OJB™
Online Journal of Bioinformatics©

8 (1) : 99-114, 2007

Binding Modes, binding Affinities and ADME Screening of HIV-1 NNRTI Inhibitor: Efavirnez and its analogues.

Sengupta D, Verma D, Naik PK

Department of Bioinformatics and Biotechnology, Jaypee University of Information Technology, Waknaghat, Distt.-Solan, Himachal Pradesh, India (Pin-173215).

ABSTRACT

Sengupta D, Verma D, Naik PK Binding Modes, Binding Affinities and ADME Screening of HIV-1 NNRTI Inhibitor: Efavirnez and its analogues, Online Journal of Bioinformatics, 8(1):99-114, Synthetic analogues of Efavirnez have been used to create efficient safer anti-HIV drugs. Forty seven analogues using combinatorial design with structural modifications at X, Y and R of the parent Efavirnez structure are herein described. Molecular interactions and binding affinities with Reverse Transcriptase 1 (RT) using docking-MM-GB/SA screening based on ADME properties are illustrated. Results showed that these analogues docked in a similar position and orientation on the active site of RT. A linear correlation (r^2 = 0.9948) was observed between the calculated free energy of binding (FEB) and pIC_{50} for the inhibitors, suggesting that the docked structure orientation and interaction energies were accurate. Three H-bonds between Efavirnez analogues and RT were observed. The electrostatic energy estimated by GB/SA predicted binding affinity (R^2 = 17.2 %). However, few Efavirnez analogues showed high binding affinity and activity with RT compared with the co-crystallized compound. This work describes modifications to the X, Y and R substitutes in Efavirenz.

Key words: Reverse transcriptase, Efavirnez, Docking, Glide, FEB, pIC_{50}, ADME

INTRODUCTION

The Reverse Transcriptase of Human Immunodeficiency Virus type 1 (HIV-1 RT) is a unique enzyme in the virus, which transcribes a single-stranded viral RNA genome into double-stranded DNA, which is subsequently integrated into the host cell genome by an integrase enzyme (Arnold et al., 1991; Le Grice SFJ, 1993; Goff P, 1990; Whitcomb JM and Hughes SH, 1992). As a consequence, it is an important target for developing effective drugs against Acquired Immune Deficiency Syndrome (AIDS). RT is a protein dimer consisting of two related chains: 66kDa (p66) and 51kDa (p51) (Smerdon SJ et al., 1994; Garg R et al., 1999; Arnold E et al., 1996; Larder BA, 1993; De Clercq E, 1995; Ding J et al., 1997; Pedersen OS and Pedersen EB, 1999). The p66 and p51 subunits are composed of subdomains, which are named thumb, palm, fingers and connection. The polymerase domain and the RNaseH domains are located on the p66 subunit. Several drugs that inhibit this enzyme have been approved by the FDA and are clinically used to treat AIDS (Huang H et al., 1998). Two categories of inhibitors targeting the enzyme RT have been developed. One is a nucleoside analogue, which can be incorporated into DNA as a nucleoside to stop the strand extension further. The other is a non-nucleoside RT inhibitor (NNRTI) (Tantillo C et al., 1994; Kohlstaedt LA et al., 1992; Smerdon SJ et al., 1994; Ding J et al., 1995; Jacobo-Molina A et al., 1993), which binds in a hydrophobic pocket located in the palm area of p66 sub domain of RT. Unfortunately; the virus rapidly develops resistance to the existing drugs through mutation. Thus, it is necessary to find new and more effective drugs that remain active across these virus mutations.

Efavirnez and other NNRTIs have demonstrated good activity to inhibit RT. Several crystal structures of NNRTIs/RT complexes have been solvated now. They show that all NNRTIs bind at the non-nucleoside inhibitor binding site in the p66 palm subdomain. However, the binding mode of many other NNRTI derivatives including efavirenz analogues is not yet known. Although it is assumed that these also bind in a similar mode, there are subtle differences among them (Koup RA et al., 1991; Richman D et al., 1991). Thus, the exploration of the binding mode will provide valuable information to understand the inhibition mechanism of the inhibitors to RT and to help in the design of more potent inhibitors.

Modern approaches for finding new leads for therapeutic targets are increasingly based on 3-Dimensional information about receptors. An effective way to predict binding structure of a substrate in its receptor is docking simulation, which has been successfully used in many applications. Some docking methods have demonstrated promising power to predict a reasonable binding structure. Combinations of the method with other methods, such as MD simulation, free energy binding calculation, comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) enable to get a lot of insights on biological systems and to help rational drug design. Several ways to calculate free energy of binding (FEB) have been suggested and used in different applications. Jorgensen et al. (2000) have successfully applied Monte Carlo and Linear Response Equation (LRE) on many systems to calculate binding affinities. Wang et al. (2001) developed MM-PBSA and apply the method in the predication of activity of 12 TIBO-like inhibitors. In a free energy calculation, the accurate predictions of solvation and entropy contributions to a binding process are always challenging. Many works have shown that Poisson-Boltzmann and generalized-Born models are good ways to estimate the electrostatic part of solvation effect in a binding process. Although some work demonstrates that normal mode analysis can be used to estimate the entropy effect in a process, it is very time-consuming.

Recognition by the pharmaceutical industry that undesirable absorption, distribution, metabolism and excretion (ADME) properties of new drug candidates are the cause of many clinical phase drug development failures (Smith PA et al., 2004). This has resulted in a paradigm shift to identify such problems early in the drug discovery process. Thus, in vitro approaches are now widely used to investigate the ADME properties of new chemical entities and, more recently, computational (*in silico*) modeling has been investigated as a tool to optimize selection of the most suitable candidates for drug development.

In the present work Efavirnez and its 47 other structural derivatives were used to study the binding modes and binding affinities in the receptor; also being analyzed was the free energy of binding. We try to use a flexible docking (Glide) approach to predict the "preferable" binding structure of a ligand in RT. To study the association of the ligands with the receptor further, we use the automated mechanism of Multi-Ligand Bimolecular Association with Energetics (eMBrAcE). It uses traditional MM methods to calculate ligand-receptor interaction energies (*G*ele, *G*vdW), with a Gaussian Smooth Dielectric Constant function method (GBSA) for electrostatic part of solvation energy (Guvench O, 2002; Still WC, 1990, Qiu D et al., 1997) and solvent-accessible surface for the nonpolar part of solvation energy (Hasel W et al., 1988). The approach is simple, fast and straightforward. It benefits the calculation of relative binding affinity needed to evaluate the activity of large set of molecules in rational drug design. The final screening of the efavirnez & its analogs for the probable absorption, distribution, metabolism and excretion (ADME) properties are calculated using Qikprop program designed by Professor William L. Jorgensen (2000) (Schrödinger).

Figure 1 shows the binding site of RT complexed with efavirnez (1JKH) (Ren J et al., 2001). However, the effectiveness of other derivatives of efavirnez is yet to be determined. The need to determine their binding structures in the active site of RT and explore the interactions for these new RT analogues is essential in order to improve the design of second generation inhibitors. Hence, in this study we have set out to study the binding mode using docking approach and evaluating their binding affinities using MM-GB/SA and overall screening based on their ADME properties, leading to successful drug development.

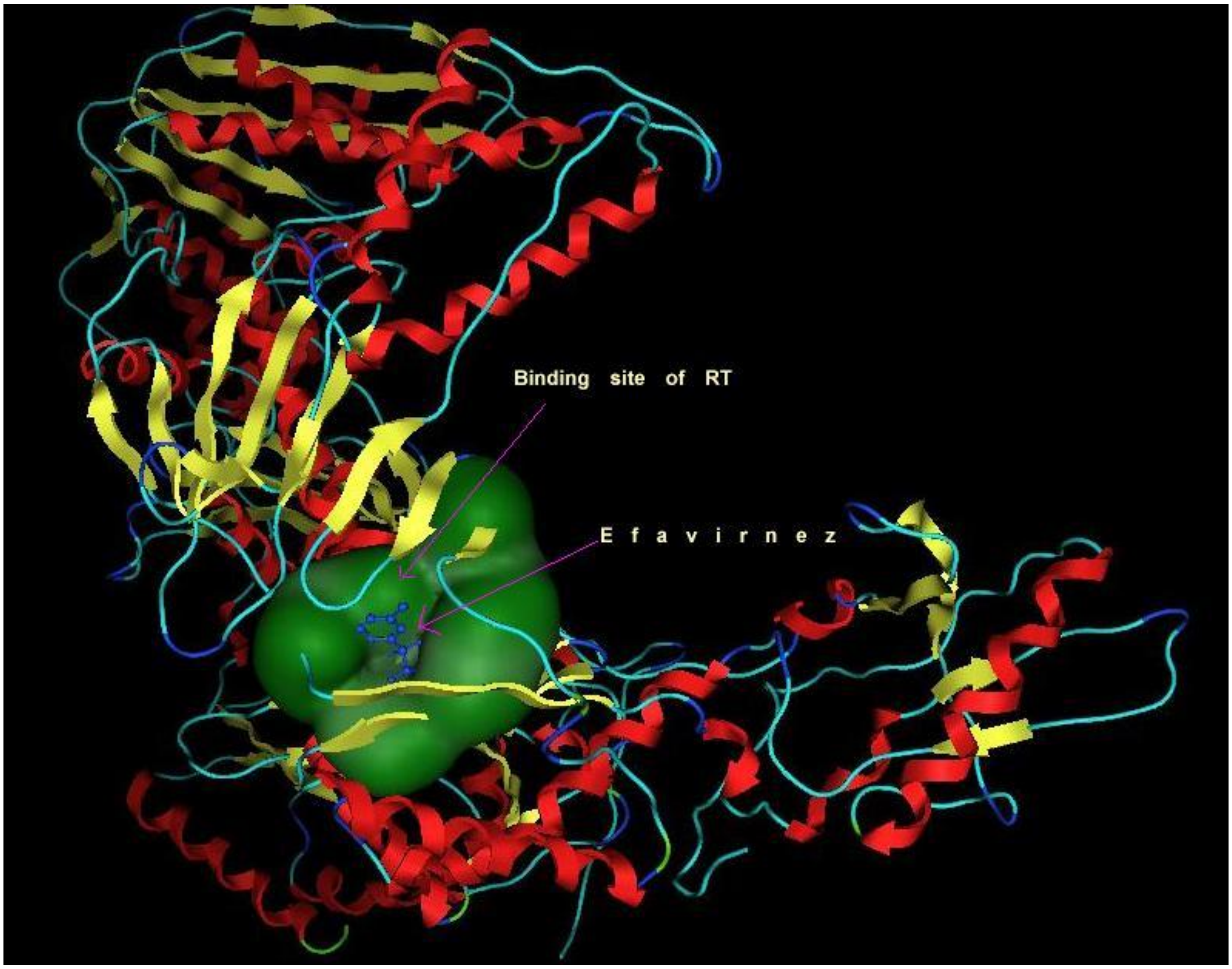


Figure 1. Binding Site of Reverse Transcriptase 1 complexed with efavirnez.

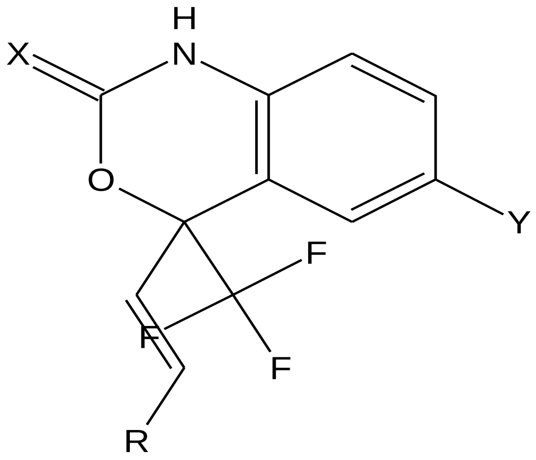
MATERIALS AND METHODS

Preparation of protein target structure: The starting coordinates of the HIV-1 RT [1JKH] was taken from the Protein Data Bank (www.rcsb.org) and further modified to be used for Glide docking (Schrödinger). The complex was imported to maestro, the co-crystallized ligands were identified and removed from the structure and the protein was minimized using the protein preparation wizard (shipped by Schrödinger) by applying OPLS_2005 force field by application of the autoref.pl script. Progressively weaker restraints (tethering force constants 3, 1, 0.3, 0.1) were applied to non-hydrogen atoms only. This refinement procedure is recommended by Schrödinger (technical notes for version 1.8), because Glide uses the full OPLS-AA force field at an intermediate docking stage and is claimed to be more sensitive towards geometric details than other docking tools. Water molecules were removed and H-atoms were added to the structure. Most likely positions of hydroxyl and thiol hydrogen's; protonation states and tautomers of His residues; and Chi "flip" assignments for Asn, Gln and His residues were selected by the protein assignment script. Minimizations were performed until the average root mean square deviation of the non-hydrogen atoms reached 0.3Å.

Preparation of compound libraries**:** The coordinates of efavirnez was obtained from the X-ray structure (1JKH). The other inhibitors for the target protein HIV-1 RT, p66 domain were built using the efavirnez as a template. The analogues library was generated by modifying the respective functional groups with stearically and conformatically allowed substituents using reagent database and combinatorial design module (Schrödinger) (**Table 1, next page**). Each structure was assigned an appropriate bond order using ligprep script shipped by Schrödinger. The inhibitors were converted to mae format (Maestro, Schrödinger Inc.) and optimized by means of the MMFF94 force field using default setting.

Glide Docking & Scoring function**:** Glide calculations were performed with Impact version v18007 (Schrödinger, Inc.) (Halgren TA et al., 2004; Krovat EM et al., 2005; Friesner RA et al., 2004). It performs Grid-based ligand docking with energetics and searches for favorable interactions between one or more typically small ligand molecules and a typically larger receptor molecule, usually a protein. Schrödinger recommends the performance of test calculations with different scaling factors for the receptor and ligand atom van der waal radii, because steric repulsive interactions might otherwise be overemphasized, leading to rejection of overall correct binding modes of active compounds. After ensuring that protein and ligands are in correct form for docking, the receptor-grid files were generated using grid-receptor generation program. To soften the potential for nonpolar parts of the receptor, we scaled van der waal radii of receptor atoms by 1.00 with partial atomic charge 0.25. A grid box of size 56 x 56 x 56 Å with coordinates X = 3.4860, Y = -36.804 and Z = 22.3858 was generated at the centroid of the active site consisting of residues; Pro-95, Leu-100, Lys-101, Val-106, Val-179, Cys-181, Tyr-188, Trp-229 (the residue information obtained from the literature) and the size of ligands to be docked was selected from the workspace. . The ligands were docked with the active site using the "xtra precision" Glide algorithm. Glide generates conformations internally and passes these through a series of filters. The first places the ligand center at various grid positions of a 1 Å grid and rotates it around the three Euler angles. At this stage, crude score values and geometric filters weed out unlikely binding modes. The next filter stage involves a grid-based force field evaluation and refinement of docking solutions including torsional and rigid-body movements of the ligand. The OPLS-AA force field is used for this purpose. A small number of surviving docking solutions can then be subjected to a Monte Carlo procedure to try to minimize the energy score. The final energy evaluation is done with GlideScore and a single best pose is generated as the output for a particular ligand.

Table 1. Library of HIV-1 RT Inhibitory Activity of Efavirnez analogues used in this Work.



(Parental structure of Efavirnez)

Analogues	**X**	**Y**	**R**	**Analogues**	**X**	**Y**	**R**
S1(Original)	O	Cl	CYCLOPROPYL	S25	O	H	CYCLOPROPYL
S2	O	Cl	CYCLOPROPYL METHYL	S26	O	H	CYCLOPROPYL METHYL
S3	O	Cl	DIMETHYL ALLYL	S27	O	H	DIMETHYL ALLYL
S4	O	Cl	CH2-COO=CH2	S28	O	H	CH2-COO=CH2
S5	O	Cl	CH2-C(CH3)=CH2	S29	O	H	CH2-C(CH3)=CH2
S6	O	Cl	CH2C(CH=CH2)=CH2	S30	O	H	CH2C(CH=CH2)=CH2
S7	O	Cl	CH2COOCH3	S31	O	H	CH2COOCH3
S8	O	Cl	CH2CH2CH3	S32	O	H	CH2CH2CH3
S9	O	Cl	CH2CH2CH2CH3	S33	O	H	CH2CH2CH2CH3
S10	O	Cl	CH2CH2CH=CH2	S34	O	H	CH2CH2CH=CH2
S11	O	Cl	CH2CH=CHCH3	S35	O	H	CH2CH=CHCH3
S12	S	Cl	CH2C(CH3)=CHCH3	S36	O	H	CH2C(CH3)=CHCH3
S13	S	Cl	CYCLOPROPYL	S37	S	H	CYCLOPROPYL
S14	S	Cl	CYCLOPROPYL METHYL	S38	S	H	CYCLOPROPYL METHYL
S15	S	Cl	DIMETHYL ALLYL	S39	S	H	DIMETHYL ALLYL
S16	S	Cl	CH2-COO=CH2	S40	S	H	CH2-COO=CH2
S17	S	Cl	CH2-C(CH3)=CH2	S41	S	H	CH2-C(CH3)=CH2
S18	S	Cl	CH2C(CH=CH2)=CH2	S42	S	H	CH2C(CH=CH2)=CH2
S19	S	Cl	CH2COOCH3	S43	S	H	CH2COOCH3
S20	S	Cl	CH2CH2CH3	S44	S	H	CH2CH2CH3
S21	S	Cl	CH2CH2CH2CH3	S45	S	H	CH2CH2CH2CH3
S22	S	Cl	CH2CH2CH=CH2	S46	S	H	CH2CH2CH=CH2
S23	S	Cl	CH2CH=CHCH3	S47	S	H	CH2CH=CHCH3
S24	S	Cl	CH2C(CH3)=CHCH3	S48	S	H	CH2C(CH3)=CHCH3

GScore = a * vdW + b * Coul + Lipo + Hbond + Metal + BuryP + RotB + Site
Where, vdW => van der Waals energy; Coul => Coulomb energy; Lipo => Lipophilic contact term; HBond => Hydrogen-bonding term; Metal => Metal-binding term; BuryP => Penalty for buried polar groups; RotB => Penalty for freezing rotatable bonds; Site => Polar interactions in the active site; and the coefficients of vdW and Coul are: a = 0.065, b = 0.130.

MM and Binding Free Energies: For the calculation of free energy of binding (FEB) of the ligands with RT only the glide-XP docking results have been taken and only the best scoring pose for each ligand was taken into consideration. Bimolecular Association with Energetics (eMBrAcE) developed by Schrödinger was used for physics based rescoring procedure (Guvench *et al.,* 2002}. For each ligand, the protein-ligand complex ($E_{\text{lig-prot}}$), the free protein (E_{prot}), and the free ligand (E_{lig}) were all subjected to energy minimization in implicit solvent (generalized Born). eMBrAcE uses the OPLS-AA all-atom force field with the surface generalized Born implicit solvent model (Wu X et al., 2003; Todorov NP et al., 2003). It uses traditional MM methods to calculate ligand-receptor interaction energies (*Gele, GvdW, Gsolv*), with a Gaussian Smooth Dielectric Constant Function method (GB/SA) (Reynold's CH, 1995) for electrostatic part of solvation energy and solvent-accessible surface for the nonpolar part of solvation energy. A conjugate gradient minimization protocol was used in all minimization. eMBrAcE minimization calculations were performed using Energy Difference Mode, in which energy changes upon association are estimated, taking as input the complexes obtained after docking analysis (Glide outputs). The energy difference was calculated using the equation:

$$\blacktriangle E = Ecomplex - Eligand - Eprotein$$

The full effects of relaxation and solvation are also included in this mode.

ADME screening: The QikProp (Duffy EM and Jorgensen WL, 2000) program has been used to obtain the absorption, distribution, metabolism, and excretion (ADME) properties of the analogues. It predicts both physically significant descriptors and pharmaceutically relevant properties. All the analogues were neutralized before being used by Qikprop. The neutralizing step is essential, as in normal mode QikProp is unable to neutralize a structure and no properties will be generated. The program was processed in normal mode, predicting 44 properties for the 48 molecules which consisted of principal descriptors and the physiochemical properties with a detailed analysis of the log P (Octanol/Water), QP%, and log HERG. It also evaluates the acceptability of the analogues based on the Lipinski's rule of 5 which is essential for rational drug design.

RESULTS AND DISCUSSION

Molecular docking of efavirnez and its analogues: The original crystal structure of 1JKH was used to validate the Glide-XP Dock for the HIV-1 RT system. This was done by moving the inhibitor outside of active site entrance and then docking it back into the active site. The top 5 configurations after docking were taken into consideration to validate the result. The RMSD was calculated for each one in compare to the co-crystallized efavirnez and the value was found in between 0.431 – 1.012 (**Table 2, next page**).

Table 2. Docking result of five lowest configurations of Efavirnez in Reverse Transcriptase (1JKH).

Rank	Ligand No.	Glide Score	Emodel*	Rank	Ligand No.	Glide Score	Emodel*
1	S12	-19.01	-63.5	25	S39	-15.73	-58.2
2	S6	-18.97	-63.5	26	S42	-15.71	-57.6
3	S2	-18.90	-61.3	27	S48	-15.64	-54.2
4	S27	-18.88	-63.4	28	S38	-15.29	-49.1
5	S30	-18.78	-63.3	29	S41	-15.27	-53.9
6	S26	-18.72	-52.8	30	S47	-15.18	-53.9
7	S5	-18.62	-58.8	31	S45	-15.15	-55.3
8	S9	-18.56	-61.2	32	S24	-15.14	-58.8
9	S3	-18.56	-60.2	33	S18	-15.05	-55.6
10	S11	-18.53	-57.9	34	S46	-15.03	-54.2
11	S36	-18.51	-56.3	35	S37	-14.95	-50.3
12	S10	-18.50	-59.8	36	S15	-14.83	-54.7
13	S34	-18.42	-58.2	37	S44	-14.80	-49.8
14	S35	-18.30	-59.3	38	S43	-14.78	-62.4
15	S29	-18.28	-58.7	39	S14	-14.69	-48.6
16	S33	-18.22	-58.9	40	S23	-14.63	-54.9
17	S1	-18.18	-55.2	41	S40	-14.60	-52.0
18	S7	-18.14	-68.1	42	S22	-14.46	-54.5
19	S8	-18.02	-56.5	43	S21	-14.43	-52.5
20	S4	-17.84	-56.5	44	S17	-11.53	-50.5
21	S32	-17.70	-54.5	45	S19	-11.35	-58.5
22	S28	-17.46	-54.1	46	S20	-11.28	-50.9
23	S35	-17.24	-47.7	47	S16	-11.26	-50.6
24	S31	-16.91	-65.1	48	S13	-11.20	-44.4

*Emodel is a specific combination of GScore, CvdW, and the internal torsional energy of the ligand conformer

This proved that the docked molecules were bound with similar orientation and conformation with the active site in RT.

To study the molecular basis of interaction and affinity of binding of the efavirnez and its analogues, all the ligands have been docked into the active site of RT. The docking result of these ligands is given in **Table 3**.

Table 3. The docking results of Efavirnez and its analogues in the Original Crystal Structure of RT (1JKH) using Glide-XP (kcal/mol).

Configuration	Glide Score (kcal/mol)	[a]E (kcal/mol)	[b]ΔE (kcal/mol)	[c]RMSD (Å)
1	-18.18	-64.5	0.0	0.430991
2	-18.1	-63.6	-0.9	0.545687
3	-18.1	-62.6	-1.9	0.618864
4	-12.71	-21.7	-42.8	0.919901
5	-12.53	-21.1	-43.4	1.012559

[a]E = a specific combination of GScore, CvdW, and the internal torsional energy of the ligand conformer.

[b]$\Delta E = E_i - E_{lowest}$

[c]RMSD between docked and crystallographic structures

Table 3. The docking results of Efavirnez and its analogues in the Original Crystal Structure of RT (1JKH) using Glide-XP (kcal/mol).

The ranking of ligands was done based on the glide score. Out of 48 docking simulations, 46 have more than 70 accepted poses and the energy difference among them is also very small (±2.15KJ/mol). The result demonstrates that the docking simulation can dock all the efavirnez analogues into the same binding site as well (**Figure 2**).

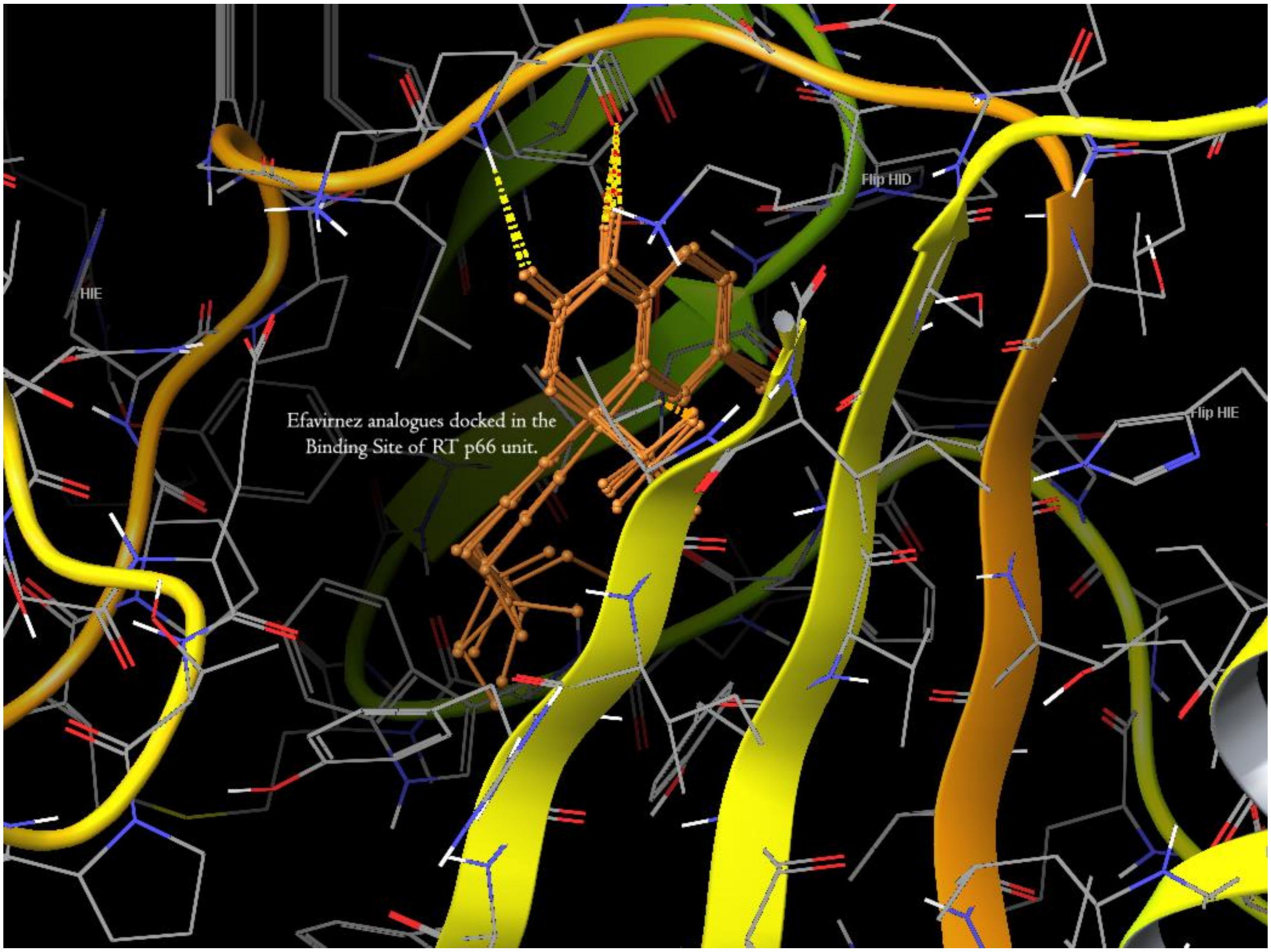


Figure 2. Showing the docked efavirnez analogues in the active site of Reverse Transcriptase 1.

By superposing the best binding configuration (with the lowest glide score), of all the analogues it is seen that these analogues bind in the same orientation and similar position in terms of the common structure (double-ring part) (**Figure 3, next page**).

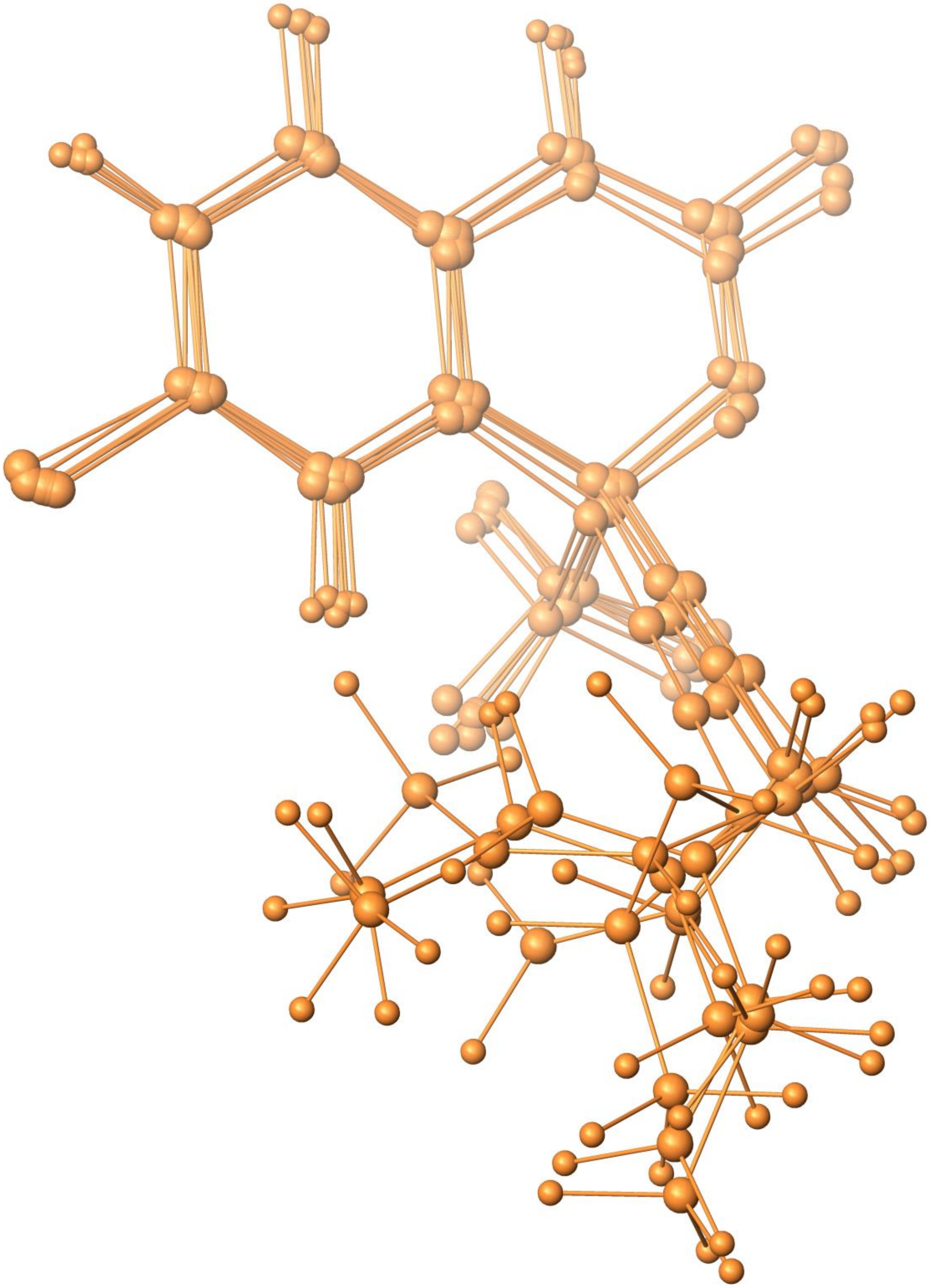


Figure 3. Superposition of six lowest energy configurations of efavirnez analogues.

As these molecules have the same backbone structure of the double -ring, it is obvious that they bind in similar pattern in NNRT1 active site of RT in nature. All the 47 efavirnez analogues were found to be good binder with RT. The docking score using Glide varies from -11.20 to a minimum of -19.01, whereas the original efavirenz (complexed with RT) inhibitor has docked score of -18.18. This proved that its analogues could be the potential drugs for second generation drug development. The amino acids putatively involved in the binding of efavirnez analogues Leu-100, Lys-101,Cys-181, Tyr-188, Trp-229, which could collectively contribute three stabilizing hydrogen bonds to the drug **(Figure 4).**

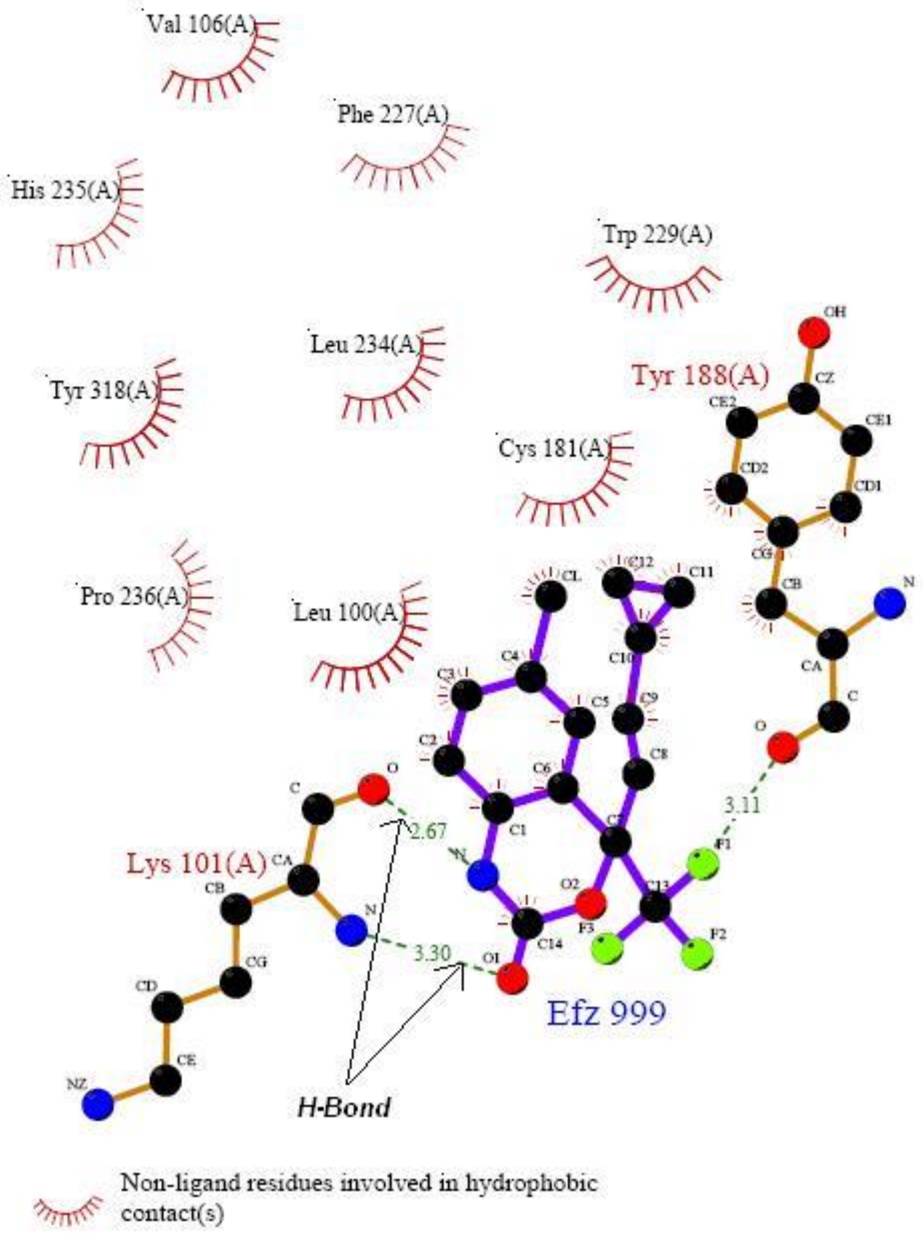


Figure 4. Ligand plot of efavirnez analogue showing the hydrogen bonding with the residues of Reverse Transcriptase 1.

Calculated free energy of binding versus activity: One docking structure with best poses (lowest glide score) from each molecule docking result was taken into consideration for calculation of free energy of binding (FEB) using eMBrAcE (Schrödinger) and predicting IC_{50}. The interaction energy was calculated after a minimization was performed on a docked ligand in which atoms within 7.5Å from the ligand were free to move (other atoms were fixed). The interaction energy includes an implicit solvation (H_2O) term. A vdW, solvation and electrostatic energy as well as solvent accessible surface area (SASA) were calculated for each minimized complex (**Table 4).**

Table 4. Calculated energies and estimated free energy of binding (FEB) of Efavirnez & its analogues (KJ/mol).

Analogues	Gvdw	Gele	Gsolv	SASA	Gcald[1]	pIC_{50}[2]	Analogues	Gvdw	Gele	Gsolv	SASA	Gcald[1]	pIC_{50}[2]
S1	-58.55	18.24	13.94	503.05	-26.37	10.57	S25	-66.52	31.05	-9.85	473.87	-32.32	12.95
S2	-49.77	-0.70	29.29	528.71	-21.18	8.49	S26	-51.55	-28.22	43.29	499.61	-36.48	14.62
S3	-61.85	18.74	13.59	551.51	-29.52	11.83	S27	-56.35	-28.32	43.14	520.92	-32.53	13.04
S4	-74.38	18.78	9.79	502.38	-25.81	10.34	S28	-75.75	18.63	5.10	475.67	-32.02	12.83
S5	-52.80	-5.56	24.73	521.84	-33.63	13.48	S29	-61.32	33.82	-9.13	494.42	-36.63	14.68
S6	-59.39	17.12	12.97	536.93	-29.30	11.74	S30	-49.66	8.82	5.84	508.01	-35.00	14.03
S7	-72.86	-21.73	32.28	540.62	-33.31	13.35	S31	-71.97	-46.92	57.41	498.20	-35.48	14.22
S8	-71.29	-2.63	38.42	516.36	-35.50	14.23	S32	-65.52	27.45	-11.85	488.83	-33.74	13.52
S9	-78.29	-7.69	42.02	548.39	-33.96	13.61	S33	-71.01	-5.47	26.90	521.50	-29.58	11.85
S10	-73.40	6.36	20.97	536.48	-33.07	13.25	S34	-51.27	-21.33	48.56	508.01	-24.04	9.63
S11	-67.42	-29.65	49.10	536.76	-31.97	12.81	S35	-56.67	2.26	27.30	510.83	-27.11	10.86
S12	-75.46	-22.61	45.12	548.71	-35.10	14.07	S36	-55.29	-27.48	11.56	520.30	-33.21	13.31
S13	-63.99	2.49	27.00	517.59	-34.50	13.83	S37	-65.98	-17.32	34.01	491.48	-29.29	11.74
S14	-74.86	-21.63	29.28	542.49	-35.21	14.11	S38	-50.23	-9.23	21.54	516.37	-23.92	9.59
S15	-79.73	9.98	13.16	569.64	-26.59	10.66	S39	-76.13	-0.05	29.32	545.28	-26.86	10.76
S16	-54.85	-28.50	46.04	520.22	-37.31	14.95	S40	-40.65	-12.52	32.47	495.95	-20.70	8.29
S17	-73.88	-20.73	32.22	543.14	-33.39	13.38	S41	-58.76	15.66	20.17	512.10	-22.93	9.19
S18	-57.02	7.87	6.42	558.40	-22.73	9.11	S42	-71.63	-13.52	31.87	532.87	-33.28	13.34
S19	-69.86	-20.43	30.28	540.01	-34.01	13.63	S43	-62.55	4.69	21.65	516.50	-36.21	14.51
S20	-64.68	35.65	3.51	531.32	-25.52	10.23	S44	-61.96	21.72	17.29	505.56	-22.95	9.20
S21	-61.83	36.35	0.99	564.04	-24.49	9.81	S45	-76.32	-16.85	43.29	537.81	-34.88	13.98
S22	-53.27	-3.43	21.08	553.82	-35.62	14.28	S46	-36.06	5.06	3.18	525.14	-27.82	11.15
S23	-61.13	14.82	2.77	554.46	-23.54	9.43	S47	-56.35	20.38	7.40	530.21	-28.57	11.45
S24	-70.04	-31.25	47.87	562.18	-38.42	15.40	S48	-55.09	11.24	22.28	534.23	-21.57	8.64

1. Calculated free energy of binding, Δ*G*cald is calculated from optimized linear combination of Δ*G*ele, Δ*G*vdW, Δ*G*solv, and SASA from regression.
2. Predicted pIC50 is estimated from Δ*G*cald using the following relationship:
   Δ*G*binding = RTlnKdissociated ≈ RTlnIC50 = -RTpIC50, where 300Kelvin is used in the work for temperature T.

A similar scheme to linear response was used to develop a free energy of binding (FEB) relationship based on these energies which in turn used to predict the activity (pIC_{50}) of efavirnez analogues. Theoretically, FEB can be partitioned into several components: vdW, electrostatic, solvation and entropy energy term. The entropy contribution is most difficult to calculate. However, several methods (Wang W et al., 2001; Zhou Z, 2003) have been suggested to estimate the entropy contribution. To relative rigid molecules, the entropy is relatively small and normally is ignored or cancelled in relative free energy calculation. Further, in the rational drug design, the calculation of relative FEB rather than absolute FEB is important. Several works have been reported, in which reasonable correlation between calculated FEB and activity for a small set of molecules (Zhou Z, 2003). The plot of the FEB and log ($1/pIC_{50}$) reveals a significant relationship ($R^2 = 0.9948$) between these two parameters (**Figure 5).** The linear trend in the plot indicated that the docking calculation produces reasonable binding modes. Based on correlation study, it is seen that electrostatic energy (Gele) has most significant correlation to the activity (pIC_{50}) (R^2 = 17.2 %) followed by vdW energy (R^2 = 12.5%), solvation energy (R^2 = 7.9 %) and SASA has less significant correlation (R^2 = 0.1%) to the activity. It indicates that in the binding of efavirnez analogues, electrostatic energy estimated by GB/SA may be a major driving force to their binding and contribution to their activity. The calculated free energy of binding (Gcald) among the ligands varies in between -20.70 to -38.42 KJ/mol and the overall difference is also very small (±5.07 KJ/mol).It revealed that all these molecules bind in RT with high affinity and showed activity (pIC_{50}) in between 8.29 and 15.40. The co-crystalized ligand efavirnez having FEB -26.37 KJ / mol and pIC_{50} 10.57 proved to be less potent than that its ligands taken into consideration.

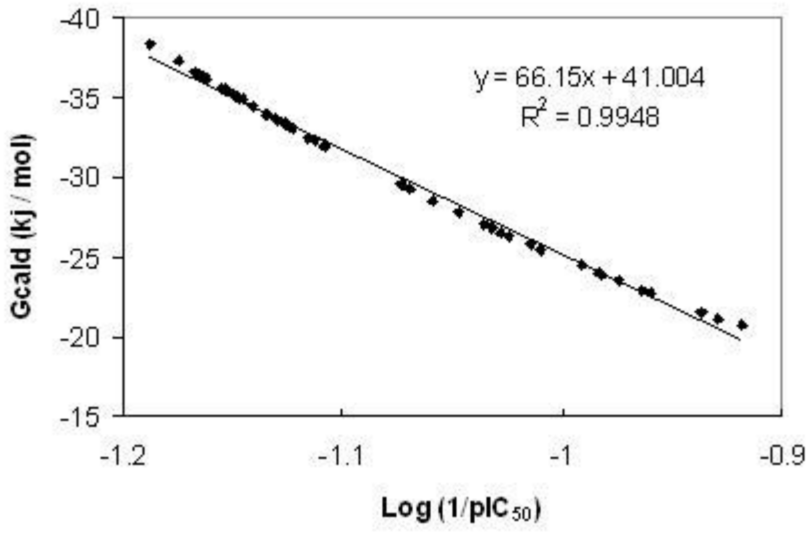


Figure 5. Linear regression plot between Gcald and log($1/pIC_{50}$) of 48 efavirnez analogues used in the study.

ADME screening: We have analyzed 44 physically significant descriptors and pharmaceutically relevant properties of efavirnez & its analogues, among which were molecular weight, polarizability (Å3), log P (octanol/gas), log P (water/gas), log P (octanol/water), log BB (brain/blood), log P MDCK, log Kp (skin permeability), log Khsa (serum protein binding) & their screening in accordance to Lipinski's rule of 5. For the log P (octanol/water), QP%, and log HERG, if the value for a utilized descriptor exceeded the range for the experimental training set, it was flagged. In this study out of 48 ligands, 29 structures showed significant values for the properties analyzed and showed drug like characteristic's based on Lipinski's rule of 5. Whereas, analogues like S13-S18, S20-S24, S38, S39, S41, S42, S45-S48 showed a violation each for drug like characteristics (**Table 5**).

Table 5. Screening of ADME properties for Efavirnez and its analogues using Qikprop simulation

Ligand No.	QP logPo/w	QP %	Log HERG	QPP Caco	QPP MDCK	Rule Of 5	Ligand No.	QP logPo/w	QP %	Log HERG	QPP Caco	QPP MDCK	Rule Of 5
S1	3.53	86.04	-4.44	1724.73	6864.04	0	S25	3.04	89.05	-4.48	1722.24	2789.61	0
S2	3.89	86.01	-4.66	1721.66	6847.34	0	S26	3.40	89.02	-4.72	1720.09	2784.20	0
S3	4.12	86.29	-4.27	1734.96	6099.60	0	S27	3.75	88.67	-4.49	1726.40	2982.43	0
S4	3.64	88.42	-4.61	1729.98	7077.92	0	S28	3.17	91.49	-4.73	1728.16	2947.71	0
S5	4.05	85.98	-4.73	1729.98	7215.59	0	S29	3.42	90.81	-4.62	1728.75	2758.37	0
S6	3.71	85.95	-4.56	1726.50	7198.77	0	S30	3.72	92.88	-4.96	1728.37	2806.17	0
S7	2.79	74.64	-4.72	619.78	2379.31	0	S31	2.31	77.56	-4.73	619.91	968.40	0
S8	4.20	89.83	-4.82	1728.83	6596.84	0	S32	3.21	88.93	-4.59	1726.54	2928.79	0
S9	3.90	87.74	-4.51	1727.98	6559.06	0	S33	3.56	88.91	-4.77	1726.25	2932.88	0
S10	4.04	88.82	-4.92	1739.21	7257.08	0	S34	3.55	91.81	-5.00	1738.43	2991.51	0
S11	3.98	86.90	-4.62	1729.94	6881.51	0	S35	3.53	89.82	-4.89	1726.92	3023.95	0
S12	4.19	85.99	-4.35	1730.53	6689.55	0	S36	3.75	88.93	-4.69	1727.85	3006.88	0
S13	5.13	100	-4.63	6135.66	10000	1	S37	4.63	100	-4.68	6132.04	10000	0
S14	5.50	100	-4.84	6127.38	10000	1	S38	5.00	100	-4.91	6129.09	10000	1
S15	5.73	100	-4.47	6162.80	10000	1	S39	5.37	100	-4.76	6144.91	10000	1
S16	5.24	100	-4.79	6151.55	10000	1	S40	4.76	100	-4.92	6148.62	10000	0
S17	5.50	100	-4.69	6150.65	10000	1	S41	5.02	100	-4.84	6147.47	10000	1
S18	5.79	100	-4.98	6152.93	10000	1	S42	5.33	100	-5.16	6148.89	10000	1
S19	4.37	88.56	-4.88	2198.79	10000	0	S43	3.87	91.47	-4.90	2197.04	10000	0
S20	5.31	100	-4.74	6150.19	10000	1	S44	4.81	100	-4.79	6151.05	10000	0
S21	5.66	100	-4.90	6153.53	10000	1	S45	5.16	100	-4.95	6149.30	10000	1
S22	5.64	100	-5.08	6171.77	10000	1	S46	5.15	100	-5.18	6165.25	10000	1
S23	5.57	100	-4.73	6145.12	10000	1	S47	5.12	100	-5.05	6141.21	10000	1
S24	5.79	100	-4.49	6155.52	10000	1	S48	5.37	100	-4.87	6148.37	10000	1

QP logPo/w - QP log P for octanol/water. **QP % -** % Human Oral Absorption in GI (+-20%) (<25% is poor). **log HERG** - HERG K+ Channel Blockage (concern below -5). **QPP Caco** - Apparent Caco-2 Permeability (nm/sec) (<25 poor, >500 great). **Qpp MDCk** - Apparent MDCK Permeability (nm/sec) (<25 poor, >500 great). **Rule of 5 -** Lipinski Rule of 5 Violations (maximum is 4).

Biology indications of docking structure and ADME screening: The natural and prepared compounds (Table 1) were evaluated using docking approach and their binding energy with RT. The docking results show that the structurally homologous inhibitors bind in a very similar position and orientation in RT, which suggest that the homologous inhibitors have similar binding patterns and interaction modes in RT, and further have similar inhibitory mechanism. Furthermore, the most potent inhibitor should have the best interaction with RT.

The docking structures of all compounds showed that they bind in very similar pattern with the active site of RT, as it is evident for the superposition of 6 most effective analogues: for better resolution all the analogues have not been superposed in the figure. The calculated FEB of the 47 efavirnez analogues in RT-based docked structures demonstrates a linear correlation ($R^2 = 0.9944$) with their log ($1/pIC_{50}$) value. This concludes that the structural modification implemented in this study is significantly related to their activity. Also this proved the reasonability and reliability of the docking results. It can be seen that substitution of functional groups at position X, Y & R leads to increase in binding affinity of modified analogues even more intense than that of co-crystalized ligand. ADME studies provided a peered analysis for the final selection of the potential candidates from the compound library generated showing effective docking score & binding modes. Based on the overall analysis we can say analogues S12 & S6 are the most potent analogues which could be used for second generation of drug development. Both have exhibited effective binding in the active site of RT, showed an optimal predicted IC_{50} values and even qualify the Lipinki's rule of five.

CONCLUSIONS

A FEB calculation on the binding affinity of 48 non-nucleoside HIV-1 RT inhibitors consisting of efavirnez and its 47 analogue structures is described. The binding structures of these ligands in RT were predicted by flexible docking simulations. The docking result demonstrated that the docking simulation could satisfactorily reproduce a binding structure from a crystal structure of a RT/efavirnez complex. Superposition of the binding structure of whole set of ligands from docking simulations shows that these structurally similar ligand bind in a very similar pattern in RT. They bind in a similar position inside the NNRTI active site of RT to try to fit the binding pocket well.

The calculated FEB for these ligands reasonably predicted the activity of this set of ligands. The result shows that the linear combination of four energy terms: vdW, electrostatic, solvation (electrostatic part), and nonpolar energies optimized by regression has power to express the binding affinity of large set of ligands in receptor. The Dock-MM-GBSA combination demonstrates a good ability on the binding structure prediction and binding energy determination to produce reasonable energies. The GBSA method predicted a reasonable solvation energy terms to enable a satisfactory FEB expression was build. In the work, it is noticed that among these energy terms, the solvation effect (electrostatic and nonpolar) plays a major role on the determination of activity of NNRTI's.

This work suggests that in the relative FEB calculation, which is major interest in drug design, the contribution of different energy terms can be scaled by a set of weight factors to reach a good correlation. In practice, it is know that same energy term plays different role in different type of systems. This is one of reasons that a reasonable activity model can be gotten just based on some energy terms.

The calculation of solvation effect upon a ligand binding in a protein is a challenge work. This work and many others have shown that solvation effect is any important driving force on ligand binding and a key factor in expression of activity of a set of ligands. In the work, GB and SASA methods were used to estimate the electrostatic and the nonpolar parts of solvation and produced satisfactory results in terms of good correlation with experimental activity. LogPo/w is another way to express solvation effect and has succeeded used in many applications. To find a better way directly to calculate solvation energy is still remained as an important and challenge work.

REFERENCES

Arnold, E., Das, K., Ding, J., Yadav, P., Hsiou, Y., Boyer, P. L., Hughes,S. H. (1996). Targeting HIV reverse transcriptase for anti-AIDS drug design: structural and biological considerations for chemotherapeutic strategies, Drug Design and Discovery, 13, 29-47.

De Clercq E (1995). Antiviral therapy for human immunodeficiency virus infections. Clinical Microbiology Reviews, 8, 200-239.

Ding, J., Das, K., Moereels, H., Koymans, L., Andries, K., Janssen, P. A.,Hughes, S., Arnold, E. (1995). Structure of HIV-1 RT/TIBO R 86183 reveals similarity in the binding of diverse non nucleoside inhibitors, Nature Structural and Molecular Biology, 2, 407-415.

Ding, J., Das, K., Tantillo, C., Zhang, W., Jr., A. D. C., Jessen, S., Lu, X., Hsiou, Y., Jacobo-Molona, A., Andries, K., Pauwels, R., Moereels, H., Koymans, L., Janssen P., Smith, R., Koepke, M. K., Michejda, C., Hughes, S. H., Arnold, E. (1995). Structure of HIV-1 reverse transcriptase in a complex with the nonnucleoside inhibitor a-APA R 95845 at 2.8Å resolution, Structure, 3, 365-379.

Ding J et al (1997). Structural studies of HIV-1 reverse transcriptase and implications for drug design, Structure-Based Drug Design, 41-82.

Duffy EM and Jorgensen WL (2000). Prediction of Properties from Simulations: Free Energies of Solvation in Hexadecane, Octanol, and Water, Journal of the American Chemical Society, 122, 2878-88.

Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D.T., Repasky, M. P., Knoll, E. H., Shaw, D. E., Shelley, M., Perry, J. K., Francis, P., Goff, P. (1990). Retroviral reverse transcriptase: Synthesis, structure and function, J. Acquired Immune Deficiency Syndromes, 3, 817-831.

Garg, R., Gupta, S. P., Gao, H., Babu, M. S., Debnath, A. K., Hansch, C. (1999). Comparative Quantitative Structure minus sign activity relationship studies on Anti HIV Drugs, Chemical Reviews, 99, 3525-3602.

Guvench, O (2002). Application of the Frozen Atom Approximation to the GB/SA Continuum Model for Solvation Free Energy, Journal of Computaional Chemistry, 23, 214 221.

Halgren, T. A., Murphy, R. B., Friesner, R. A., Beard, H. S., Frye, L. L., Pollard, W.T., Banks, J. L. (2004). Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening, Journal of Medicinal Chemistry, 47, 1750-1759.

Hasel, W., Hendrickson, T. F., Still, W. C. (1988). A Rapid Approximation to the Solvent Accessible Surface Areas of Atoms, Tetrahedron Comput. Method., 1, 103.

Huang, H., Chopra, R., Verdine, G. L., Harrison, S. C. (1998). Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance, Science, 282, 1669-1675.

Jacobo-Molina, A. and Arnold, E. (1991). HIV reverse transcriptase structure-function relationships, Biochemistry, 30, 6351-6361.

Jacobo-Molina, A., Ding, J., Nanni, R. G., Jr., A. D. C., Lu, X., Tamtilo, C., Williams, R. L., Kamer, G., Ferris, A. L., Clark, P., Hizi, A., Hughes, S. H. (1993). Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double stranded DNA at 3.0Å resolution shows bent DNA, Proceedings of the National Academy of Sciences, 90, 6320 -6324.

Jorgensen, W.L., Kroeger, S., Lamb, M.L., Tirado-Rives, J., Michejda, C.J., Ruby, S.K., Smith, R.H. (2000). Monte Carlo calculations on HIV-1 reverse transcriptase complexed with the non-nucleoside inhibitor 8-Cl TIBO: contribution of the L100I and Y181C variants to protein stability and biological activity, Protein Engineering, 13, 413-421.

Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A., Steitz, T. A. (1992), Crystal structure at 3.5Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor, Science, 256, 1783-1790.

Koup, R. A., Merluzzi, V. J., Hargrave, k. D., Adems, J., Grozinger, K., Eckner, R.J., Saivan, J. L. (1991). Inhibition of human immunodeficiency virus type1 replication by the dipyridocliazepinone BI-RG-587, Journal of infectious Diseases, 163, 966-970.

Krovat, E. M., Steindl, T., Langer, T. (2005). Recent Advances in Docking and Scoring, Current Computer Aided Drug Design, 1, 93-102.

Larder BA (1993). Inhibitors of HIV reverse transcriptase as antiviral agents and drug Resistance, Coldspring Harbor Laboratory Press: Plainview, NY, pp 163-191.

Le Grice, S.F.J. (1993). Human immunodeficiency virus reverse transcriptase, Cold spring Harbor Laboratory Press: Plainview, NY, 163-191.

Pauwels, R., Janssen, P. A. J., Arnold, E. (1994). Locations of anti-AIDS drug binding sites and resistance mutations in the three-dimensional structure of HIV-1 reverse transcriptase: implications for mechanisms of drug inhibition and resistance, Journal of Molecular Biology, 243, 369- 387.

Pedersen OS and Pedersen EB (1999). Non-nucleoside reverse transcriptase inhibitors, Antiviral Chemistry and Chemotherapy, 10, 285-314.

Qiu, D., Shenkin, P.S., Hollinger, F.P., Still, W.C. (1997). The GB/SA Continuum Mode for Solvation. A Fast Analytical Method for the Calculation of Approximate Born Radii", Journal of Physical Chemistry A, 101, 3005-3014.

Ren J et al (2001), Structural mechanisms of drug resistance for mutations at codons 181 and 188 in HIV-1 reverse transcriptase and the improved resilience of second generation non-nucleoside inhibitors, Journal of Molecular Biology, 312,795-805.

Reynold's CH (1995), Estimating Liphophilicity using GB/SA Continum Solvation Model, J of Chemical Information. Comput. Sci., 35, 738-742.

Richman, D., Rosenthal, A. S., Shoog, M., Eckner, R. J., Chou, T. C., Sabo, J. P., Merluzzi, V. J. (1991). BI-RG-587 in active against zidovudine-pesisant human Immunodeficiency virus type 1 and synegistic with zidovudine, Antimicrobial Agents and chemotherapy, 35, 305-308.

Shenkin, P. S. (2004). Glide: A New Approach for Rapid, Accurate Docking and Scoring: Method and Assessment of Docking Accuracy", Journal of Medicinal Chemistry, 47, 1739-1749.

Smerdon, S. J., Jager, J., Wang, J., Kohlstaedt, L. A., Chirino, A. J., Friedman, J. M., Rice, P. A., Steitz, T. A. (1994). Structure of the binding site for nonnucleoside inhibitors of reverse transcriptase of human immunodeficiency virus type 1, Proceedings of the National_Academy of Sciences, 91, 3911-3915.

Smith PA et al (2004). Towards integrated ADME prediction: past, present and future directions for modeling metabolism by UDP-glucuronosyltransferases”, Journal of Molecular Graphics & Modeling, 22, 507-517.

Still WC (1990). Semianalytical Treatment of Solvation for Molecular Mechanics and Dynamics, Journal of the American Chemical Society, 112, 6127.

Tantillo, C., Ding, J.,Jacobo-Molina, A., Nanni, R. G., Boyer, P. L., Hughes, S. H.,
Todorov, N. P., Mancera, R. L., Monthoux, P. H. (2003). A new quantum stochastic tunneling optimization method for protein-ligand docking. Chemical Physics Letters Chemical Physics Letters, 369, 257-263.

Wang, W., Lim, W. A., Jakalian, A., Wang, J., Wang, J., Luo, R., Bayly, C. I., Kollman, P. A. (2001). An analysis of the interactions between the Sem-5 SH3 domain and its ligands using molecular dynamics, free energy calculation and sequence analysis, Journal of the American Chemical Society, 123, 3986-3994.

Wang, J., Morin, P., Wang, W., Kollman, P.A.(2001). Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA, Journal of the American Chemical Society, 123(22), 5221-30

Whitcomb, J.M. and Hughes, S.H. (1992). Retroviral reverse transcription and integration: Progress and problems, Annual Review of Cell Biology, 8, 275-306.

Wu, X., Milne, J. L. S., Borgnia, M. J., Rostapshov, A. V., Subramaniam, S., Brooks, B. R. (2003). A core-weighted fitting method for docking atomic structures into low-resolution maps: application to cryo-electron microscopy. Journal of Structural Biology, 141, 63-76.

Zhou Z (2003). Computational Studies on Interaction, Activity and Mechanism of HIV-1 RT Inhibitors and other with enzymes, PhD thesis, Center for Computational Science, Pittsburgh.

Zhou, Z., Fisher, D., Spidel, J., Greenfield, J., Patson, B., Fazal, A., Wigal, C., Moe, O. A., Madura, J. D. (2003). Kinetic and Docking Studies of the Interaction of Quinones with the Quinone Reductase Active Site, Biochemistry, 42, 1985-1994.

MAIN