

Received: 20.09.2024

Accepted: 31.12.2024

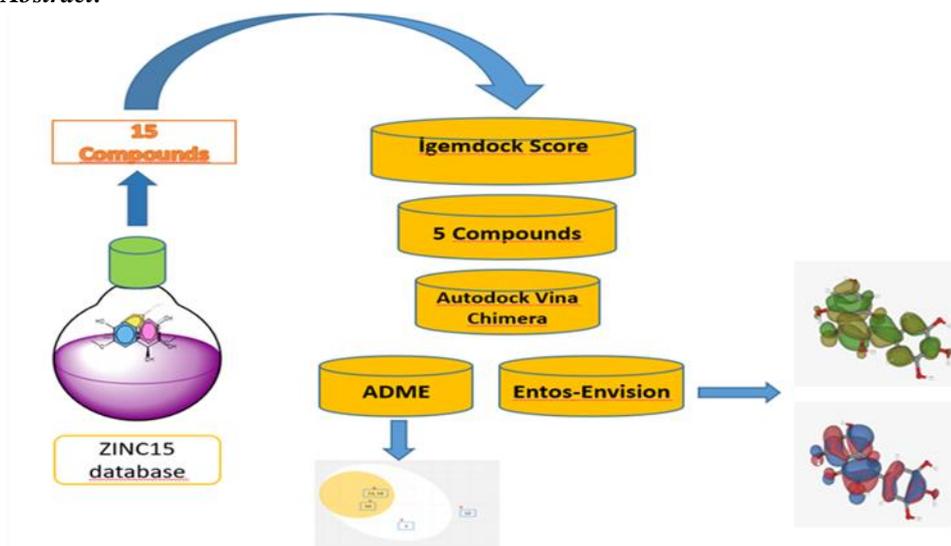
Research Article

Investigation of the Effect of Vanillin-Based Compound Derivatives Against Breast Cancer by Molecular Docking Method and Computational Chemistry, ADME Analysis

Özlem İŞCAN¹

Bartın University, Project and Technology Office, Bartın-Türkiye

Graphical Abstract:



Abstract: In this study, 15 vanillin-derived compounds were downloaded from the ZINC15 database in mol2 format. A molecular docking study was performed using the Igemdock program with PDB ID: 6QXG protein using the Igemdock program. As a result of these processes, the five lowest energy compounds were identified. Accordingly, the best binding molecules are 13, 12, 14, 3 and 10, respectively. Molecular docking was performed with five selected molecules, and Autodock Vina embedded Chimera 1.17.3 programs. The interactions of the molecules with the protein were analysed. ADME studies were also performed. According to the in silico results, this molecule seems promising in terms of its potential to be turned into a medication in the future. However, more detailed in vivo and in vitro studies with these molecules are needed. As a result of these processes, some quantum chemical properties of the lowest energy molecule were analysed with the help of the Entos-Envision web server.

Keywords: 6QXG, ADME, Igemdock, Molecular Docking, Entos-Envision.

1. Introduction

Breast cancer is a disease in which abnormal breast cells grow out of control and form tumours. If left unchecked, the tumours can spread throughout the body and become fatal. Breast cancer cells begin

inside the milk ducts and the breast's milk-producing lobules. The earliest form (in situ) is not life-threatening and can be detected in early stages. Cancer cells can spread into nearby breast tissue (invasion). This creates tumours that cause lumps

¹ Corresponding Authors

e-mail: oiscan@bartin.edu.tr

Özlem İŞCAN

or thickening. In 2022, there were 2.3 million women diagnosed with breast cancer and 670,000 deaths globally [1]. Vanillin is a member of the class of benzaldehydes carrying methoxy and hydroxy substituents at the 3rd and 4th positions, respectively. It has effects as a plant metabolite, an anti-inflammatory agent, a flavouring agent, an antioxidant and an anticonvulsant. It is a member of phenols, a monomethoxybenzene and a member of benzaldehydes. 2D and 3D representations of the vanillin compound are presented (Figure 1) [2]. A study on human cancer cell lines (MCF-7 and A549) showed that vanillin decreased cell viability and the percentage of viable cells in the colony [3]. TS has become a hot topic as a chemotherapeutic agent due to its direct involvement in DNA synthesis. This enzyme causes the conversion of

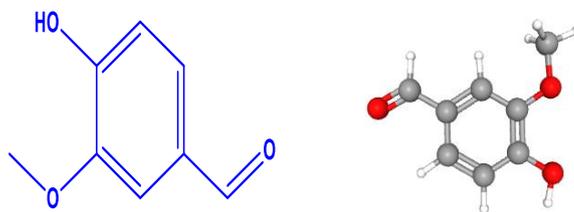


Figure 1. Vanillin compound 2D and 3D structure

In this study, considering the anticancer effects of vanillin on cancer cells, 15 vanillin-derived compounds were selected from the ZINC database [6] to identify potential compounds with anticancer activity. The preferred compounds were subjected to a preliminary screening process using iGEMDOCK [7] software, and the five compounds with the lowest binding affinity were evaluated as potential medication candidates. The suitability and efficacy of these candidate compounds for medication development were further analyzed using methods such as molecular docking and ADME studies. The thymidylate synthase protein with PDB ID 6QXG is a promising target for breast cancer treatment as it disrupts DNA synthesis, halts cancer cell growth, and induces apoptosis [8]. Computational docking techniques were employed to elucidate the binding affinity of the compounds to their target proteins.

As a result of the analyses, the compound with the lowest binding energy and the highest potential as a drug candidate (Compound 12) was further evaluated in detail for its quantum chemical properties using the Entos-Envision software [9].

deoxyuridine monophosphate (dUMP) to deoxythymidine monophosphate (dTTP), which leads to dTTP, a precursor for DNA synthesis in further phosphorylation. Therefore, blocking TS causes DNA damage, leading to cell cycle arrest and induction of apoptosis [4]. In a study, the synthesis, characterisation, pharmacokinetic properties, and molecular docking analyses of novel benzimidazole hybrids derived from the natural product vanillin were conducted to evaluate their anticancer potential comprehensively. Compound 6 emerged as a promising drug candidate due to its high aromatase inhibitory activity ($IC_{50} = 0.064 \mu M$) and potent antiproliferative effects against breast cancer cells [5].

2. Computational Method

2.1. Selection of Ligands

The chemical structures of 15 vanillin derivative ligands were downloaded from the ZINC database in mol2 format [6]. The downloaded compound structures A are presented in the material document.

2.2. Igemdock score

The target protein was selected from the PDB Data Bank, and 15 compounds were obtained from the ZINC database [6]; docking analysis of vanillin ring-containing compounds with the 6QXG protein was performed using the IGEMDOCK program [7]. As a result of this analysis, the five compounds with the lowest binding energies were identified.

2.3. Ligand and receptor preparation

Following the scoring process performed with IGEMDOCK [7], five ligands with the highest binding affinities were selected. The ligand and receptor were prepared using UCSF Chimera software [10]. This process involved sequential steps, including removing non-standard residues, eliminating water molecules, adding hydrogen

Özlem İŞCAN

bonds, and assigning charges. Finally, a minimisation procedure was performed to determine the most stable structure. Molecular docking studies of the five selected ligands were conducted using AutoDock Vina [11].

2.4. ADME studies

The SMILES codes of the ligands obtained from the ZINC database were submitted to the SwissADME [12] server for toxicity prediction. This process predicted the pharmacokinetic, toxicity, and drug-likeness properties of the ligands. Additionally, compliance with critical criteria such as Lipinski's Rule of Five, Veber's Rule, Egan's Rule, and polar surface area (TPSA) was evaluated, along with rotatable bonds and ADME (Absorption) assessments.

2.5. Entos Envision

Following these procedures, quantum chemical analysis was performed on the identified candidate compound using Entos Envision software [9]. As part of the study, molecular structures were optimised to determine their most stable geometries. Bond orders were calculated to evaluate the strength and character of chemical

bonds and HOMO-LUMO energy levels were chosen to analyse electronic properties and reactivity potential. Molecular electrostatic potential (MEP) maps were generated to visualise charge distribution and identify reactive regions within the molecule. All calculations were carried out using Entos Envision software, employing parameters optimised to meet the study's requirements, and the molecule's chemical properties were comprehensively analysed.

3. Results and discussion

3.1. Igemdock Score

The IGEMDOCK program screened vanillin ring-containing compounds against the 6QXG protein. As a first step, 15 vanillin ring-containing compounds were downloaded from the ZINC15 database in mol 2 format. These screened compounds were then scored with Igemdock [7]. Five compounds exhibiting the highest docking scores were recorded as -115.442 kcal/mol, -111.881 kcal/mol, -106.256 kcal/mol, -94.0344 kcal/mol, and -90.1735 kcal/mol, respectively. Igemdock score results are shown (Table 1 and Figure 2).

Table 1. iGEMDOCK scores for virtual screening of 6QXG and ZINC compounds (kcal/mol)

| S NO | Compound | Energy | VDW | HBond | Elec |
|------|---------------------------------|-----------|----------|----------|----------|
| 1 | cav6qyg_UFP-351312441prep-0.pdb | -80.6502 | -53.3171 | -27.3331 | 0 |
| 2 | cav6qyg_UFP-168233797prep-1.pdb | -75.315 | -37.6865 | -37.6291 | 0 |
| 3 | cav6qyg_UFP-168238909prep-0.pdb | -94.0344* | -48.245 | -37.8906 | -7.89884 |
| 4 | cav6qyg_UFP-168248689prep-1.pdb | -78.8419 | -64.5978 | -14.2441 | 0 |
| 5 | cav6qyg_UFP-174975526prep-1.pdb | -80.640 | -60.9791 | -19.6614 | 0 |
| 6 | cav6qyg_UFP-342719641prep-0.pdb | -83.2689 | -59.4009 | -23.868 | 0 |
| 7 | cav6qyg_UFP-175018247prep-0.pdb | -79.291 | -57.9819 | -21.3091 | 0 |
| 8 | cav6qyg_UFP-175017455prep-0.pdb | -80.066 | -54.4744 | -25.5919 | 0 |
| 9 | cav6qyg_UFP-342723043prep-1.pdb | -78.065 | -64.5663 | -13.4988 | 0 |
| 10 | cav6qyg_UFP-346499641prep-0.pdb | -90.1735* | -72.0137 | -18.1598 | 0 |
| 11 | cav6qyg_UFP-168232309prep-1.pdb | -75.023 | -47.1633 | -27.8602 | 0 |
| 12 | cav6qyg_UFP-406721180prep-1.pdb | -111.881* | -71.4985 | -40.3824 | 0 |
| 13 | cav6qyg_UFP-571753276prep-1.pdb | -115.442* | -84.1048 | -31.3376 | 0 |
| 14 | cav6qyg_UFP-571821700prep-0.pdb | -106.256* | -84.347 | -21.909 | 0 |
| 15 | cav6qyg_UFP-574410954prep-1.pdb | -86.9764 | -60.2265 | -26.7499 | 0 |

*lowest five molecule binding site energy

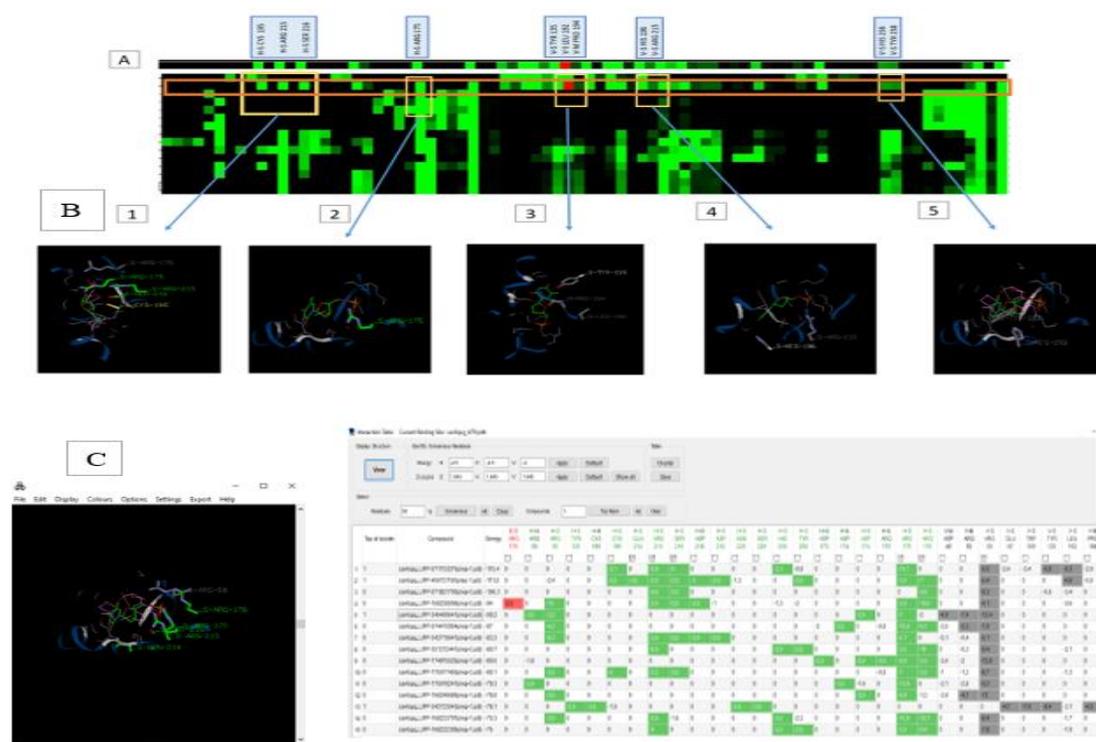


Figure 2. Result of Igemdock (A)Protein-compound interaction profiles (1-5) Some 6qyg- compound interactions B) the interactions and (C) visualisations of pharmacological interactions in the post-screening analysis interface.

3.2. Molecular Docking

Molecular docking studies of the compounds (3, 10, 12-14) were carried out with the help of the AutodockVina 1.1.2 software [11]. UCSF Chimera 1.17.2 [8] and BIOVIA Discovery Studio Visualizer [15] software visualised all obtained results (Table 2). The size of the grid box in AutoDockVina was kept at $20 \text{ \AA} \times 20 \text{ \AA} \times 20 \text{ \AA}$ for 6QXG; 45.00, -36.00, 24.00 \AA along x, y, and z directions, respectively. The resulting file was saved, and visualisation and RMSD analysis were performed with BIOVIA Discovery Studio Visualizer. The RMSD image of the reference compound was calculated by re-docking. PLIP tool [13] was used to identify non-covalent interactions between 6QXG and its ligands and calculated by re-docking 0.18 (Figure 3). Non-covalent interactions are presented (Figure 4 and Tables 3-7). The non-covalent interactions between ligands (3, 10, 12, 13, 14) and the 6QXG protein are presented in detail. These include hydrophobic interactions, hydrogen bonds, π - π stacking, and salt bridges. Hydrophobic interactions between the ligand and protein residues

were analysed, as detailed in Table 3. Key interactions were observed with residues such as LEU, PHE, TRP, and ASP, with distances ranging from 3.51 \AA to 3.97 \AA , highlighting significant binding contributions

3.3. ADME studies

The new SwissADME web tool, which provides free access to a repository of fast but robust predictive models, was used to investigate the ADME properties of ligands. The compounds' polarity and lipophilicity were assessed through analysis with the BOILED-Egg model. [12]. Lipinski's five rules are essential for rational medication development. Any medication molecule that violates even one of the rules has low permeability or poor absorption [14]. It is seen that not all of the compounds whose ADME properties were investigated violate Lipinski's rules. ADME studies are presented (Figure 5).

The ADME analysis of the compounds revealed distinct pharmacokinetic profiles. The gastrointestinal tract absorbs compound 3 but cannot be excreted by cell membrane proteins, with

Özlem İŞCAN

favourable indices except for an exceeded saturation value. Compound 10 can cross the blood-brain barrier, is not excreted by cell membrane proteins, and adheres to Lipinski's rule of five as well as the rules of Ghose, Egan, Veber, and Muegge, with suitable rotatable bonds, size, and polarity. Compound 12 cannot cross the blood-brain barrier or be excreted by cell membrane

proteins; it exhibits good solubility in fats and lipids, suitable size, and rotatable bonds but has insufficient carbon saturation and polarity. Compounds 13 and 14 can cross the blood-brain barrier, are not excreted by cell membrane proteins, and have appropriate rotatable bonds, size, and polarity within carbon atom count and lipophilicity limits.

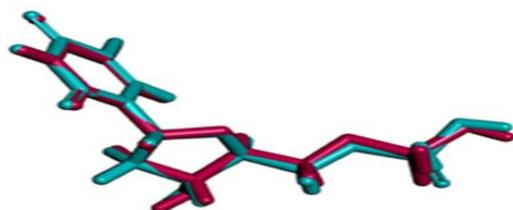


Figure 3. RMSD image of the 6QXG compound

Table 2. AutoDock results

| 6QXG-Compound interaction | | Binding energy (kcal/mol) |
|---------------------------|--|---------------------------|
| <p>3</p> | | -5.2 |
| <p>10</p> | | -5.4 |
| <p>12</p> | | -7.8 |

Özlem İŞCAN

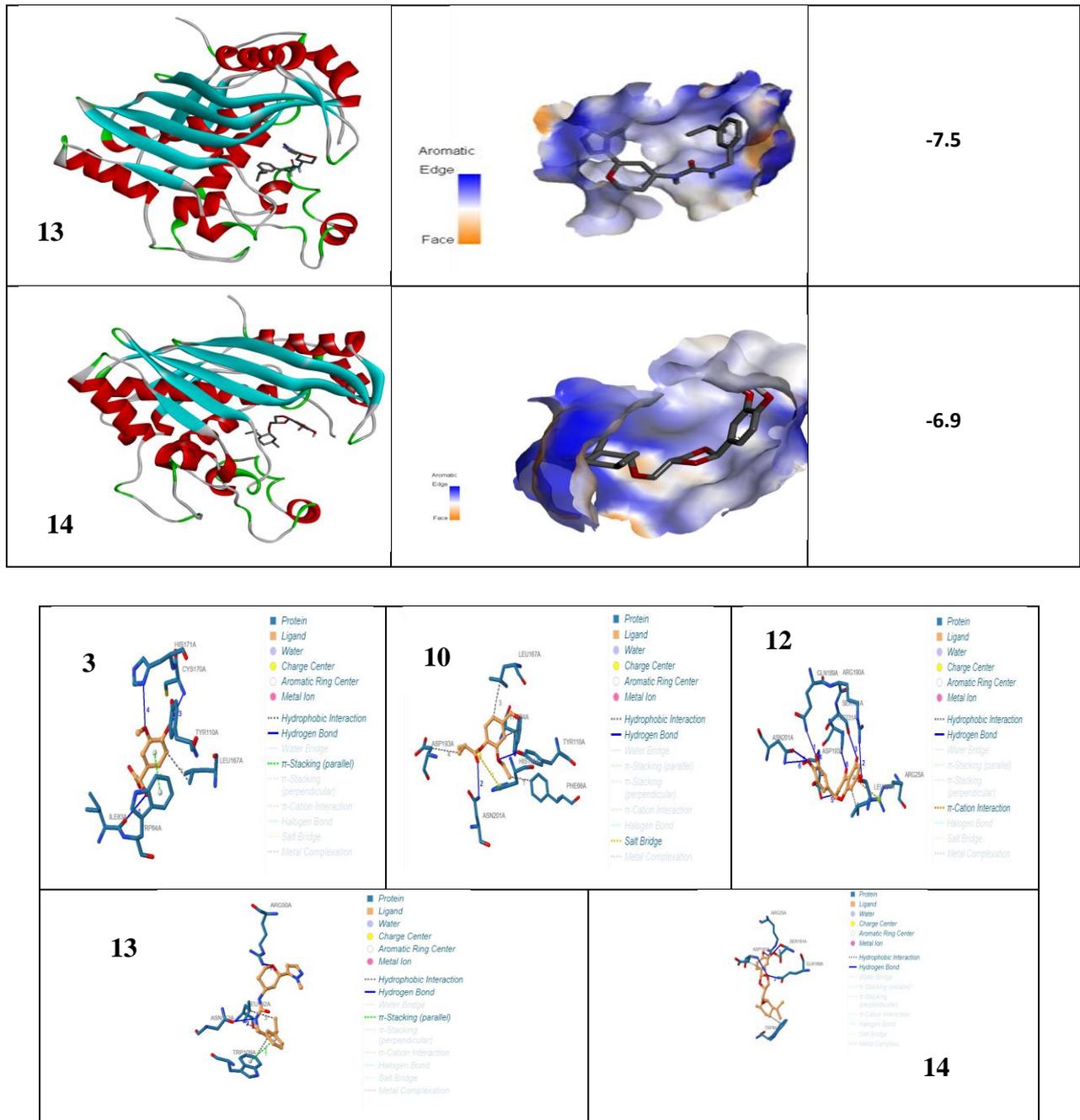


Figure 4. Demonstration of 6QXG-ligand non-covalent interactions

Table 3. Hydrophobic Interaction

| S NO | S NO | S NO | S NO | S NO | S NO | S NO |
|------|------|------|------|------|------|------|
| 3 | 1 | 167A | LEU | 3.76 | 2323 | 1349 |
| 10 | 1 | 66A | PHE | 3.51 | 2333 | 544 |
| | 2 | 84A | TRP | 3.87 | 2320 | 680 |
| | 3 | 167A | LEU | 3.87 | 2322 | 1347 |
| | 4 | 193A | ASP | 3.97 | 2329 | 1551 |
| 12 | 1 | 167A | LEU | 3.64 | 2324 | 1348 |
| | 2 | 193A | ASP | 3.68 | 2333 | 1551 |
| 13 | 1 | 109A | TRP | 3.57 | 2331 | 680 |
| | 2 | 192A | LEU | 3.92 | 2331 | 1349 |

Özlem İŞCAN

| | | | | | | |
|----|---|------|-----|------|------|------|
| | 3 | 192A | LEU | 3.65 | 2332 | 1347 |
| 14 | 1 | 84A | TRP | 3.54 | 2331 | 678 |
| | 2 | 193A | ASP | 3.65 | 2336 | 1551 |

Table 4. Hydrogen Bonds

| S NO | Index | Residue | AA | Distance H-A | Distance D-A | Donor Angle | Protein Donor | Side Chain | Donor Atom | Acceptor Donor |
|------|-------|---------|-----|--------------|--------------|-------------|---------------|------------|------------------------|----------------|
| 3 | 1 | 83A | ILE | 2.31 | 3.22 | 156.89 | x | x | 2326 [O3] | 662 [O2] |
| | 2 | 110A | TYR | 2.23 | 3.11 | 156.85 | ✓ | ✓ | 878 [O3] | 2332 [O3] |
| | 3 | 170A | CYS | 2.22 | 3.17 | 162.21 | ✓ | x | 1364[Nam] | 2332 [O3] |
| | 4 | 171A | HIS | 3.58 | 3.92 | 102.70 | ✓ | ✓ | 1376 [Nar] | 2330 [O3] |
| 10 | 1 | 110A | TYR | 3.44 | 4.07 | 127.17 | ✓ | ✓ | 878 [O3] | 2331 [O3] |
| | 2 | 201A | ASN | 3.01 | 3.65 | 123.89 | ✓ | ✓ | 1611[Nam] | 2327 [O2] |
| 12 | 1 | 189A | GLN | 2.59 | 3.08 | 110.89 | ✓ | ✓ | 1525[Nam] | 2338 [O3] |
| | 2 | 190A | ARG | 2.51 | 3.21 | 128.73 | ✓ | ✓ | 1536 [Ng+] | 2323 [O2] |
| | 3 | 191A | SER | 2.06 | 2.88 | 140.98 | ✓ | ✓ | 1542 [O3] | 2319 [O2] |
| | 4 | 193A | ASP | 2.03 | 3.00 | 167.72 | ✓ | x | 1547[Nam] | 2338 [O3] |
| | 5 | 193A | ASP | 2.38 | 3.16 | 140.74 | ✓ | ✓ | 1554 [O3] | 2329 [O2] |
| | 6 | 201A | ASN | 2.88 | 3.43 | 116.41 | ✓ | ✓ | 1611[Nam] | 2339 [O3] |
| | 7 | 201A | ASN | 2.10 | 2.70 | 120.99 | x | ✓ | 2339 [O3] | 1610 [O2] |
| | 8 | 231A | HIS | 2.45 | 3.09 | 125.47 | x | ✓ | 2341 [O2] | 1839 [Nar] |
| 13 | 1 | 50A | ARG | 2.64 | 3.09 | 108.46 | ✓ | ✓ | 205 [Ng+] | 2336 [O3] |
| | 2 | 112A | ASN | 2.39 | 3.03 | 122.29 | x | ✓ | 2319[Nam] | 700 [O2] |
| 14 | 1 | 25A | ARG | 2.55 | 2.96 | 104.68 | ✓ | ✓ | 205 [Ng ⁺] | 2342 [O3] |
| | 2 | 189A | GLN | 3.21 | 3.84 | 123.88 | ✓ | ✓ | 1525[Nam] | 2320 [O3] |
| | 3 | 191A | SER | 2.39 | 3.03 | 123.35 | ✓ | ✓ | 1542 [O3] | 2343 [O3] |
| | 4 | 193A | ASP | 2.90 | 3.86 | 164.35 | ✓ | X | 1547[Nam] | 2320 [O3] |

Table 5. π -Stacking

| S No | Index | Residue | AA | Distance | Angle | Offset | Stacking Type | Ligands Atoms |
|------|-------|---------|-----|----------|-------|--------|---------------|------------------------------------|
| 13 | 1 | 109A | TRP | 4.14 | 19.80 | 1.12 | P | 2324, 2325, 2326, 2327, 2328, 2329 |
| 14 | 1 | 84A | TRP | 4.15 | 14.30 | 1.47 | P | 2319, 2320, 2321, 2322, 2323, 2324 |

Table 6. π -Cation Interactions

| S No | Index | Residue | AA | Distance | Offset | Protein Charged? | Ligand Group | Ligand Atoms |
|------|-------|---------|-----|----------|--------|------------------|--------------|------------------------------------|
| 12 | 1 | 25A | ARG | 4.50 | 0.75 | ✓ | Aromatic | 2320, 2321, 2322, 2324, 2325, 2331 |

Table 7. Salt Bridges

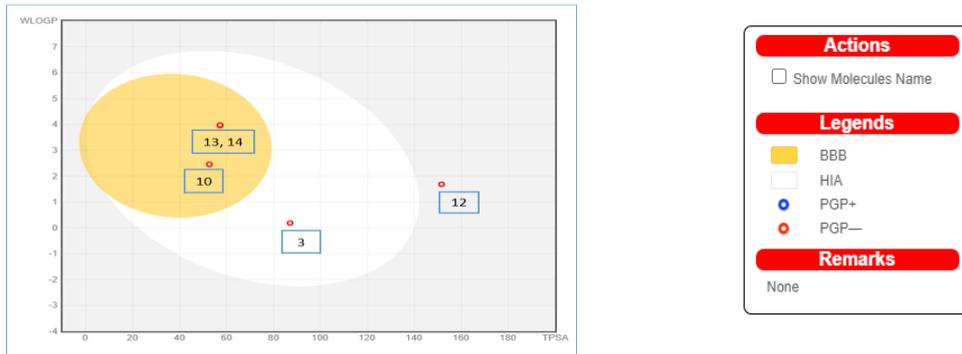
| S No | Index | Residue | AA | Distance | Protein Positive? | Protein Charged? Ligand Group | Ligand Atoms |
|------|-------|---------|-----|----------|-------------------|-------------------------------|--------------|
| 4 | 1 | 171A | HIS | 4.83 | | Carboxylate | 2325, 2327 |

3.4. Entos-Envision

Compound 12, the best of the five compounds screened with auto dock vina, was analysed using the Entos-Envision web service [9]. With this tool, besides

HOMO, LUMO, and MEP poses, atom-centered charges, bond orders, ESP at nuclei energy and Fukui reactivity properties were also examined in Table 8 [16,17].

Boiled Egg



Bioavailability Radar

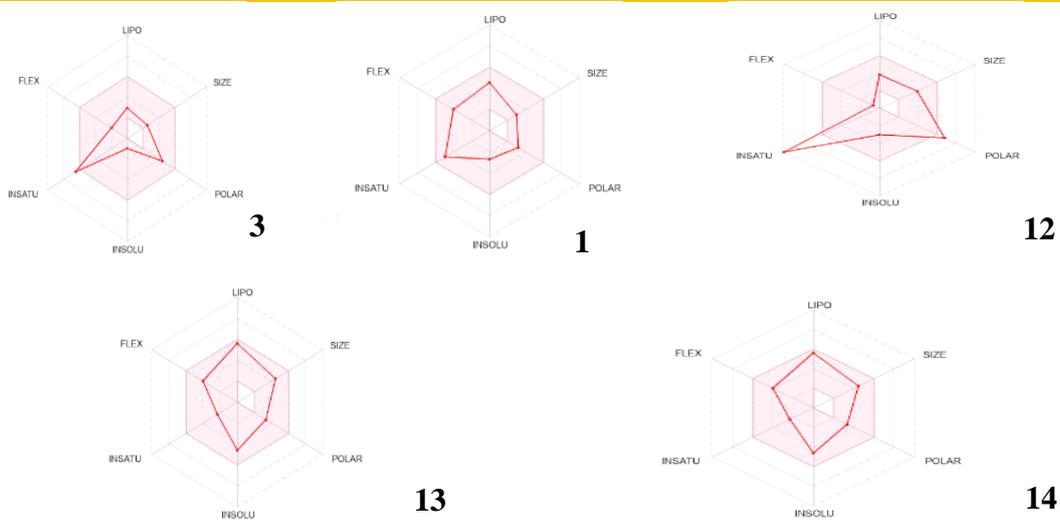
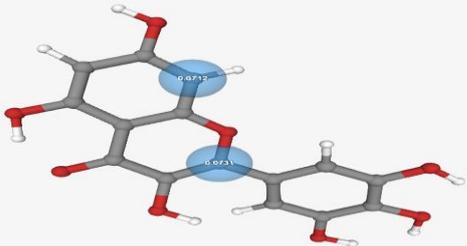
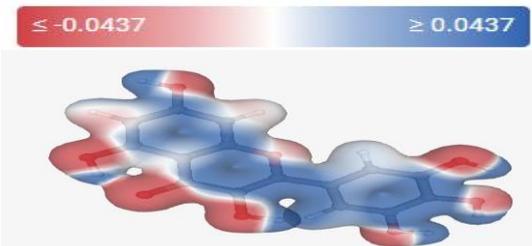
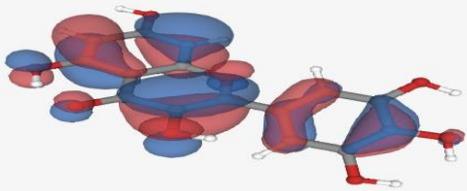
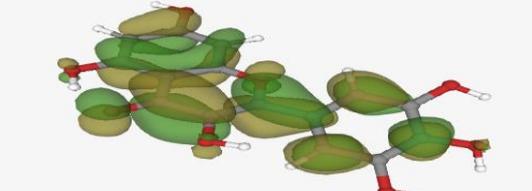


Figure 5. BOILED-Egg and Bioavailability radar plots of the studied ligand

Table 8. Blue: positive colour, red: negative colour.

| | |
|-------------------------------------|---------------------------|
| | |
| <p>Atom centered charges</p> | <p>Bond Orders</p> |
| | |

| | |
|--|--|
| <p style="text-align: center;">ESP at Nuclei</p>  | <p style="text-align: center;">Energy: -196152 kJ/mol</p>  |
| <p style="text-align: center;">Fukui Reactivity</p>  | <p style="text-align: center;">Electrostatic Potential</p>  |
| <p style="text-align: center;">HOMO (-11.1981 eV)</p> | <p style="text-align: center;">LUMO (-7.61227 eV)</p> |

4. Conclusions

Molecular modelling is a multidisciplinary tool that combines theoretical and computational sciences to simulate macromolecule-ligand interactions [18]. Analysis of Autodock Vina and Igemdock results revealed that compounds 12 and 13 were the most effective. Compound 12 demonstrated favourable pharmacokinetic properties for lipid-based absorption due to its solubility in fats and lipids, appropriate molecular size, and acceptable overall solubility. However, its inability to cross the blood-brain barrier and evade cellular efflux by membrane proteins poses limitations for central nervous system-targeted applications. Structural modifications to improve carbon saturation and polarity could address these shortcomings, thereby expanding its therapeutic applications. The higher number of hydrogen bonds in compound 12 supports its characterisation as the most stable structure. Additionally, its HOMO-LUMO energies and other physical properties were calculated using Entos Envision, a platform for rapid semi-empirical molecular simulations. These findings highlight the potential of compound 12 as a promising candidate for therapeutic applications, with further structural optimisations required to overcome its limitations and enhance its pharmacokinetic and pharmacodynamic properties.

References

- [1] <https://www.who.int/news-room/factsheets/detail/breast-cancer>
- [2] National Center for Biotechnology Information. PubChem Compound Summary for CID 1183, Vanillin. Retrieved March 29, (2024) from

<https://pubchem.ncbi.nlm.nih.gov/compound/Vanillin>.

- [3] M. Yousuf, A. Shamsi, A. Queen, M. Shahbaaz, P. Khan, Hussain, A., M. Imtaiyaz Hassan, Targeting cyclin-dependent kinase six by vanillin inhibits breast and lung cancer cell proliferation: Combined computational and biochemical studies, *Journal of Cellular Biochemistry*, 122(8), (2021) 897-910.
- [4] Z. M. M. Alzhrani, M. M. Alam, T. Neamatallah, & S. Nazreen, Design, synthesis and in vitro antiproliferative activity of new thiazolidinedione-1, 3, 4-oxadiazole hybrids as thymidylate synthase inhibitors. *Journal of Enzyme Inhibition and Medicinal Chemistry*, 35(1), (2020) 1116-1123.
- [5] A. M. Malebari, S. Nazreen, S. E. I. Elbehairi, A. A. Elhenawy, A. A. Abdelhamid, A. Alfarsi, M. M. Alam, Synthesis, Characterization, Anticancer, Pharmacokinetic, and Docking Studies of Vanillin-Benzimidazole Derivatives as Aromatase Inhibitors. *Polycyclic Aromatic Compounds*, 44(10), (2024) 6845-6861.
- [6] ZINC 15–ligand discovery for everyone, *Journal of Chemical Information and Modeling*, 55(11), (2020) 2324-2337.
- [7] K. C. Hsu, Y. F. Chen, S. R. Lin, J. M. Yang, iGEMDOCK: a graphical environment enhancing IGEMDOCK using pharmacological interactions and post-

Özlem İŞCAN

- screening analysis, BMC Bioinformatics, 12, (2011) 1-11.
- [8] M. M. Alam, A. M. Malebari, N. Syed, T. Neamatallah, A. S. Almalki, A. A., Elhenawy, Alsharif, M. A. Design, synthesis and molecular docking studies of thymol based 1, 2, 3-triazole hybrids as thymidylate synthase inhibitors and apoptosis inducers against breast cancer cells. Bioorganic & medicinal chemistry, 38, (2021) 116136.
- [9] <https://www.entos.ai/envision> (Accessed : 2022-07-07).
- [10] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera A visualisation system for exploratory research and analysis, J. Comput. Chem., 25 (2004) 1605–1612.
- [11] O. Trott, A.J. Olson, AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimisation, and multithreading, J. Comput. Chem., 31 (2010). 455–461.
- [12] A., Daina, O., Michielin, V. Zoete, SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Scientific reports, 7(1), (2017) 42717.
- [13] M. F. Adasme, K. L. Linnemann, S. N. Bolz, F. Kaiser, S. Salentin, V. J. Haupt, M. Schroeder, PLIP Expanding the scope of the protein-ligand interaction profiler to DNA and RNA. Nucleic acids research, 49(W1), (2021) W530-W534.
- [14] C.A., Lipinski, Lead-and drug-like-compounds: the rule-of-five revolution. DrugDiscovery Today: Technologies 1(4), (2004) 337-341.
- [15] BIOVIA Discovery Studio Visualizer, version 21.1.0.20298. Dassault Systèmes, San Diego, CA (2021).
- [16] J. Lu, I. Paci, & D. C. Leitch, A broadly applicable quantitative relative reactivity model for nucleophilic aromatic substitution (SN Ar) using simple descriptors. Chemical Science, 13(43), . (2022) 12681-12695.
- [17] J. Lu, I. Paci, & D. Leitch, (2022). A generally applicable quantitative reactivity model for nucleophilic aromatic substitution built from simple descriptors
- [18] M. Saeb, M. F. Mahdi, & F. A. Al-saady, (n.d.). In silico Molecular Docking, Molecular Dynamic Simulation and ADME Study of New (2-Methyl Benzimidazole-1-yl)-N- Derivatives with Potential Antiproliferative Activity. Turkish Computational and Theoretical Chemistry, 9(1), (2024) 115-128. <https://doi.org/10.33435/tcandtc.1442921>.