Cell. Mol. Biol. (ISSN: 1165-158X)

Journal Homepage: www.cellmolbiol.org

Cellular and Molecular Biology

Original Article

Identification of small-molecule inhibitors targeting HER2 for breast cancer therapy using structure-based virtual screening, ADMET profiling, and molecular dynamics simulations

Hanadi M Baeissa, Israa Jameel Hakeem*

Department of Biological Sciences, College of Science, University of Jeddah, Jeddah, Postal Code 21589 Saudi Arabia

Article Info

Article history:

Received: November 17, 2025
Accepted: January 23, 2026
Published: March 31, 2026

Use your device to scan and read the article online

Abstract

Human epidermal growth factor receptor 2 (HER2) is one of the key drivers of oncogenesis in HER2+ breast cancer. The overexpression resulted in uncontrollable cell division and additional tumor growth. Despite the current clinical therapeutics such as Tucatinib, Lapatinib, and Neratinib, the demand for novel inhibitors still exists. The research aimed to employ a multi-step computational framework for the discovery of highly effective small molecule inhibitors. The crystal structure of HER2 (PDB: 3PPO) was retrieved, and subsequently, virtual screening of small molecules was carried forward with the Mcule library. The screening resulted in filtering 100 compounds, out of which 45 were found to be non-toxic and had good ADME properties. BOILED-egg analysis predicted favorable HIA and BBB permeability properties surpassing the performance of clinically approved HER2 inhibitors. The docking studies singled out 11 compounds that had a high affinity towards the protein. On filtering those, the top two compounds achieved the best binding energies of -9.9 kcal/mol and -9.8 kcal/mol, respectively, outperforming the reference inhibitors. Moving ahead, 100 ns MD simulations confirmed enhanced stability, lower RMSF values, higher compactness, better solvation free energy, and persistent hydrogen bonding throughout the simulation period, showing the comparable results to the inhibitor. In comparison, findings reveal that Compound 1 hits structurally and energetically outperforming the existing therapeutics. Compound 2 showed a little less impact than compound 1, but can be considered as well as surpassing all the parameters successfully. Thereby, making the potential lead candidates for further experimental validation in breast cancer therapy.

Keywords: HER2+, Breast cancer, Therapeutics, Structure-based drug designing, Inhibitors, BOILED-Egg model, Molecular dynamics.

1. Introduction

The initial point for the formation of complex multicellular organisms made up of thousands to millions of cells is single-fertilized eggs. Each cell lineage accumulates mutations as a result of several mutational mechanisms during the lifetime of the organism [1]. Different mutational mechanisms cause the accumulation of genetic changes in cell lineages during the lifetime of the organism. While many of these mutations are innocuous, some allow cells to evade regular cell cycle regulators, resulting in uncontrolled growth, proliferation, and eventually cancer.

Cancer is described not only as one but as a multi-stage disease, two aspects of a series of genetic changes that are necessary both to start and further the tumor [2, 3]. Larger and longer-living animals are statistically more prone to cancer because each cell division is a chance for mutation. In many cases, tumor cells have DNA & RNA almost identical to an organism, thus they escape the immune system, especially if it is immune compromised [4, 5]. By 2020, cancer became the leading cause of death worldwide, with an estimated 10 million fatalities [6].

Breast cancer is the most common malignancy in women worldwide and a major health concern in developing countries such as India. In 2020 alone, 2.3 million women were diagnosed with the disease, accounting for 685,000 fatalities [7]. The disease is extremely diverse, beginning in either the glandular tissues (lobules and ducts) or the stromal supporting tissues of the breast. Breast cancer classification has progressed from early history markers like estrogen receptor expression in the 1980s to modern molecular profiling, such as HER2 [8].

Current molecular subtypes are identified using DNA microarrays that include Luminal A, B, HER-2, and Triple Negative breast cancer. Luminal A (ER+, PR+, and HER2-), the most common subtype that carries the prognosis characterized as being Hormone Receptor Positive. These tumors tend to grow more slowly. Luminal B (ER+, PR-, and HER2-), are hormone receptor positive but are differentiated by being HER2 negative, containing higher expression profiles. HER- 2 + (ER-, PR-, and HER2+) tumors are described by overexpression of the human epidermal growth

* Corresponding author.
E-mail address: ijhakeem@uj.edu.sa; israajhakeem@gmail.com (I. J. Hakeem).
Doi: http://dx.doi.org/10.14715/cmb/2025.72.3.6

factor receptor 2 protein and are typically ER-. Last, Triple Negative Breast Cancer (TNBC) is characterized by the absence of ER-, PR-, and HER2-. These tumors have worse prognosis than any other molecular subtype [9–12].

Human epidermal growth factor receptor 2 (HER2) is a 185kDa protein belonging to the EGFR family. It's made up of an extracellular domain with four sub-domains, a transmembrane region, and an intracellular tyrosine kinase domain. The major distinguishing feature of HER2 from other members of the family is that it does not have a ligand [13, 14]. It relies on homodimerization or heterodimerization to initiate it signaling. Amplification or overexpression of the HER2 gene results in an excessive number of growth-promoting proteins, activation of downstream pathways such as PI3K/AKT and Ras/Raf/MEK/MAPK, and inhibition of apoptosis, which in turn leads to tumor invasion, local progression, and distant metastasis[15, 16]. The systematic diagram of tumor progression has been illustrated in Figure 1.

The current therapeutic landscape for HER2+ breast cancer has been transformed by the development of agents that target HER2 from various angles, as HER2+ tumors are driven by the overproduction of the protein that instructs them to grow uncontrollably. The current scenario includes a dual blockage approach for the deactivation of signaling pathways. The incorporation of monoclonal antibodies, such as pertuzumab, used to block extracellular dimerization and tyrosine kinase inhibitors (TKIs) such as lapatinib and neratinib to reduce intracellular autophosphorylation[17, 18]. This study aims to identify novel small-molecule HER2 inhibitors from the MCULE libraries using advanced computational techniques, including Structure-Based Virtual Screening (SBVS), ADMET profiling, multi-score molecular docking, and molecular dynamics (MD) simulations.

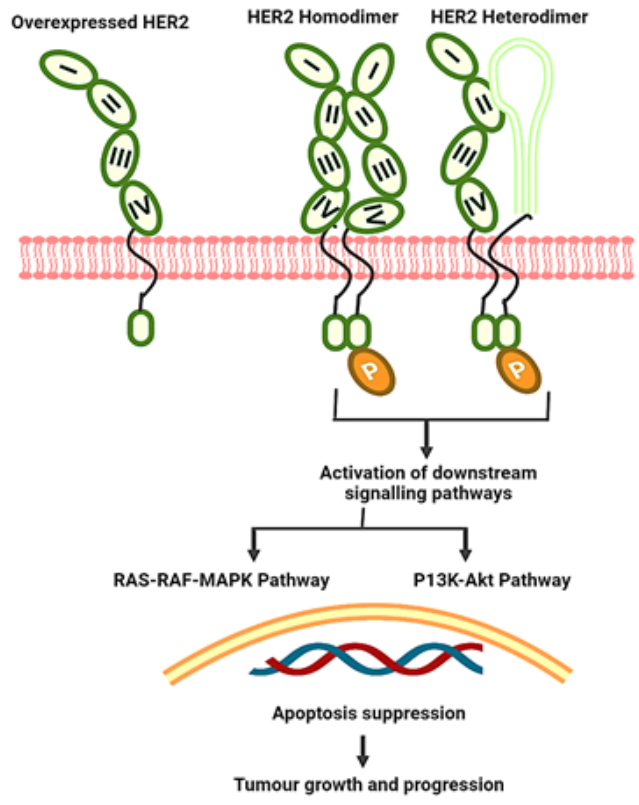


Fig. 1. Pathway of Overexpression of the HER2 gene.

2. Materials and methodology

2.1. Retrieval and preparation of HER2 protein

3D crystal structure (2.25 Å) of the motor domain of human HER2 (PDB: 3PPO) was retrieved from the Research Collaboratory for Structural Bioinformatics (RCSB) protein data bank (PDB) RCSB.org [19]. The apoprotein was prepared for adopted docking tools by removing undesired heteroatoms, ions, and molecules, e.g., adenosine-5'-diphosphate, Mg+2, and H2O. The CHARMm force field was then assigned to optimize and minimize the 3D structure. Further, the Prop Ka https://www.ddl.unimi.it/vegaol/propka.htm utility tool was used to keep the protonation states of the protein at pH 7.4 [20–22].

2.2. Structure-based virtual screening

The Mcule https://mcule.com/, an interactive drug discovery portal provides a rapid and cost-effective way to identify novel drug candidates [23]. It was used for high-throughput virtual screen of small molecules from its library containing over a hundred million synthetically accessible ligand molecules. A systematic query was set in the SBVS input search box. Lipinski rule of five (RO5: MW ≤ 500 Da; HBD ≤ 5; HBA ≤ 10; LogP ≤ 5), polar surface area (PSA: 140 Å^2), molar refractivity (40 ≤ MR ≤ 130), rotatable bonds (RBs) ≤ 10, N and O atoms <=10, and number of rings (NRs) ≤ 7 filtrations were considered in the Mcule's SBVS workflow. Sample size, diversity selection, and similarity search cutoff value for the query string were respectively taken as 1000, 100, and 0.85.

The Open Babel linear fingerprint search algorithm was employed to execute the SBVS run as an advanced option [24, 25]. The remaining search parameters were assigned default values of the SBVS workflow.

2.3. Retrieval of small molecules

The 2D structure of known inhibitors, viz., Lapatinib (CID: 208908), Neratinib (CID: 9915743), and Tucatinib (CID: 51039094) were accessed in structure data format (SDF) with the help of the PubChem database https://pubchem.ncbi.nlm.nih.gov/ [26, 27].

2.4. Conversion of small molecules & optimization

The retrieved molecules require conversion from 2D to 3D to proceed further. The BIOVIA Discovery Studio was utilized for the conversion of these molecules to the required 3D structure format. All selected inhibitors were energetically optimized and minimized similarly to the HER2 protein.

2.5. Toxicity assessment

To remove the unnecessary elements and prevent the spread of other infections or diseases, the need for checking toxicity before clinical trials is a major requirement. To follow up on the requirement, the toxicity checker facility of MCULE was used to examine the presence of substructures, scaffolds, or moieties, also referred to as toxophores, found in chemicals that are profoundly correlated with unwanted properties commonly associated with human and environmental toxicity. Toxicity Checker uses the robust SMARTS (SMILES arbitrary target specification) toxic matching rules. SMARTS is an extension of SMILES, an acronym of Simplified Molecular Input Line Entry System, used to represent the computer-ready chemical structure of molecules [28, 29].

2.6. Profiling of ADME

At the early stage of the drug discovery process, moving with ADME-satisfied compounds is imperative to find desired results in wet-lab experiments. The filtered ligands from toxicity profiling criteria were assessed on ADME parameters, including physicochemical properties, lipophilicity, pharmacokinetics, drug-likeness, and medicinal chemistry attributes. The consignment was held with the freely available web tool Swiss-ADME [30, 31].

2.7. Molecular interaction studies

The PDB structure of HER2 was imported to the AutoDock Vina (ADV) interface. The binding box was set at a grid size of 40 Å in each x, y, and z was drawn to cover the protein binding site area. The dimensions of the grid box were: In x (12.2935 Å), y (24.3815 Å), and z (34.9715 Å) axes. ADV parameters for the maximum number of binding modes per ligand and exhaustiveness were taken as the default, i.e., 2 and 1. Minimum free energy of binding (ΔG) was chosen as a selective parameter to find one of the best binding affinities and poses for each of the ligand molecules docked into the active site of HER2 [32–35].

2.8. Molecular dynamics simulation studies

The MD simulation studies enable trapping the motion of every atom moving in a defined system by providing the necessary information. On the basis of several parameters, the top two compounds were selected for simulation studies with the known inhibitor (Control) at 300K. The simulation was accomplished by the GROMACS (GROningen Machine for Chemical Simulations) with version 5.1.2.at the MM level using GROMACS 5.1.2 [36, 37]. The gmx grep module was used to extract the desired ligands from the complexes. Server CGenFF (CHARMm general force field) integrated the hit compounds with control. All docked complexes were solvated within a dodecahedron box of water molecules with a margin of 10 Å TIP3 (Transferable Intermolecular Potential). Additionally, the required boundaries were applied, and charges were neutralized accordingly by integrating NA+ & Cl- ions. The system underwent energy minimization for 25,000,000 steps using the steepest descent algorithm. Gradually, the system temperature was raised from 0 to 300 K during their equilibration of 100 ps duration at constant NVT (Number of particles at constant volume & constant temperature) and NPT (Number of particles at constant pressure and temperature). Following the equilibration stage, the particle mesh was combined with the Ewald method [38, 39]. 100ns was applied to conduct the simulation studies as the last step. Different trajectories, such as root-mean square deviation, root-mean square fluctuation, solvent accessible surface area, radius of gyration and free energy solvation, were performed individually for extracting the final conclusion of the investigation [40-42].

3. Results

3.1. Structure-based virtual screening

The structure-based virtual screening to filter and screen out the best molecules from a library of a million molecules extracted from the Mcule's library. A total of 100 molecules were selected based on the best fit into the desired binding pockets of our target HER2 protein. These molecules successfully fulfilled the criteria of RO5 and were then subjected to further steps like toxicity check and so on. The selected top molecules are listed in Table S1 (Supplement).

3.2. Profiling of toxicity

The toxicity alerts are an indication that depicts which structure may be harmful for human use, as it may lead to several serious problems. The molecules that had poor pharmacokinetics and risk of having mutagenic moieties were filtered out to prevent any serious risk. The profiling was conducted via online web-based Mcule's toxicity risk assessment [43]. The toxicity checker uses SMARTS algorithms to investigate patterns [44]. As a result, out of 100 compounds, only 49 compounds stood out. The safe filtered compounds are listed in Table 1.

3.3. Predicting ADME parameters

After the toxic filtration, the next pivotal step in drug designing is to check for the ADME (Absorption, Distribution, Metabolism, Excretion) parameters. To proceed, the BOILED or intestinal egg model of SwissADME was employed, which predicts the threshold of two key ADME parameters, HIA (Human Intestinal Absorption) and BBB (Blood Brain Barrier) [45–47]. The intestinal model replicates the original egg consist of a yellow and white part. The yellow color indicates the physicochemical locations for BBB permeation, and white represents gastrointestinal tract absorption, respectively. Fortunately, all hits were recorded within BBB permeation and HIA absorption. Notably, these compounds resulted in better than selected inhibitors. Lapatinib and Neratinib inhibitors were observed outside the model, reflecting negligible absorption but apparently successful in involving BBB permeation. Whereas, Tucatinib showed some absorption but negligible BBB permeation. The detailed boiled egg illustration is shown in Figure 2.

WLOGP ≤ 5.88 and TPSA ≤ 131.6, for lipophilicity and apparent polarity, were accessed. Blue and red dots represent P–gp+ and P-gp- molecules. The blue hits that were rejected by BBB were backout. 45 molecules fulfilled the criteria of ADME filtration.

3.4. Analysis of physicochemical properties

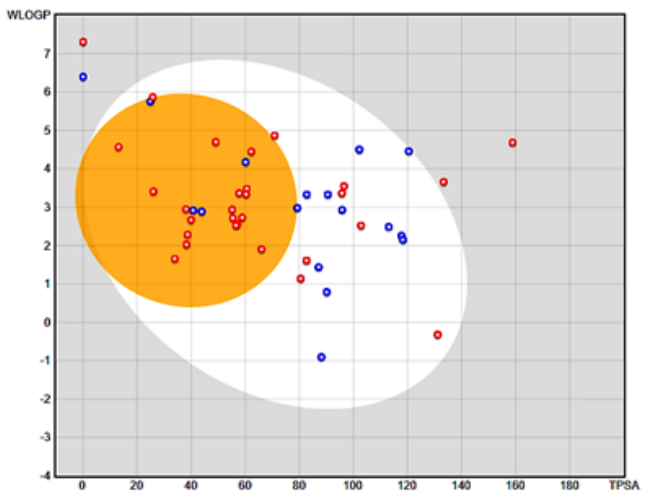


Fig. 2. Evaluation of passive gastrointestinal absorption (HIA) and blood-brain barrier (BBB) access in the function of the position of ligand hits and selected inhibitors using the BOILED-Egg model.

Table 1. Safe hits molecules filtered from Mcule toxicity checker

S. No.	Ligands	S. No.	Ligands	Toxicity Criteria
1	MCULE-9386350598	26	MCULE-2132659354	Pass
2	MCULE-9083675198	27	MCULE-7256289530	Pass
3	MCULE-2082061687	28	MCULE-4042431771	Pass
4	MCULE-9514587152	29	MCULE-8752115040	Pass
5	MCULE-8016227151	30	MCULE-5180351602	Pass
6	MCULE-4420319124	31	MCULE-1328505063	Pass
7	MCULE-7002197612	32	MCULE-7045636868	Pass
8	MCULE-3840783663	33	MCULE-5613555964	Pass
9	MCULE-4544985274	34	MCULE-7443327368	Pass
10	MCULE-9433162170	35	MCULE-1355912164	Pass
11	MCULE-8028742001	36	MCULE-2860715238	Pass
12	MCULE-4638340739	37	MCULE-1986483046	Pass
13	MCULE-2116801361	38	MCULE-6081257779	Pass
14	MCULE-5671860004	39	MCULE-4616972123	Pass
15	MCULE-6264951729	40	MCULE-4297750503	Pass
16	MCULE-5984253496	41	MCULE-4862665716	Pass
17	MCULE-8109420073	42	MCULE-1028628572	Pass
18	MCULE-7057324381	43	MCULE-8084678579	Pass
19	MCULE-8782640423	44	MCULE-1633822804	Pass
20	MCULE-1558627453	45	MCULE-5302780192	Pass
21	MCULE-2454948452	46	MCULE-4063820706	Pass
22	MCULE-7093625408	47	MCULE-1726158340	Pass
23	MCULE-6232409085	48	MCULE-9917566553	Pass
24	MCULE-8046841408	49	MCULE-7280319389	Pass
25	MCULE-9038810560	-	-	Pass

- NA

The prediction of physicochemical properties depicts several parameters. The hits showed favorable potentials for oral bioavailability sticks to Lipinski's rule of five. The bioavailability score is a determinant of the oral absorption of a drug molecule [48, 49]. The score for bioavailability score was recorded as 0.55 for all compounds showing excellent absorption. The maximum molecular weight recorded for the molecules (MCULE-8782640423) is 400.58 g/mol. The molecular saturation is powered by the fraction of $F\ sp^3$. The ideal range considered for the same is > or equal to 0.25 [50, 51]. However, the result showed that 30 compounds exhibited a value less than 0.25. Similarly, TPSA values also showed favorable values within 20-130 A^o range, presenting good membrane permeability. However, the rotatable bond was depicted quite low. Whereas, all the inhibitors exhibited higher MW and rotatable bonds.

3.5. Lipophilicity & solubility

The lipophilicity can be stated as a partition coefficient between n-octanol and water (Log Po/w), describing the most important physical property of a compound. The mathematical formula for lipophilicity is: Log Po/w = log $(ligand)_{octanol}$ / $(Ligand)_{water}$. The evaluation is based on three descriptors by SwissADME XlogP3, WlogP, & MlogP, having the reference range of -2 to 5, -0.4 to 5.8, and (≤ 4.15) [52, 53]. Supportively, the majority of the compounds gave favorable results. However, the violations had been noted for the inhibitor Lapatinib. Tucatinib remained within the range limit. Neratinib also showed a less lipophilic profile than other compounds.

In the case of solubility, the three descriptors by SwissADME were ESOL, Ali, and SILICOS-IT models were employed. The range orders are <−10 insoluble, < -6 poorly soluble, < -4 moderately soluble, <-2 soluble, and < 0 highly soluble [45, 54]. The model ESOL showed the least moderate solubility in compounds. MCULE-7093625408, MCULE 6081257779, and MCULE-9038810560 reached the moderately soluble stage, suggesting favorable properties. Lapatinib, Neratinib classified as insoluble, and Tucatinib showed poor solubility by the SILICOS-IT model. The detailed analysis is shown in Table S2.

3.6. Molecular docking

For molecular interaction studies, AutoDock Vina by Mcule is used to identify the best hits compared with the reference inhibitors. All 45 ligand hits, along with the three inhibitors Lapatinib, Neratinib, and Tucatinib, were docked to our target protein. A total of 11 hits were obtained out of 45, containing the highest energies recorded. The highest energy obtained by MCULE-9083675198 was -9.9 kcal/mol, while all three inhibitors showed less energy comparatively. Among the inhibitors, Tucatinib showed highest binding energy -8.22 kcal/mol. These top molecules were further studied for simulation studies. The best binding energies evaluated by the top 11 compounds are listed in Table 2.

The highest docked complexes were visualized to check their binding pocket with the help of BIOVIA Discovery Studio. The top two compounds, MCULE-9083675198

Table 2. Molecular interaction analysis of top eleven ligand hits with reference inhibitors

S. No.	Ligand	Binding Energies (kcal/mol)
	MCULE- 9083675198	-9.9
	MCULE-2082061687	-9.8
	MCULE-9514587152	-9.7
	MCULE-8016227151	-9.5
	MCULE-4420319124	-9.1
	MCULE-7002197612	-9.0
	MCULE-3840783663	-8.9
	MCULE-4544985274	-8.9
	MCULE-9433162170	-8.8
	MCULE-8028742001	-8.8
	MCULE-4638340739	-8.8
	Tucatinib- inhibitor	-8.22
	Neratinib- inhibitor	-8.12
	Lapatinib- inhibitor	-7.76

and MCULE-2082061687, were picked up with the best inhibitor, Tucatinib. The 2D and 3D diagrams of all such complexes were prepared and shown in Figure 3A-F. Different interactions were observed in complexes such as Van Der Waals, Conventional hydrogen bond, carbon-hydrogen bond, alkyl, pi-alkyl, and some pi-sigma and pi-alkyl bonds.

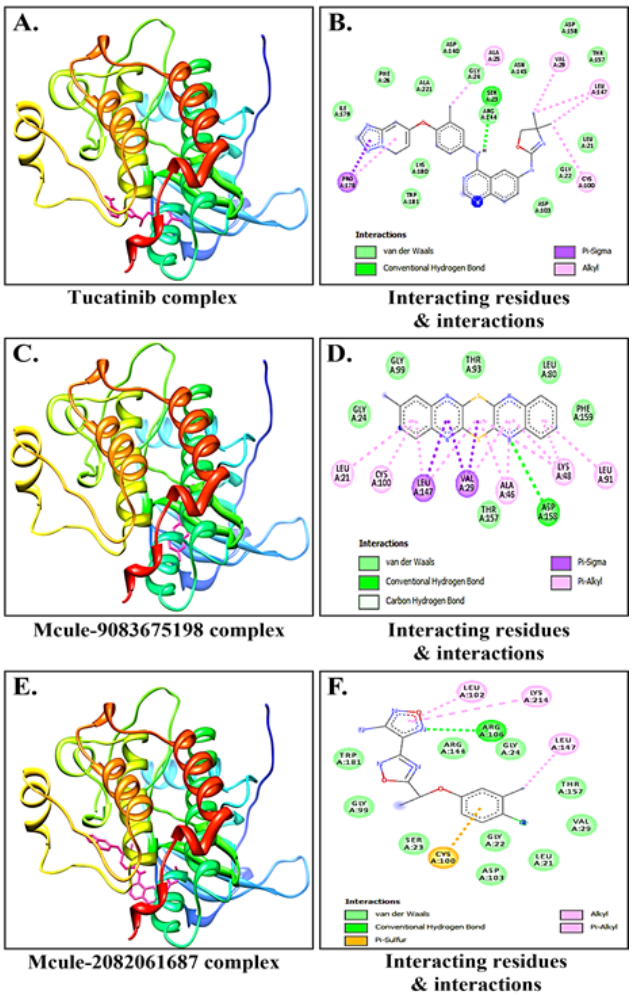


Fig. 3. (A) 3D view of the Tucatinib-bound complex **(B)** 2D diagram showing interacting residues and interactions. **(C)** 3D view of the Mcule-9083675198 -bound complex **(D)** 2D diagram showing interacting residues and interactions **(E)** 3D view of the MCULE-2082061687-bound complex **(F)** 2D diagram showing interacting residues and interactions.

3.7. Molecular dynamics studies

The final hits MCULE-9083675198 (Compound 1) and MCULE-2082061687 (Compound 2), having the energies -9.9 kcal/mol and -9.8 kcal/mol, respectively, obtained from several parameters, went for dynamic studies along with the known inhibitors. The studies were conducted at a duration of 100ns using the GROMACS server. Different trajectories were depicted to analyze the complex.

3.8. Stability analysis: Root-Mean Square Deviation

To determine the protein stability, the trajectory RMSD was employed**.** The study revealed the support of the compounds rather than the inhibitor. The inhibitor showed various spikes about 1.0nm-1.4nm, while compound 1 showed stability around 0.35 nm, and compound 2 showed maintaining stability and peaking near 0.90 nm, displaying the adjustments. The RMSD plot has been illustrated below (Figure 4A). Different colors were used to mark compounds for better analysis.

3.9. Fluctuation analysis: Root-Mean Square Fluctuations

The parameter analyzes how much the compound residues fluctuated over time, depicting the stability of a complex. The peaks in the plot (Figure 4B) indicate high mobility or disordered regions. The average fluctuations noted for compound 1 were 0.20 nm displaying notable

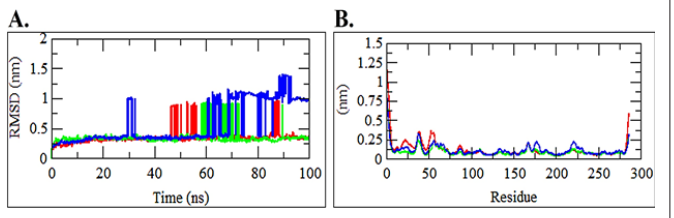


Fig. 4. (A) Stability of complexes, Red represents compound 1, Green represents compound 2, and Blue represents Inhibitor **(B)** RMSF Plot representing flexibility of the protein. peaks indicates flexible loop regions. C1 Is represented in red color, C2 is represented in green color, and the inhibitor is displayed in blue color.

peak at the loop regions, 0.15 nm for compound 2, exhibiting the lowest fluctuations, and 0.75 nm and 0.50 nm for the inhibitor, exhibiting the highest sharp peak, indicating less control and instability.

3.10. Analysis of solvent-accessible surface area and solvation free energy

The compactness of protein-ligand complexes was determined by SASA plots. All the compounds remained consistent, ranging between 10nm^2 and 16nm^2. C1 showed the average lowest value of ~12.5 nm^2, showing compactness compared to other compounds. C2 showed the range value of 13.0 nm^2. While the inhibitor showed an average value of 13.2 nm^2, showing the highest value. The overlapping confirms that the core remains protected from the solvent. Furthermore, the energy change is predicted by the delta G (Solvation free energy). The negative values indicate favorable solvation stability. C1 and C2 showed good stability, fluctuating between -17.5 and -16.8 kJ/mol. The sharp dips reaching -24.5 and -22 kJ/mol indicate high energetic stability. On the other hand, the most fluctuating compound was the inhibitor having the highest average solvation energy. Fluctuation is depicted close to -15.5 kJ/mol, showing less energetically favorable than that of Mcule hits. The plots are illustrated in the diagram (Figure 5).

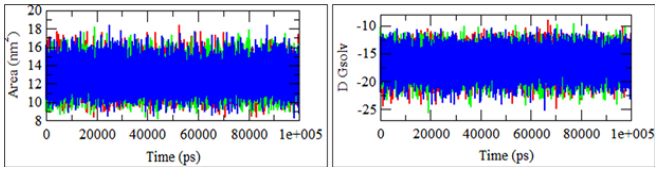


Fig. 5. SASA & solvation plots representing compactness and solvation free energy of the protein. C1 is represented in red color, C2 is represented in green color, and the inhibitor is displayed in blue color.

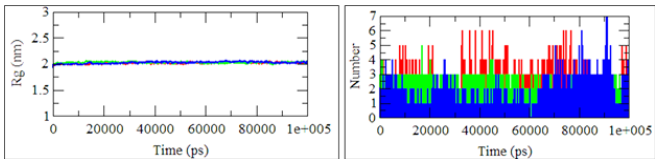


Fig. 6. **(A)** Radius of gyration (Rg) of complexes illustrating green as C1, red as C2, and blue as inhibitor maintains compactness and conformations throughout the simulation. **(B)** Hydrogen bond plot of complexes illustrating the evolution of bond counts. Red as C1, green as C2, and blue as inhibitor over the 100ns of the simulation.

3.11. Radius of gyration

The Rg indicates the compactness and overall structure integrity of a protein during the simulation studies over the period of time. All the compounds exhibited stability. C1 and Inhibitor was recorded around 2 nm maintain high stability. C2 showed marginally but maintains stability at 2.02 nm. The lack of significant drift toward higher values above 2.5 nm for any compound confirms that the protein remains in a compact. The illustrated plot depicts clears visualization of compounds (Figure 6A).

3.12. Hydrogen bonds

Throughout the 100ns simulation, the stability of complexes and consistency of these bonds provide the specificity of compounds. Based on the analysis, C1 maintained a high and consistent number of hydrogen bonds, fluctuating between 3 & 6, showing the highest bonds of 6. C2 in green color exhibits a steady profile exceeding at 4. Adding on, the reference inhibitor shows fluctuations of bonds between 0 and 4 and drops frequently, depicting structural shifting. (Figure 6B).

4. Discussion

The present research employed a comprehensive structure-based computational drug design pipeline for the identification of novel small molecules as inhibitors targeting the HER2 domain, which is an oncogenic key driver in breast cancer. Multiple parameters, such as SBVS, toxicity assessment, ADME profiling using BOILED-Egg, multimolecular docking, and molecular dynamics simulation studies, have been unified to uncover the potential lead candidates with improved binding, feasibility, and stability compared with clinically proven inhibitors [55]. Mcule library was assessed for SBVS containing millions of compounds to refine with 100 molecules exhibiting optimal best fitting with our target protein PDB: 3PPO. Three inhibitors were taken, namely Tucatinib, Lapatinib, and Neratinib.

In all of these compounds, several compounds demonstrated stronger predicted binding energies compared to the known inhibitors. The initial step of screening enriched the ligands capable of favorable interactions with the binding pockets of the target protein. This early enrichment of high-affinity ligands demonstrated the effectiveness of the virtual screening protocol with strong target specificity. Additionally, toxicity profiling using SMART-based filters further focused on refining the hit list by excluding carcinogenic molecules. Almost half of the compounds were dumped out, and only safe hit molecules were retained. These safe hits prioritize the safety evaluation at an initial stage of the drug discovery process.

Evaluating ADME profiling by the SWISS-ADME BOILED-Egg model provides critical insights into the parameters, such as distribution, absorption, metabolism, and excretion, to explore the potential of shortlisted compounds. This method predicted that all the retained hit molecules were found positioned within the favorable region for HIA (Human Intestinal Absorption), mainly demonstrating BBB (Blood Brain Barrier) permeability. Unfortunately, the notable remarks were observed by inhibitors depicting pharmacokinetic limitations, which are essential attributes for therapeutics. Poor gastrointestinal absorption was observed by Lapatinib and Neratinib, while Tucatinib restricted BBB penetration. Apart from that, physicochemical analysis profiling of hit compounds further confirmed the BOILED-Egg results. All the compounds had molecular weight, TPSA, and bioavailability scores (0.55) within the desired satisfactory range, suggesting good bioavailable drugs. However, some compounds had a low fraction of SP values, but praising the balance between molecular rigidity and flexibility maintained in comparison with inhibitors was notable. The results of lipophilicity and solubility showed that most compounds were within an acceptable logP range and moderate solubility. On the other hand, Lapatinib was detected as an outlier exceeding the recommended limits, with non-uniform bioavailability and poorly soluble or completely insoluble. On an obser-

ving note, MCULE-9083675198 & MCULE-2082061687 remained outperforming the investigations. Further, molecular docking studies support the investigation of these two compounds as the strongest binders (-9.9 kcal/mol and -9.8 kcal/mol) to our targeted protein, exhibiting the best energies compared to inhibitors (Table 2). Detailed interaction analysis revealed the formation of stable hydrogen bonds, Van Der Waals interactions with key residues closely resembled or exceeded those of reference inhibitors, highlighting the suppression of activity (Figure 3).

Now, to confirm the research studies, these molecules were further validated through molecular dynamics simulation studies at 100 ns. Different trajectories were analyzed, such as RMSD, RMSF, SASA, Solvation, Rg, and hydrogen bonds (Figure 4-8). Compounds maintained good stability, whereas the reference inhibitor exhibited notable fluctuations. SASA and solvation free energy supported these findings. The radius of gyration also confirmed the compactness throughout the simulation, with no unfolding events marked. The hydrogen bonding analysis strongly contributes to the findings. The consistent superiority of these compounds across all the parameters underscores the potential of HER2-targeted therapies, providing a strong foundation for experimental validation.

The study implied a comprehensive structure-based computational study to screen for small-molecule inhibitors of the HER2 kinase domain. The studies implemented structure-based virtual screening, toxicity, and ADME profiling via the BOILED Egg model, docking studies, and molecular dynamics studies, resulting in obtaining the two suitable candidates, MCULE-9083675198 & MCULE-2082061687. The egg model predicted favorable HIA and BBB permeability for the screened hits successfully retrieved by the Mcule library, highlighting the superior pharmacokinetics profile, gastrointestinal profile, and minimum BBB penetration compared with known inhibitors. Molecular docking results demonstrated that the top two compounds had the best binding energies, whereas the dynamics studies confirmed the conformational stability, minimal fluctuations, favorable solvation energy, and hydrogen bonding interactions. In view of that, MCULE-9083675198 gave a consistently robust performance across all the findings. Thereby, contribute as a promising next therapeutic HER2+ inhibitor for breast cancer therapy and provide foundations for wet-lab experimental validation in breast cancer therapy.

Abbreviation

HER2: Human epidermal growth factor receptor; **TNBC**: Triple Negative Breast Cancer; **TKIs:** Tyrosine kinase inhibitors; **SBVS**: Structure-Based Virtual Screening; **MD:** Molecular dynamics; **RCSB:** Research Collaboratory for Structural Bioinformatics; **PDB:** Protein data bank; **NRs:** Number of rings; **ADV:** AutoDock Vina; **ADME:** Absorption, Distribution, Metabolism, Excretion; **HIA**: Human Intestinal Absorption; **BBB:** Blood Brain Barrier.

Conflict of interests: The authors declared no conflict of interest.

Consent for publications

The author read & approved the final manuscript for publication.

Ethics approval and consent to participate

This research did not include any experiments involving humans or animals.

Acknowledgements

The authors sincerely thank the Department of Biological Sciences, College of Science, University of Jeddah, Saudi Arabia, for its academic support and research facilities. They also appreciate the assistance of the faculty members, technical staff, and colleagues in completing this work.

Generative AI Statement

The author used generative AI tools, including Grammarly and QuillBot, to improve language and clarity, and accepts full responsibility for the accuracy and integrity of the content.

Authors' contributions

Hanadi MB conceived and designed the study, analyzed the data, and prepared the final draft of the manuscript. Israa JH analyzed the results and prepared the initial draft. All authors read and approved the final version of the manuscript.

References

1. Brunet T, King N (2017) The origin of animal multicellularity and cell differentiation. Dev Cell 43: 124-140 doi: 10.1016/j.devcel.2017.09.009
2. Brown JS, Amend SR, Austin RH, Gatenby RA, Hammarlund EU, Pienta KJ (2023) Updating the definition of cancer. Mol Cancer Res 21: 1142-1147 doi: 10.1158/1541-7786.MCR-23-0411
3. Mercadante AA, Kasi A (2025) Genetics, Cancer Cell Cycle Phases. StatPearls Publishing, Treasure Island (FL)
4. Seluanov A, Gladyshev VN, Vijg J, Gorbunova V (2018) Mechanisms of cancer resistance in long-lived mammals. Nat Rev Cancer 18: 433-441 doi: 10.1038/s41568-018-0004-9
5. Vincze O, Minias P, Corthay A, Ronget V, Lemaitre JF, Gaillard JM (2025) Immunological surveillance against cancer across mammals. Nat Commun 16: 1058 doi: 10.1038/s41467-025-56134-7
6. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F (2021) Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 71: 209-249 doi: 10.3322/caac.21660
7. Arnold M, Morgan E, Rumgay H, Mafra A, Singh D, Laversanne M, Vignat J, Gralow JR, Cardoso F, Siesling S, Soerjomataram I (2022) Current and future burden of breast cancer: global statistics for 2020 and 2040. Breast 66: 15-23 doi: 10.1016/j.breast.2022.08.010
8. Harbeck N, Penault-Llorca F, Cortes J, Gnant M, Houssami N, Poortmans P, Ruddy K, Tsang J, Cardoso F (2019) Breast cancer. Nat Rev Dis Primers 5: 66 doi: 10.1038/s41572-019-0111-2
9. Orrantia-Borunda E, Anchondo-Nuñez P, Acuña-Aguilar LE, Gómez-Valles FO, Ramírez-Valdespino CA (2022) Subtypes of breast cancer. In: Breast Cancer. Exon Publications, Brisbane, pp 31-42 doi: 10.36255/exon-publications-breast-cancer-subtypes
10. Uikey G, Ahirwar H, Kori R, Shivhare S, Agrawal S, Verma M (2022) Metastatic trends in breast cancer according to molecular subtypes, estrogen receptor, progesterone receptor and human epidermal growth factor receptor 2neu status. J Med Sci Health 8: 100-105 doi: 10.46347/jmsh.2022.v08i02.004

11. El Fatemi H, Chahbouni S, Jayi S, Mrabti H, Errihani H, El Gueddari BK, Amarti A (2012) Luminal B tumors are the most frequent molecular subtype in breast cancer of North African women: an immunohistochemical profile study from Morocco. Diagn Pathol 7: 170 doi: 10.1186/1746-1596-7-170
12. Al-thoubaity FK (2020) Molecular classification of breast cancer: a retrospective cohort study. Ann Med Surg (Lond) 49: 44-48 doi: 10.1016/j.amsu.2019.11.021
13. Ye Z, Yu H, Zhu L, Xu Y, Wang J, Chen X (2025) HER2 and epigenetic modifications in tumorigenesis: mechanisms and therapeutic implications. Front Oncol 15: 1523456 doi: 10.3389/fonc.2025.1523456
14. Rubin E, Shan K, Dalal S, Hua S, Chu D, Liu X, Liu J (2024) Molecular targeting of the human epidermal growth factor receptor-2 (HER2) genes across various cancers. Int J Mol Sci 25: 1064 doi: 10.3390/ijms25021064
15. Pan L, Li J, Xu Q, Huang M, Wang Y, Liu S (2024) HER2/PI3K/AKT pathway in HER2-positive breast cancer: a review. Medicine (Baltimore) 103: e38508 doi: 10.1097/MD.0000000000038508
16. Raftery NB, Ward M, Ravi N, O'Connell F, Creavin B, O'Sullivan J (2025) Modulation of the tumour microenvironment by HER2 in oesophagogastric adenocarcinoma: implications for tumour progression, therapeutic resistance, and clinicopathological outcomes. Cancers (Basel) 17: 3987 doi: 10.3390/cancers17243987
17. Mercogliano MF, Bruni S, Mauro FL, Schillaci R (2023) Emerging targeted therapies for HER2-positive breast cancer. Cancers (Basel) 15: 1987 doi: 10.3390/cancers15071987
18. Drittone D, Lucci C, Esposito L, De Laurentiis M, Giuliano M, Arpino G (2025) Impact of anti-HER2 therapies on overall survival in patients with HER2-positive metastatic breast cancer: focusing on intracranial efficacy of emerging treatments. Cancers (Basel) 17: 3520 doi: 10.3390/cancers17213520
19. Berman HM (2000) The Protein Data Bank. Nucleic Acids Res 28: 235-242 doi: 10.1093/nar/28.1.235
20. Khan FI, Lai D, Anwer R, Shoaib M, Hu X, Wei DQ (2020) Identifying novel sphingosine kinase 1 inhibitors as therapeutics against breast cancer. J Enzyme Inhib Med Chem 35: 172-186 doi: 10.1080/14756366.2019.1692830
21. Bas DC, Rogers DM, Jensen JH (2008) Very fast prediction and rationalization of pKa values for protein-ligand complexes. Proteins 73: 765-783 doi: 10.1002/prot.22102
22. Khan S, Ali S, Rahman I, Rehman A (2021) Identification of potential lead molecules against dibenzo[a,l]pyrene-induced mammary cancer through targeting cytochrome P450 1A1, 1A2, and 1B1 isozymes. Biointerface Res Appl Chem 12: 1096-1109 doi: 10.33263/BRIAC121.10961109
23. Kiss R, Sandor M, Szalai FA (2012) http://Mcule.com: a public web service for drug discovery. J Cheminform 4(Suppl 1): P17 doi: 10.1186/1758-2946-4-S1-P17
24. Shakil S (2019) Molecular interaction of investigational ligands with human brain acetylcholinesterase. J Cell Biochem 120: 11820-11830 doi: 10.1002/jcb.28463
25. Morris GM, Goodsell DS, Huey R, Olson AJ (1996) Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4. J Comput Aided Mol Des 10: 293-304 doi: 10.1007/BF00124499
26. Kim S (2021) Exploring chemical information in PubChem. Curr Protoc 1: e217 doi: 10.1002/cpz1.217
27. Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, Han L, He J, He S, Shoemaker BA, Wang J, Yu B, Zhang J, Bryant SH (2016) PubChem Substance and Compound databases. Nucleic Acids Res 44: D1202-D1213 doi: 10.1093/nar/gkv951
28. Sliwoski G, Kothiwale S, Meiler J, Lowe EW (2014) Computational methods in drug discovery. Pharmacol Rev 66: 334-395 doi: 10.1124/pr.112.007336
29. Mriganka N, Subhasish Tech GB (2020) Toxicity detection in drug candidates using simplified molecular-input line-entry system. In: Advances in Intelligent Computing and Communication. Springer, Singapore, pp 123-132
30. Ancuceanu R, Lascu BE, Drăgănescu D, Dinu M (2025) In silico ADME methods used in the evaluation of natural products. Pharmaceutics 17: 1002 doi: 10.3390/pharmaceutics17081002
31. Daina A, Michielin O, Zoete V (2017) SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep 7: 42717 doi: 10.1038/srep42717
32. Khan MKA, Akhtar S, Arif JM (2018) Structural insight into the mechanism of dibenzo[a,l]pyrene and benzo[a]pyrene-mediated cell proliferation using molecular docking simulations. Interdiscip Sci 10: 653-673 doi: 10.1007/s12539-017-0221-3
33. Muteeb G, Rehman MT, AlAjmi MF, Almansour NM, Alqahtani AS (2022) Identification of a potential inhibitor (MCULE-8777613195-0-12) of New Delhi metallo-β-lactamase-1 (NDM-1) using in silico and in vitro approaches. Molecules 27: 5930 doi: 10.3390/molecules27185930
34. Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 31: 455-461 doi: 10.1002/jcc.21334
35. Attique SA, Hassan M, Usman M, Atif RM, Mahboob S, Al-Ghanim KA, Bilal M, Nawaz MZ (2019) A molecular docking approach to evaluate the pharmacological properties of natural and synthetic treatment candidates for use against hypertension. Int J Environ Res Public Health 16: 923 doi: 10.3390/ijerph16060923
36. Salo-Ahen OMH, Alanko I, Bhadane R, Bonvin AMJJ, Honorato RV, Hossain S, Juffer AH, Kabedev A, Lahtela-Kakkonen M, Larsen AS, Lescrinier E, Marimuthu P, Mirza MU, Mustafa G, Nunes-Alves A, Pantsar T, Saadabadi A, Singaravelu K, Vanmeert M (2020) Molecular dynamics simulations in drug discovery and pharmaceutical development. Processes 9: 71 doi: 10.3390/pr9010071
37. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl E (2015) GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 1-2: 19-25 doi: 10.1016/j.softx.2015.06.001
38. Wells BA, Chaffee AL (2015) Ewald summation for molecular simulations. J Chem Theory Comput 11: 3684-3695 doi: 10.1021/acs.jctc.5b00493
39. Stenberg S, Stenqvist B (2020) An exact Ewald summation method in theory and practice. J Phys Chem A 124: 3943-3946 doi: 10.1021/acs.jpca.0c02391
40. Vanommeslaeghe K, Hatcher E, Acharya C, Kundu S, Zhong S, Shim J, Darian E, Guvench O, Lopes P, Vorobyov I, MacKerell AD Jr (2010) CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. J Comput Chem 31: 671-690 doi: 10.1002/jcc.21367
41. Vanommeslaeghe K, MacKerell AD Jr (2012) Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. J Chem Inf Model 52: 3144-3154 doi: 10.1021/ci300363c
42. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC (2005) GROMACS: fast, flexible, and free. J Comput Chem 26: 1701-1718 doi: 10.1002/jcc.20291
43. Khan MKA, Ahmad S, Rabbani G, Shahab U, Khan MS (2022) Target-based virtual screening, computational multiscoring docking and molecular dynamics simulation of small molecules as promising drug candidate affecting kinesin-like protein KIFC1.

Cell Biochem Funct 40: 451-472 doi: 10.1002/cbf.3707
44. Zhang J, Li H, Zhang Y, Wang Z, Chen Y (2025) Computational toxicology in drug discovery: applications of artificial intelligence in ADMET and toxicity prediction. Brief Bioinform 26: bbaf012 doi: 10.1093/bib/bbaf012
45. Daina A, Zoete V (2016) A BOILED-egg to predict gastrointestinal absorption and brain penetration of small molecules. ChemMedChem 11: 1117-1121 doi: 10.1002/cmdc.201600182
46. Chedik L, Mias-Lucquin D, Bruyere A, Fardel O (2017) In silico prediction for intestinal absorption and brain penetration of chemical pesticides in humans. Int J Environ Res Public Health 14: 708 doi: 10.3390/ijerph14070708
47. Daina A, Zoete V (2016) A BOILED-Egg to predict gastrointestinal absorption and brain penetration of small molecules. ChemMedChem 11: 1117-1121 doi: 10.1002/cmdc.201600182
48. Wu K, Kwon SH, Zhou X, Fullerton J, Li Y (2024) Overcoming challenges in small-molecule drug bioavailability: a review of key factors and approaches. Int J Mol Sci 25: 13121 doi: 10.3390/ijms252313121
49. Benet LZ, Hosey CM, Ursu O, Oprea TI (2016) BDDCS, the Rule of 5 and drugability. Adv Drug Deliv Rev 101: 89-98 doi: 10.1016/j.addr.2016.05.007
50. Roskoski R (2023) Rule of five violations among the FDA-approved small molecule protein kinase inhibitors. Pharmacol Res 191: 106774 doi: 10.1016/j.phrs.2023.106774
51. Krzyzanowski A, Pahl A, Grigalunas M, Waldmann H (2023) Spacial score—a comprehensive topological indicator for small-molecule complexity. J Med Chem 66: 12739-12750 doi: 10.1021/acs.jmedchem.3c01023
52. Mannhold R, Poda GI, Ostermann C, Tetko IV (2009) Calculation of molecular lipophilicity: state-of-the-art and comparison of LogP methods on more than 96,000 compounds. J Pharm Sci 98: 861-893 doi: 10.1002/jps.21494
53. Sun Y, Hou T, He X, Wang J, Li Y (2023) Development and test of highly accurate endpoint free energy methods. 2: prediction of logarithm of n-octanol-water partition coefficient (logP) for druglike molecules using MM-PBSA method. J Comput Chem 44: 1300-1311 doi: 10.1002/jcc.27099
54. Gallardo A, Gutierrez MR, Gomez LA, Villalobos P, Sanchez M (2024) A comparative analysis on the potential anticancer properties of tetrahydrocannabinol, cannabidiol, and tetrahydrocannabivarin compounds through in silico approach. Asian Pac J Cancer Prev 25: 839-856 doi: 10.31557/APJCP.2024.25.3.839
55. Sait KHW, Alam Q, Anfinan N, Al-Ghamdi O, Malik A, Noor R, Jahan F, Tarique M (2019) Structure-based virtual screening and molecular docking for the identification of potential novel EGFR kinase inhibitors against ovarian cancer. Bioinformation 15: 287-294 doi: 10.6026/97320630015287