

Conserved Water Molecule-dependent Docking Strategy and Atom-Based 3D QSAR Studies to Design Heat Shock Protein 90 Inhibitors

S. D. GUPTA^{1,2*}, C. V. S. SUBRAHMANYAM¹, N. L. GOWRISHANKAR³ AND N. M. RAGHAVENDRA¹

¹Department of Pharmaceutical Chemistry, Gokaraju Rangaraju College of Pharmacy, Hyderabad-500 090, ²R&D Centre, Department of Pharmaceutical Sciences, Jawaharlal Nehru Technological University, Hyderabad-500 085, ³Prime College of Pharmacy, Erattayal, Palakkad-678 551, India

Gupta *et al.*: Atom-Based 3D-QSAR Design of Heat Shock Protein 90 Inhibitors

In this study a methodology was described to recognize the conserved water molecules required for efficient binding of ligands with heat shock protein. Subsequently, a substantiated procedure for the determination of an effective docking methodology was generated using various programs of Schrodinger and SYBYL softwares. In order to identify the essential structural features responsible for heat shock protein inhibitions, an atom based 3D-quantitative structure-activity relationship model with an excellent surmising ability in both external and internal validation was established. The results of 3D-quantitative structure-activity relationship and docking studies correlated well with each other. A combined analysis of docking results and quantitative structure-activity relationship analysis indicates: The key amino acids and water molecules in the active pocket of heat shock protein. The important structural requirement in ligands that will lead to enhanced biological activity. The results provide a set of simple and effective guidelines for rationally designing new heat shock protein inhibitors prior to their synthesis.

Key words: Heat shock protein, cancer, docking, conserved water molecules, atom based 3D-QSAR, Sybyl, Schrodinger

The use of existing anticancer drugs is hampered due to development of resistance against them^[1]. Resistance is attributed to rapid mutation of the target protein encoding genes which in turn alters the morphology of the active site^[1,2]. Additionally, the wide-ranging side effects associated with the present cancer chemotherapeutic agents further restricts their usage. These adverse effects, such as gastrointestinal distress, hair loss and neutropenia were attributed to the damage of normal cells that divide rapidly similar to the cancer cells^[3,4]. Therefore, targeted cancer therapy has gained momentum in recent years. The most promising target for the development of novel, potent and safer anticancer agents needs to satisfy the following two criteria, multiple oncogenic proteins along with the ones promoting resistance must be dependent upon the target for maintaining their structure and function. The target protein should be over-expressed in cancer cells as compared to the healthy ones.

Heat shock protein 90 (Hsp90) is one such exciting anticancer target, which is an ATP-dependent

molecular chaperone with an average molecular weight of 90-kDa^[5]. There are two major types of cytosolic isoforms of Hsp90 in eukaryotes, Hsp90 α and Hsp90 β . Hsp90 α is the inducible form, which is produced during stress conditions of the cells, such as elevated temperature, exposure to UV radiation, infection and inflammation^[6]. Hsp90 β is the constitutive form, which constitutes 1-2 % of total cellular protein under non-stress conditions. Moreover Hsp90 α is over expressed in cancer cells as an activated complex with co-chaperone proteins Hsp40, Hsp70, Cdc37, p23, Tpr2, Hip and Hop, while the β form of normal cells resides in a free state^[7]. Therefore, it gave rise to the hypothesis that inhibitors against Hsp90 would not affect normal cells and will be devoid of toxic side effects associated

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms

***Address for correspondence**

E-mail: sayanduttagupta1@rediffmail.com

Accepted 18 January 2020

Revised 19 Decemehr 2019

Received 02 July 2019

Indian J Pharm Sci 2020;82(2):341-355

with cancer chemotherapy. The expression of Hsp90 is associated with cancers of breast, pancreas, blood, colon, colorectal, kidney, brain and prostate^[8]. It helps in the folding, maturation, stability, trafficking, intracellular disposition and proteolytic turnover of proteins responsible for cancer cell development and survival. Such proteins are commonly referred to as the client proteins. These client proteins appear to be addicted to Hsp90 for survival. Till today, more than 300 client proteins have been identified^[5], some of which are dangerous oncogenes like Akt, Raf-1, cdk4 and Src. Few are well established cancer targets, Her-2 (trastuzumab), Bcr-Abl (imatinib), estrogen receptor (tamoxifen), androgen receptor (bicalutamide) and the vascular growth factor endothelial receptor (sunitinib)^[9]. Some mutant proteins like mutant p53 and imatinib-resistant Bcr-Abl are also dependent on Hsp90 for preserving their structure and function^[10,11]. It was also revealed that many proteins responsible for mutation like PDGFR β , COT, IGFR1, CRAF, ARAF, S6, cyclin D1, and AKT are also clients of Hsp90^[10,12]. Hence inhibition of Hsp90 will lead to arrest of multiple proteins involved in the development of cancer and resistance. Thus, the two major drawbacks of cancer therapy, drug resistance and toxic side effects could be overcome by developing agents against Hsp90.

Hsp90 has three functional domains, first of which is the N-terminal domain of ~ 25 kDa that fulfills all the structural requirement for ATP binding. Structural alterations facilitated by the hydrolysis of ATP to ADP in this domain are thought to play an essential role in the chaperoning activity of the protein. Blocking of ATPase activity hampers the repairing function of Hsp90. Hence, majority of Hsp90 inhibitors competitively bind

to this site of the protein. Second functional domain is C-terminal domain of ~ 12 kDa, which contain binding sites for co-chaperones and the third domain is the protein binding domain of ~ 12 kDa whose contact with the N-terminal domain is necessary for the latter's proper functioning^[13]. Further, small angle X-ray scattering (SAXS), TROSY NMR and electron microscopy (EM) studies have shown that Hsp90 is highly flexible as a single unit along with the sequences connecting its domains. However, individual domains of the protein are less flexible compared to the entire protein. It was experimentally proven that small molecule inhibitors, ATP, co-chaperones and client proteins along with various post-translational modifications regulate the flexibility of the Hsp90 chaperone^[14,15].

One of the earliest Hsp90 inhibitors discovered were natural products radicicol (1a) and geldanamycin (1b) (fig. 1)^[16]. Radicicol, obtained from the fungus *Monosporium bonorden* was found to be inactive *in vivo*^[17]. Geldanamycin is a type of ansamycin isolated from *Streptomyces hygroscopicus*. The highly reactive 17-methoxyl group of geldanamycin and its high toxic profile hindered further development of geldanamycin as such. Therefore, the 17-allylgeldanamycin, 17-AAG (tanespimycin, 2a (fig. 2) was developed, which reached phase I clinical trials but eventually abandoned due to severe hepatotoxicity, poor solubility, limited bioavailability and extensive metabolism. The other analogues of geldanamycin (reduced form of 17-AAG, retaspimycin, 2b and 17-DMAG, alvespimycin, 2c) shown in fig. 2 developed to increase water solubility and bioavailability were also rejected after reaching phase-II clinical study due to unsatisfactory safety profile^[18,19]. It is believed that the

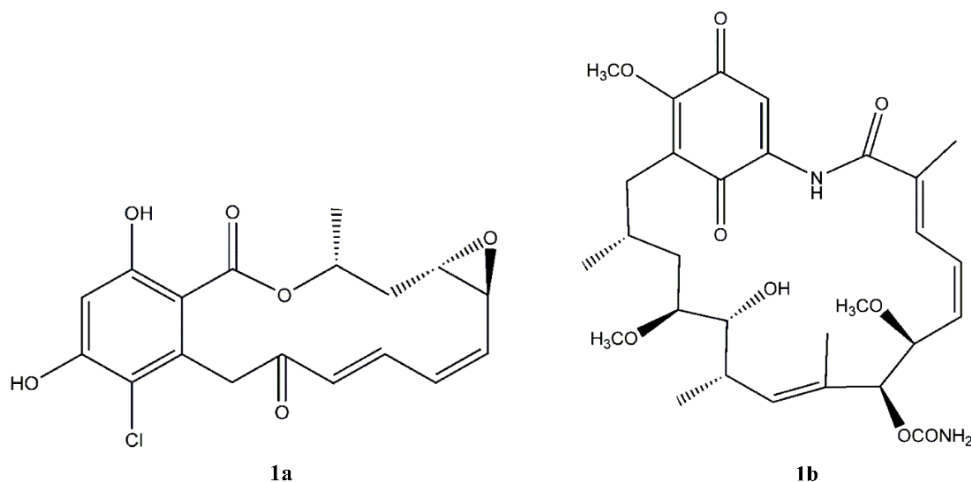


Fig. 1: First Hsp90 inhibitors discovered
1a. radicicol and 1b. geldanamycin

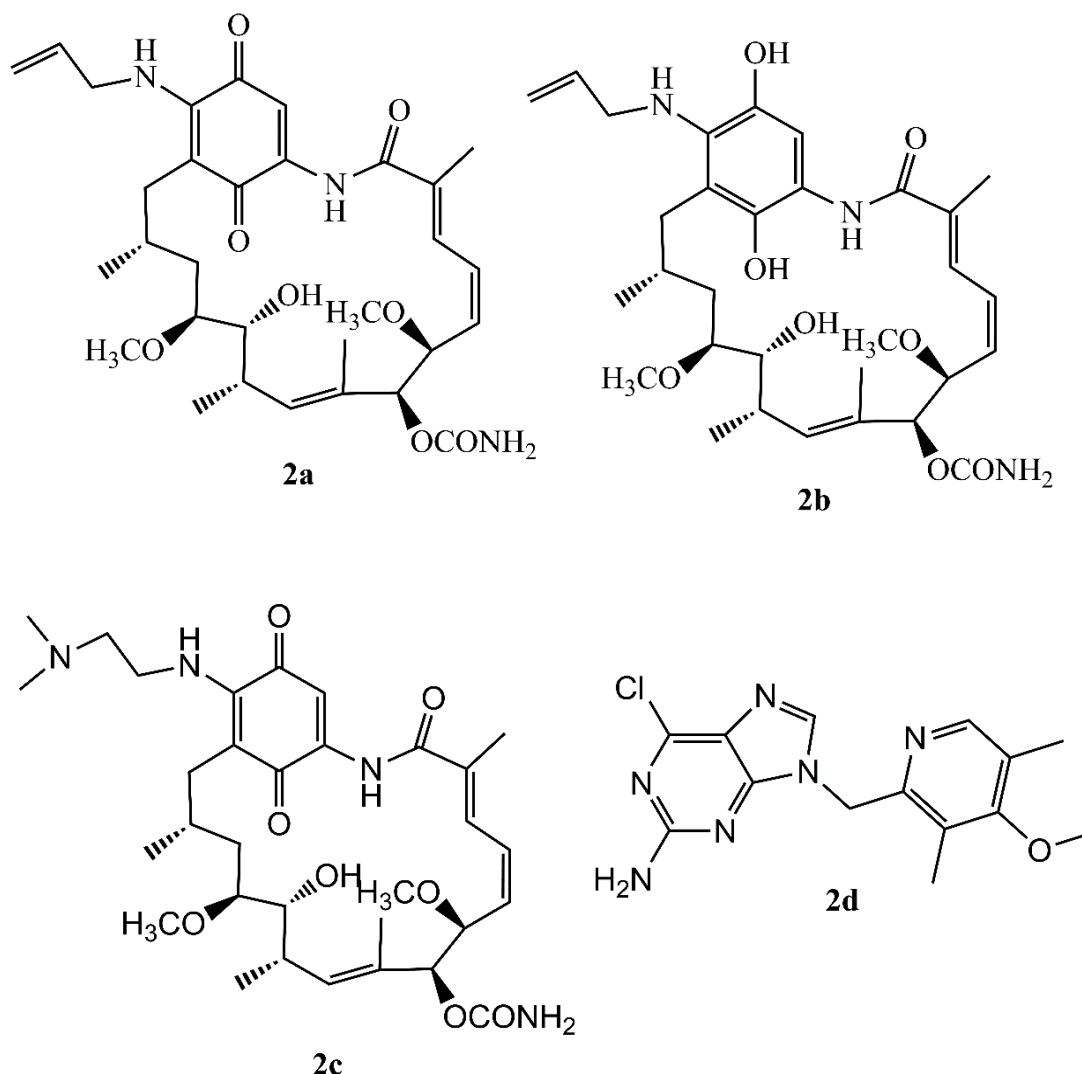


Fig. 2: Hsp90 inhibitors that reached various phases of clinical trials
2a. Tanespimycin, 2b. retaspimycin, 2c. alvespimycin and 2d. CNF2024

benzoquinone ring and the chemical complexity of the anamycins contributed to its off target toxicity. The purine derivative, BIlB021 (CNF2024, 2d, fig. 2) was one of the synthetic Hsp90 inhibitor to enter clinical trials in 2006. It reached up to clinical phase-II trials but did not exhibit any promising result^[20]. Till date 17 candidates are in clinical trials either alone or in combination with other anticancer agents, but none of those has reached approval stage for a market launch^[21].

Water plays a crucial role in determining the structure and dynamics of biomolecules. In some cases, hydrogen bond forming groups on the protein surface does not form any H-bond with the ligands due to more distance ($> 5\text{\AA}$). This distance can be bridged by some water molecules commonly referred to as the conserved water molecules. These water molecules form hydrogen bond with the protein as well as with the ligands. Thus, a protein-water-ligand complex is formed, which results

in more effective binding. The role of conserved water molecules for more effective ligand-protein binding through hydrogen bonding is already established for Hsp90 as well as for many other proteins^[22,23]. Moreover, the high rate of failure from designing a drug molecule to its marketing stage warrants the development and validation of a proper docking method by retaining these water molecules.

Computational drug design is one of the powerful techniques used to understand the key features of molecules that turn them active against certain targets. Several reports are available in this area, which are focused on Hsp90 inhibitor development^[24-26]. Although these studies have disclosed some of the key properties that turn few compounds into good Hsp90 inhibitors, there is no systematic study that has considered all the water molecules of the protein. Furthermore, no significant attempts have been made to compare various

docking methodologies and correlate all of these with a QSAR model. In view of the above, it appeared of interest to develop a conserved water molecule-dependent validated docking algorithm, which would aid in the discovery of novel Hsp90 inhibitors. The findings of the docking analysis were further correlated with a 3D-QSAR model based on the structural features of known ligands.

MATERIALS AND METHODS

Schrodinger docking, protein and grid preparation:

Docking was carried out with Schrodinger docking software (Maestro, 9.1 versions) installed on Dell Precision T-1500 workstation (Intel(R) Core(TM) i7 CPU 860 @ 2.80GHz 2.79GHz; 12.0 GB Ram, 1 TB Hard disk). The coordinates of the human Hsp90 in complex with an inhibitor [(2,4-dihydroxyphenyl) (S)-2-(pyridin-2-yl)pyrrolidin-1-yl)methanone were obtained from Protein Data Bank (PDB, PDB ID: 3EKR, resolution of 2 Å)^[27,28]. The protein file was prepared for docking by filling missing side chains and loops using Prime program of the software, deleting the analogous B chain, removing the undesirable phosphate and water molecules, followed by addition of polar hydrogens. The protein with ligand having the lowest ionization penalty state was optimized for rendering visual inspection of hydrogen bonds. Further the protein–ligand complex was submitted to restrained molecular mechanics refinement using the optimized potential for liquid simulations (OPLS) 2005 force field. The grid (virtual active site) was generated by extracting the native ligand from its binding site. The GLIDE program's Grid Generation application was utilized for the above grid generation process^[29].

Ligand preparation and docking:

The ligands were optimized for docking employing the ligprep program, which creates relevant correctly protonated 3D forms of the molecule at pH 7±2 by taking into consideration the tautomeric forms and chirality of the molecules^[30]. The ligand molecules were then docked with standard precision (SP) and extra precision (XP) approaches using the GLIDE program of the software. The XP mode of scoring includes special recognition terms for hydrophobic contacts, improved hydrogen bonding, pi-pi/pi-cation stacking and detection of buried polar groups. These extra features are absent in the SP mode of GLIDE^[29].

The Schrodinger's Glide program computes the binding efficiency of the molecules in terms of

Glide Score (G Score). The G Score is calculated as follows, $G\ Score = 0.065 \times vdW + 0.130 \times Coul + Lipo + Hbond + Metal + BuryP + RotB + Site$, where, vdW =van der Waals energy, $Coul$ =coulomb energy, $Lipo$ =lipophilic term derived from hydrophobic grid potential, $H\ bond$ =hydrogen bonding term, $Metal$ =metal binding term, $BuryP$ =penalty for buried polar groups, $RotB$ =penalty for freezing rotatable bonds and $Site$ =polar interactions in the active site.

SYBYL docking and protein preparation:

The docking was performed using SYBYL X-1.2 version software installed on Dell Precision T-1500 workstation (Intel(R) Core(TM) i7 CPU 860 @ 2.80GHz 2.79GHz; 12.0 GB Ram, 1 TB Hard disk). The same co-crystal structure of the protein and ligand as utilized in Schrodinger docking was obtained from PDB^[27]. The protein was processed for binding studies by deleting the B chain, removing the undesirable phosphate/water molecules and addition of polar hydrogen. The hydrogens of the protein and ligands were optimized by AMBER7 FF99 force field. The protomol (virtual active site) was developed from hydrogen-bearing protein mol2 file by keeping a threshold factor that determined the extent to which the protomol penetrated the protein, of 0.5 and a bloat used to dilate the protomol and included the neighbouring cavities of 0 Å^[31,32].

Ligand preparation and docking:

The ligands were first drawn in Chem draw, saved as mol files and then converted into SD file format using Schrodinger software. The structures were then cleaned by sanitize utility involving filling valencies, removing duplicates and producing only one molecule per input structure. It was then followed by a general clean-up, which filtered out compounds containing undesirable substructures, metals, or isotopes and compounds not obeying Lipinski's rule of five. A single 3D conformation for each input structure was generated using the Concord program (random perturbation of the position and conformation of the supplied ligand followed by bump relaxation)^[33]. The SYBYL docking software comprises of 4 different methods for the purpose of docking. The SYBYL Surflex dock method docks compounds without preliminary screening of the ligands. The SYBYL Surflex dock screen method takes into account 3 poses per ligand before docking. The SYBYL Surflex dock Geom and Geom X algorithm computes docking score by considering 20 poses per ligand. However, the Geom X method is

more exhaustive because it takes in to consideration additional 6 conformations whereas the Geom method takes only 4. Moreover, the Geom X mode utilizes a superior alignment method than that of Geom^[34]. The binding affinity of the ligands is predicted by the software in terms of total score, which is expressed as $-\log K_d$, where K_d is binding constant. A high value of total score indicates good protein-ligand binding.

Atom-based 3D-QSAR studies, data sets:

Phase 3.0 of Schrodinger Maestro 9.1 software version was utilized to establish the 3D-QSAR models for Hsp90 inhibitors. The experimental information of 33 molecules involved in this study was collected from the literature. The inhibitory concentrations of the compounds (IC_{50}) were transformed into respective inhibitory activity, pIC_{50} ($-\log IC_{50}$) values which were used as dependent variables in the 3D-QSAR study. The 2D structure was built using Chem draw software and saved as mol files. The geometry of the molecules was converted into 3D form for the purpose of QSAR studies by utilizing the LigPrep module of the software.

Molecular alignment:

In the 3D-QSAR studies, ligand-based molecular alignment was employed to obtain an authentic and dependable model. The most potent^[33] was chosen as the reference molecule to alienate with rest of the training and test set compounds. The above process was aided by flexible ligand alignment tool of the software. The common structure of the ligands is shown in fig. 3 and the alienated molecules are depicted in fig. 4.

3D-QSAR analysis:

Atom-based 3D-QSAR analysis was executed based on the molecular alignment as outlined above. The test set molecules were chosen by considering that they represent a wide range of ATPase activities (0.02-200 μ M) similar to that of the training set. Additionally, the training set was utilized to generate the 3D-QSAR paradigm by partial least square (PLS) method. Subsequently, the test set was engaged in verifying the quality of the models using leave-one-out (LOO) method. Sometimes internal cross-validation is not sufficient to measure the accuracy of a model in predicting the activities of new compounds^[35]. Hence, external validation with the help of Phase tool of Schrodinger software was performed which utilizes a true test set (radicicol, geldanamycin, tanespimycin, retaspimycin, alvespimycin and purine analogue B11021) whose chemical structures and

biological activities were not taken into consideration in the developed QSAR paradigm.

RESULTS AND DISCUSSION

The most comprehensive methods in each software, i.e. Schrodinger's Glide XP and SYBYL's Surflex dock Geom X were utilized to identify the water molecules vital for effective ligand protein binding. The above objective was achieved by including all the water molecules and then running the docking engine. It was observed that the compounds employed for verifying docking methodologies formed hydrogen bonds with water molecule 902, 903, 981 and 1026 as per Schrodinger XP method and with water number 902, 903, 981 according to SYBYL Geom X method. Hence these four water molecules (902, 903, 981 and 1026) were considered important for computing docking scores.

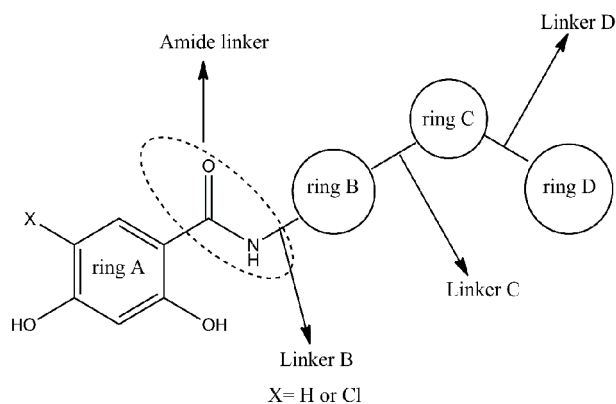


Fig. 3: Skeleton structure for analogs utilized in deriving the QSAR model

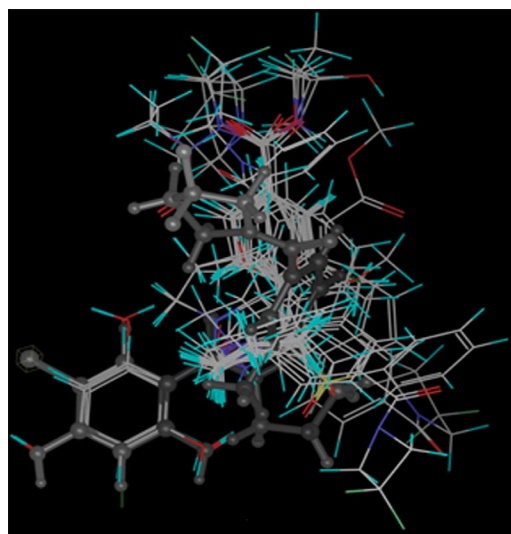


Fig. 4: Alignment of compounds represented as line model Alignment of compounds represented as line model with respect to the most potent molecule 33, which is highlighted in white ball and stick

The intensive comparison of the docking scores calculated by retaining the 4 prominent conserved water molecules (902, 903, 981 and 1026) were carried out among Schrodinger Glide XP and SP methodologies, SYBYL Surflex dock, Surflex dock Screen, Surflex dock Geom and Surflex dock Geom X methods. The 12 most potent compounds (IC_{50} values of 0.02) as determined by cell based Akt-Lum assay^[27] were docked with all the other less potent ligands using various docking methods. The accuracy of the methodologies was validated by evaluating whether the 12 most potent compounds are the same as per docking score rankings. The conclusions of the comparison are summarized in Tables 1, 2 and fig. 5. From the aforementioned tables and figure it is evident that the SYBYL Geom

and Geom X mode of scoring yielded enrichments superior to other methods. The docking protocol was further validated by comparing the software generated ligand Hsp90 binding interactions with the previously described crystallographic analysis. In case of SYBYL Geom X method, it was observed that the 2'-hydroxyl group of the native ligand, (2,4-dihydroxyphenyl)(R)-2-(pyridin-2-yl)pyrrolidin-1-yl)methanone interacted with Asp 93 and water molecule 902 via hydrogen bond. The 4'-hydroxyl group formed hydrogen bond interaction with amino acid Asn 51 and water molecule 903. The other prominent hydrogen bond contacts detected were between the carbonyl oxygen and amino acid Thr 184. Further analysis indicated that the pyrrole moiety occupied a hydrophobic area formed

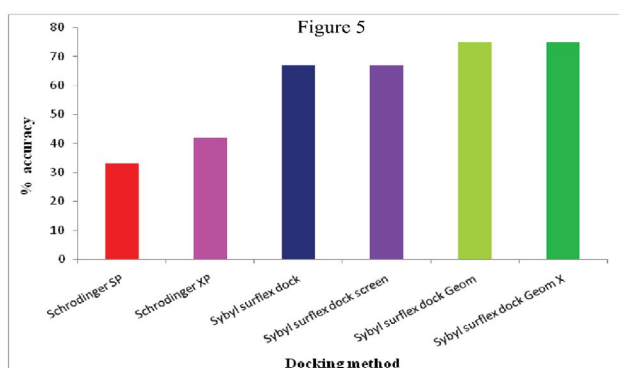
TABLE 1: DOCKING RESULTS VS IC_{50} (HSP90 INHIBITION ASSAY) VALUES OF REPORTED COMPOUNDS

| Ligand Code | I | II | III | IV | V | VI | IC_{50}^* (μ M) |
|-------------|-------|--------|------|------|------|-------|------------------------|
| 1 | -8.10 | -10.56 | 6.15 | 6.15 | 6.22 | 6.27 | 0.40 |
| 2 | -7.99 | -11.38 | 6.12 | 6.12 | 6.12 | 6.34 | 1.70 |
| 3 | -8.33 | -12.07 | 3.67 | 3.67 | 5.32 | 4.28 | 200 |
| 4 | -7.01 | -11.56 | 4.01 | 3.57 | 6.55 | 6.49 | 74.0 |
| 5 | -8.69 | -11.71 | 7.25 | 7.25 | 7.11 | 7.26 | 3.20 |
| 6 | -9.19 | -12.21 | 5.29 | 5.29 | 7.38 | 8.36 | 1.40 |
| 7 | -8.65 | -11.66 | 4.62 | 4.62 | 5.06 | 5.02 | 0.06 |
| 8 | -8.35 | -13.75 | 8.33 | 8.33 | 9.19 | 9.30 | 0.50 |
| 9 | -9.33 | -12.77 | 6.86 | 6.86 | 7.12 | 9.12 | 0.30 |
| 10 | -9.16 | -12.77 | 8.95 | 8.60 | 8.95 | 9.00 | 0.10 |
| 11 | -9.01 | -12.82 | 7.40 | 7.40 | 8.00 | 8.20 | 0.08 |
| 12 | -9.27 | -8.96 | 8.31 | 8.31 | 9.12 | 9.13 | 0.40 |
| 13 | -9.27 | -12.60 | 9.04 | 4.56 | 7.85 | 8.14 | 2.40 |
| 14 | -9.27 | -12.53 | 8.87 | 4.56 | 8.93 | 9.86 | 0.06 |
| 15 | -8.94 | -12.60 | 4.56 | 8.88 | 7.85 | 8.14 | 5.60 |
| 16 | -8.51 | -12.78 | 8.00 | 8.00 | 9.72 | 8.95 | 0.02 |
| 17 | -7.42 | -8.94 | 7.49 | 7.49 | 8.01 | 7.71 | 0.30 |
| 18 | -6.89 | -12.44 | 6.11 | 6.11 | 7.72 | 8.65 | 1.40 |
| 19 | -8.57 | -12.64 | 8.34 | 8.34 | 9.12 | 9.58 | 0.02 |
| 20 | -6.67 | -12.66 | 9.40 | 9.40 | 9.87 | 10.08 | 0.02 |
| 21 | -6.60 | -12.34 | 7.92 | 7.92 | 8.41 | 8.95 | 0.04 |
| 22 | -8.99 | -12.87 | 8.88 | 8.95 | 9.46 | 9.61 | 0.02 |
| 23 | -7.35 | -12.88 | 8.15 | 8.15 | 8.06 | 8.36 | 0.20 |
| 24 | -6.60 | -12.11 | 9.64 | 9.64 | 9.97 | 10.19 | 0.02 |
| 25 | -6.50 | -12.04 | 9.84 | 9.84 | 9.97 | 10.37 | 0.02 |
| 26 | -6.89 | -13.76 | 8.88 | 8.88 | 9.98 | 10.56 | 0.02 |
| 27 | -5.65 | -8.94 | 9.47 | 9.47 | 9.97 | 10.14 | 0.02 |
| 28 | -8.72 | -8.94 | 8.88 | 9.04 | 9.78 | 9.76 | 0.02 |
| 29 | -8.91 | -11.51 | 9.00 | 9.00 | 9.87 | 9.77 | 0.02 |
| 30 | -5.56 | -12.04 | 8.60 | 9.04 | 9.38 | 9.61 | 1.40 |
| 31 | -8.71 | -8.31 | 4.56 | 4.56 | 7.48 | 8.14 | 0.02 |
| 32 | -8.91 | -9.17 | 3.67 | 8.88 | 9.87 | 8.79 | 0.30 |
| 33 | -8.90 | -11.43 | 4.56 | 4.01 | 7.85 | 8.31 | 0.02 |

I. Glide SP docking score. II. Glide XP docking score. III. Total score as per SYBYL Dock method. IV. Total score as per SYBYL Dock screen method. V. Total score as per SYBYL Geom method. VI. Total score as per SYBYL Geom X method. *As per cell-based Akt-lum Hsp90 inhibition assay

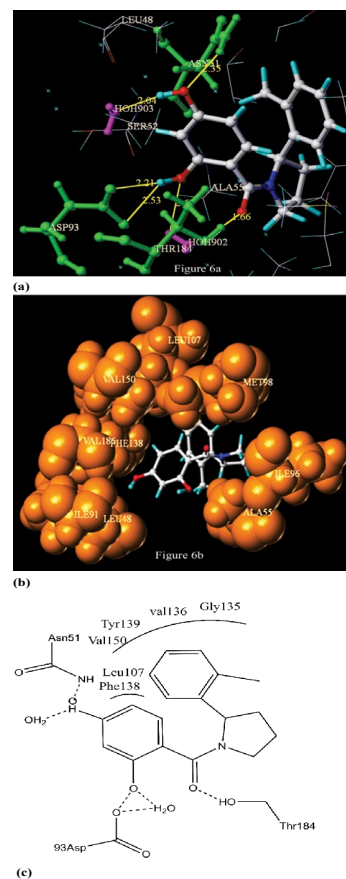
TABLE 2: VALIDATION OF DOCKING SCORES BY DIFFERENT METHODS

| Ligand code | Glide SP rank (25)* | Glide XP rank (42)* | Surflex dock rank (67)* | Surflex dock screen rank (67)* | Surflex dock Geom rank (75)* | Surflex dock Geom X rank (75)* |
|-------------|------------------------|------------------------|----------------------------|-----------------------------------|---------------------------------|-----------------------------------|
| 16 | 18 | 6 | 17 | 17 | 9 | 16 |
| 19 | 17 | 10 | 13 | 13 | 14 | 11 |
| 20 | 27 | 9 | 4 | 4 | 5 | 5 |
| 22 | 8 | 4 | 8 | 8 | 10 | 9 |
| 24 | 28 | 17 | 2 | 2 | 3 | 3 |
| 25 | 30 | 20 | 1 | 1 | 2 | 2 |
| 26 | 25 | 1 | 9 | 9 | 1 | 1 |
| 27 | 31 | 31 | 3 | 3 | 4 | 4 |
| 28 | 13 | 32 | 10 | 5 | 8 | 8 |
| 29 | 10 | 24 | 6 | 7 | 7 | 7 |
| 31 | 14 | 33 | 28 | 30 | 25 | 25 |
| 33 | 12 | 25 | 29 | 31 | 22 | 22 |

**Fig. 5: Percent accuracy of different docking methods employed**

by amino acid Met 98, Ile 96, Ala 55 and Lys 58. The other significant hydrophobic affinity was observed between phenyl ring C and amino acid Tyr, Val 136 and Gly 135. The amino acid residue Leu 107, Val 150 and Gly 135 formed a hydrophobic region where the resorcinol moiety was embedded (fig. 6a and 6b). These interactions were found matching with the reported crystallographic results (fig. 6c)^[27]. In other methods of docking, several important hydrogen bonding and hydrophobic interactions were not revealed. This was particularly prominent with SP and XP method of Schrodinger docking where the output for hydrophobic interaction for all the compounds was calculated as zero. The above discussion indicated that the Geom X of SYBYL is superior to all the other methods.

In this atom-based 3D-QSAR study, the compounds were split into a training set of 21 and a test set of 12 molecules. The most potent compound (33) was chosen as a reference to sketch the rest of the molecules. Atom-based QSAR models were generated by applying PLS regression to a broad set of binary-valued variables that encrypted whether or not atoms of the ligands occupy several cube-shaped part of space. The veracity

**Fig. 6: Ligand interactions with Hsp90 protein (PDB ID:3EKR)**

(a). Hydrogen bonding of ligand (ball and stick model) and Hsp90 protein as per SYBYL surflex X Geom method. The amino acids and the water molecules forming hydrogen bonds are highlighted in green and violet color respectively. The yellow color lines indicate the hydrogen bonds. (b). Hydrophobic interaction of ligand (ball and stick model) and Hsp90 protein as per SYBYL surflex X Geom X method. The orange colored spheres denote hydrophobic amino acids. (c). Hydrogen bonding indicated by dotted lines and hydrophobic interactions designated by semicircle arch of native ligand with Hsp90 protein

of the models improved with the increasing number of PLS factors until over-fitting starts to occur. In this study, the best QSAR model with 5 PLS factors was considered and the QSAR visualization was shown by cubic 3D grids mapped by atoms of the ligands.

The statistical specifications of the developed QSAR model were as follows: The training group achieved R^2 (coefficient of determination, i.e. percent variance accounted by the model in the observed activity data) of 0.9869 with an SD (standard deviation of the regression) of 0.1472. The test batch obtained Q^2 (value of Q^2 for the predicted activities) of 0.7346 with an RMSE (ROOT-mean-square error in the test prediction) of 0.43 and Pearson-R (Pearson R value for the correlation between the predicted and observed activity for the test set) of 0.8697. The constant F (ratio of the model variance to the observed activity variance) of 286.69 and p value (indication of degree of confidence) of 3.35e-017 further substantiates the QSAR protocol. The stability of the model predictions

to changes in the training set composition was found to be 0.738. The plots for predicted vs. actual pIC_{50} for training and test set are shown in figs 7a and 7b. The original versus forecasted activities of the training and test set molecules are listed in Table 3. The estimated activity of the molecules used for external validation (Table 4) was similar to that of their actual values. This particular aspect further exemplifies the perfectness of the derived QSAR model^[36-38].

The 3D-QSAR visualization is generated as cubes by Phase, in which the green cubes indicate the positive structural features and the red ones denote the negative structural elements responsible for the biological activity spectrum. The 3 parameters that were utilized to evaluate the activity profile of the compounds through atom-based 3D-QSAR model are hydrogen bond donor, hydrophobic/non-polar and hydrophobicity. Fig. 8 indicates the areas where hydrogen bond donors in the ligands increase or decrease ATPase activity. The green cubes near the 2-hydroxyl group of the resorcinol moiety (ring A) for all the molecules suggested that any H-bond donor substituent in these 2 positions is favorable for activity. Hence, it is proposed that the resorcinol group could be retained in the future design of Hsp90 inhibitors. Additionally, it was revealed that any H bond group at C-5 of ring A, C-2 of ring B (pyrrolidine) and C4 of ring C (phenyl) is detrimental for Hsp90 inhibition.

The favorable and unfavorable cubes for hydrophobic property of the ligands are presented in fig. 9. The two red cubes at C2 and C4 of ring A and C4-C6 of ring C stresses that hydrophobic substitution at these positions decreases inhibitor affinity. Further, the model demonstrated that a large hydrophobic region around ring B and C is suitable for designing more potent compounds. Therefore, it is suggested that replacing the pyrrolidine ring with bulkier hydrophobic rings like phenyl and naphthalene might be beneficial for activity. This would particularly aid in the design of potent small molecular inhibitors of Hsp90. Six hydrophobic cubes at C5 of ring A implied that a hydrophobic group at this position is favored. This proposition is verified by compound 16 with one chlorine at 4 position of ring A, exhibiting better potency than compound 10. This region can be explored in future for increased activity by substituting a bulkier hydrophobic group like ethyl or isopropyl. The contribution of electron withdrawing groups in the ligands towards activity data is depicted in fig. 10. The 2 green cubes at C2 of ring A indicated that the hydroxyl group at this part can

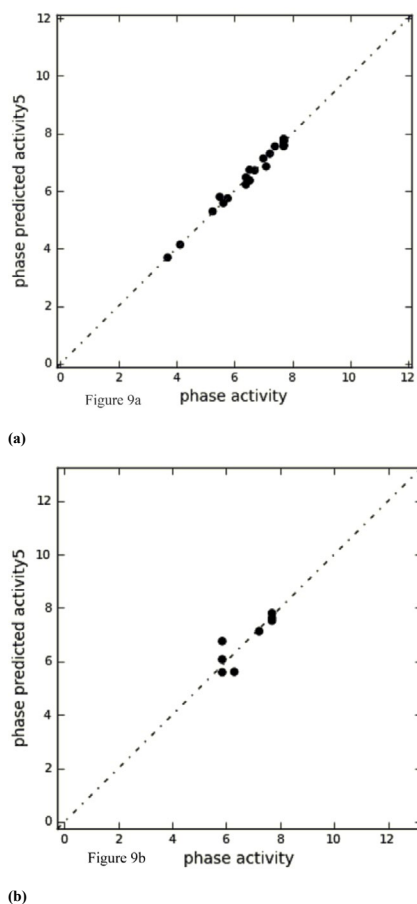
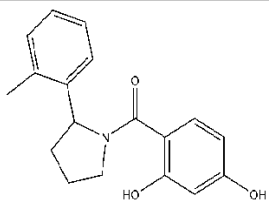
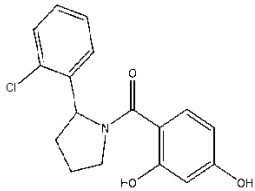
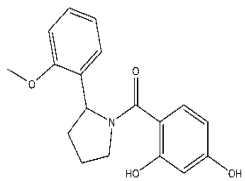
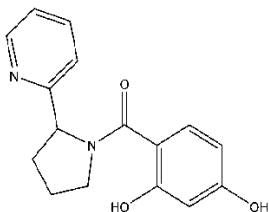
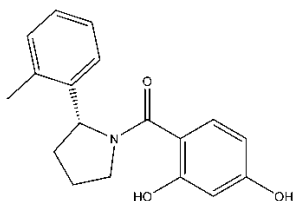
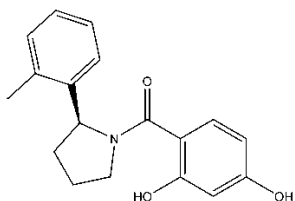
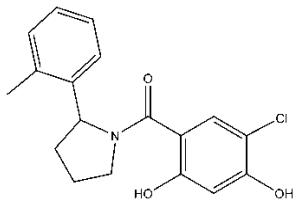
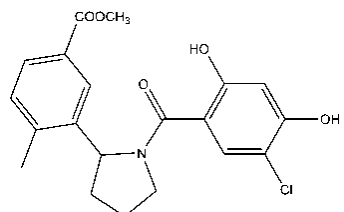


Fig. 7: Experimental and atom-based 3D-QSAR model predicted Hsp90 inhibition values
Relation between experimental and predicted Hsp90 inhibitory activity values of a) a training set molecules and b) a test set molecules using atom-based 3D-QSAR model

TABLE 3: PREDICTED VERSES REPORTED PIC_{50} VALUES OF TRAINING AND TEST SET MOLECULES

| Compound | Structure | QSAR set | Reported activity (PIC_{50})* | Phase predicted PIC_{50} * |
|----------|-----------|----------|-----------------------------------|------------------------------|
| 1 | | Training | 6.397 | 6.222 |
| 2 | | Training | 5.769 | 5.745 |
| 3 | | Training | 3.698 | 3.691 |
| 4 | | Training | 4.130 | 4.136 |
| 5 | | Training | 5.494 | 5.799 |
| 6 | | Test | 5.850 | 6.075 |
| 7 | | Training | 7.221 | 7.291 |
| 8 | | Test | 6.301 | 5.608 |
| 9 | | Training | 6.522 | 6.365 |

| Compound | Structure | QSAR set | Reported activity (pIC ₅₀)* | Phase predicted pIC ₅₀ * |
|----------|---|----------|--|-------------------------------------|
| 10 |  | Training | 7.000 | 7.131 |
| 11 |  | Training | 7.096 | 6.806 |
| 12 |  | Training | 6.397 | 6.471 |
| 13 |  | Training | 5.619 | 5.578 |
| 14 |  | Test | 7.221 | 7.134 |
| 15 |  | Training | 5.251 | 5.290 |
| 16 |  | Training | 7.698 | 7.559 |
| 17 |  | Training | 6.522 | 6.736 |

| Compound | Structure | QSAR set | Reported activity (pIC ₅₀)* | Phase predicted pIC ₅₀ * |
|----------|-----------|----------|--|-------------------------------------|
| 18 | | Test | 5.853 | 6.757 |
| 19 | | Training | 7.698 | 7.562 |
| 20 | | Test | 7.698 | 7.525 |
| 21 | | Training | 7.397 | 7.549 |
| 22 | | Training | 7.698 | 7.740 |
| 23 | | Training | 6.698 | 6.684 |

| Compound | Structure | QSAR set | Reported activity (pIC ₅₀)* | Phase predicted pIC ₅₀ * |
|----------|-----------|----------|---|-------------------------------------|
| 24 | | Training | 7.698 | 7.758 |
| 25 | | Training | 7.698 | 7.594 |
| 26 | | Test | 7.698 | 7.623 |
| 27 | | Training | 7.698 | 7.738 |
| 28 | | Test | 7.698 | 7.718 |
| 29 | | Training | 7.698 | 7.603 |
| 30 | | Test | 5.853 | 5.471 |
| 31 | | Training | 7.698 | 7.808 |
| 32 | | Training | 6.522 | 6.364 |
| 33 | | Training | 7.698 | 7.725 |

*pIC₅₀ = -log IC₅₀. IC₅₀ is expressed as μM

be substituted with electron withdrawing hydrophilic group like COOH and CONH₂. Subsequently, it was disclosed that substitution of C2 portion of ring B, C-4, C-5 part of ring C and C2-C5 region of ring C with an electron withdrawing group could lead to molecules

with enhanced activity. The study further highlighted the importance of alignment of molecules according to their stereochemistry. Compound 14 (the R isomer) aligned in a better fashion than the corresponding S isomer (15) which resulted in enhanced potency. This

TABLE 4: PREDICTED AND REPORTED ACTIVITY (pIC_{50}) VALUES OF COMPOUNDS USED FOR EXTERNAL VALIDATION

| Compound | Reported activity (pIC_{50})* | Phase predicted pIC_{50} * |
|------------------------|-----------------------------------|------------------------------|
| Radicicol | 7.22 | 7.19 |
| Geldanamycin | 6.55 | 6.45 |
| Tanespimycin, | 5.90 | 6.05 |
| Retaspimycin, | 6.30 | 6.25 |
| Alvespimycin | 6.34 | 6.45 |
| Purine analogue B11021 | 5.76 | 5.68 |

* $pIC_{50} = -\log IC_{50}$. IC_{50} is expressed as μM

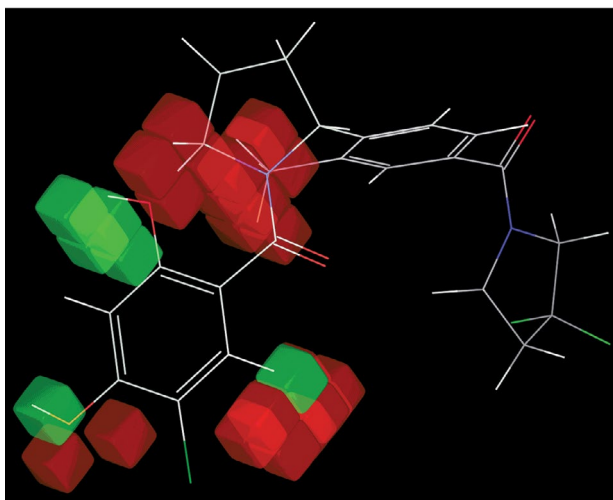


Fig. 8: Atom-based 3D-QSAR model visualized hydrogen bond donor property of molecule 33

Atom-based 3D-QSAR model visualized hydrogen bond donor properties of the most potent molecule 33. The favored regions are depicted in green cubes and disfavored regions are shown in red cubes

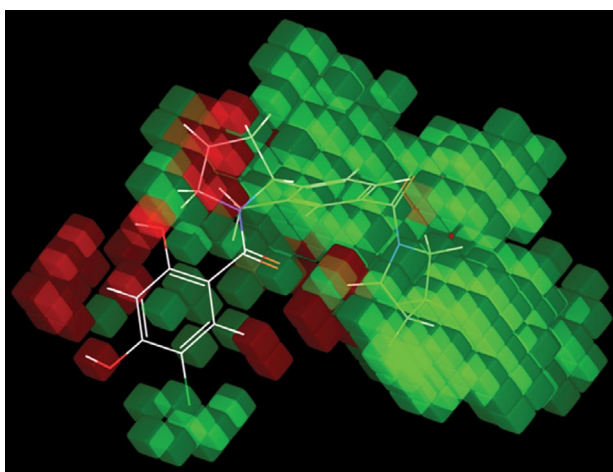


Fig. 9: Visualization by atom-based 3D-QSAR model of hydrophobicity parameters of molecule 33

Visualization by atom-based 3D-QSAR model of hydrophobicity parameter with reference to the most active molecule 33. The favorable regions are indicated in green cubes and the unfavorable portion is highlighted in the form of red cubes

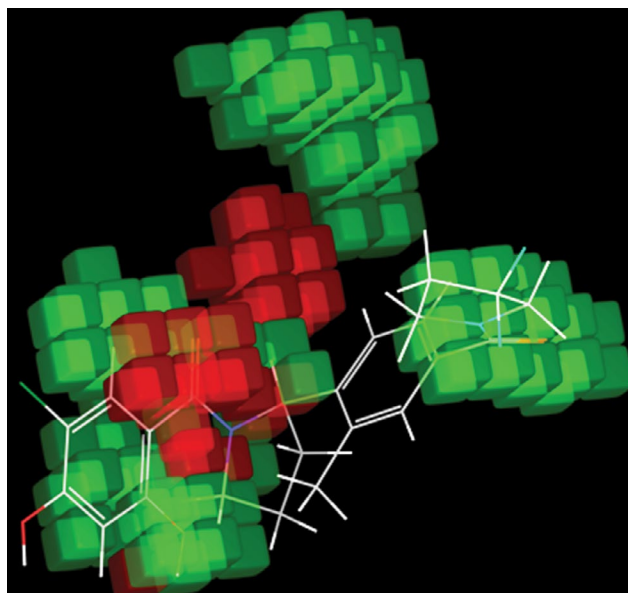


Fig. 10: Atom-based 3D-QSAR model prediction of electron withdrawing features of molecule 33

Atom-based 3D-QSAR model prediction of electron withdrawing features for the most active molecule 33. Green and red cubes indicate the favorable and unfavorable regions

was further explained by the activity of compound 28, which is more potent than the corresponding S isomer (30) and equipotent as the R isomer (31).

In conclusion, QSAR studies were performed with significant structural features of various linker regions within the molecules. It was observed that hydrophobic, electron withdrawing and H bond donor groups/regions between ring A and B plays no significant role towards enhanced potency of the molecules. Hence, it was suggested that this amide connector region can be substituted with easily synthesizable linkers like imines, enamines, diazines, esters, \sim and so on. Consequently, it was revealed that hydrophobic region is effective as a connector between ring B and C whereas H-bond donor and electron withdrawing group in this region has no major role in determining the Hsp90 inhibition potential. For example, compounds 6 and 8 have linker ($-SO_2$ and $-COOCH_2-$) in this part which demonstrated lower suppression potential than compound 19. A definite conclusion was not possible for the linker between ring C and D as it depends on the alignment of the molecules. For example, compounds 24-29, 31 and 33 are equipotent. It was observed that for few compounds (24, 26-29 and 31), the linker region was surrounded by red cubes in terms of electron withdrawing and H-bond acceptor properties whereas in case of 25 and 33 it was shown to be insignificant. However, all of them (24-33) demonstrated a large

number of green cubes for hydrophobicity in this region. Therefore, it is recommended that a hydrophobic linker group in this region could lead to potent compounds.

The above work not only highlighted the significance of conserved water molecules in designing docking studies but also provided a rapid protocol for identifying them. Another significant contribution of the study is the development of a robust technique for comparison and identification of an effective validated docking procedure using 6 different programs. This process could be extended to interaction studies of other drug targets with their ligands. The atom-based 3D-QSAR model developed by us highlighted the important structural features of molecules necessary for appropriate Hsp90 inhibition.

Thus, from the target point of view, the designed strategy identified the important amino acids (Asp 93, Thr 184, Lys 58, Asn 51 and Gly 97) and conserved water molecules (902, 903, 981 and 1026) required for Hsp90 suppression via hydrogen bond interactions with the ligands. Further the work also identified the hydrophobic amino acids (Phe 138, Val 186, Ile 91, Met 98, Leu 48, Leu 107, Ala 55, Ile 96 and Val 136) essential for effective protein-ligand binding. From ligand standpoint, the work exemplifies the significant structural requirement of the compounds for suppressing ATPase function of the chaperone. The biologically active ligands were also predicted to be potent by the docking and QSAR studies. Thus the receptor dependent studies (docking) correlated well with the ligand based drug designing approach (3D QSAR). Our research effort will particularly aid in the rational design of small molecules with greater specificity and potency against Hsp90.

ACKNOWLEDGEMENTS

Authors thank the DST (Fast track Scheme: SR/FT/CS- 079/2009) and AICTE (RPS Scheme: 8023/BOR/RID/RPS-102/2009-10) for providing drug designing software and other financial assistance for this project. The authors also thank the President, Gokaraju Rangaraju Educational Trust for providing high speed workstations along with excellent infrastructure for carrying out this work. The authors are grateful to the Indian Institute of Chemical Technology (IICT), Hyderabad, India for carrying out NMR and mass spectral studies. The authors are also grateful to the University of Buenos Aires, Argentina and IBYME-CONICET, Buenos Aires, Argentina for providing infrastructure to carry out the biological activity studies.

Conflicts of interest

All authors declare no conflict of interest.

REFERENCES

- Gottesman MM, Lavi O, Hall MD, Gillet J-P. Toward a Better Understanding of the Complexity of Cancer Drug Resistance. *Annu Rev Pharmacol Toxicol* 2016;56:85-102.
- Mansoori B, Mohammadi A, Davudian S, Shirjang S, Baradaran B. The Different Mechanisms of Cancer Drug Resistance: A Brief Review. *Adv Pharm Bull* 2017;7:339-48.
- Ohnishi S, Takeda H. Herbal medicines for the treatment of cancer chemotherapy-induced side effects. *Front Pharmacol* 2015;6:14.
- Tao JJ, Visvanathan K, Wolff AC. Long term side effects of adjuvant chemotherapy in patients with early breast cancer. *Breast* 2015;24(Suppl 2):S149-53.
- Calderon BP, Beck R, Glorieux C. Targeting hsp90 family members: a strategy to improve cancer cell death. *Biochem Pharmacol* 2019;164:177-87.
- Chatterjee S, Burns TF. Targeting Heat Shock Proteins in Cancer: A Promising Therapeutic Approach. *Int J Mol Sci* 2017;18:E1978.
- Taldone T, Wang T, Rodina A, Pillarsetty NVK, Digwal CS, Sharma S, *et al.* A Chemical Biology Approach to the Chaperome in Cancer-HSP90 and Beyond. *Cold Spring Harb Perspect Biol* 2020;12: a034116.
- Calderwood SK. Heat shock proteins and cancer: intracellular chaperones or extracellular signalling ligands? *Philos Trans R Soc Lond B Biol Sci* 2018;373: 20160524.
- Taipale M, Krykbaeva I, Koeva M, Kayatekin C, Westover KD, Karras GI, *et al.* Quantitative analysis of HSP90-client interactions reveals principles of substrate recognition. *Cell* 2012;150:987-1001.
- Zabinsky RA, Mason GA, Queitsch C, Jarosz DF. It's not magic- Hsp90 and its effects on genetic and epigenetic variation. *Semin Cell Dev Biol* 2019;88:21-35.
- Gorre ME, Ellwood-Yen K, Chiosis G, Rosen N, Sawyers CL. BCR-ABL point mutants isolated from patients with imatinib mesylate-resistant chronic myeloid leukemia remain sensitive to inhibitors of the BCR-ABL chaperone heat shock protein 90. *Blood* 2002;100:3041-4.
- Paraiso KHT, Haarberg HE, Wood E, Rebecca VW, Chen YA, Xiang Y, *et al.* The HSP90 inhibitor XL888 overcomes BRAF inhibitor resistance mediated through diverse mechanisms. *Clin Cancer Res* 2012;18:2502-14.
- Petrikaite V, Matulis D. Binding of natural and synthetic inhibitors to human heat shock protein 90 and their clinical application. *Medicina (Kaunas)* 2011;47:413-20.
- Gupta SD. Hsp90 Flexibility and Development of its Inhibitors for the Treatment of Cancer. *Current Chemical Biology*. 2018;12(1):53-64.
- Schopf FH, Biebl MM, Buchner J. The HSP90 chaperone machinery. *Nature Rev Molecular Cell Biol*. 2017;18(6):345-60.
- Thepchattri P, Eliseo T, Cicero DO, Myles D, Snyder JP. Relationship among ligand conformations in solution, in the solid state, and at the Hsp90 binding site: geldanamycin and radicicol. *J Am Chem Soc* 2007;129:3127-34.
- Soga S, Neckers LM, Schulte TW, Shiotsu Y, Akasaka K, Narumi H, *et al.* KF25706, a novel oxime derivative of radicicol, exhibits in vivo antitumor activity via selective

- depletion of Hsp90 binding signaling molecules. *Cancer Res* 1999;59:2931-8.
18. Kim T, Keum G, Pae AN. Discovery and development of heat shock protein 90 inhibitors as anticancer agents: a review of patented potent geldanamycin derivatives. *Expert Opin Ther Pat* 2013;23:919-43.
 19. Taldone T, Gozman A, Maharaj R, Chiosis G. Targeting Hsp90: small-molecule inhibitors and their clinical development. *Curr Opin Pharmacol* 2008;8:370-4.
 20. Biamonte MA, Van de Water R, Arndt JW, Scannevin RH, Perret D, Lee W-C. Heat Shock Protein 90: Inhibitors in Clinical Trials. *J Med Chem* 2010;53:3-17.
 21. Yuno A, Lee M-J, Lee S, Tomita Y, Rekhtman D, Moore B, *et al.* Clinical Evaluation and Biomarker Profiling of Hsp90 Inhibitors. *Methods Mol Biol* 2018;1709:423-41.
 22. Shimomura A, Yamamoto N, Kondo S, Fujiwara Y, Suzuki S, Yanagitani N, *et al.* First-in-Human Phase I Study of an Oral HSP90 Inhibitor, TAS-116, in Patients with Advanced Solid Tumors. *Mol Cancer Ther* 2019;18:531-40.
 23. Haider K, Huggins DJ. Combining solvent thermodynamic profiles with functionality maps of the Hsp90 binding site to predict the displacement of water molecules. *J Chem Inf Model* 2013;53:2571-86.
 24. Yan A, Grant GH, Richards WG. Dynamics of conserved waters in human Hsp90: implications for drug design. *J R Soc Interface* 2008;5 Suppl 3:S199-205.
 25. Kung P-P, Sinnema P-J, Richardson P, Hickey MJ, Gajiwala KS, Wang F, *et al.* Design strategies to target crystallographic waters applied to the Hsp90 molecular chaperone. *Bioorg Med Chem Lett* 2011;21:3557-62.
 26. Verkhivker GM. Computational Modeling of the Hsp90 Interactions with Cochaperones and Small-Molecule Inhibitors. *Methods Mol Biol* 2018;1709:253-73.
 27. Kung P-P, Funk L, Meng J, Collins M, Zhou JZ, Johnson MC, *et al.* Dihydroxyphenyl amides as inhibitors of the Hsp90 molecular chaperone. *Bioorg Med Chem Lett* 2008;18:6273-8.
 28. Gupta SD, Revathi B, Mazaira GI, Galigniana MD, Subrahmanyam CVS, Gowrishankar NL, *et al.* 2,4-dihydroxy benzaldehyde derived Schiff bases as small molecule Hsp90 inhibitors: rational identification of a new anticancer lead. *Bioorg Chem* 2015;59:97-105.
 29. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, *et al.* Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem* 2004;47:1739-49.
 30. Cross JB, Thompson DC, Rai BK, Baber JC, Fan KY, Hu Y, *et al.* Comparison of several molecular docking programs: pose prediction and virtual screening accuracy. *J Chem Inf Model* 2009;49:1455-74.
 31. Jain AN. Surflex-Dock 2.1: robust performance from ligand energetic modeling, ring flexibility, and knowledge-based search. *J Comput Aided Mol Des* 2007;21:281-306.
 32. Spitzer R, Jain AN. Surflex-Dock: Docking benchmarks and real-world application. *J Comput Aided Mol Des* 2012;26:687-99.
 33. Guedes IA, de Magalhaes CS, Dardenne LE. Receptor-ligand molecular docking. *Biophys Rev* 2014;6:75-87.
 34. Wang R, Lu Y, Wang S. Comparative evaluation of 11 scoring functions for molecular docking. *J Med Chem* 2003;46:2287-303.
 35. Zhou N, Xu Y, Liu X, Wang Y, Peng J, Luo X, *et al.* Combinatorial Pharmacophore-Based 3D-QSAR Analysis and Virtual Screening of FGFR1 Inhibitors. *Int J Mol Sci* 2015;16:13407-26.
 36. Ge J, Normant E, Porter JR, Ali JA, Dembski MS, Gao Y, *et al.* Design, synthesis, and biological evaluation of hydroquinone derivatives of 17-amino-17-demethoxygeldanamycin as potent, water-soluble inhibitors of Hsp90. *J Med Chem* 2006;49:4606-15.
 37. Howes R, Barril X, Dymock BW, Grant K, Northfield CJ, Robertson AGS, *et al.* A fluorescence polarization assay for inhibitors of Hsp90. *Anal Biochem* 2006;350:202-13.
 38. Lundgren K, Zhang H, Brekken J, Huser N, Powell RE, Timple N, *et al.* BIIB021, an orally available, fully synthetic small-molecule inhibitor of the heat shock protein Hsp90. *Mol Cancer Ther* 2009;8:921-9.