International Journal of Sciences Research Article
(ISSN 2305-3925) Volume 2, Issue May 2013
http://www.ijSciences.com

Molecular Modeling Studies and ADME Screening on HIV-1 Integrase Inhibitors as Anti-Viral Agents

Sibi Narayanan[1], D. Velmurugan[1,2]

[1]Centre of Advanced Study in Crystallography and Biophysics, University of Madras, Guindy Campus, Chennai – 600 025, Tamil Nadu, India
[2]Bioinformatics Infrastructure Facility, University of Madras, Guindy Campus, Chennai – 600 025, Tamil Nadu, India

Abstract: Human immunodeficiency virus (HIV) is a lentivirus, a member of the retrovirus family, which leads to acquired immunodeficiency syndrome (AIDS). HIV is present as both free virus particles and virus within infected immune cells. Integration of viral DNA into the host chromosome is the key process in the HIV replication cycle and therefore Integrase has served as an attractive target for antivirals. HIV-1 Integrase (HIV-1 IN) is an essential enzyme in the life cycle of the virus, responsible for catalyzing the insertion of the viral genome into the host cell chromosome and it provides an attractive target for antiviral drug design. Antiretroviral treatment reduces both the mortality and the morbidity of HIV infection. *Raltegravir*, also known as *Isentress* and *MK-0518*, is a well known integrase inhibitor. This prevents the virus from making new copies of HIV. Structures similar to this compound were identified and molecular modeling studies has been performed with HIV-1 IN as target molecule using Schrödinger Suite 2007. PASS (prediction of activity spectra for substances) prediction of the compounds was carried out and all compounds exhibited anti-viral activity. These compounds show favorable interactions with the amino acid residues and the metal ion (Mg^{2+}) at the active site of HIV-1 IN thereby substantiating their proven efficacy as anti-viral compounds. Further, their ADME (Absorption, Digestion, Metabolism and Excretion) screening was also carried out in order to check their potency to be used for second-generation drug development. The work demonstrates that molecular modeling, activity prediction and screening of the compounds for their subsequent ADME properties is a promising approach to predict the binding activity of compounds to the receptor.

Keywords: HIV, HIV-1 IN, antiviral, raltegravir, Schrödinger Suite 2007, PASS, ADME

1. Introduction

HIV is a member of the genus Lentivirus, part of the Retroviridae family. The Lentivirus can lead to *acquired immunodeficiency syndrome* (AIDS), a condition wherein the human immune system begins to fail which leads to life-threatening opportunistic infections (Asante-Appiah & Skalka, 1999). They are transmitted as single-stranded, positive-sense, enveloped RNA viruses. After entering into the target cell, the viral RNA genome is converted into double-stranded DNA by a virally encoded reverse transcriptase, which is present in the virus particle (Dyda *et. al.*, 1999). The viral DNA is integrated into the cellular DNA by a virally encoded integrase, along with host cellular co-factors, which help in the gene transcription process. Infections with HIV occur by the transfer of blood, semen, vaginal fluid or breast milk. HIV is present as both free virus particles and virus within infected immune cells (Eijkelenboom *et. al.*, 1999). HIV primarily infects vital cells in the human immune system such as helper T cells, macrophages and dendritic cells. Two species of HIV infect humans: HIV-1 and HIV-2, the former one being more virulent, relatively easily transmitted and is the cause of the majority of HIV infections globally and the latter one is less transmittable than HIV-1 (Engelman *et. al.*, 1991). Currently, it is estimated that about 0.6% of the world's population is infected with HIV (Goldgur *et. al.*, 1999).

Integration of viral DNA into the host chromosome is the key process in the HIV replication cycle and

D. Velmurugan (Correspondence)
shirai2011@gmail.com

therefore Integrase has served as an attractive target for design of antivirals (Goldgur, *et. al.,* 1998). Integrase uses a single active site to accommodate two different configurations of DNA substrates that may constrain the ability of HIV to develop drug resistance to integrase inhibitors (Greenwald, *et. al.,* 1999). HIV-1 IN is a 32-kDa enzyme that carries out DNA integration in a series of reaction steps. The first step is termed as the *3' processing step*, wherein the linear double-stranded *viral DNA* with sequence specific 3' ends is synthesized by reverse transcription from the viral RNA genome. The second step is the *strand transfer*. Here, the integrase protein joins the previously processed 3' ends to the 5' ends of the strands of target DNA at the site of integration. Then the 5' ends are produced by *integrase-catalyzed* staggered cuts, 5bp apart. A *Y shaped*, gapped, recombination intermediate results, within the 5' ends of the viral DNA strands and the 3' ends of target DNA strands remains unjoined, creating a gap of 5bp. The final step is termed *disintegration step*, where integrase may catalyze the excision of viral DNA. This also involves the *host DNA repair synthesis*, where the 5bp gaps between the unjoined strands are filled in and then litigated. Since this process occurs at both cuts flanking the HIV genome, a 5bp duplication of host DNA is produced at the ends of HIV-1 integration (Vink, *et. al*., 1993; Lodi, *et. al*., 1995; Ni, *et. al.,* 2001).

HIV-1 IN has three domains, namely; the Catalytic core, the C-terminal and N-terminal domains. Catalytic core domain (CCD) contains the active site responsible for catalysis of all integration/disintegration reactions (Gerton & Brown, 1997). The C-terminal has the capacity to bind both viral and host DNA. The structures of the catalytic core and C-terminal domains have been determined separately. The structure of catalytic core domain of HIV-1 IN consists of a central five-stranded β-sheet with six surrounding α-helices (Chen *et. al*., 2000). Three amino acids in the CCD are highly conserved among retrotransposon and retroviral integrases. Mutation of these residues (Asp 64, Asp 116 & Glu 152) leads to a loss of all catalytic activities of the proteins and therefore is thought to be essential components of the integrase active site (Chiu & Davies, 2004). The interactions between the inhibitor and integrase partially mimic the normal interactions with viral DNA substrate during the 3'processing reaction (Christoph *et. al*., 2000). Site-directed mutagenesis and photo-crosslinking experiments have identified several residues near the active site, including Lys 156, Lys 159, Gln 148 and Tyr 143, which are critical for binding viral DNA substrate and the inhibitor. Thus it has been one of the important targets for the development of effective drugs against AIDS.

One strategy to minimize the emergence of drug resistance may be the design of compounds that would interact with amino acids of cofactors that are essential for catalysis (De Clercq, 2004). The development of effective inhibitors of HIV replication targeted to reverse transcriptase and protease has demonstrated the potential effectiveness of antiviral therapy for the treatment of AIDS. Drugs targeting integrase would also be a valuable complement to reverse transcriptase and protease inhibitors. The method of therapy with integrase inhibitor is still under experimental state with only one approved drug by FDA in the class of integrase inhibitors (Esposito & Craigie, 1999; Hazuda, *et. al.,* 2000; De Clercq, 2005). This integrase/inhibitor complex will provide a platform for the design of an additional class of HIV inhibitors.

Three-dimensional structural information about receptors is one of the most modern approaches used in order to find new leads for therapeutic targets. Docking simulation is an effective way to predict the binding structure of a substrate in its receptor, which is being used successfully in many applications. Basic aim of docking procedures is to identify the correct conformation of ligands in the binding pocket of a protein and to predict the affinity between the ligand and the protein. How two molecules fit together in a three-dimensional space describes the process well (Dixon & Blaney, 1998; Leroy *et. al.,* 2001). Identification of innovative small molecular scaffolds exhibiting high binding affinity and selectivity for the target together with a favorable absorbtion, distribution, metabolism and excretion (ADME) output is one of the main goals in drug discovery. Computational modeling has been investigated as a tool to optimize selection of the most suitable candidates for drug development all over the world (Smith *et. al.,* 2004).

Raltegravir, also known as *Isentress* and *MK-0518*, is a well known integrase inhibitor produced by *Merck & Co* (ter Heine *et. al*., 2009). Integrase inhibitors work by blocking integrase, a protein that HIV needs to insert its viral genetic material into the genetic material of an infected cell. This prevents the virus from making several new copies of HIV (Zheng *et. al.,* 1996). Raltegravir helps to control HIV, even when it is resistant to other medications. It contains raltegravir potassium, a HIV integrase strand transfer inhibitor. It has shown to effectively lower the amount of virus in the blood to undetectable levels in 62 percent of people taking it in combination with other anti-HIV medicines (Kuo *et. al.,* 2004). Structures similar to this compound were identified

using NCBI Pubchem and Drugbank. All these compounds were docked at the CCD of the HIV-1 IN enzyme. Molecular modeling studies of these compounds have been carried out using Induced Fit Docking module (flexible docking) of Schrödinger Suite 2008 with HIV-1 IN as target.

2. Methods

(2.1) Target Protein structure preparation

The X-ray crystal structure of HIV-1 IN complex (PDB ID: 1QS4) was obtained from the RCSB Protein Data Bank (PDB). Water molecules of crystallization were removed from the complex, and hydrogen atoms were added to the structure. The protein was optimized. Partial atomic charges were also assigned according to the force field. Minimizations were performed until the average root mean square deviation of the non-hydrogen atoms reached 0.3Å using OPLS-AA force field to remove the steric hindrance using the Protein Preparation Wizard.

(2.2) Ligand structure preparation

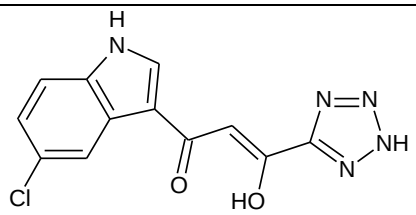
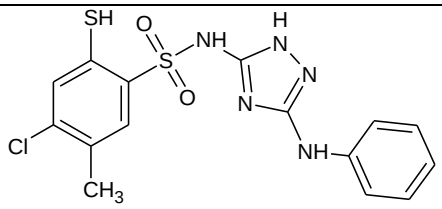
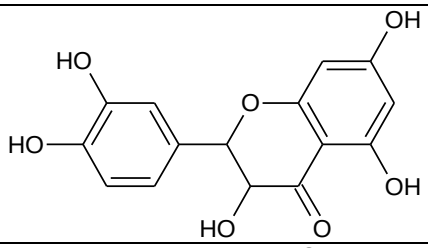
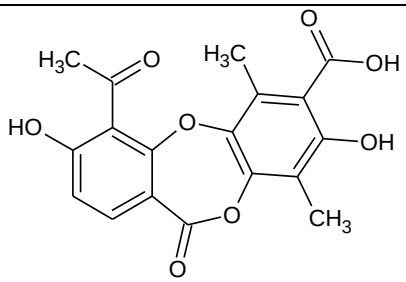
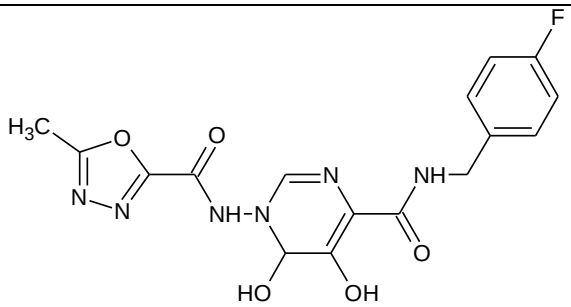
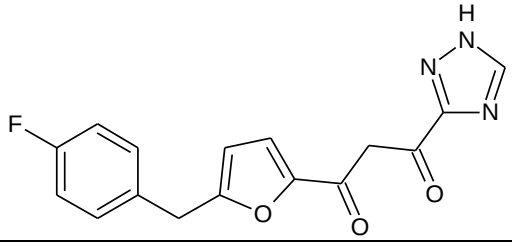
In all, seven compounds were taken for our present study. Figure 1 shows the schematic representation of the compounds with their respective *IUPAC* names. The coordinates of raltegravir were obtained from the PubChem database. The compounds exhibiting similar moiety were selected from the database. All compounds were imported into the software's workspace and energy minimized using two algorithms: steepest descent and conjugate gradient, of the Impact module having default force field, OPLS-AA. Poses with an rmsd of less than 0.5Å and a maximum atomic displacement of less than 1.3Å were eliminated as redundant in order to increase diversity in the retained ligand/compound poses. Compounds were then subjected to PASS (Prediction of Activity Spectra for Substances), which predicts the biological activity spectrum for a compound on the basis of its structural formula. PASS predicts 3678 pharmacological effects, mechanisms of action, mutagenicity, carcinogenicity, teratogenicity and embryotoxicity among the compounds from in-house and commercial data bases. The approach used in PASS is based on the suggestion that ***Activity=f(Structure).*** By "comparing" the structure of a new compound with structures of well-known biologically active substance it is possible to estimate if a new compound may have a particular effect. However, PASS operates with many thousands of substances from the training set, so provides more objective estimate if a compound is active or not for any kind of activity. Training set consists of about 46,000 of biologically active compounds. They include about 16,000 already launched drugs and 30,000 drug-candidates under clinical or advanced preclinical testing now.

(2.3) Docking protocol

All docking calculations were performed using the Induced Fit Docking module of the Schrödinger suite. It performs flexible protein-ligand docking and searches for favorable interactions between one typically small ligand molecule and a typically larger receptor molecule, usually protein. Docking process is divided into three steps. Initial Glide Docking, wherein protein preparation constrained refinement is carried out with a maximum of 20 poses. Prime Induced Fit, wherein the side chains are optimized and refinement of residues takes place, if the ligand poses are within 5.0Å. Step 3 consists of the Glide redocking stage using Standard Precision (SP) mode. Upon completion of each docking calculation, at most 20 poses per compound were generated. The best-docked structure/pose was chosen using three criterias: ***Glidescore (Gscore)*** function, ***Glide Energy*** and the number of ***residual matches (hydrogen bonds)*** with the original drug complex. All computational work was performed using various modules of Schrödinger Suite 2007 using Red Hat Enterprise Linux 5.0 interface running on Pentium D workstation.

(2.4) ADME screening

The Qikprop program was used to obtain the ADME properties of the analogues. It helps in predicting both the physically significant descriptors and pharmaceutically relevant properties. The program predicted 53 properties for the molecules taken into consideration, which consisted of principal descriptors and physiochemical properties with a detailed analysis of the CNS activity, molecular weight, QPlogS (solubility), QplogBB (Blood-Brain Barrier permeability), QplogKp (skin permeability), percentation of human oral absorption & SASA (Solvent Accessible Surface Area). We also evaluated the acceptability of the inhibitors based on the Lipinski's rule of 5, which is essential for rational drug design.

Compound Name	Structure of the Compound	IUPAC NAME
Drug Compound		(2*Z*)-1-(5-chloro-1*H*-indol-3-yl)-3-hydroxy-3-(2*H*-tetrazol-5-yl)prop-2-en-1-one
Compound 1		*N*-(3-anilino-1*H*-1,2,4-triazol-5-yl)-4-chloro-2-mercapto-5-methylbenzenesulfonamide
Compound 2		2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxy-2,3-dihydro-4*H*-chromen-4-one
Compound 3		4-acetyl-3,8-dihydroxy-6,9-dimethyl-11-oxo-11*H*-dibenzo[*b*,*e*][1,4]dioxepine-7-carboxylic acid
Compound 4 (Raltegravir)		*N*-(4-fluorobenzyl)-5,6-dihydroxy-1-{[(5-methyl-1,3,4-oxadiazol-2-yl)carbonyl]amino}-1,6-dihydropyrimidine-4-carboxamide
Compound 5		1-[5-(4-fluorobenzyl)furan-2-yl]-3-(1*H*-1,2,4-triazol-3-yl)propane-1,3-dione

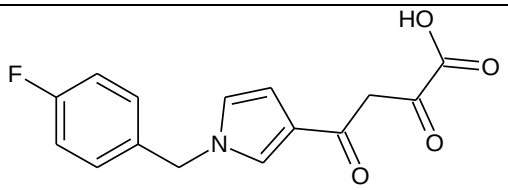
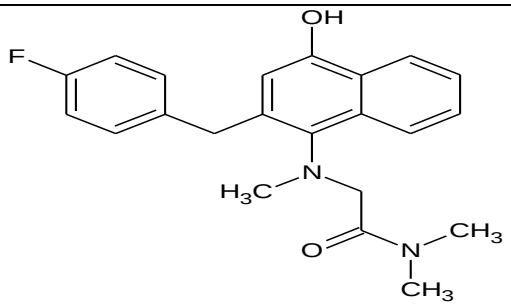
Compound 6		4-[1-(4-fluorobenzyl)-1*H*-pyrrol-3-yl]-2,4-dioxobutanoic acid
Compound 7		2-{[2-(4-fluorobenzyl)-4-hydroxynaphthalen-1-yl](methyl)amino}-*N*,*N*-dimethylacetamide

Figure 1. Schematic representation of compounds taken into study

3. Results and Discussion

So far, drug design efforts have mostly relied upon pharmacophore hypothesis derived from the structures of known inhibitors. Pharmacophore models were successful in discovering some new inhibitors by database searches, uncertainties remain about the details of the protein-ligand interaction and the effects of the molecular framework not covered by the pharmacophore model. We expand our study on compound binding by IN to address the potential binding modes of various other IN inhibitors for which no experimental structures are available, in an attempt to derive guidelines for the design of new inhibiting compounds. With recent improvement in search algorithms and energy functions, computational docking methods have become a valuable tool to probe the interaction between the protein and their inhibitors in the absence of detailed experimental data and can very well contribute to the understanding of its structural and energetic basis.

The central catalytic core domain (CCD) of the 288-residue HIV-1 IN protein is composed of amino acids 50-212. For the IN CCD, the structure obtained by X-ray analysis of the complex with 5CITEP was used (PDB ID: 1QS4) (Goldgur *et. al.,* 1999). For the purpose of docking, Subunit A was selected, which is the only monomer in the asymmetric unit where the position of the ligand could be determined. The missing residues at positions (141-144) in this subunit were incorporated from monomer B of the IN structure (PDB ID: 1BIS) after superposition of the backbones of residues (135-140) and (145-150). This domain contains the active site and shares structural homology with a large family of polynucleotidyl transferases. Domain swaps among integrases from different retroviral species indicate that the core domain determines target site preferences as well as specificity for the species-specific features of the viral DNA end. The domain is composed of a mixed α-helix and β-sheet motif and contains an absolutely conserved D, D-35-E motif characterized by three acidic residues, Asp 64, Asp 116 and Glu 152. IN represents an important, but yet unexploited target for drugs that could complement the combination therapy focused on reverse transcriptase & protease. The docking studies discussed within were performed on only the core domain of HIV-1 IN. Here we present computational docking studies which support this hypothesis and reveal preferred binding modes in the absence of the crystal environment as refined starting points for further design efforts. To verify that the crystallographic environment actually influences the binding mode of the inhibitor, docking studies were carried out for both the crystallographic dimer with one pre-bound inhibitor molecule as well as the isolated monomeric protein as receptor for the ligand. The simulations along with the docking studies of other known/unknown inhibitors should help to obtain a consensus view of inhibitor interactions at the active site and will provide heavy support to SBDD efforts as long as the current difficulties in obtaining experimental structures of IN-inhibitors complexes persist. *Raltegravir*, a well known integrase inhibitor and structures similar to it have been taken from PubChem for docking studies. To study the molecular basis of interaction and affinity of binding of raltegravir and its analogues, all the compounds were docked at the active site of HIV-1 IN. The protein and drug compound (5-CITEP)

were separately minimized and re-docking was carried out at the active site of the CCD. The ranking of the ligand was based on its glide score and glide energy. The difference in glide score (score obtained by the calculation of lipophilic, hydrophobic, hydrophilic, energy, coulombic forces) among all the compounds taken into study was minimal.

Using the online web server, activity prediction was carried out for checking the anti-viral activity and strength of the compounds in inhibiting HIV-1 IN. Energy minimized compounds using various algorithms were uploaded into the PASS server. All compounds showed anti-viral activity and proved to be HIV-1 Integrase inhibitors (Figures (2-9)). A ligand binding at the interface of the protein should always interact directly with the catalytic residues and the metal ion present there. The active site of the CCD consists of a metal ion (Mg^{2+}). Amino group of the indole ring present in the drug compound interacts well with the oxygen atom of Asp 64 as well as Glu 152. Adjacent residue, His 67 also maintains interaction with the oxygen atom of the aliphatic chain in the compound at a distance of 2.81Å. We can also find two interactions, with the metal ion (Mg^{2+}) present in the active site at distances 2.35 Å and 2.37 Å, respectively (Figure 10). Drug compound maintains a glide score of -5.690 and a glide energy (binding energy) of -61.140Kcal/mol, respectively. Amino group of the triazole ring present in compound 1 interacts with polar Lys 156 at a distance of 3.30 Å and metal ion at 2.05 Å. Oxygen atom of the Glu 152 interacts well with the amino group of compound 1. Hydroxyl group of the polar Thr 66 interacts with the oxygen of the compound 1 at a distance of 2.79 Å. Amino group of aliphatic chain also interacts with the metal ion at a range of 2.07 Å (Figure 11). Compound 1 exhibits a glide score better than the Drug (-8.356) and an energy of -65.178Kcal/mol. No significant interactions are observed with the other two catalytic residues, Asp 64 and Asp 116. Hydroxyl group of the chromene ring present in compound 2 interacts with the oxygen atom of the Thr 66 well. Oxygen atom of catalytic residue Asp 64 and Gln 148 is placed at distances of 2.94 Å and 2.34 Å, respectively. The compound also exhibits interaction with side chains of Lys 159 and metal ion at the active site (Figure 12). Compound 2 has a good glide score of -6.175 and a low glide energy (-67.127Kcal/mol) than the previous compound. Oxygen of the dioxepane ring present in the compound 3 interacts with the side chain of Asn 155. Polar residues Lys 156 and Lys 159 also maintain interactions with the compound. Metal ion bound interaction was found with the compound at a range of 2.16 Å (Figure 13). Of all the compounds taken into study, compound 3 has got the least glide score and glide energy, -8.469 and -68.036 Kcal/mol, respectively. But the compound doesn't show any interaction with the catalytic residues. *Raltegravir* (Compound 4) has also been taken into analysis. Compound maintains interactions with a list of polar residues present in the active site region and also with metal ion (Mg^{2+}) (Figure 14). *Raltegravir* exhibits the highest glide score of the lot with least glide energy. Nitrogen present in the triazole ring of compound 5 interacts with the Lys 159 at a distance of 2.65 Å. A metal ion bound interaction (Mg^{2+}) is found with the CH group of the compound. There are also interactions between the side chains of Glu 152 and Thr 66 and the compound (Figure 15). Of all the inhibitors taken into consideration, compound 5 has got the highest glide energy and a considerably high glide score. We can find two metal bound interactions with compound 6. Chlorine present in the phenyl ring of the compound also interacts with the Gln 148 at a distance of 2.98 Å. Interactions with the catalytic neighbor residues (Asn 155 and His 67) are also found in the region (Figure 16). The compound exhibits a good glide score and glide energy too. Side chain of His 67 and Thr 66 maintains interaction with the hydroxyl and amino group of naphthalene ring present in compound 7. Interactions with Lys 159 and metal ion Mg^{2+} are also seen (Figure 17). Compound 7 exhibits a glidescore of -5.867 and a glide energy of -67.631Kcal/mol. Table 1 shows the interactions of the respective compounds with amino acids and metal ion Mg^{2+} at the active site of the target HIV-1 IN. Table 2 shows the docking scores and docking energies of the docked compounds.

Activity Prediction
43 Substructure descriptors; 0 new.
Antiviral
Antiviral (HIV)
HIV-1 integrase inhibitor
HIV-1 integrase (3'-Processing) inhibitor
HIV-1 integrase (Strand Transfer) inhibitor
HIV-1 reverse transcriptase inhibitor
86 Possible activities at Pa > 70%
Pa Pi for Activity:

0,620	0,008	5 Hydroxytryptamine release inhibitor
0,596	0,003	HIV-1 integrase inhibitor
0,519	0,037	5 Hydroxytryptamine 6 agonist
0,496	0,215	5 Hydroxytryptamine 3A antagonist
0,422	0,146	3'-Demethylstaurosporine O-methyltransferase inhibitor
0,399	0,049	Antiviral

Figure 2: Activity prediction of Drug compound using online PASS prediction tool

Activity Prediction
46 Substructure descriptors; 0 new.
Antiviral
Antiviral (HIV)
HIV-1 integrase inhibitor
HIV-1 integrase (3'-Processing) inhibitor
HIV-1 integrase (Strand Transfer) inhibitor
60 Possible activities at Pa > 50%
Pa Pi for Activity:

0,982	0,001	HIV-1 integrase (Strand Transfer) in
0,981	0,001	HIV-1 integrase (3'-Processing) inhil
0,885	0,002	HIV-1 integrase inhibitor
0,750	0,004	Antiviral
0,749	0,004	Antiviral (HIV)
0,521	0,080	RET inhibitor
0,537	0,107	CYP2C9-Cys144 substrate

Figure 3: Activity prediction of compound 1 using online PASS prediction tool

Activity Prediction
31 Substructure descriptors; 0 new.
Antiviral
Antiviral (HIV)
HIV-1 integrase inhibitor
HIV-1 integrase (3'-Processing) inhibitor
HIV-1 integrase (Strand Transfer) inhibitor
96 Possible activities at Pa > 70%
Pa Pi for Activity:

0,973	0,002	Membrane integrity agonist
0,962	0,001	Antineurotoxic
0,962	0,003	CYP1 substrate
0,961	0,003	HIV-1 integrase inhibitor
0,952	0,004	CYP1A2 substrate
0,940	0,004	Antiviral
0,927	0,004	Chlordecone reductase inhibitor

Figure 4: Activity prediction of of compound 2 using online PASS prediction tool

Activity Prediction
38 Substructure descriptors; 0 new.
Antiviral
Antiviral (HIV)
HIV-1 integrase inhibitor
HIV-1 integrase (3'-Processing) inhibitor
HIV-1 integrase (Strand Transfer) inhibitor
HIV-1 reverse transcriptase inhibitor
91 Possible activities at Pa > 50%
Pa Pi for Activity:

0,883	0,004	Kinase inhibitor
0,851	0,006	Antiinflammatory
0,774	0,004	HIV-1 integrase inhibitor
0,791	0,027	Chlordecone reductase inhibitor
0,755	0,004	Mediator release inhibitor
0,770	0,047	Membrane integrity agonist
0,735	0,033	Antiviral (HIV)
0,692	0,004	Lipid peroxidase inhibitor
0,692	0,009	HCV IRES inhibitor

Figure 5: Activity prediction of compound 3 using online PASS prediction tool

Activity Prediction
54 Substructure descriptors; 4 new.
Antiviral
Antiviral (HIV)
HIV-1 integrase inhibitor
HIV-1 integrase (3'-Processing) inhibitor
HIV-1 integrase (Strand Transfer) inhibitor
HIV-1 reverse transcriptase inhibitor
55 Possible activities at Pa > 30%
Pa Pi for Activity:

0,794	0,020	Hypokalemia
0,611	0,015	Antidiarrheal
0,519	0,034	HIV-1 integrase inhibitor
0,546	0,099	Antineoplastic (lymphoma)
0,521	0,112	Antineoplastic (lung cancer)
0,502	0,119	Antiviral
0,466	0,109	Collagen inhibitor
0,451	0,150	RET inhibitor
0,427	0,149	Platelet derived growth factor receptor kinase inhibitor

Figure 6: Activity prediction of compound 4 using online PASS prediction tool

Activity Prediction
39 Substructure descriptors; 0 new.
Antiviral
Antiviral (HIV)
HIV-1 integrase inhibitor
HIV-1 integrase (3'-Processing) inhibitor
HIV-1 integrase (Strand Transfer) inhibitor
HIV-1 reverse transcriptase inhibitor
47 Possible activities at Pa > 30%
Pa Pi for Activity:

0,743	0,002	HIV-1 integrase inhibitor
0,678	0,002	HIV-1 integrase (Strand Transfer) inhibitor
0,537	0,005	Antiviral
0,534	0,019	5 Hydroxytryptamine release inhibitor
0,494	0,004	Antiviral (HIV)
0,422	0,031	Endothelial growth factor antagonist
0,312	0,005	Guanylate cyclase stimulant
0,440	0,149	CXC chemokine 1 receptor antagonist
0,346	0,115	Ligase inhibitor

Figure 7: Activity prediction of of compound 5 using online PASS prediction tool

Activity Prediction

38 Substructure descriptors; 0 new.

Antiviral

Antiviral (HIV)

HIV-1 integrase inhibitor

HIV-1 integrase (3'-Processing) inhibitor

HIV-1 integrase (Strand Transfer) inhibitor

HIV-1 reverse transcriptase inhibitor

45 Possible activities at Pa > 50%

Pa Pi for Activity:

0,894	0,004	Gluconate 2-dehydrogenase (acceptor) inhibitor
0,631	0,003	HIV-1 integrase inhibitor
0,594	0,021	Collagen inhibitor
0,573	0,019	Antiviral
0,539	0,003	HIV-1 integrase (Strand Transfer) inhibitor
0,615	0,090	Ankylosing spondylitis treatment
0,520	0,041	Neurotransmitter agonist
0,500	0,074	Carboxypeptidase Taq inhibitor

Figure 8: Activity prediction of compound 6 using online PASS prediction tool

Activity Prediction

42 Substructure descriptors; 0 new.

Antiviral

Antiviral (HIV)

HIV-1 integrase inhibitor

HIV-1 integrase (3'-Processing) inhibitor

HIV-1 integrase (Strand Transfer) inhibitor

HIV-1 reverse transcriptase inhibitor

85 Possible activities at Pa > 30%

Pa Pi for Activity:

0,802	0,012	Antiischemic, cerebral
0,660	0,021	HIV-1 integrase inhibitor
0,570	0,010	Antiarrhythmic
0,563	0,021	Cardiotonic
0,587	0,080	Membrane integrity agonist
0,578	0,097	Antiviral
0,518	0,067	CXC chemokine 1 receptor antagonist
0,453	0,021	Sodium channel blocker

Figure 9: Activity prediction of compound 7 using online PASS prediction tool

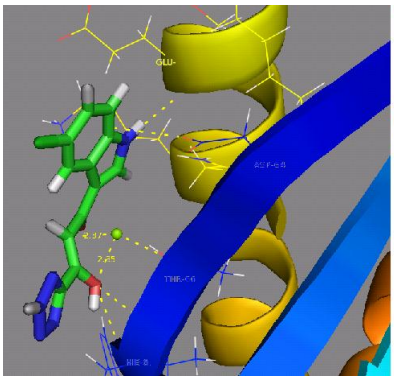


Figure 10: Interaction of Drug Compound with amino acids and Mg^{2+} (metal ion) at the active site of HIV-1 IN

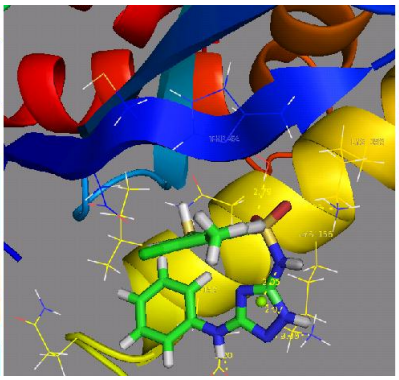


Figure 11: Interaction of compound 1 with amino acids and Mg^{2+} (metal ion) at the active site of HIV-1 IN

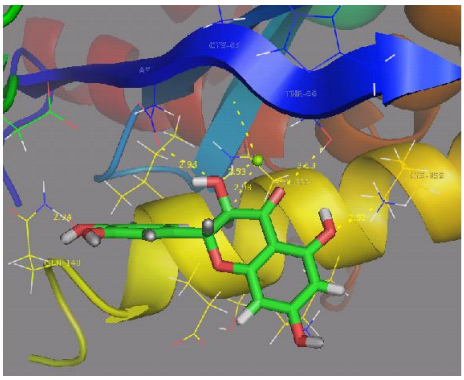


Figure 12: Interaction of compound 2 with amino acids and Mg^{2+} (metal ion) at the active site of HIV-1 IN

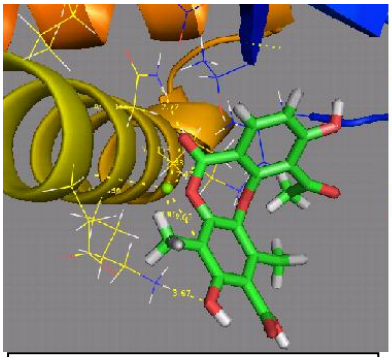


Figure 13: Interaction of compound 3 with amino acids and Mg^{2+} (metal ion) at the active site of HIV-1 IN

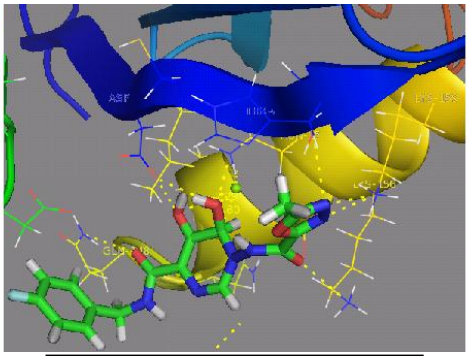


Figure 14: Interaction of compound 4 with amino acids and Mg^{2+} (metal ion) at the active site of HIV-1 IN

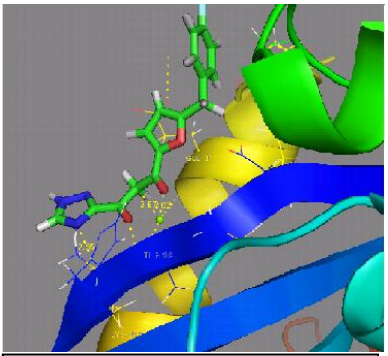


Figure 15: Interaction of compound 5 with amino acids and Mg^{2+} (metal ion) at the active site of HIV-1 IN

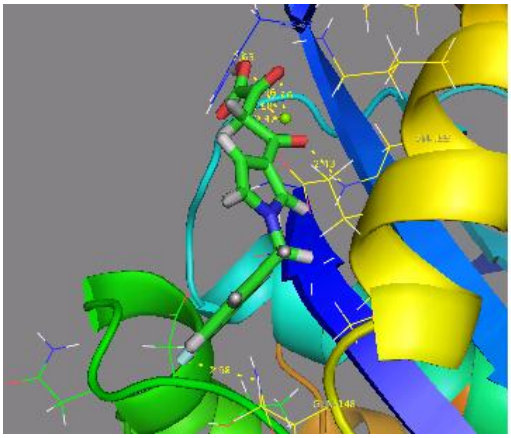


Figure 16: Interaction of compound 6 with amino acids and Mg^{2+} (metal ion) at the active site of HIV-1 IN

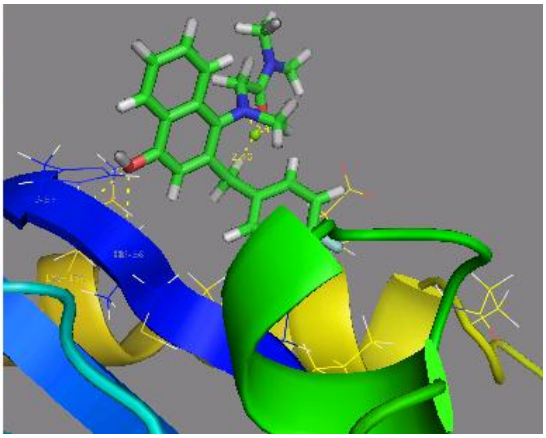


Figure 17: Interaction of compound 7 with amino acids and Mg^{2+} (metal ion) at the active site of HIV-1 IN

Table 1. Interactions of the compounds with amino acids at the active site of HIV-1 Integrase

Compound with HIV-1 Integrase	**Interaction at the active site**	**Distance between Donor and Acceptor (Å)**
Drug Compound (5-CITEP)	(N-H…O) GLU 152	2.71
	(N-H…O) ASP 64	2.92
	HIS 67 (N-H…O)	2.81
	O-H…Mg^{2+}	2.35
	C-H… Mg^{2+}	2.37
Compound 1	THR 66 (O-H…O)	2.79
	LYS 159 (N-H…O)	3.03
	(N-H...N) LYS 156	3.30
	(N-H...O) GLU 152	3.20
	N-H... Mg^{2+}	2.07
	N-H... Mg^{2+}	2.05
Compound 2	LYS 159 (N-H…O)	2.52
	(O-H…O) THR 66	3.13
	(N-H…O) ASP 64	2.94
	(O-H…O) GLN 148	2.34
	O-H… Mg^{2+}	2.53
Compound 3	LYS 159 (N-H…O)	2.33
	LYS 156 (N-H…O)	3.07
	ASN 155 (N-H…O)	2.72
	C-H… Mg^{2+}	2.16
Compound 4 (*Raltegravir*)	LYS 159 (N-H…O)	2.72
	LYS 156 (N-H…O)	2.46
	ASN 155 (N-H…O)	2.58
	GLN 148 (N-H…O)	2.87
	(O-H…O) ASP 64	2.95
	THR 66 (N-H…O)	2.68
	O-H… Mg^{2+}	2.29
	C-H… Mg^{2+}	2.30
Compound 5	LYS 159 (N-H…N)	2.65
	(N-H…O) GLU 152	2.91
	THR 66 (N-H…O)	2.55
	C-H… Mg^{2+}	2.27

Compound 6	ASN 155 (N-H…O)	2.43
	HIS 67 (N-H…O)	2.63
	GLN 148 (N-H…CL)	2.98
	O-H… Mg^{2+}	2.18
	C-H… Mg^{2+}	2.47
Compound 7	(N-H…O) HIS 67	2.58
	(N-H…O) THR 66	2.82
	LYS 159 (N-H…O)	2.97
	O-H… Mg^{2+}	2.18
	C-H… Mg^{2+}	2.40

Table 2. Docking Score and Docking Energies of the compounds docked

Compound	Glide Score	Glide Energy (Kcal/Mol)
Drug Compound (5-CITEP)	-5.690	-61.140
Compound 1	-8.356	-65.178
Compound 2	-6.175	-67.127
Compound 3	-8.469	-68.036
Compound 4 (Raltegravir)	-4.004	-56.466
Compound 5	-4.171	-40.038
Compound 6	-7.300	-61.177
Compound 7	-5.867	-67.631

53 physically significant descriptors and pharmaceutically relevant properties of raltegravir and their analogues were calculated using Qikprop module (ADME). Among them were, Molecular weight, H-bond donors, H-bond acceptors, log Kp (skin permeability), Percentation of human oral absorption, log BB (blood-brain barrier entry) and SASA (total solvent accessible surface area). Lipinski's thumb rule of 5 is a rule to evaluate the drug likeness, or to determine if a chemical compound with a certain pharmacological or biological activity has properties that would make it a likely orally active drug in humans. The rule describes the molecular properties for the drug's pharmacokinetics in the human body, including its ADME. The only non-supporting factor of the rule is that, it does not predict if a compound is pharmacologically active. Although a large search region was used for the docking simulations, the ligands investigated here are found to bind preferably in similar ways close to the active site. Except *Raltegravir,* all compounds abide the Lipinski's Rule. *Raltegravir* maintains high Molecular Weight of 450 (range->130-725), Solvent Accessible Surface Area, Volume and Number of Hydrogen Bond Donors due to the size of the molecule. Compounds 1 and 7 show very good Percentation of Human Oral Absorption, 92% and 100%, respectively. Compound 3 exhibits the least Human Oral Absorption rate. All the compounds satisfy the physicochemical properties generated by the software. Table 3 shows the principal descriptors calculated for *raltegravir* and its analogues by Qikprop simulation. A compound can be termed that it has got drug-likeability when, it exhibits similar or more hydrogen bond interactions, better glidescore and energy than the original drug present in the market. Glidescore (Gscore) resembles an experimental binding mode as determined by X-ray crystallography. Binding energy is the energy released due to formation of complex between an inhibitor and a protein. The complex is considered more stable than the drug, when it is formed with the least glide score and glide energy possible. Ranking analysis of all compounds have been carried w.r.t to glidescore, glide energy, hydrogen bond interactions and ADME properties of commercial drug (*Raltegravir*). Compound 1, 6 and 7 satisfies all the above requirements and have drug-likeness ability. The result shows that all compounds are indeed HIV-1 IN inhibitors, as they showed similar interactions with Lys 156, Lys 159, catalytic residue Glu 152 and also exhibited better glidescore and glide energy than the commercial drug Raltegravir.

Table 3. Principal descriptors calculated for all compounds using Qikprop simulation

Ligand Name	Molecular Weight (130-725)	SASA (300-1000)	Volume (500-2000)	HB Donors (0-6)	HB Acceptors (2-20)
Drug Compound	289.6	502	831.7	1.000	4.250
Compound 1	395.8	646	1103	2.000	6.000
Compound 2	304.2	517	877.4	4.000	6.450
Compound 3	360.2	529	950.8	2.000	6.500
Compound 4	444.4	740	1326	1.000	9.500
Compound 5	313.2	566	964.8	0.000	6.500
Compound 6	289.2	536	912.4	1.000	6.500
Compound 7	366.4	646	1170	1.000	4.750

Ligand Name	QPlogS (-6.5 to 0.5)	QPlogBB (-3.0 to 1.2)	QPlogKp (-8.0 to -1.0)	Percentation of Human Oral Absorbtion (>80%-high <25%-poor)	Rule of Five (Max 4)
Drug Compound (5-CITEP)	-3.462	-1.719	-4.663	68.175	0
Compound 1	-5.385	-1.055	-2.654	92.675	0
Compound 2	-2.715	-2.235	-5.356	52.560	0
Compound 3	-3.133	-2.213	-6.185	36.310	0
Compound 4 (Raltegravir)	-5.063	-1.988	-3.847	66.065	1
Compound 5	-2.782	-1.292	-3.067	79.427	0
Compound 6	-3.022	-1.351	-3.442	69.234	0
Compound 7	-5.101	-0.206	-0.206	100.000	0

The coherent pictures of possible interactions of compounds/ligands at the active site provides as improved basis for Structure Based Ligand Design. The recurring motif of tight interaction with the two lysine residues 156 and 159 is suggested to be of prime importance. Docking of the compounds shown in Figures (10-17) revealed a consistent set of recurring binding modes. The top clusters are always associated with highest frequency of occurrence, which suggest good convergence behaviour of the search algorithms. The most important interactions are summarized in table and graphical representations of the binding modes of all compounds taken into study are given in separate figures.

The active site is clearly dominated by polar residues and the challenge for the design of tighter binding compounds is to make maximum use of the available interaction partners and minimize the conflicts, avoiding especially the frustration of potential hydrogen binding sites of the ligand. Results showed that structurally similar ligands/compounds bind in a very similar position to the original drug in the active site of the target protein. The most potent inhibitor should have the best interaction with HIV-1 IN. The combined approach of docking-ADME screening-PASS prediction used in this work helps in expressing the binding affinity of a set of ligands in the receptor well and also validates them as potential candidates for second generation drug discovery. The most straightforward method of evaluating the accuracy of a docking procedure is to determine how closely the lowest energy pose (binding conformation) is predicted by the object scoring function. In the current study, the final binding energy (glide energy) of the complexes after docking, show that the new compounds are more favorable for their anti-viral activities.

Conclusions

The inhibition of HIV-1 IN promises to be a favorable therapeutic approach in the treatment of viral diseases. All the compounds have shown direct interactions with the amino acids and the metal ion (Mg^{2+}) present at the active site, the target HIV-1 IN. The compounds are found to interact extremely well, which is essential for substrate binding in case of HIV-1 IN. The active site of the target HIV-1 IN is

enriched with residues ranging from 50-212 and it has been found that the compounds exhibit interactions with all of them. Further, ADME screening and PASS online prediction of the ligands have proved to be fruitful in the study as they helped in suggesting that the ligands taken into study can be used for second generation drug development. The interactions at the CCD active site of HIV IN suggest that they are potential antagonists of viral diseases. Therefore, these compounds would be suitable inhibitors targeting HIV-1 Integrase.

Acknowledgement

DV and SN thank UGC, Government of India for financial support for this research and to purchase Schrödinger software. DV thanks DST-FIST and UGC-SAP for funding facilities to the Centre for Advanced Study in Crystallography and Biophysics. Bioinformatics Infrastructure Facility provided to the University of Madras by the Department of Biotechnology, Government of India is gratefully acknowledged.

References

Asante-Appiah, E., Skalka, A.M. (1999). HIV-1 integrase: structural organization, conformational changes, and catalysis. *Adv. Virus Res.*, 52, 351-369. http://dx.doi.org/10.1016/S0065-3527(08)60306-1

Chen, J.C., Krucinski, J., Miercke, J., Finer Moore, J.S., Tang, A.H., Leavitt, A.M., Stroud, R.M. (2000). Crystal structure of the HIV-1 integrase catalytic core and C-terminal domains: a model for viral DNA binding. *Proc. Natl. Acad. Sci. USA*, 97, 8233-8238. http://dx.doi.org/10.1073/pnas.150220297

Chiu, T.K., Davies, D.R. (2004). Structure and function of HIV-1 integrase. *Curr. Top. Med. Chem.*, 4, 965-977. http://dx.doi.org/10.2174/1568026043388547

Christoph, A.S., Haihong, N., McCammon, J.A. (2000). HIV-1 Integrase Inhibitor Interactions at the Active Site: Prediction of Binding Modes Unaffected by Crystal Packing. *J. Am. Chem.*, 122, 6136-6137. http://dx.doi.org/10.1021/ja001152x

De Clercq, E. (2004). Antiviral drugs in current clinical use. *J. Clin. Virol.*, 30, 115-133. http://dx.doi.org/10.1016/j.jcv.2004.02.009

De Clercq, E. (2005). New approaches toward anti-HIV chemotherapy. *J. Med. Chem.*, *48*, 1297-313. http://dx.doi.org/10.1021/jm040158k

Dixon, J.S., Blaney, J.M. (1998). Designing Bioactive Molecules: Three-dimensional techniques and applications. Edited by Martin, Y.C., Willet, P., Washington. DC, *American Chemical Society*, 175-197.

Dyda, F., Hickman, A.B., Jenkins, T.M., Engelman, A., Craigie, R., Davies, D.R. (1999). Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science*, 266, 1981-1986. http://dx.doi.org/10.1126/science.7801124

Eijkelenboom, A.P., Sprangers, R., Hard, K., Puras Lutzke, R.A., Plasterk, R.H., Boelens, R., Kaptein, R. (1999). Refined solution structure of the C-terminal DNA-binding domain of human immunovirus-1 integrase. *Proteins*, 36, 556-564. http://dx.doi.org/10.1002/(SICI)1097-0134(19990901)36:4<556::AID-PROT18>3.0.CO;2-6

Engelman, A., Mizuuchi, K., Craigie, R. (1991). HIV-1 DNA integration: Mechanism of viral DNA cleavage and DNA strand transfer. *Cell*, 67, 1211-1221. http://dx.doi.org/10.1016/0092-8674(91)90297-C

Esposito, D., Craigie, R. (1999). HIV integrase structure and function. *Adv. Virus Res.*, 52, 319-333. http://dx.doi.org/10.1016/S0065-3527(08)60304-8

Gerton, J.L., Brown, P.O. (1997). The Core Domain of HIV-1 Integrase Recognizes Key Features of Its DNA Substrates. *The Journal of Biological Chemistry*, 272, 25809-25815. http://dx.doi.org/10.1074/jbc.272.41.25809

Goldgur, Y., Dyda, F., Hickman, A.B., Jenkins, T.M., Craigie, R., Davies, D.R. (1998). Three new structures of the core domain of HIV-1 integrase: an active site that binds magnesium. *Proc. Natl. Acad. Sci. USA*, 95, 9150-9154. http://dx.doi.org/10.1073/pnas.95.16.9150

Goldur, Y., Craigie, R., Cohen, G.H., Fujiwara, T., Yoshinaga, T., Fujishita, T., Sugimoto, H., Endo, T., Murai, H., Davies, D.R. (1999). Structure of the HIV-1 integrase catalytic domain complexed with an inhibitor: A platform for antiviral drug design. *Proc. Natl. Acad. Sci. U. S. A.*, 96, 13040. http://dx.doi.org/10.1073/pnas.96.23.13040

Greenwald, J., Le, V., Butler, S.L., Bushman, F.D., Choe, S. (1999). The mobility of an HIV-1 integrase active site loop is correlated with catalytic activity. *Biochemistry*, 38, 8892-8898. http://dx.doi.org/10.1021/bi9907173

Hazuda, D.J., Felock, P., Witmer, M., Wolfe, A., Stillmock, K., Grobler, J.A., Espeseth, A., Gabyelski, L., Schleif, W., Blau, C., Miller, M.D. (2000). Inhibitors of strand transfer that prevent integration and inhibit HIV-1 replication in cells. *Science*, 287, 646-650. http://dx.doi.org/10.1126/science.287.5453.646

Kuo, C.L., Assefa, H., Kamath, S., Brzozowoski, Z., Slawinski, J., Saczewski, F., Buolamwini, J.K., Neamati, N. (2004). Application of CoMFA and CoMSIA 3D-QSAR and Docking Studies in Optimization of Mercaptobenzenesulfonamides as HIV-1 Integrase Inhibitors. *J. Med. Chem.*, 47, 385-399. http://dx.doi.org/10.1021/jm030378i

Leroy, D., Kajava, A.V., Frei, C., Gasser, S.M. (2001). Analysis of etoposide binding to subdomains of human DNA topoisomerase IIa in the absence of DNA. *Biochemistry*, 40, 1624-1634. http://dx.doi.org/10.1021/bi0019141

Lodi, P.J., Ernst, J.A., Kuszewski, A., Hickman, A.B., Engelman, A., Craigie, R., Chlore, G.M., Gronenborn, A.M. (1995). Solution structure of the DNA binding domain of HIV-1 integrase. *Biochemistry*, 34, 9826-9833. http://dx.doi.org/10.1021/bi00031a002

Ni, H., Sotrifier, C.A., McCammon, J.A. (2001). Ordered water and ligand mobility in the HIV-1 integrase-5CITEP complex: A molecular dynamics study. *J. Med. Chem.*, 44, 3043-3047. http://dx.doi.org/10.1021/jm010205y

Smith, P.A., Sorich, M.J., Low, L.S.C., McKinnen, R.A., Miners, J.O. (2004). Towards integrated ADME prediction: past, present and future directions for modelling metabolism by UDP-glucuronosyltransferases. *J. Mol. Graphics Modeling*, 22, 507-517. http://dx.doi.org/10.1016/j.jmgm.2004.03.011

ter Heine, R., Hillebrand, M.J.X., Rosing, H., van Gorp, E.C.M., Mulder, J.W., Beijnen, J.H., Huitema, A.D.R. (2009). Quantification of the HIV-integrase inhibitor raltegravir and detection of its main metabolite in human plasma, dried blood spots and peripheral blood mononuclear cell lysate by means of high-performance liquid chromatography tandem mass spectrometry. *J. Pharm. Biomed. Analysis*, 49, 451-458. http://dx.doi.org/10.1016/j.jpba.2008.11.025

Vink, C., Oude Groeneger, A.M., Plasterk, R.H. (1993). Identification of the catalytic and DNA-binding region of the human immunodeficiency virus type 1 integrase protein. *Nucleic Acids Res.*, 21, 1419-1425. http://dx.doi.org/10.1093/nar/21.6.1419

Zheng, R., Jenkins, T.M., Craigie, R. (1996). Zinc folds the N-terminal domain of HIV-1 integrase, promotes multimerization, and enhances catalytic activity. *Proc. Natl. Acad. Sci. U. S. A.*, 93, 13659. http://dx.doi.org/10.1073/pnas.93.24.13659