

Identification of Novel Scaffolds with Dual Role as Antiepileptic and Anti-Breast Cancer

Shailima Rampogu^{1#}, Ayoung Baek^{1#}, Rohit Bavi¹, Minky Son¹, Guang Ping Cao¹, Raj Kumar¹, Chanin Park¹, Amir Zeb¹, Rabia Mukhtar Rana¹, Seok Ju Park^{2*} and Keun Woo Lee^{1*}

¹Division of Applied Life Science (BK21 Plus Program), Systems and Synthetic Agrobiotech Center (SSAC), Plant Molecular Biology and Biotechnology Research Center (PMBBRC), Research Institute of Natural Science (RINS), Gyeongsang National University (GNU), 501 Jinju-daero, Jinju, 52828, Republic of Korea.

²Department of Internal Medicine, College of Medicine, Busan Paik Hospital, Inje University, Republic of Korea

#These authors have contributed equally to this work

*Corresponding authors:

psjgenesis@hanmail.net

kwlee@gnu.ac.kr

Abstract

Aromatase inhibitors with an IC_{50} value ranging from 1.4 to 49.7 μ M are known to act as antiepileptic drugs besides being potential breast cancer inhibitors. The aim of the present study is to identify novel antiepileptic aromatase inhibitors with higher activity exploiting the ligand-based pharmacophore approach utilizing the experimentally known inhibitors. The resultant Hypo1 consists of four features and was further validated by using three different strategies. Hypo1 was allowed to screen different databases to identify lead molecules and were further subjected to Lipinski's Rule of Five and ADMET to establish their drug-like properties. Consequently, the obtained 68-screened molecules were subjected to molecular docking by GOLD. Furthermore, the compounds with the highest dock scores were assessed for molecular interaction. Later, the MD simulation was applied to evaluate the protein backbone stabilities and binding energies adapting GROMACS 5.0.6 and MM/PBSA which was followed by the density functional theory (DFT), to analyze their orbital energies and further the energy gap between them. Eventually, the number of Hit molecules was culled to three projecting Hit1, Hit2, and Hit3 as the potential lead compounds based on their highest dock scores, hydrogen bond interaction, lowest energy gap and the least binding energies and stable MD results.

Keywords: Aromatase inhibitors; epilepsy; antiepileptic drugs; molecular docking; MD simulations; density functional theory.

1 Introduction

Aromatase Inhibitors (AIs) are the drug molecules which have gained a wider recognition for their role in the treatment of breast cancer [1]. However, these inhibitors are known to be ineffectual in post-menopausal women. Specifically, the aromatase inhibitors exert their action by hindering the enzyme aromatase that is considered vital in catalyzing the final rate-limiting step during the estrogen synthesis [2]. Heretofore, three generations of aromatase inhibitors are available with aminoglutethimide belonging to the first generation. Nevertheless, due to its poor selectivity between the CYP450s and increased toxic effects, it was later withdrawn from being marketed. However, it serves as a prototype for the discovery of other AIs. The second-generation drugs, fadrozole, with an imidazole group, are relatively highly selective and more potential upon comparison with the first generation AIs, though continues to be effective on other CYP450 enzymes and apparently lack specificity. Meanwhile, the third generation drugs were developed which demonstrated a higher therapeutic index and reduced toxicity in both the early and the advanced breast cancer cases [3] [4] [5] and were hence, successful in gaining approval from the Food and Drug Administration (FDA). The drug molecules anastrozole and letrozole that belong to the later generation act as competitive inhibitors for androstenediene [6] [7] [8]. Aromatase enzymes are involved in the synthesis of estrogen by aromatizing the ring A of androgens [9] by a three step mechanism aided by one mole of NADPH and one mole of O_2 , respectively [10] [11]. Comprising of 503 amino acid residues and a haem group, it has an androgen specific cleft at its binding site [12] and is considered to be one of the most potential drug targets for breast cancer incidences and also in the hormonal therapy.

Additionally, it was also noted that AIs have a role to be played as an antiepileptic drug (AEDs). Currently known AEDs are attributed to several negative effects such as depression, birth defects and cognitive impairment [13] [14] [15] [16]. This warrants the necessity for the development and designing of more effective and less toxic AEDs and further focuses the research towards the development and designing of novel drug molecules.

Several reports exist which emphasize on the use of AIs in treating epilepsy [13] [14] [15] [16] [17] [18] [19]. In fact, in the epileptic males, the levels of estradiol were observed to be reasonably higher which also reinforces the clinical studies [13] [20]. Moreover, AIs were employed to treat men with

sexual and reproductive dysfunctions [21] [22] [23]. Nevertheless, the seizure recurrence remains common between both the genders. Besides these, estrogen has an ability to enhance the neural membrane excitability, seizure discharges and can also reduce the seizure threshold, pronouncing oestrogens as epileptogenic [24] [25]. Indeed, considering everything, the relationship between oestrogens and epilepsy can be vindicated and the screening of efficient AIs could simultaneously lead to the identification of novel AEDs. Currently the AIs that exhibit an IC_{50} values ranging between 1.4 to $49.7\mu M$ are used as antiepileptic drugs. Therefore, the objective of the present study is to perform ligand-based pharmacophore search to identify potential lead molecules against aromatase that could be beneficial in treating both the breast cancer and the epilepsy.

2 Materials and methods

In order to retrieve the potential lead candidates, the ligand based pharmacophore approach was adapted and the methodology is depicted in figure 1.

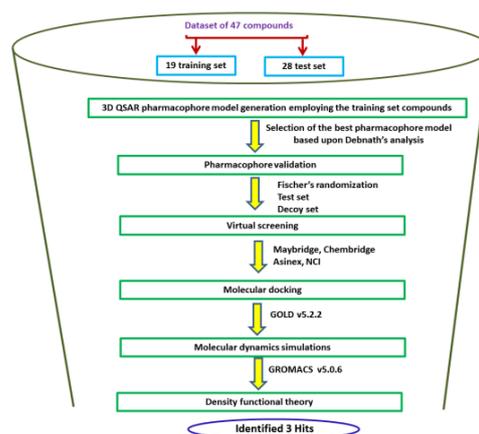


Figure 1. Pictorial representation of the overall methodology employed.

2.1 Preparation of the dataset

Selection of the compounds for the subsequent generation of the pharmacophore is the most important avenue in drug designing. For the current investigation, 47 compounds were preferred which exhibited a wide range of inhibitory activities reported from the literature [26] [27] [28] [29] [30] [31] [32] and were grouped as the dataset. The selection of these 47 compounds was performed in a systematic manner. The compounds were initially checked for the duplicates and were subsequently deleted in order to avoid the repetition of the

compound and further to produce accurate results. Furthermore, based upon the range of IC_{50} values they exhibited and the structural diversity, they were further grouped in the training set and test set [33]. Additionally, the training set compounds should represent the broad class of compounds by demonstrating structural diversity and an activity range of four order magnitudes. For the generation of the 3D QSAR pharmacophore, a minimum of 16 compounds were to be present in the training set. Additionally, the most active compounds have to be included into the training set. Accordingly, 47 compounds were thus, divided into the training set and the test set compounds. The test set was chosen employing the same protocol as the training set. Test set consisted of 28 compounds that exhibited a range of activity values [33] and structurally diverse from the training set. Training set constituted 19 assorted compounds with inhibitory activities (IC_{50}) ranging from 0.5 nmol/L to 18,000 nmol/L, figure 2. The test set comprised of the remaining 28 compounds that were structurally distinct from the training set. The training set compounds and the test set compounds were further classified based upon their IC_{50} values. The compounds that displayed an inhibitory activities less than or equal to 100 nmol/L were labeled as most active compounds (+++), compounds displaying an activity range between 100 nmol/L and 6,500 nmol/L were grouped into moderately active compounds (++) and the compounds with the IC_{50} value above 6,500 nmol/L were labelled as least active compounds (+). The rationale behind this is that the training set was employed to generate a pharmacophore model, while the test set was enrolled to further validate it. The 2D structures were illustrated by means of ChemSketch (<http://www.acdlabs.com/resources/freeware/chemsketch/>) [34] and were eventually converted to their 3D structures on the Discovery Studio (DS) v4.5.

2.2 Generation of the pharmacophore model

The 3D QSAR pharmacophore was generated based on the activity values of the training set compounds utilizing the *Catalyst HypoGen* algorithm [35] available on the Discovery Studio v4.5 (DS, www.accelrys.com, USA) that could evidently generate conformations of lower energy rendered by the *Best* algorithm, provided, the energy threshold value and the uncertainty value were fixed to 20 kcal/mol [33] and 3, respectively, retaining the other parameters as default. Adapting the *Feature Mapping* protocol present on the DS that critically probes into the crucial chemical features that exists within the training set, was

exploited to develop the hypothesis with maximum features of five and a minimum of zero. However, relies on the hydrogen bond donors (HBD), hydrogen bond acceptor (HBA), hydrophobic aliphatic (Hy-Ali), hydrophobic regions (HyP) and ring aromatic (RA) which were harvested in the generation of 10 qualitative hypothesis along with their statistical parameters. The best and the efficient pharmacophore was analyzed according to the Debnath's method [36].

2.3 Validation of the hypothesis

Statistically unique pharmacophore model should precisely predict the activity of the molecules with an aptness of retrieving the active compounds from the databases. The best pharmacophore model was evaluated and checked employing the Fischer's randomization, test set and decoy set methods, respectively. The fundamental idea behind validating the pharmacophore model was to ascertain its ability to discriminate the active compounds coupled with its predictive ability towards the corresponding molecules. Moreover, this method nullifies the chance correlation while the generation of the pharmacophore model and further secures that the model was not generated randomly. In Fischer's randomization method, 95% confidence was selected and was run alongside the hypothesis generating 19 random spreadsheets. The test set and the decoy set validations were employed to assess if the generated pharmacophore model was able to choose the molecules based upon their activities corresponding to the training set active molecules, which distinguishes potential inhibitors from the other compounds. The test set comprises of about 28 structurally variant chemical compounds to ensure the predictability of the pharmacophore. The decoy set was formulated by calculating the 1D property of 8 active compounds among 782 unknown inactive compounds and subsequently, their Enrichment Factor (EF) and the Goodness of Hit (GH) values were calculated by applying the formulae

$$EF = \left(\frac{Ha}{Ht}\right) \div \left(\frac{A}{D}\right)$$

$$GH = \left[\left(\frac{Ha}{4HtA}\right)\right] (3A + Ht) \times [1 - (Ht - Ha) \div (D - A)]$$

The generated model was considered good, if it's GH value ranges between 0.6-0.8 [37] [38]. Here, Ha refers to the total number of active compounds, Ht indicates the retrieval number of Hits from the database, A denotes the total number of active compounds in the database, and D represents the

total number of molecules in the database. The *Ligand Pharmacophore Mapping* module available on the DS, preferring the *Best* algorithm and the *Flexible fit* options were employed.

2.4 Virtual screening of the databases for retrieving the lead candidates

Pharmacophore-based database screening was pursued to determine the novel lead compounds that could inhibit aromatase. It is mightily essential to use the pharmacophore model that imbibes the chemical and functional qualities required by the potential prospective drugs and hence can be deployed for screening novel drugs. Four chemical databases were thus screened, taking the validated pharmacophore model as a 3D query. Herein, the databases screened were NCI, Asinex, Chembridge, and Maybridge. The *Ligand Pharmacophore Mapping* protocol implemented on the DS was employed for screening the databases electing the *Fast Flexible* search method. The retrieved molecules from the databases should be probed for obeying the requirements of the chemical moieties and further, should equate with the features present within the 3D query molecule.

2.5 Drug-likeness prediction

In order to ensure the retrieved compounds possess good pharmacokinetics properties, all the Hit compounds were subjected to ADMET and Rule of Five (RoF) developed by Lipinski. The predominant motive behind the ADMET is to access if the compounds were able to pass across the Blood – Brain – Barrier (BBB), exhibits low toxicity, good solubility and good human intestinal absorption. Moreover, it has to be noted that the bioavailability, ability to penetrate through the BBB and low hepatotoxicity are pivotal filters while addressing the drugs in connection with the central nervous system. Additionally, the Lipinski's Rule of Five [39] quantifies a drug to be well absorbed when it has a molecular weight ≤ 500 , number of hydrogen bond acceptors ≤ 10 , hydrogen bond donors ≤ 5 and $\log P \leq 5$. All those compounds, which satisfied these properties, were subjected to molecular docking mechanism.

2.6 Molecular docking mechanism

Ligand- based molecular docking has emerged as the most promising tool in the field of drug discovery and designing. This method allows the ligand molecule to bind in the active site space available within the protein. Thereon, the protein ligand interactions were evaluated by different scoring functions. Generally, docking delineates on

all the information of the protein and eventually estimates their binding affinities. For the present investigation, the docking program Genetic Optimization for Ligand Docking (GOLD) [40] [41] was adapted which operates on genetic algorithm. For the target structure, the aromatase with PDB ID: 3EQM complexed with 4-anderostene-3-17-dione was downloaded from the protein data bank (www.rcbs.org). The protein was prepared by removing all the water molecules and the addition of hydrogen atoms. Further, the protein was subjected to "minimization" with CHARMm force field after performing the "clean protein" accessible on the DS. The training set molecules along with the 68 screened compounds were also prepared for docking by subjecting them to minimization. The protein binding sites were determined across all the atoms that fall within the radius 15Å around the 4-anderostene-3-17-dione. For expounding the binding affinities between the protein and the ligand, the GoldScore was used, while the ChemScore was recruited for rescoring. For each ligand, 50 poses were allowed to be generated and were selected based upon the highest GoldScore. Withal, the generated docked poses were also examined for their molecular interactions and the hydrogen bonds between the ligand and the binding sites to secure the potential Hits.

2.7 Molecular dynamics simulations

MD simulations were performed employing GROMACS 5.0.6 using CHARMm27 force field, [42] [43] for the reference and the Hit compounds obtained from the docking studies. The ligand topologies were generated recruiting SwissParam [44] and the system was solvated with TIP3P water model in a dodecahedron box as described earlier [37][45]. Further, the counter ions were added to neutralize the systems. The bad contacts were removed by energy minimization using steepest descent algorithm and subsequently equilibration processes were conducted employing NVT (Berendsen thermostat) and the NPT during which the Parrinello-Rahman barostat was utilized to maintain the pressure [46] [47]. During the equilibration process, the solvent molecules were allowed to move while the protein backbone was restrained. Furthermore, to constrain the bond length and to maintain the geometry of the molecule, LINCS algorithm [48] was employed while SETTLE algorithm [49] was applied to constrain the geometry of the water molecules [49]. Particle Mesh Ewald (PME) [50] was recruited to calculate the long range electrostatic interactions, adjusting the cut-off distance of 9 Å and 10 Å for Coulombic and van der Waals interactions,

respectively. Additionally, the periodic boundary conditions were implemented to avoid edge effects. The time step of 2 fs was used throughout the simulation and coordinate data was stored at every picosecond (ps). All the analysis of MD simulations was carried out by VMD [51] and DS software.

2.8 Binding free energy calculations

Molecular Mechanics/Poisson–Boltzmann Surface Area (MM/ PBSA) methodology has been employed for calculating the binding free energy (ΔG_{bind}) of a given system retrieving 30 snapshots. For the current study, the methodology for calculating the MM/PBSA was computed as delineated earlier [52] [53] [54].

The binding free energy for protein ligand complex in the solvent was represented as

$$\Delta G_{\text{binding}} = G_{\text{complex}} - (G_{\text{protein}} + G_{\text{ligand}})$$

Herein, G_{complex} refers to the total free energy of the complex and G_{protein} and G_{ligand} indicates the separate protein and ligand in the solvent. Their free energies can be computed by

$$G_X = E_{\text{MM}} + G_{\text{solvation}}$$

Where, X can be a protein, ligand or its complex. E_{MM} represents the average molecular mechanics potential energy in vacuum, while the $G_{\text{solvation}}$ interprets the free energy present in the solvation.

Additionally, molecular mechanics potential energy in vacuum can be evaluated by adapting the equation

$$E_{\text{MM}} = E_{\text{bonded}} + E_{\text{non-bonded}} = E_{\text{bonded}} + (E_{\text{vdw}} + E_{\text{elec}})$$

E_{bonded} represents the bonded interactions, while the non-bonded interactions were denoted by $E_{\text{non-bonded}}$. ΔE_{bonded} is generally regarded as zero [55].

The solvation free energy ($G_{\text{solvation}}$) was expressed by the sum of electrostatic solvation free energy (G_{polar}) and apolar solvation free energy ($G_{\text{non-polar}}$) and is given as following

$$G_{\text{solvation}} = G_{\text{polar}} + G_{\text{non-polar}}$$

G_{polar} is computed recruiting the Poisson-Boltzmann (PB) equation [56] while $G_{\text{non-polar}}$ is computed from the solvent-accessible surface area (SASA) and can be written as below

$$G_{\text{non-polar}} = \gamma \text{SASA} + b$$

Here, the γ is the coefficient of the surface tension of the solvent, whereas, b is its fitting parameter, whose values are $0.02267 \text{ kJ/mol/\AA}^2$ or 0.0054

kcal/mol/\AA^2 and 3.849 kJ/mol or 0.916 kcal/mol , respectively.

2.9 Density functional theory

MD optimized systems were forwarded to density functional theory (DFT) studies. DFT by far is the most logical and highly productive methods for calculating the orbital energies[56] in terms of Highest Occupied Molecular Orbital (HOMO) and Lowest Unoccupied Molecular Orbital (LUMO). Additionally, it was also noted that the ionization potential (electron donor) was governed by the HOMO, while the electron affinity (electron acceptor) by LUMO. The preferred MD optimized structures along with the known inhibitors were subjected to DFT that was utilized to explore the energy transfer and stability of small molecules in the binding site. A lower energy gap between the Hit molecules demonstrates that the molecules are highly reactive while the high-energy gap implies low reactivity. The Dmol3 and B3LYP, with DNP basis set with SCF density convergence of $1.0\text{e-}6$ that was available on the DS were employed for computing the energy values.

3 Results

Generation of the ligand-based pharmacophore model is credited with one of the most beneficial techniques in the development of novel drug molecules. Conventionally, the ligand-based pharmacophore model development works by considering the common features that exist in their 3D structures, provided, they are known to bind with their respective target macromolecule. However, this is feasible only when the molecules are enumerated, and most importantly, this approach discloses the common features that exist within the diverse ligands that could be recruited as a 3D query for the establishment of new Hit compounds.

3.1 Generation of HypoGen based pharmacophore model

HypoGen algorithm, accessible on the DS was employed for the generation of the pharmacophore model, which ably analogizes with the chemical structures of the aromatase inhibitors to their corresponding biological activities. The training set composed of 19 compounds, figure 2, were utilized for the pharmacophore generation, which comprised of varied activity ranges between 0.5 nmol/L - 18000 nmol/L . The 3D QSAR pharmacophore generation protocol available on the DS was selected with the preferred features being HBA, HyP, RA, HY-Ali and HBD. Selecting these features, the pharmacophore run was initiated

and towards the end of the run, 10 pharmacophore hypotheses were generated. Amongst which, the pharmacophore model Hypo1 was chosen as the best hypothesis based on Debnath's analysis, such as the high cost difference 54.32, good correlation coefficient of 0.965, least RMSD 0.7 and a good fit value of 8.4 with four features namely 2HBA, HyP and RA, figure 3A. Herein, table1, the RMSD value refers to the variation that exists between the predicted activity values from that of the experimental value. The cost difference depicts the difference between the null and total cost of hypothesis. Besides these, the correlation coefficient relies on the linear regression, the greater is the predictive capacity, when lower the RMSD, less is the divergence between the predictive activity from that of the experimental activity. Accordingly, Hypo1 was preferred to evaluate the inhibitory activities of the training set compound. The most active compound $IC_{50} = 0.5$ nmol/L and the least active compound $IC_{50} = 18000$ nmol/L of the training set have aligned with four and three features of the Hypo1 correspondingly, figure 3B and 3C.

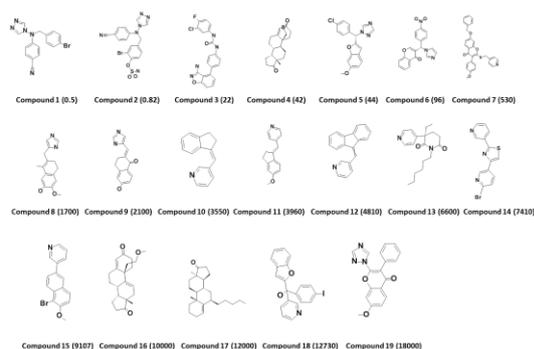


Figure 2. 2D Structures of 19 training set compounds with their corresponding IC_{50} values (nmol/L) in parenthesis.

Table: 1. Statistical and predictive significance presented in cost values for top 10 hypotheses generated due to 3D QSAR pharmacophore modelling.

| Hypo no | Total Cost | Cost difference ¹ | RM SD | Correlation | Features ² | Max Fit |
|---------|------------|------------------------------|-------|-------------|-----------------------|---------|
| Hypo 1 | 89.1 | 54.32 | 0.7 | 0.96 | 2HBA, HyP, RA | 8.4 |
| Hypo 2 | 90.4 | 53.09 | 0.8 | 0.96 | 2HBA,HyP,RA, Hy-Ali | 9.1 |
| Hypo 3 | 91.6 | 51.87 | 0.9 | 0.94 | 2HBA,Hy-Ali, RA | 7.4 |
| Hypo 4 | 91.6 | 51.83 | 0.8 | 0.95 | 2HBA,HyP, RA | 9.0 |
| Hypo 5 | 92.0 | 51.47 | 0.8 | 0.95 | 2HBA,Hy-Ali, RA | 9.5 |

| | | | | | | |
|---------|------|-------|-----|------|---------------------|-----|
| Hypo 6 | 92.3 | 51.21 | 0.9 | 0.94 | 2HBA,Hy-Ali, RA | 8.7 |
| Hypo 7 | 92.7 | 50.73 | 0.9 | 0.94 | 2HBA,Hy-Ali, RA | 8.5 |
| Hypo 8 | 93.1 | 50.33 | 0.9 | 0.94 | 2HBA,HyP, RA | 9.1 |
| Hypo 9 | 93.3 | 50.21 | 1.0 | 0.93 | 2HBA,HyP, RA | 8.4 |
| Hypo 10 | 93.3 | 50.12 | 0.9 | 0.93 | 2HBA,HyP,RA, Hy-Ali | 8.9 |

¹ Cost difference between the null and the total cost. The null cost, the fixed cost and the configuration cost were found to be 143.52, 83.73 and 18.68, respectively.

² HB- Hydrogen Bond, HyP- Hydrophobic, RA- Ring Aromatic, Hy-Ali-Hydrophobic Aliphatic.

Furthermore, Hypo1 could successfully predict the inhibitory activities of the compounds present in the training set, table 2. Based upon their activity values, the training set compounds were classified. The compounds that displayed an inhibitory activities less than or equal to 100 nmol/L were labeled as most active compounds (+++), compounds displaying an activity range between 100 nmol/L and 6,500 nmol/L were grouped into moderately active compounds (++) and the compounds with the IC_{50} above 6,500 nmol/L were labelled as least active compounds (+). However, it was noticed that the two least active compounds were over estimated as moderately active and two moderately active compounds were underestimated as least active compounds. Nevertheless, the inhibitory values of all the active compounds were envisaged in the same order of magnitude. Thus, Hypo1 successfully estimated the activities of the compounds present in the training set. Thereafter, to assess the robustness of the pharmacophore model, it was further subjected to multiple methods of validations such as the test set method, Fischer's randomization and decoy set method.

Table 2. Experimental and predicted activity of the 19 training set molecules based on the pharmacophore model Hypo1

| No | Fit Val | Exp IC_{50} nmol/L | Pre IC_{50} nmol/L | Error ^a | Exp Scale ^e | Pred Scale ^b |
|----|---------|----------------------|----------------------|--------------------|------------------------|-------------------------|
| | | | | | | |

| | | | | | | |
|----|------|-------|-------|------|-----|-----|
| 1 | 8.37 | 0.5 | 0.36 | -1.4 | +++ | +++ |
| 2 | 7.7 | 0.82 | 1.7 | 2.1 | +++ | +++ |
| 3 | 6 | 22 | 86 | 3.9 | +++ | +++ |
| 4 | 5.68 | 42 | 180 | 4.3 | +++ | ++ |
| 5 | 6.33 | 44 | 40 | -1.1 | +++ | +++ |
| 6 | 6.26 | 96 | 47 | -2 | +++ | +++ |
| 7 | 5.7 | 530 | 170 | -3.1 | ++ | ++ |
| 8 | 5.12 | 1700 | 650 | -2.6 | ++ | ++ |
| 9 | 4.23 | 2100 | 5100 | 2.4 | ++ | ++ |
| 10 | 4.01 | 3500 | 8500 | 2.4 | ++ | + |
| 11 | 3.99 | 4000 | 8700 | 2.2 | ++ | + |
| 12 | 4.19 | 4800 | 5500 | 1.1 | ++ | ++ |
| 13 | 4.2 | 6600 | 5400 | -1.2 | ++ | ++ |
| 14 | 4.24 | 7400 | 4900 | -1.5 | ++ | ++ |
| 15 | 4.09 | 9100 | 6900 | -1.3 | + | ++ |
| 16 | 4.24 | 10000 | 4900 | -2 | + | ++ |
| 17 | 4.24 | 12000 | 4900 | -2.5 | + | ++ |
| 18 | 3.53 | 13000 | 25000 | 2 | + | + |
| 19 | 4.2 | 18000 | 15400 | -3.3 | + | + |

^aError, ratio of the predicted activity (Pred IC₅₀) to the experimental activity (Exp IC₅₀) or its negative inverse if the ratio is <1

^bIC₅₀ values ≤ 100 nmol/L are most active (+++), IC₅₀ values between 100 nmol/L ~ 6,500 nmol/L are moderately active (++) and IC₅₀ values > 6,500 nmol/L are least active compounds (+).

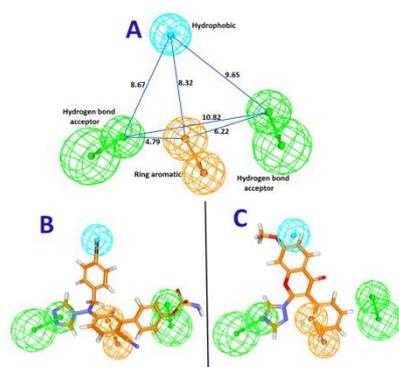


Figure 3. Selected pharmacophore, Hypo 1 with its features. Figure 3A demonstrates the best pharmacophore model with its corresponding

geometry with four features, Green-Hydrogen bond acceptor, HBA (2); Cyan- Hydrophobic, HP (1); Orange- Aromatic Ring, RA (1). Figure 3B represents the mapping of best pharmacophore model Hypo1 to the most active compound (IC₅₀ 0.5nmol/L) in the training set. The compound was seen to align with four features 2HBA, 1HP, and 1RA. Figure 3C refers to the most inactive compound (IC₅₀-18,000 nmol/L) from the training set. The compound was found to align with only three features, 1HBA, 1HP and 1RA.

3.2 Validation of Hypo1

The significance of each of the generated pharmacophore model, Hypo1, was estimated by test set, Fischer's randomization and decoy set method.

3.2.1 Test set method

28 structurally distinct compounds from the dataset, apart from the training set compounds were grouped into test set. The fundamental necessity of the test set is to probe into the ability of the Hypo1 in assessing the external compounds with the same activity range. It was however noticed that, one active compound and one moderately active compound were underestimated as moderately active and least active compounds, respectively table 3 and figure 4A, repeating the protocol used for training set compounds. The correlation coefficient (r) was calculated to be 0.94 and was categorized on the similar basis as that of the training set. The test set validation is an indicative of the fact that Hypo1 could judge the active compounds from the moderately active and the inactive compounds.

Table 3. Experimental and predicted IC₅₀ values of 28 test set molecules using Hypo1

| No | Fit Val | Exp IC ₅₀ nmol/L | Pred IC ₅₀ nmol/L | Error ^a | Exp Scale | Pred Scale ^b |
|----|---------|-----------------------------|------------------------------|--------------------|-----------|-------------------------|
| 1 | 7.41 | 2.3 | 8.5 | 3.7 | +++ | +++ |
| 2 | 7.01 | 17 | 21 | 1.3 | +++ | +++ |
| 3 | 7.24 | 20 | 13 | -1.6 | +++ | +++ |
| 4 | 7.01 | 30 | 21 | -1.4 | +++ | +++ |
| 5 | 5.76 | 37 | 380 | 10 | +++ | ++ |
| 6 | 6.83 | 39 | 33 | -1.2 | +++ | +++ |
| 7 | 6.87 | 40 | 30 | -1.3 | +++ | +++ |
| 8 | 6.64 | 44 | 51 | 1.2 | +++ | +++ |
| 9 | 5.25 | 530 | 1200 | 2.3 | ++ | ++ |
| 10 | 5.54 | 550 | 640 | 1.2 | ++ | ++ |
| 11 | 5.49 | 570 | 710 | 1.2 | ++ | ++ |
| 12 | 5.57 | 590 | 590 | 1 | ++ | ++ |
| 13 | 5.58 | 600 | 580 | -1 | ++ | ++ |

| | | | | | | |
|----|------|-------|-------|------|----|----|
| 14 | 5.59 | 600 | 570 | -1.1 | ++ | ++ |
| 15 | 5.55 | 1000 | 620 | -1.6 | ++ | ++ |
| 16 | 5.56 | 1100 | 610 | -1.9 | ++ | ++ |
| 17 | 5.15 | 1200 | 1600 | 1.3 | ++ | ++ |
| 18 | 5.19 | 4700 | 1400 | -3.3 | ++ | ++ |
| 19 | 4.12 | 5100 | 17000 | 3.2 | ++ | + |
| 20 | 4.56 | 5700 | 6000 | 1.1 | ++ | ++ |
| 21 | 4.81 | 6000 | 3400 | -1.8 | ++ | ++ |
| 22 | 4.17 | 6700 | 15000 | 2.2 | + | + |
| 23 | 4.63 | 6800 | 5100 | -1.3 | + | + |
| 24 | 4.81 | 7100 | 3400 | -2.1 | + | ++ |
| 25 | 4.32 | 7800 | 11000 | 1.3 | + | + |
| 26 | 5.38 | 9200 | 930 | -9.9 | + | ++ |
| 27 | 4.33 | 13000 | 10000 | -1.2 | + | + |
| 28 | 4.11 | 18000 | 17000 | -1.1 | + | + |

^aError, ratio of the predicted activity (Pred IC₅₀) to the experimental activity (Exp IC₅₀) or its negative inverse if the ratio is <1

^bIC₅₀ values ≤ 100 nmol/L are most active (+++), IC₅₀ values between 100 nmol/L ~ 6,500 nmol/L are moderately active (++) and IC₅₀ values > 6,500 nmol/L are least active compounds (+).

3.2.2 Fischer's randomization method

In order to statistically validate Hypo1, Fischer's randomization method was performed preferring the confidence level of 95%. Consequently, 19 random spreadsheets were produced and the magnitude of hypothesis was calculated adapting the formula

$$S = [1 - (1 + X) / Y] \times 100$$

Where, 'X' refers to the total hypothesis possessing total cost lower than the actual hypothesis. 'Y' refers to the total number of HypoGen runs (initial+random)

Upon comparing the Hypo1 with the total cost of 19 random, it was understood that Hypo1 was of far finer quality and was not generated by chance, figure 4B and suggests that Hypo1 could be imbibed with all the essential features required to inhibit the aromatase activity.

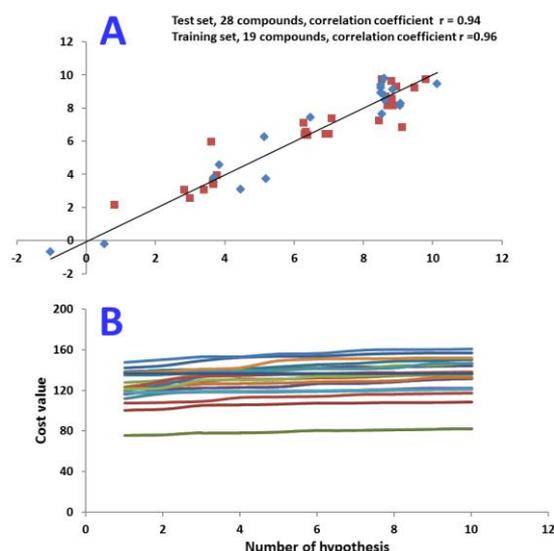


Figure 4. Validation of the pharmacophore model, Hypo1. Figure 4A refers to the correlation plot as predicted by Hypo1 between experimental and the predicted activities of the training set and the test set compounds. Figure 4B evaluates the difference in cost between 19 scrambled runs selecting 95% confidence level. Hypo1 had displayed a lower cost value.

3.2.3 Decoy set validation

As an eventual and conclusive validation, the decoy set method was performed, recruiting the "best flexible" module on the DS. With an aim of assessing the effectiveness of the Hypo1, different parameters were recorded, such as, the enrichment factor (EF) [38], goodness of fit score (GF), false positives and false negatives, respectively. The GH and the EF were calculated taking into account, the Hits (Ht), number of actives percent yield (%Y), false positives, false negatives and the percentage ratio of actives in the retrieved Hit lists (%A). The GF and the EF scores were calculated to be 0.61 and 69.8, respectively, which bespeaks the ability [57] of Hypo1 towards recognizing the true positives, table 4.

Table 4. Decoy set parameters employed to validate the pharmacophore model Hypo1.

| Parameters | Values |
|--|--------|
| Total number of molecules in database (D) | 782 |
| Total number of actives in database (A) | 8 |
| Total number of Hit molecules from the database (Ht) | 7 |

| | |
|---|------|
| Total number of active molecules in Hit list (Ha) | 5 |
| % Yield of active [(Ha/Ht) X 100] | 71.4 |
| % Ratio of actives [(Ha/A) X 100] | 62.5 |
| Enrichment Factor (EF) | 69.8 |
| False negatives (A-Ha) | 3 |
| False Positives (Ht-Ha) | 2 |
| Goodness of fit score (GF) | 0.61 |

3.3 Virtual screening and ADMET studies

Using the validated pharmacophore as a 3D query, screening for the novel scaffolds was performed across different databases such as Chembridge, Maybridge, Asinex and NCI comprising of 50000, 59652, 213262 and 238819 compounds each. Those compounds with a fitness value of > 8 were subjected to Lipinski's Rule of Five and ADMET studies. ADMET studies were conducted to delineate on the pharmacokinetic properties of a drug within the human body, as well as the values correspond to the Blood Brain Barrier penetration (BBB), solubility, hepatotoxicity, human intestinal absorption (HIA), cytochrome P450, 2D6 inhibition and plasma protein binding (PPB). The highest values selected correspondingly for BBB, solubility and absorption were 3, 3, and 0. Those molecules that qualified the ADMET were proceeded further to Lipinski's Rule of Five which ensures a given molecule has less than 5 logP, molecular weight of less than 500, less than 5 hydrogen bond donors and less than 10 hydrogen bond acceptors, with rotatable bonds less than 10. Eventually, the procured 68 filtered molecules, figure 5, were further subjected to the molecular docking studies.

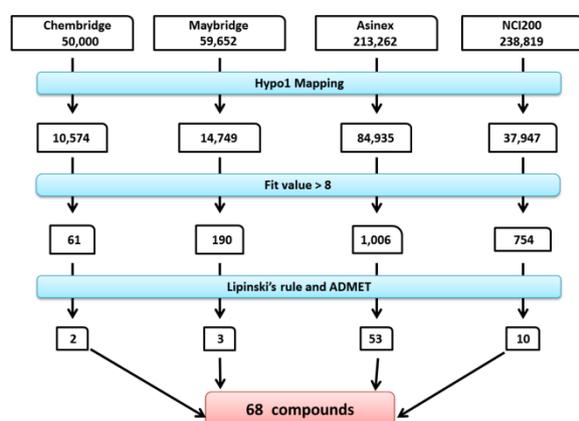


Figure 5. Schematic representation of virtual screening of the databases. Different filters were employed to redeem the most efficient dug-like candidate compounds.

3.4 Molecular docking mechanism

Training set compounds along with 68 screened compounds retrieved after virtual screening were subjected to the molecular docking mechanism using "Genetic Optimization for Ligand Docking" (GOLD). This is one of the potential molecular docking techniques available currently that can perform screening, lead optimization along with an ability to identify precise binding modes of active molecules, and renders high consistency in performance when subjected to a wide range of receptor types. Molecular docking was carried out using the default parameters, however, the active site radius was set to 15Å, around the crystal structure with 50 conformations generated and saved for each ligand. GOLD operates by GoldScore fitness function, while the ChemScore was used as the rescoring function. For the present study, the protein 3EQM was imported onto the DS from Protein Data Bank which had a resolution of 2.0Å. Hereinafter, for analyzing the dock results, the most active compound from the training set was labeled as the reference compound. Hit molecules were evaluated for the hydrogen bond interactions with amino acid residues such as Arg115, Ala306, Asp309, Val370, Leu372, Met374, and Leu477, respectively. Interactions apart from the aforementioned residues were ignored, a process adapted to identify the superlative potential leads.

Following this, the compounds that generated a GoldScore above 59.74 were considered for further evaluation as their scores were noted to be above the reference compound. Furthermore, the reference compound had demonstrated a Chemscore of -13.5. ChemScore was employed as the rescoring function that effectively evaluates the free energy change upon the ligand binding. Accordingly, the lead compounds were chosen based upon the highest GOLD score and lowest Chemscore as compared to the reference molecule, table 5, along with an appropriate ligand conformation that displays interactions with the active site residues. Subsequently, three Hits were selected as they fulfilled the aforementioned criteria and were found to map well with the pharmacophore, figure 6A.

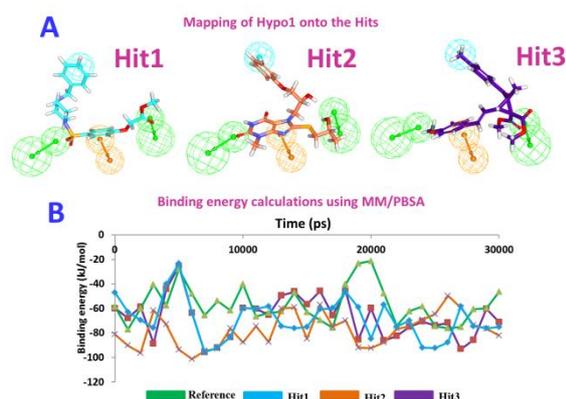


Figure 6A Mapping of Hypo1 onto the Hit compounds (Hit, Hit2, and Hit3). The Hits have mapped with all the features of Hypo1. Figure 6B MM/PBSA binding free energy calculations of the Hits and reference compound. The Hits have rendered lower binding energy than the reference compound.

Table 5. Molecular Docking results, Hit molecules with the highest dock scores than the most active reference compound.

| Systems | Gold fitness score | Chemscore |
|-------------------|--------------------|-----------|
| Protein Reference | 59.74 | -13.5 |
| Protein+Hit1 | 73.37 | -13.3 |
| Protein+Hit2 | 73.20 | -18.6 |
| Protein+Hit3 | 70.84 | -14.6 |

3.5 Molecular Dynamics simulations

MD simulations were initiated for four systems considering the best docked conformations as the initial structures so as to understand their conformational alteration, their binding stability and further their behaviour in the active site and was conducted for 30 ns. The protein stability during the simulations were monitored through the RMSD and potential energy of the corresponding systems. The RMSD values were found to be between 0.12 nm and 0.27 nm throughout the simulation, implying that the systems were converged optimally. The average RMSD were recorded to be 0.2 nm, 0.22 nm, 0.24 and 0.24 nm respectively for the reference, Hit1, Hit2 and Hit3, figure 7A. The potential energy demonstrates similar results portraying the stability of the

systems and further displaying no abnormal nature throughout the simulations, figure 7B.

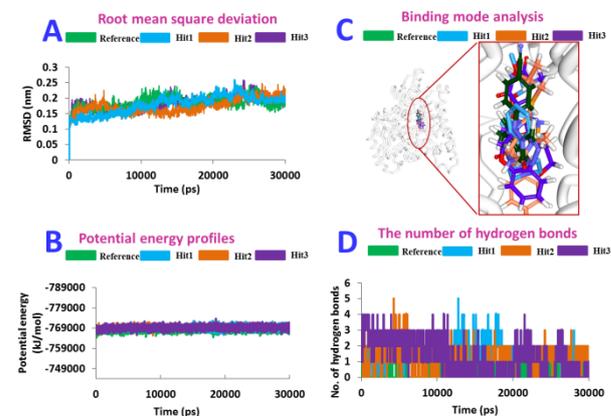


Figure 7. Graphical representation of the molecular dynamics simulation studies conducted during 30 ns. Figure 7A represents the RMSD profiles for the backbone atoms of four systems. Figure 7B demonstrates the potential energies plotted for each system. Both the figures demonstrate that the systems were well converged showing no aberrant behaviour. Figure 7C Binding pattern of reference and the Hits within the binding pocket of 3EQM. The Hits and reference have displayed a similar binding pattern. Figure 7D Number of intermolecular hydrogen bonds between protein and Hits during whole simulation. The Hits have demonstrated a higher number of hydrogen bonds.

The binding mode analysis of the representative structures was assessed from the last 2 ns trajectories. Upon superimposition, it was observed that the binding mode of the Hits was in the similar fashion as with the reference compound, figure 7C. Interrogating the intermolecular interactions revealed that the reference molecule has produced hydrogen bond with the Met374 residue between the NH atom of the residue and N21 atom of the ligand with a bond distance of 2.7 Å. Met374 is the key residue as represented in the crystal structure. Hit1 formed a hydrogen bond interaction with Met374 residue with a bond distance of 2.4 Å. It was noted that the bond existed between the NH atom of the residue and the O21 atom of the ligand. Furthermore, Hit2 interacted with three hydrogen bonds, Thr310, Met374, and Leu477 residues, respectively. The NH atom Met374 joined to the O26 of the ligand with a bond distance of 2.2 Å. The O atom of Lue477 interacted with H41 atom of the ligand with a bond distance of 2.9 Å. Furthermore, the O18 atom of the ligand has joined to HG1 atom of Thr310 residue with a distance of 3.0 Å. On the other hand, Hit3 generated two

hydrogen bonds with Met374 and Thr310, respectively. NH atom of Met374 interacted with the O17 atom of the ligand with a bond length of 1.7Å. O atom of Thr310 joined with the H37 atom of the ligand with a distance of 2.8 Å. Further details of the interactions are tabulated in table 6. These interaction notify that the selected Hits have obeyed similar binding modes and further generated an acceptable bond distance, figure 7C and figure 8. Moreover, Val370 and Cys437 held the Haem group as was observed in the crystal structure. Additionally, the number of hydrogen bonds was monitored throughout the simulations and the average number of hydrogen bonds was found to be, 1.4, 1.2 and 1.3, respectively for Hit1, Hit2 and Hit3, while the reference has represented only an average of 0.2. This demonstrates that the inhibitors have displayed more hydrogen bonds than the reference molecule, figure 7D.

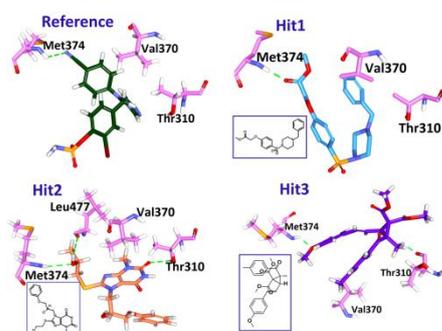


Figure 8. Hydrogen bond interactions and the binding pattern of the reference and the Hits. Green dashed lines represent the hydrogen bond interaction. The corresponding 2D structures of the Hit are represented in purple box.

Table 6. Comprehensive intermolecular interactions between the ligands (reference and Hits) and the protein.

| Na me | Hydrog en bonds | Van der Walls interacti ons | π - Sigma interacti ons | Alkyl/ π -alkyl interacti ons | π - sulphur interacti ons |
|-------------|-----------------------|--------------------------------------|--------------------------------------|--|--|
| Ref (A) | Met37 4 | Trp224 | Thr310 | Val313 | Cys437 |
| | | Ala306 | | Val370 | |
| | | Ala307 | | Val313 | |
| Hit1 (B) | Met37 4 | Thr310 | -- | Val369 | - |
| | | Val370 | | Leu477 | |
| | | Leu372 | | Met374 | |

| | | Ile395 | | | |
|-------------|------------|--------|--------|--------|--------|
| Hit2 (C) | Met37 4 | Ala306 | Val370 | Val370 | Met374 |
| | Thr310 | | | Leu477 | |
| | Leu477 | | | | |
| Hit3 (D) | Met37 4 | -- | Val370 | Val313 | -- |
| | Thr310 | | | Val375 | |

3.6 Binding free energy analysis

Binding free energy calculations conducted after the MD simulations is a time-demanding step, which considers the fluctuations that occur within the protein and the ligand conformation for securing an appropriate lodging of the ligand within the binding pocket [58]. Thus, to validate the binding modes of the protein - ligand complex, a series of calculations have been computed using MM/PBSA [52] [53]. The generated results have rendered favourable ΔG values that range between -20 kJ/mol ~ -120kJ/mol, figure 6B. Upon comparing the binding energies of the reference compound and the Hits, it was disclosed that the Hit compounds have demonstrated a higher binding energy than the reference molecule, such as, -163.8 kJ/mol, -193.4 kJ/mol and -185.3 kJ/mol, while that of the reference compound has rendered -95.82 kJ/mol, respectively. This is indicative of the superior efficacy of the Hits.

3.7 Density functional theory

The molecular orbital energies were calculated with regard to the HOMO and LUMO that are liable for the transfer of charges in a given chemical reaction [59] [60] which further describes the molecules to be attacked by the electrophiles and nucleophiles, respectively [61]. Additionally, the band gap found between the HOMO and LUMO demonstrates the reactivity of the molecules corresponding to smaller gap being more reactive and wider gap implies less reactive and therefore, the molecules with smaller band gap were considered. For the current study, the MD optimized lead compounds along with the reference were subjected to the HOMO and LUMO analysis. The Hit molecules were selected based on the least energy gap and the lowest binding energies as compared to the reference compound, table 7. Additionally, the electrostatic potential maps were computed to probe into the structural aspects of a molecule as it essentially plays an important role in the receptor

ligand interactions and are computed for a set of points in the molecule. Typically, the electron density was plotted by the intensity of the colour that reflects the characteristic feature of a molecule. Subsequently, the red colour refers to the high negatively charged region and thus corresponds to high charge accumulation and the blue represents the charge depletion and is positively charged region. The intermediate colours, such as orange, yellow and green demonstrate the charges mid-way between both the extremes. The order of the colour magnitude can be demonstrated as

Highly negative ---- red < orange < yellow < green < blue ---- highly positive

Reviewing the electron map of the reference compound reveals that the molecule has three highly negative points located at N4, N5 and N21, while the highly positive points were found between H2, H7, H8 and H12. The highly electronegative N21 participated in interacting with the protein Met374 residue as hydrogen bond interaction. Hit1 has displayed five electronegative features; however, O21 was involved in the hydrogen bond interactions with the Met374. Furthermore, the two benzene rings that represented an intermediate colour code, that are present in the Hit1 were found to be involved in the π -interactions. Majority of the hydrogens have represented to be electropositive. In the Hit2, the electro negative O26 and electro positive O18 and H41 have involved in the hydrogen bond interactions. The atoms that exhibited the intermediate colour codes formed additional bonds. Similarly, in the Hit3, the highly negative O17 and the highly positive H37 have been found to be involved in the hydrogen bond interactions, while the benzene rings have been involved in the π interactions. Furthermore, the Molecular Electrostatic Potential maps, (MEPs), figure 9, states that the O atom present towards the methionine residues are crucial in forming the hydrogen bond interaction and thus inhibiting the aromatase. Moreover, it can also be deduced that the lead molecules have higher number of reactive groups as compared with the reference molecule, making them the potential drug candidates.

Table 7. Binding free energies and the energy gap ΔE (band gap) between the reference and the Hit compounds.

| Name | HOMO (eV) | LUMO (eV) | Band Gap ΔE (eV) | Average Binding Energy (kJ/mol) |
|------|-----------|-----------|--------------------------|---------------------------------|
| Hit1 | -0.169 | -0.097 | 0.072 | |

| | | | | |
|-----------|--------|--------|-------|--------|
| | | | | -163.8 |
| Hit2 | -0.199 | -0.102 | 0.096 | -193.4 |
| Hit3 | -0.204 | -0.097 | 0.106 | -185.3 |
| Reference | -0.202 | -0.088 | 0.113 | -95.82 |

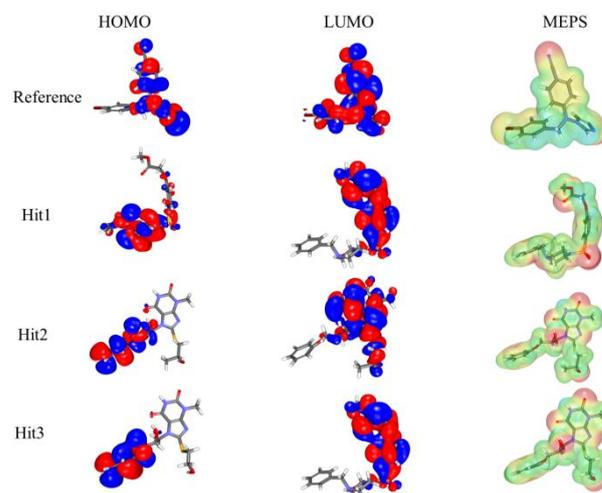


Figure 9. Molecular electrostatic potential (MEPs) maps of the reference and the Hits.

4 Discussion

Several hormones display a neuroendocrine activity and correspondingly influence the seizures more specifically; this can be noticed with testosterone, oestrogen, progesterone. Particularly, the estrogens are epileptogenic as they influence the neural membrane excitability and also reduces the seizure threshold [62] [63] [64]. Additionally, in both men and women reproductive and endocrine disabilities are commonly seen associated [65] [66]. These reports affirm the relationship associated between hormones and seizures. Additionally, reports exist that prove the antiepileptic nature of aromatase inhibitors, besides serving as effective breast cancer drugs [13]. We therefore aimed at identifying the lead anti-aromatase candidates that can simultaneously act against epilepsy.

For the current investigation, the pharmacophore model was generated from a variety of known inhibitors, which could act simultaneously both as breast cancer inhibitors and as antiepileptic drugs. For this reason, the ideal possibility of developing the new drugs is through the ligand-based drug designing approach. With this pursuit a pharmacophore model, Hypo1 was generated

which exhibited 2HBA, 1HyP and 1RA features. Since the model displayed greater correlation values, least total cost value, lower RMSD, the obtained model was considered the best prospective model that was then subjected to validation process such as Fischer's randomization, test set and decoy set methods. The three validation methods ensured Hypo1 could efficiently segregate between the active and the inactive inhibitors. The effectiveness of the model was reaffirmed by its GF value [57]. By taking Hypo1 as the 3D query, the database screening has been performed and 68 prospective Hits were selected after subjecting to ADMET and Lipinski's Rule of Five. Based upon the highest dock score values and the hydrogen bond interaction with the pivotal amino acid residues present within the active sites, the preferred molecules were forwarded to molecular dynamics simulations to evaluate the stability of the systems and further their binding energies were computed. MD results showed that the lead systems were stable throughout the simulations with no major aberrations, projecting the lead molecules to be ideal prospective drug candidates. Furthermore, in order to quantify their reactive groups, the DFT and MEPs were conducted. Subsequently, their orbital energies were assessed to ensure the potency of the inhibitors, with the HOMO and the LUMO energy gap. Finally, three Hit molecules, Hit1, Hit2, and Hit3 were determined as the lead molecules, which satisfied all the necessary pharmacophore features of the Hypo1 and could be considered as the prospective potential aromatase inhibitors. The identified Hit compounds were retrieved from the AXN database. The selected lead compounds could also be marked as probable antiepileptic drugs with increased efficacy, because the lead compounds were chosen based upon the reference molecule with an IC_{50} value of 0.5 nmol/L. This value is far lower than the values of the current AIs that are antiepileptic in nature, thereby, projecting the identified Hit molecules as the prospective drug candidate for both the diseases.

5 Conclusion

Owing by the strong literature evidence that supports the establishment of the aromatase inhibitors as antiepileptic drugs, the current study successfully identifies three lead candidates from the large databases. These candidates have shown greater binding similarity with that of the reference and strong molecular interactions with the key residues as compared with the reference compound. Additionally, they displayed stable MD profiles and rendered lower energy gaps and binding free

energies than the reference compounds. Further, the identified Hits may be tested for their efficacy in some known models as published earlier [67] [68] as these models are reported to be appropriate for studying the effect of aromatase inhibitors that could possibly ameliorate the postmenopausal breast cancer cases. Taken together, we suggest that the identified Hits might be effective against breast cancer and further epilepsy.

6 Acknowledgement

This research was supported by Next-Generation BioGreen 21 Program from Rural Development Administration, Republic of Korea (Grant no. PJ01106202), and supported by the Pioneer Research Center Program through the National Research Foundation (NRF) funded by the Ministry of Science, ICT and Future Planning (NRF-2015M3C1A3023028).

7 References

- [1] J. A. Files, M. G. Ko, and S. Pruthi, "Managing aromatase inhibitors in breast cancer survivors: Not just for oncologists," *Mayo Clin. Proc.*, vol. 85, no. 6, pp. 560–566, 2010.
- [2] A. S. Bhatnagar, "The discovery and mechanism of action of letrozole," *Breast Cancer Res. Treat.*, vol. 105, no. Suppl 1, pp. 7–17, Oct. 2007.
- [3] C. J. Fabian, "The what, why and how of aromatase inhibitors: Hormonal agents for treatment and prevention of breast cancer," *International Journal of Clinical Practice*, vol. 61, no. 12, pp. 2051–2063, 2007.
- [4] S. Hong, A. Didwania, O. Olopade, and P. Ganschow, "The expanding use of third-generation aromatase inhibitors: What the general internist needs to know," *Journal of General Internal Medicine*, vol. 24, no. SUPPL. 2, 2009.
- [5] J. Geisler, "Differences between the non-steroidal aromatase inhibitors anastrozole and letrozole--of clinical importance?," *Br. J. Cancer*, vol. 104, no. 7, pp. 1059–66, 2011.
- [6] Y. Hong, R. Rashid, and S. Chen, "Binding features of steroidal and nonsteroidal inhibitors," in *Steroids*, 2011, vol. 76, no. 8, pp. 802–806.

- [7] J. A. Mayer, C. M. Chuong, and R. Wideltz, "Rooster feathering, androgenic alopecia, and hormone-dependent tumor growth: What is in common?," *Differentiation*, vol. 72, no. 9–10, pp. 474–488, 2004.
- [8] M. Milani, G. Jha, and D. a Potter, "Anastrozole Use in Early Stage Breast Cancer of Post-Menopausal Women.," *Clin. Med. Ther.*, vol. 1, pp. 141–156, 2009.
- [9] D. C. Márquez-Garbán, H.-W. Chen, L. Goodglick, M. C. Fishbein, and R. J. Pietras, "Targeting aromatase and estrogen signaling in human non-small cell lung cancer.," *Ann. N. Y. Acad. Sci.*, vol. 1155, pp. 194–205, 2009.
- [10] I. E. Smith and M. Dowsett, "Aromatase inhibitors in breast cancer.," *N. Engl. J. Med.*, vol. 348, no. 24, pp. 2431–2442, 2003.
- [11] J. Ji *et al.*, "Discovery of novel aromatase inhibitors using a homogeneous time-resolved fluorescence assay.," *Acta Pharmacol. Sin.*, vol. 35, no. 8, pp. 1082–1092, 2014.
- [12] D. Ghosh, J. Griswold, M. Erman, and W. Pangborn, "Structural basis for androgen specificity and oestrogen synthesis in human aromatase.," *Nature*, vol. 457, no. 7226, pp. 219–23, 2009.
- [13] Y. Muftuoglu and G. Mustata, "Aromatase inhibitors and antiepileptic drugs: a computational systems biology analysis," *Reprod. Biol. Endocrinol.*, pp. 1–9, 2011.
- [14] A. G. Marson and Z. A. Kadir, "New antiepileptic drugs: a systematic review of their efficacy and tolerability," *Br. Med. J.*, 1996.
- [15] M. J. Morrell, "Epilepsy in women," *o*, vol. 66, no. 8, pp. 1489–1494, 2002.
- [16] A. P. Aldenkamp, M. De Krom, and R. Reijls, "Newer antiepileptic drugs and cognitive issues.," *Epilepsia*, vol. 44 Suppl 4, no. 9, pp. 21–29, 2003.
- [17] S. Herman, "Sex Hormones and Epilepsy: No Longer Just for Women," *Epilepsy Curr.*, vol. 8, no. 1, pp. 6–8, Jan. 2008.
- [18] D. S. Reddy, "Role of hormones and neurosteroids in epileptogenesis.," *Front. Cell. Neurosci.*, vol. 7, pp. 1–20, 2013.
- [19] M. Heiry, P. Afra, F. Matsuo, J. E. Greenlee, and S. L. Clardy, "Improvement of GAD65-associated autoimmune epilepsy with testosterone replacement therapy," *Neurol. Neuroimmunol. Neuroinflammation*, vol. 2, no. 5, p. e142, Oct. 2015.
- [20] G. Montouris and G. L. Morris, "Reproductive and sexual dysfunction in men with epilepsy.," *Epilepsy Behav.*, vol. 7 Suppl 2, pp. S7–14, 2005.
- [21] S. K. Singh, "Aromatase inhibitors in male sex.," *Indian J. Endocrinol. Metab.*, vol. 17, pp. S259–61, 2013.
- [22] M. Y. Roth, J. K. Amory, and S. T. Page, "Treatment of male infertility secondary to morbid obesity.," *Nat. Clin. Pract. Endocrinol. Metab.*, vol. 4, no. 7, pp. 415–9, 2008.
- [23] L. E. Crosnoe-Shipley, O. O. Elkelay, C. D. Rahnema, and E. D. Kim, "Treatment of hypogonadotropic male hypogonadism: Case-based scenarios.," *World J. Nephrol.*, vol. 4, no. 2, pp. 245–253, 2015.
- [24] A. G. Herzog, "Catamenial epilepsy: Definition, prevalence pathophysiology and treatment," *Seizure*, vol. 17, no. 2, pp. 151–159, 2008.
- [25] A. G. Herzog, "Psychoneuroendocrine aspects of temporolimbic epilepsy. Part II: Epilepsy and reproductive steroids.," *Psychosomatics*, vol. 40, no. 2, pp. 102–8, 1999.
- [26] M. Numazawa, T. Kamiyama, M. Tachibana, and M. Oshibe, "Synthesis and structure-activity relationships of 6-substituted androst-4-ene analogs as aromatase inhibitors.," *J. Med. Chem.*, vol. 39, no. 11, pp. 2245–52, 1996.
- [27] B. Su, J. C. Hackett, E. S. Díaz-Cruz, Y. W. Kim, and R. W. Brueggemeier, "Lead optimization of 7-benzyloxy 2-(4'-pyridylmethyl)thio isoflavone aromatase inhibitors," *Bioorganic Med. Chem.*, vol. 13, no. 23, pp. 6571–6577, 2005.
- [28] M. Numazawa, K. Yamada, S. Nitta, C.

- Sasaki, and K. Kidokoro, "Role of Hydrophilic Interaction in Binding of Hydroxylated 3-Deoxy C19 Steroids to the Active Site of Aromatase," *J. Med. Chem.*, vol. 44, no. 24, pp. 4277–4283, Nov. 2001.
- [29] L. W. L. Woo *et al.*, "First Dual Aromatase-Steroid Sulfatase Inhibitors," *J. Med. Chem.*, vol. 46, no. 15, pp. 3193–3196, Jul. 2003.
- [30] A. Cavalli *et al.*, "Enantioselective nonsteroidal aromatase inhibitors identified through a multidisciplinary medicinal chemistry approach," *J. Med. Chem.*, vol. 48, no. 23, pp. 7282–7289, 2005.
- [31] S. Ulmschneider *et al.*, "Synthesis and Evaluation of Imidazolylmethylenetetrahydronaphthalene s and Imidazolylmethyleneindanes: Potent Inhibitors of Aldosterone Synthase," *J. Med. Chem.*, vol. 48, no. 6, pp. 1796–1805, Mar. 2005.
- [32] D. Lesuisse *et al.*, "Structure–Activity Relationships of a New Family of Steroidal Aromatase Inhibitors. 1. Synthesis and Evaluation of a Series of Analogs Related to 19-[(Methylthio)methyl]androstenedione (RU54115)," *J. Med. Chem.*, vol. 39, no. 3, pp. 757–772, Jan. 1996.
- [33] S. Sakkiah, S. Thangapandian, S. John, Y. J. Kwon, and K. W. Lee, "3D QSAR pharmacophore based virtual screening and molecular docking for identification of potential HSP90 inhibitors," *Eur. J. Med. Chem.*, vol. 45, no. 6, pp. 2132–2140, 2010.
- [34] S. Rampogu, "Role of quassinoids as potential antimalarial agents: An in silico approach," *Anc. Sci. Life*, vol. 35, no. 2, pp. 85–89, 2015.
- [35] H. Li, J. Sutter, and R. Hoffman, "HypoGen: An Automated System for Generating 3D Predictive Pharmacophore Models," in *Pharmacophore Perception, Development and Use in Drug Design SE - IUL Biotechnology Series*, O. Guner, Ed. La Jolla, CA: International University Line, 2000.
- [36] A. K. Debnath, "Pharmacophore Mapping of a Series of 2,4-Diamino-5-deazapteridine Inhibitors of Mycobacterium avium Complex Dihydrofolate Reductase," *J. Med. Chem.*, vol. 45, no. 1, pp. 41–53, Jan. 2002.
- [37] S. Rampogu, M. Son, C. Park, H.-H. Kim, J.-K. Suh, and K. Lee, "Sulfonanilide Derivatives in Identifying Novel Aromatase Inhibitors by Applying Docking, Virtual Screening, and MD Simulations Studies," *Biomed Res. Int.*, vol. 2017, pp. 1–17, 2017.
- [38] J. Fei, L. Zhou, T. Liu, and X.-Y. Tang, "Pharmacophore modeling, virtual screening, and molecular docking studies for discovery of novel Akt2 inhibitors," *Int. J. Med. Sci.*, vol. 10, no. 3, pp. 265–75, 2013.
- [39] P. J. Lipinski, C.A.; Lombardo, F.; Dominy, B.W.; Feeney, "Experimental and computational approaches to estimate solubility and permeability in drug discovery and development setting," *Adv. Drug Deliv. Rev.*, vol. 23, pp. 3–25, 1997.
- [40] G. Jones, P. Willett, R. C. Glen, A. R. Leach, and R. Taylor, "Development and validation of a genetic algorithm for flexible docking," *J. Mol. Biol.*, vol. 267, no. 3, pp. 727–748, Apr. 1997.
- [41] M. L. Verdonk, J. C. Cole, M. J. Hartshorn, C. W. Murray, and R. D. Taylor, "Improved protein-ligand docking using GOLD," *Proteins Struct. Funct. Genet.*, vol. 52, no. 4, pp. 609–623, 2003.
- [42] E. A. Cino, W.-Y. Choy, and M. Karttunen, "Comparison of Secondary Structure Formation Using 10 Different Force Fields in Microsecond Molecular Dynamics Simulations," *J. Chem. Theory Comput.*, vol. 8, no. 8, pp. 2725–2740, Aug. 2012.
- [43] S. S. Mallajosyula, S. Jo, W. Im, and A. D. MacKerell, "Molecular Dynamics Simulations of Glycoproteins using CHARMM," *Methods Mol. Biol.*, vol. 1273, pp. 407–429, 2015.
- [44] V. Zoete, M. A. Cuendet, A. Grosdidier, and O. Michielin, "SwissParam: A fast force field generation tool for small organic molecules," *J. Comput. Chem.*, vol. 32, no. 11, pp. 2359–2368, 2011.
- [45] S. Rampogu, A. Baek, A. Zeb, and K. W. Lee, "Exploration for novel inhibitors showing back-to-front approach against

- VEGFR-2 kinase domain (4AG8) employing molecular docking mechanism and molecular dynamics simulations,” *BMC Cancer*, vol. 18, no. 1, p. 264, Mar. 2018.
- [46] M. Parrinello, “Polymorphic transitions in single crystals: A new molecular dynamics method,” *J. Appl. Phys.*, vol. 52, no. 12, p. 7182, 1981.
- [47] M. A. Anwar and S. Choi, “Structure-Activity Relationship in TLR4 Mutations: Atomistic Molecular Dynamics Simulations and Residue Interaction Network Analysis,” *Sci. Rep.*, vol. 7, p. 43807, Mar. 2017.
- [48] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije, “LINCS: A linear constraint solver for molecular simulations,” *J. Comput. Chem.*, vol. 18, no. 12, pp. 1463–1472, 1997.
- [49] S. Miyamoto and P. A. Kollman, “Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models,” *J. Comput. Chem.*, vol. 13, no. 8, pp. 952–962, 1992.
- [50] T. Darden, D. York, and L. Pedersen, “Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems,” *J. Chem. Phys.*, vol. 98, no. 12, p. 10089, 1993.
- [51] W. Humphrey, A. Dalke, and K. Schulten, “VMD: Visual molecular dynamics,” *J. Mol. Graph.*, vol. 14, no. 1, pp. 33–38, 1996.
- [52] R. Kumari, R. Kumar, and A. Lynn, “g_mmpbsa—A GROMACS Tool for High-Throughput MM-PBSA Calculations,” *J. Chem. Inf. Model.*, vol. 54, no. 7, pp. 1951–1962, Jul. 2014.
- [53] N. A. Baker, D. Sept, S. Joseph, M. J. Holst, and J. A. McCammon, “Electrostatics of nanosystems: Application to microtubules and the ribosome,” *Proc. Natl. Acad. Sci.*, vol. 98, no. 18, pp. 10037–10041, Aug. 2001.
- [54] S. Verma *et al.*, “Hydrophobic Interactions Are a Key to MDM2 Inhibition by Polyphenols as Revealed by Molecular Dynamics Simulations and MM/PBSA Free Energy Calculations,” *PLoS One*, vol. 11, no. 2, p. e0149014, Feb. 2016.
- [55] N. Homeyer and H. Gohlke, “Free energy calculations by the Molecular Mechanics Poisson-Boltzmann Surface Area method,” *Mol. Inform.*, vol. 31, no. 2, pp. 114–122, 2012.
- [56] A. Nakata and T. Tsuneda, “Density functional theory for comprehensive orbital energy calculations,” *J. Chem. Phys.*, vol. 139, no. 6, 2013.
- [57] S. John, S. Thangapandian, S. Sakkiah, and K. W. Lee, “Potent BACE-1 inhibitor design using pharmacophore modeling, in silico screening and molecular docking studies,” *BMC Bioinformatics*, vol. 12 Suppl 1, no. Suppl 1, p. S28, 2011.
- [58] J. Wang, Y. Deng, and B. Roux, “Absolute Binding Free Energy Calculations Using Molecular Dynamics Simulations with Restraining Potentials,” *Biophys. J.*, vol. 91, no. 8, pp. 2798–2814, Oct. 2006.
- [59] C. Ai *et al.*, “Investigation of binding features: Effects on the interaction between CYP2A6 and inhibitors,” *J. Comput. Chem.*, vol. 31, no. 9, pp. 1822–1831, Jul. 2010.
- [60] S. Sakkiah and K. W. Lee, “Pharmacophore-based virtual screening and density functional theory approach to identifying novel butyrylcholinesterase inhibitors,” *Acta Pharmacol. Sin.*, vol. 33, no. 7, pp. 964–78, 2012.
- [61] S. Rampogu *et al.*, “Computational Exploration for Lead Compounds That Can Reverse the Nuclear Morphology in Progeria,” *Biomed Res. Int.*, vol. 2017, pp. 1–15, 2017.
- [62] T. B??ckstr??m, “EPILEPTIC SEIZURES IN WOMEN RELATED to PLASMA ESTROGEN and PROGESTERONE DURING the MENSTRUAL CYCLE,” *Acta Neurol. Scand.*, vol. 54, no. 4, pp. 321–347, 1976.
- [63] D. E. WOOLLEY and P. S. TIMIRAS, “The Gonad-Brain Relationship: Effects of Female Sex Hormones on Electroshock Convulsions in the Rat,” *Endocrinology*, vol. 70, no. 2, pp. 196–209, Feb. 1962.

- [64] H. E. Scharfman and N. J. MacLusky, "The Influence of Gonadal Hormones on Neuronal Excitability, Seizures, and Epilepsy in the Female," *Epilepsia*, vol. 47, no. 9, pp. 1423–1440, Sep. 2006.
- [65] J. Isojärvi, "Disorders of reproduction in patients with epilepsy: Antiepileptic drug related mechanisms," *Seizure*, vol. 17, no. 2, pp. 111–119, 2008.
- [66] A. G. Herzog, "Disorders of reproduction in patients with epilepsy: Primary neurological mechanisms," *Seizure*, vol. 17, no. 2, pp. 101–110, 2008.
- [67] W. Yue, D. Zhou, S. Chen, and a Brodie, "A new nude mouse model for postmenopausal breast cancer using MCF-7 cells transfected with the human aromatase gene.," *Cancer Res.*, vol. 54, pp. 5092–5095, 1994.
- [68] R. R. Tekmal and V. R. Durgam, "A novel in vitro and in vivo breast cancer model for testing inhibitors of estrogen biosynthesis and its action using mammary tumor cells with an activated int-5/aromatase gene," *Cancer Lett.*, vol. 118, no. 1, pp. 21–28, 1997.