Pharmacophore based 3D-QSAR modeling and free energy analysis of VEGFR-2 inhibitors

Muthukumaran Rajagopalan, Sangeetha Balasubramanian, Amutha Ramaswamy & Premendu Prakash Mathur

informa
healthcare

RESEARCH ARTICLE

Pharmacophore based 3D-QSAR modeling and free energy analysis of VEGFR-2 inhibitors

Muthukumaran Rajagopalan, Sangeetha Balasubramanian, Amutha Ramaswamy, and Premendu Prakash Mathur

Centre for Bioinformatics, School of Life Sciences, Pondicherry University, R.V. Nagar, Kalapet, Puducherry, India

Abstract

VEGFR-2, a transmembrane tyrosine kinase receptor is responsible for angiogenesis and has been an attractive target in treating cancers. The inhibition mechanism of structurally diverse urea derivatives, reported as VEGFR-2 inhibitors, was explored by pharmacophore modeling, QSAR, and molecular dynamics based free energy analysis. The pharmacophore hypothesis AADRR, resulted in a highly significant atom based 3D-QSAR model (r^2 = 0.94 and q^2 = 0.84). Binding free energy analysis of the docked complexes of highly active and inactive compounds, after 7 ns MD simulation, revealed the importance of van der Waals interaction in VEGFR-2 inhibition. The decomposition of binding free energy on a per residue basis disclosed that the residues in hinge region and hydrophobic pocket play a role in discriminating the active and inactive inhibitors. Thus, the present study proposes a pharmacophore hypothesis representing the identified interactions pattern and its further application as a template in screening databases to identify novel VEGFR-2 inhibitor scaffolds.

Keywords: VEGFR-2, Pharmacophore, 3D-QSAR, free energy analysis, MMGBSA, docking

Introduction

Protein kinases (PKs) play a vital role in cellular processes such as cell signaling, survival and proliferation. Approximately 518 protein kinases have been identified in the human kinome and their importance in disease progression led to their identification as novel structural targets[1,2]. Vascular endothelial growth factor (VEGF) is an important angiogenic agent that mediates the formation of new blood vessels and enhances the growth, proliferation, migration and differentiation of endothelial cells[3]. The pathogenicity of angiogenesis in disorders such as cancer, proliferative retinopathies and rheumatoid arthritis has been proved experimentally[4]. VEGF family of secretory proteins modulate the cellular effects by interacting specifically with high-affinity transmembrane tyrosine kinase receptors (RTKs): VEGFR-1 (Flt-1), VEGFR-2 (KDR human/Flk-1 mouse) and VEGFR-3 (Flt-4) which then triggers effective downstream cell proliferation leading to angiogenesis[5]. Expression of VEGFR-1 and VEGFR-2 is responsible for angiogenesis while the expression of VEGFR-3 is responsible for lymphangiogenesis. The level of expression of VEGFR-2 in cell proliferation and differentiation is found to be higher when compared to that of VEGFR-1. Hence the inhibition of VEGFR-2 signaling pathway by small molecule inhibitors has become an attractive strategy for the treatment of cancers[6].

VEGFR-2 is a transmembrane receptor with 1356 amino acids and secreted by the *VEGFR-2* gene located on chromosome 4q11-q12. The extracellular region consists of seven immunoglobulin-like domains (46–753 aa) followed by a short transmembrane domain (765–789 aa) and an intracellular kinase domain (834–1162 aa). The binding of VEGF with the extracellular immunoglobulin (Ig)-like domains 2 and 3 induces receptor dimerization and autophosphorylation of tyrosine residues 1054 and 1059 in the kinase domain[7,8]. Crystal structure of complete VEGFR-2 including the extracellular, transmembrane and intracellular kinase domains has not been solved yet. Individual structures of Ig-like C2 type-2, type-3 and

the intracellular kinase domain have been crystallized by various groups[9–12]. The kinase domain possesses an N-terminal and C-terminal lobes which are connected by the residues Glu915, Phe916, Cys917. The important regions of the kinase domain like glycine rich region (GXGXXG), HRD motif, ATP binding site and DFG loop influence the substrate and inhibitor binding[7]. Studies have also shown that inhibitors interacting with this hinge region through H-bonds are biologically active.

The Federal drug administration of US has approved sorafenib (Bay 43–9006), sunitinib (SU-11248), pazopanib (GW786034), and vandetanib (ZD6474) as inhibitors of VEGFR-2. A series of small molecules with antiangiogenic property is in clinical studies, which includes cerdiranib (AZD-2171)], brivanib (BMS-582664), axitinib (AG013736), tivozanib (KRN-951) and vatalanib (pTK787)[13,14]. All these inhibitors possess a broad range of activities towards protein kinases. Therefore, it is necessary to develop new inhibitors with high selectivity towards VEGFR-2 kinase and also to overcome drug resistance.

The effort by Huang et al.[15,16] to develop small-molecular ATP-competitive VEGFR-2 inhibitors has led to the discovery of 4-aminopyrimidine-5-carbaldehyde oximes. The 4-aminopyrimidine-5-carbamate which is considered as an open form of quinazoline forms the backbone for these structures. Substitution of a phenoxy group at the 6th position of this pyrimidine increased the activity. Substitution of a urea group at the 1st position of the phenoxy ring further enhanced the activity.

An in-house database screening by Kubo et al.[17] against a database of tyrosine kinase inhibitors led to the identification of a series of N-Phenyl-N'-{4-(4-quinolyloxy) phenyl} urea derivatives with high specificity to VEGFR-2 than other kinases (PDGFR α, c-KIT FGFR-2 EFGR, HGFR). The 4-aminopyrimidine group in the former series has been replaced by a quinoline group in this series where urea and phenoxy group substitutions were studied and are tested for VEGFR-2 inhibitory activity. Among all these VEGFR-2 inhibitors, the N-Phenyl-N'-{4-(4-quinolyloxy) phenyl} urea derivatives were found to be the most active compounds.

Ligand based drug design methods like Quantitative Structure Activity Relationship (QSAR) and pharmacophore modeling have proven their efficiency in (i) designing/predicting the activity of new compounds and (ii) searching chemical databases to identify novel lead scaffolds[18–20]. The pharmacophore model is a collection of chemical features which are common to a set of compounds with similar inhibition mechanism, and are necessary for their inhibitory activity against a particular target[21–23]. The chemical features like H-bond acceptors and donors, charged or ionisable groups, hydrophobic or aromatic rings along with their spatial arrangement in terms of distance, angles and dihedrals constitute the pharmacophore model/hypothesis[24].

A number of computational methods are available to calculate the binding free energy of a protein-ligand complex, such as thermodynamic integration (TI), linear response (LR), free energy perturbation (FEP), fluctuation–dissipation theorem (FDT), molecular mechanics Poisson–Boltzman surface area (MMPBSA) and molecular mechanics generalized Born surface area (MMGBSA)[25]. The MMGBSA method has been applied successfully in the recent years, for the estimation of binding free energy[26–31]. *In silico* drug design techniques aid the experimental evidence with a clear picture of biomolecular sensing at the atomic level and methods like docking and molecular dynamics claim their merit. Docking methods define the stable conformation of a ligand with respect to the receptor's binding cavity and use only the static image of the bound conformation to define the binding energy. In free energy analysis, a conformational sampling from an unbound to bound state is performed to generate thermodynamic averages. MMGBSA, an end-point method, which calculates the binding free energy as a difference between the energies of complex, receptor and ligand has been used in addition with molecular dynamics simulation.

In this study, we attempt to develop a pharmacophore model for a set of 81 structurally diverse VEGFR-2 inhibitors belonging to N-Phenyl-N'-{4-(4-quinolyloxy) phenyl} urea and 4-aminopyrimidine-5-carbaldehyde oxime using the PHASE module of Schrödinger. Molecular docking of inhibitors into the binding cavity of VEGFR-2 has been performed to study their mode of interaction. Molecular dynamics simulation for a period of 7 ns was performed using AMBER11 package. The binding free energy of VEGFR-2 in complex with active and inactive compounds was calculated as an average over a series of snapshots obtained from the molecular dynamics simulation trajectory using the MMPBSA program available in AmberTools 1.5.

Materials and methods

Data set
A set of 81 novel compounds belonging to 4-amino-pyrimidine-5-carbaldehyde oxime and N-Phenyl-N'-{4-(4-quinolyloxy) phenyl} urea were used for this analysis. The experimentally reported pIC_{50} (i.e. the negative logarithm of the measured IC_{50}) values of these compounds measured for the inhibition of VEGFR-2 in a cell-based assay were obtained from Huang et al.[15,16] and Kubo et al.[17]. A flowchart depicting the computational approaches used in this study has been shown as Supplementary Figure S1.

Pharmacophore modeling
Pharmacophore based alignment and atom based 3D-QSAR studies were performed using PHASE[32] module of Schrödinger. The 3D structures of 81 compounds were generated using Maestro's Model Builder and prepared using LigPrep 2.5 (Schrödinger Inc.) module of PHASE. The activity (pIC_{50}) of these ranges from 4.706 to 9.699 and were categorized into active (above 8.70) and

inactive (below 6.50) compounds. All possible conformers were generated for the prepared structures by confgen method. Around 100 possible conformations per rotatable bond were generated using a rapid torsional search, followed by 100 steps of pre-processing and 50 steps of post-processing minimization using OPLS-2005 force field in a continuum solvent model[33]. The conformers possessing RMSD less than 1.2 Å were eliminated.

Generation of common pharmacophore hypothesis

PHASE identifies six different pharmacophoric features (specified as SMARTS queries) defined by the chemical structure patterns and are hydrogen bond acceptor (A), hydrogen bond donor (D), hydrophobic group (H), negatively charged group (N), positively charged group (P) and aromatic ring (R). The pharmacophores that are common in all the active compounds were identified using a tree based partitioning algorithm and studied further. A tree depth of 5 and the tolerance to match a pharmacophore (defined in terms of grid size), set as 1 Å was used to identify the common five variant pharmacophores.

The obtained Common Pharmacophore Hypotheses (CPHs) were aligned with the active molecules and were scored for matching the actives (survival score) with an RMSD cut-off, 1.2 Å. The scoring process was performed with default parameters, except the ligand activity which was incorporated with a weight of 0.3 using the equation shown below

$$S = W_{site}S_{site} + W_{vec}S_{vec} + W_{vol}S_{vol} + W_{sel}S_{sel} + W_{rew}^m - W \, \Delta E + W_{act}A \tag{1}$$

Where W's are the weights and S's are the scores.

The derived CPHs were subsequently scored for matching the inactives similar to the calculation of survival score of the actives. In addition to the inactive score, an adjusted survival score is calculated as a difference between the survival scores of actives and inactives using a weight of 1.0 for the inactives score.

$$S_{adj} = S_{actives} - W_{inactives}S_{inactives} \tag{2}$$

3D QSAR model

Pharmacophore based QSAR model considers the ligand features that are used to derive the pharmacophore variants. In such situations, an atom based 3D-QSAR is more advantageous. The selected 81 compounds were randomly classified as training (60 compounds) and test (21 compounds) set and were ensured to have proper distribution of chemical and structural diversity. The training set compounds are aligned based on the best scoring CPH and the aligned compounds are enclosed in a rectangular grid. Every atom of the ligand is classified into any one of the six classes namely: (i) D – hydrogen bond donor (hydrogens bonded to N, O, P, S), (ii) H – hydrophobic/non-polar (C, H–C, Cl, Br, F, I), (iii) N – negative ionic (formal negative charge), (iv) P – positive ionic (formal

positive charge), (v) W – electron-withdrawing (N, O), and (vi) X – miscellaneous (all other types of atoms). The rectangular grid is divided into uniformly-sized cubes of dimension 1 Å which is occupied by van der Waals spheres whose radii depend on the atom type in a ligand. Each ligand is represented by a set of bit values (0 or 1) to indicate the occupancy of the cube from different class of atoms. These occupancies of cubes and atom classes can be used as independent variables in the generation of partial least squares (PLS) regression based QSAR model. PHASE evaluates the generated QSAR models by leave-n-out (LNO) cross-validation method and gives a stability value to indicate the sensitivity of the generated model towards changes in the training set. The parameters defining the statistical significance of the training (r^2) and test set (q^2) are calculated using equation (3).

$$r^2 = 1 - \left[\sum (Y_{pred} - Y_{obs})^2 / \sum (Y_{obs} - Y_{mean})^2 \right] \tag{3}$$

Where, Y_{pred}, Y_{obs} and Y_{mean} are the predicted, observed and mean biological activity values. The value obtained from $\sum(Y_{pred}-Y_{obs})^2$ is the predicted residual sum of squares (PRESS). The predictive correlation coefficient (r^2_{pred}) calculated using equation (4) for the test set compounds reflects the reliability of the prediction.

$$r^2_{pred} = 1 - \left[\sum (Y_{pred} - Y_{obs})^2 / \sum (Y_{obs} - Y_{mean}^{training})^2 \right] \tag{4}$$

Molecular dynamics simulation

The crystal structure of VEGFR-2 kinase domain (PDB entry: 1YWN)[11] complexed with the inhibitor 4-amino-furo [2, 3-d] pyrimidine was taken as the starting structure for docking and molecular dynamics simulation studies. The missing residues in the VEGFR-2 crystal structure were added using the procedure adopted by Papakyriakou et al.[34]. The protein and inhibitors were prepared for docking using Schrödinger. The prepared ligands were docked using the GLIDE[35] module of Schrödinger with default parameters. The docked conformation of the most active (compound 51) and inactive (compound 75) compounds were extracted and their conformational stability/time evolution was analyzed through all atom molecular dynamics simulation for a period of 7 ns. Molecular dynamics simulations were carried out using PMEMD module of AMBER 11[36]. Amberff99SB force field was used for proteins[37]. The ligand charges derived at AM1 level were used by the antechamber[38] program to generate a Generalized Amber Force Field (GAFF) compatible force field for the ligand[39]. The hydrogen atoms present in the protein-ligand complex were removed, and new hydrogen atoms were added using tleap program. The complex was solvated using TIP3P[40] water box that extends upto 10 Å in all directions from the solute. The protein-ligand complex was minimized in two steps. First, the water molecules were relaxed by restraining the complex, which is followed by the minimization of the entire system. The long-range electrostatic interactions

were governed using particle-mesh Ewald procedure[41]. The system was heated to 300 K for 50 ps at constant volume with harmonic restraints of 2 kcal/mol/Å^2 on the protein-ligand complex. The temperature was maintained at 300 K using the Langevin dynamics with a collision frequency of 2.0 ps^{-1}. The hydrogen bonds were constrained using the SHAKE algorithm[42]. Density of the system was equilibrated gradually for a period of 80 ps. Then the entire system was equilibrated and simulated for a period of 7 ns.

MM-GBSA calculation

The MM-PBSA program implemented in AmberTools 1.5 package was used for the binding free energy analysis of the complex[43]. Electrostatic interactions were governed using Generalized Born method developed by Onufriev et al.[44]. From the molecular dynamics trajectory, conformational frames were extracted from the last 3 ns with an interval of 10 ps, and the binding free energy was calculated as an average of these frames. The equation used to calculate the binding free energy (ΔG) is:

$$\Delta G_{bind} = \Delta G_{complex} - \left(\Delta G_{receptor} + \Delta G_{ligand} \right) \tag{5}$$

where,

$$\Delta G = \Delta E_{mm} + \Delta G_{sol} - T\Delta S$$

$$\Delta E_{mm} = \Delta E_{int} + \Delta E_{ele} + \Delta E_{vdW}$$

$$\Delta G_{sol} = \Delta G_{pol} + \Delta G_{nonpol}$$

$$\Delta G_{nonpol} = \gamma SASA + \beta$$

The average free energies of complex, receptor and ligand are represented as $\Delta G_{complex}$, $\Delta G_{receptor}$, and ΔG_{ligand}, respectively. ΔE_{mm} is the energy derived using gas phase molecular mechanics. ΔG_{sol} is the solvation energy with polar and non polar terms. The polar contribution to the binding energy was calculated using Generalized Born approximation with a grid spacing of 0.5 Å and the dielectric constants, 1.0 and 80.0 were used for solute and solvent, respectively. The non-polar contribution is calculated from solvent accessible surface area (SASA) using a probe of radius, 1.4 Å.

Results and discussion

A ligand based drug design approach was carried out for a diverse set of 81 novel VEGFR-2 inhibitors (Supplementary Table S1) to study their mode of binding and interaction. All possible conformers of these inhibitors were generated using the confgen module of Macromodel. Pharmacophore modeling was carried out by classifying these 81 inhibitors into 16 active, 56 moderately active and nine inactive compounds based on their pIC$_{50}$ value. Pharmacophore features were identified for these inhibitors and aligned with all 16 active molecules

Table 1. Summary of the statistical parameters obtained from the PLS analysis of five best CPHs are listed below.

Statistical parameters	CPH1	CPH2	CPH3	CPH4	CPH5
Variants	AADRR	AAADR	AADRR	ADRRR	AADRR
Survival score	6.67	6.67	6.66	6.69	6.66
SD	0.25	0.25	0.26	0.28	0.27
r^2	0.94	0.94	0.94	0.93	0.93
F	193	183	168	143	164
P	3.851E-33	1.405E-32	1.269E-31	7.375E-30	2.439E-31
RMSE	0.41	0.39	0.45	0.40	0.45
Q^2	0.84	0.86	0.8	0.85	0.81
Pearson-R	0.92	0.93	0.90	0.93	0.91

using a tree-based partitioning algorithm. A total of 1784 five-featured CPHs were generated and subjected to scoring function analysis with the default parameters for site, vector and volume and a weight of 0.3 on the ligand activity. Subsequently, the hypotheses were also scored with respect to the nine inactives, and an adjusted survival score was calculated. The top 10% of the hypotheses which survived this scoring process were selected to build an atom based 3D-QSAR model. The statistical parameters of five best CPHs are listed in Table 1. The CPH1 (AADRR) with two hydrogen bond acceptors (A), one hydrogen bond donor (D) and two aromatics rings (R) as pharmacophoric features is considered for further studies due to its statistical significance. Figure 1 illustrates the distance between the various pharmacophore features. The two hydrogen bond acceptor features of CPH1 maps onto the nitrogen of quinoline group and an oxygen of urea group. The hydrogen bond donor feature is located on the urea nitrogen and the aromatic feature maps onto the middle phenyl group which connects the quinoline and urea moieties Figure 2A and 2B.

Statistical parameters such as r^2, cross validation co-efficient (q^2), Standard Deviation (SD), root mean square error (RMSE) and F-value were used to validate the 3D-QSAR model. The 81 compounds were divided into training and test sets of 60 and 21 compounds respectively. The QSAR model built for the CPH1 using five PLS factors exhibits good statistical significance for both training ($r^2 = 0.947$, SD = 0.251, F = 193, P = 3.85E-33) and test sets ($q^2 = 0.8447$, $r^2_{pred} = 0.8453$, RMSE = 0.417, Pearson-R = 0.925). The observed and predicted biological activity of training and test set compounds are given as Supplementary Table S1 and Figure 3 depicts the agreement between the reported as well as predicted activity values.

Visualization of 3D QSAR models

3D visualization of the QSAR model in the context of active and inactive compounds provides useful insights into their inhibitory activity. PHASE generates PLS regression based contours (in the form of cubes) to display the regions important for hydrophobic, hydrogen bond donor and electron-withdrawing (hydrogen bond
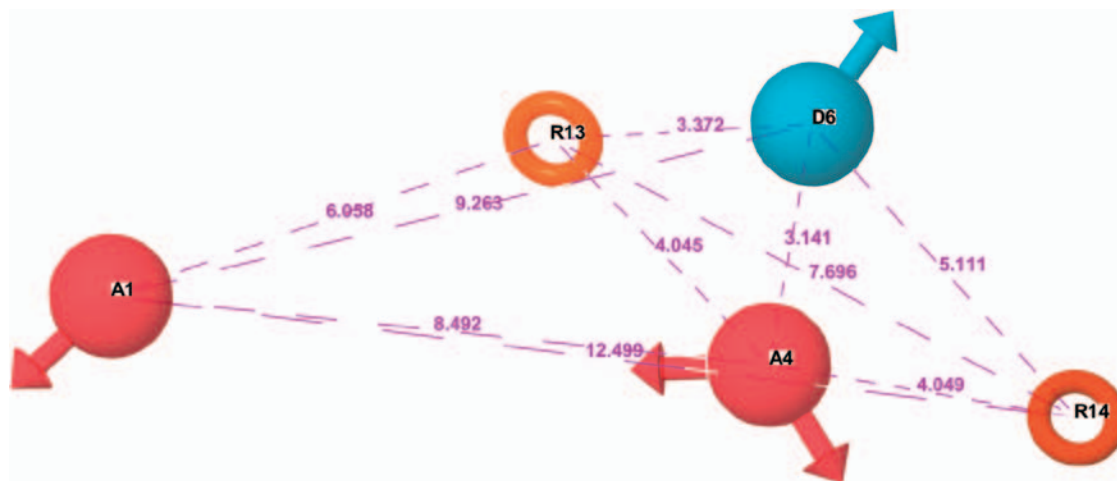
Figure 1. Distance between the pharmacophore features of CPH1. The features like H-bond donor & electron withdrawing group and the aromatic rings are denoted as D, A and R, respectively.

acceptor) properties. Figure 2 depicts these contours generated for active compound 51 and inactive compound 75.

The favorable and unfavorable regions for hydrogen bond donor property are shown in Figure 2C and 2D as blue and red cubes, respectively. In the most active compound 51, favorable regions are found close to the substituted NH of the urea group. Whereas the inactive compound, 75 inherits unfavorable regions due to the lack of urea group and the improper orientation of phenyl substituted NH. Also the alignment of active and inactive compounds shows that the amino group substituted in the pyrimidine ring is located in an electrophilic environment and hence shows an unfavorable region.

The nature of electron withdrawing property and its influence on various regions of the inhibitor is shown in the Figure 2E and 2F. The favorable regions (blue cubes) present near the linker O and the O and NH groups of urea connected to methyl-phenyl indicate that H-bond acceptor features at these positions favor the activity. On the other hand, the inactive compound 75 does not show any favorable regions for the electron withdrawing property.

Figure 2G illustrates the sterically favorable (blue cubes) and unfavorable (red cubes) hydrophobic/non polar interactions for the active compound 51. The terminal substitution at urea group is highly favorable for the activity, which is also supported by the other active compounds 58, 52, 56, 60, 61 and 19 having bulkier substitutions at this position. Substitution at 3rd position of middle phenyl ring is unfavorable for the activity. The same is true in the case of inactive compound 75 (Figure 2H) and additional unfavorable contours are also observed at the 6th position of middle phenyl ring. In the active compound 51, the acceptor pharmacophoric feature A1 is positioned at the quinozoline N-atom and the same A1 is observed at the methyloxime O-atom in the inactive compound. The active compound has no unfavored steric sites in the vicinity of this acceptor,

whereas, in the inactive compound, the methyloxime group expresses unfavored steric sites around the N-O atoms. The favorable regions observed at the methyl group suggest that hydrophobic substitutions at this site would favor for the activity.

In general, the favorable and unfavorable sites identified for H-bond donor-acceptor and hydrophobic regions depict the participation of functional groups in distinguishing active and inactive compounds. Electrophilic substitution at the phenyl substituted -NH and terminal substitutions at the urea group significantly favors the activity. In the closed form, quinazoline ring (compound 51) shows neither favorable nor unfavorable nature whereas its open form (compound 75) reveals unfavorable regions.

Molecular dynamics simulation

In addition to the pharmacophore model prediction of the 81 VEGFR-2 inhibitors, docking and MD simulation studies were performed to understand their binding mode and stability with VEGFR-2. The inhibitors were docked at the ATP binding site of VEGFR-2 using GLIDE. The obtained docking scores revealed a good agreement with the trend in experimental activity and are reported in Supplementary Table S1. The complexes of most active and inactive compounds (51 and 75, respectively) docked at the ATP binding site of VEGFR-2 were subjected to molecular dynamics simulation for a period of 7 ns to understand their stable interaction pattern and the free energy of binding. The plot of atomic root mean square deviation (RMSD) of protein backbone shown in Figure 5A depicts the stable dynamics of the protein-ligand complex formed by active (black lines) and inactive (red lines) compound during the course of simulation. From this RMSD plot, it is observed that the complex-51 (active compound) stabilizes after 2 ns and fluctuates around 2.25 ± 0.25 Å, whereas the complex-75 (inactive compound) attains equilibrium only after 2.5 ns and fluctuates as 2.75 ± 0.25 Å. The root mean square
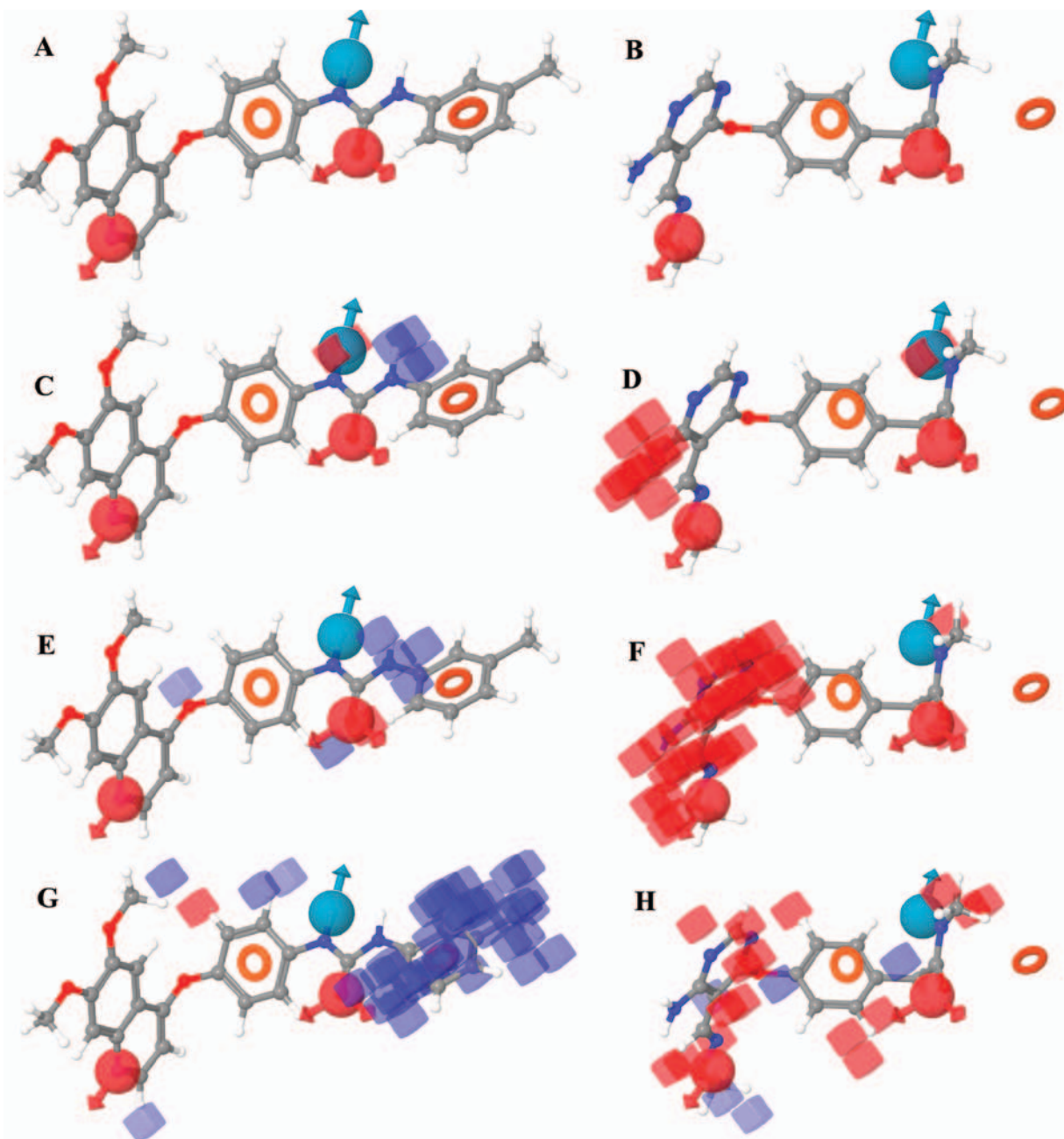
Figure 2. Visualization of 3D contours generated using CPH1 for the most active and inactive compounds (51 and 75). The subsets A and B depict the mapping of CPH1 onto the compounds 51 and 75, respectively. Subsets C–H depict the favorable (blue) and unfavorable (red) regions derived from the atom based 3D QSAR model for the H-bond donor (C, D), electron withdrawing (E, F) and hydrophobic properties (G, H) respectively.

fluctuation (RMSF) calculated as a function of residue number (Figure 5B) for both complexes reveals higher fluctuation in the complex-75 than complex-51. This RMS fluctuation also discloses the existing compact interaction within the active complex and the free residual motions in the inactive complex. The complex-51 shows an RMS fluctuation of 7 ± 1 Å whereas the complex-75 varies as 8 ± 3 Å.

Non-covalent interactions such as H-bond and π-interactions play an important role in stabilizing the protein-inhibitor complex and are analyzed for both the active and inactive complexes. The binding cavity of

VEGFR-2 is formed by Leu838, Val846, Ala864, Lys866, Leu887, Val897, Gly915, Phe916, Cys917 and Leu1033 residues. Analysis of non-covalent interactions in active and inactive complexes revealed the binding of both inhibitors to this cavity and express similar site-specific interactions (Figure 4). The residues involved in polar contacts are depicted in magenta caret and the residues participating in non-polar interactions are shown in green caret.

The compound 51 formed four H-bonds and four π-interactions when docked against VEGFR-2, and the compound 75 expressed four H-bonds and two
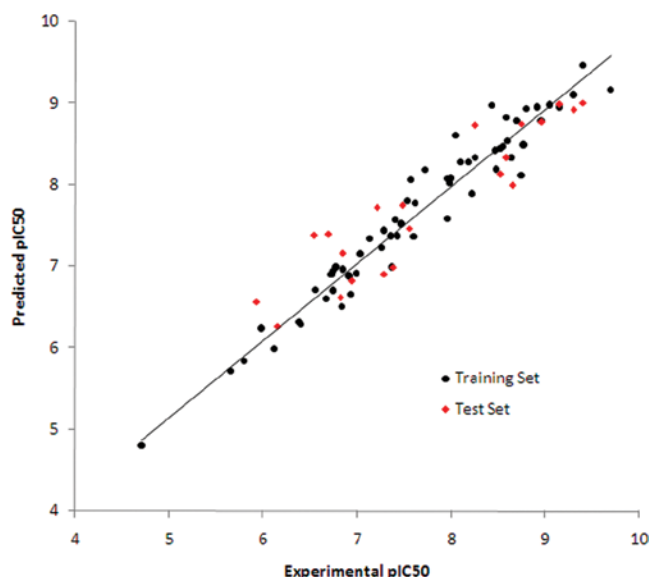
Figure 3. Plot of experimental activities against predicted biological activities of VEGFR-2 inhibitors. The traning and test set compounds are in spheres and diamonds, respectively.

π-interactions in the docked conformation. During MD, the inhibitor is effectively accommodated in the binding cavity via the structural changes of active site residues. The observed π-interaction of middle aromatic ring with Phe1045 in docked active complex disappeared during simulation. The simulated inactive complex failed to express interactions like (i) π-interaction between Lys866 and middle aromatic ring and (ii) H-bond with Asp1044 that were observed while docking. By allowing such ligand induced-fit changes during dynamics, the observed stable protein-inhibitor interactions are discussed. Figure 4 illustrates the 2D-interaction pattern observed before and after simulation (of 7 ns) for both active and inactive complexes.

In the complex-51, several H-bonding interactions are observed. The N atom (in quinolone group) and O atom (in urea group) forms H-bonds with the main chain NH of the residues Cys917 and Asp1044, respectively. Both the N atoms of urea interact with the side chain of Glu883 through a bifurcated H-bond. All these H-bond forming moieties of the inhibitor has been recognized as the pharmacophoric features (denoted as A1, A4, D6 in Figure 1) in the CPH1.
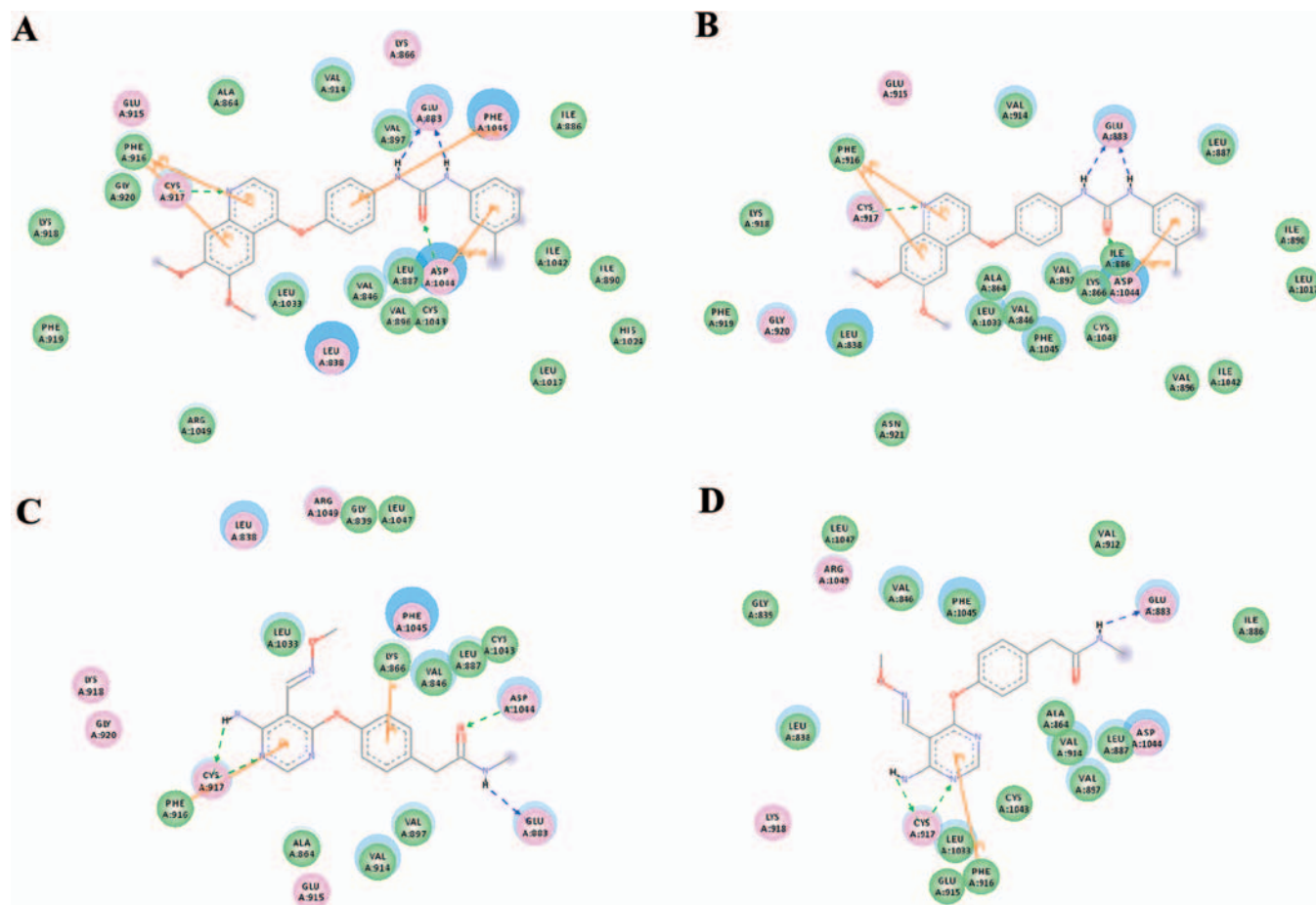

Figure 4. 2D representation of the interaction pattern of VEGFR-2 with active (A, B) and inactive (C, D) compounds before (A,C) and after (B,D) a simulation of 7 ns.
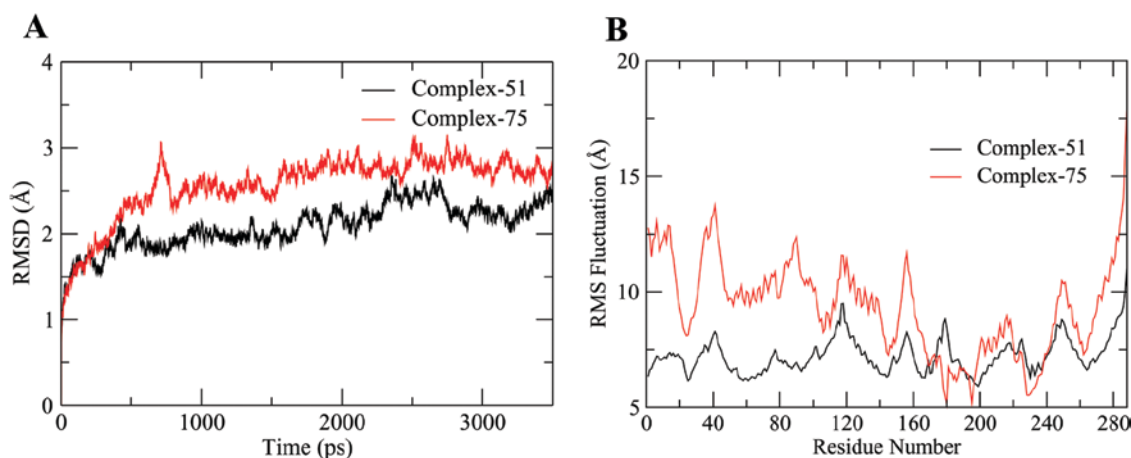
Figure 5. The RMSD (A) and RMSF (B) of a 7 ns long MD trajectory of VEGFR-2 complexes with active (black lines) and inactive (red lines) compounds, respectively.

Table 2. Binding free energy and its individual components in kcal/mol calculated by the MMGBSA method.

Inhibitors	ΔE_{ele}	ΔE_{vdw}	ΔG_{nonpol}	ΔG_{pol}	$\Delta G_{ele+pol}$	ΔG_{bind}	ΔG_{exp}*
Complex-51	–35.57	–60.88	–7.70	44.26	8.69	–59.89	–13.30
Complex–75	–24.96	–47.22	–5.96	41.34	16.38	–36.81	–6.45

*ΔG_{exp} is calculated as $RT\ln(IC_{50})$.

In addition to these H-bonding interactions, the terminal aromatic rings involve in π-interactions. Both the aromatic rings of quinolone moiety involve in π-π interaction with Phe916. The phenyl ring substituted at the urea group, which is sandwiched between Asp1044 and the hydrophobic pocket (lined by the residues Ile886, Leu887 and Ile890) expresses a σ-π interaction with Asp1044. The aromatic substitution at the urea group is a common feature of the active compounds and is also recognized as a pharmacophore feature in the CPH1.

The inactive compound 75 possesses an open quinozoline moiety. This compound interacts with Cys917 through a bifurcated H-bond and with Glu883 via a single H-bond. A π-π interaction between the pyrimidine ring and Phe916 is also observed. This inactive compound fails to express strong interactions due to the absence of urea and the subsequent groups which are important pharmacophore features. Though the two compounds bind to the same pocket, the DFG loop involves in different type of interaction mechanisms like a stronger σ-π interaction with active compound and a weaker hydrophobic interaction with inactive compound. This analysis is further extended to investigate the binding free energy and its decomposition in a per-residue basis to understand the role of active site residues in stabilizing the protein-inhibitor interactions.

Binding free energy

The binding free energy of complexes 51 and 75 calculated using the MM/GBSA method is –59.89 and –36.81 kcal/mol, respectively. The electrostatic, van der Waals, non-polar and polar components of binding free energy are listed in Table 2. Among the calculated energy components of the active and inactive compounds, the E_{vdW}

value is highly negative (–60.88 and –47.22 kcal/mol, respectively) and explains the significant role of van der Waals interaction during inhibition. The electrostatic interaction –35.57 and –24.96 kcal/mol, respectively) is the second important factor influencing the binding free energy, which is almost half of the van der Waals interaction. The non-polar interaction (–7.7 and –5.96 kcal/mol, respectively) derived by the solvent accessible surface area contributes about 5 times lesser than the electrostatic interaction. The highly positive polar interaction (44.26 and 41.34 kcal/mol, respectively) screens the favorable electrostatic interaction into a net electrostatic energy of 8.69 and 16.38 kcal/mol, respectively. The strength of favoring binding free energy components involved in VEGFR-2 inhibition by both active and inactive compound is ordered as $\Delta E_{vdw} > \Delta E_{ele} > \Delta G_{nonpol} > \Delta G_{pol}$. Overall, this observation clearly shows that the inhibition activity is highly modulated by van der Waals interactions.

Decomposition of the binding free energy

The decomposition of binding free energy on a per residue basis of the VEGFR-inhibitor complexes provides a detailed insight into the role of non-covalent interactions on VEGFR-2 inhibition. This quantitative information could reveal the significance of interacting residues and provide useful insights in designing new VEGFR-2 specific inhibitors.

Figure 6A shows a plot of total interaction energy of the inhibitor-residue pairs. It is clear that both 51 and 75 bind to the same binding cavity lined by the residues Leu838, Val846, Glu883, Val 897, Phe916, Cys917, Gly920, Cys1043, Asp1044, and Phe1045, in which the residue Glu883 interacts strongly when compared to others. Glu883 contributes total interaction energy of
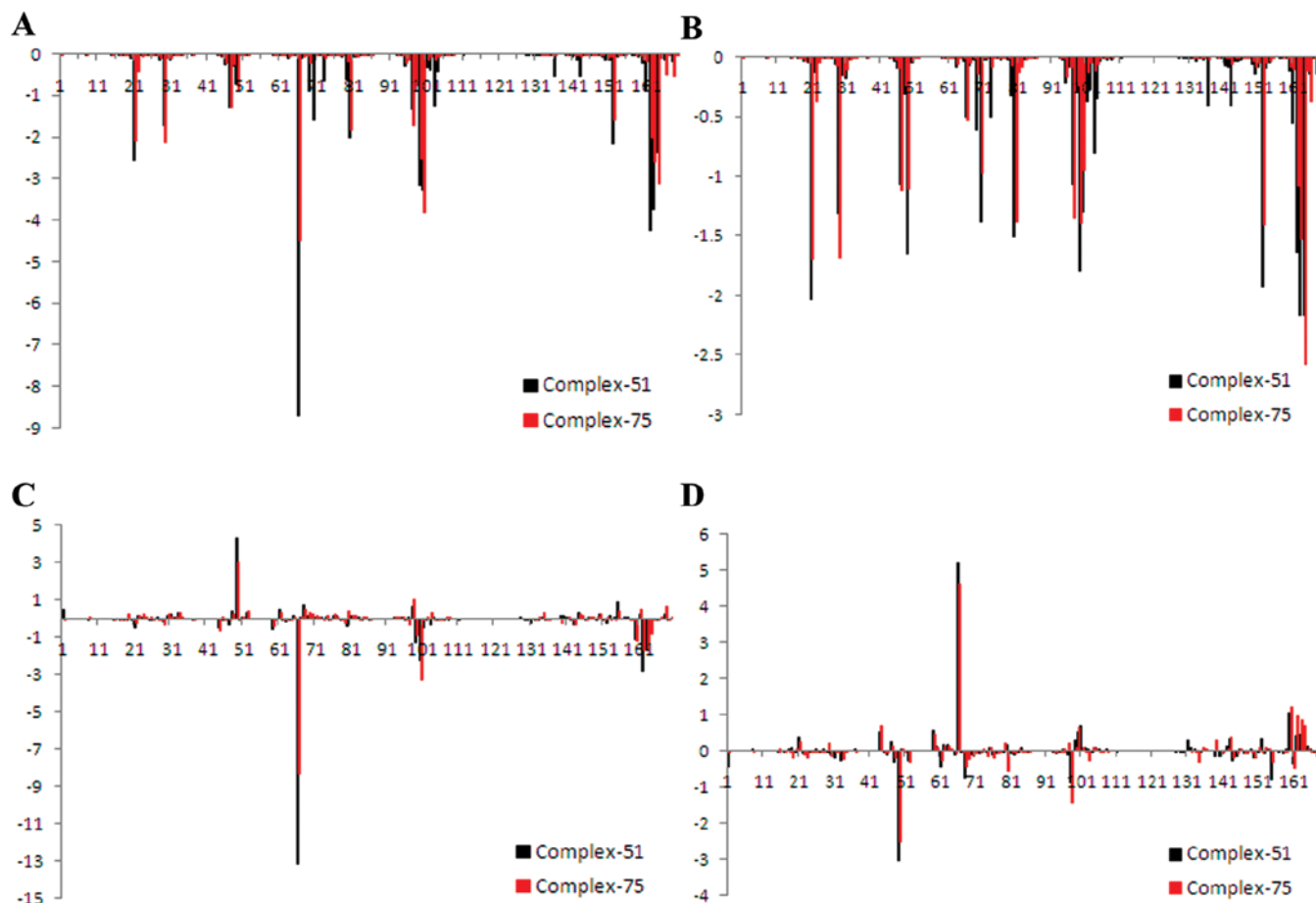
Figure 6. The decomposition of binding free energy into inhibitor-residue pairs. The contribution of total (A), vdW (B), electrostatic (C) and polar solvation (D) energies in the active (black lines) and inactive (red lines) complexes, respectively.

–8.72 kcal/mol when complexed with active compound and in the case of inactive complex, it is –4.51 kcal/mol. Such high energy of binding with 51 is attributed by the formation of a bifurcated H-bond whereas, 75 interacts via only one H-bond. The other residues such as Leu838, Glu883, Phe916, Gly920, Cys1043 and Asp1044 express significant interaction energies viz. –2.56, –8.72, –3.17, –1.25, –4.26 and –3.73 kcal/mol, respectively in the complex-51 than that with complex-75 (–2.08, –4.51, –2.51, –0.13, –2.02, –2.6 kcal/mol, respectively). In contrast to this, the residues Cys917 and Phe1045 show weaker interaction with 51 (–3.26 and –2.39 kcal/mol, respectively) than 75 (–3.81 and –3.14 kcal/mol, respectively).

The contribution of van der Waals energy for the binding of active (black lines) and inactive (red lines) compounds is shown in Figure 6B. The compounds 51 and 75 show van der waals interaction with the residues of VEGFR-2 like Leu838, Val846, Ala864, Lys866, Leu887, Val897, Val914, Phe916, Cys917, Leu1033, Cys1043, Asp1044 and Phe1045. The strength of vdW interaction is quite high in complex-51, when compared to complex-75 except the residues Val846, Val914 and Phe1045 (–1.69, –1.34 and –2.58 kcal/mol, respectively). In general, E_{vdw} shows favorable interaction with both the compounds, but with different amplitudes.

Electrostatic interaction plays a major role in determining the activity of a compound when complexed with the receptor. The role of electrostatic interactions in the inhibition of VEGFR-2 can be evidenced by H-bond and π-π interactions (Figure 6C). In this analysis, the residue Glu883 of VEGFR-2 is found to impart higher attractive energy (–13.17 kcal/mol) when complexed with 51 than 75 (–8.39 kcal/mol). The other favoring sites such as Phe916, Cys917, Cys1043, Asp1044 and Phe1045 contribute –1.36, –2.28, –2.81, –1.74 and –0.55 kcal/mol, respectively in the complex-51. The contribution of these residues is comparatively less in the inactive complex, except Cys917, which contributes higher. The Lys866 contributes unfavorable electrostatic energy with both the compounds which is 1.25 kcal/mol higher in the active complex.

As the ligand binding site of VEGFR-2 also possesses polar residues, it is note-worthy to understand the role of polar interactions. The polar solvation free energy analyzed from the ΔG_{pol} reveals that the favorable electrostatic interactions are screened by the unfavorable polar solvation energies and vice versa (Figure 6D). Despite this polar screening, the net electrostatic contribution of the residues Glu883, Phe916, Cys917, Leu1033, Cys1043, Asp1044 and Phe1045 is considerably higher to influence electrostatic interaction.

Binding free energy analysis using molecular dynamics simulations finds merit in designing novel VEGFR-2 inhibitors based on the knowledge of receptor's binding specificity. Further decomposition of binding free energy into per residue analysis provides a clear understanding on the active participation of binding site residues in inhibiting receptor function. The active site of VEGFR-2 is hydrophobic in nature and its role in stabilizing the ligand is quantitatively highlighted by ΔE_{vdw} component of binding free energy, which is also envisaged by the per residue contribution. In addition to this interaction, the compound 51 forms four H-bonds and three π interactions with VEGFR-2 whereas the compound 75 interacts via three H-bonds and one π interaction. The amino acid residues Glu883, Phe916, Cys917 and Asp1044 that express strong electrostatic interactions play a major role in ligands activity. The identified acceptor and donor pharmacophoric features of CPH1 interact with these important residues via hydrogen bond. The compound 51 is identified as an active ligand since the urea group participates in explicit interactions (like hydrogen bonds and π-interactions) as well as encompasses the derived pharmacophoric features (A4 and D6) of CPH1. On the other hand, the features A4 and D6 are not accommodated by the inactive compound and this reveals the ability of CPH1 in discriminating active and inactive compounds.

Precisely, the present analysis explains the importance of binding site residues like Glu883, Phe916, Cys917 and Asp1044 and proposes to design the inhibitors that could establish electrostatic interactions with these residues for efficient inhibition against VEGFR-2. The generated phramacophore model (CPH1) also represents these interactions pattern and demonstrate as a statistically significant model with high predictivity and hence, could serve as a template model in screening databases to identify novel inhibitor scaffolds.

Conclusion

In this present work, an attempt has been made to derive a 3D-QSAR model for VEGFR-2 inhibitors using pharmacophore based alignment. A five featured hypothesis, AADRR has been identified as the best hypothesis as it distinguishes the active and inactive compounds. The H-bond donor-acceptor and hydrophobic properties identified using this pharmacophore model witnessed the role of favorable and unfavorable sites of interaction. The QSAR model built using this hypothesis showed statistical significance in determining the most active and inactive compounds and were docked into VEGFR-2. Molecular dynamics simulations for 7 ns were performed to investigate the structural dynamics and complex stability. The binding free energy analysis revealed that the binding affinity of compound 51 (–59.89 kcal/mol) with VEGFR-2 is more than compound 75 (–36.81 kcal/mol). The decomposition of free energy into ligand-residue interaction pairs revealed Glu883, Phe916, Cys917 and Asp1044

as the significant interacting residues. From Figure 6, it is clear that the complexation with VEGFR-2 is highly driven by van der Waals interactions and the only residue imparting electrostatic interaction is Glu883 (–13.17 kcal/mol). Thus, this pharmacophore modeling, docking and simulation studies highlight the important pharmacophoric features that mediate VEGFR-2 inhibition and provide further insights in designing new VEGFR-2 inhibitory scaffolds with higher inhibition activity.

Acknowledgement

Declaration of interest

The authors report no declarations of interest.

References

1. Rahimi N. Vascular endothelial growth factor receptors: molecular mechanisms of activation and therapeutic potentials. Exp Eye Res 2006;83:1005–1016.
2. Iwata H, Imamura S, Hori A, Hixon MS, Kimura H, Miki H. Biochemical characterization of a novel type-II VEGFR2 kinase inhibitor: comparison of binding to non-phosphorylated and phosphorylated VEGFR2. Bioorg Med Chem 2011;19:5342–5351.
3. Ferrara N, Gerber HP, LeCouter J. The biology of VEGF and its receptors. Nat Med 2003;9:669–676.
4. Schenone S, Bondavalli F, Botta M. Antiangiogenic agents: an update on small molecule VEGFR inhibitors. Curr Med Chem 2007;14:2495–2516.
5. Holmes K, Roberts OL, Thomas AM, Cross MJ. Vascular endothelial growth factor receptor-2: structure, function, intracellular signalling and therapeutic inhibition. Cell Signal 2007;19:2003–2012.
6. Manley PW, Bold G, Brüggen J, Fendrich G, Furet P, Mestan J et al. Advances in the structural biology, design and clinical development of VEGF-R kinase inhibitors for the treatment of angiogenesis. Biochim Biophys Acta 2004;1697:17–27.
7. Roskoski R Jr. VEGF receptor protein-tyrosine kinases: structure and regulation. Biochem Biophys Res Commun 2008;375:287–291.
8. Kendall RL, Rutledge RZ, Mao X, Tebben AJ, Hungate RW, Thomas KA. Vascular endothelial growth factor receptor KDR tyrosine kinase activity is increased by autophosphorylation of two activation loop tyrosine residues. J Biol Chem 1999;274:6453–6460.
9. Leppänen VM, Prota AE, Jeltsch M, Anisimov A, Kalkkinen N, Strandin T et al. Structural determinants of growth factor binding and specificity by VEGF receptor 2. Proc Natl Acad Sci USA 2010;107:2425–2430.
10. Franklin MC, Navarro EC, Wang Y, Patel S, Singh P, Zhang Y et al. The structural basis for the function of two anti-VEGF receptor 2 antibodies. Structure 2011;19:1097–1107.
11. Miyazaki Y, Matsunaga S, Tang J, Maeda Y, Nakano M, Philippe RJ et al. Novel 4-amino-furo[2,3-d]pyrimidines as Tie-2 and VEGFR2 dual inhibitors. Bioorg Med Chem Lett 2005;15:2203–2207.
12. Hasegawa M, Nishigaki N, Washio Y, Kano K, Harris PA, Sato H et al. Discovery of novel benzimidazoles as potent inhibitors of TIE-2 and VEGFR-2 tyrosine kinase receptors. J Med Chem 2007;50:4453–4470.

13. Lee K, Jeong KW, Lee Y, Song JY, Kim MS, Lee GS et al. Pharmacophore modeling and virtual screening studies for new VEGFR-2 kinase inhibitors. Eur J Med Chem 2010;45:5420–5427.

14. Morabito A, De Maio E, Di Maio M, Normanno N, Perrone F. Tyrosine kinase inhibitors of vascular endothelial growth factor receptors in clinical trials: current status and future directions. Oncologist 2006;11:753–764.

15. Huang S, Li R, Connolly PJ, Xu G, Gaul MD, Emanuel SL et al. Synthesis and biological study of 4-aminopyrimidine-5-carboxaldehyde oximes as antiproliferative VEGFR-2 inhibitors. Bioorg Med Chem Lett 2006;16:6063–6066.

16. Huang S, Li R, LaMontagne KR, Greenberger LM, Connolly PJ. 4-aminopyrimidine-5-carbaldehyde oximes as potent VEGFR-2 inhibitors. Part II. Bioorg Med Chem Lett 2011;21:1815–1818.

17. Kubo K, Shimizu T, Ohyama S, Murooka H, Iwai A, Nakamura K et al. Novel potent orally active selective VEGFR-2 tyrosine kinase inhibitors: synthesis, structure-activity relationships, and antitumor activities of N-phenyl-N'-{4-(4-quinolyloxy)phenyl} ureas. J Med Chem 2005;48:1359–1366.

18. Dhanachandra Singh Kh, Karthikeyan M, Kirubakaran P, Nagamani S. Pharmacophore filtering and 3D-QSAR in the discovery of new JAK2 inhibitors. J Mol Graph Model 2011;30:186–197.

19. Dixon SL, Smondyrev AM, Rao SN. PHASE: a novel approach to pharmacophore modeling and 3D database searching. Chem Biol Drug Des 2006;67:370–372.

20. John S, Thangapandian S, Sakkiah S, Lee KW. Discovery of potential pancreatic cholesterol esterase inhibitors using pharmacophore modelling, virtual screening, and optimization studies. J Enzyme Inhib Med Chem 2011;26:535–545.

21. Ebalunode JO, Zheng W, Tropsha A. Application of QSAR and shape pharmacophore modeling approaches for targeted chemical library design. Methods Mol Biol 2011;685:111–133.

22. Yang SY. Pharmacophore modeling and applications in drug discovery: challenges and recent advances. Drug Discov Today 2010;15:444–450.

23. Khedkar SA, Malde AK, Coutinho EC, Srivastava S. Pharmacophore modeling in drug discovery and development: an overview. Med Chem 2007;3:187–197.

24. Langer T, Wolber G. Pharmacophore definition and 3D searches. Drug Discov Today. 2004;1:203–207.

25. Gilson MK, Zhou HX. Calculation of protein-ligand binding affinities. Annu Rev Biophys Biomol Struct 2007;36:21–42.

26. Hu G, Wang D, Liu X, Zhang Q. A computational analysis of the binding model of MDM2 with inhibitors. J Comput Aided Mol Des 2010;24:687–697.

27. Stoica I, Sadiq SK, Coveney PV. Rapid and accurate prediction of binding free energies for saquinavir-bound HIV-1 proteases. J Am Chem Soc 2008;130:2639–2648.

28. Rastelli G, Del Rio A, Degliesposti G, Sgobba M. Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. J Comput Chem 2010;31:797–810.

29. Hou T, Wang J, Li Y, Wang W. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. J Chem Inf Model 2011;51:69–82.

30. Hu G, Zhang Q, Chen LY. Insights into scFv:drug binding using the molecular dynamics simulation and free energy calculation. J Mol Model 2011;17:1919–1926.

31. Muthukumaran R, Sangeetha B, Amutha R, Mathur PP. Development of anti-HIV activity models of lysine sulfonamide analogs: a QSAR perspective. Curr Comput Aided Drug Des 2012;8:70–82.

32. Dixon SL, Smondyrev AM, Knoll EH, Rao SN, Shaw DE, Friesner RA. PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results. J Comput Aided Mol Des 2006;20:647–671.

33. Banks JL, Beard HS, Cao Y, Cho AE, Damm W, Farid R et al. Integrated Modeling Program, Applied Chemical Theory (IMPACT). J Comput Chem 2005;26:1752–1780.

34. Papakyriakou A, Katsarou ME, Belimezi M, Karpusas M, Vourloumis D. Discovery of potent vascular endothelial growth factor receptor-2 inhibitors. ChemMedChem 2010;5:118–129.

35. 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;47:1739–1749.

36. Case DA, Cheatham TE 3rd, Darden T, Gohlke H, Luo R, Merz KM Jr et al. The Amber biomolecular simulation programs. J Comput Chem 2005;26:1668–1688.

37. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins 2006;65:712–725.

38. Wang J, Wang W, Kollman PA, Case DA. Automatic atom type and bond type perception in molecular mechanical calculations. J Mol Graph Model 2006;25:247–260.

39. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. J Comput Chem 2004;25:1157–1174.

40. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Kleins ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983, 79, 926–993.

41. Darden T, York D, Pedersen L. Particle mesh Ewald: An N. log(N) method for Ewald sums in large systems. J Chem Phys 1993, 98, 10089–10092.

42. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. J Comput Phys 1977, 23, 327–341.

43. Gohlke H, Case DA. Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. J Comput Chem 2004;25:238–250.

44. Onufriev A, Bashford D, Case DA. Exploring protein native states and large-scale conformational changes with a modified generalized born model. Proteins 2004;55:383–394.