Journal of Plant Development
Vol. **32**, Dec **2025**: 157-162
doi: 10.47743/jpd.2025.32.1.976
Research Article

ISSN 2065-3158 print /e-ISSN 2066-9917
Available online: www.plant-journal.uaic.ro

PREDICTION OF 3D STRUCTURE OF MITOGEN-ACTIVATED PROTEIN KINASE KINASE 2 USING ALPHAFOLD 2 AND SIMULATION OF ITS INTERACTION WITH MITOGEN-ACTIVATED PROTEIN KINASE 6 IN *ARABIDOPSIS THALIANA*

Sameer KHAN[1], Tim XING[1*]

[1] Department of Biology and Institute of Biochemistry, Carleton University, Ottawa ON – Canada.
[*] Corresponding author. E-mail: tim.xing@carleton.ca, ORCID: 0000-0002-8325-0835

Abstract: *Arabidopsis thaliana* mitogen-activated protein kinase (MPK or MAPK) signaling network plays significant roles in various cellular processes. The three-dimensional structure of mitogen-activated protein kinase kinase 2 (MKK2), an upstream kinase in the MAP kinase cascade, was predicted using AlphaFold 2, and protein-protein docking simulations were performed between MPK6 and MKK2. The docking analysis identified important residues mediating their interaction. This structural prediction and protein docking analysis provide a further understanding at protein structure level.

Keywords: HADDOCK, interface visualization, kinase cascade, protein docking, stress.

Introduction

A mitogen-activated protein kinase cascade consists of three main tiers of protein kinases: MPK kinase kinases (MAPKKKs), MPK kinases (MKKs), and MPKs, which are activated through sequential phosphorylation events in response to various extracellular and intracellular signals [LEE & al. 2008; POPESCU & al. 2009]. In *Arabidopsis thaliana*, there are 60 MAPKKKs, 10 MKKs and 20 MPKs, thus leaving a multitude of possibilities to form cascades, even when bearing in mind that not all combinations occur [XING & FOROUD, 2021]. Recent advances in protein structure prediction, particularly with the emergence of AlphaFold 2, have revolutionized the field of computational biology and protein research. AlphaFold 2, a deep learning-based approach, has demonstrated remarkable success in accurately predicting protein structures [PAKHRIN & al. 2021]. The accuracy of AlphaFold 2 has been reported to be close to that of experimental determination techniques, signifying a significant leap in the reliability of predicted protein structures [WANG & al. 2022]. AlphaFold 3 is a major advance over AlphaFold 2 and it broadens the scope from protein-only structures to multi-molecule complexes [ABRAMSON & al. 2024]. Ever since its publication, AlphaFold 2 began to pave the way for the convergence of structural bioinformatics and artificial intelligence, with ongoing efforts to establish standardized models for protein structure prediction [SZELOGOWSKI, 2023].

AlphaFold 2 is a highly accurate deep learning algorithm developed to predict the three-dimensional (3D) structure of proteins from their amino acid sequences [JUMPER & al. 2021; MA & al. 2022). This algorithm has demonstrated remarkable accuracy in predicting protein structures, as evidenced by its success in the 14th Community Wide Experiment on the Critical Assessment of Techniques for Protein Structure Prediction (CASP14) [TAKEI & ISHIDA, 2022]. The success of AlphaFold 2 in predicting the 3D structures of single protein

Received: 10 September 2025 / ***Revised***: dd Month yyyy / ***Accepted***: dd Month yyyy – Editorial editing space

chains has raised questions about its future role in the field of protein structure prediction [KWON & al. 2021]. One of the key advantages of AlphaFold 2 is its ability to predict the 3D structures of proteins even when their native structures are unknown. This is achieved through the use of structure-based prediction methods, such as homology modeling or the application of its deep learning capabilities [PAK & IVANKOV, 2022]. The neural network-based method of AlphaFold 2 has allowed for the prediction of 3D structures for a significant portion of the human proteome, making these predicted structures publicly available [ZWECKSTETTER, 2021]. In our study here, we used the amino acid sequence of *Arabidopsis thaliana* MKK2 as a query in ColabFold, an extension of AlphaFold 2, to generate the top five ranked structural models.

The MKK2-MPK6 interaction plays an important role in plant responses to cold and salt stress signaling TEIGE & al. (2004). In this study, we also explored the molecular basis of this interaction through *in silico* docking analysis using HADDOCK. This software enables us to model the protein-protein interface between MKK2 and MPK6, providing insight into how specific residues mediate their interaction. By understanding these molecular contacts, we aim to shed light on the structural underpinnings of the MKK2-MPK6 signaling mechanism.

Material and methods

MKK2 structure prediction and visualization by ColabFold

ColabFold is an extension of AlphaFold 3 that focuses on predicting protein complexes, offering both accuracy and speed in its predictions [CHANG & al. 2024]. By combining fast homology search with AlphaFold 2 and RoseTTAFold, ColabFold can efficiently predict large protein complexes [JUSSUPOW & KAILA, 2023]. This implementation significantly reduces computation time, making it possible to predict protein-peptide complexes within a few minutes, depending on the size of the system [CHANG & al. 2024]. There have been various applications of ColabFold such as modeling protein structures and predicting changes in protein structure associated with genetic effects on traits and disorders [EINSON & al. 2022; HARIO & al. 2024]. Additionally, ColabFold has been used to build complexes with specific peptides, demonstrating its versatility in various protein modeling tasks [SALIMINASAB & al. 2023; ZLOBIN & al. 2023].

In this study, we used the amino acid sequence of *Arabidopsis thaliana* MKK2 as a query in ColabFold, an extension of AlphaFold 2, to generate the top five ranked structural models. The top four models were subsequently selected for further analysis. For the identification of domains and functional sites, we used ScanProSite (https://prosite.expasy.org/scanprosite) and InterProScan (https://www.ebi.ac.uk/interpro/about/interproscan). We used ChimeraX for the visualization and analysis of the protein models [PETTERSEN & al. 2021]. Additionally, Root Mean Square Deviation (RMSD) values were calculated between different models to assess their structural divergence.

MPK6 structure retrieval from Protein Data Bank

The experimental 3D structure of MPK6 was originally determined through x-ray diffraction as reported by PUTARJUNAN & al. (2019) and was also retrieved from Protein Data Bank (PDB, https://www.rcsb.org).

MKK2-MPK6 docking analysis with HADDOCK

HADDOCK (High Ambiguity Driven protein-protein DOCKing) is a computational tool widely used for modeling protein-protein interactions. It specializes in utilizing experimental data such as NMR, mutagenesis, or bioinformatics predictions to drive the docking process by integrating biochemical and biophysical information. The strength of HADDOCK lies in its ability to handle ambiguous interaction restraints, allowing it to predict complexes with high accuracy even when precise details of the interaction are unknown [DOMINGUEZ & al. 2003]. This makes it particularly useful in the study of complex biological mechanisms and the design of therapeutic molecules.

Visualization and analysis

For visualization and analysis of the docking simulations, PyMOL (https://www.pymol.org) was used to examine the interaction interfaces This included a detailed investigation of bond types and interacting residues across each of the four simulations conducted.

Results and discussions

AlphaFold 2 was accessed through the ColabFold interface, where the amino acid sequence of MKK2 was inputted. This resulted in five top-ranking models, with the top four selected for further analysis. Various visualizations of these models are presented in Figure 1. The analysis of the structural models revealed significant variations among the top four MKK2 models as indicated by Root Mean Square Deviation (RMSD) values. The RMSD between the first and second models was 6.470 Å, between the first and third models was 5.264 Å, and between the first and fourth models was 5.699 Å. Comparatively, the RMSD values between the second and third models, and between the second and fourth models were higher, at 8.112 Å and 7.624 Å, respectively, while the smallest deviation was observed between the third and fourth models at 3.565 Å. In regard to the structural variance among the MKK2 models, the RMSD values suggest that although there is a notable consistency between some models, significant disparities exist, especially between models two and three.

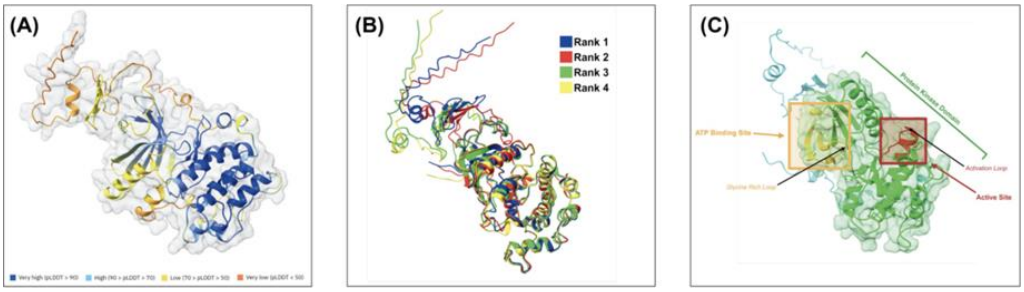

Figure 1. (A) Top-ranked MKK2 model colour-coded by predicted Local Distance Difference Test (plDDT) scores. (B) Superimposed structures of the MKK2 models, ranks 1-4. (C) Top-ranked MKK2 model with key functional sites labeled, identified using ScanProSite and InterProScan.

Functional site analysis predicted the presence of a protein kinase domain spanning residues 70 to 330, an ATP binding site across residues 76 to 99, and an active site between residues 188 to 200. These predictions highlight the critical regions potentially involved in the catalytic function and substrate interactions of the kinase. The analysis also reinforces the

importance of specific residues in kinase activity and ATP binding, crucial for the functional integrity for the protein. These structural insights are foundational for subsequent analysis on where these four models are docked with MPK6. We then performed a docking study, which aimed to analyze the interaction dynamics and the molecular mechanisms underlying the MKK2-MPK6 signaling pathway.

The outcomes of the docking simulations between MKK2 and MPK6 are shown in Figure 2, highlighting the specific residues involved from both proteins. The docking simulation results are summarized in Table 1.

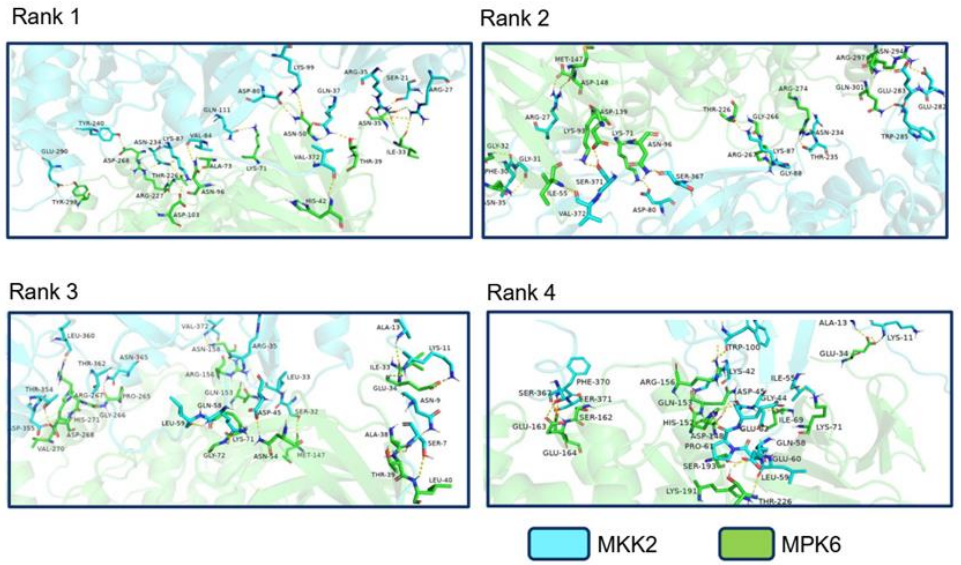

Figure 2. 3D Visualization of the MKK2-MPK6 interface for each MKK2 model being docked with MPK6, indicating the involved residues from both proteins.

Table 1. Summary of MKK2-MPK6 docking simulation results

	Binding Energy	Buried Surface Area	Desolvation Energy
#1	-475.02	3802.2	-7.71484
#2	-338.541	3796.74	-7.85969
#3	-445.666	3740.65	7.90779
#4	-318.102	4005.41	-14.3504

Discussion

In this study, we explored the protein structure of MKK2 and the molecular basis of MKK2-MPK6 interaction. In regard to the structural variance among the MKK2 models, the RMSD values suggest that although there is a notable consistency between some models, significant disparities exist, especially between models two and three.

The functional site analysis reinforces the importance of specific residues in kinase activity and ATP binding, crucial for the functional integrity of proteins. These structural

insights are foundational for subsequent analysis on where these four models are docked with MPK6. This docking study should further illustrate the interaction dynamics and the molecular mechanisms underlying the MKK2-MPK6 signaling pathway.

The MKK2-MPK6 docking analysis reveals the critical residues and binding interfaces involved in this essential signaling pathway. The interaction sites identified suggest that these proteins form a stable complex that likely ensures efficient and accurate signal transmission. This reinforces the hypothesis that MKK2-MPK6 binding is a significant step in the activation of downstream stress response genes.

The structural insights gained from this analysis could be instrumental in developing strategies for improving plant resilience. Understanding the specific residues involved in the MKK2-MPK6 interaction can help in the design of targeted mutations or small molecules that could enhance or disrupt this complex. This could be particularly useful for engineering plants with increased resistance to harsh environmental conditions.

Our current work has indicated the importance of integrating computational and experimental approaches in unraveling the complexities of biological systems. Despite the modeling approach, our docking analysis is limited by the static nature of *in silico* models. *In vivo* or *in vitro* studies are necessary to confirm the interaction dynamics and the role of specific residues identified in this study. Future research should focus on validating these findings through mutagenesis or protein interaction assays, as well as expanding the analysis to other related signaling pathways to uncover broader regulatory networks.

Note: Our data are valid even when AlphaFold 3 [ABRAMSON & al. 2024] was introduced after the completion of our work due to improvement in specific areas of molecular interactions.

Acknowledgements

This work was supported by a research grant to Tim XING from Bayer Crop Science Inc. and Carleton University, Canada.

References

ABRAMSON J., ADLER J., DUNGER J., EVANS R., GREEN T., PRITZEL A., RONNEBERGER O., WILLMORE L., BALLARD A. J., BAMBRICK J., BODENSTEIN S. W., EVANS D. A., HUNG C. C., O'NEILL M., REIMAN D., TUNYASUVUNAKOOL K., WU Z., ŽEMGULYTĖ A., ARVANITI E., BEATTIE C., BERTOLLI O., BRIDGLAND A., CHEREPANOV A., CONGREVE M., COWEN-RIVERS A. I., COWIE A., FIGURNOV M., FUCHS F. B., GLADMAN H., JAIN R., KHAN Y. A., LOW C. M. R., PERLIN K., POTAPENKO A., SAVY P., SINGH S., STECULA A., THILLAISUNDARAM A., TONG C., YAKNEEN S., ZHONG E. D., ZIELINSKI M., ŽÍDEK A., BAPST V., KOHLI P., JADERBERG M., HASSABIS D. & JUMPER J. M. 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*. **630**(8016): 493-500. https://doi.org/10.1038/s41586-024-07487-w

CHANG L., MONDAL A, SINGH B., MARTÍNEZ-NOA Y. & PEREZ A. 2024. Revolutionizing peptide-based drug discovery: advances in the post-AlphaFold era. *Wiley Interdisciplinary Reviews Computational Molecular Science*. **14**(1): e1693. https://doi.org/10.1002/wcms.1693

DOMINGUEZ C., BOELENS R. & BONVIN A. M. 2003. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *Journal of the American Chemical Society*. **125**(7): 1731-1737. https://doi.org/10.1021/ja026939x

EINSON J., MINAEVA M., RAFI F. & LAPPALAINEN T. 2024. The impact of genetically controlled splicing on exon inclusion and protein structure. *PLoS One*. **19**(3): e0291960. https://doi.org/10.1371/journal.pone.0291960

HARIO S., LE G. N. T., SUGIMOTO H., TAKAHASHI-YAMASHIRO K., NISHINAMI S., TODA H., LI S., MARVIN J. S., KURODA S., DROBIZHEV M., TERAI T., NASU Y. & CAMPBELL R. E. 2024. High-performance genetically encoded green fluorescent biosensors for intracellular l-lactate. *ACS Central Science*. **10**(2): 402-416. https://doi.org/10.1021/acscentsci.3c01250

JUMPER J., EVANS R., PRITZEL A., GREEN T., FIGURNOV M., RONNEBERGER O., TUNYASUVUNAKOOL K., BATES R., ŽÍDEK A., POTAPENKO A., BRIDGLAND A., MEYER C., KOHL S. A. A., BALLARD A. J., COWIE A., ROMERA-PAREDES B., NIKOLOV S., JAIN R., ADLER J., BACK T., PETERSEN S., REIMAN D., CLANCY E., ZIELINSKI M., STEINEGGER M., PACHOLSKA M., BERGHAMMER T., BODENSTEIN S., SILVER D., VINYALS O., SENIOR A. W., KAVUKCUOGLU K., KOHLI P. & HASSABIS D. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature*. **596**(7873): 583-589. https://doi.org/10.1038/s41586-021-03819-2

JUSSUPOW A. & KAILA V. R. I. 2023. Effective molecular dynamics from neural network-based structure prediction models. *Journal of Chemical Theory and Computation*. **19**: 1965-1975. https://doi.org/10.1021/acs.jctc.2c01027

KWON S., WON J., KRYSHTAFOVYCH A. & SEOK C. 2021. Assessment of protein model structure accuracy estimation in CASP14: Old and new challenges. *Proteins*. **89**(12): 1940-1948. https://doi.org/10.1002/prot.26192

LEE J. S., HUH K. W., BHARGAVA A. & ELLIS B. E. 2008. Comprehensive analysis of protein-protein interactions between *Arabidopsis* MAPKs and MAPK kinases helps define potential MAPK signalling modules. *Plant Signaling & Behavior*. **3**(12): 1037-1041. https://doi.org/10.4161/psb.3.12.6848

MA W., ZHANG S., LI Z., JIANG M., WANG S., LÜ W., BI X., JIANG H., ZHANG H. & WEI Z. 2022. Enhancing protein function prediction performance by utilizing AlphaFold-predicted protein structures. *Journal of Chemical Information and Modeling*. **62**: 4008-4017. https://doi.org/10.1021/acs.jcim.2c00885

PAK M. & IVANKOV D. 2022. Best templates outperform homology models in predicting the impact of mutations on protein stability. *Bioinformatics*. **38**: 4312-4320. https://doi.org/10.1093/bioinformatics/btac515

PAKHRIN S. C., SHRESTHA B., ADHIKARI B. & KC D. B. 2021. Deep learning-based advances in protein structure prediction. *International Journal of Molecular Sciences*. **22**(11): 5553. https://doi.org/10.3390/ijms22115553

PETTERSEN E. F., GODDARD T. D., HUANG C. C., MENG E. C., COUCH G. S., CROLL T. I., MORRIS J. H. & FERRIN T. E. 2021. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science*. **30**(1): 70-82. https://doi.org/10.1002/pro.3943

POPESCU S. C., POPESCU G. V., BACHAN S., ZHANG Z., GERSTEIN M., SNYDER M. & DINESH-KUMAR S. P. 2009. MAPK target networks in *Arabidopsis thaliana* revealed using functional protein microarrays. *Genes & Development*. **23**(1): 80-92. https://doi.org/10.1101/gad.1740009

PUTARJUNAN A., RUBLE J., SRIVASTAVA A., ZHAO C., RYCHEL A. L., HOFSTETTER A. K., TANG X., ZHU J. K., TAMA F., ZHENG N. & TORII K. U. 2019. Bipartite anchoring of SCREAM enforces stomatal initiation by coupling MAP kinases to SPEECHLESS. *Nature Plants*. **5**(7): 742-754. https://doi.org/10.1038/s41477-019-0440-x

SALIMINASAB M., YAMAZAKI Y., PALMATEER A., HARRIS A., SCHUBERT L., LANGNER P., HEBERLE J., BONDAR A. N. & BROWN L. S. 2023. A proteorhodopsin-related photosensor expands the repertoire of structural motifs employed by sensory rhodopsins. *The Journal of Physical Chemistry B*. **127**(37): 7872-7886. https://doi.org/10.1021/acs.jpcb.3c04032

SZELOGOWSKI D. 2023. Deep learning for protein structure prediction: advancements in structural bioinformatics. *bioRxiv*. https://doi.org/10.1101/2023.04.26.538026

TAKEI Y. & ISHIDA T. 2022. How to select the best model from AlphaFold 2 structures? *bioRxiv*. https://doi.org/10.1101/2022.04.05.487218

TEIGE M., SCHEIKL E., EULGEM T., DÓCZI R., ICHIMURA K., SHINOZAKI K., DANGL J. L. & HIRT H. 2004. The MKK2 pathway mediates cold and salt stress signaling in *Arabidopsis*. *Molecular Cell*. **15**(1): 141-152. https://doi.org/10.1016/j.molcel.2004.06.023

WANG G., WU Z., FANG X., XIANG Y., LIU Y. YU D. & MA Y. 2022. Efficient AlphaFold 2 training using parallel evoformer and branch parallelism. *arXiv*: 2211.00235. https://doi.org/10.48550/arXiv.2211.00235

XING T. & FOROUD N. A. 2021. *A network of MAP kinase pathways in pant defense regulation: what we have learned from genetic modification studies*? Pp. 393-407. In: PANDEY G. K. (ed.). *Protein kinases and stress signaling in plants*: *functional genomic perspective*. John Wiley & Sons Inc.

ZLOBIN A., BELYAEVA J. & GOLOVIN A. 2023. Challenges in protein QM/MM simulations with intra-backbone link atoms. *Journal of Chemical Information and Modeling*. **63**(2): 546-560. https://doi.org/10.1021/acs.jcim.2c01071

ZWECKSTETTER M. 2021. NMR hawk-eyed view of AlphaFold 2 structures. *Protein Science*. **30**(11): 333-2337. https://doi.org/10.1002/pro.4175

How to cite this article:
KHAN S. & XING T. 2025. Prediction of 3D structure of mitogen-activated protein kinase kinase 2 using AlphaFold 2 and simulation of its interaction with mitogen-activated protein kinase 6 in *Arabidopsis thaliana*. *J. Plant Develop*. **32**: 157-162. https://doi.org/10.47743/jpd.2025.32.1.976