In Silico Design: Those Accentuate Assembly of HIV-1 Capsid

Tamalapakula V, Balabadra SK, Munnaluri RK and Manga V*

Department of Chemistry, Osmania University, India

Submission: August 28, 2017; **Published:** October 24, 2017

*Corresponding author: Manga V, MMMC Group, Department of Chemistry, University College of Science, Osmania University, Hyderabad-500007, Telangana, India, Email: vijjulathamanga@gmail.com

Abstract

Considering the significance of RNA genome in the HIV infection, its release can be forbidden by inhibiting encapsulated Capsid (CA) "core". In this context, design of new molecules that interact with HIV-1 CA assembly was made by conducting 3D-QSAR and Pharmacophore studies on a stockpile of fifty eight molecules. The supramolecular interactions of these inhibitors with the receptors active site amino acids Glu 35, Gly 60, His62, Gln 63, Ala 65, and Val 136 has been characterized using docking studies. The combined five point Pharmacophore hypothesis AHHRR, and 3D-QSAR CoMFA (Comparative molecular field analysis) and CoMSIA (Comparative molecular similarity indices analysis) has shown good statistical partial least square factors. The association of generated 3D-QSAR and PHASE pharmacophore hypothesis has provided structural insights to explore new molecules with enhanced CA assembly activity.

Keywords: Human immunodeficiency Virus type-1 (HIV-1); Capsid protein (CA); Common pharmacophore hypothesis (CPHs); Pharmacophore alignment and scoring engine (PHASE); Comparative molecular field analysis (CoMFA); Comparative molecular similarity indices analysis (CoMSIA) emphasize

Introduction

Capsid protein (CA) of Human immunodeficiency Virus type-1 (HIV-1) is a fullerene cone accomplished with approximately 250 hexamers and 12 pentamers that plays a prominent role in the initial stages of post host cell entry and in virus assembly [1-4]. Proteolytic cleavage of Gag (Matrix, Capsid and Nucleo capsid) polyprotein releases CA which reassembles to form a conical "core" that encloses the viral RNA genome [5-7]. The assembly process of CA protein is associated with the N-terminal domain (CANTD residues 1–145) and C-terminal domain (CACTD residues 151-231). The CANTD aid in virion maturation by self associating into hexameric rings and incorporation of the cellular protein cyclophilin A (CypA). While CACTD contributes to Gag-Gag interactions [8-10]. Once the virus enters the cell, the core disassembles and releases ribonucleoprotien complex, resulting in the formation of reverse transcription complexes [11-14]. However, virions with deformed core structured are defective in initiating reverse transcription and exhibit reduced infectivity [15]. It is believed that the N-terminus of CA can refold to form a new γ-hairpin helix by cleavage of CA from matrix protein (MA), which is stabilized by a buried salt bridge between Proline-1 of CA and the carboxyl side chain of Asp51. The renovation from tubular to spherical forms can be induced by changing the pH from 7.0 to 6.8 [16,17]. Several small molecule inhibitors like CAP-1, NYAD-1 a peptide inhibitor, bevirimat, a triterpene derivative have been reported to interfere with CA function on binding to the N-terminal domain of CA in both its mature and immature forms [18-25]. Pharmacophore mapping was performed by using Pharmacophore Alignment and Scoring Engine (PHASE) to evaluate features necessary for the ligand to interact with a receptor [26]. The pharmacophore may be used as a query in searching 3D databases containing "drug like" small organic molecules and also aid in exploring and engineering of new scaffold with high potency [27]. To investigate the mode of recognition and interaction mechanism of HIV-1 CA inhibitors with CA protein, molecular docking studies are performed. Further 3D-QSAR studies were performed by using CoMFA and CoMSIA analysis.

Methodology

Ligand construction and preparation

A stockpile of fifty eight 1,5-dihydrobenzo[b][1,4]diazepine-2,4-dione derivatives [28-30] (Figure 1) with their experimental information were collected from the literature. Outlining the three dimensional structures of these ligands using Maestro build panel in Schrödinger Suite, all possible low energy states are generated at physiological pH range of 7 +/- 2 in the Ligprep module of Schrödinger.
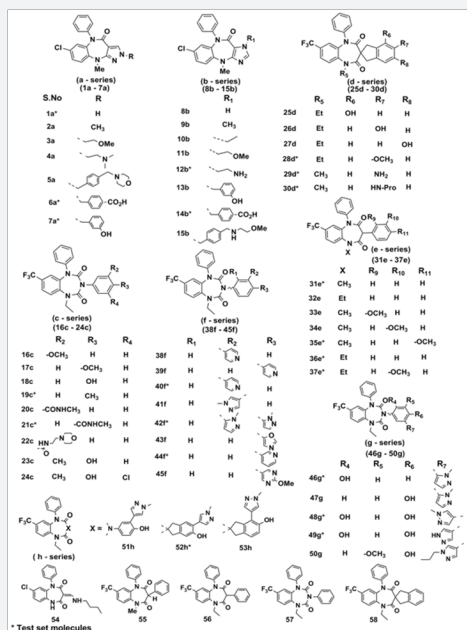
Figure 1: Structures of HIV-1 Capsid assembly inhibitors acting on N-terminal domain.

Generation of the common pharmacophore hypothesis (CPH)

The pharmacophore hypothesis and alignment were carried out by PHASE [31] (version 2.5, 2011; Schrodinger, LLC, New York, NY). Considering the significance of EC_{50} in biological activity [32,33] and we have converted E C_{50} into the corresponding pEC_{50} (-logE C_{50}), and randomly selected thirty five molecules as training set to generate Pharmacophore models. PHASE defines chemical features of ligand that facilitate non covalent bonding between the ligand and target receptor. PHASE models are created by applying Partial Least Squares (PLS) regression. The accuracy of the models improves with an increase in the number of PLS factors [34]. The PLS regression analysis was carried out using three PLS factors and possible statistical parameters are evaluated. The test set predictions were described as Q^2 root mean squared error and Pearson correlation coefficient (R) value. The regression value is calculated using the formula;

$$R^2 = \frac{\left[\sum \left(Y_i - \overline{Y_o}\right)\left(\overline{Y_i} - \overline{Y_p}\right)\right]}{\sqrt{\sum \left(Y_i - \overline{Y_o}\right)\sum \left(\overline{Y_i} - \overline{Y_p}\right)^2}}$$

Where $\overline{Y_o}$, $\overline{Y_p}$ are the average observed and predicted pEC50 values of the test set molecules.

(Figure1).

Docking studies

The X-ray structure of CANTD in complex with the benzodiazepine inhibitor (PDB: 4E91) [35] was retrieved from Protein Data Bank (http://www.rcsb.org/). Protein was prepared by using protein preparation wizard in Schrödinger. [36] GLIDE 5.6 (Grid-based ligand docking with energies) docking was performed by generating a grid (10 Å x10 Å x10 Å) around the active site of capsid assembly inhibitor. Rigid receptor docking protocols of Glide, namely Standard Precision (SP) and Extra Precision (XP) were employed to gain insight into the binding modes of all inhibitors. Further, to account for the receptor flexibility, a computationally intensive Induced Fit Docking (IFD) was performed. To validate the docking protocol, co-crystallized ligand benzodiazepinedione was re-docked and its atomic root mean square deviation (RMSD) was calculated [37-39].

Prime/MM-GBSA calculations

The relative binding free energies for best ranking molecules in XP mode are calculated by Molecular Mechanics Generalized Born Surface Area (MM/GBSA) using Prime. The dock pose features that are locally optimized features from glide were minimized from prime and complex energies were calculated using OPLS_2005 force field. The relative binding free energy ΔGbind was estimated according to equation.

ΔGbind=Ecomplex (minimized)-[Eligand (unbound, minimized) + Ereceptor (unbound, minimized)]

ΔGbind is the calculated relative free energy which includes both ligand and receptor strain energy. Ecomplex (minimized) is MM-GBSA energy of the minimized complex, Eligand (unbound, minimized) is the MM-GBSA energy of the ligand after removing it from the complex and allowing it to relax. Ereceptor (unbound, minimized) is the MM-GBSA energy of protein after separating it from the ligand.

CoMFA and CoMSIA Models

CoMFA and CoMSIA studies are performed by aligning data set molecules and are evaluated for thier steric, electrostatic and hydrogen bond effect on bioactivities using SYBYLX-2.1[40,41]. Gasteiger Huckle charges were assigned for all the inhibitors [42]. For CoMFA, the overlapped molecules were placed in a rectangular grid points separated by 2 Å. The van der Waals potential and columbic terms, representing the steric and electrostatic fields, were calculated using standard Tripos force field [43]. The regression analysis was carried out using set of variables, and cross validated the PLS method (leave-one-out). The final model (noncross-validated conventional analysis) was developed with the optimal number of components which poses the highest q^2 value.

For CoMSIA, five physicochemical properties namely steric, electrostatic, hydrophobic, hydrogen bond donor and acceptor were calculated. The CoMFA or CoMSIA descriptors were used as independent variables and pEC50 values as dependent variables in partial least square analysis. [34] The predictive correlation coefficient (r2pred) based on the test molecules, is computed

with the formula r^2pred = (SD-PRESS)/SD, where SD is the sum of the squared deviations between the biological activities of test set and mean activities of the training set molecules. PRESS is the sum of squared deviation between predicted and observed activity for each molecule in the test set.

ADME prediction

The ADME (absorption, distribution, metabolism and excretion) properties of designed molecules were evaluated computationally using Qik Prop module of Schrodinger [44]. The physically significant descriptors and pharmaceutically relevant properties of organic molecules with compliance to Lipinski's rule of five were calculated by Qik prop.

Results and Discussion

The hexameric architecture of capsid relay on three basic type of interactions. (i) CANTD-CANTD six fold symmetric interfaces that create the hexameric rings [45]. (ii) CANTD-CACTD intermolecular interface that reinforces the hexamer [8-10] and (iii) CACTD-CACTD dimeric interface that links the hexameric rings to form the lattice [46-50]. The current study is mainly focused on hexamer formation of CA by disrupting the intermolecular interaction of CANTD and hence in silico methods employed are restricted to N-terminal domain only. The X-ray crystal structure (4E91) of HIV-1 CANTD is a single chain containing the N-terminal domain. Sundquist reported intermolecular N-terminal domain interactions between Phe-32, Glu35, Gly60, His62, Gln63, and Ala65 residues of helices 1 and 2 [51]. The perturbation of intermolecular N-terminal domain interaction involves in disruption of NTD-NTD interaction and hence inhibits CA assembly [52].
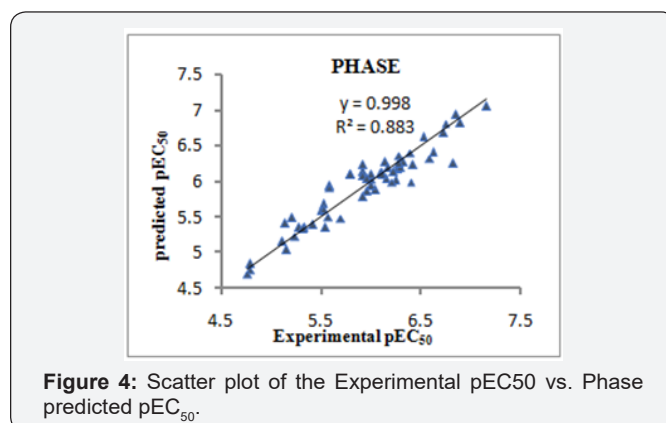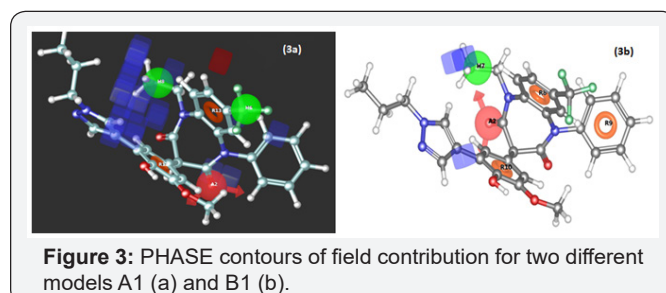
Pharmacophore generation

Figure 2: Pharmacophore hypothesis for best two models (2a) AHHRR (A1) and (2b) AHRRR (B$_1$).
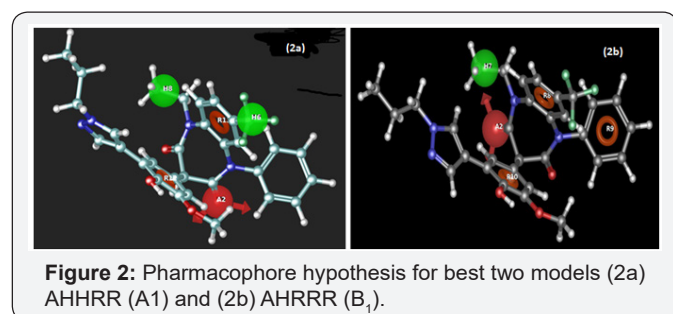
PHASE analysis for a set of 58 molecules derived five point CPHs belonging to AHHRR, AHRRR. Thirty five molecules in the training set were aligned on these CPHs and evaluated for PLS analysis using three PLS factors. The predictivity of each hypothesis was cross validated by the test set of twenty three molecules. The variants named AHHRR (model A$_1$) and AHRRR (model B1) in Figure 2 showed better statistical significance than AHHRR (A$_2$), AHHRR (A$_3$) and AHHRR (A$_4$). Since all the five pharmacophore hypothesis capitulate a statistically significant data as shown in Table 1, the hypothesis AHHRR (A$_1$) with good survival score 3.549, with significant R2 of 0.926 and Q^2 of

0.824 was considered in this work. From the above results it is confirmed that the obtained PHASE model ($r^2 > 0.5$, $q^2 > 0.6$, [(r^2-r0^2)/$r^2 < 0.1$, $0.85 \leq k \leq 1.15$ and rm2 > 0.5) was in the acceptable range.

The hydrogen bond acceptor A$_2$, and the phenyl ring R$_{12}$ was found to be significant for the activity. The carbonyl group at second position can form hydrogen bond interaction with His 62 residue in the active site, while the aromatic ring at third position contributes to hydrophobicity. The hypothesis AHHRR (A$_1$) was fit into the molecular skeleton of highest active molecule of capsid assembly inhibitor 50g. The distance and angles between the five features of the best two models AHHRR (A$_1$) and AHRRR (B$_1$) are shown in Table 1 & 2 (provided in supplementary data). The pharmacophore model implies that molecules which would fit the AHHRR (A$_1$) sites may consequence in stronger interactions with the active site residues. The distance and the angle maps of the best two models are represented in Figure S$_1$ given in supplementary data. The field contribution (blue and red cube contours) of the best two models are represented in Figure 3. The blue cubes represent that substitution of hydrophobic favored groups while red cubes represent disfavored region. Scatter plot of experimental and phase predicted pEC50 is shown in (Figure 2-4) (Table 1).


Figure 3: PHASE contours of field contribution for two different models A1 (a) and B1 (b).

Figure 4: Scatter plot of the Experimental pEC50 vs. Phase predicted pEC$_{50}$.

Table 1: Results of the best Pharmacophore hypothesis obtained through PHASE

Hypothesis	AHHRR (A1) (Model-1)	AHRRR (B1) (Model-2)	AHHRR (A2)	AHHRR (A3)	AHHRR (A4)
Survival score	3.549	3.191	3.037	2.970	2.316

SD	0.164	0.147	0.15	0.15	0.148
R2	0.930	0.942	0.940	0.940	0.942
F	134.3	148.9	189.8	189.3	195.3
P	7.587e-18	1.066e-20	4.114e-22	4.286e-22	2.539e-22
Q2	0.824	0.860	0.770	0.779	0.771
RMSE	0.216	0.264	0.2304	0.2265	0.2307
Pearson-R	0.926	0.922	0.887	0.889	0.885

A-Hydrogen bond acceptor, H- Hydrophobic/non-polar group, R-Aromatic ring, R2-correlation coefficient, F- variance ratio, Q^2-predicted activities for training set, SD- standard deviation of regression.

Molecular docking studies and Prime MM/GBSA calculations

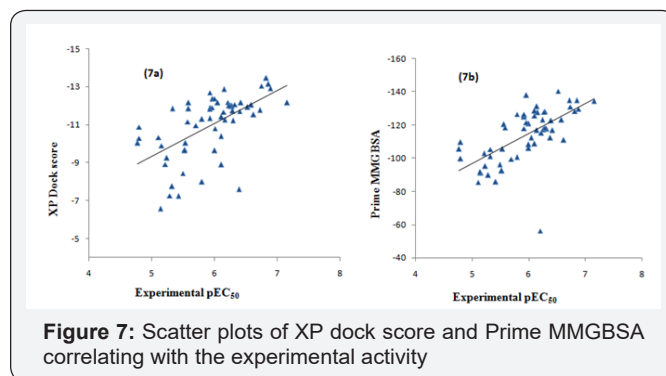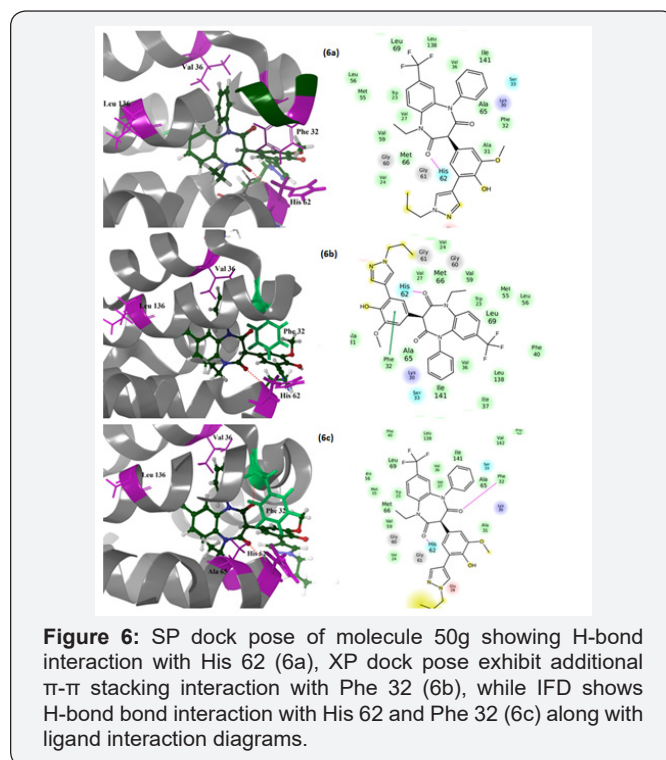**Figure 5:** Superimposition of dock poses to calculate its root mean square deviation.
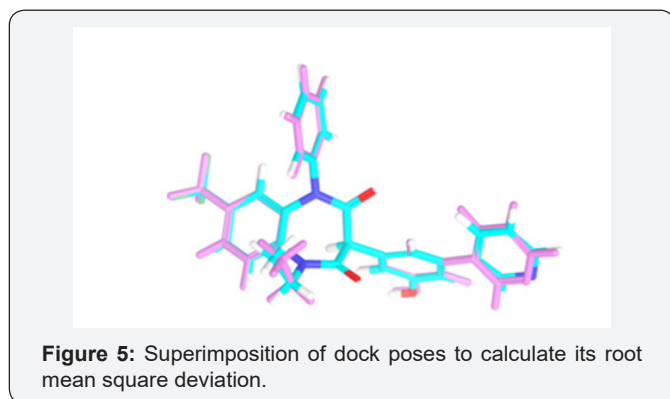
To evaluate the validity of docking protocol, root mean square deviation (RMSD) calculations were performed. The crystal ligand was haul out from the complex (PDB id: 4E91) and subsequently re-docked into the receptor. Best dock pose from docking studies differ from the original conformation by 0.216 Å, substantiating the robustness of docking protocol to enumerate the experimental binding mode. Figure 5 shows the overlay of the X-ray crystal structures (colored in cyan) of capsid assembly inhibitor and the re-docked pose (colored in plum) (Figure 5).

Fifty eight molecules with benzodiazepine scaffold which were reported to exhibit inhibitory effect on assembly process of HIV-1 capsid (Figure 1) are chosen for docking studies. The binding orientation of these has shown hydrogen bond interactions with Glu35, Gly60, His 62, Gln 63, Ala 65, Val 135 and Tyr 145 amino acid residues and are said to exhibit HIV-1 replication with a novel mechanism of action [52]. This compounds bind to the N-terminal domain (NTD) of the viral capsid protein (CA) and is believed to prevent its ensemble into conical core that is trivial for virion maturation and viral infectivity [53].

The hydrogen bond interaction with His 62 and Phe 32 residues of helices 1 and 2 are plays a crucial role since they hinder the hexameric lattice organization by interrupting intermolecular N-terminal interaction. SP dock pose of highest active molecule 50g display polar H-bond interaction with His 62 (Figure 6a), while XP dock pose exhibit additional π-π stacking interaction with Phe 32. (Figure 6b) The IFD dock pose for the same 50g molecule reveal polar hydrogen bonding interaction with Phe 32 and His 62 residues (Figure 6c). Application of different docking protocols i.e., SP, XP and IFD resulted in variable dock scores, of which IFD has shown more refined scores. Increase in dock scores in IFD protocol is possibly due to additional flexibility attributed to the protein, which allow more interactions with active site residues. However, some deviations have been observed in IFD scores of 45f and 53h which were lower than XP dock scores which might be due to steric clashes.

Figure 6: SP dock pose of molecule 50g showing H-bond interaction with His 62 (6a), XP dock pose exhibit additional π-π stacking interaction with Phe 32 (6b), while IFD shows H-bond bond interaction with His 62 and Phe 32 (6c) along with ligand interaction diagrams.

Figure 7: Scatter plots of XP dock score and Prime MMGBSA correlating with the experimental activity

The relative binding affinities of protein ligand complexes were assessed by Molecular mechanics with generalized born surface area (MM/GBSA). The ΔGbind value of highest active molecule 50g was found to be -134.50 that is slightly higher than the lead molecule 54 which was screened through CA assembly assay. It is imperative to evaluate statistical correlation between the different computational parameters and experimental pEC50 values to validate the applicability of in silico studies in predicting the HIV-1 CA inhibitors. These inhibitors showed a

statistically significant correlation of -0.579 for XP docking, -0.614 for prime MMGBSA (Figure 7). The list of inhibitors along with their pEC50, SP, XP and IFD dock scores, ΔGbind values from PRIME along with predicted CoMFA and CoMSIA values have been tabulated in Table 2.

Table 2: The list of inhibitors with their biological and predicted activities along with their dock scores (SP, XP, IFD) and binding energy (ΔGbind).

S.No	SP Dock score	XP Dock score	IFD score	ΔGbind	EC$_{50}$	pEC$_{50}$	PHASE Pred. activity	Pred. activity	
								CoMFA	CoMSIA
1a*	-9.38	-10.33	-13.39	-85.453			5.16	5.107	5.13
2a	-9.96	-9.89	-12.97	-91.272	7	5.154	5.04	5.044	5.106
3a	-10	-10.88	-13.4	-99.564	16	4.795	4.84	4.906	4.87
4a	-10.12	-10.29	-13.33	-109.74	16	4.795	4.76	4.819	4.825
5a	-10.89	-12.16	-14.83	-112.37	0.9	6.045	5.89	5.509	5.615
6a*	-9.86	-10.96	-13.17	-99.461	2	5.698	5.48	5.555	5.349
7a*	-11.21	-11.85	-14.87	-105.26	4.7	5.327	5.35	5.414	5.382
8b	-9.73	-7.24	-12.58	-85.714	3.8	5.42	5.39	5.443	5.372
9b	-9.33	-10.02	-12.54	-105.52	2.9	5.537	5.36	5.386	5.346
10b	-9.72	-7.26	-12.82	-89.899	5.2	5.283	5.36	5.334	5.319
11b	-10.05	-9.25	-13.09	-95.067	5.8	5.236	5.23	5.511	5.456
12b*	-9.64	-7.74	-13.47	-101.02	4.8	5.318	5.34	5.319	5.402
13b	-9.7	-6.58	-14.66	-91.524	7.3	5.136	5.41	5.393	5.42
14b*	-11.08	-8.91	-13.79	-102.97	6.1	5.214	5.49	5.257	5.189
15b	-10.11	-7.98	-13.89	-100.58	0.9	5.795	6.1	6.07	6.136
16c	-11.4	-12.69	-13.84	-125.01	1.2	5.92	6.13	5.846	5.943
17c	-10.83	-11.66	-14.53	-116.95	0.72	6.142	6.28	6.18	6.091
18c	-11.05	-12.16	-14.11	-116.95	0.38	6.42	6.24	6.184	6.396
19c*	-11.13	-12	-13.03	-117.25	0.54	6.267	6.17	6.246	6.019
20c	-11.26	-11.88	-14.07	-125.75	1.2	5.92	6.24	5.767	5.955
21c*	-10.87	-12.06	-13.83	-117.76	0.48	6.318	6.28	6.39	6.457
22c	-11.76	-12.36	-15.54	-138.11	1.1	5.958	5.86	5.956	6.009
23c	-11.25	-12.05	-14.36	-123.03	0.26	6.585	6.32	6.053	6.345
24c	-11.16	-13.45	-14.09	-128.25	0.15	6.823	6.25	6.243	6.336
25d	-11.48	-11.41	-13.65	-125.63	0.78	6.107	6.12	5.845	6.283
26d	-11.46	-11.25	-14.21	-127.77	0.69	6.161	6.2	6.019	6.216
27d	-11.44	-11.33	-13.58	-126.53	1.2	5.92	6.08	5.935	5.8816
28d*	-11.45	-11.22	-13.96	-128.26	0.51	6.292	6.2	6.205	5.98
29d*	-10.73	-7.6	-14.12	-112.24	0.41	6.387	6.4	6.435	6.275
30d*	-10.96	-11.73	-13.74	-118.42	0.52	6.283	6.36	6.184	6.393
31e*	-8.916	-9.7	-11.78	-92.415	3	5.522	5.69	5.878	5.784
32e	-11.5	-11.84	-13.32	-118.39	2.6	5.585	5.94	5.805	5.725
33e	-11.53	-11.16	-12.59	-120.38	2.7	5.568	5.5	5.766	5.777
34e	-11.36	-11.82	-12.42	-117.92	1.2	5.92	5.78	5.823	5.827
35e*	-9.13	-8.44	-12.89	-96.05	3.2	5.494	5.59	5.802	5.784
36e*	-11.47	-12.37	-14.69	-120.74	1	6	6.03	6.063	6.051
37e*	-11.68	-11.95	-13.45	-127.84	0.53	6.275	6.25	5.775	6.138
38f	-9.759	-10.39	-14.56	-108.73	0.79	6.102	6.11	6.687	6.08
39f	-11.11	-12	-13.04	-122.99	0.57	6.244	6.02	6.23	5.845

| 40f*	-9.42	-9.65	-14.2	-106.05	1	6	6.1	5.998	6.132
41f	-11.64	-12.86	-14.13	-131.37	0.71	6.148	6.04	6.537	6.147
42f*	-9.76	-10.8	-13.86	-108.16	0.99	6.004	5.94	5.766	5.743
43f	-10.95	-11.9	-13.01	-121.5	1.1	5.958	6.04	6.107	5.955
44f*	-11.1	-11.7	-13.43	-122.81	0.4	6.397	5.99	6.158	5.832
45f	-5.22	-12.18	-6.652	-56.258	0.62	6.207	5.99	6.185	6.24
46g*	-11.73	-13.05	-14.37	-131.09	0.18	6.744	6.81	6.762	6.608
47g	-11.85	-11.76	-13..62	-134.73	0.19	6.721	6.68	6.259	6.227
48g*	-12.08	-13.13	-14..64	-134.86	0.14	6.853	6.93	6.964	6.869
49g*	-11.9	-12.91	-14.05	-129.67	0.13	6.886	6.82	6.92	6.872
50g	-12.16	-12.17	-14.35	-134.5	0.07	7.154	7.05	7.142	7.191
51h	-10.38	-11.53	-14.56	-111.04	0.24	6.619	6.41	6.557	6.711
52h*	-11.95	-11.91	-14.39	-140.52	0.3	6.522	6.62	6.68	6.375
53h	-10.27	-8.89	-5.79	-128.68	0.78	6.107	6.1	6.075	6.637
54	-9.33	-10.02	-12.27	-105.52	17	4.769	4.7	5.488	5.862
55	-9.13	-9.63	-12.77	-92.518	3	5.522	5.61	5.674	5.796
56	-11.48	-12.17	-14.53	-118.39	2.6	5.585	5.91	5.76	5.705
57*	-11.16	-11.99	-12.39	-115.28	0.6	6.221	6.14	5.839	5.932
58*	-11.61	-11.29	-12.75	-126.53	1.6	5.795	6.1	5.732	5.855

(SP, XP, IFD) and binding energy (ΔGbind).

*test set molecules.

3D-QSAR model
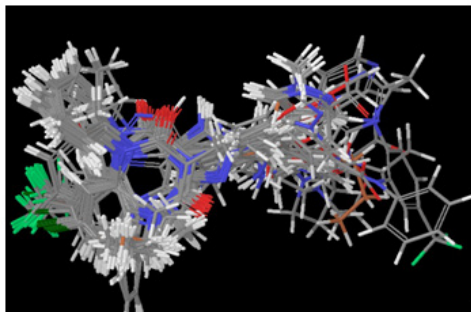
Figure 8: The alignment obtained through XP docking.

The 3D-QSAR CoMFA and CoMSIA analysis was performed by aligning the dock poses of all fifty eight molecules (Figure 8). A training set of thirty five molecules (which has been selected for generating PHASE models) were also chosen for constructing CoMFA and CoMSIA models. The best predictions were obtained with CoMFA standard model q2 value of 0.607, r2 value of 0.968 and CoMSIA with q2 value of 0.616 and r2 value of 0.954 respectively. The predictivity of the model was cross validated by the test set of twenty three molecules. The predictive correlation coefficient r2 pred of 0.576 for CoMFA and 0.528 for CoMSIA explains good predictive ability of the models. The scatter plot of actual and predicted pEC50 values for training and test set of CoMFA and CoMSIA studies are shown in Figure 9a and Figure 9b. The results of CoMFA and CoMSIA are listed in Table 3.
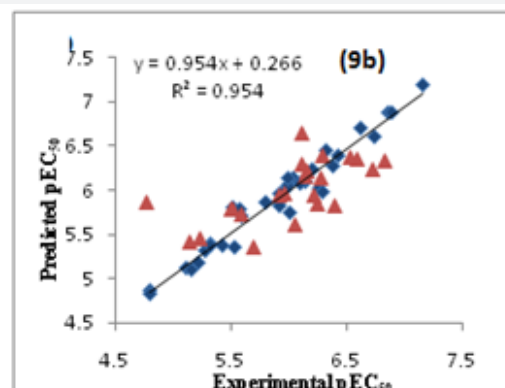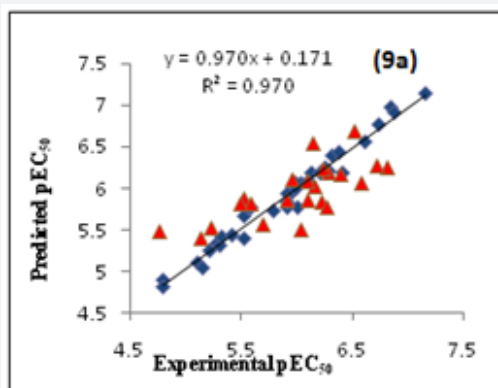
Figure 9: Scatter plot of Experimental pEC50 vs. Predicted pEC50 (a) CoMFA, (b) CoMSIA. (Squares represent training set and triangles represent test set molecules).
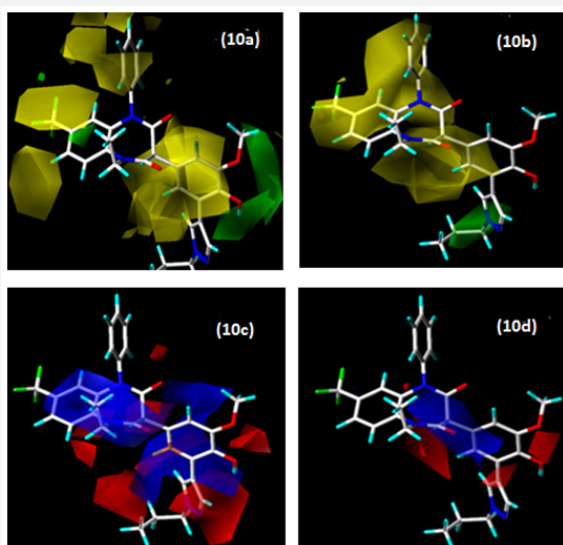
Figure 10: Steric contour maps for CoMFA (a) and CoMSIA (b) signifying Green as favored and yellow as disfavored region. Electrostatic contour maps for CoMFA (c) and CoMSIA (d) representing red as electronegative favored and blue as electronegative disfavored region.

Table 3: Significant statistical data for CoMFA and CoMSIA.

Statistical Parameters	CoMFA	CoMSIA
Molecules in the training set	35	35
Molecules in the test set	23	23
q^2	0.607	0.616
r^2	0.968	0.954
ONC	6	5
r^2Pred	0.576	0.528
SEE	0.115	0.136
F value	143.365	120.728

q^2 = correlation coefficient from leave one out method. ONC = Optimum number of components, SEE=Standard error of estimate, F = Fisher test value, r^2_{Pred} = predictive r^2 on test set.

CoMFA and CoMSIA contour maps

To visualize the information derived from the 3D-QSAR models, contour maps were generated. The contour plots are representation of the lattice points in the grid and the difference in the lattice points is strongly connected with the difference in the receptor binding affinity. Whereas molecular fields defines the favorable and unfavorable interaction energies of aligned molecules with a probe atom traversing across the lattice points, suggesting the modification required to design new potent molecules.

The CoMFA contours indicate the region in space where the molecule would favorably or unfavorably interact with the receptor, while CoMSIA contours indicate the areas within the specified region where the presence of groups with particular physicochemical property binds to the receptor. Therefore the most potent inhibitor among the series 50g was displayed on the maps for visualization.

The steric contours of CoMFA and CoMSIA Figure 10(a) and Figure 10(b) indicate the regions where sterically bulky substituent might have favorable (green) and unfavorable (yellow) effects on the activity of the inhibitor. The methoxy and hydroxyl substitutions are necessary to have good interactions with the receptor, and pyrazole ring is less sterically crowded, substitutions at this position with a bulky group facilitate increase in the activity. The electrostatic contours of CoMFA and CoMSIA Figure 10(c) and Figure 10(d) suggest that increasing the negative charge in red region will have a propensity for increase in binding affinity towards the receptor. The hydrophobic contour of CoMSIA Figure 11(a) recommends the substituent's like CF3 and phenyl are accountable for the hydrophobic activity. Any modification of CF_3 and phenyl would result in a decrease of activity. The H-bond acceptor and H-bond donor contours of CoMSIA imply the regions with cyan acceptor favored and magenta donor favored regions are liable for activity as depicted in Figure 11(b) and Figure 11(c) (Figure 6-10) (Table 2,3).

Design of new molecules
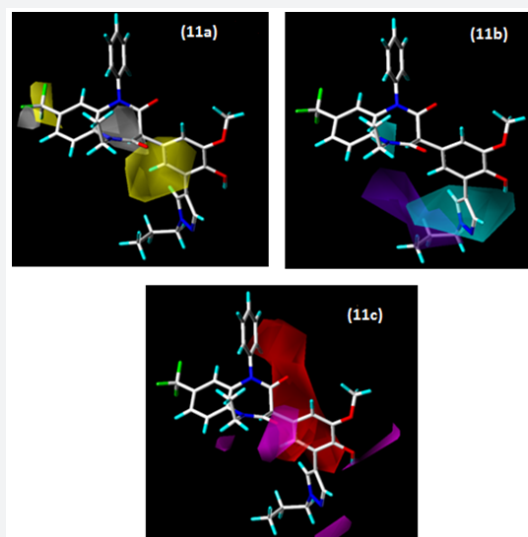
Figure 11: (a) Hydrophobic (yellow-favored, white-disfavored), (b) Hydrogen bond acceptor (cyan-favored, purple-disfavored) and (c) Hydrogen bond donor (magenta-favored, red-disfavored) contour maps for CoMSIA.

As the first hit molecule screened through Capsid assembly assay was compound 54 with an enamine side chain at 3rd position was reported to be chemically instable by Lee Fader [28], and the SAR studies were performed by replacing the enamine side chain with more polar groups like 2-methoxyethyl, 2-hydroxyethyl chains which resulted in no loss of potency thus confirmed that small substitutions at this position can be tolerated [29]. In this article designing of new molecules was attempted by substitution on phenyl ring at 5th position but results were not industrious in terms of compliance to PHASE. These observations lead us to a conclusion that the substitution at R_1, R_2 and R_3 of phenyl moiety at third position are well tolerated and therefore were targeted for isosteric substitution (Figure 11).
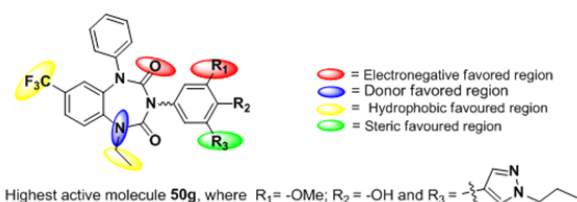
Figure 12: Structural requirements for designing new molecule as inhibitors targeting HIV-1capsid
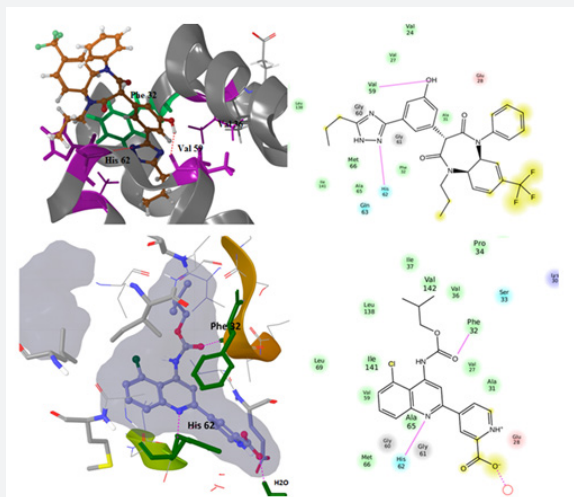
**Figure 13:** IFD dock pose of newly designed molecule N1 along with ligand interaction diagram.

The sterically favored pyrazole at R$_3$, that showed single interaction with His 62 has been substituted with 1,2,4 triazole and electronegative favored methoxy (-OMe) group in 50g has been replaced with methyl sulfonate (-SO2Me) of N10, showed equivalent dock scores and facilitate newer interactions with Val 27 and Val 59. This implies that improved electronegativity on the

substituents showed higher affinity. Increase in aliphatic carbon chain on diazapine nitrogen with propyl (N$_1$-N$_6$) and isobutyl (N$_7$-N$_{10}$) accounts for increased hydrophobicity. Considering the structural necessities depicted in Figure 12, we have designed ten new molecules which adhere to PHASE as well as CoMFA and CoMSIA. Docking studies revealed that the designed molecules Figure 13 show good interaction with the active site amino acids. The structures of newly designed molecules and their possible interactions are tinted in Figure 14. The pharmacokinetic data of newly designed molecules is illustrated in Table 3 (provided in supplementary data).

Design of new molecules by retaining quinoline scaffold. The quinolines not only exhibited antibacterial, antimalarial but also shown to inhibit HIV-1 activity by involving in assembly process. Hence new molecules are designed with a view to interact with N-terminal domain of HIV-1 Capsid, they too have shown the similar interactions as that of reported molecules. The docks pose C$_1$ with its ligand interactions are depicted in the Figure 14. The oxygen of carbonyl carbon forms hydrogen bond interaction with amine of Phe 32, where as the amine of quinoline forms hydrogen bond interaction with the His 62 amino acids within the active site.

Prediction of ADME properties

The newly designed molecules were analyzed for their drug-likeness by assessing their physiochemical properties (Table S3 provided in supplementary data) and by applying Lipinski's rule of five. This rule states that the molecule should have molecular weight < 650 Daltons, H-bond donors <5, H-bond acceptors <10, and a log P of <5. For the selected 10 molecules the partition coefficient (QPlogPo/w), water solubility (QPlogS) and MDCK cell permeability (QPPMDCK) properties have been estimated. All these pharmacokinetic parameters were found to be within the acceptable range.
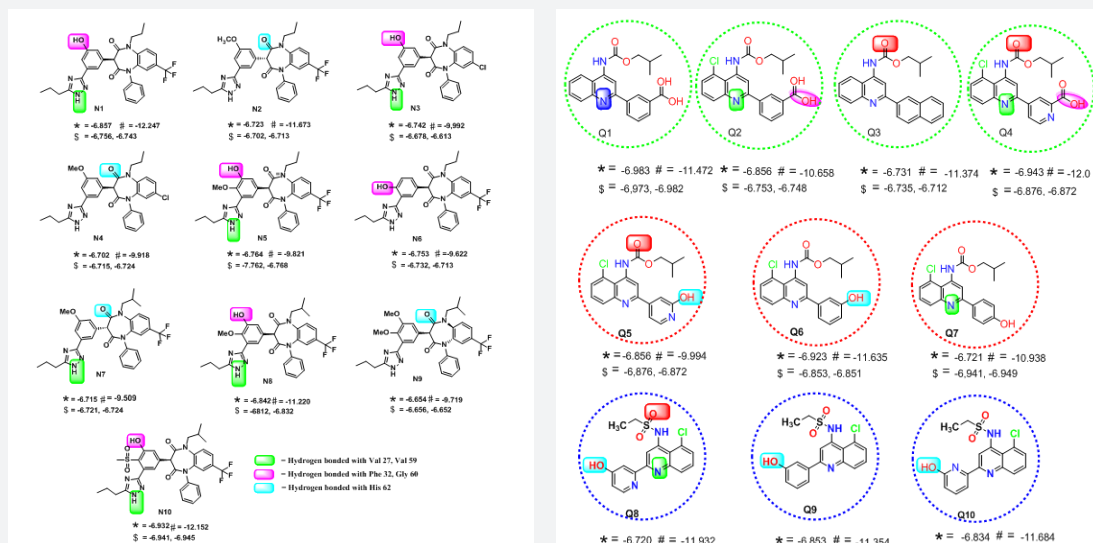
Figure 14: Structures of the newly designed molecules highlighting their interactions with the protein active site residues. Note: *,$ indicates predicted activities of PHASE and CoMFA & CoMSIA. # represents the dock scores of the newly designed molecules.

Conclusion

In summary, Docking studies guided by the receptor, predicted the binding conformations, and binding free energies for a series of fifty eight benzodiazepenedione inhibitors against HIV-1 capsid which are well correlated with their reported inhibitory activities, The docking results for the newly designed molecules provide additional hydrogen bond interactions with residues Val 27, Val 59 and Gly 60 along with His 62, Phe 32. The higher binding affinities of designed molecules N1, N2, N9 and N10 may be attributed to the additional hydrophobic interaction. Pharmacophore generation and 3D QSAR CoMFA and CoMSIA field distribution are in good conformity with the structural requirements of the active site of capsid assembly inhibitors that allows conception of a plausible template for designing novel potent inhibitors. The predicted activities of newly designed molecules for both CoMFA and CoMSIA are found to be equivalent to that of the highest active molecule 50g and expected to be active against HIV-1 Capsid assembly (Figure 14).

Acknowledgment

References

1. Freed EO (1998) HIV-1 Gag proteins: diverse functions in the virus life cycle. Virology 251(1): 1-15.

2. Vogt VM, Coffin JM, Hughes SH, Varmus HE (1997) Retroviral virions and genomes. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA, pp. 27-69.

3. Ganser BK, Li S, Klishko VY, Finch JT, Sundquist WI (1999) Assembly and analysis of conical models for the HIV-1 core. Science 283(5398): 80-83.

4. Li F, Golia-Gaur R, Salzwedel K, Kilgore NR, Reddick M, et al. (2003) a potent HIV inhibitor that disrupts core condensation by targeting a late step in Gag processing, Proc Natl Acad Sci USA 100(23): 13555-13560.

5. Henderson LE, Bowers MA, Sowder II RC, Serabyn SA, Johnson DG, et al. (1992) Gag proteins of the highly replicative MN strain of human immunodeficiency virus type 1: post translational modifications, proteolytic processings, and complete amino acid sequences. J Virol 66(4): 1856-1865.

6. Adamson CS, Freed EO (2007) Human immunodeficiency virus type- 1 assembly, release and maturation. Adv Pharmacol 55: 347-387.

7. Swanstrom R, Wills JW, Coffin JM, Hughes SH, Varmus HE (1997) Synthesis, assembly, and processing of viral proteins. Cold Spring Harbor, Newyork, USA, pp. 263-334.

8. Lanman J, Lam TT, Barnes S, Sakalian M, Emmett MR, et al. (2003) Identification of novel interactions in HIV-1 capsid protein assembly by high-resolution mass spectrometry. J Mol Biol 325(4): 759-772.

9. Lanman J, Prevelige PE (2005) Kinetic and mass spectrometry-based investigation of human immunodeficiency virus type 1 assembly and maturation. Adv Virus Res 64: 285-309.

10. Briggs JA, Grunewald K, Glass B, Förster F, Kräusslich HG, et al. (2006) The mechanism of HIV-1 core assembly: insights from three-dimensional reconstructions of authentic virions. Structure 14(1): 15-20.

11. Worthylake DK, Wang H, Yoo S, Sundquist WI, Hill CP (1999) Structures of the HIV-1 capsid protein dimerization domain at 2.6 Å resolution. Acta Crystallogr D Biol Crystallogr 55(Pt 1): 85-92.

12. Fassati A, Goff SP (2001) Characterization of intracellular reverse transcription complexes of human immunodeficiency virus type 1. J Virol 75(8): 3626-3635.

13. Tang C, Ndassa Y, Summers MF (2002) Structure of the amino-terminal core domain of the HIV-1 capsid protein. Nat Science 273: 231-235.

14. Tang C, Ndassa Y, Summers MF (2002) Structure of the N-terminal 283-residue fragment of the immature HIV-1 Gag polyprotein. Nat Struct Biol 9(7): 537-543.

15. Tang S, Murakami T, Agresta BE, Campbell S, Freed EO, et al. (2001) Human Immuno deficiency virus type 1 N-terminal capsid mutants that exhibit aberrant core morphology and are blocked in initiation of reverse transcription in infected cells. J Virol 75(19): 9357- 9366.

16. Ehrlich L, Liu T, Scarlata S, Chu B, Carter C (2001) Human Immuno deficiency virus type 1 N-terminal capsid mutants that exhibit aberrant core morphology and are blocked in initiation of reverse transcription in infected cells. J Biophys 81: 586-594.

17. Gross I, Hohenberg H, Wilk T, Wiegers K, Grattinger M (2000) N-terminal extension of Human Immunodeficiency virus capsid protein converts the in vitro assembly, phenotype from tubular to spherical particles. EMBO J 9: 103- 113.

18. Niedrig M, Gelderblom HR, Pauli G, Marz J, Bickhard H, et al. (1994) Inhibition of infectious human immunodeficiency virus type 1 particle formation by Gag protein-derived peptides. Jour Gen Virol 75(Pt 6): 1469-1474.

19. Houglund S, Su J, Reneby SS, Vegvari A, Hjerten S, et al. (2002) Tripeptide interference with human immunodeficiency virus type 1 morphogenesis. Antimicrob Agents Chemother 46(11): 3597-3605.

20. Garzon MT, Lindon-moya MC, Barrera FN, Prieto A, Gomez J, et al. (2004) The dimerization domain of the HIV-1 capsid protein binds a capsid protein-derived peptide: a biophysical characterization. Protein Sci 13(6): 1512-1523.

21. Tang C, Loelinger E, Kinde I, Kyere S, Mayo K, et al. (2003) Antiviral Inhibition of the HIV-1 Capsid Protein. J Mol Biol 327(5): 1013-1020.

22. Sakalian M, McMurtrey CP, Deeg FJ, Maloy CW, Li F, et al. (2006) 3-O-(3', 3'-dimethylsuccinyl) betulinic acid inhibits maturation of the human immunodeficiency virus type1Gag precursor assembled in vitro. J Virol 80(12): 5716-5722.

23. Sticht J, Humbert M, Findlow S, Bodem J, Muller B, et al. (2005) A Peptide inhibitor of HIV- 1 assembly in vitro. Nat Struct Mol Biol 12(8): 671-677.

24. Zhang H, Zhao Q, Battacharya S, Waheed AA, Tong X, et al. (2008) A Cell-penetrating helical peptide as a potential HIV-1 inhibitor. J Mol Biol 378(3): 565-580.

25. Bartonova V, Ignoet S, Stich J, Glass B, Habermann A, et al. (2008) Residues in the HIV-1 capsid assembly inhibitor binding site are essential for maintaining the assembly- competent quaternary structure of the capsid protein. J Biol Chem 283(46): 32024-32033.

26. Roy J, Kumar UC, Machiraju PK, Muttineni RK, Kumar BVSS, et al. (2010) Insilico studies on anthrax lethal factor inhibitors: Pharmacophore modeling and virtual Screening approaches towards designing of novel inhibitors for a killer. J Mol Graph Model 29(2): 256-265.

27. Debnath AK (2002) Pharmacophore Mapping of a Series of 2, 4-Diamino-5- deazapteridine Inhibitors of Mycobacterium avium complex Dihydrofolate Reductase. J Med Chem 45(1): 41-53.

28. Fader LD, Bethell R, Bonneau P, Bös M, Bousquet Y, et al. (2011) Discovery of a 1, 5-dihydrobenzo[b][1,4]diazepine-2,4-dione series of inhibitors of HIV-1 capsid assembly. Bioorg Med Chem Lett 21(1): 398-404.

29. Fader LD, Serge L, Sebastien M, Stephen HK, Bousquet Y, et al. (2013) Discovery of a 1, 5-dihydrobenzo[b][1,4]diazepene-2,4-dione series of hiv-1capsid assembly inhibitors 1: Addressing configurationally instability through scaffold modification. Bioorg Med Chem Lett 23(11): 3396-3400.

30. Lee DF, Serge L, Sylvie G, Sebastien M, Stephen HK, et al. (2013) Optimization of a 1,5-dihydrobenzo[b][1,4]diazepene-2,4-dione series of hiv-1capsid assembly inhibitors 2: Structure-activity relationships (SAR) of the C3- Phenyl moiety. Bioorg Med Chem Lett 23: 3401-3405.

31. Dixon SL, Smondyrev AM, Knoll EH, Rao SN, Shaw DE, et al. (2006) PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology ad preliminary results. J Comput Aided Mol Des 20(10-11): 647-671.

32. Jain SV, Ghate M, Bhadoriya KS, Bari SB, Chaudhari A, et al. (2012) 2D, 3D-QSAR and docking studies of 1,2,3-thiadiazole thioacetanilides analogues as potent HIV-1 non-nucleoside reverse transcriptase inhibitors. Org Med Chem Lett 2(1): 22.

33. Li J, Tan Z, Tang S, Hewlett I, Pang R, et al. (2009) Discovery and dual inhibitors targeting both HIV-1 capsid and human cyclohilin A to inhibit the assembly and uncoating of the viral capsid. Bioorg Med Chem 17(8): 3177-3188.

34. Wold S, Johansson A, Cochi M, Lieden (1993) 3D QSAR in drug design theory methods and application, in; Kubinyi PLS-Partial least square projection to latent structures. ESCOM: 523-550.

35. Lemke CT1, Titolo S, von Schwedler U, Goudreau N, Mercier JF, et al. (2012) Distinct affects of two HIV-1 Capsid Assembly Families That Bind the Same Site within the N-Terminal Domain of CA Protein. J Virol 86(12): 6643-6655.

36. Kaminski GA, Friesner RA, Tirado-Rives J, Jorgensen WL (2001) Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. J Phys Chem B 105(28): 6474-6487.

37. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, et al. (2004) Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem 47(7): 1739-1749.

38. Frieshner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, et al. (2006) Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein ligand complexes. J Med Chem 49: 6177-6196.

39. Glide (2011) Version 5.6, Schrödinger, LLC, Newyork, USA.

40. Patrick, An introduction to medicinal Chemistry Chapter 18, Quantitative structure activity relationships Fourth Edition: 377-405.

41. St Lious MO (2012) Tripos associates Sybyl-X 2.1.

42. Gasteiger J, Marsili M (1980) Iterative partial equalization of orbital electro negativity- a rapid access to atomic charges. Tetrahedron 36: 3219-3228.

43. Klebe G, Abraham U, Mietzner T (1994) Molecualr similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. J Med Chem 37(24): 4130-4146.

44. Qik prop (2011) version 3.4, Schrödinger, LLC, New york, USA.

45. Xu JP, Branson JD, Lawrence R, Cocklin S (2016) Identification of a small molecule HIV-1 inhibitor that targets the capsid hexamer. Bioorg Med Chem Lett 26(3): 824-828.

46. Gamble TR, Yoo S, Vajdos FF, von Schwedler UK, Worthylake DK, et al. (1997) Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein,. Science 278(5339): 849-853.

47. Ganser-Pornillos BK, von Schwedler UK, Stray KM, Aiken C, Sundquist WI (2004) Assembly properties of the human immunodeficiency virus type 1 CA protein. J Virol 78(5): 2545–2552.

48. Ganser-Pornillos BK, Cheng A, Yeager M (2007) Structure of full-length HIV-1 CA: a model for the mature capsid lattice. Cell 131(1): 70-79.

49. Von Schwedler UK, Stemmler TL, Klishko VY, Li S, Albertine KH, et al. (1998) Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. EMBO J 17(6): 1555-1568.

50. Von Schwedler UK, Stray KM, Garrus JE, Sundquist WI (2003) Functional surfaces of the human immunodeficiency virus type 1 capsid protein. J Virol 77(9): 5439-5450.

51. Sundquist WI, Hill CP (2007) How to assemble a capsid. Cell 131(1): 17-19.

52. Ganser-Pornillos BK, Yeager M, Sundquist WI (2008) The structural biology of HIV assembly. Curr Opin Struct Biol 18(2): 203-217.

53. Yoakim C, Deroy P, Duplessis M, Gagnon A, Goulet S, et al. () Determination of EC50 values by HIV-1 Replication assay. PCT Int Appl WO 2008067644: 90.