and fungi (Corratge *et al.*, 2010). This transporter was first cloned from wheat roots (Schachtman and Schroeder,1994). Heterologous expression of *At*HKT1 was also studied in *Xenopus laevis* oocytes, *Saccharomyces cerevisae* and *Escherichia coli* (Uozumi *et al.*, 2000). HKT1 function by influx and distributes the $Na^+$ into the root cells and prevents them from moving into the shoots. It has been reported that plant roots are getting tolerant to salinity on overexpression and transgenic overexpression of HKT1 (Han *et al.*, 2018). Based on the structure and ions selectivity, the HKT transporter has been divided into two sub-classes with differences in the amino acids of the pore domain of the proteins (Platten J.D. *et al.*, 2006; Ali

*et al.*, 2019). HKT class 1 is considered a symporter as it transports both sodium or potassium ions in the same direction, whereas HKT class 2 is considered an antiporter as it transports both sodium and potassium ions in the opposite direction (Horie T. *et. al.*, 2001; Platten J.D. *et al.*, 2006; Kato Y. *et al.*, 2001; Ali *et al.*, 2019).

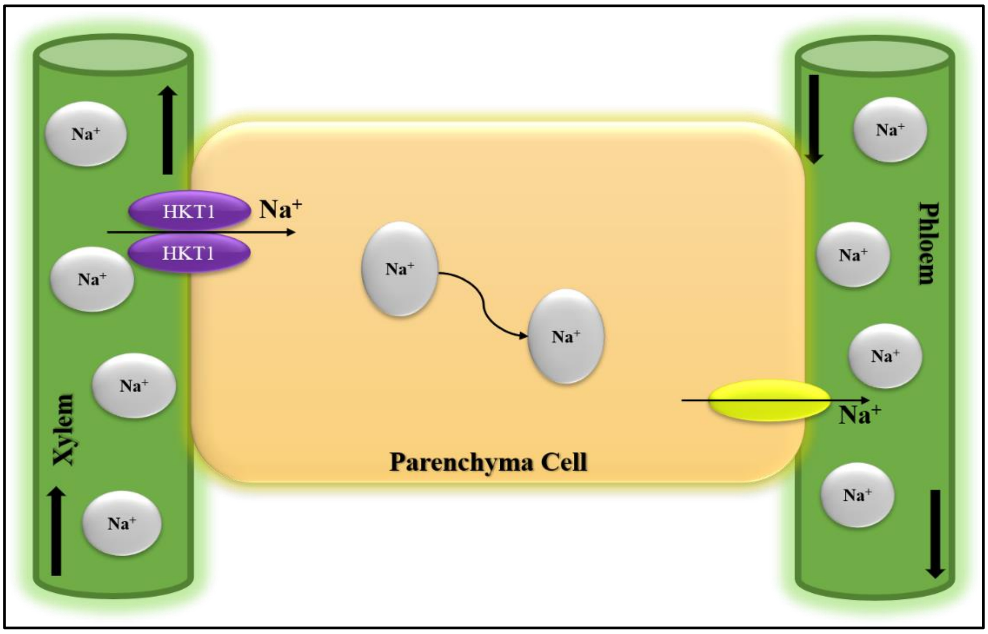


***Figure 4:*** HKT pathway

From the figure 4 we can understand that the $Na^+$ retrieve from xylem sap through shoot into the plant cell with the help of HKT1 and it experiences salinity stress then with the membrane protein $Na^+$ enrich in phloem sap and returns to root.

***Table 2:*** Different HKT1 genes studied in plants

| **Sl.No.** | **Protein Name** | **Organism Name** | **Comments** | **Reference** |
|---|---|---|---|---|
| 1 | *At*HKT1 | *Arabidopsis thaliana* | HKT1 protein efflux the sodium ions from the xylem parenchyma cells and helps to protect them from tissue damage. | Ali *et al.*,2021 |
| 2 | *Bo*HKT1 | *Brassica oleracea* | Expressing the HKT in low potassium ion concentration in a salinity environment shows a | Yang *et al.*, 2023 |

|  |  |  |  |  |
|---|---|---|---|---|
|  |  |  | positive role in maintaining the potassium ion concentration. |  |
| 3 | *Gm*HKT1 | *Glycine max* | It has been observed that on expression of HKT1, it helps to maintain the ionic ratio in roots in saline environment. | Chen *et al.*,2014 |
| 4 | *Ca*HKT1 | *Cicer arietinum* | On the expression of HKT1 could maintain homeostasis, osmolytes and enzymatic antioxidants in roots help them to tolerate the saline environment for the survival of plants. | Kaur *et al.*, 2022 |
| 5 | *Gh*HKT1 | *Gossypium hirsutum* | It has been observed that on overexpression of *Salicornia bigelovii* HKT1 (*Sb*HKT1) in *Gossypium* plant shows increase the uptake capacity of $K^+$ and ion homeostasis in saline environment. | Guo *et al.*, 2020 |
| 6 | *So*HKT1 | *Spinacia oleracea* | It shows the HKT1 helps in the tolerance of the salinity by the movement of the sodium ions. | Katschnig *et al.*, 2015 |
| 7 | *Nt*HKT1 | *Nicotiana tabacum* | It has been observed that on overexpression of the *Arabidopsis* HKT1 gene in root, stem and leaf, it reduced the sodium ion toxicity and make the plant tolerable to salinity. | Wang *et al.*,2018 |
| 8 | *Sl*HKT1 | *Solanum lycopersicum* | They demonstrate the improvement of the transport of the sodium ions in the plants. | Almeida *et al.*, 2013 |
| 9 | *St*HKT1 | *Solanum tuberosum* | Overexpression of the *At*HKT1 in Solanum tuberosum maintains the sodium ion and potassium ion concentration in the leaves. | Wang *et al.*,2019 |

| | | | | |
|---|---|---|---|---|
| 10 | *Hv*HKT1 | *Hordeum vulgare* | On the expression of the HKT gene, it was observed that Hordeum vulgare leaves show sodium or potassium homeostasis and tolerance to salinity. | Fu *et al.*,2018 |
| 11 | *Ca*HKT1 | *Capsicum annuum* | The expression of AtHKT1 controls the sodium ion influx from roots. It is distributed among the shoots of the plants in salinity stress conditions. | Yang *et al.*,2008 |
| 12 | *Ss*HKT1 | *Suaeda salsa* | They show that overexpression of *Ss*HKT1 shows the tolerability to salinity and also increases the concentration of the potassium ions. | Shao *et al.*, 2014 |
| 13 | *Vv*HKT1 | *Vitis vinifera* | It has been observed that the accumulation of salt stress forms chlorotic leaves, but on the expression of HKT1, it regulates the sodium import and retransform to other lower parts of leaves and from roots, it excretes out from plants. | Lu *et al.*,2022 |
| 14 | *Br*HKT1 | *Brassica rapa* | Expressing the HKT in low potassium ion concentration in a salinity environment shows a positive role in maintaining the potassium ion concentration. | Yang *et al.*, 2023 |
| 15 | *Ls*HKT1 | *Lactuca sativa* | It shows that it improves the growth, water holding capacity and ratio of sodium and potassium. | Santander *et al.*,2020 |

## Aim of the study

To study the SOS1 (Salt Overall Sensitivity 1) and HKT1 (High-affinity $K^+$ Transporter 1) protein of plants that is important for plant salt tolerance.

The objective of the study: -

- Literature survey with active role of on SOS1 and HKT1 in plant salinity.
- Designed of SOS1 and HKT1 based polypeptide markers using computational tools.

# Materials and Methods

## 1. SOS1 and HKT1 protein responsible for plant salinity (Literature Survey)

The literature survey was done to identify the important membrane proteins HKT1 and SOS1 which have reported role in salt tolerance in plants. Additionally, identifying different plants species effective with respective proteins we have used the BLAST server. In the BLAST server [https://blast.ncbi.nlm.nih.gov/Blast.cgi], we have used the SOS1 and HKT1 sequences from the model plant *Arabidopsis thaliana* for identifying the respective proteins in different plants. it is to be noted that in *Arabidopsis thaliana* has only one copy of SOS1 and HKT1 is present. The sequence was selected from the server on the basis of percentage identity. The FASTA format sequence was downloaded from NCBI [https://www.ncbi.nlm.nih.gov/] database and UniprotKB database [https://www. uniprot.org/] for the respective proteins in different plants. Refer to Figure- 5.

## 2. Analysis of FASTA sequences

The downloaded FASTA sequences of the various plant species were aligned using multiple sequence alignment tools. Different multiple sequence alignment tools were used such as MAFFT (Multiple Alignment using Fast Fourier Transform) server [https://www.ebi.ac.uk/Tools/msa/mafft/], CLUSTALW server [https://www.genome.jp/tools-bin/clustalw], MUSCLE server [https://www. ebi.ac.uk/jdispatcher/msa/muscle]. We downloaded the result and for easier interpretation of the conserved region, we have used the ESPript 3.0 server [https://espript.ibcp.fr/ESPript/ ESPript/]. The above server provides the conserved region in graphical enhancements. Refer to Figure- 5.

## 3. Identification of highly conserved amino acid

Conserved regions in protein sequences are sections of the protein that remain relatively unchanged or exhibit a high degree of similarity across different species or within a set of related protein sequences. We have used computational tools like MEME (Multiple Em for Motif Elicitation) software [https://meme-suite.org/meme/tools/meme] are used to identify the highly conserved motif. Refer to Figure- 5.

## 4. Identification of the motif predicted position in common plants

We have selected a few plants that show salinity tolerance. Downloaded the FASTA sequences from NCBI or UniProt. The sequences including a selected stable sequence (Motif) were aligned using multiple sequence alignment tools. Different multiple sequence alignment tools were used such as MAFFT (Multiple Alignment using Fast Fourier Transform) server. We downloaded the result and for easier interpretation of the conserved region, we have used the ESPript 3.0 server [https://espript.ibcp.fr/ESPript/ ESPript/]. Refer to Figure- 5.

## 5. Designing the polypeptide sequence based on stability and plotting the secondary structure

From the Meme software we have noted the most conserved amino acid residues and aligned those residues to form motifs. Afterwards, a different combination of motifs was done to get the polypeptide sequence. Computational tools like ExPasy-Protofam [https://robetta.bakerlab. org/] and RaptorX [http://raptorx6.uchicago.edu/] were used to identify the highly stable polypeptide sequences based on their instability index and helix percentage. After the prediction of a stable polypeptide sequence, a secondary structure was designed with the help of the Robetta server [https://robetta.bakerlab.org/] and Pymol software [https://pymol.org/] was used to visualize the predicted structure. In the Robetta server there were five results from which we have selected the less error rate structure. Refer to Figure- 5.

## 6. Prediction the 3D polypeptide structure designed

Designed structure was predicted using the Ramachandran plot [http://vadar.wishartlab. com/]. It shows the potential of the confirmation of the peptide structure. Accordingly, they plot the phi $\varphi$ and the psi $\psi$ angle in X and Y axis respectively. After designing the structure, we validate the sequence by plotting the Ramachandran Plot. They plot an angle spectrum ranging from -180° to +180°. Refer to Figure- 5.

## 7. Molecular Dynamic Simulation

We have validated the results by using the online molecular dynamics simulator servers. For these Molecular dynamics we preferred to use the UAMS Simlab (WEBGRO Macromolecular Simulation). In this server there are two types of tools ‘Protein in Water

Simulation' and 'Protein with Ligand Simulation'. They use the GROMACS simulation for predicting the solvated molecular dynamics simulations. Out of the above tools we preferred to use the 'Protein in Water Simulation' where we have just uploaded the predicted SOS1 and HKT1 structure one at a time. In input parameters, we have used the default parameters where 'Forcefield: GROMOS96 43a1'; 'Water model as SPC'; 'Box Type as triclinic'; 'Salt type as NaCl'. In Energy Minimization parameters, we have used the default parameters where 'Integrator as Steepest Descent'; 'Steps as 5000'. In Equilibration and MD run Parameters, we have used the default parameters where 'Equilibration type as NVT/NPT'; 'Temperature as 300K'; 'Pressure as 1.0bar'; 'MD integrator as Leap frog'; 'Simulation time as 50ns'; 'Approximate number of frames per simulation as 5000' and then submit to run the MD simulator. Refer to Figure- 5.

## 8. Predicting the antigenicity part in the sequence and mapping of the antigenicity on the sequences

Potential antigenicity was analyzed on the predicted sequences. It helps to predict computationally the B-cells and the T-cells in the sequence. Various computational tools were used such as ABCpred, Bcepred, BepiPred and SVMTriP.

ABCpred [https://webs.iiitd.edu.in/raghava/abcpred/] is a computational tool used to predict the B-cell epitopes from the protein sequences. These epitopes are specifically part of the proteins which can bind to the antibodies. This tool uses the artificial neural network to predict the likely part of epitopes in the predicted protein sequence.

Like ABCpred, Bcepred [http://crdd.osdd.net/raghava/bcepred/bcepred_instructions.html] and BepiPred [https://services.healthtech.dtu.dk/services/BepiPred-2.0/] is also a computational tool used for prediction of B-cell epitopes. BepiPred uses the Hidden Markov Models (HMMs) and propensity scale methods to predict the B-cell epitopes from the protein sequences.

SVMTriP [http://sysbio.unl.edu/SVMTriP/prediction.php] is a computational tool used to predict the T-cell epitopes from the protein sequences. These epitopes are specifically a part of the proteins that are identified by T-cells of the immune system. This tool uses the Support Vector Machine algorithm to predict the T-cell epitopes.

After the prediction of the antigenicity by computational methods, the area of the antigenicity was marked in the predicted peptide structure. For marking the antigenicity, the Pymol software [https://pymol.org/] was used. By marking it we can analyze the part which is more suitable for which epitopes. Refer to Figure- 5 depicts the workflow of the methodology used to design polypeptide marker for both SOS1 and HKT1.

## Workflow for the designing of the Polypeptide markers

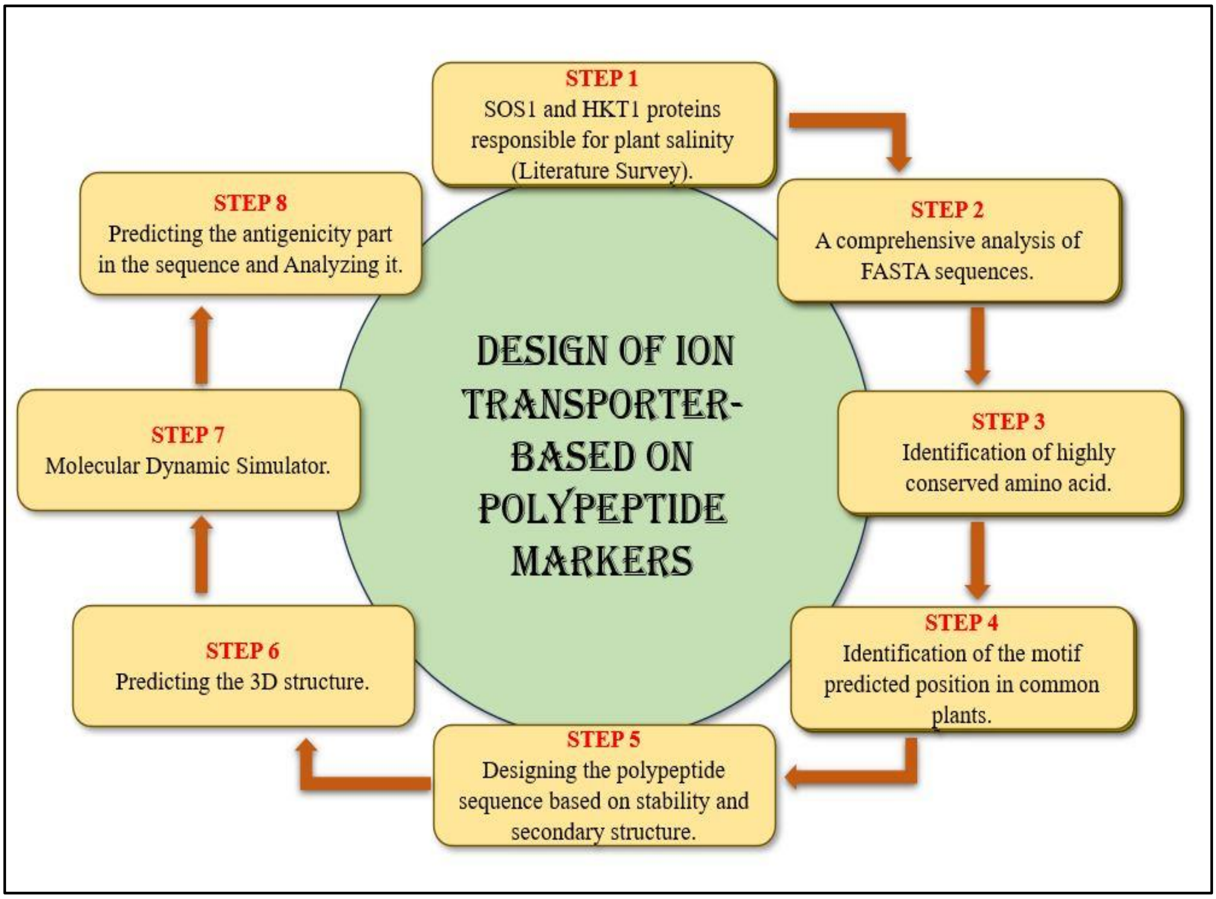


***Figure 5:*** Flowchart to design an ion transporter based on polypeptide markers

# Results

## 1. Analysis and comparison of the SOS1 and HKT1 proteins in crops

We have selected two proteins that are responsible for salinity tolerance SOS1 and HKT1. For both proteins, we have selected a common model plant *Arabidopsis thaliana.* The accession number of *Arabidopsis thaliana* SOS1 is NP_178307.2. The accession number of *Arabidopsis thaliana* HKT1 is OAO98616.1. The sequences of the above crops were used to identify other crops that are homologous to SOS1 and HKT1 proteins. Then, we go through the literature survey so that we can predict the crops with respect to the above proteins that have more tolerance to salinity. Accordingly, we have designed the table below which contains the organism's name, accession number and the homologous crops with respect to the *Arabidopsis thaliana* SOS1 (Table 3) and HKT1 (Table 4).

***Table 3:*** SOS1 proteins from different crops with homology with *Arabidopsis thaliana* SOS1

| **Sl.No.** | **Protein Name** | **Scientific Name** | **Common Name** | **Accession Number** | **Query Cover (%)** | **Percentage Identity (%)** |
|---|---|---|---|---|---|---|
| 1. | *At*SOS1 | *Arabidopsis thaliana* | Thale cress | NP_178307.2 | 100 | 100 |
| 2. | *Tt*SOS1 | *Triticum turgidum* | Durum Wheat | ACB47885.1 | 96 | 60.18 |
| 3. | *Os*SOS1 | *Oryza sativa* | Rice | ATU90113.1 | 97 | 60.31 |
| 4. | *Zm*SOS1 | *Zea mays* | Maize | QOI16621.1 | 97 | 59.54 |
| 5. | *Gh*SOS1 | *Gossypium hirsutum* | Cotton | AKN19929.1 | 96 | 62.14 |
| 6. | *Gm*SOS1 | *Glycine max* | Soybean | NP_001244939.1 | 96 | 64.85 |
| 7. | *Nt*SOS1 | *Nicotiana tabacum* | Tobacco | XP_016466679.1 | 96 | 66.01 |
| 8. | *So*SOS1 | *Spinacia oleracea* | Spinach | CDL70805.1 | 92 | 63.44 |
| 9. | *Vv*SOS1 | *Vitis vinifera* | Grape | CBI26761.3 | 99 | 65.63 |
| 10. | *Sp*SOS1 | *Sesuvium portulacastrum* | Sea purslane | AFX68848.1 | 99 | 62.28 |
| 11. | *St*SOS1 | *Solanum tuberosum* | Potato | XP_006364070.1 | 98 | 63.73 |
| 12. | *Kc*SOS1 | *Karelinia caspia* | Capsicum | QOS02253.1 | 97 | 64.56 |
| 13. | *Pe*SOS1 | *Populus euphratica* | Poplar | AQN76669.1 | 99 | 64.56 |

|  |  |  |  |  |  |  |
|---|---|---|---|---|---|---|
| 14. | *Cq*SOS1 | *Chenopodium quinoa* | Inca wheat | ABS72166.1 | 98 | 62.24 |
| 15. | *Mc*SOS1 | *Mesembryanthemum crystallinum* | Baby sun rose | ABN04858.1 | 97 | 63.85 |
| 16. | *Nt*SOS1 | *Nitraria tangutorum* | Nitre bushes | AGW30210.1 | 99 | 63.87 |
| 17. | *Lg*SOS1 | *Limonium gmelinii* | Statice | ACF05808.1 | 99 | 60.19 |
| 18. | *Sp*SOS1 | *Schrenkiella parvula* | Salt cress | ADQ43186.1 | 99 | 78.71 |

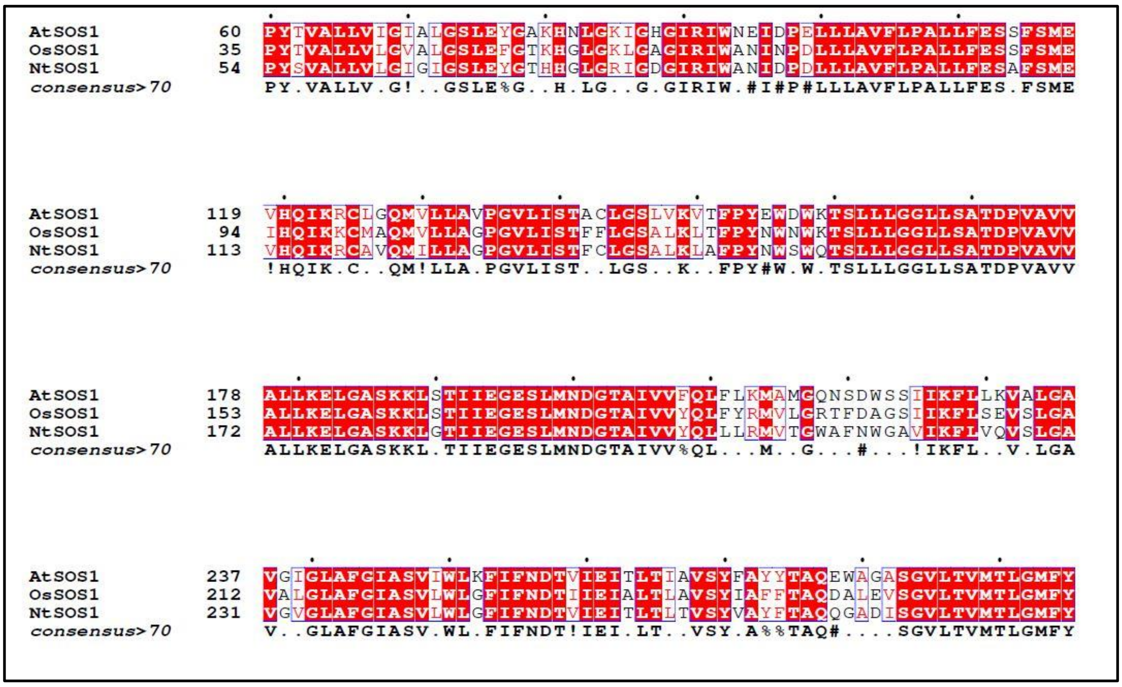
***Table 4:*** HKT1 proteins from different crops with homology with *Arabidopsis thaliana* HKT1

| **Sl.No** | **Protein Name** | **Scientific Name** | **Common Name** | **Accession Number** | **Query Cover (%)** | **Percentage Identity (%)** |
|---|---|---|---|---|---|---|
| 1. | *At*HKT1 | *Arabidopsis thaliana* | Thale cress | OAO98616.1 | 100 | 100 |
| 2. | *Gm*HKT1 | *Glycine max* | Soybean | XP_003517618.3 | 94 | 49.7 |
| 3. | *Cc*HKT1 | *Cajanus cajan* | Pegion pea | XP_020222038.2 | 94 | 51.76 |
| 4. | *Ca*HKT1 | *Cicer arietinum* | Chickpea | XP_004513109.1 | 94 | 48.5 |
| 5. | *Ah*HKT1 | *Arachis hypogaea* | Groundnut | XP_025697586.1 | 94 | 49.11 |
| 6. | *Ga*HKT1 | *Gossypium arboretum* | Cotton | XP_017644471.1 | 94 | 52.95 |
| 7. | *Si*HKT1 | *Sesamum indicum* | Sesame | XP_011094146.1 | 94 | 52.52 |
| 8. | *So*HKT1 | *Spinacia oleracea* | Spinach | XP_021850306.1 | 94 | 48.21 |
| 9. | *Nt*HKT1 | *Nicotiana tabacum* | Tobacco | XP_016459984.1 | 94 | 51.96 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 10. | *Sl*HKT1 | *Solanum lycopersicum* | Tomato | NP_001289833.1 | 94 | 49.8 |
| 11. | *St*HKT1 | *Solanum tuberosum* | Potato | XP_006359731.1 | 94 | 49.7 |
| 12. | *Ha*HKT1 | *Helianthus annuus* | Sunflower | XP_022009341.1 | 93 | 51.13 |
| 13. | *Hv*HKT1 | *Hordeum vulgare* | Barley | KAE8777465.1 | 89 | 47.87 |
| 14. | *Rc*HKT1 | *Ricinus communis* | Castor bean | XP_015576583.1 | 94 | 52.44 |
| 15. | *Ca(pepper)*HKT1 | *Capsicum annuum* | Pepper | XP_016577037.2 | 94 | 49.8 |
| 16. | *Dc*HKT1 | *Daucus carota* | Carrot | XP_017257369.1 | 94 | 48.13 |
| 17. | *Ec*HKT1 | *Eucalyptus camaldulensis* | River red gum | AAF97728.1 | 94 | 51.29 |
| 18. | *Ss*HKT1 | *Suaeda salsa* | Seepweed | AAS20529.2 | 93 | 48.01 |
| 19. | *Vv*HKT1 | *Vitis vinifera* | Grape | XP_010656787.2 | 93 | 47.62 |
| 20. | *Br*HKT1 | *Brassica rapa* | Turnip mustard | XP_009128719.1 | 99 | 77.17 |

## 2. Analysis of FASTA sequences

Multiple sequence alignments were done of the plants that were selected from the SOS1 and HKT1 sequences. It was done with the help of MAFFT, CLUSTALW, and MUSCLE tools respectively. The aligned sequences were interpreted with ESPript 3.0 tool as illustrated below for SOS1 refer to Figure 6 and for HKT1 refer to Figure 7.



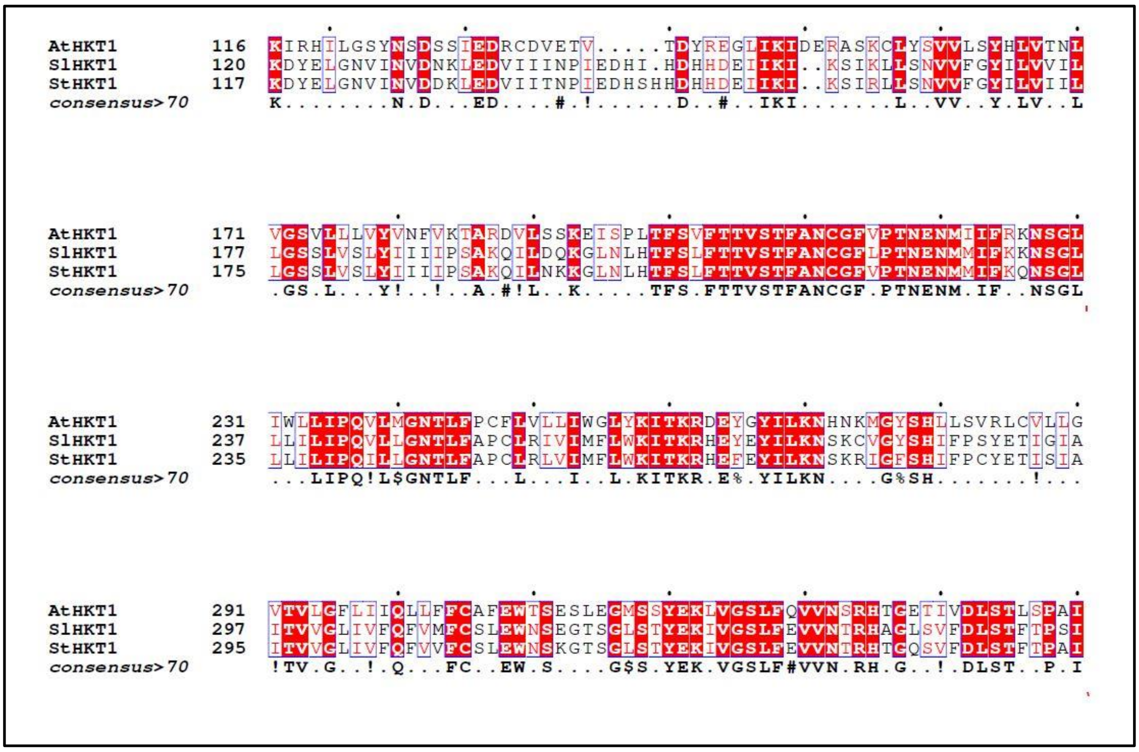
```
AtSOS1           60  PYTVALLVIGIALGSLEYGAKHNLGKIGHGIRIWNEIDPELLLAVFLPALLFESSFSME
OsSOS1           35  PYTVALLVLGVALGSLEFGTKHGLGKLGAGIRIWANINPDLLLAVFLPALLFESSFSME
NtSOS1           54  PYSVALLVLGIGIGSLEYGTHHGLGRIGDGIRIWANIDPDLLLAVFLPALLFESAFSME
consensus>70         PY.VALLV.G!..GSLE%G..H.LG..G.GIRIW.#I#P#LLLAVFLPALLFES.FSME

AtSOS1          119  VHQIKRCLGQMVLLAVPGVLISTACLGSLVKVTFPYEWDWKTSLLLGGLLSATDPVAVV
OsSOS1           94  IHQIKKCMAQMVLLAGPGVLISTFFLGSALKLTFPYNWNWKTSLLLGGLLSATDPVAVV
NtSOS1          113  VHQIKRCAVQMILLAGPGVLISTFCLGSALKLAFPYNWSWQTSLLLGGLLSATDPVAVV
consensus>70         !HQIK.C..QM!LLA.PGVLIST..LGS..K..FPY#W.W.TSLLLGGLLSATDPVAVV

AtSOS1          178  ALLKELGASKKLSTIIEGESLMNDGTAIVVFQLFLKMAMGQNSDWSSIIKFLLKVALGA
OsSOS1          153  ALLKELGASKKLSTIIEGESLMNDGTAIVVYQLFYRMVLGRTFDAGSIIKFLSEVSLGA
NtSOS1          172  ALLKELGASKKLGTIIEGESLMNDGTAIVVYQLLLRMVTGWAFNWGAVIKFLVQVSLGA
consensus>70         ALLKELGASKKL.TIIEGESLMNDGTAIVV%QL...M..G...#...!IKFL..V.LGA

AtSOS1          237  VGIGLAFGIASVIWLKFIFNDTVIEITLTIAVSYFAYYTAQEWAGASGVLTVMTLGMFY
OsSOS1          212  VALGLAFGIASVLWLGFIFNDTIIEIALTLAVSYIAFFTAQDALEVSGVLTVMTLGMFY
NtSOS1          231  VGVGLAFGIASVLWLGFIFNDTVIEITLTLTVSYVAYFTAQQGADISGVLTVMTLGMFY
consensus>70         V..GLAFGIASV.WL.FIFNDT!IEI.LT..VSY.A%%TAQ#....SGVLTVMTLGMFY
```

*Figure 6:* Multiple Sequence Alignment for SOS1 Protein



```
AtHKT1          116  KIRHILGSYNSDSSIEDRCDVETV.....TDYREGLIKIDERASKCLYSVVLSYHLVTNL
SlHKT1          120  KDYELGNVINVDNKLEDVIIINPIEDHI.HDHHDEIIKI..KSIKLLSNVVFGYILVVIL
StHKT1          117  KDYELGNVINVDDKLEDVIITNPIEDHSHHDHHDEIIKI..KSIRLLSNVVFGYILVIIL
consensus>70         K........N.D...ED....#.!......D..#..IKI.......L..VV..Y.LV..L

AtHKT1          171  VGSVLLLVYVNFVKTARDVLSSKEISPLTFSVFTTVSTFANCGFVPTNENMIIFRKNSGL
SlHKT1          177  LGSSLVSLYIIIIPSAKQILDQKGLNLHTFSLFTTVSTFANCGFLPTNENMMIFKKNSGL
StHKT1          175  LGSSLVSLYIIIIPSAKQILNKKGLNLHTFSLFTTVSTFANCGFVPTNENMMIFKQNSGL
consensus>70         .GS.L...Y!..!..A.#!L..K.....TFS.FTTVSTFANCGF.PTNENM.IF..NSGL

AtHKT1          231  IWLLIPQVLMGNTLFPCFLVLLIWGLYKITKRDEYGYILKNHNKMGYSHLLSVRLCVLLG
SlHKT1          237  LLILIPQVLLGNTLFAPCLRIVIMFLWKITKRHEYEYILKNSKCVGYSHIFPSYETIGIA
StHKT1          235  LLILIPQILLGNTLFAPCLRLVIMFLWKITKRHEFEYILKNSKRIGFSHIFPCYETISIA
consensus>70         ...LIPQ!L$GNTLF...L...I..L.KITKR.E%.YILKN....G%SH.......!...

AtHKT1          291  VTVLGFLIIQLLFFCAFEWTSESLEGMSSYEKLVGSLFQVVNSRHTGETIVDLSTLSPAI
SlHKT1          297  ITVVGLIVFQFVMFCSLEWNSEGTSGLSTYEKIVGSLFEVVNTRHAGLSVFDLSTFTPSI
StHKT1          295  ITVVGLIVFQFVVFCSLEWNSKGTSGLSTYEKIVGSLFEVVNTRHTGQSVFDLSTFTPAI
consensus>70         !TV.G..!.Q...FC..EW.S....G$S.YEK.VGSLF#VVN.RH.G..!.DLST..P.I
```

*Figure 7:* Multiple Sequence Alignment for HKT1 Protein

## 3. Identification of highly conserved amino acid

After running multiple sequence alignments using MAFFT software, we have identified the conserved motif for SOS1 and HKT1 sequences from the alignment using the MEME server. It designated five motifs as Motif 1, Motif 2, Motif 3, Motif 4, Motif 5. The motifs of the conserved amino acid were identified on the basis of the E-value. For SOS1 refer to Figure 8 and for HKT1 refer to Figure 9.

***Figure 8:*** Conserved motif for SOS1

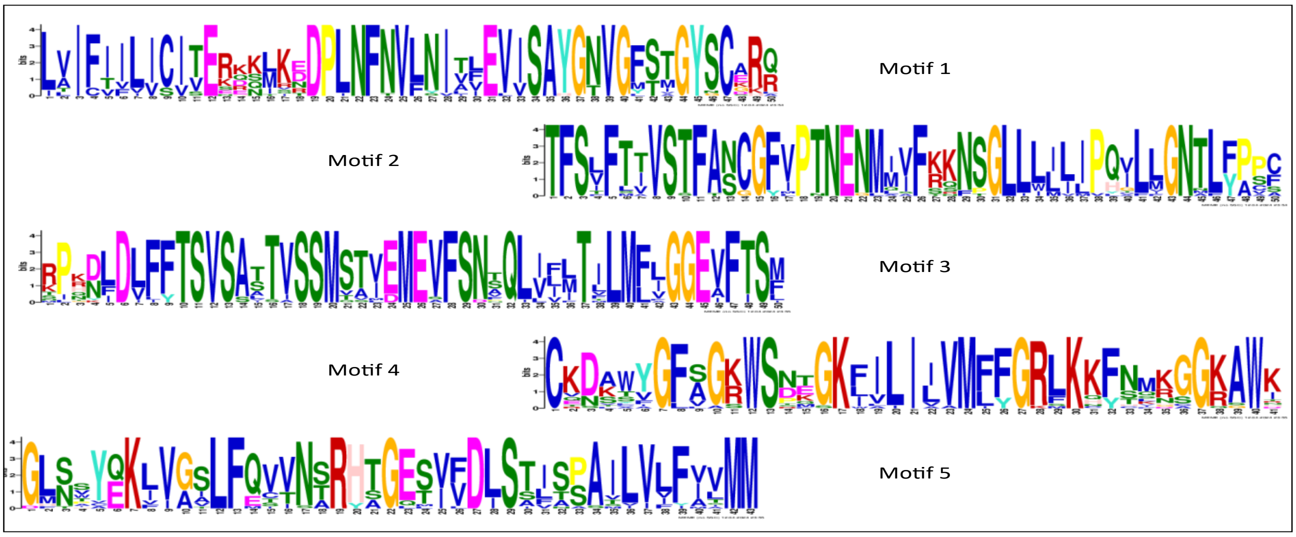
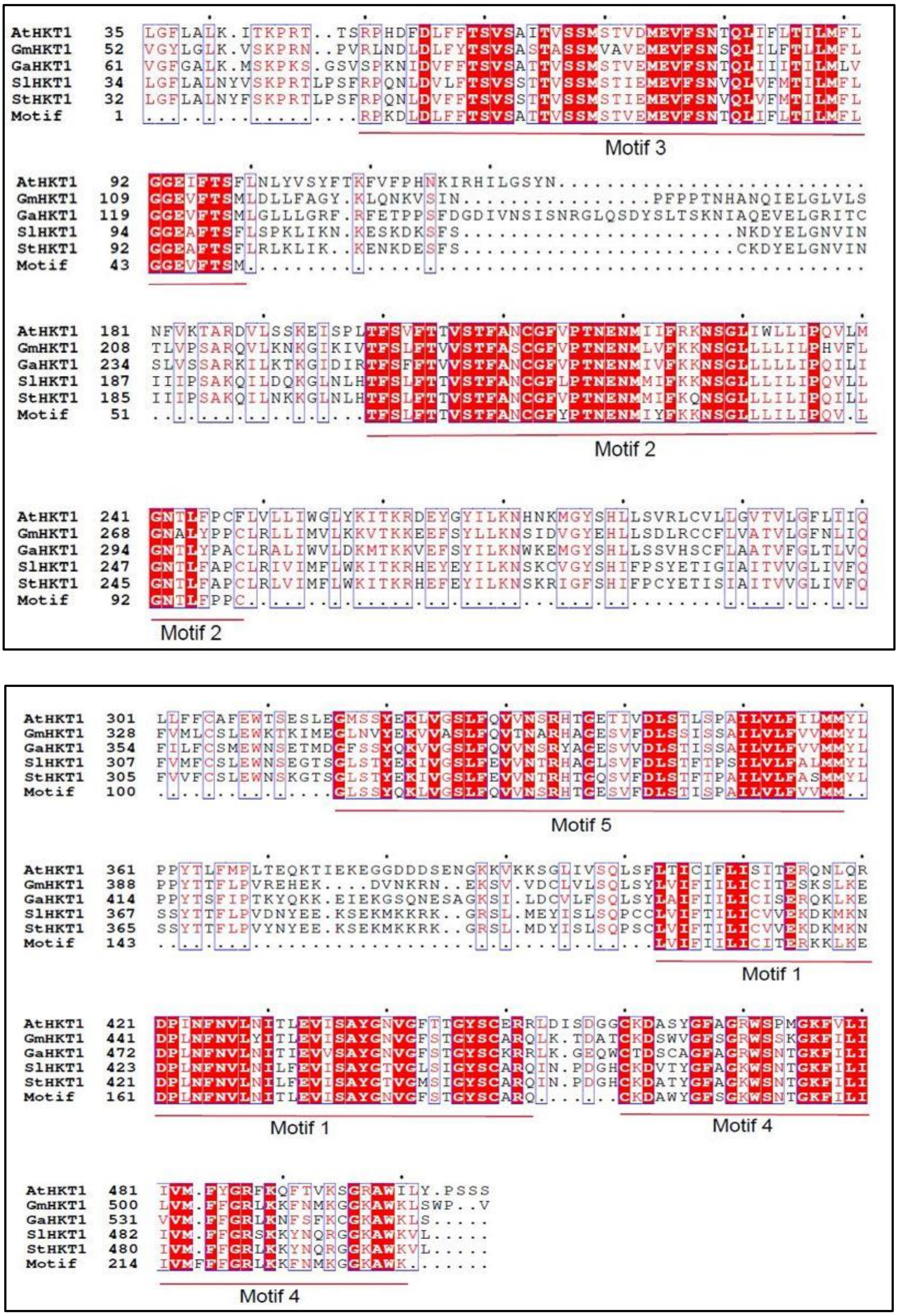


***Figure 9:*** Conserved motif for HKT1

## 4. Identification of the motif predicted position in common plants

Multiple sequence alignments were done on the plants that were selected from the SOS1 and HKT1 sequences with selected stable sequences (Motif). It was done with the help of MAFFT. The aligned sequences were interpreted in ESPript 3.0 tools. For SOS1 refer to Figure 10 and for HKT1 refer to Figure 11.

*Figure 10:* Alignment with Selected sequence SOS1



```
AtHKT1   35  LGFLALK.ITKPRT..TSRPHDFDLFFTSVSAITVSSMSTVDMEVFSNTQLIFLTILMFL
GmHKT1   52  VGYLGLK.VSKPRN..PVRLNDLDLFYTSVSASTASSMVAVEMEVFSNSQLILFTLLMFL
GaHKT1   61  VGFGALK.MSKPKS.GSVSPKNIDVFFTSVSATTVSSMSTVEMEVFSNTQLIIITILMLV
SlHKT1   34  LGFLALNYVSKPRTLPSFRPQNLDVLFTSVSSTTVSSMSTIEMEVFSNVQLVFMTILMFL
StHKT1   32  LGFLALNYFSKPRTLPSFRPQNLDVFFTSVSSTTVSSMSTIEMEVFSNVQLVFMTILMFL
Motif     1  ..................RPKDLDLFFTSVSATTVSSMSTVEMEVFSNTQLIFLTILMFL

AtHKT1   92  GGEIFTSFLNLYVSYFTKFVFPHNKIRHILGSYN..........................
GmHKT1  109  GGEVFTSMLDLLFAGY.KLQNKVSIN................PFPPTNHANQIELGLVLS
GaHKT1  119  GGEVFTSMLGLLLGRF.RFETPPSFDGDIVNSISNRGLQSDYSLTSKNIAQEVELGRITC
SlHKT1   94  GGEAFTSFLSPKLIKN.KESKDKSFS.......................NKDYELGNVIN
StHKT1   92  GGEAFTSFLRLKLIK..KENKDESFS.......................CKDYELGNVIN
Motif    43  GGEVFTSM....................................................

AtHKT1  181  NFVKTARDVLSSKEISPLTFSVFTTVSTFANCGFVPTNENMIIFRKNSGLIWLLIPQVLM
GmHKT1  208  TLVPSARQVLKNKGIKIVTFSLFTVVSTFASCGFVPTNENMLVFKKNSGLLLLILPHVFL
GaHKT1  234  SLVSSARKILKTKGIDIRTFSFFTVVSTFANCGFVPTNENMIVFKKNSGLLLLLIPQILI
SlHKT1  187  IIIPSAKQILDQKGLNLHTFSLFTTVSTFANCGFLPTNENMMIFKKNSGLLLILIPQVLL
StHKT1  185  IIIPSAKQILNKKGLNLHTFSLFTTVSTFANCGFVPTNENMMIFKQNSGLLLILIPQILL
Motif    51  ..................TFSLFTTVSTFANCGFYPTNENMIYFKKNSGLLLILIPQV.L

AtHKT1  241  GNTLFPCFLVLLIWGLYKITKRDEYGYILKNHNKMGYSHLLSVRLCVLLGVTVLGFLIIQ
GmHKT1  268  GNALYPPCLRLLIMVLKKVTKKEEFSYLLKNSIDVGYEHLLSDLRCCFLVATVLGFNLIQ
GaHKT1  294  GNTLYPACLRALIWVLDKMTKKVEFSYILKNWKEMGYSHLLSSVHSCFLAATVFGLTLVQ
SlHKT1  247  GNTLFAPCLRIVIMFLWKITKRHEYEYILKNSKCVGYSHIFPSYETIGIAITVVGLIVFQ
StHKT1  245  GNTLFAPCLRLVIMFLWKITKRHEFEYILKNSKRIGFSHIFPCYETISIAITVVGLIVFQ
Motif    92  GNTLFPPC....................................................
```

```
AtHKT1  301  LLFFCAFEWTSESLEGMSSYEKLVGSLFQVVNSRHTGETIVDLSTLSPAILVLFILMMYL
GmHKT1  328  FVMLCSLEWKTKIMEGLNVYEKVVASLFQVTNARHAGESVFDLSSISSAILVLFVVMMYL
GaHKT1  354  FILFCSMEWNSETMDGFSSYQKVVGSLFQVVNSRYAGESVVDLSTISSAILVLFVVMMYL
SlHKT1  307  FVMFCSLEWNSEGTSGLSTYEKIVGSLFEVVNTRHAGLSVFDLSTFTPSILVLFALMMYL
StHKT1  305  FVVFCSLEWNSKGTSGLSTYEKIVGSLFEVVNTRHTGQSVFDLSTFTPAILVLFASMMYL
Motif   100  ...............GLSSYQKLVGSLFQVVNSRHTGESVFDLSTISPAILVLFVVMM..

AtHKT1  361  PPYTLFMPLTEQKTIEKEGGDDDSENGKKVKKSGLIVSQLSFLTICIFLISITERQNLQR
GmHKT1  388  PPYTTFLPVREHEK....DVNKRN..EKSV.VDCLVLSQLSYLVIFIILICITESKSLKE
GaHKT1  414  PPYTSFIPTKYQKK.EIEKGSQNESAGKSI.LDCVLFSQLSYLAIFIILICISERQKLKE
SlHKT1  367  SSYTTFLPVDNYEE.KSEKMKKRK..GRSL.MEYISLSQPCCLVIFTILICVVEKDKMKN
StHKT1  365  SSYTTFLPVYNYEE.KSEKMKKRK..GRSL.MDYISLSQPSCLVIFTILICVVEKDKMKN
Motif   143  ..........................................LVIFIILICITERKKLKE

AtHKT1  421  DPINFNVLNITLEVISAYGNVGFTTGYSCERRLDISDGGCKDASYGFAGRWSPMGKFVLI
GmHKT1  441  DPLNFNVLYITLEVISAYGNVGFSTGYSCARQLK.TDATCKDSWVGFSGRWSSKGKFILI
GaHKT1  472  DPLNFNVLNITIEVVSAYGNVGFSTGYSCKRRLK.GEQWCTDSCAGFAGRWSNTGKFILI
SlHKT1  423  DPLNFNVLNILFEVISAYGTVGLSIGYSCARQIN.PDGHCKDVTYGFAGKWSNTGKFILI
StHKT1  421  DPLNFNVLNILFEVISAYGTVGMSIGYSCARQIN.PDGHCKDATYGFAGKWSNTGKFILI
Motif   161  DPLNFNVLNITLEVISAYGNVGFSTGYSCARQ.......CKDAWYGFSGKWSNTGKFILI

AtHKT1  481  IVM.FYGRFKQFTVKSGRAWILY.PSSS
GmHKT1  500  LVM.FFGRLKKFNMKGGKAWKLSWP..V
GaHKT1  531  VVM.FFGRLKNFSFKCGKAWKLS.....
SlHKT1  482  IVM.FFGRSKKYNQRGGKAWKVL.....
StHKT1  480  IVM.FFGRLKKYNQRGGKAWKVL.....
Motif   214  IVMFFFGRLKKFNMKGGKAWK.......
```

***Figure 11:*** Alignment with Selected sequence HKT1

## 5. Designing of the polypeptide sequence on the basis of the instability index and the helicity and secondary structure plotted.

For designing the SOS1 and HKT1 sequences, from the motifs the conserved amino acids were identified. After identification, the amino acid was arranged to form one sequence for SOS1 and one sequence for HKT1. The various combination of sequence was arranged and stability was checked on the basis of the instability index and the helicity. By instability index, we can predict the stability of the peptide sequence. It is calculated on the basis of the peptide composition. During prediction if the instability index is less 40 then we can estimate that the peptide sequence is stable or if the instability index is greater than 40 then we can estimate that the sequence form is unstable. Around 120 combinations of sequences were verified from the above conserved amino acid sequences. On successful verification we were able to establish the 5 stable conserved amino acid sequences for SOS1 (Table 5) and HKT1 (Table 6). Later, the secondary structure for SOS1 and HKT1 sequences was plotted on the Robetta server, which had the more stability [Refer to ANNEXURE-I SOS1 (Table 10), (Figure 27); HKT1 (Table 11), (Figure 27)]. In this server, 5 models were formed from the above most stable sequence on the basis of the energy, stability and the error rate. Out of which we have selected the more stable and less error rate-based models.

***Table 5:*** Instability index and Helix Percentage of SOS1.

| **SL.No.** | **Motif's Sequence's** | **Instability Index** | **Helix Percentage (%)** |
|---|---|---|---|
| 1 | A1, A2, A3, A5, A4 | 35.35 | 80 |
| 2 | A1, A4, A5, A3, A2 | 37.09 | 77 |
| 3 | A2, A3, A4, A5, A1 | 37.58 | 77 |
| 4 | A4, A1, A2, A3, A5 | 37.79 | 75 |
| 5 | A4, A2, A3, A1, A5 | 37.79 | 75 |

#Here **A1=** FPYDWNWKTSLLLGGLLSATDPVAVVALLKELGASKKLSTIIE GESLMND;
**A2=** VMTLGMFYAAFARTAFKGESQQSLHHFWEMVAYIANTLIF ILSGVVIAEG;
**A3=** ANIDPDLLLAVFLPALLFESSFSMEVHQIKRCMAQMV LLA GPGV;
**A4=** PRKLVTYFTVERLESACYICAAFLRAHRIARRQLHDFIGDS;
**A5=** VQAAYWEMLDEG RITQTTANILMQSVDEALDLVS.

*Table 6:* Instability index and Helix Percentage of HKT 1

| SL.No. | Motif's Sequence's | Instability Index | Helix Percentage (%) |
|---|---|---|---|
| 1 | A1, A4, A5, A3, A2 | 22.69 | 57 |
| 2 | A1, A4, A5, A2, A3 | 23.01 | 61 |
| 3 | A4, A1, A5, A3, A2 | 23.01 | 57 |
| 4 | A3, A1, A4, A5, A2 | 23.01 | 60 |
| 5 | A5, A3, A2, A4, A1 | 23.01 | 62 |

#Here **A1=** LVIFIILICITERKKLKEDPLNFNVLNITLEVISAYGNVGFSTGYSCA RQ;
**A2=** TFSLFTTVSTFANCGFYPTNENMIYFKKNSGL LLILIPQVLGNTLFPP C;
**A3=** RPKDLDLFFTSVSATTVSSMSTVEMEVFSNTQLIFLTILMFLGGEVFTS M;
**A4=** CKDAWYGFSGKWSNTGKFILIIVMFF FGRLKKFNMKGGKAWK;
**A5=** GLSS YQKLVGSLFQVVNSRHTGES VFDLSTISPAILVLFVVMM

## 6. Prediction of the 3D polypeptide structure designed

After the 3D models were created from the above selected conserved amino acid the potential of the confirmation of the peptide structure was done with the help of the Vadar server. This server mainly confirms the structure potential by plotting the "Ramachandran Plot" (Figure-12) on which it plots the phi (φ) and the psi (ψ) angle in X and Y axis respectively. On the basis of that residue in phi (φ) psi (ψ) core, residue in phi (φ) psi (ψ) allowed, residue in phi (φ) psi (ψ) generous and residue in phi (φ) psi (ψ) outside were also observed for both SOS1 (Table 7) and HKT1 (Table 8) protein.

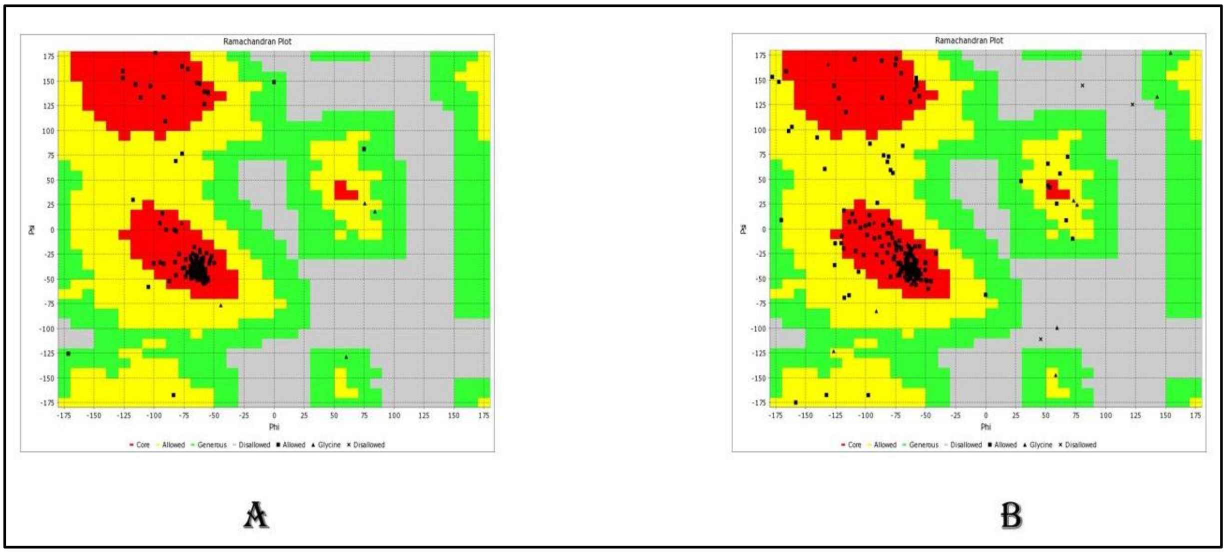


***Figure 12:*** Ramachandran plot of secondary Structure; **A.** SOS1; **B.** HKT1

***Table 7:*** Different Ramachandran plot statistics for residue in phi-psi in SOS1 structure

| Sl.No. | Ramachandran plot statistics | Observed | Expected |
|---|---|---|---|
| 1 | Residue in phi (φ) psi (ψ) core | 212 (96%) | 197 (90%) |
| 2 | Residue in phi (φ) psi (ψ) allowed | 5 (2%) | 15 (7%) |
| 3 | Residue in phi (φ) psi (ψ) generous | 2 (0%) | 2 (1%) |
| 4 | Residue in phi (φ) psi (ψ) outside | 0 (0%) | 0 (0%) |

***Table 8:*** Different Ramachandran plot statistics for residue in phi-psi in HKT1 structure

| Sl.No. | Ramachandran plot statistics | Observed | Expected |
|---|---|---|---|
| 1 | Residue in phi (φ) psi (ψ) core | 211 (90%) | 211(90%) |
| 2 | Residue in phi (φ) psi (ψ) allowed | 16(6%) | 16(7%) |
| 3 | Residue in phi (φ) psi (ψ) generous | 4(1%) | 2(1%) |
| 4 | Residue in phi (φ) psi (ψ) outside | 3(1%) | 0(0%) |

## 7. Molecular Dynamic Simulation

After predicting the SOS1 and HKT1 membrane protein structure and then predicting the 3D structure of the protein using the Ramachandran plot, we were finally able to establish a suitable structure. For further validation, we have done the molecular dynamics simulation in which we focus on the different parameters of a graph such as radius of gyration vs time, solvent accessible vs time, RMSD value vs time and RMS fluctuation vs residues. For SOS1 refer to figure 13,14; [ANNEXURE-I (Figure 28)] and for HKT1 refer to figure 15,16); [ANNEXURE-I (Figure 29)].

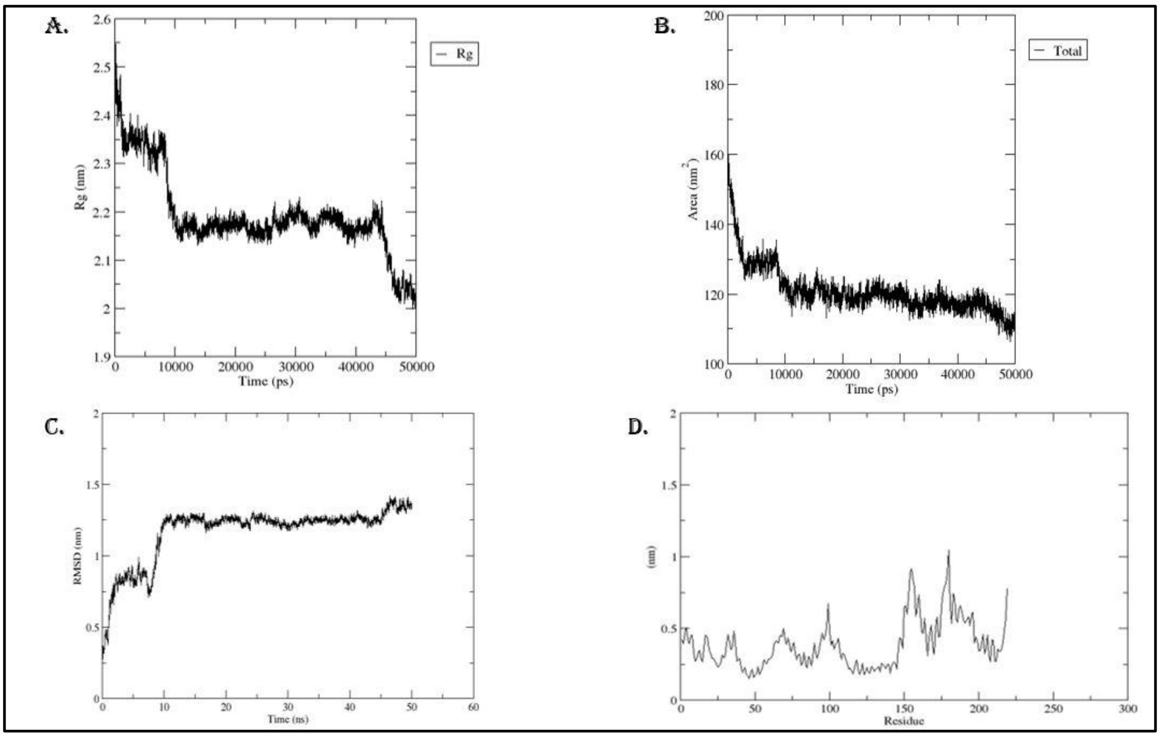


***Figure 13:*** Molecular dynamic simulation of SOS1 sequence **A.** Rg vs Time [Radius of gyration]; **B.** Area vs Time [Solvent Accessible surface]; **C.** RMSD vs Time [RMSD]; **D.** RMS fluctuation.

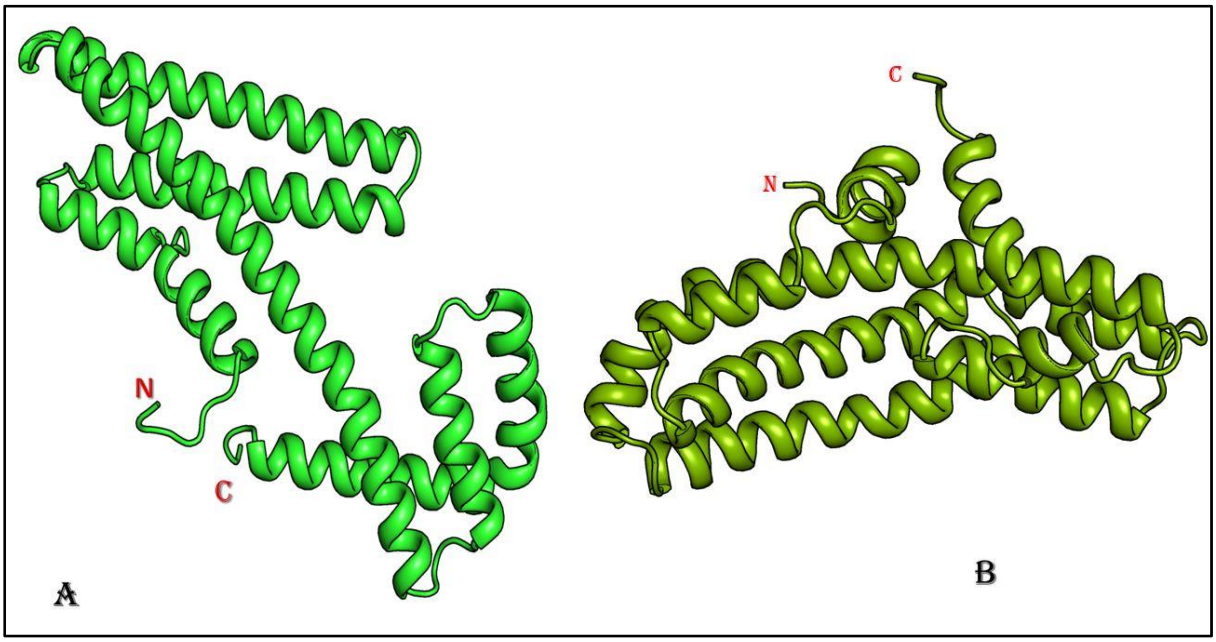


***Figure 14:*** Comparison of SOS1 peptide structure with **A.** Before Simulation; **B.** After Simulation

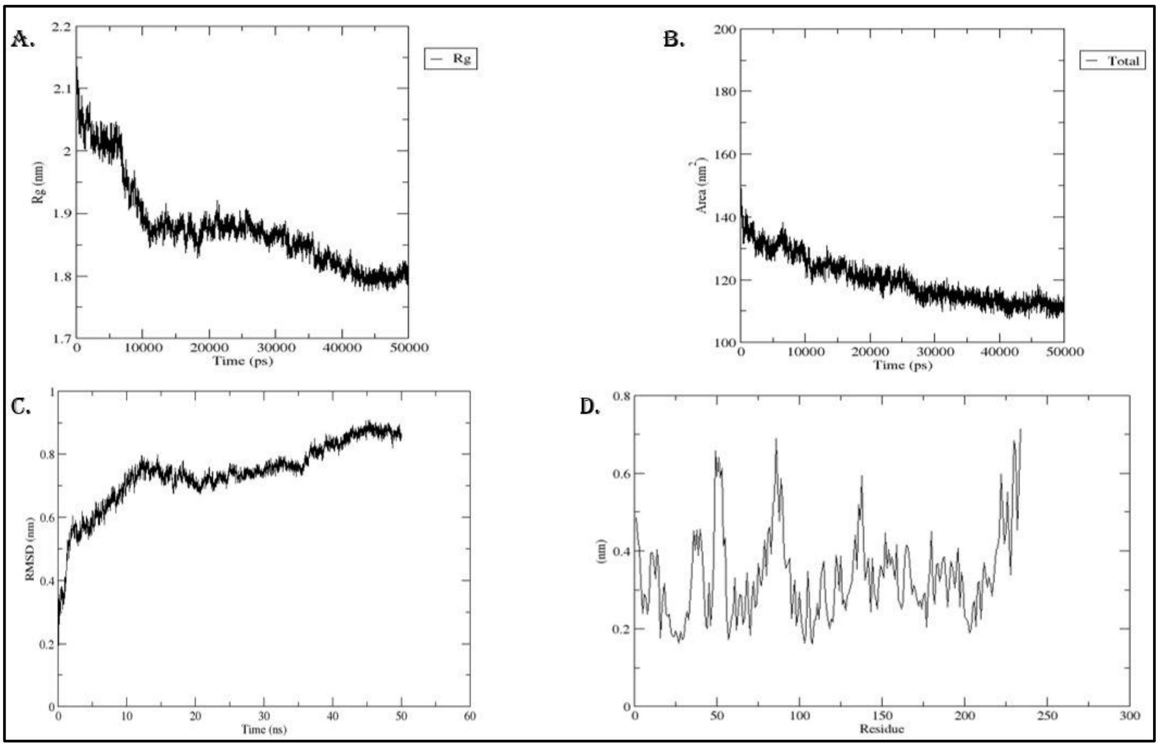


***Figure 15:*** Molecular dynamic simulation of HKT1 sequence **A.** Rg vs Time [Radius of gyration]; **B.** Area vs Time [Solvent Accessible surface]; **C.** RMSD vs Time [RMSD]; **D.** RMS fluctuation.

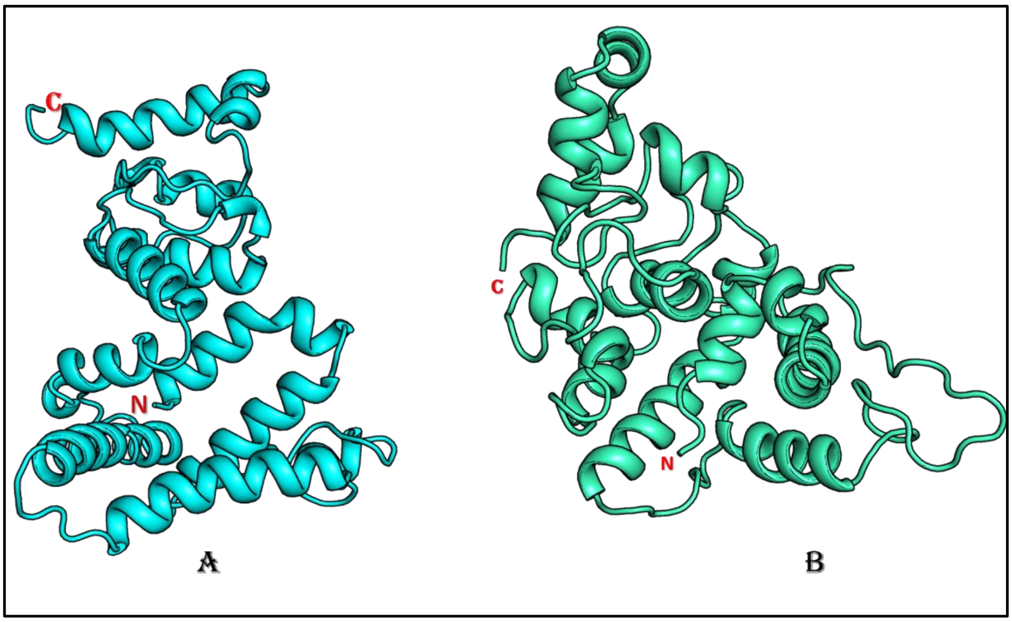


***Figure 16*:** Comparison of HKT1 peptide structure with **A.** Before Simulation; **B.** After Simulation

## 8. Predicting the antigenicity part in the sequence and mapping of the antigenicity on the sequences

Predicted sequences of SOS1 and HKT1 proteins antigenicity were analysed using the servers ABCpred, Bcepred, BepiPred and SVMTriP. Different servers were used to improve the accuracy of the antigenic parameter. On this server, we have analyzed the results on the basis of the threshold value 0.5 to predict the more proper binders. It was also observed in Bepi pred server that in SOS1 and HKT1 predicted sequences mostly had overlapping which specifies its potential antigenicity.

For SOS1, we have selected the part of the sequence on the basis of the score such as in SVMTrip (Table 9) (Figure 17,18); [ANNEXURE-I (Table 12); (Figure 29)]; we have preferred the threshold value of 1.0, in ABCpred (Figure 17,18); [ANNEXURE-I (Table 14); (Figure 30)]; we prefer the highest threshold value and in BepiPred we prefer the threshold value greater than 0.5 (Table 9); (Figure17,18), [ANNEXURE-I (Figure 31)].

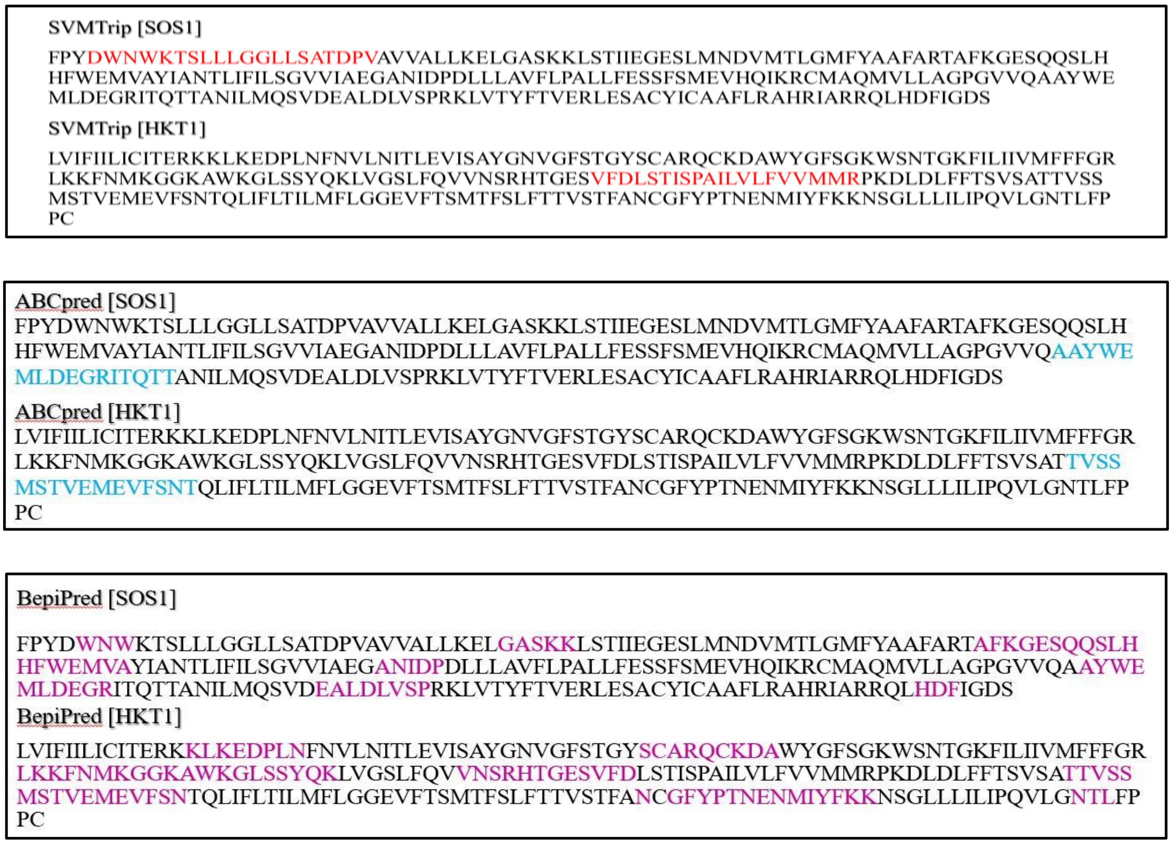
Similarly, for HKT1, we have selected the part of the sequence on the basis of the score such as in SVMTrip (Table 9) (Figure 17,18); [ANNEXURE-I (Table 13); (Figure 29)], we have preferred the threshold value of 1.0, in ABCpred (Table 9) (Figure 17,18); [ANNEXURE-I (Table 15); (Figure 30)] we prefer the highest threshold value and in BepiPred we prefer the threshold value greater than 0.5 (Table 9); (Figure17,18), [ANNEXURE-I (Figure 31)].

We have also used the BcePred in which we have predicted the B-Cell epitope on the basis of physiochemical properties such as hydrophilicity, flexibility/mobility, accessibility, polarity, exposed surface and turns. The tables are attached in ANNEXURE-I (Table 16,17).

After plotting of the secondary structure and prediction of 3D structure in the Ramachandran plot, we started to map the antigenicity part of the structure to find the most capable binding site. This part was analysed using a different server such as SVMTrip. ABCpred and BepiPred.

***Table 9:*** Selected sequence of SOS1 and HKT1 from server on the basis of score

| **Sl.No.** | **Predicting Server** | **Protein Name** | **Selected sequence** | **Score** |
|---|---|---|---|---|
| 1 | SVMTrip | SOS1 | DWNNWKTSLLLGGLLSATDPV | 1.0 |
| | | HKT1 | VFDLSTISPAILVLFVVMMR | 1.0 |
| 2 | ABCpred | SOS1 | AAYWEMLDEGRITQTT | 0.84 |
| | | HKT1 | TVSSMSTVEMEVFSNT | 0.91 |
| 3 | BepiPred | SOS1 | WNW/ GASKK/ AFKGESQQSLHHFWEMVA/ ANIDP/ AYWEMLDEGR/ EALDLVSP/ HDF | Greater than 0.5 |
| | | HKT1 | KLKEDPLN/ SCARQCKDA/ LKKFNMKGGKAWKGLSSYQK/ VNSRHTGESVFD/ TTVSSMSTVEMEVFSN/ N-GFYPTNENMIYFKK/ NTL | Greater than 0.5 |



SVMTrip [SOS1]

FPYDWNWKTSLLLGGLLSATDPVAVVALLKELGASKKLSTIIEGESLMNDVMTLGMFYAAFARTAFKGESQQSLH
HFWEMVAYIANTLIFILSGVVIAEGANIDPDLLLAVFLPALLFESSFSMEVHQIKRCMAQMVLLAGPGVVQAAYWE
MLDEGRITQTTANILMQSVDEALDLVSPRKLVTYFTVERLESACYICAAFLRAHRIARRQLHDFIGDS

SVMTrip [HKT1]

LVIFIILICITERKKLKEDPLNFNVLNITLEVISAYGNVGFSTGYSCARQCKDAWYGFSGKWSNTGKFILIIVMFFFGR
LKKFNMKGGKAWKGLSSYQKLVGSLFQVVNSRHTGESVFDLSTISPAILVLFVVMMRPKDLDLFFTSVSATTVSS
MSTVEMEVFSNTQLIFLTILMFLGGEVFTSMTFSLFTTVSTFANCGFYPTNENMIYFKKNSGLLLILIPQVLGNTLFP
PC

ABCpred [SOS1]
FPYDWNWKTSLLLGGLLSATDPVAVVALLKELGASKKLSTIIEGESLMNDVMTLGMFYAAFARTAFKGESQQSLH
HFWEMVAYIANTLIFILSGVVIAEGANIDPDLLLAVFLPALLFESSFSMEVHQIKRCMAQMVLLAGPGVVQAAYWE
MLDEGRITQTTANILMQSVDEALDLVSPRKLVTYFTVERLESACYICAAFLRAHRIARRQLHDFIGDS

ABCpred [HKT1]
LVIFIILICITERKKLKEDPLNFNVLNITLEVISAYGNVGFSTGYSCARQCKDAWYGFSGKWSNTGKFILIIVMFFFGR
LKKFNMKGGKAWKGLSSYQKLVGSLFQVVNSRHTGESVFDLSTISPAILVLFVVMMRPKDLDLFFTSVSATTVSS
MSTVEMEVFSNTQLIFLTILMFLGGEVFTSMTFSLFTTVSTFANCGFYPTNENMIYFKKNSGLLLILIPQVLGNTLFP
PC

BepiPred [SOS1]

FPYDWNWKTSLLLGGLLSATDPVAVVALLKELGASKKLSTIIEGESLMNDVMTLGMFYAAFARTAFKGESQQSLH
HFWEMVAYIANTLIFILSGVVIAEGANIDPDLLLAVFLPALLFESSFSMEVHQIKRCMAQMVLLAGPGVVQAAYWE
MLDEGRITQTTANILMQSVDEALDLVSPRKLVTYFTVERLESACYICAAFLRAHRIARRQLHDFIGDS

BepiPred [HKT1]

LVIFIILICITERKKLKEDPLNFNVLNITLEVISAYGNVGFSTGYSCARQCKDAWYGFSGKWSNTGKFILIIVMFFFGR
LKKFNMKGGKAWKGLSSYQKLVGSLFQVVNSRHTGESVFDLSTISPAILVLFVVMMRPKDLDLFFTSVSATTVSS
MSTVEMEVFSNTQLIFLTILMFLGGEVFTSMTFSLFTTVSTFANCGFYPTNENMIYFKKNSGLLLILIPQVLGNTLFP
PC

***Figure 17:*** Selected Sequence of SOS1 and HKT1 from different server

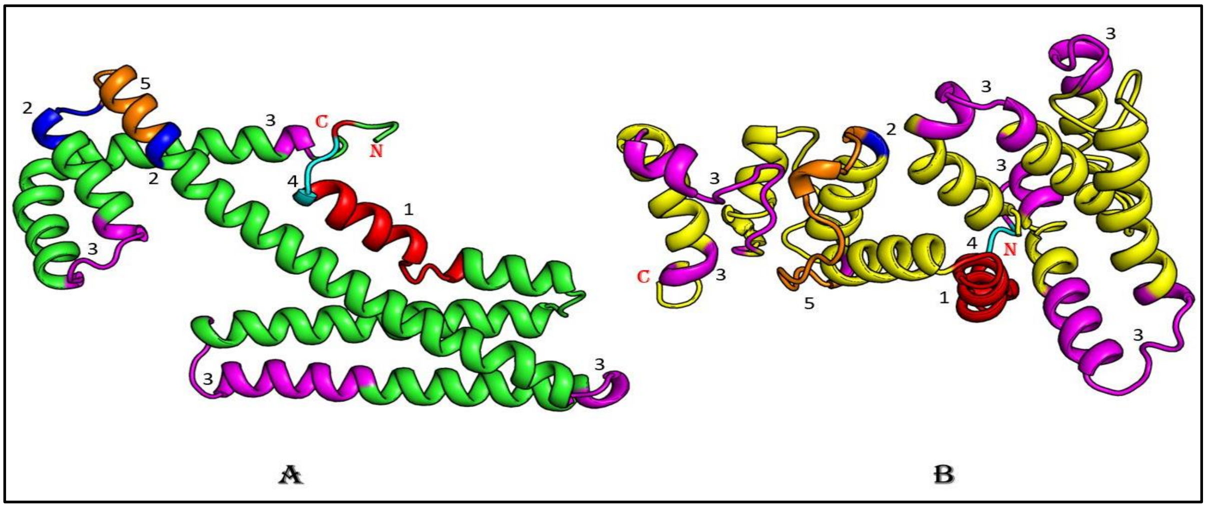


***Figure 18:*** Mapping of antigenicity by different servers in designed structure **A.** SOS1; **B.** HKT1. Both structures were mapped and labeled as 1. SVMTrip [Red]; 2. ABCpred [Blue]; 3. BepiPred [Magenta]; 4. SVMTrip + BepiPred [Cyans]; 5. ABCpred + BepiPred [Orange]

# Discussion

SOS1 is one of the major membrane proteins associated with the salinity tolerance. SOS1 protein are uses electrochemical gradients to extract the sodium ions from the internal of plant cell and extruding it, thereby maintaining low concentration of sodium in the cytoplasm of the cells. HKT1 is also a membrane protein associated with the salinity tolerance. It functions as a uniporter or symporter for the transport of potassium ions across the plasma membrane.

We have selected SOS1 and HKT1 membrane protein sequences of different plants respectively that were used for multiple sequence alignment and predicted the conserved part from the aligned sequences. The conserved motif 1, motif 2, motif 3, motif 4 and motif 5 were combined to form 120 amino acid sequences. We have also predicted the conserved motif parts position in some common plants such as in SOS1 compared plants were taken as *At*SOS1, *Tt*SOS1, *Os*SOS1, *Nt*SOS1, *St*SOS1 and for HKT1 compared plants were taken as *At*HKT1, *Gm*HKT1, *Ga*HKT1, *Sl*HKT1, *St*HKT1. This prediction was done to identify the region of conserved motif in plants.

Later, from 120 combination we have selected the most stable sequence as per the instability index and helix percentage. This prediction shows that for SOS1 sequence predicted combination were A1, A2, A3, A5, A4 whose instability index was 35.35 and helix percentage was 80%; for HKT1 sequence predicted combination were A1, A4, A5, A3, A2 whose instability index was 22.69 and helix percentage was 57%. After getting the stable sequences for SOS1 and HKT1 membrane protein, we have plotted a de novo model using the above selected sequences for SOS1 and HKT1. During modelling we get 5 models for SOS1 sequence and HKT1 sequence on the basis of the energy, stability and the error rate. Out of which we have selected the more stable and less error rate-based models.

After predicting the stable sequences and models for SOS1 and HKT1 protein, we predicted the structure using Ramachandran Plot. On the basis of the phi φ and the psi ψ angle in X and Y axis respectively and different statistics for residue in phi (φ) psi (ψ) core, residue in phi (φ) psi (ψ) allowed, residue in phi (φ) psi (ψ) generous and residue in phi (φ) psi (ψ) outside were enabled us to confirming the 3D structure for both SOS1 and HKT1 protein. This shows that we have successfully able to design the polypeptide sequences for SOS1 with length of 219 amino acid and HKT1 with length of 234 amino acid.

For further validation we have also run the molecular dynamic simulation of SOS1 and HKT1 stable structure. In this simulator, we focus on the different parameters such as radius of gyration, solvent accessibility, RMSD value and RMS fluctuation. In the designed SOS1 structure, from a radius of gyration we were able to predict from the graph that the molecular structure formed was compact; from RMSD we can predict that the structure formed is in an equilibrium state with little fluctuation by which we can presume the deviation of the atomic position with time; from RMSF graph we can predict some part of the sequence is flexible and some parts are rigid indicates about the fluctuation of each atom in its average positions. In designed HKT1 structure, from radius of gyration, we were able to predict from the graph that some part in the starting of the molecular structure is in a compact state while the rest of the parts is less compact or expansion; from RMSD graph, we can predict that most of the structure is in equilibrium state by which we can state the atomic position is stable overtime; from RMSF graph we can predict some part of the sequence is flexible and some parts are rigid indicates about the fluctuation of each atom in its average positions. They have also replotted the structure of the peptide designed where, we found that the after-simulation peptide structure formed is somehow compact structure. By this, we can assume that structure formed after simulation is more stable on the basis of the reducing the likelihood of degradation. This simulator helps to predict the different properties of the above two designed protein sequences on the basis of parameters.

For further analysis with respect to future aspects, we have gone for predicting the antigenicity part of the above designed SOS1 and HKT1 sequence. From this prediction we were able to analyse the binding site of the antibody in the above polypeptide designed sequences.  It was observed that most of the antigenic sites in SOS1 and HKT1 sequence are having overlapping sites respectively. This analysis was done to predict the accurate polypeptide binding sites, which might bind to the predicted region when antibody is produced in rabbits or mice or rat.

# Conclusion

We have selected SOS1 and HKT1 membrane protein sequences of different plants that show some functions in salinity stress. From the selected sequences we have done the multiple sequence alignment and predicted the five conserved motifs. We have also gone for multiple sequence alignment with some common plants to identify the region of conserved motif in such plants. Then the predicted conserved motifs were combined to form the one hundred twenty sequences. Out of those sequences we predicted the most stable sequences and plotted a de novo model and selected the most accurate model on the basis of the less error rate. After predicting the model and to validate it, we use Ramachandran Plot. On the basis of the phi $\varphi$ and the psi $\psi$ angle in X and Y axis respectively it established that we have successfully able to design the polypeptide sequences for SOS1 with length of 219 amino acid and HKT1 with length of 234 amino acid.

Later, we have also validated the above designed polypeptide structure using the molecular dynamic simulation on the basis of different parameters. Structure formation after simulation, we can assume that the structure is more stable. For future prospects, we have gone for predicting the antigenicity part of the above designed SOS1 and HKT1 sequence.

After various prediction, designing, interpreting and finally validating by various methods we can expect that the designer of ion transporter-based polypeptide markers can be applied in multiple plants for detection of salinity stress.

# References

- Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX, 1, 19-25.
- Adolf, V. I., Jacobsen, S. E., & Shabala, S. (2013). Salt tolerance mechanisms in quinoa (Chenopodium quinoa Willd.). Environmental and Experimental Botany, 92, 43-54.
- Ali, A., Maggio, A., Bressan, R. A., & Yun, D. J. (2019). Role and functional differences of HKT1-type transporters in plants under salt stress. International journal of molecular sciences, 20(5), 1059. https://doi.org/10.3390/ijms20051059
- Ali, A., Petrov, V., Yun, D. J., & Gechev, T. (2023). Revisiting plant salt tolerance: novel components of the SOS pathway. Trends in plant science, 28(9), 1060–1069. https://doi.org/10.1016/j.tplants.2023.04.003.
- Almeida, P., de Boer, G. J., & de Boer, A. H. (2014). Differences in shoot Na+ accumulation between two tomato species are due to differences in ion affinity of HKT1; 2. Journal of Plant Physiology, 171(6), 438-447.
- Arora, S., Singh, Y. P., Vanza, M., & Sahni, D. (2016). Bio-remediation of saline and sodic soils through halophilic bacteria to enhance agricultural production. Journal of Soil and Water Conservation, 15(4), 302-305. http://dx.doi.org/10.5958/2455-7145.2016.00027.8
- Assaha, D. V. M., Ueda, A., Saneoka, H., Al-Yahyai, R., & Yaish, M. W. (2017). The Role of Na+ and K+ Transporters in Salt Stress Adaptation in Glycophytes. Frontiers in physiology, 8, 509. https://doi.org/10.3389/fphys.2017.00509.
- Bailey, T. L., & Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in bipolymers.
- Banik, S., & Dutta, D. (2023). Membrane proteins in plant salinity stress perception, sensing, and response. The Journal of Membrane Biology, 256(2), 109-124. https://doi.org/10.1007/s00232-023-00279-9
- Bartels, D., & Sunkar, R. (2005). Drought and salt tolerance in plants. Critical reviews in plant sciences, 24(1), 23-58. http://dx.doi.org/10.1080/07352680590910410
- Bekker, H., Berendsen, H. J. C., Dijkstra, E. J., Achterop, S., Vondrumen, R. V., Vanderspoel, D., ... & Renardus, M. K. R. (1993). Gromacs-a parallel computer for

molecular-dynamics simulations. In 4th international conference on computational physics (PC 92) (pp. 252-256). World Scientific Publishing.

- Bjelkmar, P., Larsson, P., Cuendet, M. A., Hess, B., & Lindahl, E. (2010). Implementation of the CHARMM force field in GROMACS: analysis of protein stability effects from correction maps, virtual interaction sites, and water models. Journal of chemical theory and computation, 6(2), 459-466.
- Cabot, C., Sibole, J. V., Barceló, J., & Poschenrieder, C. (2014). Lessons from crop plants struggling with salinity. Plant Science, 226, 2-13.
- Chen, H., Chen, X., Gu, H., Wu, B., Zhang, H., Yuan, X., & Cui, X. (2014). GmHKT1; 4, a novel soybean gene regulating Na+/K+ ratio in roots enhances salt tolerance in transgenic plants. *Plant growth regulation*, *73*, 299-308.
- Corratgé-Faillie, C., Jabnoune, M., Zimmermann, S., Véry, A. A., Fizames, C., & Sentenac, H. (2010). Potassium and sodium transport in non-animal cells: the Trk/Ktr/HKT transporter family. *Cellular and Molecular Life Sciences*, 67(15), 2511-2532.
- Cosentino, C., Fischer-Schliebs, E., Bertl, A., Thiel, G., & Homann, U. (2010). Na+/H+ antiporters are differentially regulated in response to NaCl stress in leaves and roots of Mesembryanthemum crystallinum. New Phytologist, 186(3), 669-680.
- CRAIG PLETT, D. A. R. R. E. N., & Møller, I. S. (2010). Na+ transport in glycophytic plants: what we know and would like to know. Plant, cell & environment, 33(4), 612-626.
- Cushman, K. R., Pabuayon, I. C., Hinze, L. L., Sweeney, M. E., & de Los Reyes, B. G. (2020). Networks of physiological adjustments and defenses, and their synergy with sodium (Na+) homeostasis explain the hidden variation for salinity tolerance across the cultivated Gossypium hirsutum germplasm. Frontiers in plant science, 11, 588854.
- Dresselhaus, T., & Hückelhoven, R. (2018). Biotic and abiotic stress responses in crop plants. Agronomy, 8(11), 267.
- Dutta, D. (2023). Interplay between membrane proteins and membrane protein-lipid pertaining to plant salinity stress. Cell Biochemistry and Function, 41(4), 399-412. https://doi.org/10.1002/cbf.3798

- Ellouzi, H.; Ben Hamed, K.; Cela, J.; Munné-Bosch, S.; Abdelly, C. Early effects of salt stress on the physiological and oxidative status of Cakile maritima (halophyte) and Arabidopsis thaliana (glycophyte). Physiol. Plant. 2011.
- Fan, Y., Yin, X., Xie, Q., Xia, Y., Wang, Z., Song, J., ... & Jiang, X. (2019). Co-expression of SpSOS1 and SpAHA1 in transgenic Arabidopsis plants improves salinity tolerance. BMC plant biology, 19, 1-13.
- Farooq, M., Hussain, M., Wakeel, A., & Siddique, K. H. (2015). Salt stress in maize: effects, resistance mechanisms, and management. A review. Agronomy for Sustainable Development, 35, 461-481.
- Feki, K., Quintero, F. J., Pardo, J. M., & Masmoudi, K. (2011). Regulation of durum wheat Na+/H+ exchanger TdSOS1 by phosphorylation. Plant molecular biology, 76, 545-556.
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. E., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server (pp. 571-607). Humana press.
- Gong, D., Guo, Y., Schumaker, K.S., and Zhu, J.-K. (2004). The SOS3 family of calcium sensors and SOS2 family of protein kinases in Arabidopsis. Plant Physiol. 134 : 919–926.
- Guo, Q., Meng, L., Han, J., Mao, P., Tian, X., Zheng, M., & Mur, L. A. (2020). SOS1 is a key systemic regulator of salt secretion and K+/Na+ homeostasis in the recretohalophyte Karelinia caspia. Environmental and experimental botany, 177, 104098.
- Guo, Q., Meng, S., Tao, S., Feng, J., Fan, X., Xu, P., ... & Shen, X. (2020). Overexpression of a samphire high-affinity potassium transporter gene SbHKT1 enhances salt tolerance in transgenic cotton. *Acta physiologiae plantarum*, *42*, 1-8.
- Hamamoto, S., Horie, T., Hauser, F., Deinlein, U., Schroeder, J. I., & Uozumi, N. (2015). HKT transporters mediate salt stress resistance in plants: from structure and function to the field. *Current opinion in biotechnology*, *32*, 113-120.
- Han, X., & Yang, Y. (2021). Phospholipids in Salt Stress Response. Plants (Basel, Switzerland), 10(10), 2204. https://doi.org/10.3390/plants10102204.
- Han, Y., Yin, S., Huang, L., Wu, X., Zeng, J., Liu, X., ... & Zhang, G. (2018). A sodium transporter HvHKT1; 1 confers salt tolerance in barley via regulating tissue and cell ion homeostasis. Plant and Cell Physiology, 59(10), 1976-1989.

- Horie, T., Motoda, J., Kubo, M., Yang, H., Yoda, K., Horie, R., ... & Uozumi, N. (2005). Enhanced salt tolerance mediated by AtHKT1 transporter-induced Na+ unloading from xylem vessels to xylem parenchyma cells. *The Plant Journal,* 44(6)*,* 928-938.
- Horie, T., Yoshida, K., Nakayama, H., Yamada, K., Oiki, S., & Shinmyo, A. (2001). Two types of HKT transporters with different properties of Na+ and K+ transport in Oryza sativa. The Plant Journal, 27(2), 129-138. https://doi.org/10.1046/j.1365-313x.2001.01077.x
- Källberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H., & Xu, J. (2012). Template-based protein structure modeling using the RaptorX web server. Nature protocols, 7(8), 1511-1522.
- Katoh, K., Rozewicki, J., & Yamada, K. D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in bioinformatics*, *20*(4), 1160-1166.
- Katschnig, D., Bliek, T., Rozema, J., & Schat, H. (2015). Constitutive high-level SOS1 expression and absence of HKT1; 1 expression in the salt-accumulating halophyte Salicornia dolichostachya. Plant Science, 234, 144-154.

- Katschnig, D., Bliek, T., Rozema, J., & Schat, H. (2015). Constitutive high-level SOS1 expression and absence of HKT1; 1 expression in the salt-accumulating halophyte Salicornia dolichostachya. Plant Science, 234, 144-154.

- Kaur, G., Sanwal, S. K., Sehrawat, N., Kumar, A., Kumar, N., & Mann, A. (2022). Getting to the roots of Cicer arietinum L.(chickpea) to study the effect of salinity on morpho-physiological, biochemical and molecular traits. Saudi Journal of Biological Sciences, 29(12), 103464.
- Kumar, P., & Sharma, P. K. (2020). Soil salinity and food security in India. Frontiers in Sustainable Food Systems, 4, 533781. https://doi.org/10.3389/fsufs.2020.533781
- Kuraku, S., Zmasek, C. M., Nishimura, O., & Katoh, K. (2013). aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. Nucleic acids research, 41(W1), W22-W28.
- Liang, L., Guo, L., Zhai, Y., Hou, Z., Wu, W., Zhang, X., ... & Liu, W. (2023). Genome-wide characterization of SOS1 gene family in potato (Solanum tuberosum) and expression analyses under salt and hormone stress. Frontiers in Plant Science, 14, 1201730.

- Lin, H., Yang, Y., Quan, R., Mendoza, I., Wu, Y., Du, W., ... & Guo, Y. (2009). Phosphorylation of SOS3-LIKE CALCIUM BINDING PROTEIN8 by SOS2 protein kinase stabilizes their protein complex and regulates salt tolerance in Arabidopsis. The Plant Cell, 21(5), 1607-1619.
- Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., & Shaw, D. E. (2010). Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins: Structure, Function, and Bioinformatics, 78(8), 1950-1958.
- Liu, J., & Shi, D. C. (2010). Photosynthesis, chlorophyll fluorescence, inorganic ion and organic acid accumulations of sunflower in responses to salt and salt-alkaline mixed stress. Photosynthetica, 48, 127-134. https://doi.org/10.1007/s11099-010-0017-4
- Lu, X., Ma, L., Zhang, C., Yan, H., Bao, J., Gong, M., ... & Chen, B. (2022). Grapevine (Vitis vinifera) responses to salt stress and alkali stress: transcriptional and metabolic profiling. BMC Plant Biology, 22(1), 528.
- Ma, J., Wang, S., Zhao, F., & Xu, J. (2013). Protein threading using context-specific alignment potential. Bioinformatics, 29(13), i257-i265.
- Mansour, M. M. F., Salama, K. H., & Allam, H. Y. (2015). Role of the plasma membrane in saline conditions: lipids and proteins. The Botanical Review, 81, 416-451.
- Munns, R., & Tester, M. (2008). Mechanisms of salinity tolerance. Annu. Rev. Plant Biol., 59, 651-681. https://doi.org/10.1146/annurev.arplant.59.032607.092911
- Murthy, M., & Tester, M. (2006). Cation currents in protoplasts from the roots of a Na+ hyperaccumulating mutant of Capsicum annuum. *Journal of experimental botany*, *57*(5), 1171-1180.
- Mythili, G., & Goedecke, J. (2016). Economics of land degradation in India. Economics of land degradation and improvement–a global assessment for sustainable development, 431-469. http://dx.doi.org/10.1007/978-3-319-19168-3_15
- Oostenbrink, C., Villa, A., Mark, A. E., & Van Gunsteren, W. F. (2004). A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *Journal of computational chemistry*, *25*(13), 1656-1676.

- Parkin, I. A., Koh, C., Tang, H., Robinson, S. J., Kagale, S., Clarke, W. E., ... & Sharpe, A. G. (2014). Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid Brassica oleracea. Genome biology, 15, 1-18.
- Platten, J. D., Cotsaftis, O., Berthomieu, P., Bohnert, H., Davenport, R. J., Fairbairn, D. J., & Tester, M. (2006). Nomenclature for HKT transporters, key determinants of plant salinity tolerance. Trends in plant science, 11(8), 372-374. https://doi.org/10.1016/j.tplants.2006.06.001
- Qadir, M., Oster, J. D., Schubert, S., Noble, A. D., & Sahrawat, K. L. (2007). Phytoremediation of sodic and saline-sodic soils. *Advances in agronomy*, *96*, 197-247.
- Riedelsberger, J., Miller, J. K., Valdebenito-Maturana, B., Piñeros, M. A., González, W., & Dreyer, I. (2021). Plant HKT channels: an updated view on structure, function and gene regulation. *International journal of molecular sciences*, *22*(4), 1892.
- Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. Nucleic acids research, 42(W1), W320-W324.
- Roeber, V. M., Bajaj, I., Rohde, M., Schmülling, T., & Cortleven, A. (2021). Light acts as a stressor and influences abiotic and biotic stress responses in plants. Plant, Cell & Environment, 44(3), 645-664.
- Saha, S., & Raghava, G. P. (2007). Prediction methods for B-cell epitopes. Immunoinformatics: Predicting Immunogenicity In Silico, 387-394.
- Saha, S., & Raghava, G. P. S. (2004, September). BcePred: prediction of continuous B-cell epitopes in antigenic sequences using physico-chemical properties. In International conference on artificial immune systems (pp. 197-204). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Saha, S., & Raghava, G. P. S. (2006). Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins: Structure, Function, and Bioinformatics, 65(1), 40-48.
- Santander, C., Aroca, R., Cartes, P., Vidal, G., & Cornejo, P. (2021). Aquaporins and cation transporters are differentially regulated by two arbuscular mycorrhizal fungi strains in lettuce cultivars growing under salinity conditions. Plant Physiology and Biochemistry, 158, 396-409.

- Schachtman, D. P., & Schroeder, J. I. (1994). Structure and transport mechanism of a high-affinity potassium uptake transporter from higher plants. Nature, 370(6491), 655-658. https://doi.org/10.1038/370655a0
- Shahzad, B., Shabala, L., Zhou, M., Venkataraman, G., Solis, C. A., Page, D., ... & Shabala, S. (2022). Comparing essentiality of SOS1-mediated Na+ exclusion in salinity tolerance between cultivated and wild rice species. International Journal of Molecular Sciences, 23(17), 9900.
- Shao, Q., Han, N., Ding, T., Zhou, F., & Wang, B. (2014). SsHKT1; 1 is a potassium transporter of the C3 halophyte Suaeda salsa that is involved in salt tolerance. Functional Plant Biology, 41(8), 790-802.
- Shi, H., Ishitani, M., Kim, C., & Zhu, J. K. (2000). The Arabidopsis thaliana salt tolerance gene SOS1 encodes a putative Na+/H+ antiporter. Proceedings of the national academy of sciences, 97(12), 6896-6901.
- Soltabayeva, A., Ongaltay, A., Omondi, J. O., & Srivastava, S. (2021). Morphological, Physiological and Molecular Markers for Salt-Stressed Plants. Plants (Basel, Switzerland), 10(2), 243. https://doi.org/10.3390/plants10020243.
- Uozumi, N., Kim, E. J., Rubio, F., Yamaguchi, T., Muto, S., Tsuboi, A., ... & Schroeder, J. I. (2000). The Arabidopsis HKT1 gene homolog mediates inward Na+ currents in Xenopus laevis oocytes and Na+ uptake in Saccharomyces cerevisiae. Plant physiology, 122(4), 1249-1260.
- Vaishnav, A., Kumari, S., Jain, S., Varma, A., Tuteja, N., & Choudhary, D. K. (2016). PGPR-mediated expression of salt tolerance gene in soybean through volatiles under sodium nitroprusside. Journal of basic microbiology, 56(11), 1274-1288.
- Wang, L., Liu, Y., Feng, S., Wang, Z., Zhang, J., Zhang, J., ... & Gan, Y. (2018). AtHKT1 gene regulating K+ state in whole plant improves salt tolerance in transgenic tobacco plants. *Scientific reports*, *8*(1), 16585.
- Wang, L., Liu, Y., Li, D., Feng, S., Yang, J., Zhang, J., ... & Gan, Y. (2019). Improving salt tolerance in potato through overexpression of AtHKT1 gene. BMC plant biology, 19, 1-15.
- Wang, T. B., Gassmann, W., Rubio, F., Schroeder, J. I., & Glass, A. D. (1998). Rapid up-regulation of HKT1, a high-affinity potassium transporter gene, in roots of barley and wheat following withdrawal of potassium. Plant Physiology, 118(2), 651-659.

- Willard, L., Ranjan, A., Zhang, H., Monzavi, H., Boyko, R. F., Sykes, B. D., & Wishart, D. S. (2003). VADAR: a web server for quantitative evaluation of protein structure quality. Nucleic acids research, 31(13), 3316-3319.
- Wu, Y., Meng, K., & Liang, X. (2017). Distinct patterns of natural selection in Na+/H+ antiporter genes in Populus euphratica and Populus pruinosa. Ecology and Evolution, 7(1), 82-91.
- Xu, Y., Zheng, X., Song, Y., Zhu, L., Yu, Z., Gan, L., ... & Zhu, C. (2018). NtLTP4, a lipid transfer protein that enhances salt and drought stresses tolerance in Nicotiana tabacum. Scientific reports, 8(1), 8873.
- Yao, B., Zhang, L., Liang, S., & Zhang, C. (2012). SVMTriP: a method to predict antigenic epitopes using support vector machine to integrate tri-peptide similarity and propensity.
- Yue, Y., Zhang, M., Zhang, J., Duan, L., & Li, Z. (2012). SOS1 gene overexpression increased salt tolerance in transgenic tobacco by maintaining a higher K+/Na+ ratio. Journal of Plant Physiology, 169(3), 255-261.
- Zhang, H., Zhu, J., Gong, Z., & Zhu, J. K. (2022). Abiotic stress responses in plants. Nature Reviews Genetics, 23(2), 104-119.
- Zhang, M., Cao, J., Zhang, T., Xu, T., Yang, L., Li, X., ... & Xie, L. (2022). A putative plasma membrane Na+/H+ antiporter GmSOS1 is critical for salt stress tolerance in Glycine max. Frontiers in Plant Science, 13, 870695.
- Zhang, X. Y., Tang, L. H., Nie, J. W., Zhang, C. R., Han, X., Li, Q. Y., ... & Chen, Y. H. (2023). Structure and activation mechanism of the rice Salt Overly Sensitive 1 (SOS1) Na+/H+ antiporter. *Nature Plants*, *9*(11), 1924-1936.
- Zhu, J. K. (2016). Abiotic stress signaling and responses in plants. Cell, 167(2), 313-324.
- Zhu, L., Lu, L., Yang, L., Hao, Z., Chen, J., & Cheng, T. (2021). The full-length transcriptome sequencing and identification of Na+/H+ antiporter genes in halophyte Nitraria tangutorum Bobrov. Genes, 12(6), 836.

# ANNEXURE-I

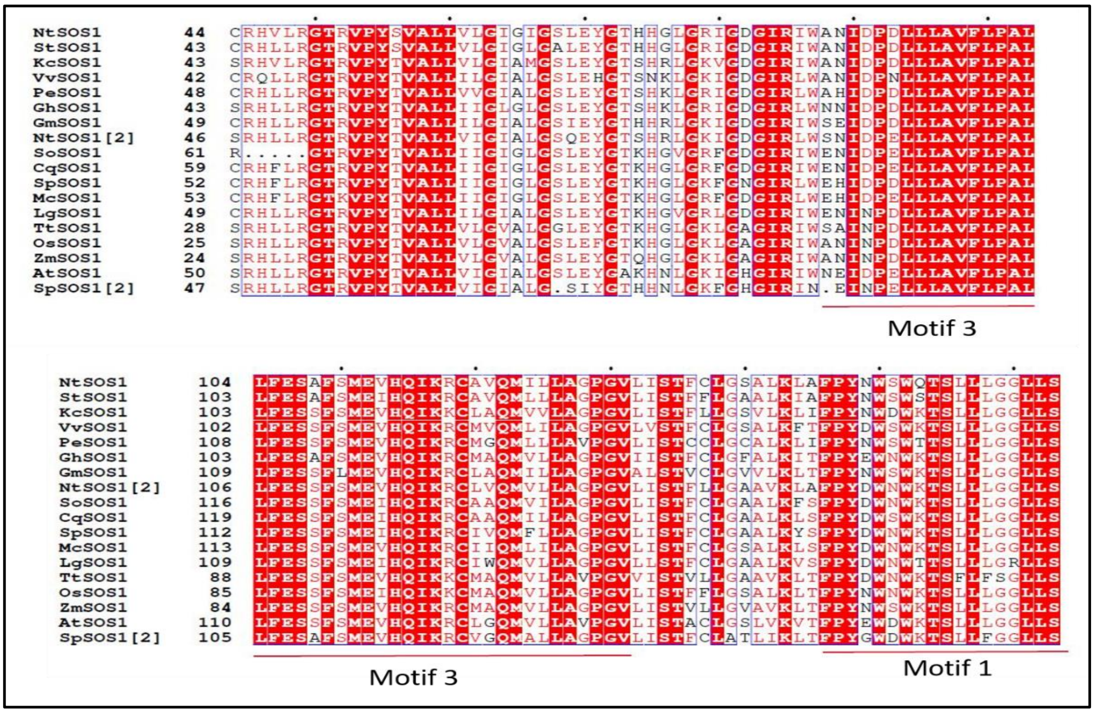


***Figure 19:*** Highly conserved amino acid motif after multiple sequence alignment of SOS1 with motif 3 and motif 1

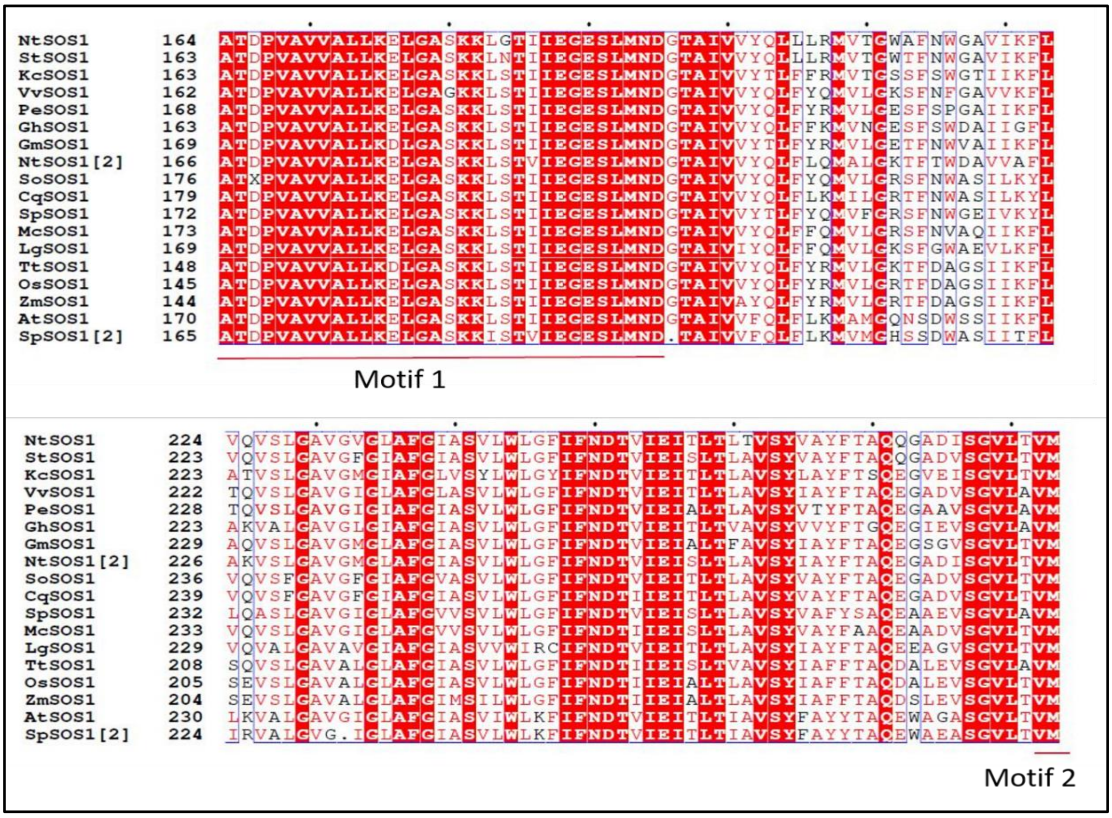
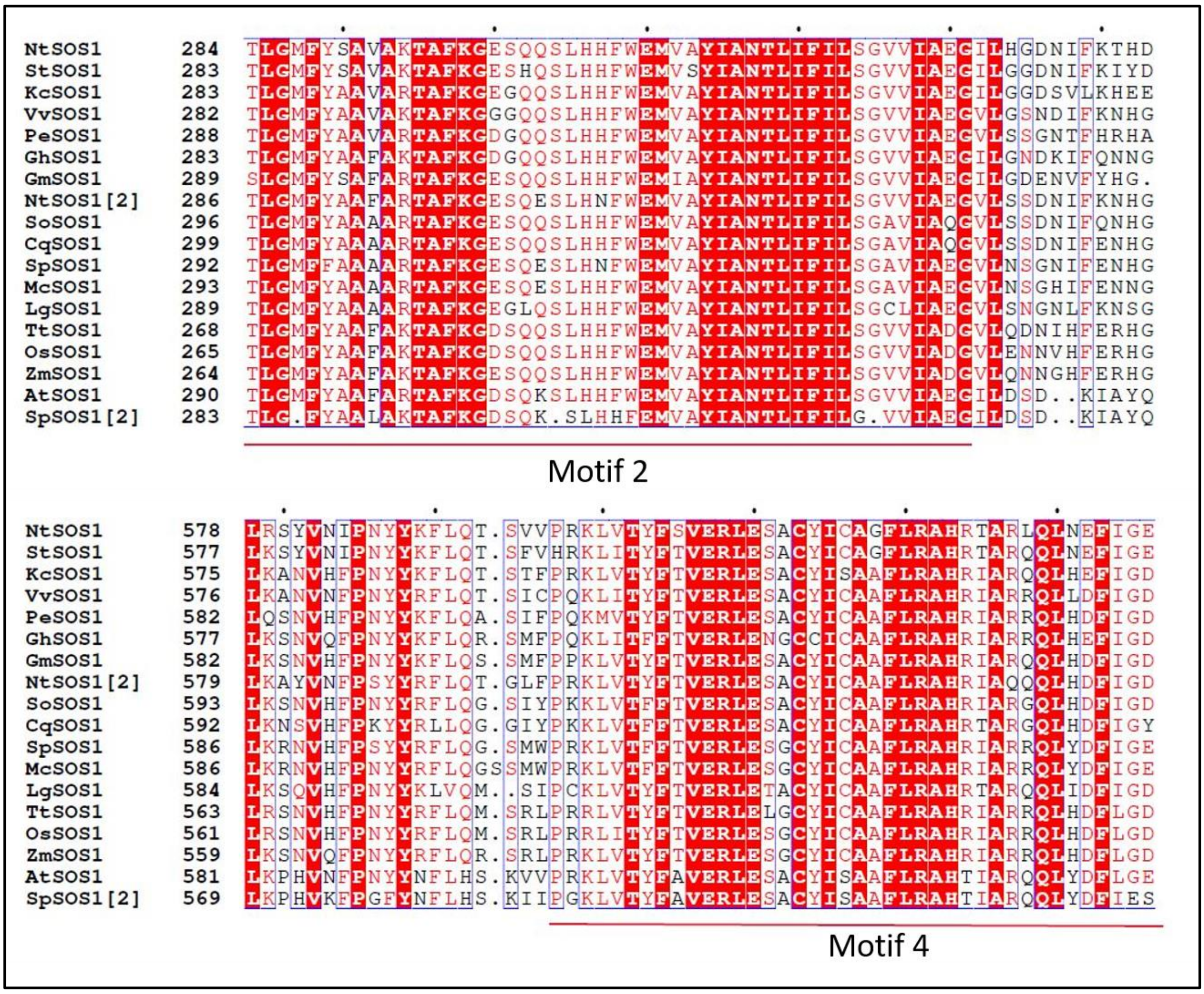


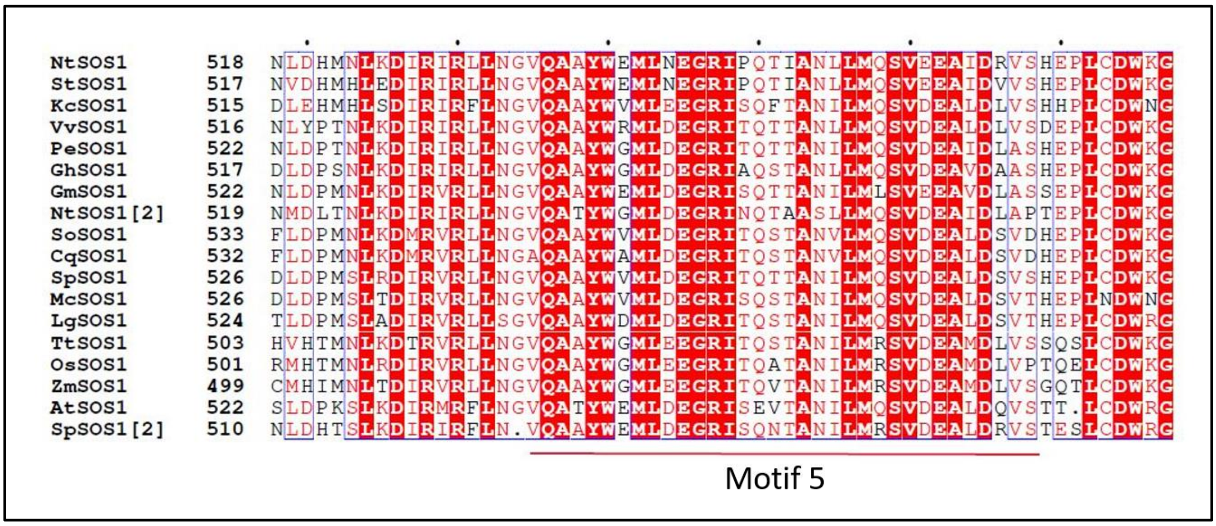
***Figure 20:*** Highly conserved amino acid motif after multiple sequence alignment of SOS1 with motif 1 and motif 2



```
NtSOS1     284  TLGMFYSAVAKTAFKGESQQSLHHFWEMVAYIANTLIFILSGVVIAEGILHGDNIFKTHD
StSOS1     283  TLGMFYSAVAKTAFKGESHQSLHHFWEMVSYIANTLIFILSGVVIAEGILGGDNIFKIYD
KcSOS1     283  TLGMFYAAVARTAFKGEGQQSLHHFWEMVAYIANTLIFILSGVVIAEGILGGDSVLKHEE
VvSOS1     282  TLGMFYAAVAKTAFKGGGQQSLHHFWEMVAYIANTLIFILSGVVIAEGVLGSNDIFKNHG
PeSOS1     288  TLGMFYAAVARTAFKGDGQQSLHHFWEMVAYIANTLIFILSGVVIAEGVLSSGNTFHRHA
GhSOS1     283  TLGMFYAAFAKTAFKGDGQQSLHHFWEMVAYIANTLIFILSGVVIAEGILGNDKIFQNNG
GmSOS1     289  SLGMFYSAFARTAFKGESQQSLHHFWEMIAYIANTLIFILSGVVIAEGILGDENVFYHG.
NtSOS1[2]  286  TLGMFYAAFARTAFKGESQESLHNFWEMVAYIANTLIFILSGVVIAEGVLSSDNIFKNHG
SoSOS1     296  TLGMFYAAAARTAFKGESQQSLHHFWEMVAYIANTLIFILSGAVIAQGVLSSDNIFQNHG
CqSOS1     299  TLGMFYAAAARTAFKGESQQSLHHFWEMVAYIANTLIFILSGAVIAQGVLSSDNIFENHG
SpSOS1     292  TLGMFFAAAARTAFKGESQESLHNFWEMVAYIANTLIFILSGAVIAEGVLNSGNIFENHG
McSOS1     293  TLGMFYAAAARTAFKGESQESLHHFWEMVAYIANTLIFILSGAVIAEGVLNSGHIFENNG
LgSOS1     289  TLGMFYAAAARTAFKGEGLQSLHHFWEMVAYIANTLIFILSGCLIAEGVLSNGNLFKNSG
TtSOS1     268  TLGMFYAAFAKTAFKGDSQQSLHHFWEMVAYIANTLIFILSGVVIADGVLQDNIHFERHG
OsSOS1     265  TLGMFYAAFAKTAFKGDSQQSLHHFWEMVAYIANTLIFILSGVVIADGVLENNVHFERHG
ZmSOS1     264  TLGMFYAAFAKTAFKGESQQSLHHFWEMVAYIANTLIFILSGVVIADGVLQNNGHFERHG
AtSOS1     290  TLGMFYAAFARTAFKGDSQKSLHHFWEMVAYIANTLIFILSGVVIAEGILDSD..KIAYQ
SpSOS1[2]  283  TLG.FYAALAKTAFKGDSQK.SLHHFEMVAYIANTLIFILG.VVIAEGILDSD..KIAYQ
```

```
NtSOS1     578  LRSYVNIPNYYKFLQT.SVVPRKLVTYFSVERLESACYICAGFLRAHRTARLQLNEFIGE
StSOS1     577  LKSYVNIPNYYKFLQT.SFVHRKLITYFTVERLESACYICAGFLRAHRTARQQLNEFIGE
KcSOS1     575  LKANVHFPNYYKFLQT.STFPRKLVTYFTVERLESACYISAAFLRAHRIARQQLHEFIGD
VvSOS1     576  LKANVNFPNYYRFLQT.SICPQKLITYFTVERLESACYICAAFLRAHRIARRQLLDFIGD
PeSOS1     582  LQSNVHFPNYYKFLQA.SIFPQKMVTYFTVERLESACYICAAFLRAHRIARRQLHDFIGD
GhSOS1     577  LKSNVQFPNYYKFLQR.SMFPQKLITFFTVERLENGCCICAAFLRAHRIARRQLHEFIGD
GmSOS1     582  LKSNVHFPNYYKFLQS.SMFPPKLVTYFTVERLESACYICAAFLRAHRIARQQLHDFIGD
NtSOS1[2]  579  LKAYVNFPSYYRFLQT.GLFPRKLVTYFTVERLESACYICAAFLRAHRIAQQQLHDFIGD
SoSOS1     593  LKSNVHFPNYYRFLQG.SIYPKKLVTFFTVERLESACYICAAFLRAHRIARGQLHDFIGD
CqSOS1     592  LKNSVHFPKYYRLLQG.GIYPKKLVTFFTVERLESACYICAAFLRAHRTARGQLHDFIGY
SpSOS1     586  LKRNVHFPSYYRFLQG.SMWPRKLVTFFTVERLESGCYICAAFLRAHRIARRQLYDFIGE
McSOS1     586  LKRNVHFPNYYRFLQGSSMWPRKLVTFFTVERLESGCYICAAFLRAHRIARRQLYDFIGE
LgSOS1     584  LKSQVHFPNYYKLVQM..SIPCKLVTYFTVERLETACYICAAFLRAHRTARQQLIDFIGD
TtSOS1     563  LRSNVHFPNYYRFLQM.SRLPRRLVTYFTVERLELGCYICAAFLRAHRIARRQLHDFLGD
OsSOS1     561  LRSNVHFPNYYRFLQM.SRLPRRLITYFTVERLESGCYICAAFLRAHRIARRQLHDFLGD
ZmSOS1     559  LKSNVQFPNYYRFLQR.SRLPRKLVTYFTVERLESGCYICAAFLRAHRIARRQLHDFLGD
AtSOS1     581  LKPHVNFPNYYNFLHS.KVVPRKLVTYFAVERLESACYISAAFLRAHTIARQQLYDFLGE
SpSOS1[2]  569  LKPHVKFPGFYNFLHS.KIIPGKLVTYFAVERLESACYISAAFLRAHTIARQQLYDFIES
```

***Figure 21:*** Highly conserved amino acid motif after multiple sequence alignment of SOS1 with motif 2 and motif 4



```
NtSOS1     518  NLDHMNLKDIRIRLLNGVQAAYWEMLNEGRIPQTIANLLMQSVEEAIDRVSHEPLCDWKG
StSOS1     517  NVDHMHLEDIRIRLLNGVQAAYWEMLNEGRIPQTIANLLMQSVEEAIDVVSHEPLCDWKG
KcSOS1     515  DLEHMHLSDIRIRFLNGVQAAYWVMLEEGRISQFTANILMQSVDEALDLVSHHPLCDWNG
VvSOS1     516  NLYPTNLKDIRIRLLNGVQAAYWRMLDEGRITQTTANLLMQSVDEALDLVSDEPLCDWKG
PeSOS1     522  NLDPTNLKDIRIRLLNGVQAAYWGMLDEGRITQTTANILMQSVDEAIDLASHEPLCDWKG
GhSOS1     517  DLDPSNLKDIRIRLLNGVQAAYWGMLDEGRIAQSTANLLMQSVDEAVDAASHEPLCDWKG
GmSOS1     522  NLDPMNLKDIRVRLLNGVQAAYWEMLDEGRISQTTANILMLSVEEAVDLASSEPLCDWKG
NtSOS1[2]  519  NMDLTNLKDIRIRLLNGVQATYWGMLDEGRINQTAASLLMQSVDEAIDLAPTEPLCDWKG
SoSOS1     533  FLDPMNLKDMRVRLLNGVQAAYWVMLDEGRITQSTANVLMQSVDEALDSVDHEPLCDWKG
CqSOS1     532  FLDPMNLKDMRVRLLNGAQAAYWAMLDEGRITQSTANVLMQSVDEALDSVDHEPLCDWKG
SpSOS1     526  DLDPMSLRDIRVRLLNGVQAAYWVMLDEGRITQTTANILMQSVDEALDSVSHEPLCDWKG
McSOS1     526  DLDPMSLTDIRVRLLNGVQAAYWVMLDEGRISQSTANILMQSVDEALDSVTHEPLNDWNG
LgSOS1     524  TLDPMSLADIRVRLLSGVQAAYWDMLDEGRITQSTANILMQSVDEALDSVTHEPLCDWRG
TtSOS1     503  HVHTMNLKDTRVRLLNGVQAAYWGMLEEGRITQSTANILMRSVDEAMDLVSSQSLCDWKG
OsSOS1     501  RMHTMNLRDIRVRLLNGVQAAYWGMLEEGRITQATANILMRSVDEAMDLVPTQELCDWKG
ZmSOS1     499  CMHIMNLTDIRVRLLNGVQAAYWGMLEEGRITQVTANILMRSVDEAMDLVSGQTLCDWKG
AtSOS1     522  SLDPKSLKDIRMRFLNGVQATYWEMLDEGRISEVTANILMQSVDEALDQVSTT.LCDWRG
SpSOS1[2]  510  NLDHTSLKDIRIRFLN.VQAAYWEMLDEGRISQNTANILMRSVDEALDRVSTESLCDWRG
```

***Figure 22:*** Highly conserved amino acid motif after multiple sequence alignment of SOS1 with motif 5

```
AtHKT1      31  FFSFLGFLALKITKPR......TTSRPHDFDLFFTSVSAITVSSMSTVDMEVFSNTQLIF
BrHKT1      33  TFSILGFLALKISKPR......ITSRPHDFDLFFTSVSALTISSMSTIDMEVFSNTQLIF
GaHKT1      57  VLSLVGFGALKMSKPKSG..SVSP...KNIDVFFTSVSATTVSSMSTVEMEVFSNTQLII
RcHKT1      49  LISLLGYLALKISTPRTT..SSSSSKIKDIDIFFTSVSATTVSSMSTVEMEAFSNTQLII
SlHKT1      30  TLSLLGFLALNY.VSKPR..TLPSFRPQNLDVLFTSVSSTTVSSMSTIEMEVFSNVQLVF
StHKT1      28  TLSFLGFLALNY.FSKPR..TLPSFRPQNLDVFFTSVSSTTVSSMSTIEMEVFSNVQLVF
CaHKT1[2]   38  IFSFLGFLALN..VTKPR..TLPSFRPHNLDVFFTSVSATTVSSMSTVEMEVFSNAQLVF
NtHKT1      38  ILSLLGLLALT..ISKPR..TLPSFRPQNLDVFFTSVSAATVSSMSTVEMEVFSNAQLVF
SiHKT1      57  TLSSLGFLALN..ASNPR..TTTTFKPKYIDLFFTSVSASTVSSMSTVEMEVFSNTQLVL
DcHKT1      61  LFSAAGFLVLKV.SSSS......ISKPGDLDLFFTSVSAVSVSSMSTVEMEVFSDFQLVV
GmHKT1      48  ILSLVGYLGLKVSKPRN......PVRLNDLDLFYTSVSASTASSMVAVEMEVFSNSQLIL
CcHKT1      41  ILSMLGYFSLKLSKPRT......RVRPNDLDLFFTSVSASTVSSMVAVEMEVFSNSQLIL
CaHKT1      45  ILSLFGYLGLKFSKPRT......CFRPKDFDLFYTSVSASTVSSMTSIEMEVFSNFQLIL
AhHKT1      35  IMSVFGYLGLKVSKPKT......YVKPKNFDLFYTSVSASTVSSMTAIEMEVFSNSQLVL
SOHKT1      56  LVSSLGFFILKSLNP....NHNNNYNPKDLDLFFTSVSATTVSSMSTMEMEVFSNSQIVI
SsHKT1      56  LVSSFGFLFLKTLIPKSSNNNNNNNSINDLDLFFTSVSATTVSSMSTLEMEVFSNSQLIV
HaHKT1      40  FISTLGFLILHSLDVR.....TTTFTPRNLDLFFTSVSAATVSSMSTVEMEVFSNAQLII
LsHKT1      55  FISFLGFLILNAMDPR.....TPSFTPRKLDLFFTSVSASTVSSMSVIEMEVFSNAQLII
EcHKT1      57  FLSFLGFWVLKALGPR.....TDSFRPRDLDLFFTSVSATTVSSMSTVEMEVFSNSQLVV
VvHKT1      61  CLSFQGFLVLKALKPR.....TYSFRPRNLDLFFTSVSAATVSSMSTVEMEVFSNNQLLI
HvHKT1       1  .....................MKTSMPTDLDLIFTSVSATTVSSMQAVEMESFSNPQLLL
```

```
AtHKT1      85  LTILMFLGGEIFTSFLNLYVSYFTKFVFPHNK............................
BrHKT1      87  LTILMFLGGEVFTSFLNLYFSHFTKFVYPHNK............................
GaHKT1     112  ITILMLVGGEVFTSMLGLLLGRFRFETPPSFDGDIVNSISNRGLQ..SDYSLTSKNIAQE
RcHKT1     107  MTTLMLVGGEVFTSMLGLRLERSKFRTPKPKP.........................SNS
SlHKT1      87  MTILMFLGGEAFTSFLSPKLIKNKESKDKSFS........................NKDY
StHKT1      85  MTILMFLGGEAFTSFLRLKLIK.KENKDESFS........................CKDY
CaHKT1[2]   94  MTILMFLGGEAFTSFLRLQLMKLKLVKKHSVLQNNDAFD..............YSNSSNG
NtHKT1      94  ITMLMFLGGEAFTSFLRLQLMKSKLSMEKTVLQNKIETSNSI.....ISIDSHDSSNSPR
SiHKT1     113  LTILMLVGGEVFTSMLGLHLIRAKIYLQKNGLITHRTVS..............DSSISSD
DcHKT1     114  LTVLMLIGGEVFISIFKLQFSKSMISKNHIVYGRSSLDSNS..........EDKIDVNQI
GmHKT1     102  FTLLMFLGGEVFTSMLDLLFAGYKLQN...........KVSIN.......PFPPTNHANQ
CcHKT1      95  LTLLMFLGGEVFTSMLDLVLARYKVTTN.........PFSPTS.......KTGTPFKANQ
CaHKT1      99  LTFLMLAGGEVFTSMLQLFLARFSFTQNQCVKNEYSSTTLSLK.......HDPSHNQINQ
AhHKT1      89  ITFLMFFGGEVFISLLELVFARFNNNSFTSMKNDYSVKVSTNN.......SLPIKKVNDQ
SOHKT1     112  ITILMFIGGEVFTSMVGLHFSASKLKYTPLKARSRVNSVASLP............LPPES
SsHKT1     116  LTILMFIGGEVFTSMVGLHFSASKLVYTPLQARSRVNSVASLP............LPPES
HaHKT1      95  MTILMFIGGEIFTSLVGHYIRTFFVHDSPKDGNKVDFCDTQS............SFSDKI
LsHKT1     110  MTILMFIGGEVFTSMVGLLLRSCSRHSSIKEGGRVNDSMSDLG........IPPNFHDQL
EcHKT1     112  MTVLMFIGGEVFISLVGLHLRKSKLRWRIRTEDKVASADGN.....LCPSAPTNDIVDHI
VvHKT1     116  LTLLMFVGGEVFTSMVELHLRRSKLEKPWIAENKIDSVTSINGSAPLSPGKPFGQVELRV
HvHKT1      40  LTLLMLLGGEVFTSMLGMYFTYVKSKKKEAQAPHDDGAKVKPAP................
```

***Figure 23:*** Highly conserved amino acid motif after multiple sequence alignment of HKT1 with motif 3

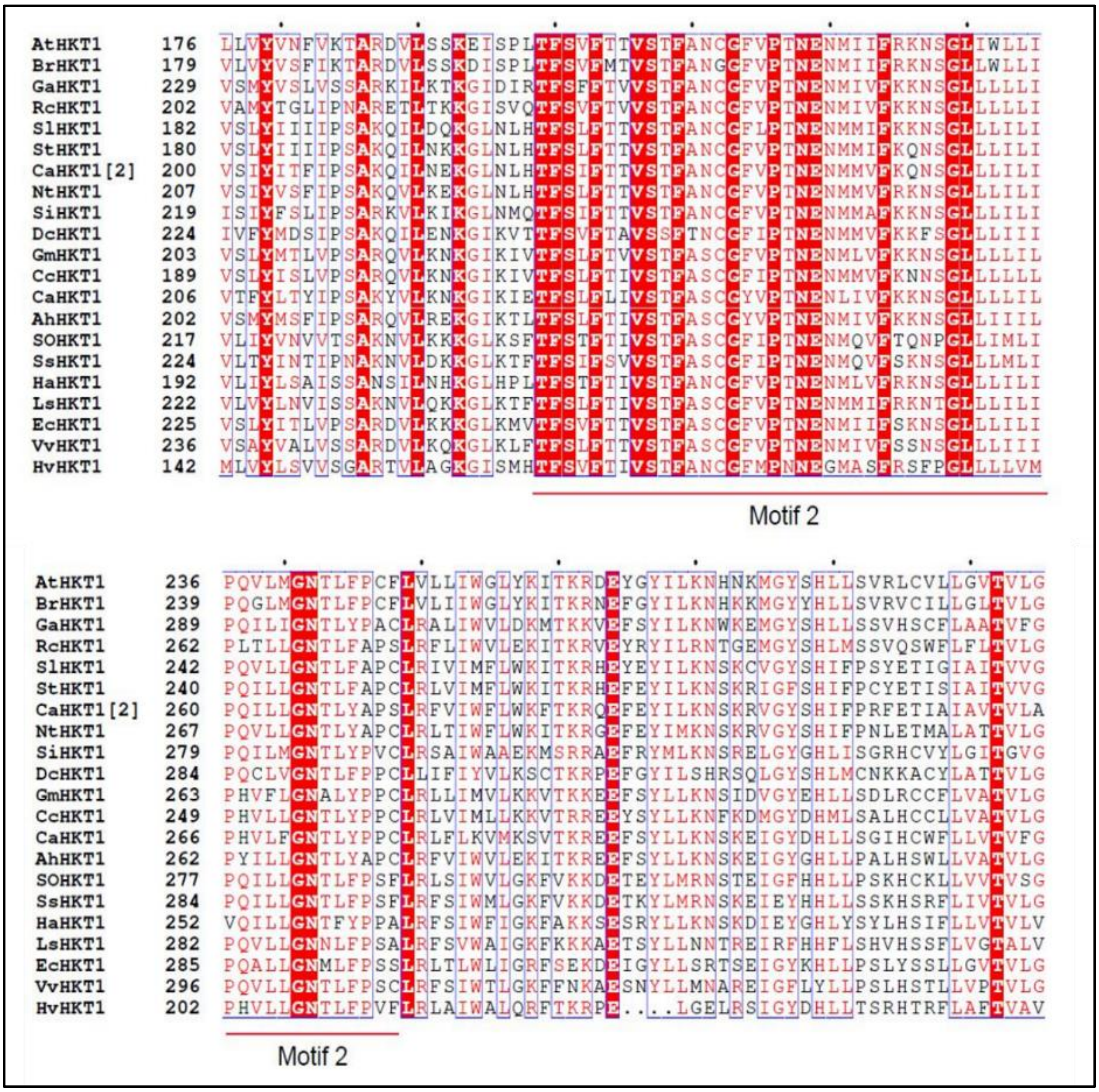
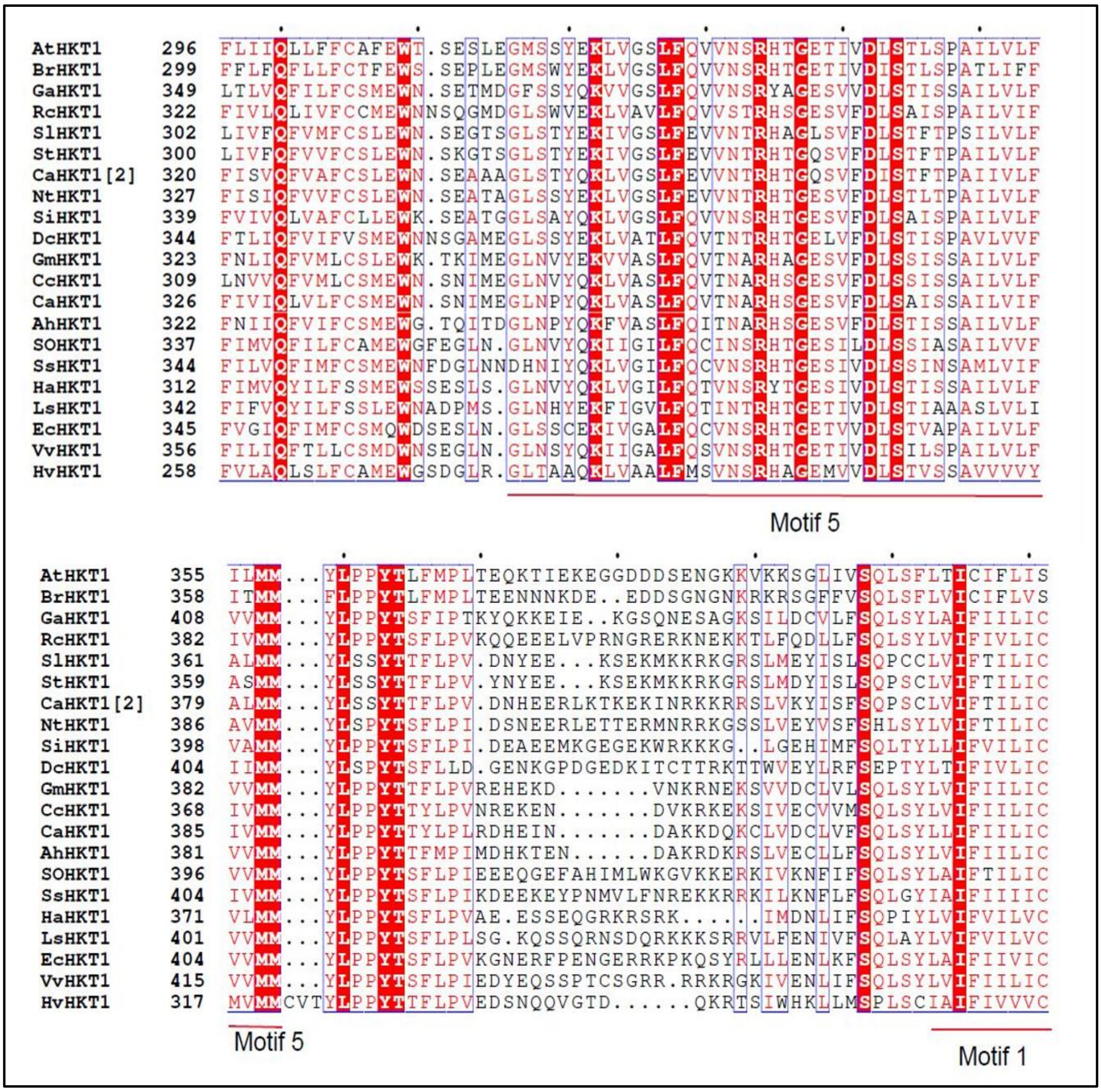


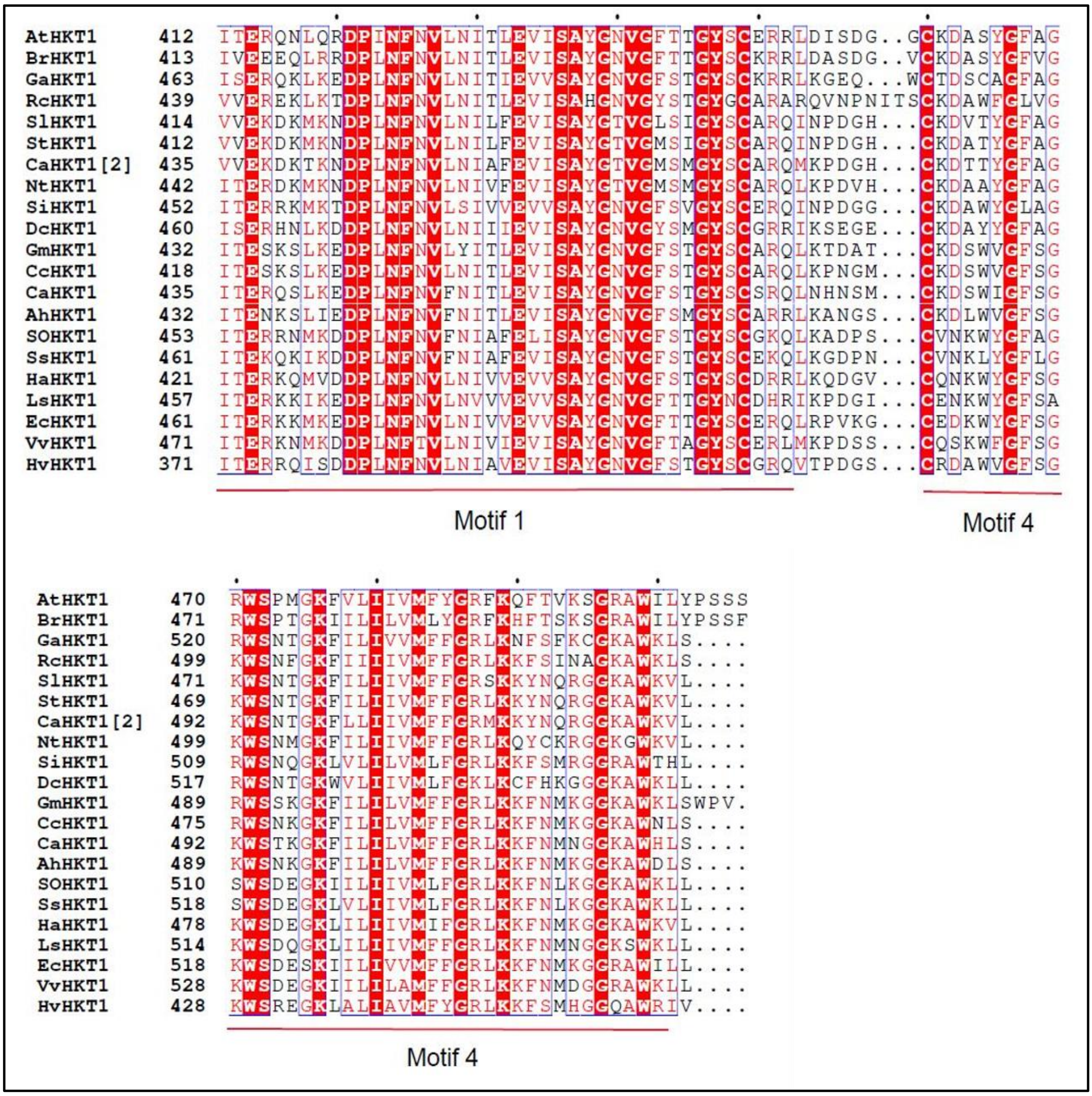
***Figure 24:*** Highly conserved amino acid motif after multiple sequence alignment of HKT1 with motif 2



```
AtHKT1      296 FLIIQLLFFCAFEWT.SESLEGMSSYEKLVGSLFQVVNSRHTGETIVDLSTLSPAILVLF
BrHKT1      299 FFLFQFLLFCTFEWS.SEPLEGMSWYEKLVGSLFQVVNSRHTGETIVDISTLSPATLIFF
GaHKT1      349 LTLVQFILFCSMEWN.SETMDGFSSYQKVVGSLFQVVNSRYAGESVVDLSTISSAILVLF
RcHKT1      322 FIVLQLIVFCCMEWNNSQGMDGLSWVEKLVAVLFQVVSTRHSGESVFDLSAISPAILVIF
SlHKT1      302 LIVFQFVMFCSLEWN.SEGTSGLSTYEKIVGSLFEVVNTRHAGLSVFDLSTFTPSILVLF
StHKT1      300 LIVFQFVVFCSLEWN.SKGTSGLSTYEKIVGSLFEVVNTRHTGQSVFDLSTFTPAILVLF
CaHKT1[2]   320 FISVQFVAFCSLEWN.SEAAAGLSTYQKLVGSLFEVVNTRHTGQSVFDISTFTPAIIVLF
NtHKT1      327 FISIQFVVFCSLEWN.SEATAGLSSYEKLVGSLFEVVNTRHTGESVFDLSTLTPAILVLF
SiHKT1      339 FVIVQLVAFCLLEWK.SEATGGLSAYQKLVGSLFQVVNSRHTGESVFDLSAISPAILVLF
DcHKT1      344 FTLIQFVIFVSMEWNNSGAMEGLSSYEKLVATLFQVTNTRHTGELVFDLSTISPAVLVVF
GmHKT1      323 FNLIQFVMLCSLEWK.TKIMEGLNVYEKVVASLFQVTNARHAGESVFDLSSISSAILVLF
CcHKT1      309 LNVVQFVMLCSMEWN.SNIMEGLNVYQKLVASLFQVTNARHSGESVFDLSSISSAILVLF
CaHKT1      326 FIVIQLVLFCSMEWN.SNIMEGLNPYQKLVASLFQVTNARHSGESVFDLSAISSAILVIF
AhHKT1      322 FNIIQFVIFCSMEWG.TQITDGLNPYQKFVASLFQITNARHSGESVFDLSTISSAILVLF
SOHKT1      337 FIMVQFILFCAMEWGFEGLN.GLNVYQKIIGILFQCINSRHTGESILDLSSIASAILVVF
SsHKT1      344 FILVQFIMFCSMEWNFDGLNNDHNIYQKLVGILFQCVNSRHTGESIVDLSSINSAMLVIF
HaHKT1      312 FIMVQYILFSSMEWSSESLS.GLNVYQKLVGILFQTVNSRYTGESIVDLSTISSAILVLF
LsHKT1      342 FIFVQYILFSSLEWNADPMS.GLNHYEKFIGVLFQTINTRHTGETIVDLSTIAAASLVLI
EcHKT1      345 FVGIQFIMFCSMQWDSESLN.GLSSCEKIVGALFQCVNSRHTGETVVDLSTVAPAILVLF
VvHKT1      356 FILIQFTLLCSMDWNSEGLN.GLNSYQKIIGALFQSVNSRHTGETIVDISILSPAILVLF
HvHKT1      258 FVLAQLSLFCAMEWGSDGLR.GLTAAQKLVAALFMSVNSRHAGEMVVDLSTVSSAVVVVY
```

```
AtHKT1      355 ILMM...YLPPYTLFMPLTEQKTIEKEGGDDDSENGKKVKKSGLIVSQLSFLTICIFLIS
BrHKT1      358 ITMM...FLPPYTLFMPLTEENNNKDE..EDDSGNGNKRKRSGFFVSQLSFLVICIFLVS
GaHKT1      408 VVMM...YLPPYTSFIPTKYQKKEIE..KGSQNESAGKSILDCVLFSQLSYLAIFIILIC
RcHKT1      382 IVMM...YLPPYTSFLPVKQQEEELVPRNGRERKNEKKTLFQDLLFSQLSYLVIFIVLIC
SlHKT1      361 ALMM...YLSSYTTFLPV.DNYEE...KSEKMKKRKGRSLMEYISLSQPCCLVIFTILIC
StHKT1      359 ASMM...YLSSYTTFLPV.YNYEE...KSEKMKKRKGRSLMDYISLSQPSCLVIFTILIC
CaHKT1[2]   379 ALMM...YLSSYTTFLPV.DNHEERLKTKEKINRKKRRSLVKYISFSQPSCLVIFTILIC
NtHKT1      386 AVMM...YLSPYTSFLPI.DSNEERLETTERMNRRKGSSLVEYVSFSHLSYLVIFTILIC
SiHKT1      398 VAMM...YLPPYTSFLPI.DEAEEMKGEGEKWRKKKG..LGEHIMFSQLTYLLIFVILIC
DcHKT1      404 IIMM...YLSPYTSFLLD.GENKGPDGEDKITCTTRKTTWVEYLRFSEPTYLTIFIVLIC
GmHKT1      382 VVMM...YLPPYTTFLPVREHEKD.......VNKRNEKSVVDCLVLSQLSYLVIFIILIC
CcHKT1      368 IVMM...YLPPYTTYLPVNREKEN.......DVKRKEKSIVECVVMSQLSYLVIFIILIC
CaHKT1      385 IVMM...YLPPYTTYLPLRDHEIN.......DAKKDQKCLVDCLVFSQLSYLLIFIILIC
AhHKT1      381 VVMM...YLPPYTTFMPIMDHKTEN......DAKRDKRSLVECLLFSQLSYLVIFIILIC
SOHKT1      396 VVMM...YLPPYTSFLPIEEEQGEFAHIMLWKGVKKERKIVKNFIFSQLSYLAIFTILIC
SsHKT1      404 IVMM...YLPPYTSFLPIKDEEKEYPNMVLFNREKKRRKILKNFLFSQLGYIAIFIIIIC
HaHKT1      371 VLMM...YLPPYTSFLPVAE.ESSEQGRKRSRK......IMDNLIFSQPIYLVIFVILVC
LsHKT1      401 VVMM...YLPPYTSFLPLSG.KQSSQRNSDQRKKKSRRVLFENIVFSQLAYLVIFVILVC
EcHKT1      404 VVMM...YLPPYTSFLPVKGNERFPENGERRKPKQSYRLLLENLKFSQLSYLAIFIIVIC
VvHKT1      415 VVMM...YLPPYTSFLPIEDYEQSSPTCSGRR.RRKRGKIVENLIFSQLSYLVIFIILIC
HvHKT1      317 MVMMCVTYLPPYTTFLPVEDSNQQVGTD......QKRTSIWHKLLMSPLSCIAIFIVVVC
```

***Figure 25:*** Highly conserved amino acid motif after multiple sequence alignment of HKT1 with motif 5 and motif 1



```
AtHKT1     412 ITERQNLQRDPINFNVLNITLEVISAYGNVGFTTGYSCERRLDISDG..GCKDASYGFAG
BrHKT1     413 IVEEEQLRRDPLNFNVLNITLEVISAYGNVGFTTGYSCKRRLDASDG..VCKDASYGFVG
GaHKT1     463 ISERQKLKEDPLNFNVLNITIEVVSAYGNVGFSTGYSCKRRLKGEQ...WCTDSCAGFAG
RcHKT1     439 VVEREKLKTDPLNFNVLNITLEVISAHGNVGYSTGYGCARARQVNPNITSCKDAWFGLVG
SlHKT1     414 VVEKDKMKNDPLNFNVLNILFEVISAYGTVGLSIGYSCARQINPDGH...CKDVTYGFAG
StHKT1     412 VVEKDKMKNDPLNFNVLNILFEVISAYGTVGMSIGYSCARQINPDGH...CKDATYGFAG
CaHKT1[2]  435 VVEKDKTKNDPLNFNVLNIAFEVISAYGTVGMSMGYSCARQMKPDGH...CKDTTYGFAG
NtHKT1     442 ITERDKMKNDPLNFNVLNIVFEVISAYGTVGMSMGYSCARQLKPDVH...CKDAAYGFAG
SiHKT1     452 ITERRKMKTDPLNFNVLSIVVEVVSAYGNVGFSVGYSCERQINPDGG...CKDAWYGLAG
DcHKT1     460 ISERHNLKDDPLNFNVLNIIIEVISAYGNVGYSMGYSCGRRIKSEGE...CKDAYYGFAG
GmHKT1     432 ITESKSLKEDPLNFNVLYITLEVISAYGNVGFSTGYSCARQLKTDAT...CKDSWVGFSG
CcHKT1     418 ITESKSLKEDPLNFNVLNITLEVISAYGNVGFSTGYSCARQLKPNGM...CKDSWVGFSG
CaHKT1     435 ITERQSLKEDPLNFNVFNITLEVISAYGNVGFSTGYSCSRQLNHNSM...CKDSWIGFSG
AhHKT1     432 ITENKSLIEDPLNFNVFNITLEVISAYGNVGFSMGYSCARRLKANGS...CKDLWVGFSG
SOHKT1     453 ITERRNMKDDPLNFNVFNIAFELISAYGNVGFSTGYSCGKQLKADPS...CVNKWYGFAG
SsHKT1     461 ITEKQKIKDDPLNFNVFNIAFEVISAYGNVGFSTGYSCEKQLKGDPN...CVNKLYGFLG
HaHKT1     421 ITERKQMVDDPLNFNVLNIVVEVVSAYGNVGFSTGYSCDRRLKQDGV...CQNKWYGFSG
LsHKT1     457 ITERKKIKEDPLNFNVLNVVVEVVSAYGNVGFTTGYNCDHRIKPDGI...CENKWYGFSA
EcHKT1     461 ITERKKMKEDPLNFNVLNIVVEVVSAYGNVGFTTGYSCERQLRPVKG...CEDKWYGFSG
VvHKT1     471 ITERKNMKDDPLNFTVLNIVIEVISAYGNVGFTAGYSCERLMKPDSS...CQSKWFGFSG
HvHKT1     371 ITERRQISDDPLNFNVLNIAVEVISAYGNVGFSTGYSCGRQVTPDGS...CRDAWVGFSG
```

```
AtHKT1     470 RWSPMGKFVLIIVMFYGRFKQFTVKSGRAWILYPSSS
BrHKT1     471 RWSPTGKIILILVMLYGRFKHFTSKSGRAWILYPSSF
GaHKT1     520 RWSNTGKFILIVVMFFGRLKNFSFKCGKAWKLS....
RcHKT1     499 KWSNFGKFIIIIVMFFGRLKKFSINAGKAWKLS....
SlHKT1     471 KWSNTGKFILIIVMFFGRSKKYNQRGGKAWKVL....
StHKT1     469 KWSNTGKFILIIVMFFGRLKKYNQRGGKAWKVL....
CaHKT1[2]  492 KWSNTGKFLLIIVMFFGRMKKYNQRGGKAWKVL....
NtHKT1     499 KWSNMGKFILIIVMFFGRLKQYCKRGGKGWKVL....
SiHKT1     509 RWSNQGKLVLILVMLFGRLKKFSMRGGRAWTHL....
DcHKT1     517 RWSNTGKWVLIIVMLFGKLKCFHKGGGKAWKLL....
GmHKT1     489 RWSSKGKFILILVMFFGRLKKFNMKGGKAWKLSWPV.
CcHKT1     475 RWSNKGKFILILVMFFGRLKKFNMKGGKAWNLS....
CaHKT1     492 KWSTKGKFILILVMFFGRLKKFNMNGGKAWHLS....
AhHKT1     489 KWSNKGKFILILVMFFGRLKKFNMKGGKAWDLS....
SOHKT1     510 SWSDEGKIILIIVMLFGRLKKFNLKGGKAWKLL....
SsHKT1     518 SWSDEGKLVLIIVMLFGRLKKFNLKGGKAWKLL....
HaHKT1     478 KWSDEGKLILIIVMIFGRLKKFNMKGGKAWKVL....
LsHKT1     514 KWSDQGKLILIIVMFFGRLKKFNMNGGKSWKLL....
EcHKT1     518 KWSDESKIILIVVMLFGRLKKFNMKGGRAWILL....
VvHKT1     528 KWSDEGKIILILAMFFGRLKKFNMDGGRAWKLL....
HvHKT1     428 KWSREGKLALIAVMFYGRLKKFSMHGGQAWRIV....
```

***Figure 26:*** Highly conserved amino acid motif after multiple sequence alignment of HKT1 with motif 1 and motif 4

***Table 10*****:** Instability index and Helix Percentage of SOS1

*This protein combination are unstable

| **SL.No.** | **Motif's Sequence's** | **Instability Index** | **Helix Percentage (%)** |
|---|---|---|---|
| 1. | A1, A2, A3, A5, A4 | 35.35 | 80 |
| 2. | A1, A4, A5, A3, A2 | 37.09 | 77 |
| 3. | A2, A3, A4, A5, A1 | 37.58 | 77 |
| 4. | A4, A1, A2, A3, A5 | 37.79 | 75 |
| 5. | A4, A2, A3, A1, A5 | 37.79 | 75 |
| 6. | A5, A2, A3, A1, A4 | 37.79 | 77 |
| 7. | A2, A3, A5, A1, A4 | 37.79 | 77 |
| 8. | A1, A4, A2, A3, A5 | 37.79 | 77 |
| 9. | A4, A5, A1, A2, A3 | 37.79 | 78 |
| 10. | A5, A1, A4, A2, A3 | 37.79 | 78 |
| 11. | A4, A1, A5, A2, A3 | 37.79 | 80 |
| 12. | A4, A5, A2, A3, A1 | 37.79 | 81 |
| 13. | A4, A2, A3, A5, A1 | 37.79 | 81 |
| 14. | A2, A3, A1, A4, A5 | 37.79 | 81 |
| 15. | A1, A4, A5, A2, A3 | 37.79 | 81 |
| 16. | A3, A4, A5, A1, A2 | 37.97 | 73 |
| 17. | A5, A3, A2, A1, A4 | 38.17 | 69 |
| 18. | A4, A3, A5, A2, A1 | 38.17 | 71 |
| 19. | A4, A2, A5, A1, A3 | 38.17 | 71 |
| 20. | A3, A2, A4, A1, A5 | 38.17 | 72 |
| 21. | A3, A1, A4, A2, A5 | 38.17 | 72 |
| 22. | A4, A3, A2, A5, A1 | 38.17 | 73 |
| 23. | A4, A3, A1, A5, A2 | 38.17 | 73 |
| 24. | A2, A5, A1, A4, A3 | 38.17 | 73 |
| 25. | A2, A4, A3, A5, A1 | 38.17 | 73 |
| 26. | A5, A2, A4, A3, A1 | 38.17 | 74 |
| 27. | A4, A5, A2, A1, A3 | 38.17 | 74 |
| 28 | A3, A5, A1, A4, A2 | 38.17 | 74 |

|  |  |  |  |
|---|---|---|---|
| 29. | A3, A2, A4, A5, A1 | 38.17 | 74 |
| 30. | A5, A1, A4, A3, A2 | 38.17 | 75 |
| 31. | A4, A5, A3, A1, A2 | 38.17 | 75 |
| 32. | A4, A1, A3, A5, A2 | 38.17 | 75 |
| 33. | A4, A1, A5, A3, A2 | 38.17 | 75 |
| 34. | A3, A1, A4, A5, A2 | 38.17 | 75 |
| 35. | A5, A3, A1, A4, A2 | 38.17 | 76 |
| 36. | A5, A3, A1, A2, A4 | 38.17 | 76 |
| 37. | A5, A1, A3, A2, A4 | 38.17 | 76 |
| 38. | A4, A2, A1, A3, A5 | 38.17 | 76 |
| 39. | A4, A1, A2, A5, A3 | 38.17 | 76 |
| 40. | A2, A4, A5, A1, A3 | 38.17 | 76 |
| 41. | A4, A5, A1, A3, A2 | 38.17 | 77 |
| 42. | A4, A5, A3, A2, A1 | 38.17 | 77 |
| 43. | A4, A3, A1, A2, A5 | 38.17 | 77 |
| 44. | A4, A2, A1, A5, A3 | 38.17 | 77 |
| 45. | A3, A5, A2, A1, A4 | 38.17 | 77 |
| 46. | A3, A2, A5, A1, A4 | 38.17 | 77 |
| 47. | A2, A5, A3, A1, A4 | 38.17 | 77 |
| 48. | A3, A2, A1, A4, A5 | 38.17 | 78 |
| 49. | A5, A3, A2, A4, A1 | 38.17 | 78 |
| 50. | A5, A2, A1, A4, A3 | 38.17 | 78 |
| 51. | A5, A1, A2, A4, A3 | 38.17 | 78 |
| 52. | A4, A2, A5, A3, A1 | 38.17 | 78 |
| 53 | A3, A5, A1, A2, A4 | 38.17 | 78 |
| 54. | A3, A1, A5, A2, A4 | 38.17 | 78 |
| 55. | A2, A4, A5, A3, A1 | 38.17 | 78 |
| 56. | A1, A5, A3, A2, A4 | 38.17 | 78 |
| 57. | A2, A4, A3, A1, A5 | 38.17 | 79 |
| 58. | A1, A4, A3, A2, A5 | 38.17 | 79 |
| 59. | A2, A1, A4, A5, A3 | 38.17 | 80 |
| 60. | A4, A3, A5, A1, A2 | 38.17 | 80 |

|  |  |  |  |
|---|---|---|---|
| 61. | A4, A1, A3, A2, A5 | 38.17 | 80 |
| 62. | A3, A5, A2, A4, A1 | 38.17 | 80 |
| 63. | A2, A4, A5, A1, A3 | 38.17 | 80 |
| 64. | A2, A4, A1, A3, A5 | 38.17 | 80 |
| 65. | A2, A1, A4, A3, A5 | 38.17 | 80 |
| 66. | A1, A4, A3, A5, A2 | 38.17 | 80 |
| 67. | A3, A1, A2, A4, A5 | 38.17 | 81 |
| 68. | A4, A3, A2, A1, A5 | 38.17 | 81 |
| 69. | A5, A2, A4, A1, A3 | 38.17 | 81 |
| 70. | A1, A2, A4, A3, A5 | 38.17 | 81 |
| 71. | A1, A5, A2, A4, A3 | 38.17 | 82 |
| 72. | A1, A4, A2, A5, A3 | 38.17 | 82 |
| 73. | A1, A3, A5, A2, A4 | 38.17 | 82 |
| 74. | A1, A3, A2, A4, A5 | 38.17 | 83 |
| 75. | A1, A2, A4, A5, A3 | 38.17 | 84 |
| 76. | A2, A3, A4, A1, A5 | 38.67 | 79 |
| 77. | A1, A2, A3, A4, A5 | 38.67 | 80 |
| 78. | A5, A2, A3, A4, A1 | 38.67 | 80 |
| 79. | A1, A5, A2, A3, A4 | 38.67 | 80 |
| 80. | A5, A1, A2, A3, A4 | 38.76 | 84 |
| 81. | A3, A4, A5, A2, A1 | 39.05 | 66 |
| 82. | A2, A5, A3, A4, A1 | 39.05 | 68 |
| 83. | A3, A4, A2, A1, A5 | 39.05 | 71 |
| 84. | A3, A4, A1, A5, A2 | 39.05 | 72 |
| 85. | A5, A1, A3, A4, A2 | 39.05 | 73 |
| 86. | A3, A4, A2, A5, A1 | 39.05 | 73 |
| 87. | A1, A5, A3, A4, A2 | 39.05 | 73 |
| 88. | A1, A3, A4, A2, A5 | 39.05 | 73 |
| 89. | A5, A3, A4, A2, A1 | 39.05 | 75 |
| 90. | A5, A2, A1, A3, A4 | 39.05 | 76 |
| 91. | A3, A4, A1, A2, A5 | 39.05 | 76 |
| 92. | A2, A5, A1, A3, A4 | 39.05 | 76 |

|  |  |  |  |
|---|---|---|---|
| 93. | A1, A2, A5, A3, A4 | 39.05 | 77 |
| 94. | A2, A1, A5, A3, A4 | 39.05 | 77 |
| 95. | A1, A3, A4, A5, A2 | 39.05 | 78 |
| 96. | A2, A1, A3, A4, A5 | 39.05 | 81 |
| 97. | A5, A3, A4, A1, A2 | 39.05 | 83 |
| 98. | A2, A3, A1, A5, A4 | 39.79 | 76 |
| 99. | A5, A4, A1, A2, A3 | 39.79 | 77 |
| 100. | A5, A4, A2, A3, A1 | 39.79 | 78 |
| 101. | A2, A3, A5, A4, A1 | 39.79 | 78 |
| 102. | A1, 45, A4, A2, A3 | 39.79 | 82 |
| 103. | A2, A1, A3, A5, A4 | 40.18* | 73 |
| 104. | A3, A1, A5, A4, A2 | 40.18* | 73 |
| 105. | A1, A5, A4, A3, A2 | 40.18* | 75 |
| 106. | A3, A5, A4, A1, A2 | 40.18* | 76 |
| 107. | A1, A3, A5, A4, A2 | 40.18* | 76 |
| 108. | A1, A2, A5, A4, A3 | 40.18* | 76 |
| 109. | A5, A4, A2, A1, A3 | 40.18* | 77 |
| 110. | A2, A5, A4, A3, A1 | 40.18* | 77 |
| 111. | A5, A4, A1, A3, A2 | 40.18* | 78 |
| 112. | A3, A2, A5, A4, A1 | 40.18* | 78 |
| 113. | A3, A1, A2, A5, A4 | 40.18* | 78 |
| 114. | A5, A4, A3, A2, A1 | 40.18* | 79 |
| 115. | A2, A1 A5, A4, A3 | 40.18* | 79 |
| 116. | A1, A3, A2, A5, A4 | 40.18* | 79 |
| 117. | A5, A4, A3, A1, A2 | 40.18* | 80 |
| 118. | A3, A2, A1, A5, A4 | 40.18* | 81 |
| 119. | A3, A5, A4, A2, A1 | 40.18* | 81 |
| 120. | A2, A5, A4, A1, A3 | 40.18* | 83 |

#Here **A1=** FPYDWNWKTSLLLGGLLSATDPVAVVALLKELGASKKLSTIIEGESLMND;

**A2=** VMTLGMFYAAFARTAFKGESQQSLHHFWEMVAYIANTLIFILSGVVIAEG;

**A3**= ANIDPDLLLAVFLPALLFESSFSMEVHQIKRCMAQMV LLAGPGV;

**A4=** PRKLVTYFTVERLESACYIAAFLRAHRIARRQLHDFIGDS;

**A5**=VQAAYWEMLDEG RITQTTANILMQSVDEALDLVS.

***Table 11***: Instability index and Helix Percentage of HKT1

*This protein combination are unstable

| **SL.No.** | **Motif's Sequence's** | **Instability Index** | **Helix Percentage (%)** |
|---|---|---|---|
| 1. | A1, A4, A5, A3, A2 | 22.69 | 57 |
| 2. | A1, A4, A5, A2, A3 | 23.01 | 61 |
| 3. | A4, A1, A5, A3, A2 | 23.01 | 57 |
| 4. | A3, A1, A4, A5, A2 | 23.01 | 60 |
| 5. | A5, A3, A2, A4, A1 | 23.01 | 62 |
| 6. | A3, A5, A1, A4, A2 | 23.05 | 59 |
| 7. | A5, A3, A1, A2, A4 | 23.05 | 58 |
| 8. | A2, A3, A1, A4, A5 | 23.14 | 59 |
| 9. | A4, A5, A3, A1, A2 | 23.14 | 60 |
| 10. | A4, A1, A2, A5, A3 | 23.14 | 55 |
| 11. | A2, A4, A5, A3, A1 | 23.14 | 54 |
| 12. | A1, A2, A4, A5, A3 | 23.14 | 48 |
| 13. | A2, A5, A3, A4, A1 | 23.14 | 44 |
| 14. | A5, A3, A4, A1, A2 | 23.14 | 56 |
| 15. | A5, A3, A1, A4, A2 | 23.18 | 55 |
| 16. | A2, A5, A3, A1, A4 | 23.18 | 52 |
| 17. | A1, A4, A2, A5, A3 | 23.18 | 42 |
| 18. | A4, A1, A5, A2, A3 | 23.34 | 56 |
| 19. | A4, A5, A2, A3, A1 | 23.34 | 56 |
| 20. | A3, A2, A4, A1, A5 | 23.34 | 63 |
| 21. | A3, A2, A4, A5, A1 | 23.34 | 56 |
| 22. | A4, A1, A3, A5, A2 | 23.34 | 60 |
| 23. | A4, A5, A1, A3, A2 | 23.34 | 57 |
| 24. | A4, A1, A3, A2, A5 | 23.34 | 60 |
| 25. | A3, A5, A2, A4, A1 | 23.34 | 61 |
| 26. | A5, A2, A4, A1, A3 | 23.34 | 50 |
| 27. | A1, A3, A2, A4, A5 | 23.34 | 60 |
| 28. | A5, A2, A3, A4, A1 | 23.34 | 58 |

|  |  |  |  |
|---|---|---|---|
| 29. | A3, A4, A1, A5, A2 | 23.34 | 66 |
| 30. | A1, A3, A4, A5, A2 | 23.34 | 64 |
| 31. | A5, A4, A1, A3, A2 | 23.34 | 61 |
| 32. | A3, A2, A5, A4, A1 | 23.34 | 60 |
| 33. | A5, A2, A3, A1, A4 | 23.38 | 56 |
| 34. | A3, A2, A5, A1, A4 | 23.38 | 49 |
| 35. | A1, A5, A3, A2, A4 | 23.38 | 52 |
| 36. | A2, A3, A4, A5, A1 | 23.46 | 45 |
| 37. | A4, A1, A2, A3, A5 | 23.46 | 58 |
| 38. | A4, A5, A1, A2, A3 | 23.46 | 61 |
| 39. | A3, A4, A5, A1, A2 | 23.46 | 62 |
| 40. | A2, A4, A5, A1, A3 | 23.46 | 56 |
| 41. | A2, A4, A5, A1, A3 | 23.46 | 56 |
| 42. | A2, A4, A1, A3, A5 | 23.46 | 55 |
| 43. | A3, A1, A2, A4, A5 | 23.46 | 63 |
| 44. | A2, A3, A4, A1, A5 | 23.46 | 50 |
| 45. | A1, A2, A3, A4, A5 | 23.46 | 58 |
| 46. | A3, A4, A1, A2, A5 | 23.46 | 60 |
| 47. | A5, A4, A1, A2, A3 | 23.46 | 61 |
| 48. | A2, A3, A5, A4, A1 | 23.46 | 58 |
| 49. | A3, A5, A4, A1, A2 | 23.46 | 59 |
| 50. | A2, A5, A4, A1, A3 | 23.46 | 55 |
| 51. | A2, A3, A5, A1, A4 | 23.50 | 50 |
| 52. | A1, A4, A2, A3, A5 | 23.50 | 53 |
| 53. | A5, A1, A4, A2, A3 | 23.50 | 56 |
| 54. | A3, A1, A4, A2, A5 | 23.50 | 58 |
| 55. | A4, A2, A5, A3, A1 | 23.50 | 49 |
| 56. | A1, A5, A3, A4, A2 | 23.50 | 55 |
| 57. | A1, A2, A5, A3, A4 | 23.50 | 39 |
| 58. | A5, A2, A1, A3, A4 | 23.52 | 52 |
| 59. | A2, A1, A4, A5, A3 | 23.64 | 49 |
| 60. | A5, A1, A3, A2, A4 | 23.70 | 64 |

|  |  |  |  |
|---|---|---|---|
| 61. | A3, A1, A5, A2, A4 | 23.70 | 58 |
| 62. | A1, A3, A5, A2, A4 | 23.70 | 62 |
| 63. | A1, A5, A2, A3, A4 | 23.70 | 54 |
| 64. | A1, A3, A2, A5, A4 | 23.70 | 59 |
| 65. | A1, A2, A3, A5, A4 | 23.82 | 51 |
| 66. | A4, A2, A3, A1, A5 | 23.82 | 52 |
| 67. | A4, A2, A3, A5, A1 | 23.82 | 52 |
| 68. | A4, A2, A5, A1, A3 | 23.82 | 57 |
| 69. | A3, A5, A1, A2, A4 | 23.82 | 62 |
| 70. | A5, A1, A2, A3, A4 | 23.82 | 58 |
| 71. | A5, A1, A3, A4, A2 | 23.82 | 66 |
| 72. | A3, A4, A2, A5, A1 | 23.82 | 58 |
| 73. | A1, A3, A4, A2, A5 | 23.82 | 58 |
| 74. | A2, A5, A1, A3, A4 | 23.82 | 49 |
| 75. | A2, A3, A1, A5, A4 | 23.82 | 56 |
| 76. | A5, A4, A2, A3, A1 | 23.82 | 60 |
| 77. | A1, 45, A4, A2, A3 | 23.82 | 49 |
| 78. | A3, A1, A5, A4, A2 | 23.82 | 58 |
| 79. | A1, A3, A5, A4, A2 | 23.82 | 60 |
| 80. | A3, A1, A2, A5, A4 | 23.82 | 57 |
| 81. | A4, A5, A3, A2, A1 | 23.84 | 56 |
| 82. | A3, A2, A1, A4, A5 | 23.84 | 52 |
| 83. | A5, A3, A2, A1, A4 | 23.88 | 45 |
| 84. | A4, A5, A2, A1, A3 | 24.16 | 52 |
| 85. | A3, A4, A5, A2, A1 | 24.16 | 57 |
| 86. | A3, A5, A2, A1, A4 | 24.20 | 55 |
| 87. | A2, A1, A3, A4, A5 | 24.28 | 44 |
| 88. | A4, A2, A1, A5, A3 | 24.32 | 58 |
| 89. | A5, A3, A4, A2, A1 | 24.32 | 60 |
| 90. | A2, A1, A5, A3, A4 | 24.32 | 52 |
| 91. | A3, A2, A1, A5, A4 | 24.52 | 59 |
| 92. | A4, A2, A1, A3, A5 | 24.65 | 49 |

| | | | |
|---|---|---|---|
| 93. | A3, A4, A2, A1, A5 | 24.65 | 64 |
| 94. | A2, A1, A3, A5, A4 | 24.65 | 55 |
| 95. | A5, A4, A2, A1, A3 | 24.65 | 59 |
| 96. | A3, A5, A4, A2, A1 | 24.65 | 57 |
| 97. | A5, A1, A4, A3, A2 | 24.77 | 59 |
| 98. | A1, A4, A3, A2, A5 | 24.77 | 49 |
| 99. | A1, A4, A3, A5, A2 | 24.77 | 59 |
| 100. | A2, A5, A1, A4, A3 | 24.89 | 46 |
| 101. | A4, A3, A2, A5, A1 | 25.09 | 54 |
| 102. | A4, A3, A1, A5, A2 | 25.09 | 59 |
| 103. | A5, A2, A4, A3, A1 | 25.09 | 58 |
| 104. | A1, A5, A2, A4, A3 | 25.09 | 44 |
| 105. | A1, A5, A4, A3, A2 | 25.09 | 53 |
| 106. | A2, A4, A3, A5, A1 | 25.22 | 50 |
| 107. | A4, A3, A1, A2, A5 | 25.22 | 60 |
| 108. | A5, A1, A2, A4, A3 | 25.22 | 62 |
| 109. | A2, A4, A3, A1, A5 | 25.22 | 50 |
| 110. | A4, A3, A5, A1, A2 | 25.22 | 63 |
| 111. | A1, A2, A4, A3, A5 | 25.22 | 54 |
| 112. | A1, A2, A5, A4, A3 | 25.22 | 51 |
| 113. | A2, A5, A4, A3, A1 | 25.22 | 55 |
| 114. | A5, A4, A3, A1, A2 | 25.22 | 57 |
| 115. | A5, A2, A1, A4, A3 | 25.59 | 54 |
| 116. | A2, A1, A4, A3, A5 | 25.72 | 51 |
| 117. | A4, A3, A5, A2, A1 | 25.92 | 62 |
| 118. | A4, A3, A2, A1, A5 | 25.92 | 60 |
| 119. | A5, A4, A3, A2, A1 | 25.92 | 64 |
| 120. | A2, A1 A5, A4, A3 | 26.04 | 52 |

#Here **A1=** LVIFIILICITERKKLKEDPLNFNVLNITLEVISAYGNVGFSTGYSCARQ;

**A2=** TFSLFTTVSTFANCGFYPTNENMIYFKKNSGLLLILIPQVLGNTLFPPC;

**A3=** RPKDLDLFFTSVSATTVSSMSTVEMEVFSNTQLIFLTILMFLGGEVFTSM;

**A4=** CKDAWYGFSGKWSNTGKFILIIVMFFFGRLKKFNMKGGKAWK;

**A5=** GLSS YQKLVGSLFQVVNSRHTGESVFDLSTISPAILVLFVVMM

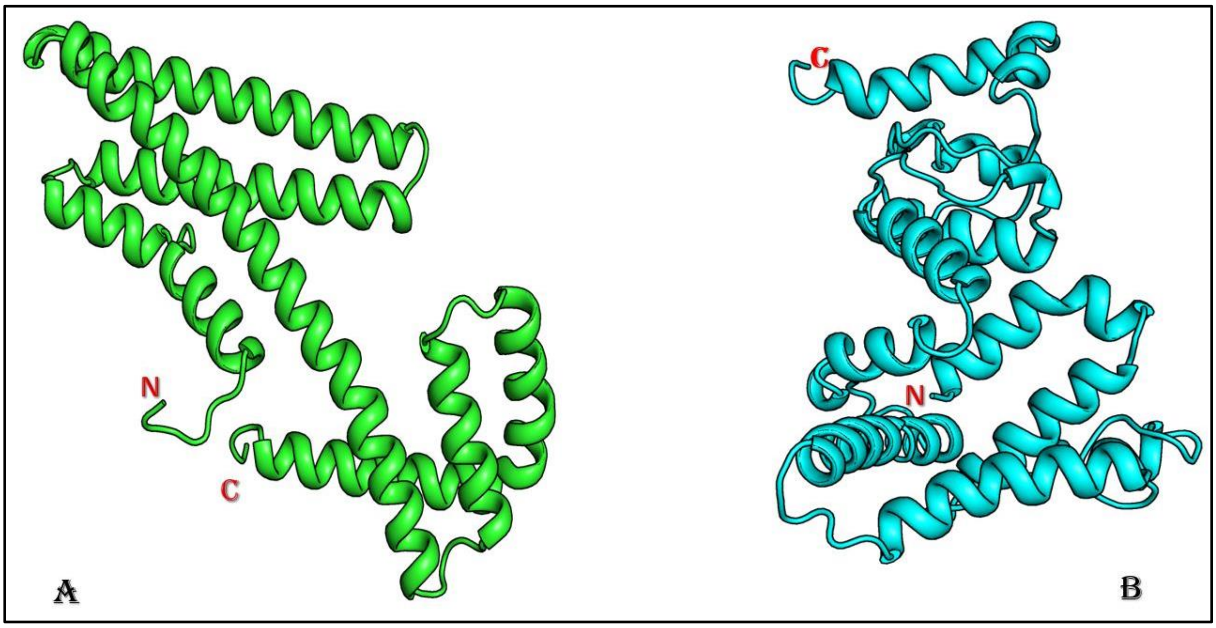


***Figure 27*: Peptide design of secondary structure before molecular dynamic simulation.**
**A.** SOS1; **B.** HKT1

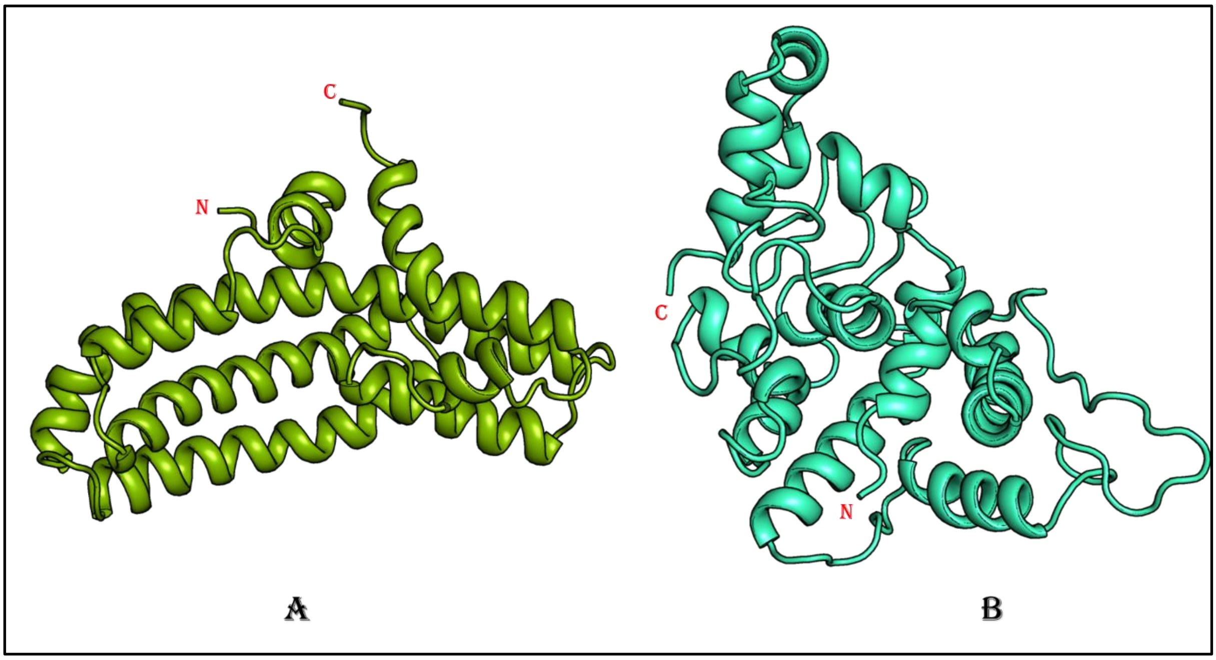


***Figure 28*: Peptide design of secondary structure after molecular dynamic simulation.**
**A.** SOS1; **B.** HKT1

***Table 12*****:** Antigenicity of SOS1 by SVMTrip

*Recommend Score

| **Sl.No.** | **Location** | **Epitope** | **Score** |
|---|---|---|---|
| 1. | 117-136 | VFDLSTISPAILVLFVVMMR | 1.0* |
| 2. | 46-65 | SCARQCKDAWYGFSGKWSNT | 0.84 |
| 3. | 211-230 | FKKNSGLLLILIPQVLGNTL | 0.708 |
| 4. | 74-93 | MFFFGRLKKFNMKGGKAWKG | 0.571 |

***Table 13*****:** Antigenicity of HKT1 by SVMTrip

*Recommend Score

| **Sl.No.** | **Location** | **Epitope** | **Score** |
|---|---|---|---|
| 1. | 4-23 | DWNWKTSLLLGGLLSATDPV | 1.0* |
| 2. | 52-71 | MTLGMFYAAFARTAFKGESQ | 0.773 |
| 3. | 84-103 | IANTLIFILSGVVIAEGANI | 0.623 |
| 4. | 194-213 | ACYICAAFLRAHRIARRQLH | 0.622 |
| 5. | 143-162 | GVVQAAYWEMLDEGRITQTT | 0.474 |

***Table 14*****:** Antigenicity of SOS1 by ABCPred

*Recommend Score

| **Sl.No.** | **Rank** | **Sequence** | **Start position** | **Score** |
|---|---|---|---|---|
| 1. | 1 | AAYWEMLDEGRITQTT | 147 | 0.84* |
| 2. | 2 | STIIEGESLMNDVMTL | 39 | 0.83 |
| 3. | 3 | LLSATDPVAVVALLKE | 16 | 0.82 |

|  |  |  |  |  |
|---|---|---|---|---|
| 4. | 4 | AHRIARRQLHDFIGDS | 204 | 0.8 |
| 5. | 4 | VDEALDLVSPRKLVTY | 170 | 0.8 |
| 6. | 5 | KGESQQSLHHFWEMVA | 67 | 0.79 |
| 7. | 6 | VIAEGANIDPDLLLAV | 96 | 0.78 |
| 8. | 7 | FYAAFARTAFKGESQQ | 57 | 0.77 |
| 9. | 7 | HQIKRCMAQMVLLAGP | 127 | 0.77 |
| 10. | 8 | EGRITQTTANILMQSV | 155 | 0.75 |
| 11. | 9 | DWNWKTSLLLGGLLSA | 4 | 0.69 |
| 12. | 9 | ACYICAAFLRAHRIAR | 194 | 0.69 |
| 13. | 9 | MVLLAGPGVVQAAYWE | 136 | 0.69 |
| 14. | 10 | VSPRKLVTYFTVERLE | 177 | 0.65 |

***Table 15*:** Antigenicity of HKT1 by ABCPred

*Recommend Score

| **Sl.No.** | **Rank** | **Sequence** | **Start position** | **Score** |
|---|---|---|---|---|
| 1. | 1 | TVSSMSTVEMEVFSNT | 151 | 0.91* |
| 2. | 2 | IFIILICITERKKLKE | 3 | 0.88 |
| 3. | 3 | AWYGFSGKWSNTGKFI | 54 | 0.87 |
| 4. | 4 | TGYSCARQCKDAWYGF | 43 | 0.86 |
| 5. | 5 | GNVGFSTGYSCARQCK | 37 | 0.85 |
| 6. | 6 | CITERKKLKEDPLNFN | 9 | 0.84 |
| 7. | 6 | GGEVFTSMTFSLFTTV | 178 | 0.84 |
| 8. | 7 | FVVMMRPKDLDLFFTS | 131 | 0.79 |
| 9. | 8 | YPTNENMIYFKKNSGL | 202 | 0.78 |
| 10. | 8 | FFTSVSATTVSSMSTV | 143 | 0.78 |
| 11. | 9 | GKAWKGLSSYQKLVGS | 88 | 0.76 |
| 12. | 10 | TTVSTFANCGFYPTNE | 191 | 0.75 |
| 13. | 11 | VVNSRHTGESVFDLST | 107 | 0.73 |

|  |  |  |  |  |
|---|---|---|---|---|
| 14. | 12 | LLLILIPQVLGNTLFP | 217 | 0.7 |
| 15. | 13 | KKFNMKGGKAWKGLSS | 81 | 0.67 |
| 16. | 13 | LEVISAYGNVGFSTGY | 30 | 0.67 |
| 17. | 14 | GKFILIIVMFFFGRLK | 66 | 0.66 |

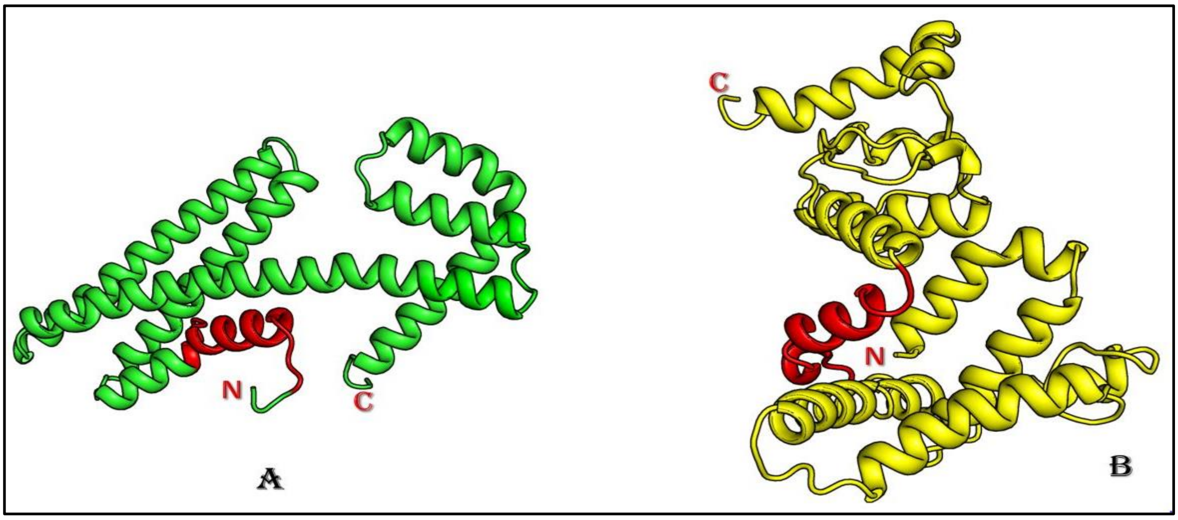


***Figure 29:*** Antigenicity mapping of highest score **A.** SOS1 and **B.** HKT1 by SVMTrip server

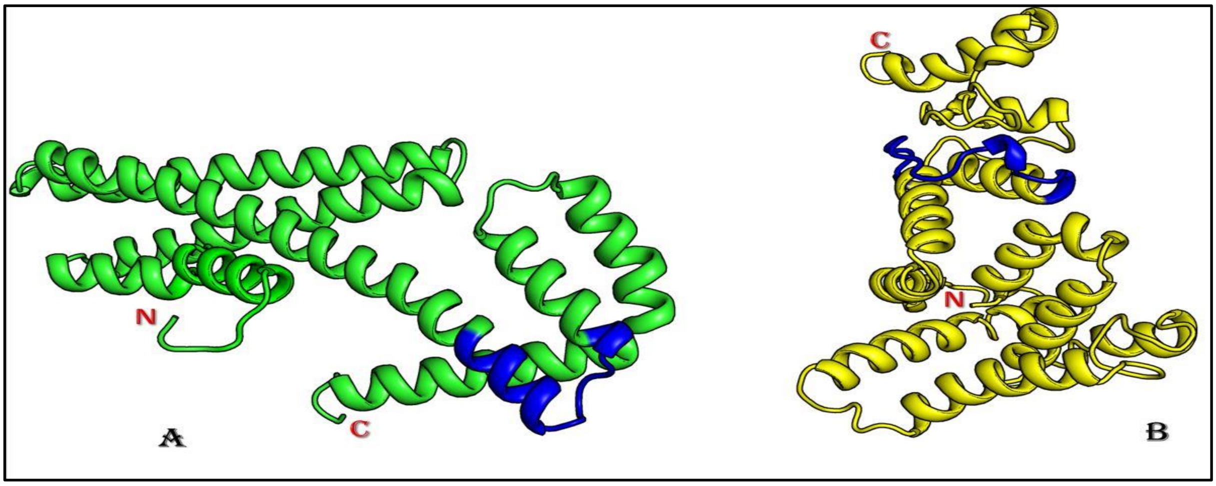


***Figure 30:*** Antigenicity mapping of highest score **A.** SOS1 and **B.** HKT1 by ABCPred server

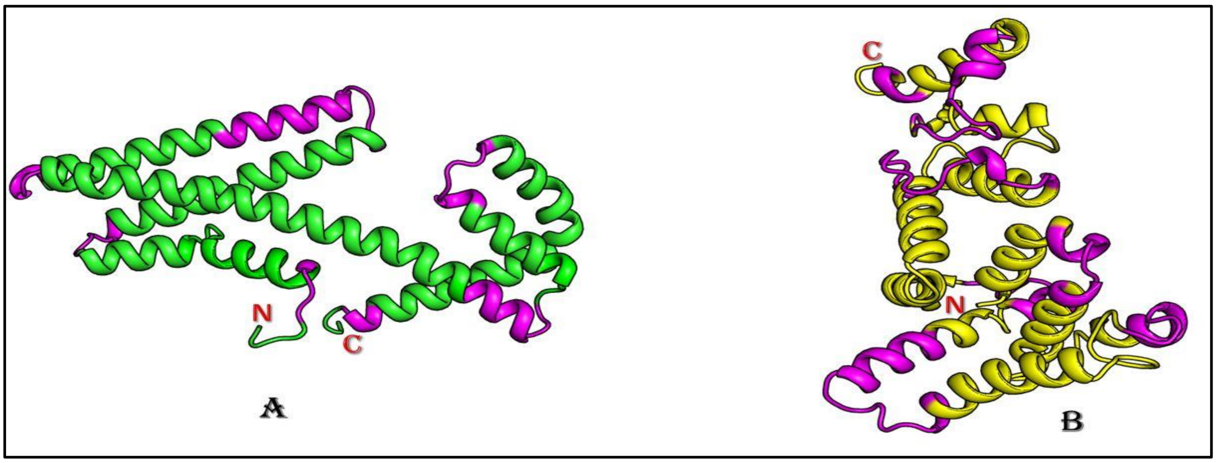


***Figure 31:*** Antigenicity mapping of highest score **A.** SOS1 and **B.** HKT1 by BepiPred server

***Table 16*****:** Shows Physio-chemical properties of SOS1 by Bcepred server

| **Sl.No.** | **Sequence** | **Different properties** | **Selected Sequence** |
|---|---|---|---|
| 1. | FPYDWNWKTSLLLGGLLSATDPVAVVALLKELGASKKLSTIIEGESLMNDVMTLGMFYAAFARTAFKGESQQSLHHFWEMVAYIANTLIFILSGVVIAEGANIDPDLLLAVFLPALLFESSFSMEVHQIKRCMAQMVLLAGPGVVQAAYWEMLDEGRITQTTANILMQSVDEALDLVSPRKLVTYFTVERLESACYICAAFLRAHRIARRQLHDFIGDS | Hydrophilicity | KGESQQS |

| | | | |
|---|---|---|---|
| 2. | FPYDWNWKTSLLLGGLLSATDPVAVVALLKELGASKKLSTIIEGESLMNDVMTLGMFYAAFARTAFKGESQQSLHHFWEMVAYIANTLIFILSGVVIAEGANIDPDLLLAVFLPALLFESSFSMEVHQIKRCMAQMVLLAGPGVVQAAYWEMLDEGRITQTTANILMQSVDEALDLVSPRKLVTYFTVERLESACYICAAFLRAHRIARRQLHDFIGDS | Flexibility | LGASKKL/<br>TAFKGESQ |
| 3. | FPYDWNWKTSLLLGGLLSATDPVAVVALLKELGASKKLSTIIEGESLMNDVMTLGMFYAAFARTAFKGESQQSLHHFWEMVAYIANTLIFILSGVVIAEGANIDPDLLLAVFLPALLFESSFSMEVHQIKRCMAQMVLLAGPGVVQAAYWEMLDEGRITQTTANILMQSVDEALDLVSPRKLVTYFTVERLESACYICAAFLRAHRIARRQLHDFIGDS | Accessibility | PYDWNWKTS/<br>KELGASKKLST/<br>RTAFKGESQQSLH/<br>EVHQIKR/<br>DEGRITQT/<br>DLVSPRKLVTY/<br>RAHRIARRQLHDF |
| 5. | FPYDWNWKTSLLLGGLLSATDPVAVVALLKELGASKKLSTIIEGESLMNDVMTLGMFYAAFARTAFKGESQQSLHHFWEMVAYIANTLIFILSGVVIAEGANIDPDLLLAVFLPALLFESSFSMEVHQIKRCMAQMVLLAGPGVVQAAYWEMLDEGRITQTTANILMQSVDEALDLVSPRKLVTYFTVERLESACYICAAFLRAHRIARRQLHDFIGDS | Polarity | RTAFKGE/<br>SLHHFWEMV/<br>MEVHQIKRCM/<br>EMLDEGR/<br>FTVERLESAC/<br>AFLRAHRIARRQLHDF |

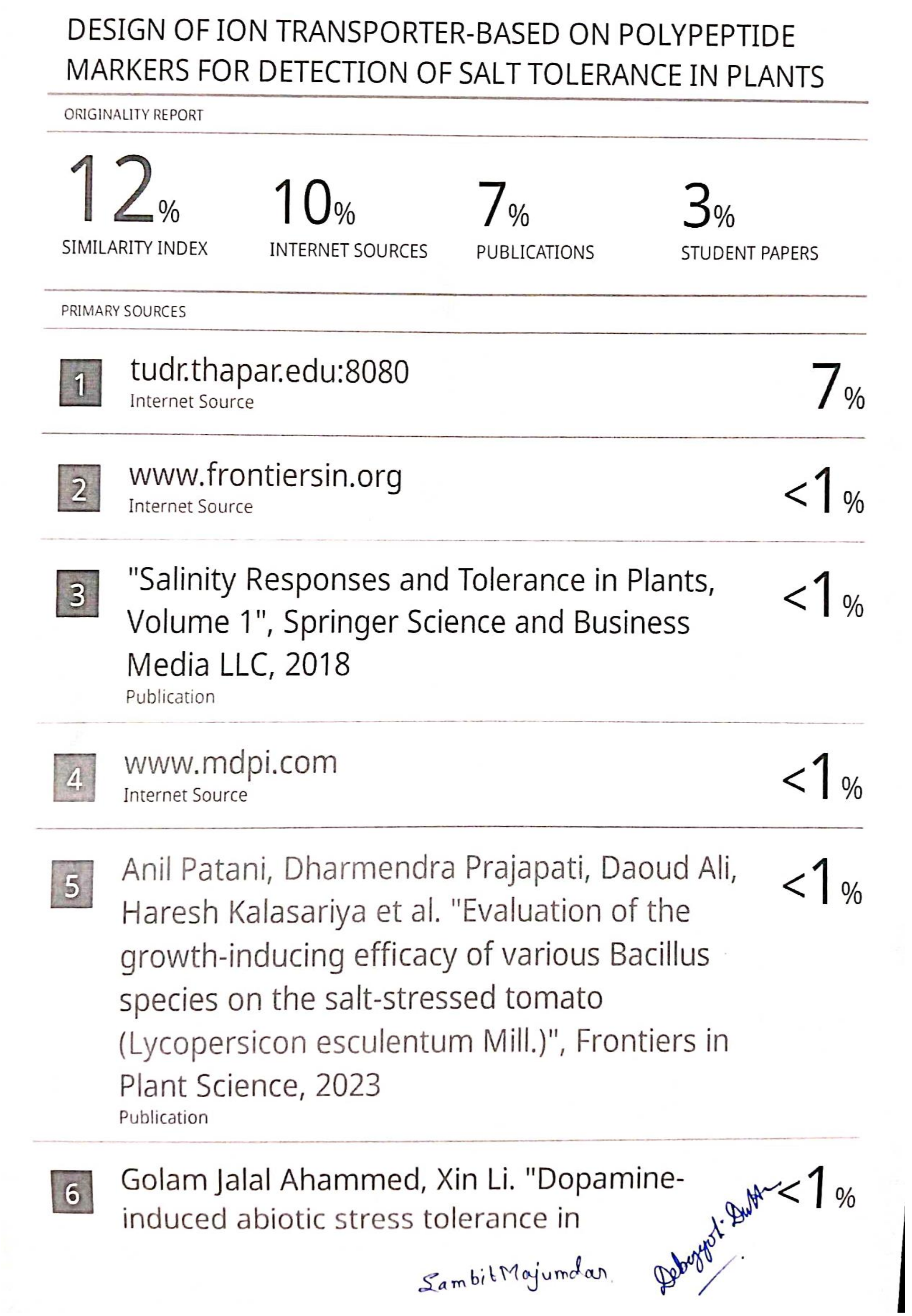
| | | | |
|---|---|---|---|
| 6. | FPYDWNWKTSLLLGGLLSATDPVAVVALLKELGASKKLSTIIEGESLMNDVMTLGMFYAAFARTAFKGESQQSLHHFWEMVAYIANTLIFILSGVVIAEGANIDPDLLLAVFLPALLFESSFSMEVHQIKRCMAQMVLLAGPGVVQAAYWEMLDEGRITQTTANILMQSVDEALDLVSPRKLVTYFTVERLESACYICAAFLRAHRIARRQLHDFIGDS | Antigenic propensity | SLLLGGLLS/<br>TLIFILSGVVI/<br>IDPDLLL/<br>VHQIKRC/<br>LDLVSPRKLVTYFTVE |

***Table 17*:** Shows Physio-chemical properties of HKT1 by Bcepred server

| **Sl.No.** | **Sequence** | **Different properties** | **Selected Sequence** |
|---|---|---|---|
| 1. | LVIFIILICITERKKLKEDPLNFNVLNITLEVISAYGNVGFSTGYSCARQCKDAWYGFSGKWSNTGKFILIIVMFFFGRLKKFNMKGGKAWKGLSSYQKLVGSLFQVVNSRHTGESVFDLSTISPAILVLFVVMMRPKDLDLFFTSVSATTVSSMSTVEMEVFSNTQLIFLTILMFLGGEVFTSMTFSLFTTVSTFANCGFYPTNENMIYFKKNSGLLLILIPQVLGNTLFPPC | Hydrophilicity | ARQCKDA/<br>NSRHTGES |
| 2. | LVIFIILICITERKKLKEDPLNFNVLNITLEVISAYGNVGFSTGYSCARQCKDAWYGFSGKWSNTGKFILIIVMFFFGRLKKFNMKGGKAWKGLSSYQKLVGSLFQVVNSRHTGESVFDLSTISPAILVLFVVMMRPKDLDLFFTSVSATTVSSMSTVEMEVFSNTQLIFLTILMFLGGEVFTSMTFSLFT | Flexibility | ICITERKKLK/<br>YGFSGKW/<br>KKFNMKG/<br>QVVNSRHTG/<br>MIYFKKNS |

| | | | |
|---|---|---|---|
| | TVSTFANCGFYPTNENMIYFKKNSGLLLILIPQVLGNTLFPPC | | |
| 3. | LVIFIILICITERKKLKEDPLNFNVLNITLEVISAYGNVGFSTGYSCARQCKDAWYGFSGKWSNTGKFILIIVMFFFGRLKKFNMKGGKAWKGLSSYQKLVGSLFQVVNSRHTGESVFDLSTISPAILVLFVVMMRPKDLDLFFTSVSATTVSSMSTVEMEVFSNTQLIFLTILMFLGGEVFTSMTFSLFTTVSTFANCGFYPTNENMIYFKKNSGLLLILIPQVLGNTLFPPC | Accessibility | CITERKKLKEDPLNFN/ ARQCKDAW/ KWSNTGK/ GRLKKFNMKGGKAWK/ NSRHTGES/ MMRPKDLDL/ FYPTNENMIYFKKNSG |
| 4. | LVIFIILICITERKKLKEDPLNFNVLNITLEVISAYGNVGFSTGYSCARQCKDAWYGFSGKWSNTGKFILIIVMFFFGRLKKFNMKGGKAWKGLSSYQKLVGSLFQVVNSRHTGESVFDLSTISPAILVLFVVMMRPKDLDLFFTSVSATTVSSMSTVEMEVFSNTQLIFLTILMFLGGEVFTSMTFSLFTTVSTFANCGFYPTNENMIYFKKNSGLLLILIPQVLGNTLFPPC | Exposed surface | CITERKKLKEDPL/ LKKFNMKG |
| 5. | LVIFIILICITERKKLKEDPLNFNVLNITLEVISAYGNVGFSTGYSCARQCKDAWYGFSGKWSNTGKFILIIVMFFFGRLKKFNMKGGKAWKGLSSYQKLVGSLFQVVNSRHTGESVFDLSTISPAILVLFVVMMRPKDLDLFFTSVSATTVSSMSTVEMEVFSNTQLIFLTILMFLGGEVFTSMTFSLFTTVSTFANCGFYPTNENMIYFKKNSGLLLILIPQVLGNTLFPPC | Polarity | ICITERKKLKEDPL/ GRLKKFNMKG/ NSRHTGESV/ MRPKDLDL |

| | | | |
|---|---|---|---|
| 6. | LVIFIILICITERKKLKEDPLNFNVLNITLEVISAYGNVGFSTGYSCARQCKDAWYGFSGKWSNTGKFILIIVMFFFGRLKKFNMKGGKAWKGLSSYQKLVGSLFQVVNSRHTGESVFDLSTISPAILVLFVVMMRPKDLDLFFTSVSATTVSSMSTVEMEVFSNTQLIFLTILMFLGGEVFTSMTFSLFTTVSTFANCGFYPTNENMIYFKKNSGLLLILIPQVLGNTLFPPC | Antigenic propensity | LVIFIILICITE/ ITLEVISA/ KFILIIVMFFF/ LSSYQKLVGSLFQVVNS/ FDLSTI/ ILVLFVVMMR/ LFFTSVS/ QLIFLTILMFL/ FSLFTTV/ NSGLLLILIPQVLG |



# DESIGN OF ION TRANSPORTER-BASED ON POLYPEPTIDE MARKERS FOR DETECTION OF SALT TOLERANCE IN PLANTS

ORIGINALITY REPORT

| 12% | 10% | 7% | 3% |
|---|---|---|---|
| SIMILARITY INDEX | INTERNET SOURCES | PUBLICATIONS | STUDENT PAPERS |

PRIMARY SOURCES

| | | |
|---|---|---|
| 1 | tudr.thapar.edu:8080<br>Internet Source | 7% |
| 2 | www.frontiersin.org<br>Internet Source | <1% |
| 3 | "Salinity Responses and Tolerance in Plants, Volume 1", Springer Science and Business Media LLC, 2018<br>Publication | <1% |
| 4 | www.mdpi.com<br>Internet Source | <1% |
| 5 | Anil Patani, Dharmendra Prajapati, Daoud Ali, Haresh Kalasariya et al. "Evaluation of the growth-inducing efficacy of various Bacillus species on the salt-stressed tomato (Lycopersicon esculentum Mill.)", Frontiers in Plant Science, 2023<br>Publication | <1% |
| 6 | Golam Jalal Ahammed, Xin Li. "Dopamine-induced abiotic stress tolerance in | <1% |

Sambit Majumdar Debjyoti Dutta