Simulation Studies 3D QSAR and Molecular Docking, on a Point Mutation of Protein Kinase B with Flavonoids Targeting Ovarian Cancer

Suchitra Ajjarapu
Govind Ballabh Pant University of Agriculture and Technology
Apoorv Tiwari
Department of Computational biology and bioinformatics, Jacob Institute of Biotechnology and Bio-Engineering, Sam Higginbottom University of Agriculture, Technology and Sciences, Allahabad
Gohar Taj
Govind Ballabh Pant University of Agriculture and Technology
Dev Singh
Department of Biotechnology, Institute of Biosciences and Biotechnology, Chhatrapati Shahu Ji Maharaj University, Kanpur, Uttar Pradesh
Sakshi Singh
Department of Molecular and Human Genetics, Banaras Hindu University, Varanasi
Sundip Kumar (malik.sundip@gmail.com)
Govind Ballabh Pant University of Agriculture and Technology

Research Article

Keywords: AKT1, point mutation, ADME, QSAR, virtual docking, dynamic simulations

Posted Date: February 1st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-143480/v1

Abstract

Cancer is the world's dreaded disease and its prevalence is expanding globally. The study of integrated molecular networks is crucial for the basic mechanism of cancer cells and its progression. During the present investigation we have examined different flavonoids that targets protein kinases B (AKT1) protein which exerts their anticancer efficiency intriguing the role in cross talk cell signalling, by metabolic processes through *in-silico* approaches. Molecular dynamics simulation (MDS) was performed to analyse and evaluate the stability of the complexes under physiological conditions and the results were congruent with molecular docking. This investigation revealed the effect of a point mutation (W80R), considered based on their frequency of occurrence, with AKT1. The ligand with high docking scores and favourable behaviour on dynamic simulations are proposed as potential W80R inhibitors. A virtual screening analysis was performed with 12000flavonoids satisfying the Lipinski's rule of 5 according to which drug-likeness is predicted based on its pharmacological and biological properties to be active and taken orally. The pharmacokinetic ADME (adsorption, digestion, metabolism and excretion) studies featured drug likeness. Subsequently, a statistical significant 3D-QSAR model of high correlation coefficient (R2) with 0.992 and cross validation coefficient (Q2) with 0.6132 at 4 component PLS (partial least square) were used to verify accuracy of the models. The molecular dynamics simulation of this study showed that the compound is Taxifolin (I-UPAC namely2-(3,4-dihydroxyphenyl)-3,4 –dihydro-2H-chromene-5,7-diol of C15H14O5, of CID ID-443637 evidenced a better interaction with docking score (-9.63Kcal/mol) exhibited the binding affinity with W80R mutant protein thus reflecting that natural inhibitor can be considered for experimental evaluation which provides targeted insights for new combination of drugs in forming a network in pharmacology.

1 Introduction

Ovarian cancer marks the most lethal gynaecological malignancy which ranks the fifth leading cause of cancer deaths in females[1].It is estimated that there are 22530 cases with a mortality rate of approximately 13980 deaths in the United States in 2019[1] .Ovarian cancers are categorized into 3 types based on cell origin: epithelial, stromal and germ cell[2]. The low survival rate and poor prognosis of ovarian cancer is due to a lack of screening methods at the early stages and ineffective treatments for advanced stages of disease[3]. Moreover it is very crucial to dissect the role of tumor causing microenvironment during early stage, proliferation, and metastasis. Thus, it becomes paramount to understand the root cause from different views of its molecular pathogenesis, histological subtypes, hereditary factors, epidemiology, methods of treatment and diagnostic perspectives. The Cancer Genome Atlas (TCGA) revealed that the expression of AKT1, AKT2 and AKT3 was associated with poor patient survival[4]. The leading cause of disease is due to genetic and epigenetic changes of the cellular genome. So, numerous small drug molecules of AKT gene targeting mutations such as, FOXO, glucose metabolism(GSK3), apoptotic proteins (BAD,NF-kB, FKHR). Cell cycle arrest, apoptosis, DNA repair (MDM2) are critical in disease progression. Among various kinases, over expression of AKT1 protein and associated mutations play a deciding role in cross-talk cell signalling in causing cancer. Recent studies have introduced assorted therapeutic agents as targets specific for cancer driven factors involved in inhibition of ovarian cancer development. One such factor of kinase family is protein kinase B/serine-threonine serves as a decisive mediator of the P13K/AKT/mTOR cell signaling pathway that has distinct physiological functions such as cell growth, survival, proliferation, and metabolism[5]. Structurally AKT1 consists of three domains, including an N-terminal pleckstrin homology, a central catalytic kinases domain, and C-terminal domain[6].

AKT1 is the kinase which connects upstream signals from PI3K and mammalian targets of rapamycin complex2 (mTORC2) with downstream signals to mTORC1 and effectors such as mTOR, GSK3b along with phosphorylation cascade which acts as substrates that induce cell cycle progression, protein synthesis, lipid and protein phosphatases, glucose metabolism and cell growth [7]. AKT1 is mutated and AKT2 is amplified in about 40% AKT1 is inhibited by tumor suppressors including phosphatase and tensin homology (PTEN) and inositol polyphosphate 4-phosphatase type 2(INPP4B)[8,9,10]. Therefore, targeting ATP binding cleft of AKT gene by inhibitors (natural/synthetic) has become attractive strategy for treating patients in ovarian cancer. Interestingly, AKT1 gene inhibitors showed strong binding affinity with mutant forms when compared to the native form. However, the emergence of acquired drug resistance in patients found to limit its usage in last phase of clinical trials. In ovarian cancer, overexpression of AKT is associated with advanced-stage platinum resistance[11]. As an isoform of the AKT family, AKT1 is observed to be expressed unduly in a wide assortment of many human cancers including breast and ovarian cancers[12,13]. This study scrutinizes substitution mutation from tryptophan to arginine at 80 residue position. The underlying molecular mechanism is assumed to cause conformational changes in native protein structure (AKT1) which modify covalent bond interaction by limiting their practical application. On that account, there is need to search and develop novel as well as regimes that can counteract the drug resistance induced by AKT1 gene. However, the molecular interactions and atomic stability for the W80R have been considered as the novel and taken as a crucial platform for the present study.

W80R results in increased repression of FOXO 3 compared to wild type AKT1 in an invitro assay which then predicted to result in a gain of AKT1 protein function. FOXO is a transcription factor in the nucleus induces CGN2 transcription in epithelial ovarian cancer cells with enhanced catenin activity. The absence of Wnt ligand dissociates catenin from the destruction complex and translocates to the nucleus where it acts with the FOXO3 factor which is known to play a role in the W80R protein pathway. Abnormal activation of this pathway leads to hyperactivation of catenin, which has been reported in ovarian cancers. W80R is one of the reported mutants of AKT1 cancer which cause missense driver mutation with 238T > C of the coding sequence, also CDS (change in the nucleotide sequence as a result of mutation, where the syntax here used is identical to the method used for the peptide sequence) mutation c.238T > A with gene location 14q32.33 [14] in the uterus section causing endometrial cancer. It has been proved that W80R contains highly conserved residues damaged by polyphen2, targeting through PI3K/AKT1/mTOR pathway of substitution-missense variant type affecting exon of protein domain PH (the UniProt Consortium 2019) and SIFT prediction as 3[12]. The mutant W80R-Q79K on combination found to be displayed a very strong membrane localization and hyperactivation in transfected HeLa cells in both presence and absence of serum under fluorescence microscopy[15]. The previous studies of AKT1 co-occurring mutations(like Q79K-W80R) found to be hyperactive equal to E17K mutant widely distributed in different tissues such as endometrium (homozygous and heterozygous), large intestine (caecum), prostrate (with heterozygosity condition) breast cancers involving cross-talk signaling pathways[16]. The deleterious mutations of AKT1 (E17K and W80R) concluded to be of functional relevance exclusively in myxoidtumors [17]. The altering mutations promote growth factor independent cell proliferation as compared to wild type AKT1[18]. AKT1 gene alterations account for most of the genetic drive contributing to the

pulmonary sclerosing haemangioma which is a benign tumor development [19]. It was observed in the patients receiving genomically targeted therapy that W80R mutant found to be in clinical benefit of SD 4 mo+(stable disease), working efficiently with synthetic drugs temsirolimus and ixabepilone targeting ovary granulosa cell[20]. In line with, the inhibition of AKT1 or its mutant proteins has been recognized as a compelling strategy for the treatment of cancers with[21] induce ovarian tumor angiogenesis[22] and in immune evasion[23].

Existing chemotherapeutic drugs have developed resistance to the novel compounds along with side effects despite of enormous progress in anticancer drug discovery. Hence more targeted strategies are required to develop with sensitivity and specificity. Most of the successful anticancer compounds were originated from natural sources or as their analogues. Flavonoids are naturally occurring secondary metabolites consisting of polyphenols having therapeutic benefits in multiple ways. These are low-molecular-weight compounds with non-nitrogenous properties consisting of C6-C3-C6 as a backbone with different classes[24] and their activities are structure-dependent. Chemically, flavonoids depend on their structural class, degree of hydroxylation, substitutions, and conjugations, and degree of polymerization[25]. Several mechanisms have been proposed for the effect of flavonoids at the initiation and promotion stages of the carcinogenicity including influences on development and hormonal activities[26]. Flavonoids falls under 6 different categories based on the functional group flavones (luteolin, apigenin), flavonols (quercetin, kaempferol), flavanones (naringenin), flavanonol (taxifolin), isoflavones, and flavan-3-ols (genistein, epicatechin, catechin, wedelactone, ellagic acid, silibinin, folstein, parthenoilods, oridonin, curcumin, reservertol. The choice of this study has been relied on the compounds of family called flavonoids with tremendous variety of pharmacological and biochemical consequences including hepatoprotective, antidiabetic, cardioprotective, anti-tumor, neuroprotective, and anti-inflammatory and played a wonderful role in the preclusion of Alzheimer's disease[27]. In earlier investigation in this area has demanded series of chemical methods and animal models to synthesis lead compounds with more time, investment, and level of exposure. To overcome this issue, the computational approaches have been developed reliably in predicting the mutation both in induced drug resistance and also to design resistance evading drugs. As a result of above mentioned short falls, the present study has aimed on the dynamic simulation at molecular level and molecular docking studies on taxifolin targeting W80R mutant protein in protein kinase B/AKT1 protein of Ovarian cancer for designing therapeutic. This computational study rely on learning and pattern classification methods (phylogeny, neural systems, vector machines, and FATHMM servers) which can classify mutations, create 3D protein structures.

2 Materials And Methods

2.1 Sequence retrieval and structure analysis of selected protein

The amino acid sequence of AKT1 protein was retrieved from the Uniprot database with accession number P31749. The primary structure of the protein was elucidated using the ProtParam tool [28 29] of the Expasy server and the difference between physical and chemical properties of the AKT1 protein (wild) and mutant (W80R) were evaluated. Factors such as physicochemical properties, molecular weight, theoretical pI (isoelectric point), half-life, instability index (II), aliphatic index (AI), extinction coefficient (EI), grand average hydropathy (GRAVY), and site of origin were analyzed. The secondary structure properties prediction was carried out by the RAMPAGE server, which provides the configuration score like the total number of helices, turns, coils, predicted solvent accessibility, with the range, existed from 0 (highly buried) to 9 (exposed region) depending on the residue exposed. Normalized B-factor is measured for a selected protein as Z score which is a combination of template and profile-based prediction where residues are higher than zero are considered as less stable during experimental structures. The mutant protein W80R was edited manually at the amino acid position number and submitted to homology modelling.

2.2 Homology modelling

The 480 amino acid residue length of W80R protein was retrieved to recognize the appropriate template for structure modelling and functional prediction of the protein. This modelling depends mainly on a sequence alignment between the target and template sequence whose structure has been experimentally determined, the 3D structure of target protein using its template was visualized by PYMOL tool; based on template-target alignment. These theoretical structural models of the W80R protein were ranked based on the normalized discrete RMSD values. The model with the lowest RMSD score was considered as the best model[30].

2.3 Evaluation of the structure model

The quality of AKT1 and mutant form W80R models was assessed by many tools to test the stability and reliability of the model. PROCHECK suite[31] quantifies the residues in favourable zones of the Ramachandran plot, were used to evaluate the stereochemical quality of the model. ERRAT tool[32] find the overall quality factor of the protein and was used to check the statistics of non-bonded interactions between different atom types. The compatibility of the atomic model (3D) with its amino acid sequence was determined using the VERIFY 3D program. Swiss PDB viewer 4.1.07 was used for the energy minimization of the predicted AKT1protein along with its mutant form. The W80R model was further subjected to structural analysis and verification server to evaluate its quality, before and after energy minimization. ProSA tool[33] was employed for the refinement and validation of the minimized structure to check the native protein folding energy. The superimposition of the proposed model of AKT1 protein along with mutant form with its closest-structural homolog was carried out using chimera 1.11[34].

2.4 Selection and preparation of ligands

Natural compounds database containing more than 12,000 ligands were aimed to the AKT1 protein family were downloaded from the Pubchem library[35] and subjected to ligand preparation by ligprep wizard application of the Maestro 9.3[36]. Ligprep tool was used to prepare the high quality of ligands, such as the addition of hydrogen's, conversion of 2D to 3D structures, corrected bond angles and bond lengths, with lower energy structure, stereochemistry's, and ring

conformation followed by minimization in the optimized potential of OPLS 2005 force field[37, 38]. Properties such as ionization did not change and tautomers were not generated, specifically retained chiralities. Compounds were selected based on the lowest energy.

2.5 Preparation of protein molecule and active site prediction

The protein was modelled by using the protein preparation wizard of Schrodinger Suite; by adding hydrogen atoms, optimizing hydrogen bonds, and verifying the protonation states of His, Gln, and Asn. Energy minimization was carried out using constraint 0.3A RMSD and OPLS 2005 force field. The sitemap tool was used to identify binding pockets of W80R protein[39].

2.6 Receptor grid generation

Receptor grid generation was done by the Glide application[40]. The receptor grid for W80R was generated using active site residues which were identified Sitemap tool. Once the grid has generated, the ligands are docked to the protein (W80R) using Glide version 5.8(Grid-based Ligand Docking with Energetics) docking protocol. The scaling factor (0.25) and partial charge (1 Å) represents cut-offs of Vander Waals radius scaling.

2.7 Molecular docking

Molecular docking procedures were consistently carried out using a preparation of protein of Schrodinger[41] and defining the grid on the active site of the protein. GLIDE molecular docking tool uses computational simulation methods for evaluating particular poses and ligand flexibility. GLIDE systematic method, a new approach for rapid, accurate molecular docking and its output G-score, is found to be an empirical scoring function, is a combination of diversified attributes. G-score is calculated in Kcal/mol, encompass ligand-protein interaction energies, hydrophobic interactions, hydrogen bonds, internal energy, pi-pi stacking interactions, root mean square deviation (RMSD), and desolvation. GLIDE modules of the XP visualize analyses of the specific ligand-protein interactions. The ligands were docked using Extra Precision mode(XP) and conformers were evaluated using the Glide(G) score. The G score is calculated as follows:

$$G\,Score = a * vdW + b * Coul + Lipo + Hbond + Metal + BuryP + RotB + Site$$

where vdW denotes vanderwaals energy, Coul denotes columb energy, Lipo denotes lipophilic contact, H-bond indicates hydrogen bonding, Metal indicates metal-binding, BuryP indicates penalty for buried polar groups, RotB indicates penalty for freezing rotatable bonds, site denotes polar interactions in the active site and a = 0.065 while b = 0.130 were the coefficients of vdW and Coul.

2.8 ADME properties studies

Calculation of absorption, distribution, metabolism, excretion, and toxicity (ADME/T) properties were performed for best-docked ligand molecules by QikProp software. This software predicts various limiting factors such as QP log Po/w, QPlog BB, SASA, FOSA, FISA, PISA, WPSA, volume, donarHB, acceptorHB, dip^2/V, AC*DN*5, Caco, QlogS, rotors, rule of 5, rule of 3, the overall percentage of human oral absorption, etc[42]. Lipinski's rule of five[43] measures the drug-likeness for the prediction of a chemical compound as an orally active drug based on biological compounds and pharmacological properties.

2.9 Analysis of cancer-associated mutants

The deleterious W80R mutations that are specific for cancers were predicted using the FATHMM server (http://fathmm.biocompute.org.uk/)[44] which allows the distinct difference between cancer-promoting/driver mutations and other germline polymorphisms. The gene number identifiers (UniProt id) along with mutant form as a text were provided as the input for the prediction.

2.10 Molecular alignment and 3D QSAR studies and validation

The key component of 3D QSAR analysis is the arrangement of the molecules based on the scaffold they share which generated using the training was set of 44 molecular poses with a grid spacing of 1 Å PLS (partial least square) algorithm to establish the relationship between biological activity and different structural features. The training set was adjusted to 50%. Three models were generated by Gaussian filed extension as Gaussian steric, electrostatic, hydrophobic, hydrogen bond donor, hydrogen bond acceptor, and aromatic ring fields. CoMFA and CoMSIA are the tools employed as independent variables in PLS regression analysis. The best model was chosen based on the criteria of statistical robustness and visualized using contour map modules. The predictive power and stable models were assessed using the leave one odd (LOO) cross-validation method. The crucial aspects for the test set statistics include RMSE, Q2, SD, R2, R2CV, R2scramble, stability, F, P, Q2, Pearson's r which indicates the predictive ability of the model. A Scatter plot was generated in correlation with predicted activity on the Y-axis and observed activity on the X-axis of the data set model[46].

2.11 Contour maps visualisation

Representation of the fields as contours (surfaces) or as color intensities of the fields on the grid can be displayed in different styles. Based on the field type, the colors are designed and field intensities are shown for one field at a time. The fields with greater absolute values than the cut-off were presented at the maximum brightness.

2.12 Molecular dynamics simulation

The simulation of protein-ligand complexes was implemented by GROMACS 4.5.5(Groningen machine for Chemical Simulations) software[45]. The complex with the lowest binding energy was selected for molecular dynamics (MD) simulation. The ligand parameters were analyzed using PRODRG online server[47] in the framework of GROMACS force-field 43a1[46]. The ligand enzyme complex was solvated at a simple point charge as well as a water box under periodic boundary conditions using 1.0nm distance protein to the box faces. The system was then neutralized by Cl^- or Na^+ counter ions for the W80R complex with ligand respectively. To perform energy minimization, the complex was equilibrated under volume, constant number of particles, and temperature condition for 100ps at 300k, followed by 100ps. All the covalent bonds with hydrogen bonds were considered using a linear constraint solver algorithm. The electrostatic interactions were treated using the particle mesh Ewald method[48] Further MD simulation studies were noted for 20ns to check the accuracy and stability of the ligand-protein complexes. The potential of each trajectory produced after MD simulations were analyzed using g_rms, g_rmsf, and g_h bond of GROMACS utilities[49] the root mean square deviation (RMSD), the root mean square fluctuation (RMSF), with hydrogen bonds formed between the ligand and protein complex.

3 Results

3.1 Mutant W80R sequence analysis

The development of anticancer compounds with variegated pharmacological effects becomes a very paramount topic and hence main class of secondary metabolites, both dietary and synthetic flavonoids have been subjected to clinical trials[50]. Definite beneficial biological activities of dietary flavonoids including antioxidants[51] anticancer[52] and cardio-protective properties[53] have been identified in a series of previous studies. Flavonoids are known for their wide exposure to chemo-preventive, chemotherapeutic activities, and the availability of the compound in plant sources for the human diet in routine consumption[54].

The analysis of the mutant W80R protein sequence of the AKT1 has 480 amino acid residue which plays a very crucial role in metabolism, cell proliferation, cell survival, growth, and angiogenesis, was downloaded from Uniprot with accession number (P31750). The amino acids in the protein sequence of W80R were found to exhibit larger contents of lysine, leucine, glutamic acid, and alanine. The ProtParam tool was used for the W80R protein sequence to compute physio-chemical parameters such as molecular weight of 5565.45 kD. The W80R had a pI (isoelectric point) of 5.99 indicating its acidic nature (pI < 7.0) with an aliphatic index (AI) (71.69). The protein volume is occupied by aliphatic side chains such as lysine, leucine, glutamic acid, and alanine. The instability index of W80R measured 35.76 of the unstable nature. The grand average of hydropathicity (GRAVY) of W80R protein was lower (-0.583), which proves its high affinity with water. The comparison of statistical characteristics are showing the differences among wild AKT1 and mutantW80R using the ProtParam tool (Table 1). The comparison of sequence analysis of W80R mutant protein with AKT1(wild) at nucleotide and protein level was same with a slight difference, thus proving-T, C-G rich region, and properties such as molecular weight, amino acid composition, theoretical pI, aliphatic index, and grand average of hydropathicity (GRAVY) were found in an appropriate range of influencing the protein stability.

Table 1
Comparison of primary sequence analysis using the ProtParam tool between AKT1 (wild) and W80R (mutant)

S.No	Parameters	AKT1	W80R
1.	Molecular weight	5586.7kD	5565.4kD
2.	pI	5.75	5.99
3.	Aliphatic Index	71.69	71.69
4.	Instability Index	35.47	35.76
5.	GRAVY	-0.575	-0.583
6.	Atoms	7772	7776
7.	Total number of Asp + Glu residues in a protein content	77	76
8.	Total number of Arg + Lysresidues in a protein content	66	68

3.2 Homology modelling of W80R mutant protein

The 480 amino acid residue length of W80R protein was subjected to BLASTp analysis against RCSB PDB to identify the suitable template for comparative structural modelling and functional prediction. The result of the BLASTp search revealed a template (PDB id 3O96) of high-level identity with the target sequence of AKT1. The query coverage (100%) showed high degree of identity between two proteins (AKT1 and W80R) of 480 sequence length, and E value (2e-60) is expected value obtained by hits, percentage identity defines the extent of two sequences, Modeller 9.13 has generated 5 models of W80R, among these the lowest score is considered as stable which is thermodynamically subjected to further refinement. The lowest RMSD as 0.18 score model was considered as the best one for further validation purposes[30]. Finally, three dimensional (3D) structure of selected protein using its template was visualized by PYMOL tool.

3.3 Model assessment and validation

The stability of the protein was constructed based on the backbone of torsion angles psi and phi which were evaluated by the PROCHECK server that computes the amino acid residues in the existing zones of Ramachandran plot analysis of W80R mutant forms (Table 2). The information presented in the Table 2 depicts Ramachandran plot through RAMPAGE server where W80R mutant protein has 79.3% amino acids falls in the most favored region with located major active binding sites, while 13.8% in an allowed region and 6.9% residues in the outlier region of the plot with lesser significance. SAVES analysis was conducted to confirm the quality of the protein model followed by ProSA, RMSD assessment for a high-quality structural model for virtual screening. The quality of the predicted model of AKT1 protein and a W80R mutant was supported by a high ERRAT score of 81.99 in an acceptable protein environment. The VERIFY 3D results of W80R showed 81.88% of the residues with an average 3D-1D score > = 0.2, indicating the stability of the model. 'WHAT IF' tool examines the coarse packing quality, the model protein structure, reflecting the acceptance of good quality. The reliability of the W80R form was confirmed by ProSA (Fig. 1) which achieved a Z score of -7.92 kcal/mol compared to the wild form AKT1 having a Z score − 7.2kcal/mol, wherein the energy is negative, reflects the best quality of the model. The quality of the model was evaluated through the comparison of predicted structure with experimentally determined structure followed by superimposition and atoms RMSD assessment using Chimera 1.11, proved that the predicted model is good and quite similar to the wild protein.

Table 2
Comparison of secondary structure using RAMPAGE server between AKT1and W80R mutants.

S.No	Protein Properties	AKT1(wild)	W80R(mutant)
1.	Total amino acids	480	480
2.	Number of residues in favoured region (-98.0% expected)	388 (81.2%)	379 (79.3%)
	Number of residues in allowed region(-2.0% expected)	61 (12.8%)	66 (13.8%)
3.	Number of residues in outlier region	29 (6.1%)	33 (6.9%)

3.4 Active site and score prediction

A proven algorithm for binding site identification and evaluation of the drug ability of those sites lead to modify hit-compounds to enhance receptor complementarity. The active site was performed using a sitemap tool to assess each site by calculating attributes such as size, volume, amino acid exposure, hydrophobicity, hydrophilicity, donor/acceptor ratio. The most reliable score was obtained in the binding pockets of W80R. The predicted amino acids in the active region were LEU^{156}, GLY^{157}, GLU^{234}, MET^{281}, ASN^{279}, GLU^{278}, LYS^{276}, ASP^{274}, THR^{291}, ASP^{292}, PHE^{293}, GLY^{294}, LEU^{295}, $GLU^{298}$of site score for the selected model was 1.128, druggability score − 1.149 with Volume 384.486 and size measured was 179 for further docking analysis.

3.5 Analysis of cancer associated mutants

The mutation impact for the protein W80R was classified using the FATHMM server derived from the new FATHMM-MKL algorithm. It distinguishes between cancer-promoting/driver mutations and other germline polymorphisms. This algorithm predicts the functional, molecular, and phenotypic consequences of the missense mutation of a functional protein using hidden Markov models (HMMs), representing the alignment of homologous sequences and conserved protein domains with "pathogenicity weights", representing overall tolerance of protein/domain to mutations [44] The gene number identifier (uniprot id) along with mutant form as a text was provided as the input for the prediction based on the FATHMM server predictions with a score − 1.12 responsible for benign cancer. The functional scores for individual mutations were obtained from the FATHMM-MKL server which falls in the range of 0–1 known as single p- values fall in the range of (0–1) where the values below 0.5 are predicted as benign and above 0.5 are deleterious.

3.6 Determination of ADME profile

Molecular properties of the selected compounds were studied using Qikprop and chosen based on the Lipinski rule of five which marks the most important activity in drug discovery and development. Multifarious Insilco techniques have been employed to measure the drug-likeness for a compound based on numerous descriptors. Calculation of absorption, distribution, metabolism, excretion, and toxicity (ADME/T) properties were predicted for best-docked ligand molecules using Qikprop software. Qikprop computes almost 20 physical descriptors over a wide range of predicted properties unlike a fragment-based approach, by screening compound libraries for hits and play a lead optimization that can be used to improve predictions by fitting to experimental data and also to generate QSAR models. The detailed analyses of chemical and molecular descriptors and also solubility properties were tabulated in Table3, 4, and 5. The results of ADME properties are an important index to check the clinical candidates have reached the required standard. It is revealed that compounds in the table were ranked based on the potential drug properties. According to a previous study, ~ 40% of failures to develop medicine in the development phase are due to poor biopharmaceutical properties (pKa-dissociation constant and bioavailability)[55]. The ADME as a deal medicine has following characteristics, hydrogen bond donar < 5; hydrogen bond acceptor < 10; molecular weight < 500Da; lipid water partition coefficient < 5; water solubility partition coefficient − 6.5 < logs < 0.5; and polar surface area 7.0–20.

Table 3
Detailed analysis of ADME properties of ligands using QIKPROP software

CID id	Mol.wt	Dipole	SASA	FOSA	FISA
443639	310.392	7.125	515.633	442.178	73.455
44264122	276.334	2.395	457.915	415.498	42.417
71424203	336.387	5.934	489.544	447.123	42.421
5907705	336.387	0.667	489.55	447.131	42.419
6421929	298.338	0.002	600.799	533.867	66.932
6440365	310.392	6.34	532.097	458.644	73.453
10688261	336.387	3.009	553.24	511.992	41.249
10852057	304.388	8.606	478.22	435.803	42.417
14407192	280.366	4.742	505.615	428.846	76.769
19792482	349.256	21.006	553.576	388.72	42.774
19792563	308.419	2.685	554.716	511.942	42.774
22321203	290.358	7.377	519.093	427.419	91.673
24952793	342.393	4.9	577.054	369.818	89.296
24952797	391.25	24.947	516.006	322.823	88.262
24966389	350.413	7.928	576.902	488.639	88.262
44538447	361.224	9.432	540.471	339.66	90.009
44560954	280.366	2.771	484.13	407.362	76.769
44610342	318.412	7.987	624.49	536.048	88.441
53693682	286.37	5.469	532.126	460.735	71.391

*SASA: total solvent accessible surface area in square angstroms using a probe with a 1.4A radius; FISA: hydrophilic component of the SASA(SASA on N, O, and H on heteroatom);FOSA: hydrophobic component of the SASA(saturated carbon and attached hydrogen); CID ID: compound Id from PUBCHEM database

Table 4
Detailed analysis of ADME properties of ligands using QIKPROP software

CID ID	VOLUME	DonorHB	AcceptorHB	dip2/v	AC*DN
443639	846.643	1	2.75	0.059953	0.005333
44264122	721.579	0	4	0.007948	0
71424203	799.111	2	6	0.044058	0.017333
5907705	799.119	2	6	0.000557	0.017333
6421929	914.122	0	3.5	0	0
6440365	855.378	1	2.75	0.046994	0.005168
10688261	914.124	2	6	0.009905	0.015337
10852057	769.386	0	4	0.096272	0
14407192	805.364	1	2.75	0.02792	0.005439
19792482	891.173	0	4	0.495123	0
19792563	896.185	0	4	0.008045	0
22321203	812.517	1	2.5	0.066971	0.004816
24952793	899.071	1	4.75	0.026703	0.008232
24952797	833.649	1	4.75	0.746536	0.009205
24966389	906.417	1	4.75	0.069341	0.008234
44538447	846.213	1	4.75	0.105131	0.008789
44560954	791.867	1	2.75	0.009697	0.00568
44610342	980.133	1	4.75	0.06508	0.007606
53693682	821.324	1	2.75	0.036413	0.005168

*.Donor HB: it is the calculated number of hydrogen bonds that would be donated by the solute to water molecules in an aqueous solution, values are averages take over many configurations, so they can be non-integer; Acceptor HB: it is estimated as the number of hydrogen bonds that would be accepted by the solute from water molecules in aqueous solution; dip2/v: square of the dipole moment divided by the molecular volume. This is the key term given in Kirkwood-Onsager equation for the free energy described of solvation of a dipole moment with volume V; AC*DN: index of cohesive interaction in solids; Volume: total solvent-accessible volume in the cubic angstroms using a probe with 1.4 A radius.

Table 5
Solubility prediction parameters for molecular descriptors

CID	QPlogpC1	QPlogPW	Qplogpoct	QPlogPw	QPlogPo	QPlogS	ClQPlogS	QPlogHER	QPPCaco	QPPMDCK	QPlogKp
5318278.0	27.1	6.7	11.8	5.4	2.9	-4.5	-3.0	-4.1	1992.1	1042.1	-2.8
5469422.0	22.7	5.0	8.7	5.1	1.8	-2.8	-1.9	-4.0	3923.3	2167.8	-2.3
5469423.0	25.8	6.3	14.3	10.0	1.5	-3.3	-2.2	-3.9	3923.0	2167.6	-2.3
5907705.0	25.8	6.3	13.7	10.0	1.5	-3.3	-2.2	-3.9	3923.1	2167.7	-2.3
443637.0	30.4	6.9	10.9	5.1	3.1	-5.8	-2.6	-5.4	2297.1	1215.5	-2.8
44264122.0	27.4	6.7	11.7	5.5	3.0	-4.8	-3.0	-4.5	1992.2	1042.1	-2.8
71424203.0	30.4	7.3	15.4	10.2	2.1	-4.3	-2.2	-4.4	4024.7	2228.4	-2.3
10852057.0	24.6	5.5	10.6	5.1	2.1	-3.2	-2.3	-4.0	3923.3	2167.8	-2.3
14407192.0	25.4	6.2	10.8	5.4	2.6	-4.3	-2.6	-4.3	1853.1	963.6	-2.8
19792482.0	29.5	7.5	18.3	5.3	3.3	-5.1	-3.4	-4.6	3892.8	10000.0	-2.3
19792563.0	29.7	6.6	11.1	5.3	2.9	-4.6	-2.4	-4.6	3892.8	2149.6	-2.3
22321203.0	25.1	6.3	11.1	5.1	2.7	-4.5	-2.9	-4.6	1338.3	677.9	-3.0
24952793.0	30.3	-7.7	13.7	8.4	2.4	-4.9	-3.3	-18.3	1409.6	717.0	-2.7
24952797.0	26.6	-9.3	22.7	7.5	2.3	-4.3	-3.9	-17.7	1441.8	2759.8	-3.1
24966389.0	29.5	-8.5	13.8	7.6	2.4	-4.9	-2.9	-17.7	1441.8	734.7	-3.1
44538447.0	27.1	-8.6	14.0	7.6	2.4	-4.8	-3.4	-17.7	1387.8	2852.3	-3.1
44560954.0	24.9	-10.9	10.4	5.4	2.5	-4.0	-2.6	-17.7	1853.1	963.7	-2.8
44610342.0	32.4	-7.3	14.7	7.7	2.8	-5.8	-2.5	-17.7	1436.1	731.6	-3.1
53693682.0	26.1	-9.8	11.0	5.4	2.7	-4.8	-2.7	-17.7	2084.0	1094.1	-2.739

*QPlogPoct: predicted octanol/gas partition coefficient; QPlogPw: predicted water/gas partition coefficient; QPlogPo/w: predicted octanol/water partition coefficient;ClQPlogS: conformation –independent predicted aqueous solubility, logs. S in mol dm^{-3} is the concentration of the solute in a saturated solution that is in equilibrium with the crystalline solid; QPlogHERG: predicted IC 50value for blockage of HERG K^+ channels; QPPCaCo: predicted apparent CaCo-2 cell permeability in nm/sec; Caco-2 cells are a model for the gut blood barrier; QPlogKp: predicted skin permeability, logKp; QPlogS: Predicted aqueous solubility, log S, S in mol dm^{-3} is the concentration of solute in the saturated solution that is in equilibrium with the crystalline solid; QPPMDCK: Predicted apparent MDCK cell permeability in nm/sec, MDCK cells are considered to be a good mimic for the blood- barrier; QPlogpCl: Predicted hexadecane/gas partition coefficient.

3.7 QSAR studies and validation

A dataset of 44 ligand compounds was chosen for statistical studies and classified as the training set and test set into 50% for suitable 3D QSAR model development. The graphical interface allowed building dataset into training and testing equally for 50% by generating a correlation coefficient. The graph obtained for all/training models/test models were observed in Fig. 2A. Molecular descriptors (ligands) were divided into a training set and test set (Table 6) with parameters such as phase QSAR, phase activity, % extrapolation, predicted error, and predicted activity. QSAR built model was generated based on docking poses and substructure alignment was represented with standard deviation for the regression as 10.7913, R^2 gives 0.8226 measures the coefficient of determination, where R^2 always lies between 0 and 1, R^2C yields 0.2055 for cross-validated where R^2 is obtained to leave an N-out approach, R^2scramble (R^2 is regression or coefficient of determination) obtained as 0.4889 computing the average value obtained using scrambled activities of Fig. 2B. It measures the degree to which the molecular fields can fit random data, stability statistical measure observed to be 0.379 for the model predicting the changes obtained in the training set composition F with 92.7 measuring higher F value indicates more statistical significant regression. Pearson 5.95e-09, root mean square error predictions were to be 22.02 (RMSE), $Q^2$for predicted activities with 0.2915, Pearson-r correlated with predicted activity, and observed activity observed for test set with 0.7508. The test set was determined within the maximum range of training set.

Table 6
List of statistical analysis generated using field based QSAR

CID ID	Phase QSAR	Phase ACTIVITY	% Extrapolated	Prediction Error	Prediction Activity
368569	Test	314.424	0.653	5.16645	319.59
443637	Training	358.477	0.458	-24.7856	333.692
461867	Training	296.408	0	-0.612247	295.796
2751794	Training	314.424	0	-5.94593	308.478
44264122	Test	280.366	0.049	18.5276	298.894
6324890	Training	330.423	0	5.28888	335.712
71424203	Test	332.482	0.227	-74.0301	258.452
11566133	Test	330.466	0.055	-13.8518	5.51
14379076	Training	344.45	0	-3.96902	340.481
14633296	Test	286.456	0.246	-35.6831	250.773
15690458	Test	298.381	0.044	-2.43817	295.943
20975206	Test	236.353	0.22	9.44186	245.795
22973545	Test	300.483	0.037	-3.7.753	296.775
24207736	Training	250.38	0.039	-5.88.58	296.775
25203072	Training	286.456	0	12.9111	299.367
44254444	Test	328.407	0.044	-2.11377	326.293
44254874	Test	344.407	0.045	-0.262634	344.144

3.8 Contour visualisation

The contour maps (Fig. 3) were used to illustrate the fields required for biological activity. Field-based QSAR interface creates electrostatic, hydrophobic, and steric fields for optimization and leads discovery. The represented green contour indicates the bulky group in a favorable region. The contour map depicts hydrophobicity in the solvent-accessible hydrophobic pocket steric fields are considered as the most favorable regions with a high Glide score. The obtained results have shown the steric and Gaussian field fractions are much larger than other fields suggesting most of the binding energy has been contributed from hydrophobic interactions.

3.9 Molecular docking studies

Molecular docking is the paramount computational tool to configure (Fig. 4) all the possible active conformations of binding at the active site for the receptor molecule. Before performing the docking protocol, the co-crystallized ligand was re-docked into the crystal structure of the W80R receptor molecule to evaluate the reliability of the standard precision algorithm of the Glide. A dataset of flavonoids family along with its structural analogs comprising 7000 ligands was selected. Upon generation of Epik for suitable tautomeric states per 16 for each ligand, 12000 ligands were chosen entirely as a whole set for virtual screening with W80R mutant protein. The top three ligands with the best binding energy were considered for further analysis (Fig. 4).

Several hydrogen bond interactions were found in the docking result. The top-scoring compound belongs to CID-443637 was having the lower binding energy with the Glide score of -9.63 Kcal/mol. The hydroxyl group of SER^{208} formed a hydrogen bond with GLU^{198} and also found interacted with THR^{211}revealing the strongest stability with the receptor molecule. The three hydrogen bond interactions provide the guarantee for stable conformation of a binding ligand molecule to protein structure which influences the activity of ligand. The interaction with 1 pi ~ cation recognized as an energetically significant[56] noncovalent binding interaction proves to exist in a quite strong platform both in the gas phase and liquid media[57] which is a special hydrophobic interaction with LYS^{268}having a cationic side chain amino acid, indicating that the geometry is biased towards aromatic amino acid, one that experiences a favorable pi ~ cation interaction [58] having IUPAC name 2-(3,4-dihydroxy phenyl)-3,4 –dihydro-2H-chromene-5,7-diol of C15H14O5 (Fig. 5). The second highest molecule of CID 71424203 has the binding energy of -9.43 kcal/mol forming three hydrogen bond interactions with amino acid residues THR^{211}, MET^{227}, SER^{205} aromatic amino acid residue, and $TYR^{474}$of 1 pi ~ pi stacking interaction. The residue TRP^{80} between two aromatic amino acids has a separation of -3.35A (vDw) having IUPAC name 2,5,7–trihydroxy-3-(4-hydroxyphenyl)-2,3-dihydrochromen-4-one of $C_{15}H_{12}O_6$ (Fig. 6). The third compound CID 44264122 has the binding energy of -9.36 kcal/mol with hydrophobic contacts with residues such as $LYS^{268,}$ THR^{291} having IUPAC name 3,4 Difluoro-8,9-dihydroxbenzo[c] chromen-6-one of C13H6F2O4 (Fig. 7).After the comparison of all the three models, the compound CID with 443637 with the lowest energy is chosen for further molecular dynamics simulation studies.

Lower Glide score represents the most and highest favorable binding affinity. Hydrogen bond interactions, pi-interactions, pi staking of the best poses were visualized and interpreted using XP visualizer with descriptors (Table 7) in ascending order. It rewards the topmost ligands for hydrogen bond with lengths and angles deviating significantly from "ideal" hydrogen-bond interaction (1.65A H-A distance,180 D-H A angle)[42] .The PhobEn measures hydrophobic

enclosure reward on the protein. The lipophilic EvdW is the term for hydrophobic region lies within receptor and Ligand proximity. For the obtained data, PiCat, ClBr, PhobEnPa, penalties, HB penal, exposed penal, zprot remained at zero, whereas other properties of descriptors were exhibited accordingly.

Table 7
Top ranked hit compounds of docking with protein W80R obtained using XP visualise

CID ID	Dock Score	GScore	Lipophilic EvdW	PhobEn	HBond	Electro	RotPenl	Activity
443637	-9.63	-9.63	-3.22	-2.17	-3.02	-0.72	0	-0.08
4426412	-9.43	-9.43	-3.11	-2.11	-2.11	-0.57	0	-0.06
7142423	-9.36	-9.36	-3.98	-0.86	-3.2	-0.64	0.12	0
586089	-9.34	-9.34	-3.82	-2.55	-1.04	-0.15	0.13	-0.1
443902	-9.32	-9.32	-3.93	-1.7	-2.05	-0.45	0.09	-0.31
847733	-9.26	-9.26	-5.25	-2.7	-0.7	-0.25	0.14	-0.32
443639	-9.2	-9.2	-4.95	-2.09	-1.08	-0.7	0.1	-0.16
6901802	-9.18	-9.18	-4.5	-1.65	-2.4	-0.32	0.52	0
2157490	-9.16	-9.16	-3.81	-1.89	-2.49	-0.88	0.41	-0.84
3239870	-9.15	-9.15	-3.27	-1.97	-0.7	-0.18	0.14	-0.32
5880326	-9.14	-9.14	-4.7	-1.84	-1.87	-0.64	0.41	-0.85
4013888	-8.97	-8.97	-4.24	-1.87	-0.7	-0.23	0.14	-0.32
5942266	-8.9	-8.9	-4.73	-2.65	-1.27	-0.32	0.57	0
586088	-8.71	-8.71	-4.18	-1.98	-0.85	-0.43	0.13	-0.11
80337	-8.71	-8.71	-3.45	-1.85	-1.7	-0.35	0	0
229016	-8.66	-8.66	-4.66	-2.31	-0.67	-0.65	0.13	-0.24

*G Score-total G score along with sum of XP terms(G score = a*vdW + b*Coul + Lipo + Hbond + Metal + BuryP + RotB + Site where vdW is vanderwaals energy, Coloumb energy, Lipo is lipophilic contact, Hbondis hydrogen bonding, Metal is metal-binding, BuryPis penalty for buried polar groups, RotBis penalty for freezing rotatable bonds, site is polar interactions in the active site and a = 0.065 while b = 0.130 were the coefficients of vdW and Coul.

Dock score

-Vanderwaals + coulombic + HBonds represents potentiality of bonding. In simple rigid systems, the ligand is searched in a 6 dimensional rotational or translational space to fit in the binding site, which can serve as a lead compound for drug design[60]

Lipophilic term is derived from the hydrophobic grid potential and the fraction of the total protein ligand vdW energy, PhobEn- can be as hydrophobic enclosure reward for penalty for ligands with large hydrophobic contacts and low hydrogen bond scores phobic penal for penalty for exposed hydrophobic ligand groups, Rot Penal for rotatable bond penalty.

4.10 Molecular dynamics simulation

MD simulations were performed to W80R protein-ligand complex with least binding energy (Fig. 8a). The results of MD trajectories were evaluated by root mean square deviation (RMSD) and root mean square fluctuation (RMSF) plot which could provide significant insights into understanding structural changes in atomic details. The RMSD is a significant parameter to analyze the equilibrium in MD trajectories, which is estimated for backbone atoms of W80R protein and taxifolin ligand complex. For W80R protein complex, the fluctuations were raised about 0.3 to 0.4nm during initial stage (Fig. 8a).Clear and noticeable deviations were observed in the residues of RMSD values with increase in time from 200ps to 600ps. Majority of residues resulted to attain a stable state at 600psbetween 0.45nm to 0.5nm. At the same time, W80R protein–ligand complex fluctuated from 700ps to 900ps at 0.4nm and remained stable between 0.4nm to 0.45 nm until the end of simulation [61]

RMSF results were obtained by considering the average of all backbone residues of atoms to inspect the local variations of protein flexibility (Fig. 8b). The fluctuations observed above have an important role in protein complex flexibility and thus affect protein-ligand activity and stability. The high RMSF value shows more flexibility with a maximum level of fluctuation in the residue positions of 355 and405 at 6Åof the backbone structure, while the minimum RMSF shows very limited movements. The RMSF graph for the W80R-ligand complex was shown in Fig. 8b. The W80R-ligand complex has attained the amino acid residues at 455 and500 also show a fluctuation at 5Å of RMSF. While at positions 305 and355 at4Å indicate similar steep up graph at 5Å.The amino acid residues between 15, 55, and 105, 155, have shown medial deviation at 3Å.

To determine the residue interaction network, RING2.0 software identifies all types of non-covalent interactions in atomic levels which have wide different energies and lengths. The output has been visualized in two different ways (i) interaction network which has been visualized using different labels and (ii) structural contacts using RING_viz-script for pymol (Fig. 9). The applications of RING 2.0 have a growth in protein folding patterns, domain-domain

communication and catalytic activity, inter-intrachain interactions that combine both solvent and ligand atoms. Residue interaction network (RIN) describes the single amino acid as nodes and physicochemical properties as edges including covalent and non-covalent bonds. RIN has become common practice to explore the complexity inherent in macromolecular systems[59].

4 Discussion

The enigmatic in ovarian cancer is that in nearly 75% of patients, cancer do recurse during first two years and fail to respond to available therapeutic drugs due to acquired resistance[62,63] in addition to late diagnosis in advanced clinical stages and metastasis within the peritoneal cavity[63,64]. Therefore, there is immediate need to design novel drugs to deal with the existing problem.Numerous studies since a decade has reported that Flavinoids as candidates are meant to block, retard, or reverse the progression of carcinogenesis[80]. Although various studies have been carried out using flavinoids but the anticancer mechanisms have not been defined clearly. However, it was found that the flavonoids such as quercetin and silymarin induce anti-cancer mechanisms in ovarian cancer cells[65,66]. Consequently, the effects of apigenin, luteolin and myricetin on ovarian cancer have to uncover the link between potential mechanisms underlying their anticancer effects. Quercetin inhibits cell proliferation of ovarian cancer cell line of SKOV-3 which correlated with findings of[67] caused on concentration and time-dependent manner[68] showed to inhibit UVB induced skin cancer cell proliferation and induce apoptosis in vivo models upon apigenin treatment. Taxifolin *invitro* studies have been efficient especially in anticancer, antimicrobial activities but leaves a strong gap in the *invivo* studies at root level.

The lead compound in the current study was recognised as taxifolin which has potent to exhibit anti-cancer effects on U2OS and Saos-2 in osteosarcoma cell lines by inhibiting the proliferation and disrupting colony formation. *In vivo* studies exhibit intraperitoneal administration in nude mice bearing U2OS xenograft that resists tumor growth. This potency is known to arrest the G1 phase of the cell cycle in U2OS and Saos-2 cell lines. Taxifolin has known to function by inhibiting colon carcinogenesis by NF-kB mediated Wnt/b catenin signalling through upregulation of Nrf2 pathway while downregulation in genes such as TNF-α, COX-2, β-catenin, and cyclin-D1 were inhibited by NF-kB and Wnt signalling pathway[69]. It is also reported that injection of taxifolin has reduced the proliferative activity on wistar rats with benign prostatic hyperplasia[70]. Taxifolin also has an excellent report on antiangiogenic effect by new blood vessels and its branches per area of chick chorioallantoic membrane assay which is inhibited by tube formation on matrigel matrix in human umbilical vein of endothelial cells which were evaluated against tachyzoites in vitro with IC50 of 1.39μg/mL($p \leq 0.05$) along with pyrimethamine. Taxifolin has known to express anti-proliferative effect on cancer cell types by inhibiting cell lipogenesis and inhibits the fatty acid synthesis in cancer cell lines which is able to prevent the growth of cancer cells[79].

An extensive animal (rat) study of antioxidant activity on taxifolin acid has shown the decreased lipid peroxidation in the serum and liver levels. The presence of OH groups at position 5th and 7th together with 4-OXO function in the A and C rings were meant for scavenging effect while O-dihydroxy group in the B ring provided stability[71]. Consequently, In vivo studies on taxifolin induced in apoptosis of HCT116 and HT 29 cells revealed PARP1 over expression is responsible for ovarian cancer. AKT and catenin proved that down-regulated expression by taxifolin on HCT 116 and HT 29 cells demonstrates a decline in p-AKT and catenin in a dose of 40 μM against DMSO altering in G2 cell cycle and its regulators[72]. The expression levels of AKT, SKP-2, v-mc avian myelocytomatosis viral oncogene homolog(c-myc) and p-Ser^{473}, have reduced activity on AKT gene by taxifolin[73]. Although the above mentioned experimental outcomes have contributed for diversified pharmacological activities with AKT1 protein, we still lack the detailed and molecular changes wrt to W80R mutant protein of AKT1 family. Consequently, the marginal overview of the molecular mechanism and atomic level with W80R mutation has aimed to identify hits for optimization from large data set of compounds from the PubChem database screening of flavonoids in parallel to W80R mutant protein of AKT1 targeting ovarian cancer. The Table 1 for the receptor molecule W80R of 480 amino acid sequence provides the detailed knowledge about the stability of protein using Protoparam tools of Expasy server. The extensive evaluation on W80R sequence at nucleotide level reveals its density, while other parameters such as A-T,C-G rich region, molecular weight, amino acid composition, theoretical pI, aliphatic index, instability index and GRAVY significantly stand up for stability factor. The most favoured region by RAMPAGE server was assessed to be 79.3% (Table 2) with active site binding. Furthermore, the reliability of the protein model has been assessed by 3D or homology modelling. Therefore, Generation of 3D protein structure from sequence information, in the absence of experimentally determined structures in protein data bank through computational approaches has become topmost priority in the scientific community based on structural biology research for several decades[74,75,76]. The protein was henceforth evaluated with SAVES server (structural analysis and verification) for quality check, structural refinement through energy minimization in lowest energy state in its stable conformation, followed by ProSA (Fig. 1) and superimposition analysis with experimentally determined template structure as well as atoms and RMSD assessment to obtained a high -quality structural model for virtual screening[77]. The predicted score for 3D homology model of RMSD for the W80R protein was 0.18, the model was considered as the best one for further validation purposes.3D QSAR studies have been performed with structural similarity to predict the unknown/untested ligands for better potency by correlating mathematical and statistical values. QSAR models can prioritize ideas in virtual screening as well in the optimization of lead compounds. Thus it has gained acceptance in *in-silco* drug discovery. The scatter plot QSAR tool (Fig. 2) assessed the molecular fields for the compounds which estimate the stability and establish statistical value to be 0.379 predicting the changes obtained in the training set composition with 92.7 measured higher F indicates more statistical significant regression. The dataset of 44 ligands was classified into test and training models randomly with combined mathematical and statistical approaches for the drug candidate represents phase activity of 358.477% extrapolated for 0.458 with the predicted activity of 333.692 and predicted error of -24.7856 which was a good combination as a lead compound (Table 6). As per Lipinski's rule of five, a drug is good molecule if it possesses ADME (absorption, distribution, metabolism, and excretion) properties[43]. All the physicochemical properties and drug-likeness were listed in Table 3, 4 and 5consequently; it becomes easy for the lead compound to enter the mammalian cell to interact with proteins and regulating gene expression in metabolic pathways. The top 10 hits obtained by molecular docking were further docked into the active binding sites of protein using a sitemap tool of above score 1 and grid generation followed by XP protocol (Table 7). However, a contour map is one such tool used in the present study to determine favourable regions based on field-based QSAR which depends on steric, electrostatic, hydrophobicity in solvent-accessible pockets based on least binding energy. This application plays a vital role in combination therapies of multi-drug-resistant conditions as well in drug discovery.

The evaluated hydrophobicity gives an accurate check for the drug-ability of a compound (Fig. 3). Sitemap tool treats entire protein to locate binding sites whose size, the extent of solvent exposure is assessed based on scoring function by ranks. Active sites are ranked based on ligand propensity of binding measured by their ability to bind tightly for passively absorbed small molecules. Among the predicted combinations, active site amino acid residues of site score 1.128, drug-ability score – 1.149, volume 384.486, and size 179 (Fig. 4) were taken for further analysis. Taxifolin holds good interactions with the binding domain of W80R, highest Glide score of -9.63kcal/mol with O-H of SER ^{208} and H bond GLU ^{198} and THR ^{211} amino acid residues and one pi-cation interaction and one hydrophobic bond with LYS ^{268} (Fig. 5). The lead molecule satisfied all the surface area calculations using QIKPROP tool of SASA, FISA,FOSA, PSA and partition coefficient of Qplogpoct, QPlogPw, QPlogPo, QPlogS, ClQPlog, QPlogHER, QPPCaco, QPPMDCK, QPlogKp, wherefore, this inhibitor of the PI3K/AKT pathway has shown diverse aptitudes for anticancer activity in both preclinical and clinical experimental values and also supported through *in-silico* analysis.

It has been reported by the administration of taxifolin in colorectal cancer cell lines and in HCT 116 xenograft mouse model had shown excellent antitumor activity. The studies proved that the administration of taxifolin hindered the mRNA expression of β-catenin thus compiling anti-proliferative activity which was arbitrated by PI3K/AKT signal by jamming Wnt/ β –catenin signaling transduction through hampering the β expression[72]. The elucidation of suppression by taxifolin on nuclear factor-kB, C-Fos, and mitogen-activated protein kinase also decreased osteoclast specific gene expression including Trap, Mmp-9, Cathepsin K, C-Fos, Nfatc1, and Rank; taxifolin osteoclastogenesis via regulation of many RANKL signaling pathways was also confirmed[78,79]. Taken together, these studies demonstrated that Wnt/catenin pathway plays a crucial role in ovarian cancer development and this idea also laid a strong platform for the development of targeted curatives.

CID- 44264122 with 2 hydrogen bonds of a hydroxyl group (-OH) interacting with LYS^{268}, THR^{291}, ILE^{290}, and THR^{211} and ILE^{290} and –OH with THR ^{291} and oxy bond with residue LYS ^{268} (Fig. 6) with Glide XP score – 9.43Kcal/mol. The hydrogen bond interaction with residues of TYR^{474}, SER^{215}, THR^{211}, with 1 pi-pi interaction at TRP^{80} residue, and 1 pi-cationic interaction bonding with LYS^{265} with G score – 9.36Kcal/mol showed good hydrophobic interactions (Fig. 7). The molecular dynamics simulation was performed to obtain lowest error and data loss. The fluctuations in relative positions of atoms in protein-ligand complex explains the structural stability (RMSD) at 0.45nm to 0.50nm between 600 to 800ps (Fig. 8a). The RMSF has shown a steep up graph at 5A with a slight medial deviation and not much structural change in protein cavity was observed[80,81,82] (Fig. 8b). Residue interaction network (RINs) consider single amino acid as nodes and physio-chemical interactions as edges (Fig. 9) representing the protein structure as RINs have become common practice to explore the complexity inherent in macromolecular systems. Henceforth, the taxifolin has been suggested as a drug for human use in clinical trials.

5 Conclusions

The mutant forms of the amino acid were found to induce pathological outcomes disrupting the native conformation of a protein. The W80R mutation in the PH domain of AKT1 had been reported to cause ovarian cancer by in-vitro studies and recorded in My Cancer genome database. The synthetic drugs reported in clinical trials are being used currently. To examine the detailed molecular mechanism of W80R, we conducted molecular docking along with dynamic simulation studies to understand the stability of the mutant structure, which is known to cause a damaging effect of the mutation. Furthermore, a rise in RMSD values for stability in trajectory and conformational drifts were observed in W80R protein. The expected result supported the molecular cause in a mutant form which resulted in a gain of ovarian cancer. However, experimental evaluation or in vivo studies is recommended for further validation.

Declarations

Acknowledgment:

The authors acknowledge the financial support provided by the BTISNet, Department of Biotechnology, Government of India, to the Bioinformatics Sub-DIC G. B. Pant University of Agriculture & Technology, Pantnagar.

Author Contribution Statement:

SK and SMA conceived the idea, SMA and AT designed the study, SMA performed the docking study, SS and DBS done the molecular dynamics. AT, GT, DBS and SK helped in the critical revision of the manuscript. All authors read and approved the final manuscript.

Conflict of Interest:

Authors declare no conflict of interest.

References

1. Siegel RL, Miller KD, Jemal A. (2019) Cancer statistics, 2019. CA Cancer J Clin.;**69**,(1):7–34.
2. Gilks CB, Prat J. (2009) Ovarian carcinoma pathology and genetics: recent advances. Hum Pathol.;**40,**(9):1213–23.
3. Howlader N, Noone AM, Krapcho M, Garshell J, Neyman N, Altekruse SF, Kosary CL, Yu M, Ruhl J, Tatalovich Z, Cho H, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA, editors(2013). SEER Cancer Statistics Review, 1975–2010, National Cancer Institute. Bethesda, MD: National Cancer Institute
4. Ghoneum A, Said N. (2019) PI3K-AKT-mTOR and NFκB Pathways in Ovarian Cancer: Implications for Targeted Therapeutics. Cancers (Basel).;**11**,(7):949. Published 2019 Jul 5. doi:10.3390/cancers11070949

5. Dancey, J.E. (2006). Therapeutic targets - MTOR and related pathways. Cancer biology & therapy. 5:1065-73.
6. Du-Cuny, L., Song, Z., Moses, S., Powis, G., Mash, E.A., Meuillet, E.J., Zhang, S. (2009). Computational modelling of novel inhibitors targeting the Aktpleckstrin homology domain. Bioorg Med Chem. **17**,(19):6983-92
7. Mundi PS, Sachdev J, McCourt C, Kalinsky K. (2016) AKT in cancer: new molecular insights and advances in drug development. Br J ClinPharmacol. **82**, (4):943-956. doi:10.1111/bcp.13021
8. Martini M, De Santis MC, Braccini L, Gulluni F, Hirsch E. (2014) PI3K/AKT signaling pathway and cancer: an updated review. Ann Med.;**46**,(6):372-383. doi:10.3109/07853890.2014.912836
9. Hanrahan AJ, Schultz N, Westfal ML, et al. (2012) Genomic complexity and AKT dependence in serous ovarian cancer [published correction appears in Cancer Discov. 2(4):376. Janikariman, Manickam [corrected to Janakiraman, Manickam]]. Cancer Discov. 2012;**2**,(1):56-67. doi:10.1158/2159-8290.CD-11-0170
10. Rodgers SJ, Ferguson DT, Mitchell CA, Ooms LM. (2017) Regulation of PI3K effector signalling in cancer by the phosphoinositide phosphatases. Biosci Rep.;**37**,(1):BSR20160432. Published 2017 Feb 10. doi:10.1042/BSR20160432
11. Brasseur K, Gévry N, Asselin E. (2017) Chemoresistance and targeted therapies in ovarian and endometrial cancers. Oncotarget. 17;**8**,(3):4008. [PubMed: 28008141]
12. Liu, J., Wu, D.C., Qu, L.H., Liao, H.Q., Li, M.X. (2018). The role of mTOR in ovarian Neoplasms, polycystic ovary syndrome, and ovarian aging. Clin Ana. **31**, (6):891-898.
13. Hsu, J.H., Shi, Y., Hu, L., Fisher, M., Franke, T.F., Lichtenstein, A. (2002). Role of the AKT kinase in expansion of multiple myeloma clones: effects on cytokine-dependent proliferative and survival responses. Oncogene. **21**,(9):1391-400
14. Hart R and Prlic A. (2015) Universal Transcript Archive Repository. Version uta_20180821. San Francisco CA: Github;. https://github.com/biocommons/uta
15. Laurianne B., Anne-Laure Todeschini, AurélieAuguste, Sabine Sarnacki, Delphine Flatters, et al. (2015). A Hot-spot of In-frame Duplications Activates the Oncoprotein AKT1 in Juvenile Granulosa Cell Tumors.. EBioMedicine, Elsevier, **2**, (5):421-31.
16. Tate, J.G., Bamford, S., Jubb, H.C., ZbyslawSondka, David, M. Beare, NidhiBindal, Harry Boutselakis, Charlotte, G. Cole, CelestinoCreatore, Elisabeth Dawson, Peter Fish, BhavanaHarsha, Charlie Hathaway, Steve C Jupe, Chai Yin Kok, Kate Noble, Laura Ponting, Christopher C Ramshaw, Claire E Rye, Helen E Speedy, Ray Stefancsik, Sam L Thompson, Shicai Wang, Sari Ward, Peter J Campbell, Forbes, S.A. (2018). COSMIC: the Catalogue of Somatic Mutations In Cancer, Nucleic Acids Research, **47**,(D1):D941–D947.
17. Trautmann, M., Cyra, M., Isfort, I., Jeiler, B., Krüger, A., Grünewald I, Steinestel K, Altvater B, Rossig C, Hafner S, Simmet T, Becker J, Åman P, Wardelmann E, Huss S, Hartmann W. (2019). Phosphatidylinositol-3-kinase (PI3K)/Akt Signaling is Functionally Essential in MyxoidLiposarcoma. Mol Cancer Ther. **18**, (4):834-844.
18. Parikh, C., Janakiraman, V., Wu, WI., Foo, C.K., Kljavin, N.M., Chaudhuri, S., Stawiski, E., Lee, B., Lin, J., Li, H., Lorenzo, M.N., Yuan, W., Guillory, J., Jackson, M., Rondon, J., Franke, Y., Bowman, K.K., Sagolla M, Stinson J, Wu TD, Wu J, Stokoe D, Stern HM, Brandhuber BJ, Lin K, Skelton NJ, Seshagiri S. (2012). Disruption of PH-kinase domain interactions leads to oncogenic activation of AKT in human cancers. Proc Natl AcadSci U S A. **109**,(47):19368-73.
19. Jung, S.H., Kim, M.S., Lee, S.H., Park, H.C., Choi, H.J., Maeng, L., Min, K.O., Kim, J., Park, T.I., Shin, O.R., Kim, T.J., Xu, H., Lee, K.Y., Kim, T.M., Song, S.Y., Lee. C., Chung Y.J., Lee, S.H. (2016). Whole-exome sequencing identifies recurrent AKT1 mutations in sclerosinghemangioma of lung. Proc Natl AcadSci U S A. **113**,(38):10672-7.
20. Bryce, A.H., Egan, J.B., Borad, M.J., et al. (2017). Experience with precision genomics and tumor board, indicates frequent target identification, but barriers to delivery. Oncotarget. **8**,(16):27145-27154.
21. Craig, W. L. (2010). The Akt/PKB Family of Protein Kinases: A Review of Small Molecule Inhibitors and Progress Towards Target Validation: A Current Topics in Medicinal Chemistry, **10**,(4), 458-477.
22. Tang MKS, Yue PYK, Ip PP, et al. (2018) Soluble E-cadherin promotes tumor angiogenesis and localizes to exosome surface. Nat Commun. **9**,(1):2270. Published 2018 Jun 11. doi:10.1038/s41467-018-04695-7
23. Cannon MJ, Ghosh D, Gujja S. (2015) Signaling Circuits and Regulation of Immune Suppression by Ovarian Tumor-Associated Macrophages. Vaccines (Basel).;**3**,(2):448-466. Published 2015 May 29. doi:10.3390/vaccines3020448
24. Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., Zhang, Y. (2015). The I-TASSER Suite: protein structure and function prediction. Nat Methods. **(1):7-8.**
25. Kelly, M. G., Mor, G., Husband, A., O'Malley, D. M., Baker, L., Azodi, M.,Rutherford, T. J. (2011). Phase II evaluation of phenoxodiol in combination with cisplatin or paclitaxel in women with platinum/taxane-refractory/resistant epithelial ovarian, fallopian tube, or primary peritoneal cancers. Int J Gynecol Cancer, **21**,(4):633–639.
26. Kumar, S., Pandey, A.K. (2013). Chemistry and biological activities of flavonoids: an overview. ScientificWorldJournal. 162750:1-16.
27. Oi N, Chen H, Kim MO, Lubet RA, Bode AM, Dong Z. (2012) Taxifolin suppresses UV-induced skin carcinogenesis by targeting EGFR and PI3K. Cancer Prev Res.;**5**,(9):1103–14.
28. Khurana, R., Hate, A. T., Nath, U., Udgaonkar, J. B. (1995). pH dependence of the stability of barstar to chemical and thermal denaturation. Protein Sci. **4**, (6):1133–1144
29. Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res.;**31**,(13):3784-8.
30. Webb, B., Sali, A. (2014). Comparative Protein Structure Modelling Using MODELLER. CurrProtoc Bioinformatics. **47**,:5.6.1-32.

31. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK A programme to check the steriochemical properties of protein. J. Appl. Cryst. **26,** 283-291.
32. C and., Yeates, T.O. (1993). Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. (9):1511-9.
33. Wiederstein, M., Sippl, M.J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 35(Web Server issue):W407-10.
34. Pettersen, E.F., Goddard, T.D, Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E. (2004). UCSF Chimera-a visualization system for exploratory research and analysis. J Comput Chem., **13**,:1605-12.
35. Sayers, E.W., Beck, J., Brister, J.R., Bolton, E.E., Canese, K., Comeau, D.C., Funk, K., Ketter, A., Kim, S., Kimchi, A., Kitts, P.A., Kuznetsov, A., Lathrop, S., Lu, Z., McGarvey, K., Madden, T.L., Murphy, T.D., O'Leary, N., Phan, L., Schneider, V.A., Thibaud-Nissen, F., Trawick, B.W., Pruitt, K.D., Ostell, J. (2020). Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 8;**48**,(D1):D9-D16.
36. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shaw, D. E.; Shelley, M.; Perry, J. K.; Francis, P.; Shenkin, P. S., (2004). "Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy," J. Med. Chem., **47**,:1739–1749
37. Jorgensen, W.L. & Tirado-Rives, J. (1988). The OPLS potential function for proteins. Energy minimizations for crystals of cyclic peptides and crambin. J. Am. Chem. SOC. **l10**, (6), 1657-1666
38. Shivakumar, D., Williams J., Wu, Y., Damm, W., Shelley, J., Sherman W. (2010). Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the OPLS force field. J Chem Theory Comput, **6**:1509-1519.
39. Schrödinger Release 2020-4: Protein Preparation Wizard; Epik, Schrödinger, LLC, New York, NY, 2020.
40. Schrödinger Release 2020-4: SiteMap, Schrödinger, LLC, New York, NY, 2020
41. Schrödinger Release 2020-4: Glide, Schrödinger, LLC, New York, NY, 2020
42. Schrödinger Release 2020-2:Maestro, Schrödinger, LLC, New York, NY, 2020.
43. Lipinski, C.A., Lombardo, F., Dominy, B.W., Feeney, P.J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev. **46,**(1-3):3-26.
44. Shihab, H.A., Gough, J., Mort, M., Cooper, D.N., Day, INM, Gaunt, T.R. (2014). Ranking Non-Synonymous Single Nucleotide Polymorphisms based on Disease Concepts. Human Genomics, **8**,(1):1-11
45. Hess, B., Kutzner, C., Van Der Spoel, D., Lindahl, E. (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. Journal of Chemical Theory and Computation **4**,(3):435–447
46. Schuler, Lukas & Daura, Xavier & van Gunsteren, Wilfred. (2001). An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase. Journal of Computational Chemistry. **22**,. 1205-1218
47. Schuttelkopf, A.W., Van Aalten DM. (2004). PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. ActaCrystallogr D BiolCrystallogr. **60**,(8):1355-63
48. Perera, L, Essmann, U., and Berkowitz, M.L. (1995). A smooth particle mesh Ewald method. J. Chem. Phys. **103**, (19), 8577-8593
49. Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E., Berendsen, H.J. (2005). GROMACS: fast, flexible, and free. J Comput Chem. (16):1701-18.
50. Mamounas, E., Wieand, S., Wolmark, N., Bear, H.D., Atkins, J.N., Song, K., Jones, J., Rockette, H. (1999). Comparative efficacy of adjuvant chemotherapy in patients with Dukes' B versus Dukes' C colon cancer: results from four National Surgical Adjuvant Breast and Bowel Project adjuvant studies (C-01, C-02, C-03, and C-04). J ClinOncol. **17**,(5):1349-55.
51. Rigano, Daniela &Formisano, Carmen &Grassia, Armando &Grassia, Gianluca&Perrone, Angela &Piacente, Sonia &Vuotto, Maria &Senatore, Felice. (2007). Antioxidant Flavonoids and Isoflavonoids from Rhizomes of Iris pseudopumila. Planta medica. **73**,:93-6.
52. Bosetti, C., Gallus, S., La Vecchia, C. (2006). Aspirin and cancer risk: an updated quantitative review to 2005. Cancer Causes Control. **17**,(7):871-88.
53. Gonzalez, F.D., Gamero, E., Garrido, M., Ramírez, R., Moreno, D., Delgado, J., Valdés, E., Barriga, C., Rodríguez, A.B., Paredes, S.D. (2012). Urinary 6-sulfatoxymelatonin and total antioxidant capacity increase after the intake of a grape juice cv. Tempranillo stabilized with HHP. Food Funct. **3**,(1):34-9.
54. Karthikeyan, S., Srinivasan, R., Wani, S. A., &Manoharan, S. (2013). Chemopreventive potential of chrysin in dimethylbenzanthraceneinduced hamster buccal pouch carcinogenesis. Int J NutrPharmacolNeurol Dis., **3**,(4), 46–53
55. Gomeni, R., Goyal, N., Bressolle, F. et al.(2015). A Novel Methodology to Estimate the Treatment Effect in Presence of Highly Variable Placebo Response. Neuropsychopharmacol **40,** 2588–2595
56. Zhong, W., Gallivan, J.P., Zhang, Y., Li, L., Lester, H.A., Dougherty, D.A. (1998). From ab initio quantum mechanics to molecular neurobiology: a cation-pi binding site in the nicotinic receptor. Proc Natl AcadSci U S A. **95,**(21):12088-93.
57. Sussman, J.L., Harel, M., Frolow, F., Oefner, C., Goldman, A., Toker, L., Silman, I. (1991). Atomic structure of acetylcholinesterase from Torpedo californica: a prototypic acetylcholine-binding protein. Science. **23**;253(5022):872-9.
58. Dougherty, D.A. (1996). Cation-pi interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. Science. 271(5246):163-8.
59. Albert, R., Jeong, H., Barabasi, A.L. (2000). Error and attack tolerance of complex networks. Nature. 406 (6794):378-82.
60. Alberg DG and Schreiber SL. (1993). Structure-based design of a cyclophilin– calcineurin bridging ligand. Science**, 262,** 248–250
61. Cheng, S., Niv, M.Y. (2010). Molecular dynamics simulations and elastic network analysis of protein kinase B (Akt/PKB) inactivation. J ChemInf Model. **50,** (9):1602-10.

62. K, M. Yanokura, M. Iida, M. Adachi, K. Nakamura, Y. Nogami, K. Umene, K. Masuda, I. Kisu, H. Nomura, F. Kataoka, E. Tominaga, D. Aoki (2014) Application of MicroRNA in diagnosis and treatment of ovarian CancerBiomed Res. Int., 1-6, 10.1155/2014/232817
63. Stronach E.A., Cunnea P., C. Turner, T. Guney, R. Aiyappa, S. Jeyapalan, C.H. de Sousa, A. Browne, N. Magdy, J.B. Studd, R. Sriraksa, H. Gabra, M. El-Bahrawy (2015) The role of interleukin-8 (IL-8) and IL-8 receptors in platinum response in high grade serous ovarian carcinoma Oncotarget., **6**, 10.18632/oncotarget.3415
64. Kellenberger L.D., J.E. Bruin, J. Greenaway, N.E. Campbell, R.A. Moorehead, A.C. Holloway, J. Petrik (2010), The role of dysregulated glucose metabolism in epithelial ovarian cancerJ. Oncol. 1-13, 10.1155/2010/514310
65. Cho H.J., D.S. Suh, S.H. Moon, Y.J. Song, M.S. Yoon, D.Y. Park, K.U. Choi, Y.K. Kim,K.H. Kim, (2013) Silibinin inhibits tumor growth through downregulation of extracellularsignal-regulated kinase and Akt in vitro and in vivo in human ovarian cancer cells,J. Agric. Food Chem. **61** ,4089–4096.
66. Fan L., Y. Ma, Y. Liu, D. Zheng, G. Huang, (2014) induces cell cycle arrest andapoptosis in ovarian cancer cells, Eur. J. Pharmacol. **743** ,79–88
67. Yi L, Zongyuan Y, Cheng G, Lingyun Z, Guilian Y and Wei G (2014). Quercetin enhances apoptotic effect of tumor necrosis factor-related apoptosis-inducing ligand (TRAIL) in ovarian cancer cells through reactive oxygen species (ROS) mediated CCAAT enhancer-binding protein homologous protein (CHOP)-death receptor 5 pathway. Cancer Sci. **105**,:520–527.
68. Czyz J., Z. Madeja, U. Irmer, W. Korohoda, D.F. Hülser (2005) Flavonoid apigenin inhibits motility and invasiveness of carcinoma cells in vitro, Int. J. Cancer **114**, 12–18, https://doi.org/10.1002/ijc.20620.
69. Manigandan K., D. Manimaran, R.L. Jayaraj, N. Elangovan, V. Dhivya, A. (2015) KaphleTaxifolin curbs NF-κB-mediated Wnt/β-catenin signaling via up-regulating Nrf2 pathway in experimental colon carcinogenesisBiochimie, **119**, pp. 103-112
70. Borovskaya, T.G, et al., (2015) Dihydroquercetin effects on the morphology and antioxidant/prooxidant balance of the prostate in rats with sulpiride-induced benign hyperplasiaBull. Exp. Biol. Med., **158**, pp. 513-516
71. Salah N., N.J. Miller, G. Paganga, L. Tijburg, G.P. Bolwell, C. (1995) Rice-Evans Polyphenolic flavanols as scavengers of aqueous phase radicals and as chain-breaking antioxidantsArch. Biochem. Biophys., **322**, pp. 339-346
72. Razak, S., Afsar, T., Ullah, A. et al. (2018) Taxifolin, a natural flavonoid interacts with cell cycle regulators causes cell cycle arrest and causes tumor regression by activating Wnt/ β -catenin signaling pathway. BMC Cancer **18,** 1043 https://doi.org/10.1186/s12885-018-4959-4
73. Chen X., N. Guu, C. Xue, Li B.R (2018),.Plant flavonoid taxifolin inhibits the growth, migration and invasion of human osteosarcoma cellsMol. Med. Rep., **17**, pp. 3239-3245
74. McGuffin, L.J., Bryson, K., Jones, D.T. (2000). The PSIPRED protein structure prediction server. Bioinformatics. **16**,(4):404-5.
75. Hekkelman, M.L., TeBeek, T.A., Pettifer, S.R., Thorne, D., Attwood, T.K., Vriend, G. (2010). WIWS: a protein structure bioinformatics Web service collection. Nucleic Acids Res. W719-23.
76. Bagaria, A., Jaravine, V., Huang, Y.J., Montelione, G.T., Güntert P. (2012). Protein structure validation by generalized linear model root-mean-square deviation prediction. Protein Sci. (2):229-38.
77. Vidya, N., Vadivukkarasi, B., Manivannan, G., and Anbarasu, K. (2008) Molecular modelling and docking studies of glutamate racemase in Vibrio vulnificus CMCP6. In Silico Biol. **8,** 471–483.
78. Cai C, Liu C, Zhao L, Liu H, Li W, Guan H, Zhao L, Xiao J. (2018) Effects of Taxifolin on Osteoclastogenesis in vitro and in vivo. Front Pharmacol. 12;9:1286. doi: 10.3389/fphar.2018.01286. PMID: 30483128; PMCID: PMC6240596.
79. Brusselmans, Koen; Vrolix, Ruth; Verhoeven, Guido; Swinnen, Johannes V. (2005). "Induction of Cancer Cell Apoptosis by Flavonoids is Associated with Their Ability to Inhibit Fatty Acid Synthase Activity". Journal of Biological Chemistry. **280,** (7): 5636–5645. doi:10.1074/jbc.M408177200. PMID 15533929
80. S.E et al.,(2009) Evaluation of the polyphenol composition and antioxidant activity of African variety of Dacryodes edulis (G.Don) H.J Lam fruit J. Med. Food, **12**, (2009), pp. 1321-1325, 10.1089/jmf.2008.0215
81. Pantavos, A. et al., (2015) Total dietary antioxidant capacity, individual antioxidant intake and breast cancer risk: the Rotterdam study Int. J. Cancer, **136**, pp. 2178-2186, 10.1002/ijc.29249.
82. Mourouti,N.et al., Di et (2015) Breast cancer: a systematic review Int. J. Food Sci. Nutr., **66**, pp. 1-42, 10.3109/09637486.2014.950207.

Figures

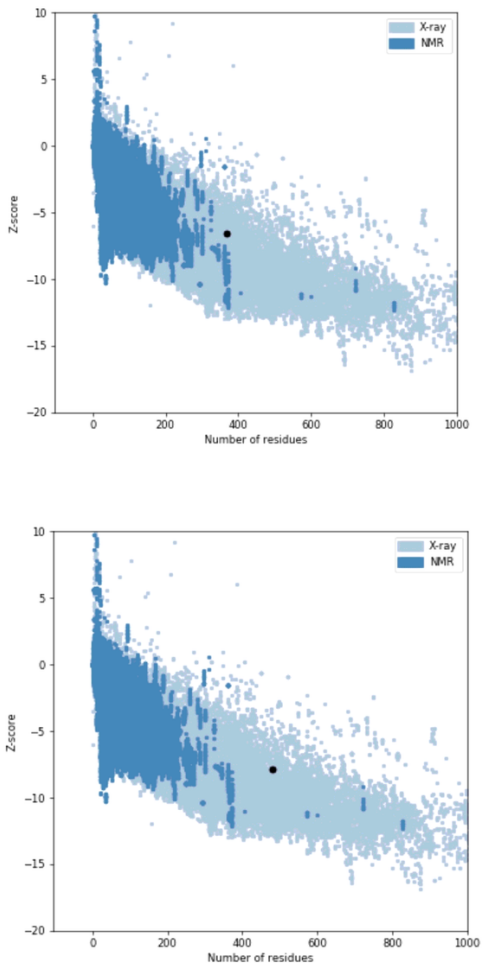

Figure 1

Protein structure analysis (ProSA) of the W80R(mutant) on the left side and AKT1(wild)on right side. (a)Overall quality of W80R model represents Z score of -7.9Kcal/mol(b)Overall quality of the wild protein AKT1 represents Z score of -7.2Kcal/mol.

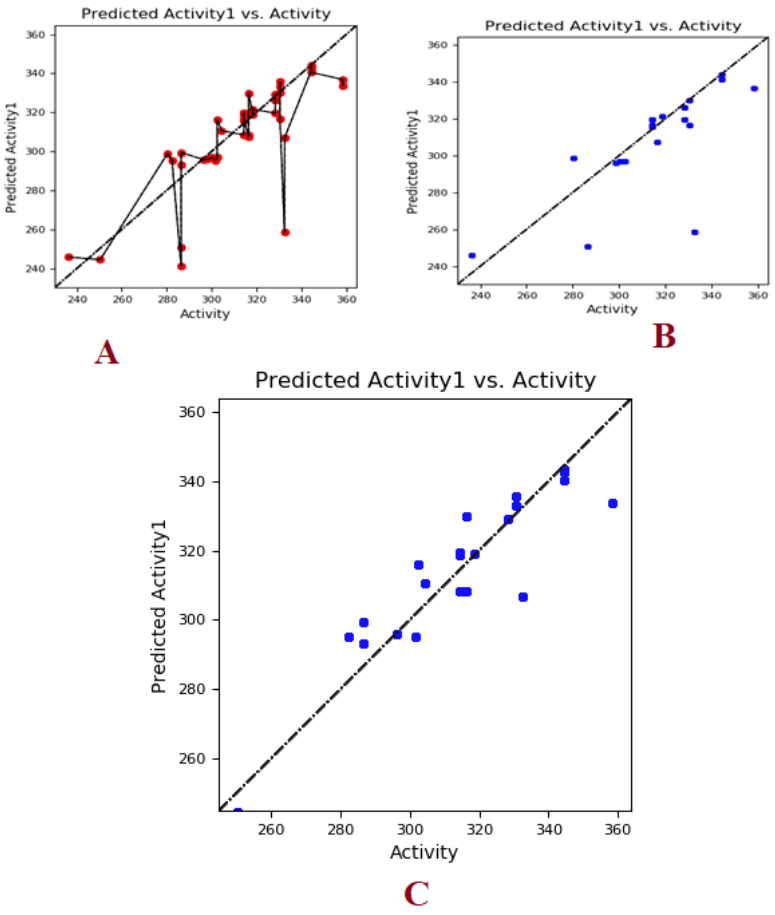

Figure 2

(A) Scatter plot diagram between actual activity and predicted activity showing QSAR results of all molecular descriptors (B) Activity predicted between only training set chemical descriptors(C) Predicted activity between test set of chemical descriptor

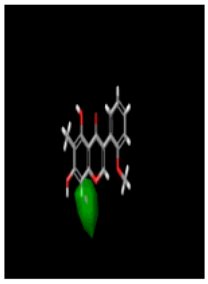

Figure 3

The CoMFA steric field with 2Å in a green hydrophobic region as grid spacing is displayed for the compound CID ID: 5482167.

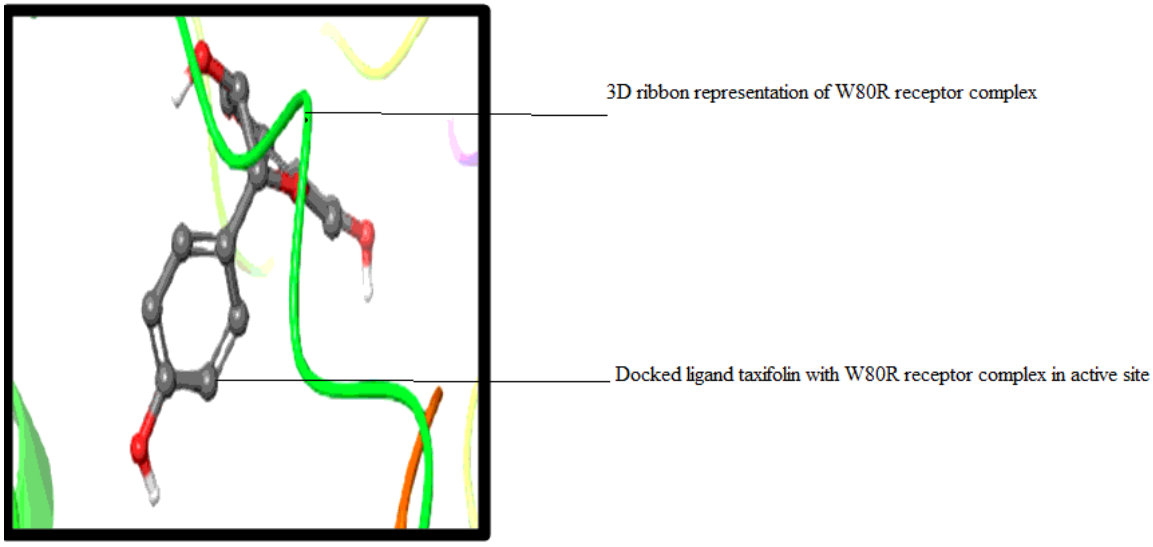

Figure 4

A docked complex of W80R protein in ribbon model with inhibitor CID ID 443637 (Taxifolin) at an active site of binding pocket with XP score -9.63 kcal/mol.

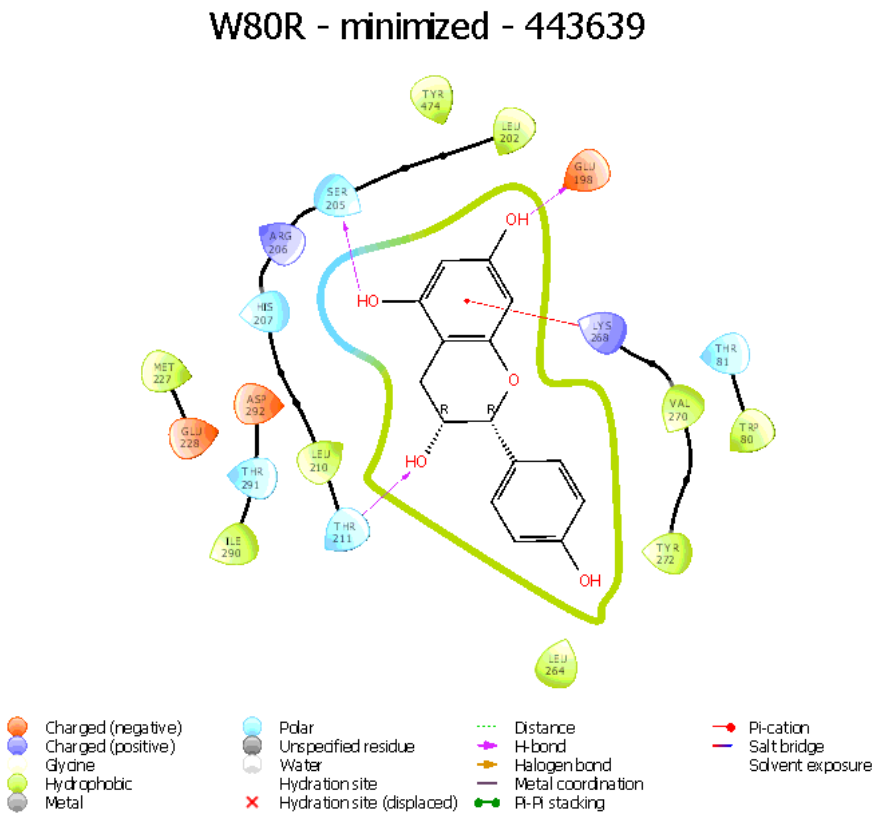

Figure 5

Representation of W80R receptor molecule with CID ID-443637 as a ligand interaction with protein residues SER205, THR211, GLU198 of a hydroxyl group (-OH) and 1 pi-cation interaction with LYS268 with noticeable solvent exposure sites observed at some residue locations with highest Glide XP score of -9.63Kcal/ mol.

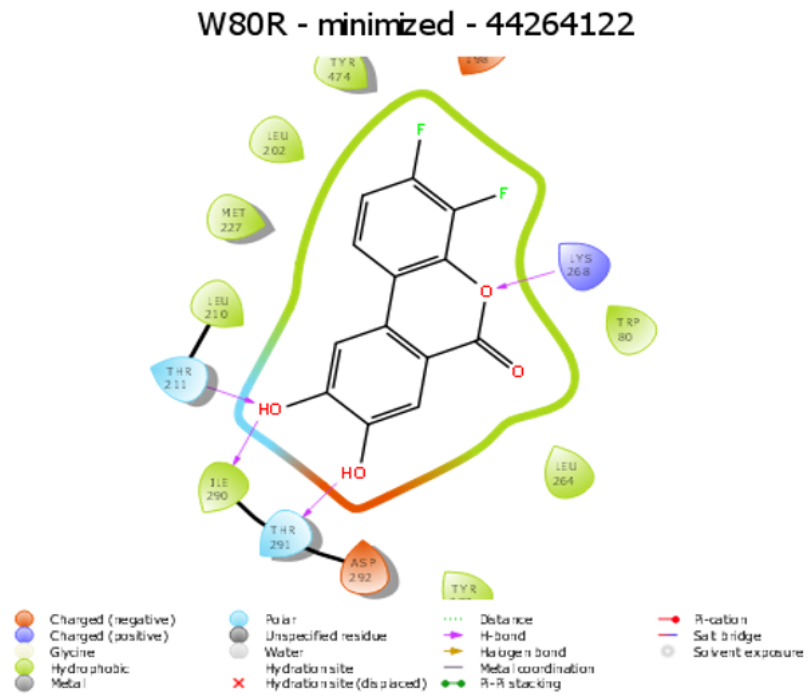

Figure 6

Representation of W80R protein-inhibitor complex of CID 44264122 with 2 hydrogen bonds of a hydroxyl group (-OH) interacting with LYS268, THR291, ILE290, and THR211 and ILE290 and –OH with THR 291 and oxy bond with residue LYS 268 with Glide XP score: -9.43Kcal/mol.

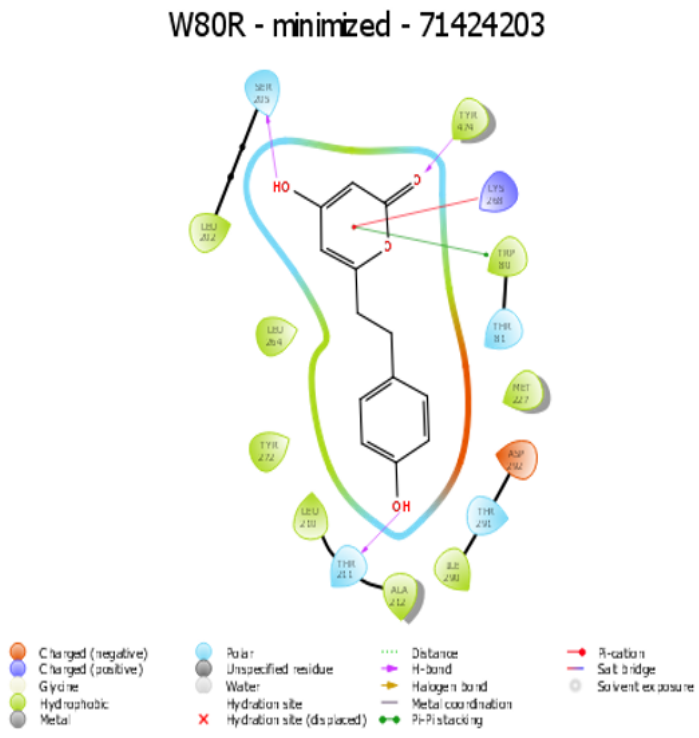

Figure 7

Representation of W80R receptor molecule with inhibitor at the active site showing protein-ligand hydrogen bond interaction with residues as TYR474, SER215, THR211, with 1 pi-pi interaction at TRP80 residue, and 1 pi-cationic interaction bonding with LYS265 with G score -9.36Kcal/mol.

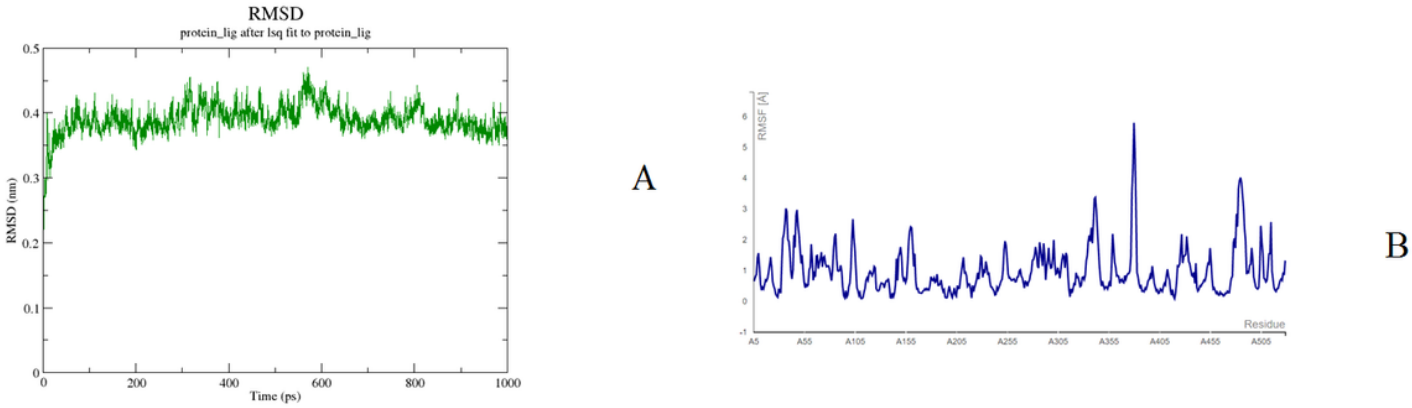

Figure 8

(a)Root mean square deviation (RMSD) of the C-alpha backbone of the W80R protein complex and ligand taxifolin (X-axis time scale in ps and Y-axis in RMSD in nm). (b)Root mean square fluctuation (RMSF) for C-alpha backbone atom of a W80R protein complex with ligand taxifolin (X-axis shows amino acid residue number and Y-axis shows RMSF in nm).

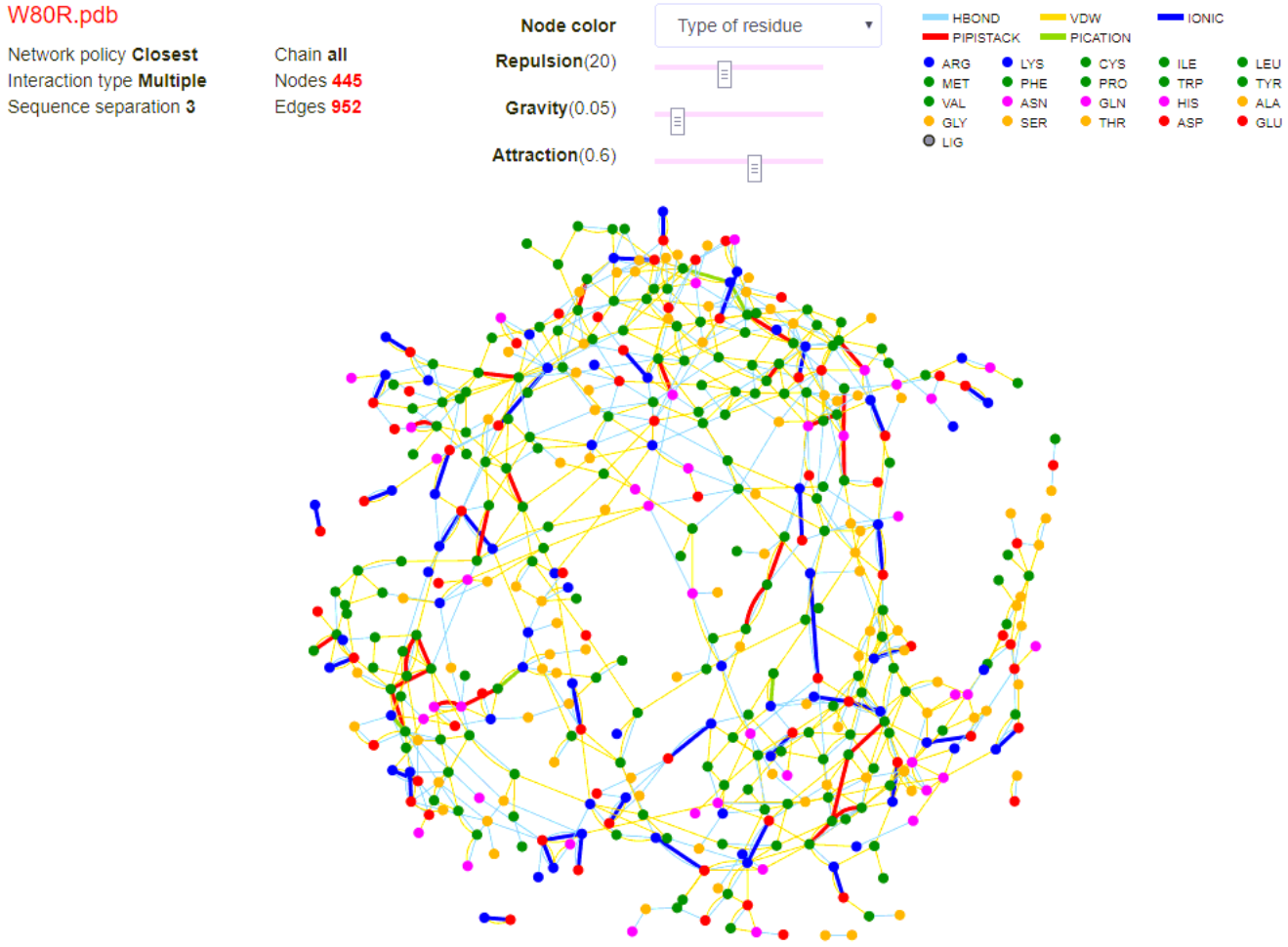

Figure 9

Visualisation of residual network of W80R protein complex.