Berkala **PENELITIAN HAYATI**

ISSN: 08526834 | E-ISSN:2337-389X
Volume 29 | No. 3 | **December** | 2023

Original Article

Comprehensive analysis of protein-protein interactions (PPIs) with structure prediction program for breast cancer determination

Febrina Margaretha[1#], Mukti Subagja[2#], Altri Diana Putri[3], Christabel Queentha Salam[2], Louis Valenthenardo[2], Priscilla Klaresza Adhiwijaya[2], Gabriella Zevania Kaitlyn[2], Arli Aditya Parikesit[1*]

[1] Department of Bioinformatics, School of Life Sciences, Indonesia International Institute for Life Sciences, Jakarta, Indonesia
[2] Department of Biotechnology, School of Life Sciences, Indonesia International Institute for Life Sciences, Jakarta, Indonesia
[3] Department of Biomedicine, School of Life Sciences, Indonesia International Institute for Life Sciences, Jakarta, Indonesia

Abstract

Breast cancer is a prevalent disease that primarily affects women, with significant implications for public health. Early detection improves survival rates, and technological advancements can enhance early diagnosis. Protein structure prediction methods can provide valuable insights into the structure of proteins involved in breast cancer, including human epidermal growth factor receptor-2 (HER2). HER2 is a known protein receptor that plays a critical role in breast cancer development and is a target for therapy. Predicting the binding affinity between HER2 and potential ligands can help identify novel treatment options. This study aimed to predict the structure of HER2 using I-TASSER and determine potential ligands using empirical graph neural network-based scoring functions. Molecular docking simulations were performed to evaluate the binding conformation and stability of the HER2-ligand complexes. The results showed the potential ligands identified by I-TASSER: AEE, UUU, and Mg2+. Afterward, their binding affinity with HER2 was assessed, yielding AEE as the best binding with the lowest vina score.

Keywords: HER2, I-TASSER, ligand, novel treatment, breast cancer, binding affinity
Received: August 21, 2023 Revised: December 5, 2023 Accepted: December 8, 2023

Introduction

The background and importance of breast cancer

Breast cancer is a disease primarily affecting the breast tissue and is characterized by the rapid and uncontrolled growth of cells (CDC, 2022). Its historical roots trace back to ancient Egypt, where it was first documented by George Ebers and Edwin Smith, signifying the long-standing significance of this ailment (Mandal, 2019) . Today, researchers categorize breast cancer into invasive and non-invasive types, each further divided based on their relation to the membrane. The non-invasive forms include DCIS and LCIS, while invasive breast cancer comprises ILC and IDC (Akram et al., 2017) . Another subset of breast cancer is what is known as HER2-positive. These types of cancer have a receptor protein called human epidermal growth factor receptor 2 (HER2) on their surface. Usually, this receptor helps in the control of cell growth. However, cancer cells that overexpress the receptor tend to proliferate quickly and are more likely to metastasize or spread to other body parts (National Cancer Institute, 2011).

According to the World Health Organization, breast cancer affected 7.8 million women in 2020, making it the most prevalent cancer and the second leading cause of cancer-related deaths among women. Therefore, raising awareness about breast cancer is paramount, as early detection significantly improves the chances of successful treatment and cure. While breast cancer mortality rates have declined, the testimonies of many women who have battled this disease underscore the critical significance of early detection. Several research studies have focused on early diagnosis, such as the one by Paramkusham et al. (2013) on using innovative image processing techniques on mammograms. Another recent approach by Brown et al. (2017) utilizes a phylogenetic tree analysis to monitor disease progression, further emphasizing the relevance of early diagnosis. The urgent need for improving early detection technologies makes breast cancer an essential topic for exploration, as it can significantly increase the likelihood of cancer remission for those affected.

Human epidermal growth factor receptor-2 (HER2) as a cancer inducer

HER2 is a transmembrane receptor from the epidermal growth factor receptor (EGFR) family that performs essential tyrosine kinase activity to regulate the growth, proliferation, differentiation, and survival of epithelial cells (Albagoush & Limaiem, 2023) . Overexpression of HER2 results in constitutive activation of the HER2 downstream signaling pathway, which results in worse biological behavior and clinical aggressiveness in breast cancer. Around 20-30% of human breast cancers and some ovarian and gastric cancers show amplification of HER2 (Iqbal & Iqbal, 2014) . HER2-amplified breast cancers have distinct biological and clinical characteristics, such as increased proliferation rates, high histologic and nuclear grade, low ER and PR levels, increased aneuploidy, proclivity to metastasize to CNS viscera, relative resistance to endocrine therapy, and increased sensitivity to doxorubicin, making HER2 an important prognostic

*Corresponding Author:
Arli Aditya Parikesit
Department of Bioinformatics, School of Life Sciences, Indonesia International Institute for Life Sciences, Jakarta, Indonesia
E-mail: arli.parikesit@i3l.ac.id
#Authors contributed equally

http://dx.doi.org/10.23869/bphjbr.29.3.20232
Published by © PBI East Java. Open Access www.berkalahayati.org

target for cancer therapy (Gutierrez & Schiff, 2011). Consequently, the prevalence and aggressive nature of HER2-positive breast cancer have led to the development and approval of HER2-targeted therapy drugs. The first breakthrough occurred 25 years ago with the monoclonal antibody drug trastuzumab (Kreutzfeldt et al., 2020). However, determining additional potential ligands for HER2 receptors can offer novel avenues for targeting and may contribute to developing innovative drug delivery methods.

Protein structure prediction (I-TASSER) and binding affinity with HER2

Numerous protein structure predictors have been developed to visualize and identify the protein's structure and alignment templates from the Protein Data Bank (PDB) (Burley et al., 2023). Much effort has been exerted to create novel algorithms to design the structure predictors. One of them is I-TASSER, which notably impacted protein and function structure prediction. The structure prediction will help to determine the target protein-protein interactions (PPIs) where I-TASSER may generate five top predicted structure models along with their spatial location (A. Roy et al., 2010). The final models are generated by decoy clustering for the structural similarity, where they are evaluated for the global topology and local secondary structure with TM-score and RMSD (Zhang et al., 2019; Zhang, 2008). Once this is done, the ligands can be determined based on the predicted structure.

The binding affinity of the protein-ligand complex is measured to compare which protein can most likely bind to HER2. One standard method for predicting protein-ligand binding affinity involves free energy simulations, such as free energy perturbation and thermodynamic integration. These methods integrate molecular dynamics simulation with free energy sampling algorithms but can be computationally expensive (Jespers et al., 2021). Another popular approach is the use of scoring functions (SFs). SFs predict binding affinity based on the structure of the protein-ligand complex without simulating the entire binding process. This makes them faster and more suitable for large-scale applications (Seo et al., 2021). SFs can be categorized into physics-based, knowledge-based, and empirical scoring functions, each with its approach and advantages. However, approximating the occurrence frequency precisely and accurately representing intermolecular interaction mechanisms remains challenging for empirical scoring functions (Meli et al., 2022).

In recent years, machine learning-based scoring functions, also known as descriptor-based scoring functions, have gained attention due to their superior performance. One of these machine learning algorithms is the Empirical Graph Neural Network for accurate protein-ligand binding affinity prediction (EGNA). EGNA constructs different graphs based on intermolecular distances and represents proteins and ligands separately. An empirical interaction representation layer is designed to capture distance information and raw atom features. It has shown superior performance in lower errors and higher correlations compared to state-of-the-art scoring functions (Xia et al., 2023).

Research that have been conducted in this domain are various, such as from Linding et al. (2003); Russell et al. (2004); and Hashemifar et al. (2018) and who reviewed the methods, challenges, and applications of PPI analysis in human diseases, with examples from cancer, neurodegeneration, and immune disorders. These studies demonstrate the potential of structure prediction programs for PPI analysis in cancer and other diseases. However, that research did not cater specific data analysis on specialized ligand such as RNA and metal-based. Moreover, they were focused on developing general interaction database on PPIs, and not on the specific ligands. Research in the protein-protein interactions (PPIs) of HER2 has the potential to uncover critical signaling pathways and potential HER2 interaction partners, providing valuable insights into its erroneous activity in breast cancer. It is hypothezied that understanding the PPIs of HER2 with molecular simulation may guide the development of novel strategies for effectively diagnosing, treating, and managing breast cancer, thus improving patient outcomes and advancing breast cancer research. Hence, this research aims to determine the PPIs of HER2 with a structure prediction program for breast cancer determination. With the hope that the investigation of PPIs may contribute to developing more effective prognostic and therapeutic tools in breast cancer research.

Methods

Sequence retrieval

The FASTA sequence of HER2 was obtained from the protein database NCBI (National Center for Biotechnology Information). The sequence retrieved is a whole gene sequence that is coming from Homo sapiens with a length of 1255 aa and a gene code ID of AAA75493.1 (Friend Tambunan & Parikesit, 2014).

Ligand determination

The sequence was uploaded to the I-TASSER server for ligand determination, where the tool can predict potential ligand binding sites on the HER2 structure. Ligand-binding pockets and potential ligands are identified based on various criteria, including shape complementarity and electrostatic interactions. Those criteria are designated for the benchmark test: RMSD (Root mean squared deviation) and TM-score (Thread modeling score).

RMSD is a way to measure the differences between two sets of proteins, commonly being the predicted protein and known reference structure. Parameters measured include atom comparison, which specifies the position of carbons, nitrogens, and hydrogens of amino acids between the structure. Distance calculation, meanwhile, measures the distance between a pair of atoms in a Euclidean manner where a straight line is drawn in 3D space. These parameters are then inputted into the formula:

$$\text{RMSD} = \sqrt{\frac{\sum_{i=1}^{N} (x - \hat{x})^2}{n}}$$

The final value is measured in angstroms and indicates the variability of distances between the atoms. A value less than 3 Å is considered good, while a value between 3-5 Å is still acceptable. Higher values than the specified range are not advisable (Wicaksono & Parikesit, 2023).

TM-score compares the value from the experimental to the native structure. It compares the structure and sizes for the overall topological similarity. TM-score is calculated with the following formula:

$$TM - Score = \frac{1}{L}\sum_{i=1}^{L} \frac{1}{1 + d_i^2/d_0^2}$$

L represents the protein structure length, and d is the distance in the aligned structures. The formula determined the mean of the inverted square distances between residues and normalized by the structure length for overall protein folding. The value will range from 0 to 1, where a higher value indicates a closer similarity to the reference. A score above 0.9 is an identical match for the structure determination.

Binding affinity

The top three predicted ligands provided by I-TASSER are selected for further analysis. The HER2 structure and each Protein Data Bank (PDB) ligand are inputted into EGNA. This web-based tool utilizes an empirical graph neural network to predict protein-ligand binding affinity. It was trained on the refined and general set of PDBbind 2016 and evaluated on the CASF-2016 and non-overlapping CSAR-HiQ. The tool constructs the ligand and their interactions based on different regions of each bound complex to estimate binding affinities. This analysis helps determine which predicted ligands have the highest binding affinity to HER2 (Zhou et al., 2022).

Comparison with molecular docking

HER2 and the ligand with the highest affinity are inputted into CB-Dock2 to see the molecular docking mechanism. CB-Dock2 is a blind docking server developed by a research lab. It utilizes a protein-surface-curvature-based cavity detection approach called CurPocket to guide the molecular docking process using AutoDock Vina (Liu et al., 2020). This analysis predicts the possible binding conformation and binding energy of the HER2-ligand complex, providing additional information about the interaction and stability of the complex.

To use CB-Dock2, the protein structure is inputted as a PDB file, while the ligand is as an SDF file. The submitted ligand is processed by adding hydrogen atoms and partial charges and generating an initial 3D conformation using RDKit. CB-Dock2 checks the submitted protein, adds any missing side-chain and hydrogen atoms, alerts the user about missing residues, and removes co-crystallized water molecules and other nonstandard groups (Liu et al., 2022) . The next step is cavity detection and docking, where CB-Dock2 will use template matching to find similar complexes in a database. It runs two pipelines if found: structure-based blind docking and template-based blind docking. Each pipeline produces a list of binding sites and poses, which are merged and ranked. If no similar complex is found, CB-Dock2 proceeds with structure-based blind docking only (Liu et al., 2022) . Finally, the visualization and analysis step presents the results to the user.

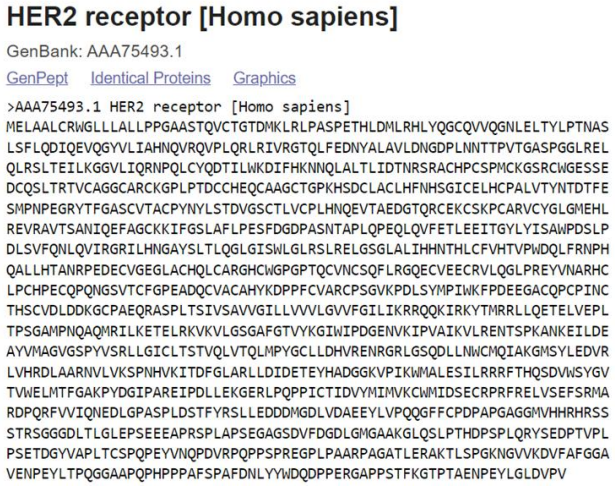
Visualization by PyMOL

Protein is visualized using PyMOL for a detailed examination of the binding site, where the residues can be selected and colored to highlight the binding pockets. The ligand interactions and protein conformation of the ligand-receptor complex can also be visualized, aiding in the interpretation and analysis of the obtained results (Jia et al., 2019).

Results and Discussion

I-TASSER

The HER2 human receptor FASTA sequence was retrieved from NCBI with an overall length of 1255 aa (amino acids) (Figure 1).


HER2 receptor [Homo sapiens]

GenBank: AAA75493.1

GenPept Identical Proteins Graphics

>AAA75493.1 HER2 receptor [Homo sapiens]
MELAALCRWGLLLALLPPGAASTQVCTGTDMKLRLPASPETHLDMLRHLYQGCQVVQGNLELTYLPTNAS
LSFLQDIQEVQGYVLIAHNQVRQVPLQRLRIVRGTQLFEDNYALAVLDNGDPLNNTTPVTGASPGGLREL
QLRSLTEILKGGVLIQRNPQLCYQDTILWKDIFHKNNQLALTLIDTNRSRACHPCSPMCKGSRCWGESSE
DCQSLTRTVCAGGCARCKGPLPTDCCHEQCAAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTFE
SMPNPEGRYTFGASCVTACPYNYLSTDVGSCTLVCPLHNQEVTAEDGTQRCEKCSKPCARVCYGLGMEHL
REVRAVTSANIQEFAGCKKIFGSLAFLPESFDGDPASNTAPLQPEQLQVFETLEEITGYLYISAWPDSLP
DLSVFQNLQVIRGRILHNGAYSLTLQGLGISWLGLRSLRELGSGLALIHHNTHLCFVHTVPWDQLFRNPH
QALLHTANRPEDECVGEGLACHQLCARGHCWGPGPTQCVNCSQFLRGQECVEECRVLQGLPREYVNARHC
LPCHPECQPQNGSVTCFGPEADQCVACAHYKDPPFCVARCPSGVKPDLSYMPIWKFPDEEGACQPCPINC
THSCVDLDDKGCPAEQRASPLTSIVSAVVGILLVVVLGVVFGILIKRRQQKIRKYTMRRLLQETELVEPL
TPSGAMPNQAQMRILKETELRKVKVLGSGAFGTVYKGIWIPDGENVKIPVAIKVLRENTSPKANKEILDE
AYVMAGVGSPYVSRLLGICLTSTVQLVTQLMPYGCLLDHVRENRGRLGSQDLLNWCMQIAKGMSYLEDVR
LVHRDLAARNVLVKSPNHVKITDFGLARLLDIDETEYHADGGKVPIKWMALESILRRRFTHQSDVWSYGV
TVWELMTFGAKPYDGIPAREIPDLLEKGERLPQPPICTIDVYMIMVKCWMIDSECRPRFRELVSEFSRMA
RDPQRFVVIQNEDLGPASPLDSTFYRSLLEDDDMGDLVDAEEYLVPQQGFFCPDPAPGAGGMVHHRHRSS
STRSGGGDLTLGLEPSEEEAPRSPLAPSEGAGSDVFDGDLGMGAAKGLQSLPTHDPSPLQRYSEDPTVPL
PSETDGYVAPLTCSPQPEYVNQPDVRPQPPSPREGPLPAARPAGATLERAKTLSPGKNGVVKDVFAFGGA
VENPEYLTPQGGAAPQPHPPPAFSPAFDNLYYWDQDPPERGAPPSTFKGTPTAENPEYLGLDVPV

Figure 1. HER2 Sequence Retrieval from NCBI

Once the FASTA sequence was obtained, it was inputted into I-TASSER for further analysis. The first step in I-TASSER's process involved threading, which is the identification of potential templates from a database of known protein structures. Threading was performed by comparing the target protein sequence with the sequences of proteins with known structures to find the best matches. These template proteins were assumed to have an overall fold and function similar to the target protein. As seen in Figure 2, I-TASSER determined eight domains of the HER2 protein based on our input for the structural modeling.

Once potential templates were identified, I-TASSER constructed an initial model of the target protein through fragment assembly. Fragments from the template proteins that match regions of the target protein sequence were aligned and assembled to create an initial structural model (Yang & Zhang, 2015). In Figure 3, the top five

models generated by I-TASSER can be seen. After the initial model was generated, I-TASSER employed a refinement step to improve the accuracy and reliability of the structure model to predict binding pockets. This step involved molecular dynamics simulations, energy minimization, and other optimization techniques to refine the protein structure and enhance the prediction of ligand-binding regions.

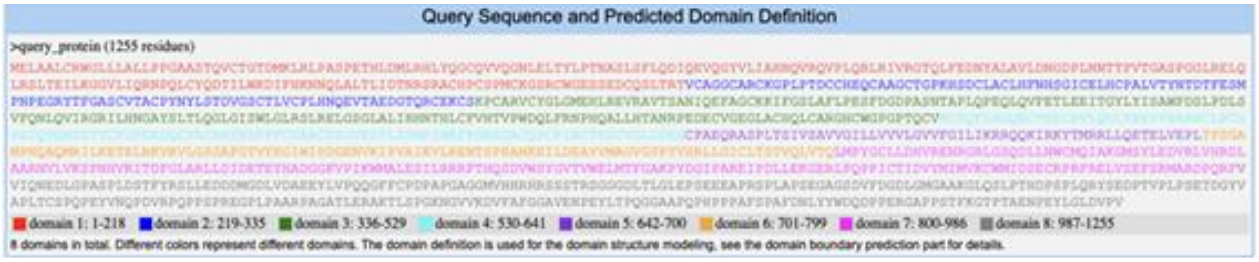


Figure 2. Domain section from I-TASSER prediction model

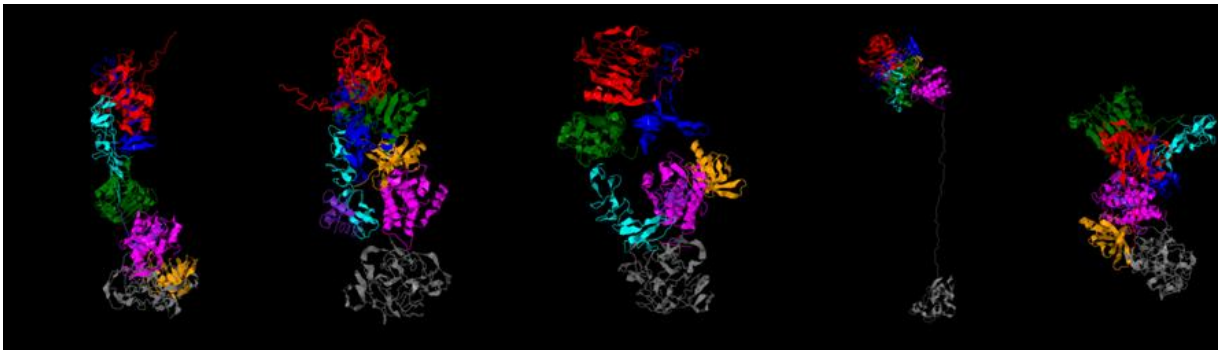


Figure 3. Top 5 I-TASSER model

The algorithm then evaluated potential binding pockets by considering cavity size, shape complementarity, and conserved residues commonly found in ligand-binding regions. Once the binding site had been identified, I-TASSER utilized various scoring functions and algorithms to assess the protein's local environment and surface properties to determine the ligands. These included solvent accessibility, electrostatic potentials, and residue interactions. The possible protein hits were generated through the structural similarity from its specific region and overall sequence with TM-Score and RMSD, respectively. These protein hits were referred to the top 5 predicted ligands and their confidence scores, as seen in Figure 4.

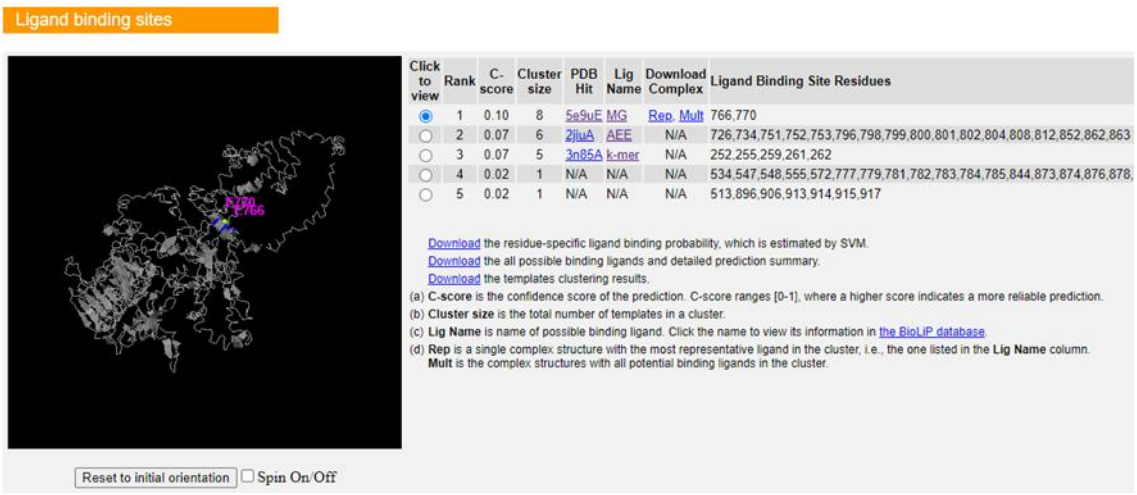


Figure 4. I-TASSER results

By analyzing the predicted binding pockets of the HER2 receptor, we identified the top three ligands along with their corresponding C-scores. The C-score measures confidence in the predicted binding affinity between the ligand and the receptor (Ain et al., 2015) . A higher C-score indicates a more vital predicted binding interaction.

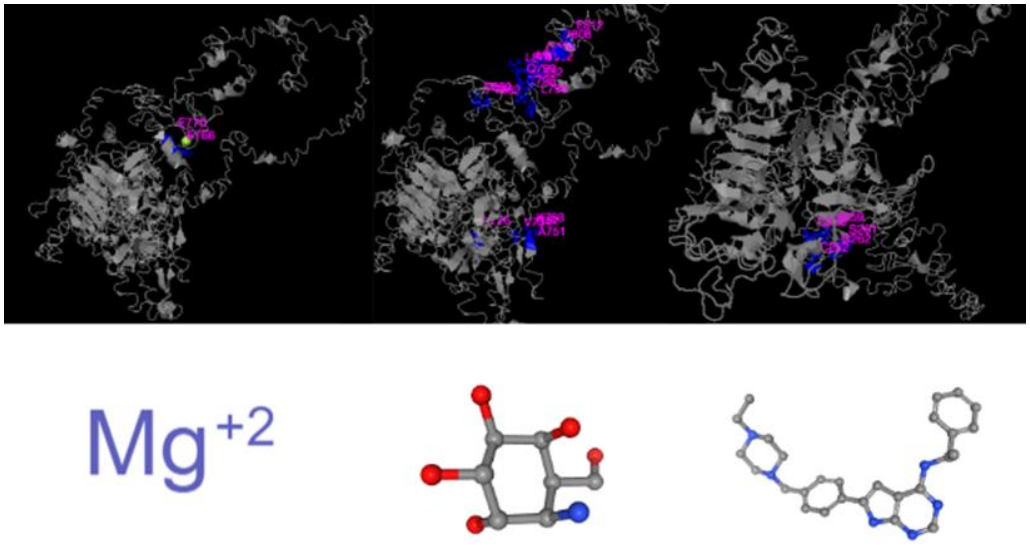


Figure 5. HER2 binding pocket and potential ligands

EGNA

To verify the results obtained from I-TASSER, the binding affinity of each ligand was assessed EGNA (Empirical Graph Neural Network). However, the binding affinity calculations could not be determined for Mg+2 and AEE ligands due to issues during the program execution. Only the UUU ligand exhibited successful binding affinity estimation, with a pKd value of 3.42 mol, as depicted in Figure 6. The binding of these three ligands was further analyzed using molecular docking to predict the quality and favored orientation of their potential interactions with HER2 (Roy et al., 2015).

Figure 6. Binding affinity of UUU ligand

CB-Dock2

The molecular docking results were evaluated using Vina scores, indicating the binding quality between proteins and small molecules. The Vina scoring function combines different types of atomic interactions, including steric, hydrophobic, and hydrogen bonding (Quiroga & Villarreal, 2016) . These interactions are calculated and adjusted based on the number of rotatable bonds, considering the entropic penalties associated with molecular flexibility. A lower (more negative) value in this scoring system signifies stronger binding (Xue et al., 2022).

The ligand AEE exhibited the strongest binding affinity with a Vina score of -10.6. This indicates a favorable interaction between the protein and AEE, suggesting its potential as a promising ligand (Figure 7A). On the other hand, the ligand UUU showed a slightly weaker binding affinity with a Vina score of -6.1. It is worth noting that the binding affinity calculation for the ligand Mg+2 encountered an error and could not be determined (Figure 7B). This suggests a limitation in accurately assessing the binding between Mg+2 and the protein using the Vina scoring method.

PyMOL

PyMOL was used to visualize the protein complex bonded with the ligand. Figures 8 depict the visualization of the HER2 complex bound with the ligand, UUU, and AEE, respectively. The residue for the chain was specified from the prediction of the ligand binding site obtained in I-TASSER. For UUU, it is specified at 252, 255, 259, 261, 262 while 726, 734, 751, 752, 753, 796, 798, 799, 800, 801, 802, 804, 808, 812, 852, 862, 863 for AEE. The color is based on their secondary structures on the atoms: yellow for the sheet region, purple for the helical region, and green for the rest of the region. The red color in the center showed the surface structure of our ligand. Although it displayed how the HER2 bound

in more positional specific atoms with the UUU ligand, the AEE ligand was still confirmed to be better based on the previous confirmation through molecular docking, as the energy of the binding affinity evaluated also considered their overall orientation to another

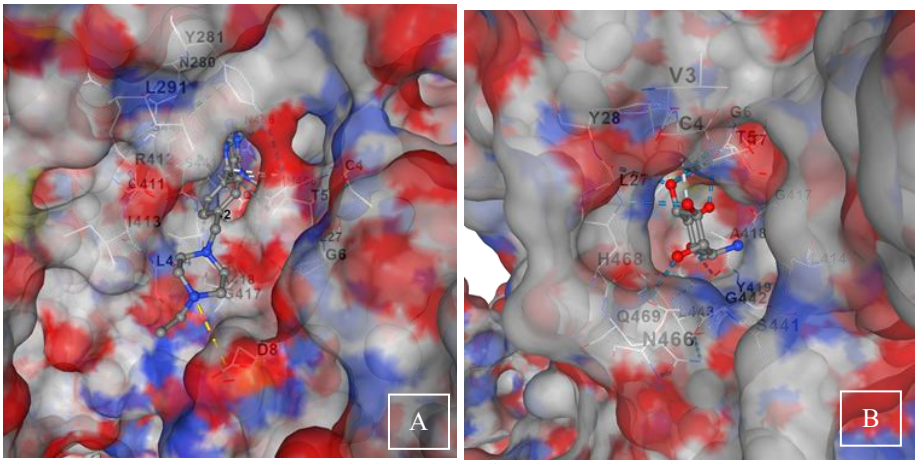


Figure 7. Molecular docking of AEE ligand (A) and UUU ligand (B)

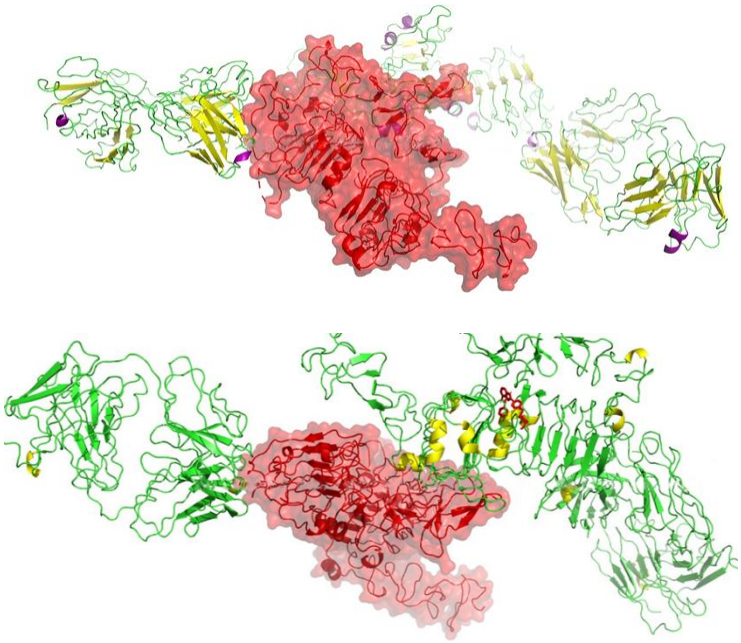


Figure 8. Visualization of HER2 Complex with UUU ligand (A) and AEE ligand (B)

Conclusion

Overall, this study utilized protein structure prediction methods, specifically I-TASSER, to identify the potential ligand of HER2. The results showed that I-TASSER successfully predicted the structure of HER2, providing valuable insights into its protein-protein interactions relevant to breast cancer developments. The predicted ligands AEE, UUU, and Mg+2 were identified based on their binding potential to HER2. Subsequent analysis of PPI and molecular docking data shows AEE as having significant potential as a ligand to the HER2 receptor. This finding can advance our understanding of the protein's role in breast cancer, contributing to the development of improved detection methods and treatments for affected individuals. Furthermore, further studies could expand upon this research by conducting molecular docking and dynamics simulations to identify potential drugs that could effectively target breast cancer. Moreover, in depth molecular simulation will eventually provide leads to the direction of wet laboratory assays. Such investigation would offer valuable insights and pave the way for developing novel therapeutic strategies.

Acknowledgement

First and above all, we praise God, the Almighty, for providing us with this opportunity and granting us the capability to proceed successfully with this research paper. This piece of work will never be accomplished without His blessings. We also would like to thank to Research and Community Service Department (LPPM) of Indonesia International Institute for Life Sciences for their heartfelt support.

References

Ain, Q. U., Aleksandrova, A., Roessler, F. D., & Ballester, P. J. (2015). Machine-learning scoring functions to improve structure-based binding affinity prediction and virtual screening. WIREs Computational Molecular Science, 5(6), 405–424. https://doi.org/10.1002/wcms.1225

Akram, M., Iqbal, M., Daniyal, M., & Khan, A. U. (2017). Awareness and current knowledge of breast cancer. Biological Research, 50(1), 33. https://doi.org/10.1186/s40659-017-0140-9

Albagoush, S. A., & Limaiem, F. (2023). HER2. PubMed; StatPearls Publishing.

Brown, D., Smeets, D., Székely, B., Larsimont, D., Szász, A. M., Adnet, P.-Y., Rothé, F., Rouas, G., Nagy, Z. I., Faragó, Z., Tőkés, A.-M., Dank, M., Szentmártoni, G., Udvarhelyi, N., Zoppoli, G., Pusztai, L., Piccart, M., Kulka, J., Lambrechts, D., … Desmedt, C. (2017). Phylogenetic analysis of metastatic progression in breast cancer using somatic mutations and copy number aberrations. Nature Communications, 8(1), 14944. https://doi.org/10.1038/ncomms14944

Burley, S. K., Bhikadiya, C., Bi, C., Bittrich, S., Chao, H., Chen, L., Craig, P. A., Crichlow, G. V, Dalenberg, K., Duarte, J. M., Dutta, S., Fayazi, M., Feng, Z., Flatt, J. W., Ganesan, S., Ghosh, S., Goodsell, D. S., Green, R. K., Guranovic, V., … Zardecki, C. (2023). RCSB Protein Data Bank (RCSB.org): delivery of experimentally-determined PDB structures alongside one million computed structure models of proteins from artificial intelligence/machine learning. Nucleic Acids Research, 51(D1), D488–D508. https://doi.org/10.1093/nar/gkac1077

CDC. (2022). What is breast cancer? Centers for Disease Control and Prevention. Https://Www.Cdc.Gov/Cancer/Breast/Basic_info/What-Is-Breast-Cancer.Htm.

Friend Tambunan, U. S., & Parikesit, A. A. (2014). Cracking the genetic code of human virus by using open source bioinformatics tools. Malaysian Journal of Fundamental and Applied Sciences, 6(1). https://doi.org/10.11113/mjfas.v6n1.175

Gutierrez, C., & Schiff, R. (2011). HER2: Biology, Detection, and Clinical Implications. Archives of Pathology & Laboratory Medicine, 135(1), 55–62. https://doi.org/10.5858/2010-0454-RAR.1

Hashemifar, S., Neyshabur, B., Khan, A. A., & Xu, J. (2018). Predicting protein-protein interactions through sequence-based deep learning. Bioinformatics (Oxford, England), 34(17), i802–i810. https://doi.org/10.1093/bioinformatics/bty573

Iqbal, N., & Iqbal, N. (2014). Human Epidermal Growth Factor Receptor 2 (HER2) in Cancers: Overexpression and Therapeutic Implications. Molecular Biology International, 2014, 1–9. https://doi.org/10.1155/2014/852748

Jespers, W., Åqvist, J., & Gutiérrez-de-Terán, H. (2021). Free Energy Calculations for Protein–Ligand Binding Prediction (pp. 203–226). https://doi.org/10.1007/978-1-0716-1209-5_12

Jia, Z., Yan, L., Ren, Z., Wu, L., Wang, J., Guo, J., Zheng, L., Ming, Z., Zhang, L., Lou, Z., & Rao, Z. (2019). Delicate structural coordination of the Severe Acute Respiratory Syndrome coronavirus Nsp13 upon ATP hydrolysis. Nucleic Acids Research, 47(12), 6538–6550. https://doi.org/10.1093/nar/gkz409

Kreutzfeldt, J., Rozeboom, B., Dey, N., & De, P. (2020). The trastuzumab era: current and upcoming targeted HER2+ breast cancer therapies. American Journal of Cancer Research, 10(4), 1045–1067.

Linding, R., Jensen, L. J., Diella, F., Bork, P., Gibson, T. J., & Russell, R. B. (2003). Protein disorder prediction: implications for structural proteomics. Structure (London, England : 1993), 11(11), 1453–1459. https://doi.org/10.1016/j.str.2003.10.002

Liu, Y., Grimm, M., Dai, W., Hou, M., Xiao, Z.-X., & Cao, Y. (2020). CB-Dock: a web server for cavity detection-guided protein–ligand blind docking. Acta Pharmacologica Sinica, 41(1), 138–144. https://doi.org/10.1038/s41401-019-0228-6

Liu, Y., Yang, X., Gan, J., Chen, S., Xiao, Z.-X., & Cao, Y. (2022). CB-Dock2: improved protein–ligand blind docking by integrating cavity detection, docking and homologous template fitting. Nucleic Acids Research, 50(W1), W159–W164. https://doi.org/10.1093/nar/gkac394

Mandal, A. (2019). History of breast cancer. News. Https://Www.News-Medical.Net/Health/History-of-Breast-Cancer.Aspx#:~:Text=Ancient%20Greece%20and%20Egypt, Breast%20that%20has%20no%20cure.

Meli, R., Morris, G. M., & Biggin, P. C. (2022). Scoring Functions for Protein-Ligand Binding Affinity Prediction Using Structure-based Deep Learning: A Review. Frontiers in Bioinformatics, 2. https://doi.org/10.3389/fbinf.2022.885983

National Cancer Institute. (2011). HER2-Positive. Https://Www.Cancer.Gov/Publications/Dictionaries/Cancer-Terms/Def/Her2-Positive.

Paramkusham, S., Rao, K. M. M., & Prabhakar Rao, B. V. V. S. N. (2013). Early stage detection of breast cancer using novel image processing techniques, Matlab and Labview implementation. 2013 15th International Conference on Advanced Computing Technologies (ICACT), 1–5. https://doi.org/10.1109/ICACT.2013.6710511

Quiroga, R., & Villarreal, M. A. (2016). Vinardo: A Scoring Function Based on Autodock Vina Improves Scoring, Docking, and Virtual Screening. PLOS ONE, 11(5), e0155183. https://doi.org/10.1371/journal.pone.0155183

Roy, A., Kucukural, A., & Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. Nature Protocols, 5(4), 725–738. https://doi.org/10.1038/nprot.2010.5

Roy, K., Kar, S., & Das, R. N. (2015). Other Related Techniques. In Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment (pp. 357–425). Elsevier. https://doi.org/10.1016/B978-0-12-801505-6.00010-7

Russell, R. B., Alber, F., Aloy, P., Davis, F. P., Korkin, D., Pichaud, M., Topf, M., & Sali, A. (2004). A structural perspective on protein-protein interactions. Current Opinion in Structural Biology, 14(3), 313–324. https://doi.org/10.1016/j.sbi.2004.04.006

Seo, S., Choi, J., Park, S., & Ahn, J. (2021). Binding affinity prediction for protein–ligand complex using deep attention mechanism based on intermolecular interactions. BMC Bioinformatics, 22(1), 542. https://doi.org/10.1186/s12859-021-04466-0

Wicaksono, A., & Parikesit, A. A. (2023). Molecular Docking and Dynamics of SARS-CoV-2 Programmed Ribosomal Frameshifting RNA and Ligands for RNA-Targeting Alkaloids Prospecting. HAYATI Journal of Biosciences, 30(6), 1025–1035. https://doi.org/10.4308/hjb.30.6.1025-1035

Xia, C., Feng, S.-H., Xia, Y., Pan, X., & Shen, H.-B. (2023). Leveraging scaffold information to predict protein–ligand binding affinity with an empirical graph neural network. Briefings in Bioinformatics, 24(1). https://doi.org/10.1093/bib/bbac603

Xue, Q., Liu, X., Russell, P., Li, J., Pan, W., Fu, J., & Zhang, A. (2022). Evaluation of the binding performance of flavonoids to estrogen receptor alpha by Autodock, Autodock Vina and Surflex-Dock. Ecotoxicology and Environmental Safety, 233, 113323. https://doi.org/10.1016/j.ecoenv.2022.113323

Yang, J., & Zhang, Y. (2015). Protein Structure and Function Prediction Using I-TASSER. Current Protocols in Bioinformatics, 52(1). https://doi.org/10.1002/0471250953.bi0508s52

Zhang, C., Lane, L., Omenn, G. S., & Zhang, Y. (2019). Blinded Testing of Function Annotation for uPE1 Proteins by I-TASSER/COFACTOR Pipeline Using the 2018–2019 Additions to neXtProt and the CAFA3 Challenge. Journal of Proteome Research, 18(12), 4154–4166. https://doi.org/10.1021/acs.jproteome.9b00537

Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. BMC Bioinformatics, 9(1), 40. https://doi.org/10.1186/1471-2105-9-40

Zhou, X., Zheng, W., Li, Y., Pearce, R., Zhang, C., Bell, E. W., Zhang, G., & Zhang, Y. (2022). I-TASSER-MTD: a deep-learning-based platform for multi-domain protein structure and function prediction. Nature Protocols, 17(10), 2326–2353. https://doi.org/10.1038/s41596-022-00728-0