Computational prediction of the potential L-type Lectin Receptor Kinases in Extracellular Adenosine Triphosphate binding

Pham Vy H.[1,2], Tran Vu Q.[1,2] and Pham An Q.[1,2*]
1. Faculty of Biology and Biotechnology, University of Science, Ho Chi Minh city, VIETNAM
2. Vietnam National University, Ho Chi Minh city, VIETNAM
*pqan@hcmus.edu.vn

Abstract

Plants can perceive extracellular adenosine triphosphate (eATP) as a signaling molecule via purinergic 2 kinase (P2K) receptors. However, (1) how eATP binds to these receptors is not well understood and (2) apart from AtP2K1 and AtP2K2, which have been experimentally confirmed in Arabidopsis thaliana, what other plant species besides Brassicaceae perceive eATP? Our report visualizes eATP in active binding sites of P2K1 and P2K2 compared with mutant p2k1^{His99Ala} and p2k2^{His99Ala}. Tomato, Solanum lycopersicum, was used to search for homologous eATP receptors outside Brassicaceae. Given that His99 at the active binding site is a necessary condition, eight SlLecRKs, conserving His99 with AtP2Ks, were modeled and examined for their ability to interact with eATP. Our model shows that eATP binds to another favorable site if the receptor active binding site is not compatible.

In addition, Solyc09g012000 and Solyc09g011060 are able to interact with eATP at energies of -8.187 Kcal/mol and -8.306 Kcal/mol respectively. Our results show the potential of computational 3D modeling in explaining how ligands bind to their receptors, as well as predicting receptor homologues.

Keywords: P2Ks, *Sl*LecRK, extracellular ATP, computational modeling, receptor.

Introduction

Adenosine triphosphate (ATP) is known as the energy molecule for all living organisms on earth. In 1962, Burnstock accidentally discovered and demonstrated that extracellular ATP can act as a neurotransmitter[11]. After 30 years of skepticism about Burnstock's discovery, the first extracellular ATP (eATP) receptors were isolated from animals, demonstrating the existence of this interesting signaling pathway[1]. The discovery of two families of eATP receptors in animals, P2X and P2Y, has also opened up a better understanding of many growth, development and stress response processes in animal cells[24].

From here, drugs that act on these receptors, are also studied to fight against cancer, inflammatory responses or neurological diseases[6]. Studies demonstrating the response of plants to eATP have been studied since the 1970s[9] and have received much attention in the 2000s.

However, proteins similar to P2Xs and P2Ys are completely absent from the genomes of plant species[22]. With a strategy of screening individuals that do not respond to eATP using random mutant populations and genome sequencing methods on the model plant Arabidopsis thaliana, Choi et al[5] announced the discovery of DORN1 (DOes not Response to Nucleotide 1) as the receptor for receiving eATP[5]. DORN1, later named P2K1, is a member of group I of a 43-member protein family, divided into 11 groups in Arabidopsis thaliana called L-type LecRK (Lectin receptor like kinase)[2]. eATP has been shown to be a damage associated molecular patterns (DAMP) signaling molecule in both animals and plants[6,21,22,24]. DAMP molecules are involved in plant defense processes such as insect defense, wound response and thereby stimulating the production of secondary metabolites[3,21].

Understanding the role of these processes helps to control the formation of secondary metabolites in economically valuable plants such as tomatoes (*Solanum lycopersicum*). While there are at least 11 P2Xs and 8 P2Ys involved in a wide range of biological processes in animal cells; with the increasing number of reports on the role of eATP in plants, the hypothesis of the existence of multiple eATP receptors in plants was strengthened when a second receptor was discovered by Pham et al[17].

However, both receptors were found to belong to group I of the L-LecRLK family, which is only found in the Brassicaceae family. Studies on the response in plant species other than the model plant, Arabidopsis thaliana, are still very limited.

Experimental search for homologous receptors is time-consuming, laborious and expensive, as it requires the generation of a large number of candidates[17]. Meanwhile, the development of bioinformatics tools has made it possible to predict protein structure and function more accurately[10]. In particular, P2K1 and P2K2 have been structurally modeled and their interactions with ATP and other ligands have been discussed using 3D computational modeling methods[4,16].

From here, we apply bioinformatics tools to model the structures of potential proteins, providing important information to predict which is the eATP receptor in tomato plants.

Material and Methods

Sequences, templates and ligand structure: The lectin domain sequences of P2K1 (Q9LSR8 - 236 amino acids) and P2K2 (Q9M1G4 - 234 amino acids) were retrieved from the UniProtKB[23]. The *Sl*LecRLK amino acid sequences were obtained from the iTAK database[27]. Template structures were obtained from the protein data bank in PDB file format. The three-dimensional structure of ATP was obtained in SDF file format from the PubChem database[12].

Pipeline for computational 3D structure of L-type LecRK ecto-domain prediction: In this study, the pipeline for predicting the 3D structure of L-type LecRLK receptors relies on the previous P2K1 and P2K2 modeling pipeline[4,16] with some modification due to no longer available softwares (Figure 1). In details, the target lectin domain sequence was retrieved from the UniProt database and uploaded to PDB to find the template structures. Candidate templates were selected based on major criteria: a high percentage of sequence identity, crystal structural quality (low resolution), sequence coverage in the alignment and ligand binding (Ca^{2+}, Mn^{2+}, sugar/adenine). The target lectin models were generated following the model-ligand module of the Modeller manual[19,26]. For each target protein, 1000 model structures were used to generate cruel model and adding ions[19,26]. As the results of this step, 10 models, with the best score, are selected for the next step.

The generated models were scored based on Modeller's probability density function with a low discrete optimized protein energy score (DOPE)[14]. The models that have normalized DOPE scores under -1 were chosen for the structural geometry using Verify3D and the Ramachandran plot embedded in the SAVEs web server[14] . The good model is checked with Verify3D scores higher than 80% and no residues in outlier regions of the Ramachandran plot. It involves calculating the energy function to explore the different conformation changes of the target structure[14]. The target of the model structure would be configured with the Charmm22 force field[15] that emerged in NAMD[18] .

Pipeline for ligands docking: The molecular docking process was carried out using AutoDock Vina and AutoDockTool (ADT), a package of the MGLtools[7]. Before carrying out the docking process, polar-hydrogen and Gasteiger partial charge were added to convert the low-energy structures in PDB file format to PDBQT file format. The final target modeling receptor and ligand were imported in to AutoDockTool. A grid box that covers 4 loops of the receptor was used for target docking to ATP active binding site as previous described in P2K1 and P2K2[4,16]. The results of docking were analyzed with Pymol or MGLtool for 3D visualization and Ligplot for 2D diagrams[8,13].

Phylogenetic analysis: These sequences were imported into Mega 11 for alignment with P2K receptor sequences to analyze their relationship. All sequences aligned with the ClustalW algorithm[20]. The phylogeny of LecRKs was performed using the maximum likelihood (ML) method and was replicated 1000 times.

Results and Discussion

With the aim of applying 3D computational modeling structure to predict potential eATP receptor in tomato (*Solanum lycopersicum*), we simulated the structures of P2K1, P2K2 and their mutants to build pipeline and control. P2K1 and P2K2 had been built by 3D computational modeling structures[4,16]. We reconstructed these structures and their mutants using a new available pipeline.

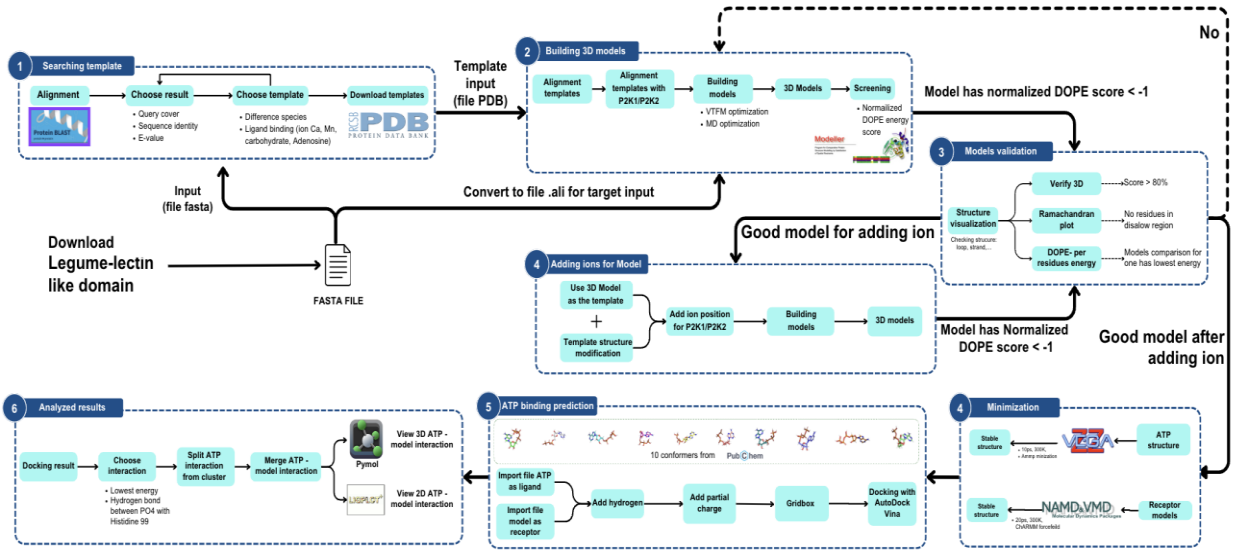

Figure 1: Workflow for homology modeling and molecular docking of ecto-domain of L-type Lectin receptor kinases. Quá trình thực hiện thông qua sáu bước chính: (1) searching template thông qua cơ sở dữ liệu PDB; (2) building 3D models bằng Modeller; (3) Models validation bằng verify 3D, Ramachadran plot và DOPE score; (4) adding ions and minimization to generate final protein structures; (5) ATP or other ligands docking using AutoDock Vina and (6) Analyzed results to visulize them.

Histidine99 playing an important role in holding eATP at the correct active binding site of P2K1 and P2K2: 3D computational modeling of P2K1 and P2K2, interacting with eATP, were successfully generated with the corresponding energy -8.287 Kcal/mol and -8.386 Kcal/mol respectively (Figures 2A and 2B). Histidine 99 of P2K2 has been experimentally demonstrated to play an important role in eATP binding[4]. However, the reason for this process has not been clearly explained. Here, *p2k1*^{his99ala} and *p2k2*^{his99ala} were simulated by us. eATP was targeted for docking at the position equivalent to His99 (Ala99). The results show that eATP interacts with the *p2k1*^{his99ala} (-5.858 Kcal/mol) and *p2k2*^{his99ala} (1.334 Kcal/mol) mutants with less stable energy than the wild-type (Figures 2C and 2D). With these results, we explain the importance of histidine 99 in the eATP-P2K binding model.

Our structures are also consistent with the energy results discussed in previous studies[4,16] indicating that our receptor structural modeling and ligand target docking pipelines are suitable for further prediction. The structures in figure 2 will be used as controls for the prediction of tomato eATP receptors. We named the pocket where eATP interacts with His99 as the active binding site.

Eight members of the *Sl*LecRK family share P2K's Histidine 99: In 2017, Wang and Bouwmesster[25] identified 22 members of the L-type Lectin receptor kinase family of Tomato (Solanum lycopersicum). Through the iTALK database, we found 24 genes in tomato with high homology to P2K1 and P2K2 (Figure 3). These genes, when translated, will produce proteins with the typical structures of LecRK proteins such as a transmembrane domain containing hydrophobic region, an intracellular domain with a conserved kinase domain and an extracellular domain capable of interacting with extracellular ligands. Through the process of amino acid blast sequencing, we found eight *Sl*LecRKs with amino acids conserved with His99 in P2K2. We decided to construct 3D computational modeling structures of these proteins and to test their ability to interact with eATP. In addition, Solyc10g047810 without His99 homology was also used as a negative sample (Table 1).

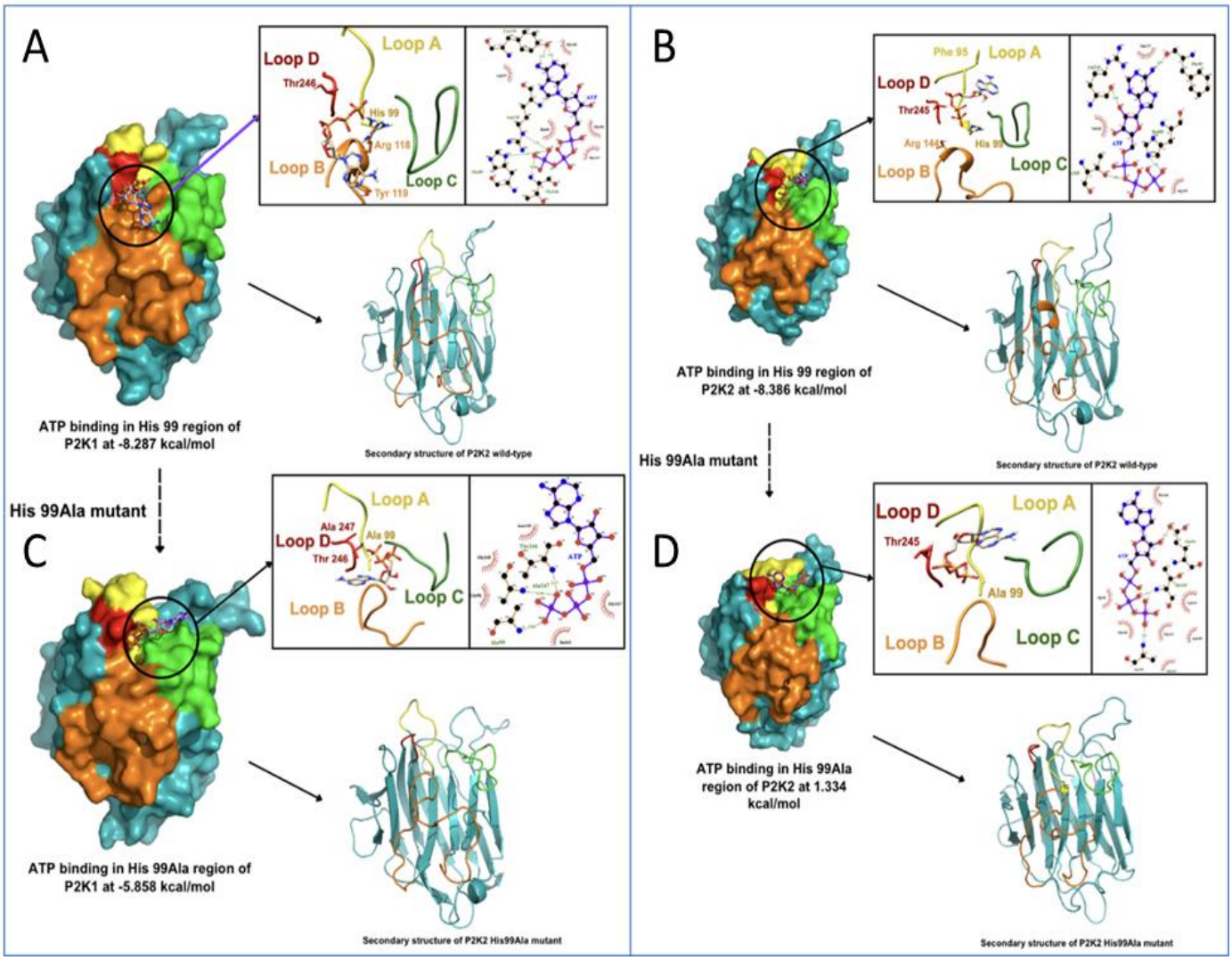

Figure 2: ATP binding at predicted active binding site of *At*P2Ks' extracellular domains. (A) Wild-type P2K1; (B) Wild-type P2K2; (C) *p2k1*^{his99ala}; and (D) *p2k2*^{his99ala}. Yellow: loopA region; orange: loopB region; green: loopC region; and red: loopD region.

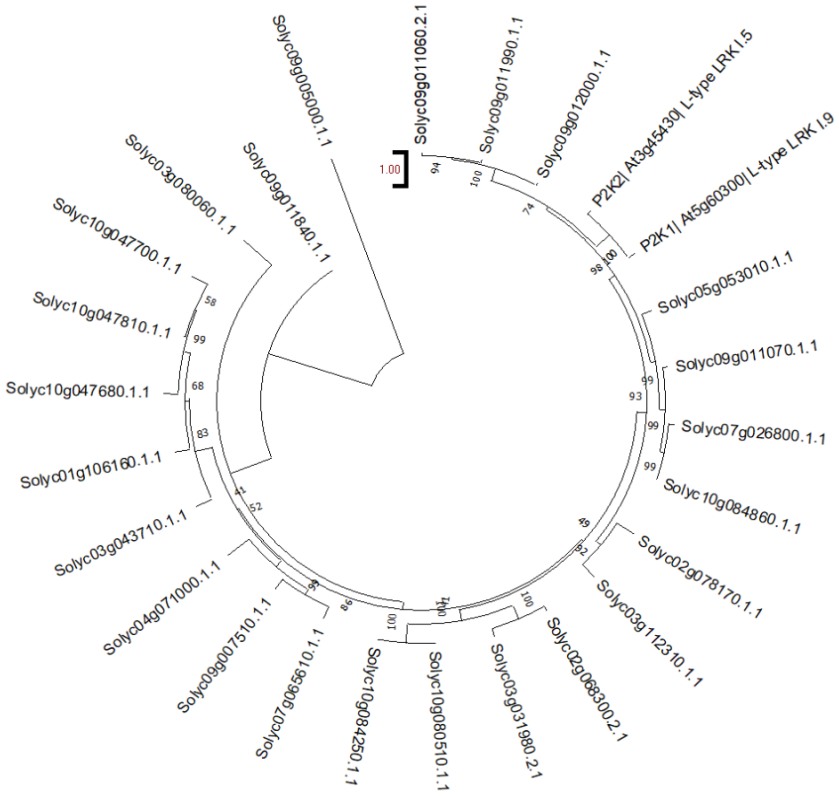

Figure 3: *Sl*LecRKs comparison to *At*P2Ks. Numbers at the nodes indicates the bootstrap values of maximum likelihood (ML) method and replicated 1000 times. Bar 1.00 represent sequence divergence.

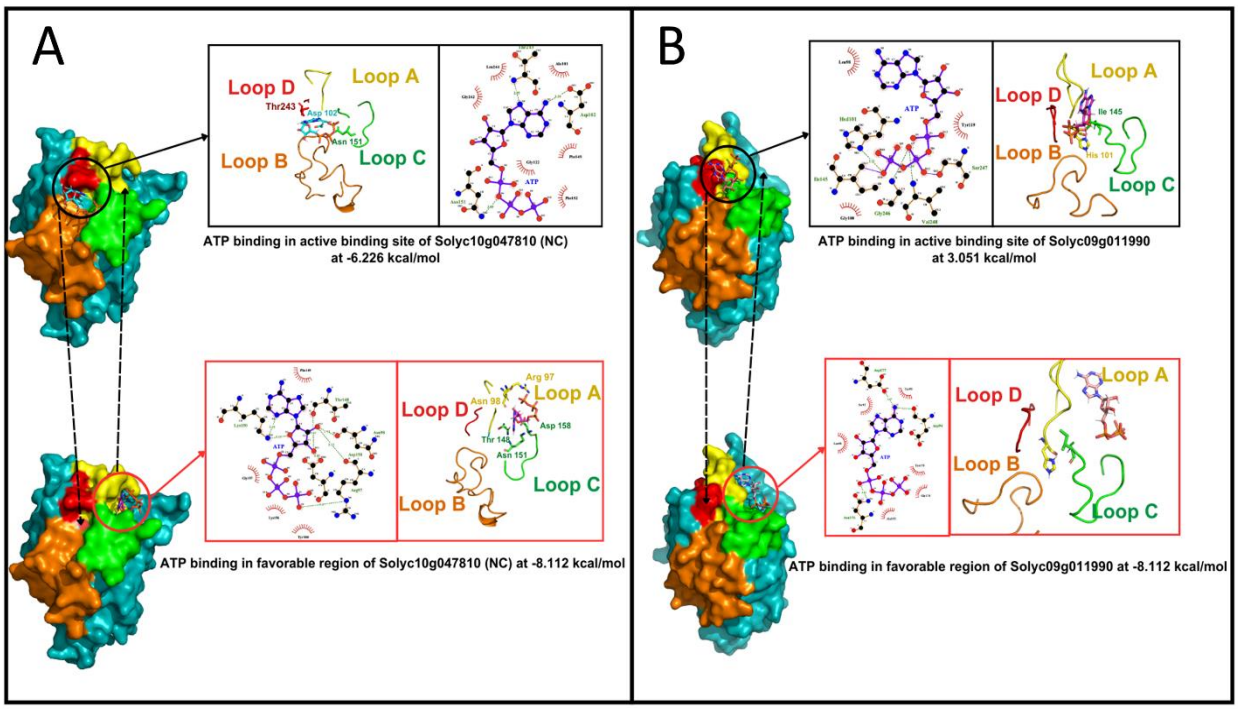

Figure 4: ATP binds to another favorable site if the receptor active binding site is not compatible. Binding of ATP in active binding site and another favorable region models of (A) Solyc10g047810 and (B) Solyc09g011990. Yellow: loopA region; orange: loopB region; green: loopC region; and red: loopD region.

Histidine 99 is a necessary but not sufficient condition: As expected, when performing target docking, Solyc10g047810, without His99 homology, had a poor interaction with ATP at the active binding site, with energy -

6.226 Kcal/mol (Figure 4A and table 1). Interestingly, Solyc09g011990, sharing His99 homology as well as having a very high sequence similarity to P2K1 and P2K2, had a very poor interaction with ATP at the active binding site, with energy 3.051 Kcal/mol (Figure 3, figure 4A and table 1).

Thus, although His99 homology is an experimentally proven key amino acid, it is not the only condition for ATP receptor prediction. In addition, when performing free ATP docking for Solyc10g047810 and Solyc09g011990, the results showed that both proteins interacted tightly with eATP at the other favorable region, biased towards loop A and not interacting with loop B (Figure 4). This favorable binding site may prevent eATP (or other ligands) from entering the active binding site, helping to increase the specificity for L-type LecRK receptors.

ATP strongly binds to the active binding sites of Solyc09g011060 and Solyc09g012000: More importantly, we found that Solyc09g011060 and Solyc09g012000 could bind ATP in the active binding site at energies of -8.306 Kcal/mol and -8.187 Kcal/mol respectively. Our results also show that other ligands interact less well than ATP at the active binding site (Figure 5 and table 1).

In particular, Solyc09g012000 has a relatively high energy for sugar binding, while still interacting well with ADP (-7.029 Kcal/mol). These characteristics are very similar to previously described Arabidopsis P2K1 and P2K2.

Table 1
Interaction ability of some SlLecRKs with ATP compared with *At*P2Ks

GeneID	Template (%identity)	ATP binding energy at active binding site (kcal/mol)	Interaction residues	
P2K1	3IPV_A (29.03%) 1FAT_A(28.98%) BJQ_F(29.44%)	-8.287	His60, Thr117, Arg118,Tyr119, Asp147, Phe 148, Gly245, Thr246	His99
p2k1^{His99Ala}		-5.858		Ala99
P2K2	1HQL_A(27.4%) 1AVB_A(28.8%) 1DBN_A(26.5%)	-8.386	Phe95, Arg118, Val143, Arg144, Thr177, Thr245	His99
p2k2^{His99Ala}		1.334		Ala99
Solyc10g047810 (NS)	3IPV_A (31.08%), 3UJO_A (29.80%), 2FMD_A (33.74%)	-6.226	Ala101, Asp102, Gly122, Phe 149, Asn151, Phe 152, Gly242, Thr243, Leu244	
Solyc09g011070	3USU_B (31.71%), 4WV8_A(28.10%), 5KXB_A(28.40%)	-6.758	Lys99, Leu100, His103, Val147, Gln148, Phe150, Gly 247, Leu248, Leu 249	
Solyc05g053010	3IPV_A(33.73%), 3USU_A(36.51%), 2FMD_A(32.61%)	-5.909	Arg61, His103, Pro120, Ala118, Ser122, Gln121, Asp149, Gly248, Leu249, Leu250	
Solyc02g078170	3UJO_A (26.74%), 1G7Y_A(27.06%), 2FMD_A(29.37%)	-7.186	Arg98, His102, Phe146, Lys147, Asn 148, Glu 258	
Solyc09g011060	3IPV_A(26.89%), 3UJO_A(26.74%), 2FMD_A(29.37%)	-8.306	Leu 97, Ser 98, His100, Asn118, His 119, Ile 144, Tyr 145, Ser 146, Phe149, Gly 245, Ser 246, Val 247	
Solyc09g011990	5AVA_A(30.92%), 1FAT_A(32.02%), 2FMD_A (30.61%)	3.051	Leu98, Gly100, His101, Tyr119, Ile145, Gly224, Val248, Ser225	
Solyc09g012000	5AVA_A(31.17%), 1FAT_A(30.65%), 2FMD_A (34.75%)	-8.187	Gln58, His100, Asn118, His119, Ile144, Tyr145, Ser146, Ala147, Asp148, Gly246, Ser247	
Solyc03g043710	3IPV_A (37.97%), 3UJO_A (35.27%), 4WV8_A (40.42%)	-7.521	His 103, Asp121, Ser120, Gly122, Gly123, Phe124, Phe148, Val151, Glu152	
Solyc10g084860	3USU_B (33.33%), 4WV8_A(31.97%), 5KXB_A(31.56%)	-7.682	Gly99, Hsd101, Val145, Arg146, Asp147, Leu247, Leu248	

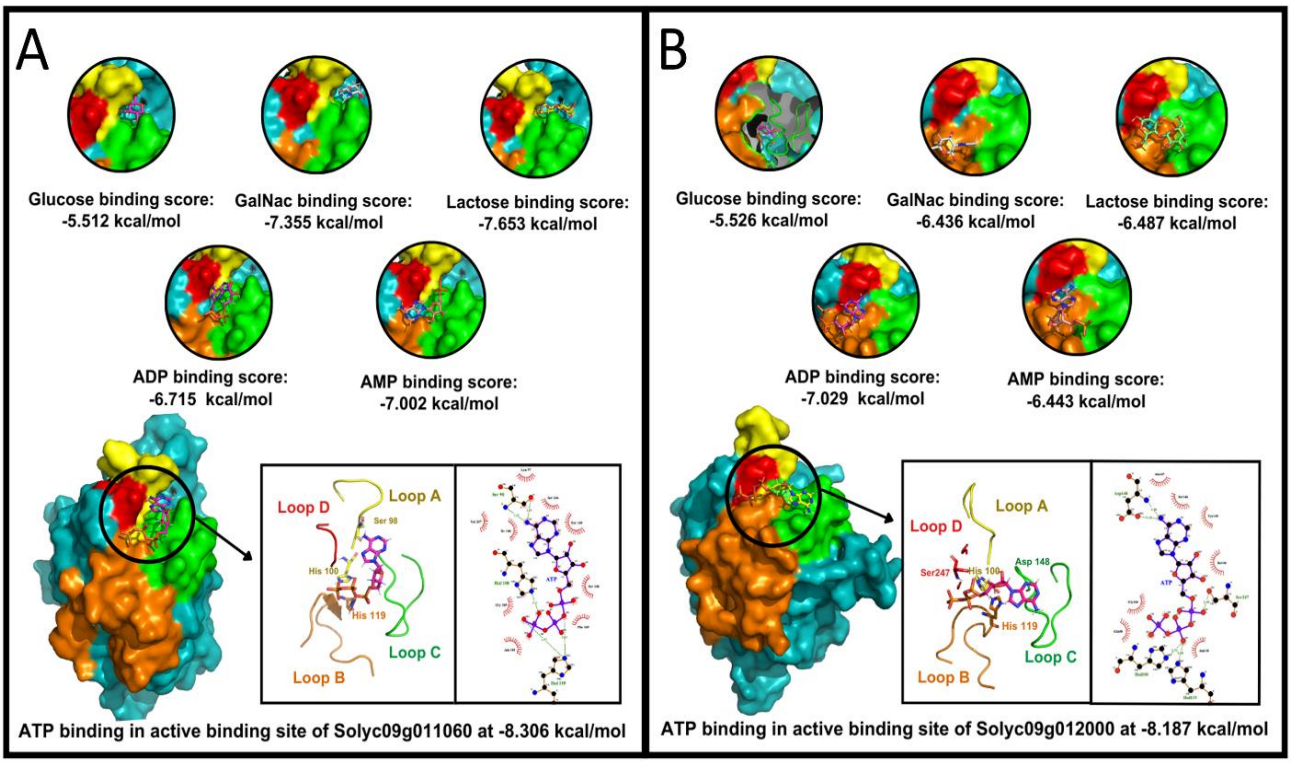

Figure 5: Solyc09g012000 and Solyc09g011060, potentially, are *Sl*P2Ks. Binding of ATP and other ligands in active binding site models of (A) Solyc09g011060 and (B) Solyc09g012000. Yellow: loopA region; orange: loopB region; green: loopC region; and red: loopD region.

Although L-type LecRKs were previously thought to belong to the sugar binding receptor family, recent studies have shown that L-type LecRKs interact with other ligands, which are involved in many aspects of plant immunity[25]. Our model contributes to this new observation.

Conclusion

We have built a new pipeline, using available tools, to model the ectodomains of L-type LecRK receptors and test their ability to interact with various ligands. Our model visualizes the active binding site as well as other favorable binding sites, explaining the specific interactions between ligands and receptors. His99 is again confirmed to play an important role in the ATP active binding site but is not a sufficient condition. This opens up the need to find other important factors in the model of ATP interacting with the active binding site. Our model predicts that Solyc09g011060 and Solyc09g012000 have a very good ability to interact with ATP at the active binding site. If this is confirmed experimentally in the future, our model can be applied to quickly find ATP receptors in other plant species.

Acknowledgement

This research is funded by The University of Science, VNU-HCM, Vietnam under grant number SH-CNSH 2023-03

References

1. Abbracchio M.P. and Burnstock G., Purinoceptors: Are there families of P2X and P2Y purinoceptors?, *Pharmacology & Therapeutics*, **64**, 445–475 **(1994)**

2. Bouwmeester K. and Govers F., Arabidopsis L-type lectin receptor kinases: phylogeny, classification and expression profiles, *Journal of Experimental Botany*, **60**, 4383–4396 **(2009)**

3. Cho S.H. et al, Activation of the plant mevalonate pathway by extracellular ATP, *Nat Commun*, **13**, 450 **(2022)**

4. Cho S.H. et al, Computational prediction and in vitro analysis of the potential ligand binding site within the extracellular ATP receptor, P2K2, *Plant Signaling & Behavior*, **18**, 2173146 **(2023)**

5. Choi J. et al, Identification of a Plant Receptor for Extracellular ATP, *Science*, **343**, 290–294 **(2014)**

6. Di Virgilio F., Sarti A.C., Falzoni S., De Marchi E. and Adinolfi E., Extracellular ATP and P2 purinergic signalling in the tumour microenvironment, *Nat Rev Cancer*, **18**, 601–618 **(2018)**

7. Forli S. et al, Computational protein–ligand docking and virtual drug screening with the AutoDock suite, *Nat Protoc*, **11**, 905–919 **(2016)**

8. Humphrey W., Dalke A. and Schulten K., VMD: Visual molecular dynamics, *Journal of Molecular Graphics*, **14**, 33–38 **(1996)**

9. Jaffe M.J., The Role of ATP in Mechanically Stimulated Rapid Closure of the Venus's Flytrap 1, *Plant Physiol*, **51**, 17–18 **(1973)**

10. Jumper J. et al, Highly accurate protein structure prediction with AlphaFold, *Nature*, **596**, 583–589 **(2021)**

11. Khakh B.S. and Burnstock G., The Double Life of Atp, *Scientific American*, **301**, 84–93 **(2009)**

12. Kim S. et al, PubChem 2023 update, *Nucleic Acids Research*, **51**, D1373–D1380 **(2023)**

13. Lourenço B.L.A. et al, Virtual Screening and Molecular Docking for Arylalkylamine-N-Acetyltransferase (aaNAT) Inhibitors, a Key Enzyme of Aedes (Stegomyia) aegypti (L.) Metabolism, *Computational Molecular Bioscience*, **5**, 35–44 **(2015)**

14. Lüthy R., Bowie J.U. and Eisenberg D., Assessment of protein models with three-dimensional profiles, *Nature*, **356**, 83–85 **(1992)**

15. Morris A.L., MacArthur M.W., Hutchinson E.G. and Thornton J.M., Stereochemical quality of protein structure coordinates, *Proteins: Structure, Function and Bioinformatics*, **12**, 345–364 **(1992)**

16. Nguyen C.T. et al, Computational Analysis of the Ligand Binding Site of the Extracellular ATP Receptor, DORN1, *PLOS ONE*, **11**, e0161894 **(2016)**

17. Pham A.Q., Cho S.H., Nguyen C.T. and Stacey G., Arabidopsis Lectin Receptor Kinase P2K2 Is a Second Plant Receptor for Extracellular ATP and Contributes to Innate Immunity1 [OPEN], *Plant Physiology*, **183**, 1364–1375 **(2020)**

18. Phillips J.C. et al, Scalable molecular dynamics with NAMD, *Journal of Computational Chemistry*, **26**, 1781–1802 **(2005)**

19. Šali A. and Blundell T.L., Comparative Protein Modelling by Satisfaction of Spatial Restraints, *Journal of Molecular Biology*, **234**, 779–815 **(1993)**

20. Tamura K., Stecher G. and Kumar S., MEGA11: Molecular Evolutionary Genetics Analysis Version 11, *Molecular Biology and Evolution*, **38**, 3022–3027 **(2021)**

21. Tanaka K., Choi J., Cao Y. and Stacey G., Extracellular ATP acts as a damage-associated molecular pattern (DAMP) signal in plants, *Frontiers in Plant Science*, **5**, https://doi.org/10.3389/fpls.2014.00446 **(2014)**

22. Tanaka K., Gilroy S., Jones A.M. and Stacey G., Extracellular ATP signaling in plants, *Trends Cell Biol*, **20**, 601–608 **(2010)**

23. The UniProt Consortium, UniProt: the Universal Protein Knowledgebase in 2023, *Nucleic Acids Research*, **51**, D523–D531 **(2023)**

24. Vitiello L., Gorini S., Rosano G. and la Sala A., Immunoregulation through extracellular nucleotides, *Blood*, **120**, 511–518 **(2012)**

25. Wang Y. and Bouwmeester K., L-type lectin receptor kinases: New forces in plant immunity, *PLOS Pathogens*, **13**, e1006433 **(2017)**

26. Waterhouse A. et al, SWISS-MODEL: homology modelling of protein structures and complexes, *Nucleic Acids Research*, **46**, W296–W303 **(2018)**

27. Zheng Y. et al, iTAK: A Program for Genome-wide Prediction and Classification of Plant Transcription Factors, Transcriptional Regulators and Protein Kinases, *Molecular Plant*, **9**, 1667–1670 **(2016)**.

(Received 09th December 2024, accepted 11th January 2025)