Perspectives on Developing Small Molecule Inhibitors Targeting HIV-1 Integrase

J.J. Tan[1], C. Liu[1], X.H. Sun[1], X.J. Cong[1], L.M. Hu[1], C.X. Wang*[,1] and X.J. Liang*[,2]

[1]*College of Life Science and Bioengineering, Beijing University of Technology, Beijing 100124, China*

[2]*Laboratory of Nanomedicine and Nanosafety, Division of Nanomedicine and Nanobiology, National Center for Nanoscience and Technology, China, and CAS Key Laboratory for Biomedical Effects of Nanomaterials and Nanosafety, Chinese Academy of Sciences, Beijing 100190, P.R. China*

Abstract: HIV-1 integrase (IN) is a crucial enzyme in the life cycle of HIV-1 and also a validated target for developing anti-HIV inhibitors. Recent progress in drug design has significantly accelerated the development of anti-AIDS IN inhibitors. A large amount of novel inhibitors that interact specifically with IN were developed along with the expanding and application of methods to drug design. This article reviewed the anti-HIV IN inhibitors discovered by the rational drug design approaches in the recent 5-year.

Keywords: *de novo* drug design, HIV, integrase inhibitors, ligand-based drug design, pharmacophore, receptor-based drug design, virtual screening.

1. INTRODUCTION

Caused by the human immunodeficiency virus type 1 (HIV-1), acquired immunodeficiency syndrome (AIDS) remains one of the most formidable and destructive diseases in the world, after being discovered for 30 years. It is therefore essential to develop novel drugs for AIDS treatment [1-2]. HIV-1 integrase (IN) is a crucial enzyme to the infection of host cells by catalyzing the insertion of viral DNA into the genome of host cells, which makes it a validated target for anti-HIV inhibitor development. The full-length IN monomer includes one peptide chain which is made of three distinct functional domains [3-4]: the N-terminal domain (NTD) (residues 1-50), which comprises a conserved HHCC-binding motif that coordinates one zinc ion; the catalytic core domain (CCD) (residues 51-211), which contains a canonical three-amino acid DDE motif corresponding to D64, D116 and E152; the C-terminal domain (CTD) (residues 212-288), which is important for IN-IN and IN-DNA interactions. MK-0518 (Fig. (**1**)), also called Raltegravir, which was developed by Merck in 2007 [5-7], is the first IN inhibitor approved by US FDA. It inhibits HIV-1 IN at low nanomolar concentrations by targeting the strand transfer or joining reaction, and represents a great breakthrough in this field. The traditional drug development strategy based on combinatorial chemistry and high throughput screening (HTS) is costly and inefficient. In addition, most lead compounds obtained *via* traditional strategy are inactive or toxic *in vivo* or in human beings. Currently it costs about $802,000,000 to bring a new drug to the market [8]. Apparently, a novel and efficient approach for drug development is urgently desirable to replace the traditional strategy. A rational drug design that targets the HIV-1 IN normally involves understandings of the structure of drug targets-receptors (IN), ligand binding kinetics, and drug/receptor interactions based on structural data. The rational drug design approaches falls into several natural categories: receptor-based drug design (RBDD), ligand-based drug design (LBDD) and other combinatorial approaches with RBDD and LBDD.

The objective of this review is to provide an overview of the basic methods about rational drug design and its application on designing novel anti-HIV IN inhibitors.

2. RECEPTOR-BASED DRUG DESIGN

RBDD refers to the drugs design based on the computerized 3D structure of a drug target (receptor) which can be an enzyme, an ion channel, a transporter, or other types of protein. The receptor structure is attained by X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy or comparative modeling methods. RBDD studies could usually be categorized into molecular docking and Quantum mechanics/molecular mechanics method. The key of both methods is the detailed atomic-level description of the binding site (sometimes with a substrate or a known drug). The application of this method has been critically discussed in recent years [9-10].

2.1. Molecular Docking and Molecular Dynamics Simulations

Molecular docking is *in silico* study computing the binding energy and geometry between targets and ligands,

*Address correspondence to these authors at the College of Life Science and Bioengineering, Beijing University of Technology, Beijing 100124, China; Tel: 86-10-67392724; Fax: 86-10-67392837; E-mail: cxwang@bjut.edu.cn
Laboratory of Nanomedicine and Nanosafety, Division of Nanomedicine and Nanobiology, National Center for Nanoscience and Technology, China, and CAS Key Laboratory for Biomedical Effects of Nanomaterials and Nanosafety, Chinese Academy of Sciences, Beijing 100190, P.R. China; Tel: 86-10-82545569; Fax: 86-10-62656765; E-mail: liangxj@nanoctr.cn
A teaser: This article presents details of the anti-HIV integrase inhibitors discovered by the rational drug design approaches.

1875-5607/12 $58.00+.00

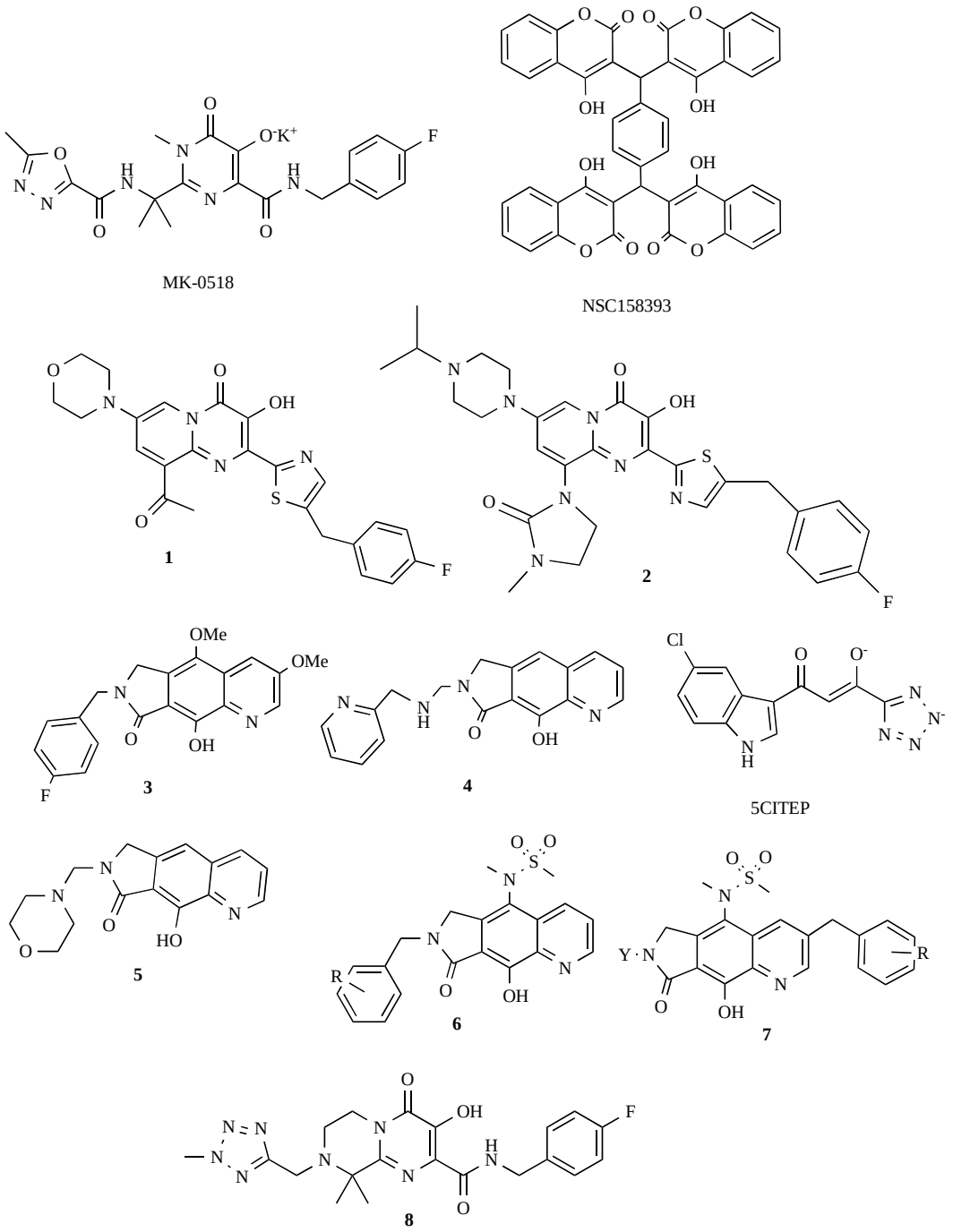

Fig. (1). The structures of MK-0518, NSC158393, 5CITEP and compound **1** to **8**.

substrates or possible drug candidates. It allows users to identify the ligands that fit into the given binding site and to predict its binding affinities. A molecular docking calculation consists of the following steps: (i) optimizing the ligand geometry, calculating pH-dependent partial charges, and identifying rotatable bonds; (ii) calculating electrostatic properties of the receptor and defining the ligand-binding region; (iii) evaluating receptor-ligand interaction by a scoring function [11-14] that includes terms and equations describing the intermolecular energies. The result of a docking calculation is a series of ligand-protein complex geometries along with the corresponding binding energies. Different algorithms are employed to accurately explore the space of possible conformations. GOLD [15] and AutoDock [16] are among the most well-known docking programs based on Genetic Algorithms (GA). Other programs such as DOCK [17], FlexX [18], and Surflex [19], use the fragment-based methods which split the ligand into pieces and dock them to the binding site in an incremental way.

Since most of the current docking programs assume the receptor to be rigid [20], molecular dynamics (MD) simulations are often performed on the protein to sample

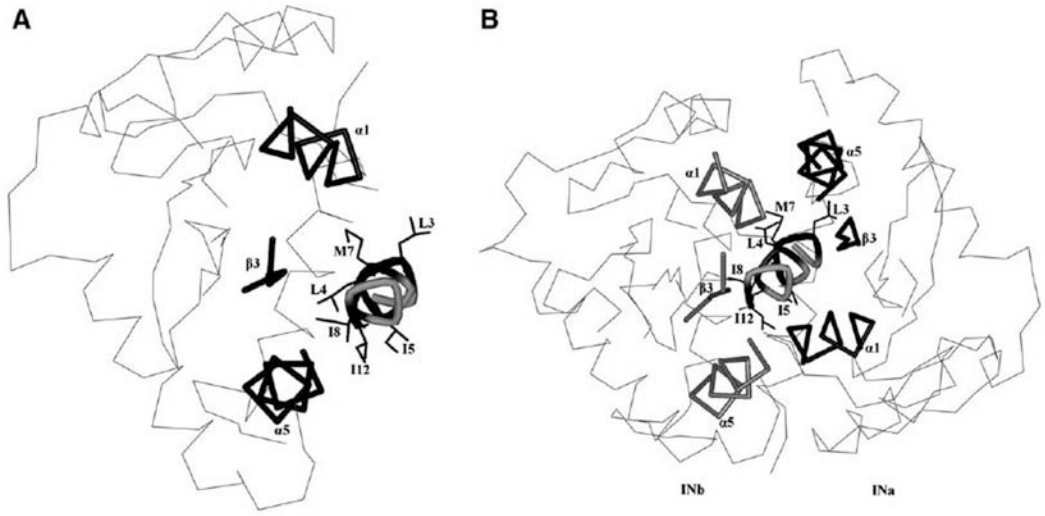

Fig. (2). The binding modes of EBR28 with (**A**) IN1 and (**B**) IN_2.

multiple conformations of drug binding site, compute binding free energy, and investigate the effect of explicit solvent molecules and binding free energy [21].

In general, molecule docking and MD simulation are performed jointly to build ligand-receptor complex model and then to explore their binding modes and inhibition mechanism, or drug resistance.

2.1.1. Binding Modes and Inhibition Mechanism

To date, the crystal structures of the isolated IN domains CCD and the CCD in complex with either the CTD or the NTD had been resolved [22-24]. Very recently, crystal structure of the tetrameric integrase of the prototype foamy virus (PFV) co-crystallized with its cognate DNA has been published in Nature [25]. It is worth mentioning that the solution of PFV intasome structure and relative mechanism studis of IN and IN inhibitors have huge consequences in this filed and challenge a lot literatures published before [26]. To have a whole perspective on IN inhibitors developing, here we introduce both results published before and after the PFV intasome. Despite the prominent contribution of PFV intasome on understanding the structure and function of IN, the crystal structure of full-length HIV-1 IN remains obscure due to its insolubility and easy polymerization. Thus, molecule modeling, especially homology modeling, of required binding complex based on docking and MD simulation is deployed extensively to improve our understanding of the interaction between the inhibitors and IN active site.

Chen *et al.* [27] positioned and developed an IN-DNA model including IN domain CCD and CTD by means of the crystallographic structure of Tn5 bacterial transposase - DNA complex. The resulted complex model including the IN active-site model, viral DNA and inhibitor showed a common binding mode and similar interaction with different inhibitors. Interestingly, it was observed that only one of two catalytic Mg^{2+} is fully coordinated by IN and viral DNA but not available for chelating an incoming inhibitor. This result is different with the binding modes obtained by our group, which requires an inhibitor to interact with both two active - site Mg^{2+} [28-29]. Besides, analyses indicate that IN is apt to form multimers during integration and many oligomeric, dimer or trimer models, were constructed to reveal the mechanism of integration [30-31].

The peptide EBR28, which exhibits a strong inhibitory activity against IN binding with viral DNA, is one of the most potential peptide anti-IN inhibitors. Binding modes of EBR28 with HIV-1 IN monomer core domain (IN_1) and dimmer core domain (IN_2) were investigated by our group, using molecular docking and MD simulation methods [32]. MD indicated that EBR28 bounds to the interfaces of IN_1 and IN_2 systems mainly through the hydrophobic interactions with β3, α1 and α5 regions of the protein (Fig. (**2**)). EBR28 may prevent the binding between IN and viral DNA through forming the IN_1-EBR28 complex which prevents the formation of IN dimmer.

2.1.2. Drug Resistance

Coumarins are a class of potent IN inhibitors maintaining good inhibitory activity. This type of compound does not contain the catechol group which is proposed to induce cytotoxicity in many inhibitors. To investigate the binding modes and drug resistance mechanisms of Coumarins, our lab [33] built three binding models of the hydroxycoumarin compound NSC158393 (Fig. (**1**)) with the wild type IN (WT) and two mutants type, W132G and C130S, using the relaxed complex molecular docking Method [34]. The obtained models were subject to MD simulations. Docking results showed that the specific binding modes for these three systems are different. In WT model, NSC158393 can effectively depress the stability of IN_2 by causing a steric hindrance around the monomer interface. However, for the two mutants, NSC158393 is not close enough to the dimer interface. Energy decomposition analysis revealed three key binding residues, W131, K136, and G134. W131 seems to be indispensable for the ligand binding. In addition, the W131 and G134 mutants were found to be correlated with the

flexibility of the 128-136 peptides and the flexibilities of the 140s loops. That may be another important reason to explain the drug resistance of IN.

Perryman *et al.* [35] described a model of the active site of this drug target which was produced by developing a restrained molecular dynamics protocol. This model ensures the presence of the catalytic DDE motif and, furthermore, could be used to explore the effects of drug resistance mutations on the dynamic flexibility and conformational penchants of HIV IN.

For the discovery and development of new IN inhibitors, the molecular characterization of resistant viruses is a critical step [36]. Le, G. *et al*. [37] reported a series of novel HIV IN inhibitors (Fig. (**1**) compound **1**, **2**) comprising a hydroxypyridinone and a thiazole ring. These compounds have higher potency against all major clinically relevant Raltegravir-resistant HIV strains at nanomolar level, including double mutants Q148HG140S and N155HE92Q.

2.2. Quantum Mechanics/Molecular Mechanics

The hybrid quantum mechanics/molecular mechanics (QM/MM) approaches are the combinational application of quantum computation and classical molecular mechanics in different regions of ligand-receptor system. In general, the system is divided into a QM region, where is the ligand or protein active site, and a MM region, where is the rest of the system such as the surrounding amino acids and solvent environment. One of the advantages of QM/MM is that ligand polarization upon binding is considered into model, which is important to the studies involving metal effects.

In 2007, Nunthaboot *et al.* [38] performed conventional MM and QM/MM simulations with the initial structure of HIV IN aligned with inhibitor 5CITEP (Fig. (**1**)) [39]. Results obtained from the two methods were different. Firstly, MM study showed that the ligand is positioned far from Lys159 residue and they cannot form salt bridge. On the contrary, there is a clear salt linkage between Lys159 and inhibitor in the QM/MM study. In addition, loop residues 140-149 undergo distinct fluctuations in two systems, which leads to variation binding modes between ligands and proteins. Meanwhile, the attractive difference is the coordination geometry of Mg^{2+}, which is octahedral and distorted octahedral in MM and QM/MM study, respectively.

Similar metal coordination for MM and QM/MM has been observed by Alves and his co-workers [40]. QM/MM simulation was carried out to investigate energy between HIV-1 IN and L-731,988, together with its 10 analogues. Compared to the results obtained from common theories, there is no chelation between Mg^{2+} and putative inhibitors in their calculation but only a favorable interaction with the metal. This novel conclusion is in accordance to that reported by Brigo *et al.* [41] Furthermore, they also first pointed out the importance of the interaction between the Mg^{2+} and aromatic ring of inhibitors.

3. LIGAND-BASED DRUG DESIGN

Ligand-based Drug Design (LBDD) is based on the fact that the compounds structurally similar to an active ligand are likely to be active. Ligand-based approaches are widely employed when three-dimensional structure information for the target protein is absent or insufficient. By analyzing a series of known inhibitors, a novel potential inhibitor could be developed from the common features of the known ones. Specific interactions with the target can also be figured out. In cases where the target structure is not available, LBDD methods provide useful chemical information for drug design [42-43].

3.1. Structure-Activity Relationships (SAR)

Jin and his team have previously studied the role of substitutions at 3- and 5-positions on quinoline ring of the pyrrolo[3,4-g]quinolin-8-one inhibitors [44-45]. They firstly obtained three analogues (Fig. (**1**) compound **3**, **4**, **5**) that displayed satisfied IC_{50} values of 0.039 μM, 0.19 μM and 0.036 μM, respectively. Then they synthesized a set of analogs of C5 aza tricyclic quinoline by modifying the fluorobenzyl moiety (Fig. (**6**) compound **6**) [46]. This SAR study revealed that the preferred substituents on the benzyl group of tricyclic moiety are hydrophobic functional groups. In addition, Metobo *et al*. [47] pointed out that the inhibitory potency would be improved when aryl portion of the "benzyl tail" is bound to the pyridine C3 position which is so-called "benzyl flipped" (Fig. (**1**) compound **7**).

Petrocchi *et al*. [48] identified a series of tetrahydro-pyrazinopyrimidine-2- carbox-amides starting from Raltegravir (IC_{50} = 15 nM, Fig. (**1**)), the first marketed HIV integrase inhibitor. Alternative cyclization, chain substitution and insertion of basic group were undertaken. As result, a 2-methyl tetrazole (Fig. (**1**) compound **8**)**,** with IC_{50} = 6 nM, was discovered as the most potent inhibitor. It indicated that tethering the N-methyl group of pyrimidones onto the amino gem-dimethyl group leads to a series of novel potent bicyclic HIV integrase inhibitors.

Donghi *et al*. [49] reported a new template of HIV-1-integrase based on his previous discovery of bicyclic pyrimidinones which were generated by linking the N1-Methyl group of the pyrimidinone scaffold (Fig. (**3**) compound **9**) into a saturated cycle (Fig. (**3**) compound **10**) . To remove stereo-centers, the new template were characterized by a pyrido[1,2-a]pyrimidin-4-one scaffold (Fig. (**3**) compound **11**) instead of 6,7,8,9-tetrahy-dropyrido[1,2-a]pyrimidin-4-one scaffold (Fig. (**3**) compound **10**). A series of derivatives with either amide substitution (Fig. (**3**) compound **12**) or oxalamide substitution at position 9 (Fig. (**3**) compound **13**) showed potent anti-viral activity at nanomolar level (18 to 185 nM for IC_{50}).

Compared to these derivatives with substitution at position 9 , Jones *et al*. [50] discovered a similar bicyclic compound (Fig. (**3**) compound **14**) with IC_{50} value of 80 nM. They subsequently explored the effect of substitution at C7 and C9 positions of **14**. By introducing an electron acceptor or an electron donor into C7 or C9, respectively, the potency in the anti-HIV assay was generally enhanced, which indicated that the IN may undergo different conformation in cell-based system.

Following the discovery of 3-Hydroxy-pyrido[1,2-a]pyrimidin-4-one HIV-1 IN inhibitors, Jones *et al.* [51]

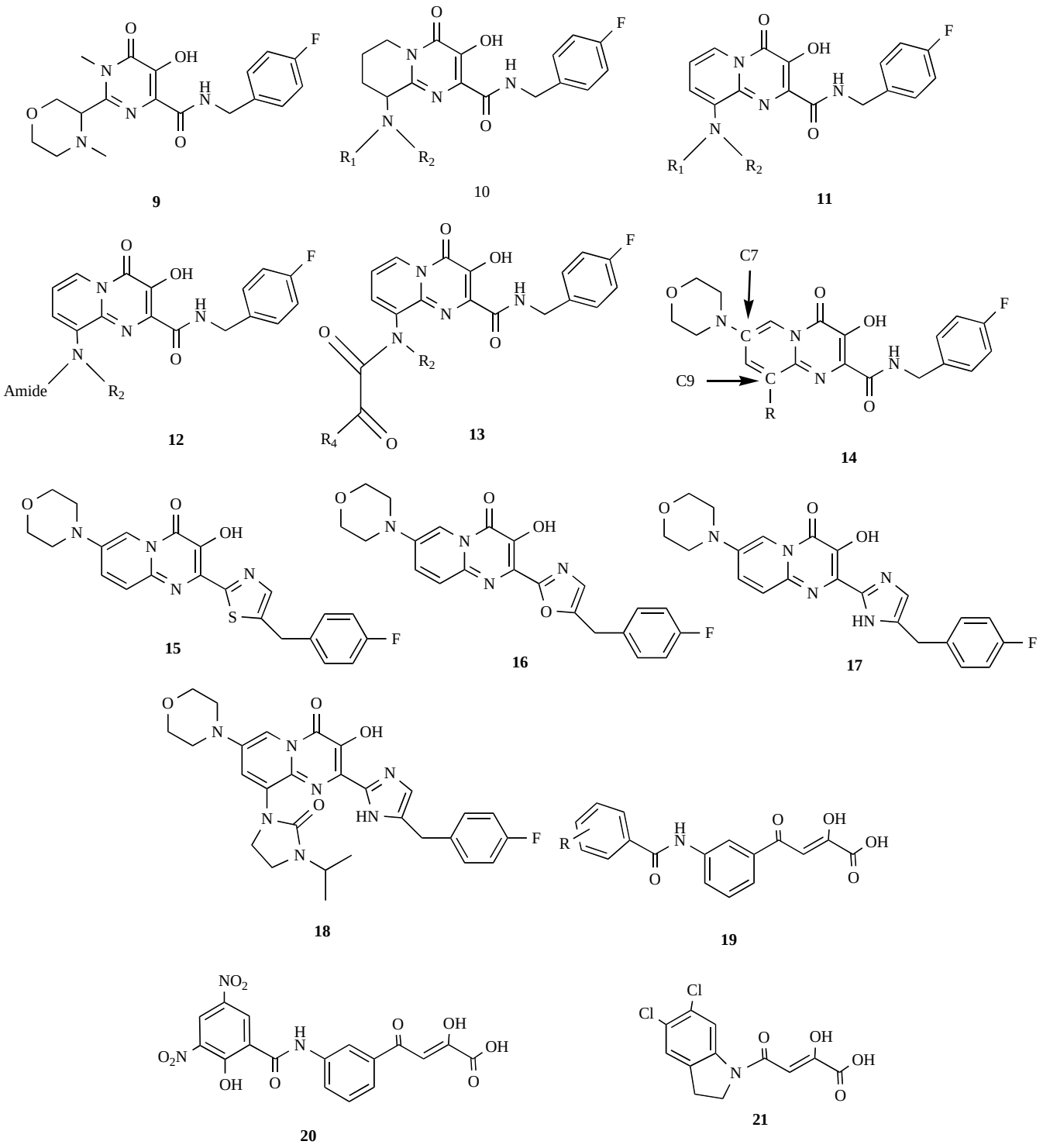

Fig. (3). The structures of compound **9**-**21**.

embarked on further improvement of these compounds by replacing the amide group by a five-member-ring azoles (thiazole, oxazole or imidazole). This effort aimed to provide an efficient chelation to divalent Mg^{2+} ions in the active site of IN. A series of azoles derivatives with replacement by thiazoles, oxazoles, imidazoles, or halogen on position 9 were synthesized and mostly showed significantly higher potency than parent compound with an IC_{50} of 775 nM and EC_{50} of 215 nM. In particular, thiazole (Fig. (**3**) compound **15**, IC_{50} = 20 nM, EC_{50} = 32 nM), oxazole (Fig. (**3**) compound **16**, IC_{50} = 59 nM, EC_{50} = 21.5 nM), and imidazole (Fig. (**3**) compound **17,** IC_{50} = 45 nM, EC_{50} = 62 nM) are outsranding for their significantly improved potent. Additionally, analog of **17** substituted by cyclic urea on position 9 (Fig. (**3**) compound **18**) shows the most promising EC_{50} value of 6 nM.

Diketo acids and its derivatives are considered as promising classes of HIV-1 IN inhibitors. The first generations of IN inhibitors are diketo acids. One of them, named as MK-0518, has been proved by US FDA as the first HIV IN inhibitors. Previous studies have revealed that an amide group in the diketo acid is well tolerated [52-53]. According to the above conclusion, Li *et al*. [54] designed and synthesized a series of novel amide-containing diketo acids (Fig. (**3**) compound **19**). SAR analysis indicated a potent inhibitor (Fig. (**3**) compound **20**). The experimental test revealed that it has a better inhibitory efficacy with the IC_{50} values of 8 μM against 3-processing and 2 μM against strand transfer process. Further modifications to the indole-based diketoacid chemotype explored by Walker *et al*. [55] gave a more active compound (Fig. (**3**) compound **21**) with IC_{50} value of 0.02 μM against strand transfer. In this compound, the indole group is replaced by secondary but not tertiary aniline-based amides without sacrificing inhibitory activity. They proposed that the difference between the activities of secondary and tertiary amides is probably due to the preferred conformation of amide bonds, s-trans for the secondary-amide and s-cis for the tertiary-amide.

The core of a diketo acid-type compound, 4-oxo-quinolizine-3-carboxylic acid, was modified by adding various bulky groups at the C-1 position to investigate the chelation of Mg^{2+} (Fig. (**4**)) [56]. It was proposed that the two-metal chelating scaffold and an appreciate side hydrophobic chain are indispensable for new types of HIV-1 IN inhibitors.

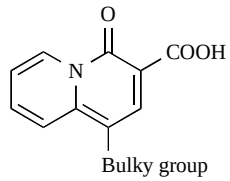

Fig. (4). The structure of 4-oxo-quinolizine-3-carboxylic acid with bulky groups at the C-1 position.

Bifunctional DKAs (BDKAs) are characterized by the presence of two diketo acid chains. Di Santo *et al.* [57] designed and synthesized a set of novel bifunctional quinolonyl diketo acid derivatives, among which compound **22** (Fig. (**5**)) showed good inhibitory ability on both 3'-processing and strand transfer reactions and exhibited low cytotoxicity (EC_{50} = 4.29 μM; CC_{50} > 200 μM). Subsequent docking study further revealed that the ligands primarily interact with IN at metal ions, K156 and/or K159 residues and the hydrophobic pocket, which means that BDKAs can bind both DNA acceptor and donor site of HIV-1 IN. Therefore, they could inhibit both ST and 3'-P processes, whereas the monofunctional DKAs are selectively effective on ST.

Long *et al*. [58] designed and synthesized three new types of aryl diketo acid (ADK) isosteres by converting the biological labile 1, 3-diketo unit into isoxazole, 1H-pyrazole, and isothiazole moieties. During the tested compounds, the best antiviral effect was showed by 5-(4-nitrophenyl)-1H-pyrazole-3-carboxylic acid (Fig. (**5**) compound **23**) and 3-(3-(benzyloxy)phenyl)isoxazole-5-carboxylic acid (Fig. (**5**) compound **24**) with an EC_{50} value of 3.6 and 7.2 μM, respectively. That means that the phenyl-substituted heteroaromatic carboxylic acids can improve antiviral potency and decrease cytotoxicity to IN inhibitors.

Johns *et al.* [59-60] found that a series of 5-substituted derivatives of the validated clinical candidate L-870,810 (Fig. (**5**) compound **25**) have promising potency as inhibitors. They designed a sequence of 5-substituted derivatives of the 1,3,4-oxadiazole (Fig. (**5**) compound **26**) and found that many of their designed molecules have potency similar to previously reported clinical compounds. These results showed that the oxadiazole can be regarded as an amino acid isostere for metal coordination.

Hadi, V. *et al*. [61] have exploited SAR studies of pyrazolone scaffold and discovered one compound exhibiting single-digit micro molar activity against IN strand transfer process. It is predicted that the pyrazolones will become the second-generation IN inhibitors.

Recently, Sechi *et al.* [62] identified a compound that is analogue of roquinimex. Based on which, a new scaffold was designed considering the fact that most inhibitors, for example diketo acids, are involved in coordination with divalent metal ions. It could facilitate the chelation of two metal ions with the active site of IN. Eight more analogues with variety substitutions at 1st (N1) and 6th-position (C6) of designed scaffold were designed and synthesized. These compounds have optimal fitness with the quinolone 3-carboxylic acid pharmacophore, which is derived from a training set of compounds including the clinical candidate GS-9137. In computation and biology test, compound **27** (Fig. (**5**)) is stand out with better inhibitory activity than lead compound (IC_{50} = 0.9). It can be treated as a promising candidate for further optimization.

Lens epithelium-derived growth factor (LEDGF/p75) is a cellular cofactor of HIV-1 integrase that promotes viral integration by tethering the preintegration complex to the chromatin. LEDGF/p75 by itself is known to act as an allosteric activator of IN activity. Christ *et al*. [63] rationally designed a series of 2-(quinolin-3-yl)acetic acid derivatives (LEDGINs) that act as potent inhibitors of the LEDGF/p75-integrase interaction and defined the 2-(quinolin-3-yl)acetic acid derivatives as the first genuine allosteric HIV-1 integrase inhibitors. The IC_{50} and EC_{50} values of the most active compound are 1.37 μM and 2.35 μM, respectively.

3.2. Quantitative Structure-Activity Relationships (QSAR)

In drug design, QSAR analyzes diverse chemical features (i.e., hydrophobic, electrostatic, polar, steric) of various ligands and build a mathematical model that are able to predict biological properties of similar ligands. The resulting models could be used to predict the target properties for new compounds. The most widely used QSAR analysis in drug design is 3D-QSAR which refers to the application of force field calculations requiring three-dimensional structures of the molecules. The process of 3D-QSAR model development includes the following steps:

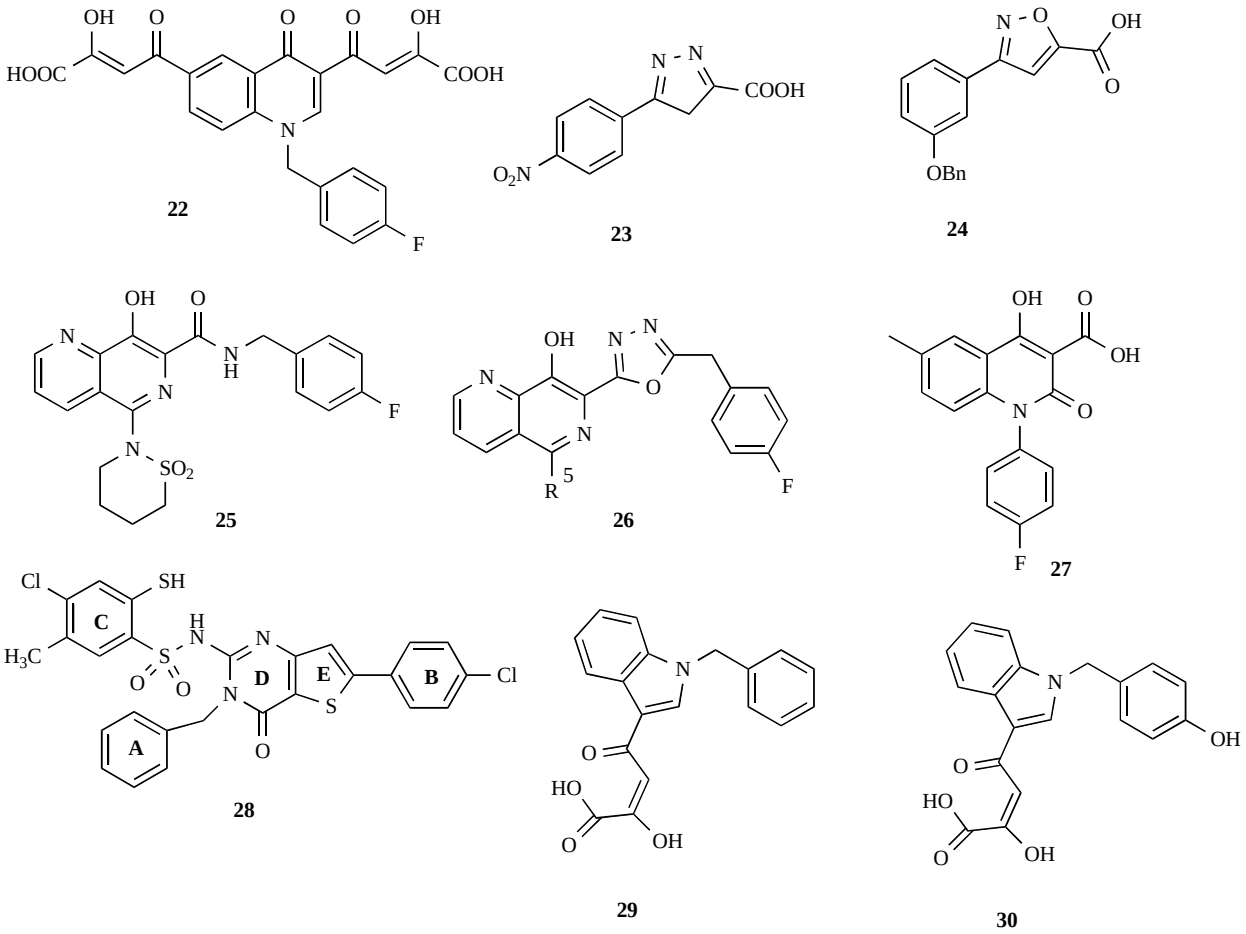

Fig. (5). The structures of compound **22-30**.

(1) Pre-processing of the structure representations. This step includes converting 2D molecular structures into 3D structures and optimizing the 3D structures since the target protein structure is unknown in some cases, it is crucial to determine the conformations of the ligands for the reliability of the final model. There is a challenge for 3D-QSAR model development. In general, the lowest energy conformers are required because they can be reproduced in latter calculations.

(2) Calculation of molecular descriptors. The purpose is to calculate numerical descriptions or characterizations for the structures of the molecules. Several strategies have been developed to calculate and/or select molecular descriptors, such as Distance Geometry (DG) [64], Molecular Shape Analysis (MSA) [65], Comparative Molecular Field Analysis (CoMFA) [66], Comparative Molecular Similarity Indices Analysis (CoMSIA) and so on. Among these methods, CoMFA and CoMSIA are the most frequently used. Recently, quantum mechanics-derived descriptors were proposed, which represent the new frontier of molecular descriptors [67].

(3) QSAR model building. This step is realized by statistically analyzing experiment data in combination with the calculated molecular descriptors. Commonly used statistical structure-property correlation techniques in QSAR include the following types: Partial Least Squares Analysis (PLS) and Partial Least Squares Discriminant Analysis (PLS-DA), Principal Component Analysis (PCA), Clustering, Factorial Analysis (FA), Genetic Algorithm (GA), Hopfiled Neural Network (HNN) or Kohonen Neural Network (KNN), and Support Vector Machine (SVM). The application of these methodologies has been discussed recently [68-72].

CoMFA and CoMSIA models were developed based on a series of 4-chloro-N-(4-oxopyrimidin-2-yl)-2-mercap-tobenzenesulfonamide and 3-aroyl-1,1-dioxo-1,4,2-benzodithiazine derivatives [73]. Total of 41 molecules were divided into training set, which is used to develop the 3D-QSAR model, and test set, which is used to test the predictive ability of obtained model. Based on the best conformation of compound **28** (Fig. (**5**)), a robust CoMFA model was developed with the cross validation $q^2 = 0.728$, non-cross validation $r^2 = 0.934$. Three more physiochemical properties, hydrophobic, hydrogen-bond donor and acceptor, were included besides steric and electrostatic properties in CoMFA. By developing different models with different combination of the five descriptors, the model SH (S and H represent steric and hydrophobic descriptors respectively) is

considered as the best CoMSIA model with $q^2 = 0.794$, $r^2 = 0.928$. Thus, steric and hydrophobic properties are considered important for the binding affinity of these molecules to the target. Subsequent Contour Maps Analysis showed that either negatively charged or hydrophobic substitutions in ring E are unfavorable, whereas electropositive group in both ring A and B improves the activity.

The calculation of molecular descriptors is the critical step of QSAR. Different descriptors adopted in calculation lead to different result of predict capacity of models. Leonard *et al*. [74] performed 3D-QSAR analysis using MSA technique based on 36 styrylquinoline derivatives as potent HIV-1 IN inhibitors. The topological and structural descriptors of these compounds were calculated and the entire data set was divided into training set (n = 26) and test set (n = 10) using K-means clustering technique. Several models were obtained from five statistical methods: stepwise regression, genetic function approximation (GFA), multiple linear regressions with factor analysis (FA-MLR), partial least squares regression with factor analysis (FA-PLS) and genetic partial least squares (G/PLS). The best q^2, r^2 and r^2_{test} are 0.658 for FA-MLR model, 0.794 for G/PLS, model and 0.699 for FA-PLS model respectively.

De Melo *et al*. [75] applied multivariate QSAR method to 33 4,5-dihydroxypyrimidine carboxamides as HIV-1 IN inhibitors and generated a model with $q^2 = 0.58$, $r^2 = 0.87$. The four descriptors being chosen to predict the inhibitory activity of investigated compounds are the energy of the highest occupied molecular orbital, the component vector to the overall polarizability in the Y plane, the total energy, and the sum of the bond electrotopological values of carbon-carbon aromatic bonds in which the carbons are not substituted. All of the four descriptors are related to the electronic distribution, which indicates a requisite relation between the HIV-1 IN inhibition and the electronic distribution of the investigated inhibitors.

Mostly, 3D-QSAR studies are ligand-based, but sometimes the information of receptor is also available to explore the structure-activity relationship of inhibitors. Dhaked *et al*. [76] developed a receptor-based 3D-QSAR method, called comparative residue interaction analysis (CoRIA), This method was employed on 81 IN inhibitors belonging to 13 structurally different classes. Three QSAR models were generated with q^2 values ranging from 0.76 to 0.80. Calculation of non-bonded interaction energies of inhibitors with individual active site residues showed that Asp64, Thr66, Val77, Asp116, Glu152 and Lys159 are the key residues influencing the binding of ligands with IN. The intension of the Van Der Waals interaction of residues Asp64 and Asp116, and the Columbic interaction of residue Thr66 may enhance binding affinity. While reducing the Columbic interaction with Val77 is favorable to the overall binding of IN inhibitors. Guided by CoRIA models, a known compound (Fig. (**5**) compound **29**) was modified to a new compound (Fig. (**5**) compound **30**) and resulted higher anti-HIV activity. The activity value (pIC_{50}) is improved from 4.92 (compound **29**) to 5.22 (compound **30**).

Multi-target QSAR modeling was conducted by Liu *et al*. [77] based on multi-task learning. They formulated the QSAR modeling of HIV-1 inhibitors as a linear ridge regression model and acquired the parameters of this model using a 10-fold cross-validation process. A SAR report was proposed based on their multiple target analysis and provided useful information for designing the potent inhibitors which could bind to HIV-1 on multiple targets.

Cheng *et al*. [78] have applied QSAR methodology for modeling using 3D-Molecular Representation of Structure based on Electron diffraction (3D-MoRSE) descriptors. They proved that combination of replacement method (RM) and support vector machine (SVM) approaches improve QSAR research of carboxylic acid derivatives targeting HIV IN. Therefore, the combination of SVM with RM comes to a useful method for developing novel carboxylic acid derivatives as HIV-1 IN inhibitors.

4. CONJUNCTION WITH RECEPTOR-BASED AND LIGAND-BASED DRUG DESIGN APPROACHES

In the context of the increasing amount of structural information of receptor complex with inhibitors and the predictive results of ligand-based drug design models, utilization of the combination of RBDD and LBDD methods has been useful on discovery of HIV-1 IN inhibitors. A 3D complex model based on the structure of IN could provide binding modes between receptor and ligand. On the other hand, the statistical models of QSAR or pharmacophore based on known active molecules could effectively screen the candidates from a huge amount of database molecules. Combinational use of these two approaches is expected more reliable as it considers both the receptor-ligand interaction and the structure-activity relationship. Recently, some achievements on novel HIV-1 IN inhibitor design were derived from LBDD&RBDD approaches.

Ferro *et al*. [79] combined RBDD and LBDD approaches to investigate the binding mechanism of HIV IN inhibitors and thus obtained a set of attractive leading compounds. Firstly, a ligand-based pharmacophore model was generated and was used for virtual screening. A series of 4-[1-(4-fluorobenzyl)-1H-indol-3-yl]-2-hydroxy-4-oxobut-2-enoic acids were discovered as potent anti HIV-1 IN agents selectively inhibiting strand transfer step of IN. Secondly, acceptor-based docking model of IN-Mg-DNA complex, including the full structure of core domain of IN and especially chelating two positive ions with DDE moiety, was developed. Finally, one compound (Fig. (**6**) compound **31)** was docked into IN by using this model to show the binding mode. Docking results showed thata large hydrophobic cavity is defined by nonpolar residues L68, I73, V75, L158, and I162 and occupied by the 4-fluorophenyl ring of the inhibitor. According to that, a series of derivatives were designed by changing the position of the fluorine or adding other substituent. Experimental test proved the activity of these analogues and showed a better compound (Fig. (**6**) compound **32**) showed a higher activity (IC_{50} = 0.03 μM) and lower toxicity (SI = 175) than the compound **31** (IC_{50} = 0.08, SI = 70).

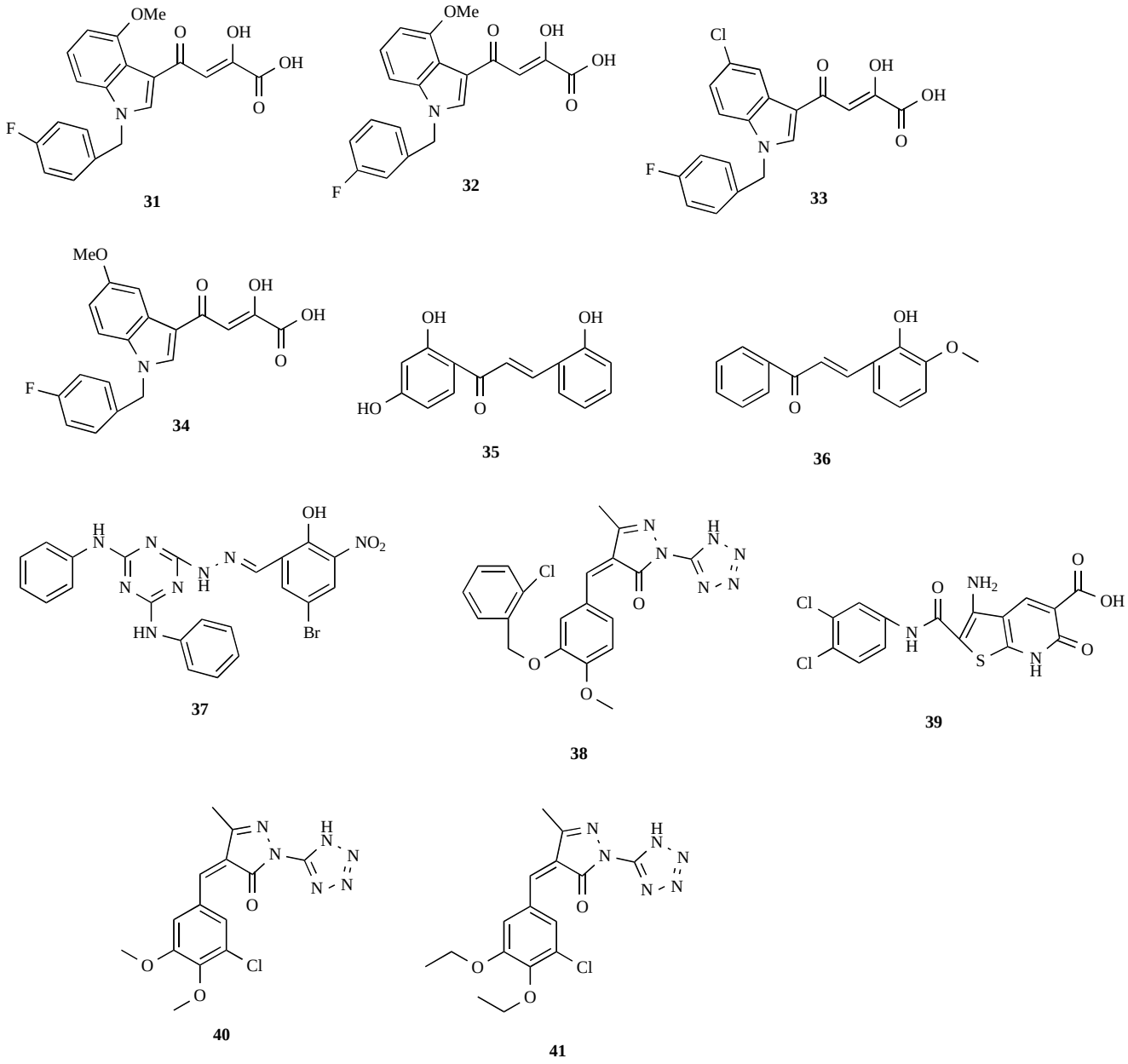

Fig. (6). The structures of compound **31-41**.

5. PHARMACOPHORE MODELING

Pharmacophore modeling combined with 3D database searching technologies plays an important role in drug discovery [80-81]. 3D pharmacophore model is prepared based on either the 3D structure of a protein active site or a set of known active compounds that bind to the same site. A representative pharmacophore model consists of key features of ligands for binding to the macromolecule, through electrostatic, hydrophobic and H-bond interactions. It is possible to search compounds that "fit" to the pharmacophore in 3D databases. Pharmacophore based database searching is successful in finding compounds not only binding to the active site but also with high degree of structural variety. A number of pharmacophore programs and databases are accessible to researchers. A comparison of pharmacophore software packages has been reported by Gillet *et al.* [82].

5.1. Ligand-Based Pharmacophore Modeling

Barreca *et al*. [83] structured 10 quantitative pharmacophore from 33 molecules acting as diketoacid (DKA)-like strand transfer selective IN inhibitors. The best hypothesis consists of one hydrophobic aromatic feature, two H-bond acceptors, and one H-bond donor feature, that are in full agreement with their previously manually developed pharmacophore [84]. It is noticeable that most of the inactive molecules of the data set are unable to map the hydrophobic aromatic feature. Base on this observation, they designed a series of new derivatives bearing a substituent that might interact with the hydrophobic aromatic feature. Two

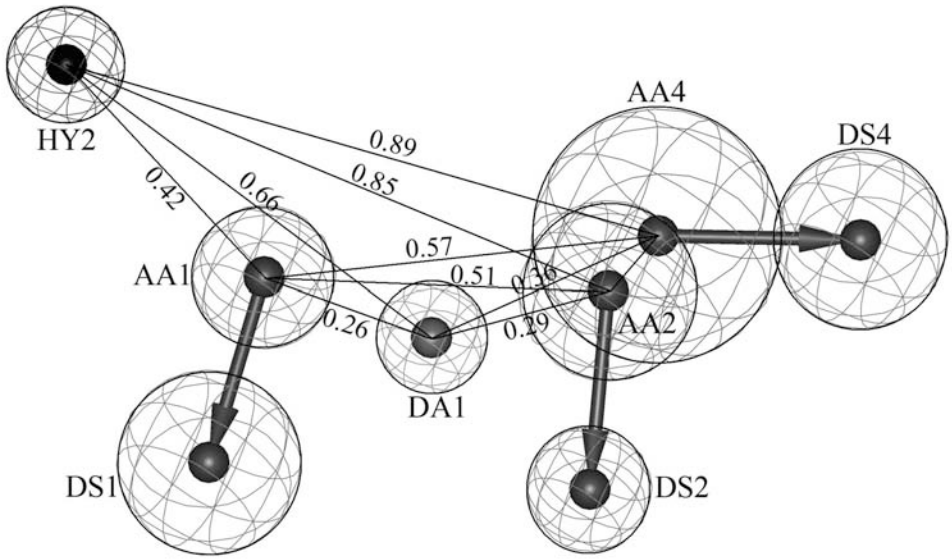

Fig. (7). The optimal pharmacophore model including one hydrophobic feature, three hydrogen pair features and one H-bond donor feature.

compounds (Fig. (**6**) compound **33, 34**) of these derivatives display an inhibition potency against strand-transfer with IC_{50} = 0.01 μM and 0.004 μM, respectively.

Based on two previously reported chalcones (Fig. (**6**) compound **35, 36**) [85], Deng *et al.* [86] developed five six-feature pharmacophore and two four-feature pharmacophore from **35** and **36** (Fig. (**6**)), respectively. And then the best model was applied to search a database, 44 selected compounds showed inhibitory potency at IC_{50} values < 100 μM. Compound **37** (Fig. (**6**)) is the most active molecule with IC_{50} values of 1.9 μM and 0.6 μM for 3-processing and strand transfer process, respectively.

Dayam *et al.* [87] generated 10 common feature pharmacophore hypotheses from a set of quinolone 3-carboxylic acid IN inhibitors, including clinical candidate GS-9137 and four analogues. The best pharmacophore model consisted of a negatively ionizable (NI) feature, an H-bond acceptor (HBA) feature, and two hydrophobic aromatic (HRA) features, in which the orientation of HBA and NI features is in agreement with proposed binding mechanism of the quinolone 3-carboxylate IN inhibitors. Using this model, 56 hits were selected from a database of 362,260 small molecules. *In vitro* assay showed that two of the 56 compounds (Fig. (**6**) compound **38**, **39**) inhibit the 3-processing and strand transfer with IC_{50} values of 14 μM and 5 μM. Substructure searching based on compound **39** retrieved two analogues (Fig. (**6**) compound **40**, **41**) with better anti-HIV IN activities. The IC_{50} values of compound **40** are 5±2 μM and 4±2 μM against 3-processing and strand transfer. And the correspondingly values of compound **41** are 7±6 μM and 4±2 μM. Unfortunately, both compounds are toxic in cytotoxicity test.

Our group [88] had developed a 3-D pharmacophore model from MK-0518 and S-1360, of which the IC_{50} value is less than 20 nmol/L. For the strong anti-HIV potency and reliable drug-like properties, we mapped inhibitor conformations into the pharmacophore model and superimposed it in their docking model with IN core domain. Thus, we got the corresponding positions between the pharmacophore model and IN residues. Pharmacophore model was modified according to our superimposition result. Finally, an optimal pharmacophore was generated including one hydrophobic feature, three hydrogen pair features and one H-bond donor feature (Fig. (**7**)).

5.2. Receptor-Based Pharmacophore Modeling

With the increasing number of available IN structure and efficient result of pharmacophore-based virtual screening, receptor-based pharmacophore generation method is applied more frequently to anti-HIV IN drug design [89-90]. Tintori *et al.* [91] applied a multistep computational protocol for the development of receptor-based pharmacophore which combines pharmacophore generation with conformational analysis, docking studies and MD simulation. Firstly, they performed a conformational search starting from the flexible loop region of HIV IN. And then those conformations were clustered into 10 clusters. Secondly, based on the best conformation found in the most populated cluster, three pharmacophores were generated containing H-bond acceptors, H-bond donors and hydrophobic features. Subsequently, three hypotheses were used alternately to screen the database with 200,000 commercially available compounds after docking calculations. Finally, one hit compound was identified and showed EC_{50} value of 30 μM, IC_{50} values of 25 μM and 3 μM toward 3-processing and strand transfer, respectively.

6. VIRTUAL SCREENING AND LIBRARY DESIGN

Virtual database screening (VS) has been proved to be an alternative and complementary approach to HTS in lead finding, with significant higher hit rates than HTS [92]. There are two broad categories of screening techniques:

receptor-based [93] and ligand-based [94] methods. Receptor-based approaches require information of 3D structure of the target and binding mode between receptor and ligand. Fast automated docking methods are used to filter large libraries of compounds in order to identify those that are most likely to bind to the target. Ligand-based approaches use pharmacophore or QSAR models for screening. Virtual screening is typically processed in the following steps: 2D similarity/dissimilarity filter, generating 3D conformations, 3D similarity/pharmacophore filter, docking and scoring. The ability to combine many filters allows researchers to reduce a huge virtual library to a manageable size.

The design of large combinatorial libraries has provided pharmacologists an advanced way of finding a needle in a haystack. Essentially, the idea involves taking a small number of starting compounds and reacting with a larger number of reagents, to create and test logically and systematically rational mixtures to obtain active lead compounds. Today, with the assistance of advanced algorithms for virtual screening and molecular design approaches, combinatorial library design can result in significantly higher hit rates [95]. Library design software are represented by HARPick program [96], MoSELECT [97], or TOPAS [98]. Most of these algorithms perform stochastic searching (Monte Carlo, Evolutionary Algorithm) to compile a library of candidates. The achievements have been reviewed comprehensively [99].

Tintori *et al.* [100] employed an innovative virtual screening approach consisting of the electron-ion interaction potential (EIIP) technique, drug like property calculation, pharmacophoric model generation, and docking studies. EIIP is a fundamental physical parameter of biological molecules that is determined by some formulas derived from the "general model pseudopotential". The VS approach was completed by the following steps. Firstly, the EIIP screening was applied to a database containing over 200,000 molecules, which leaded to 96,000 molecules to be used for further considerations. Secondly, the number of selected compounds was reduced to 40,000 according to the Lipinski's rule-of-five and a range of rotatable bonds below 10. Subsequently, a ligand-based pharmacophore was generated based on 30 compounds and was used as a query to retrieving molecules that fully fit for the model. As a result, 15,000 compounds were selected. Finally, docking simulations were performed to calculate the binding energy between selected compounds and IN active domain. According to the binding energy, structure diversity and commercially availability, 12 compounds were eventually chosen for *in vitro* assay and one compound (Fig. (**8**) compound **42**) of them displayed a reasonable activity with IC_{50} value of 69 μM. Substructure modification of **42** was employed subsequently, identifying an analogue (Fig. (**8**) compound **43**) with higher inhibitory potency with IC_{50} value of 10 μM.

Using the same virtual screening approach mentioned above, Mugnaini *et al.* [101] also found the hit compound **42** as IN binding inhibitors from databases containing 200,000 molecules, which has an overall IC_{50} value of 12 μM.

Liao *et al.* [102] initially performed a virtual screening based on a model of full-length HIV-1 IN combined with viral DNA, which they structured previously to find inhibitors for strand transfer reaction catalyzed by wild-type HIV-1 IN. 30 receptor-based and ligand-based pharmacophores were developed to search a database comprising 13.5 million available molecules; and then filtered by the Lipinski's rule-of-five. 167,479 compounds were identified as hits. Then, this model was used for docking. In experimental validation, 6 of 88 resulted compounds showed an IC_{50} values for strand transfer process ranging from 49 to 120 μM.

7. *DE NOVO* DRUG DESIGN

Different with other methods that modify molecule structures based on known ligands, *de novo* drug design methods create brand new ligands according to a well-defined binding site. The core principles of *de novo* drug design involve not only assembling possible compounds and

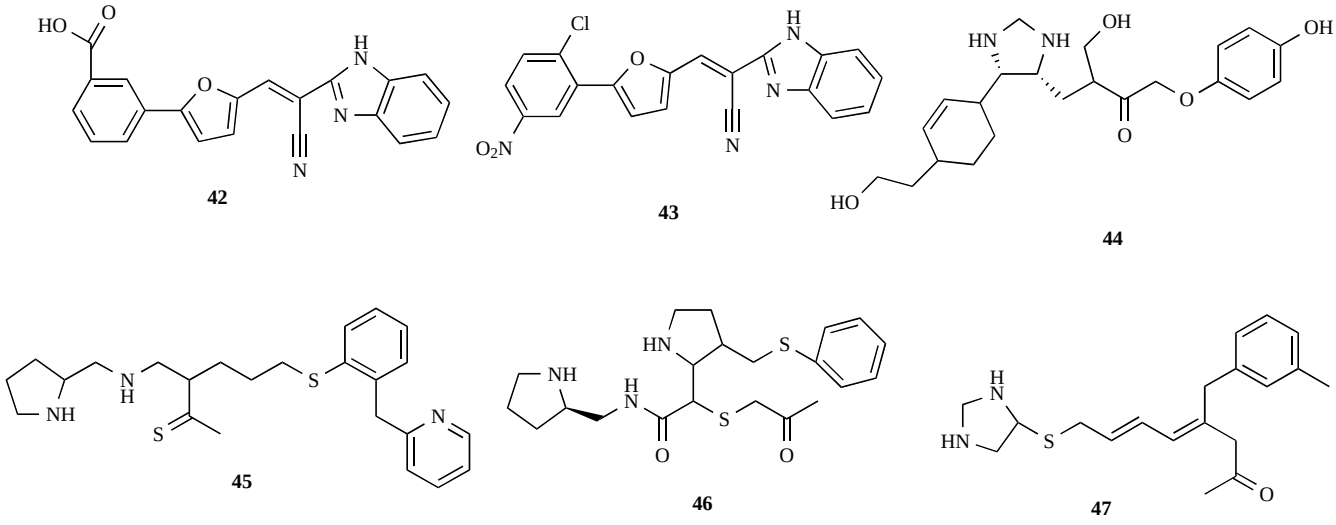

Fig. (8). The structure of compound **42-47**.

evaluating their quality but also searching the sample space for novel structures with drug-like properties. The different ways of building complete ligands include linking, growing, lattice-based sampling, random structure change, and MD-based methods. Scoring process runs simultaneously with the generation and evolution of the ligands, guiding the process to the most promising products. Well-known *de novo* design programs, such as LEGEND [103], LUDI [104], LeapFrog [105], SPROUT [106], HOOK [107], and PRO-LIGAND [108], use different scoring functions varying from simple steric constraints and H-bond placement to explicit force fields and empirical or knowledge-based scoring methods. Schneider *et al.* [109] described the various design concepts and delineated the developments of computer-based *de novo* design. In another review, Schneider *et al.* [110] focused on adaptive *de novo* drug design techniques that can be used to artificially evolve novel bioactive molecules. Evolutionary algorithms and particle swarm optimization are common methods for iterative virtual synthesis and test. Evolving compound libraries are suited for hit and lead discovery in situations where resources are limited and the complete testing of large-scale compound collection is not feasible.

However, negletion of the synthetic feasibility has been the major defect and hindrance of *de novo* design. Inaccurate evaluation of binding affinities for the designed structures is another issue that leads to low potency.

Huang *et al.* [111] performed a pharmacophore-based *de novo* design method of HIV-1 IN inhibitors. This approach started from a known pharmacophore model taken from previous article [112], which was used to choose proper fragments and install them into integrated molecule. They obtained 100 molecules as the output. Subsequent, clustering, drug-likeness and synthetic accessibility were assessed to filter the compounds and pick out the hit compounds. Finally, four compounds were identified as potential IN inhibitors (Fig. (**8**) compound **44**, **45**, **46** and **47**).

8. SUMMARY AND PERSPECTIVES

Rational drug design not only reduces the cost and time of drug development, but also, more importantly, provides crucial information, such as the binding mode, key residues, and other data that the experimental studies could not obtain. In recent years, with the assistance of rational drug design technologies, virtual screening has achieved encouraging progress on discovery of novel class of anti-HIV-1 IN inhibitors. However, there are several deficiencies in the current studies:

(1) Although there has been great breakthrough in studying on structure and mechanism of IN, as mentioned as PFV intasome, the crystal structure of full-length HIV-1 IN and the exact physiologic mechanism of IN are still unknown, which makes it difficult to design rational molecules selectively inhibiting IN.

(2) Docking and simulation technologies need to be improved to take more practical issues into account, such as the best conformation, receptor flexibility, solution effects and etc.

(3) Although numerous compounds are included in available chemical databases, few of them are marketable.

This review gives a concise summary of the recent 5-year scientific achievements in discovering anti-HIV IN inhibitors by using CADD methods. The latest technological advances, the increasing in numbers of chemical and biological databases, and the development in software are together starting a new chapter of anti-HIV IN inhibitors discovery. We are thus convinced that adaptive, multi-objective drug design approaches will increasingly gain impact in lead compounds discovery and optimization as a primary source of new anti-HIV drugs.

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGEMENTS

The authors thank Chinese Natural Science Foundation project (No.30670497 and 21173014), National Key Basic Research Program of China (2009CB930200), Beijing Natural Science Foundation (No. 7082006). We are grateful to Hao Zhu (UNC Eshelman School of Pharmacy, University of North Carolina at Chapel Hill) for excellent word assistance.

REFERENCES

[1] Tan, J. J.; Cong, X. J.; Hu, L. M.; Wang, C. X.; Jia, L.; Liang, X. J. Therapeutic strategies underpinning the development of novel techniques for the treatment of HIV infection. *Drug Discov Today*, **2010**, *15* (5-6), 186-97.

[2] Luo, Z. G.; Tan, J. J.; Zeng, Y.; Wang, C. X.; Hu, L. M. Development of integrase inhibitors of quinolone acid derivatives for treatment of AIDS: an overview. *Mini Rev Med Chem*, **2010**, *10* (11), 1046-57.

[3] Nair, V.; Chi, G. HIV integrase inhibitors as therapeutic agents in AIDS. *Rev Med Virol*, **2007**, *17* (4), 277-295.

[4] Lu, R.; Ghory, H. Z.; Engelman, A. Genetic analyses of conserved residues in the carboxyl-terminal domain of human immunodeficiency virus type 1 integrase. *J Virol*, **2005**, *79* (16), 10356-10368.

[5] Markowitz, M.; Morales-Ramirez, J. O.; Nguyen, B. Y.; Kovacs, C. M.; Steigbigel, R. T.; Cooper, D. A.; Liporace, R.; Schwartz, R.; Isaacs, R.; Gilde, L. R.; Penning, L.; Zhao, J.; Teppler, H.; Team, P. S. Antiretroviral activity, pharmacokinetics, and tolerability of MK-0518, a novel inhibitor of HIV-1 integrase, dosed as monotherapy for 10 days in treatment-naive HIV-1-infected individuals. *JAIDS-J Acq Imm Def*, **2006**, *43* (5), 509-515.

[6] Kassahun, K.; McIntosh, I.; Cui, D.; Hreniuk, D.; Merschman, S.; Lasseter, K.; Azrolan, N.; Iwamoto, M.; Wagner, J. A.; Wenning, L. A. Metabolism and disposition in humans of raltegravir (MK-0518), an Anti-AIDS drug targeting the human immunodeficiency virus 1 integrase enzyme. *Drug Metab Dispos*, **2007**, *35* (9), 1657-1663.

[7] Cocohoba, J.; Dong, B. J. Raltegravir: The First HIV Integrase Inhibitor. *Clin Ther*, **2008**, *30* (10), 1747-1765.

[8] DiMasi, J. A.; Hansen, R. W.; Grabowski, H. G. The price of innovation: new estimates of drug development costs. *J Health Econ*, **2003**, *22* (2), 151-85.

[9] Andricopulo, A. D.; Salum, L. B.; Abraham, D. J. Structure-based drug design strategies in medicinal chemistry. *Curr Top Med Chem*, **2009**, *9* (9), 771-90.

[10] Siddiqi, M. I.; Kumar, A. Review of knowledge for rational design and identification of anti-tubercular compounds. *Expert Opin Drug Dis,* **2009**, *4* (10), 1005-1015.

[11] Jain, A. N. Scoring functions for protein-ligand docking. *Curr Protein Pept Sci,* **2006**, *7* (5), 407-20.

[12] Lensink, M. F.; Mendez, R.; Wodak, S. J. Docking and scoring protein complexes: CAPRI 3rd Edition. *Proteins,* **2007**, *69* (4), 704-18.

[13] Sudol, M.; Tran, M.; Nowak, M. G.; Flanagan, J. M.; Robertson, G. P.; Katzman, M. A nonradioactive plate-based assay for stimulators of nonspecific DNA nicking by HIV-1 integrase and other nucleases. *Anal Biochem,* **2010**, *396* (2), 223-230.

[14] Rajamani, R.; Good, A. C. Ranking poses in structure-based lead discovery and optimization: current trends in scoring function development. *Curr Opin Drug Discov Devel,* **2007**, *10* (3), 308-15.

[15] Jones, G.; Willett, P.; Glen, R. C. Molecular Recognition of Receptor-Sites Using a Genetic Algorithm with a Description of Desolvation. *J Mol Biol,* **1995**, *245* (1), 43-53.

[16] Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem,* **1998**, *19* (14), 1639-1662.

[17] Ewing, T. J. A.; Kuntz, I. D. Critical evaluation of search algorithms for automated molecular docking and database screening. *J Comput Chem,* **1997**, *18* (9), 1175-1189.

[18] Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J Mol Biol,* **1996**, *261* (3), 470-489.

[19] Jain, A. N. Surflex: Fully automatic flexible molecular docking using a molecular similarity-based search engine. *J Med Chem,* **2003**, *46* (4), 499-511.

[20] Stewart, K. D.; Shiroda, M.; James, C. A. Drug Guru: A computer software program for drug design using medicinal chemistry rules. *Bioorgan Med Chem,* **2006**, *14* (20), 7011-7022.

[21] Alonso, H.; Bliznyuk, A. A.; Gready, J. E. Combining docking and molecular dynamic simulations in drug design. *Med Res Rev,* **2006**, *26* (5), 531-68.

[22] Hughes, S.; Jenkins, V.; Dar, M. J.; Engelman, A.; Cherepanov, P. Transcriptional Co-activator LEDGF Interacts with Cdc7-Activator of S-phase Kinase (ASK) and Stimulates Its Enzymatic Activity. *J Biol Chem,* **2010**, *285* (1), 541-554.

[23] Hickman, A. B.; Chandler, M.; Dyda, F. Integrating prokaryotes and eukaryotes: DNA transposases in light of structure. *Crit Rev Biochem Mol,* **2010**, *45* (1), 50-69.

[24] Chen, J. C. H.; Krucinski, J.; Miercke, L. J. W.; Finer-Moore, J. S.; Tang, A. H.; Leavitt, A. D.; Stroud, R. M. Crystal structure of the HIV-1 integrase catalytic core and C-terminal domains: A model for viral DNA binding. *P Natl Acad Sci USA,* **2000**, *97* (15), 8233-8238.

[25] Hare, S.; Gupta, S. S.; Valkov, E.; Engelman, A.; Cherepanov, P. Retroviral intasome assembly and inhibition of DNA strand transfer. *Nature,* **2010**, *464* (7286), 232-6.

[26] Maertens, G. N.; Hare, S.; Cherepanov, P. The mechanism of retroviral integration from X-ray structures of its key intermediates. *Nature,* **2010**, *468* (7321), 326-U217.

[27] Chen, X.; Tsiang, M.; Yu, F.; Hung, M.; Jones, G. S.; Zeynalzadegan, A.; Qi, X.; Jin, H.; Kim, C. U.; Swaminathan, S.; Chen, J. M. Modeling, analysis, and validation of a novel HIV integrase structure provide insights into the binding modes of potent integrase inhibitors. *J Mol Biol,* **2008**, *380* (3), 504-519.

[28] Wang, L. D.; Liu, C. L.; Chen, W. Z.; Wang, C. X. Constructing HIV-1 integrase tetramer and exploring influences of metal ions on forming integrase-DNA complex. *Biochem Bioph Res Co,* **2005**, *337* (1), 313-319.

[29] Wang, L. D.; Wang, C. X. Study on the influence of metal ions on interaction between HIV-1 integrase and thiazolothiazepine inhibitors with molecular simulation methods. *Acta Chim Sinica,* **2008**, *66* (7), 817-822.

[30] Hu, J. P.; Ke, G. T.; Chang, S.; Chen, W. Z.; Wang, C. X. Conformational Change of HIV-1 Viral DNA after Binding with Integrase. *Acta Phys-Chim Sin,* **2008**, *24* (10), 1803-1810.

[31] Hu, J. P.; Wang, C. X. Molecular Dynamics Simulation of HIV-1 Integrase Dimer Complexed with Viral DNA. *Chinese J Chem,* **2010**, *28* (1), 33-40.

[32] Hu, J. P.; Gong, X. Q.; Su, J. G.; Chen, W. Z.; Wang, C. X. Study on the molecular mechanism of inhibiting HIV-1 integrase by EBR28 peptide *via* molecular modeling approach. *Biophys Chem,* **2008**, *132* (2-3), 69-80.

[33] Liu, M.; Cong, X. J.; Li, P.; Tan, J. J.; Chen, W. Z.; Wang, C. X. Study on the Inhibitory Mechanism and Binding Mode of the Hydroxycoumarin Compound NSC158393 to HIV-1 Integrase by Molecular Modeling. *Biopolymers,* **2009**, *91* (9), 700-709.

[34] Amaro, R. E.; Baron, R.; McCammon, J. A. An improved relaxed complex scheme for receptor flexibility in computer-aided drug design. *J Comput Aid Mol Des,* **2008**, *22* (9), 693-705.

[35] Perryman, A. L.; Forli, S.; Morris, G. M.; Burt, C.; Cheng, Y. H.; Palmer, M. J.; Whitby, K.; McCammon, J. A.; Phillips, C.; Olson, A. J. A Dynamic Model of HIV Integrase Inhibition and Drug Resistance. *J Mol Biol,* **2010**, *397* (2), 600-615.

[36] Metifiot, M.; Marchand, C.; Maddali, K.; Pommier, Y. Resistance to integrase inhibitors. *Viruses,* **2010**, *2* (7), 1347-66.

[37] Le, G. A.; Vandegraaff, N.; Rhodes, D. I.; Jones, E. D.; Coates, J. A. V.; Lu, L.; Li, X. M.; Yu, C. J.; Feng, X.; Deadman, J. J. Discovery of potent HIV integrase inhibitors active against raltegravir resistant viruses. *Bioorg Med Chem Lett,* **2010**, *20* (17), 5013-5018.

[38] Nunthaboot, N.; Pianwanit, S.; Parasuk, V.; Ebalunode, J. O.; Briggs, J. M.; Kokpol, S. Hybrid quantum mechanical/molecular mechanical molecular dynamics simulations of HIV-1 integrase/inhibitor complexes. *Biophys J,* **2007**, *93* (10), 3613-3626.

[39] Goldgur, Y.; Craigie, R.; Cohen, G. H.; Fujiwara, T.; Yoshinaga, T.; Fujishita, T.; Sugimoto, H.; Endo, T.; Murai, H.; Davies, D. R. Structure of the HIV-1 integrase catalytic domain complexed with an inhibitor: A platform for antiviral drug design. *P Natl Acad Sci USA,* **1999**, *96* (23), 13040-13043.

[40] Alves, C. N.; Marti, S.; Castillo, R.; Andres, J.; Moliner, V.; Tunon, I.; Silla, E. A quantum mechanics/molecular mechanics study of the protein-ligand interaction for inhibitors of HIV-1 integrase. *Chem-Eur J,* **2007**, *13* (27), 7715-7724.

[41] Brigo, A.; Lee, K. W.; Fogolari, F.; Mustata, G. L.; Briggs, J. M. Comparative molecular dynamics simulations of HIV-1 integrase and the T66I/M154I mutant: Binding modes and drug resistance to a diketo acid inhibitor. *Proteins,* **2005**, *59* (4), 723-741.

[42] Ma, X. H.; Jia, J.; Zhu, F.; Xue, Y.; Li, Z. R.; Chen, Y. Z. Comparative analysis of machine learning methods in ligand-based virtual screening of large compound libraries. *Comb Chem High Throughput Screen,* **2009**, *12* (4), 344-57.

[43] Douguet, D. Ligand-based approaches in virtual screening. *Curr Comput-Aid Drug,* **2008**, *4* (3), 180-190.

[44] Metobo, S. E.; Jin, H. L.; Tsiang, M.; Kim, C. U. Design, synthesis and biological evaluation of novel tricyclic HIV-1 integrase inhibitors by modification of its pyridine ring. *Bioorg Med Chem Lett,* **2006**, *16* (15), 3985-3988.

[45] Jin, H. L.; Cai, R. Z.; Schacherer, L.; Jabri, S.; Tsiang, M.; Fardis, M.; Chen, X. W.; Chen, J. M.; Kim, C. U. Design, synthesis, and SAR studies of novel and highly active tri-cyclic HIV integrase inhibitors. *Bioorg Med Chem Lett,* **2006**, *16* (15), 3989-3992.

[46] Jin, H. L.; Metobo, S.; Jabri, S.; Mish, M.; Lansdown, R.; Chen, X. W.; Tsiang, M.; Wright, M.; Kim, C. U. Tricyclic HIV integrase inhibitors V. SAR studies on the benzyl moiety. *Bioorg Med Chem Lett,* **2009**, *19* (8), 2263-2265.

[47] Metobo, S.; Mish, M.; Jin, H. L.; Jabri, S.; Lansdown, R.; Chen, X. W.; Tsiang, M.; Wright, M.; Kim, C. U. Tricyclic HIV integrase inhibitors: VI. SAR studies of 'benzyl flipped' C3-substituted pyrroloquinolines. *Bioorg Med Chem Lett,* **2009**, *19* (4), 1187-1190.

[48] Petrocchi, A.; Jones, P.; Rowley, M.; Fiore, F.; Summa, V. N-(4-Fluorobenzyl)-3-hydroxy-9,9-dimethyl-4-oxo-6,7,8,9-tetrahydro-4H-pyrazino[1,2-a]pyrimidine-2-carboxamides a novel class of potent HIV-1 integrase inhibitors. *Bioorg Med Chem Lett,* **2009**, *19* (15), 4245-4249.

[49] Donghi, M.; Kinzel, O. D.; Summa, V. 3-Hydroxy-4-oxo-4H-pyrido[1,2-a]pyrimidine-2-carboxylates-A new class of HIV-1 integrase inhibitors. *Bioorg Med Chem Lett,* **2009**, *19* (7), 1930-1934.

[50] Jones, E. D.; Vandegraaff, N.; Le, G.; Choi, N.; Issa, W.; Macfarlane, K.; Thienthong, N.; Winfield, L. J.; Coates, J. A.; Lu,

L.; Li, X.; Feng, X.; Yu, C.; Rhodes, D. I.; Deadman, J. J. Design of a series of bicyclic HIV-1 integrase inhibitors. Part 1: selection of the scaffold. *Bioorg Med Chem Lett*, **2010**, *20* (19), 5913-7.

[51] Le, G.; Vandegraaff, N.; Rhodes, D. I.; Jones, E. D.; Coates, J. A.; Thienthong, N.; Winfield, L. J.; Lu, L.; Li, X.; Yu, C.; Feng, X.; Deadman, J. J. Design of a series of bicyclic HIV-1 integrase inhibitors. Part 2: azoles: effective metal chelators. *Bioorg Med Chem Lett*, **2010**, *20* (19), 5909-12.

[52] Walker, M. A.; Johnson, T.; Naidu, B. N.; Banville, J.; Remillard, R.; Plamondon, S.; Martel, A.; Li, C.; Torri, A.; Samanta, H.; Lin, Z.; Dicker, I.; Krystal, M.; Meanwell, N. A. Benzyl amide-ketoacid inhibitors of HIV-integrase. *Bioorg Med Chem Lett*, **2007**, *17* (17), 4886-4890.

[53] Pace, P.; Spieser, S. A. H.; Summa, V. 4-Hydroxy-5-pyrrolinone-3-carboxamide HIV-1 integrase inhibitors. *Bioorg Med Chem Lett*, **2008**, *18* (14), 3865-3869.

[54] Li, H. C.; Wang, C.; Sanchez, T.; Tan, Y. M.; Jiang, C. Y.; Neamati, N.; Zhao, G. S. Amide-containing diketoacids as HIV-1 integrase inhibitors: Synthesis, structure-activity relationship analysis, and biological activity. *Bioorgan Med Chem*, **2009**, *17* (7), 2913-2919.

[55] Walker, M. A.; Johnson, T.; Ma, Z. P.; Zhang, Y. H.; Banville, J.; Remillard, R.; Plamondon, S.; Pendri, A.; Torri, A.; Wong, H.; Smith, D.; Samanta, H.; Lin, Z. Y.; Deminie, C.; Terry, B.; Krystal, M.; Meanwell, N. Exploration of the diketoacid integrase inhibitor chemotype leading to the discovery of the anilide-ketoacids chemotype. *Bioorg Med Chem Lett*, **2006**, *16* (22), 5818-5821.

[56] Xu, Y. S.; Zeng, C. C.; Jiao, Z. G.; Hu, L. M.; Zhong, R. G. Design, Synthesis and Anti-HIV Integrase Evaluation of 4-Oxo-4H-quinolizine-3-carboxylic Acid Derivatives. *Molecules*, **2009**, *14* (2), 868-883.

[57] Di Santo, R.; Costi, R.; Roux, A.; Artico, M.; Lavecchia, A.; Marinelli, L.; Novellino, E.; Palmisano, L.; Andreotti, M.; Amici, R.; Galluzzo, C. M.; Nencioni, L.; Palamara, A. T.; Pommier, Y.; Marchand, C. Novel bifunctional quinolonyl diketo acid derivatives as HIV-1 integrase inhibitors: Design, synthesis, biological activities, and mechanism of action. *J Med Chem*, **2006**, *49* (6), 1939-1945.

[58] Zeng, L. F.; Zhang, H. S.; Wang, Y. H.; Sanchez, T.; Zheng, Y. T.; Neamati, N.; Long, Y. Q. Efficient synthesis and utilization of phenyl-substituted heteroaromatic carboxylic acids as aryl diketo acid isosteres in the design of novel HIV-1 integrase inhibitors. *Bioorg Med Chem Lett*, **2008**, *18* (16), 4521-4524.

[59] Johns, B. A.; Weatherhead, J. G.; Allen, S. H.; Thompson, J. B.; Garvey, E. P.; Foster, S. A.; Jeffrey, J. L.; Miller, W. H. The use of oxadiazole and triazole substituted naphthyridines as HIV-1 integrase inhibitors. Part 1: Establishing the pharmacophore. *Bioorg Med Chem Lett*, **2009**, *19* (6), 1802-1806.

[60] Johns, B. A.; Weatherhead, J. G.; Allen, S. H.; Thompson, J. B.; Garvey, E. P.; Foster, S. A.; Jeffrey, J. L.; Miller, W. H. 1,3,4-Oxadiazole substituted naphthyridines as HIV-1 integrase inhibitors. Part 2: SAR of the C5 position. *Bioorg Med Chem Lett*, **2009**, *19* (6), 1807-1810.

[61] Hadi, V.; Koh, Y. H.; Sanchez, T. W.; Barrios, D.; Neamati, N.; Jung, K. W. Development of the next generation of HIV-1 integrase inhibitors: Pyrazolone as a novel inhibitor scaffold. *Bioorg Med Chem Lett*, **2010**, *20* (22), 6854-6857.

[62] Sechi, M.; Rizzi, G.; Bacchi, A.; Carcelli, M.; Rogolino, D.; Pala, N.; Sanchez, T. W.; Taheri, L.; Dayam, R.; Neamati, N. Design and synthesis of novel dihydroquinoline-3-carboxylic acids as HIV-1 integrase inhibitors. *Bioorgan Med Chem*, **2009**, *17* (7), 2925-2935.

[63] Christ, F.; Voet, A.; Marchand, A.; Nicolet, S.; Desimmie, B. A.; Marchand, D.; Bardiot, D.; Van der Veken, N. J.; Van Remoortel, B.; Strelkov, S. V.; De Maeyer, M.; Chaltin, P.; Debyser, Z. Rational design of small-molecule inhibitors of the LEDGF/p75-integrase interaction and HIV replication. *Nat Chem Biol*, **2010**, *6* (6), 442-448.

[64] Crippen, G. M. Distance Geometry Approach to Rationalizing Binding Data. *J Med Chem*, **1979**, *22* (8), 988-997.

[65] Iyer, M.; Hopfinger, A. J. Treating chemical diversity in QSAR analysis: Modeling diverse HIV-1 integrase inhibitors using 4D fingerprints. *J Chem Inf Model*, **2007**, *47* (5), 1945-1960.

[66] Cramer, R. D.; Patterson, D. E.; Bunce, J. D. Comparative Molecular-Field Analysis (Comfa) .1. Effect of Shape on Binding of Steroids to Carrier Proteins. *J Am Chem Soc*, **1988**, *110* (18), 5959-5967.

[67] Nightingale, S. J.; Hollis, R. P.; Pepper, K. A.; Petersen, D.; Yu, X. J.; Yang, C.; Bahner, I.; Kohn, D. B. Transient gene expression by nonintegrating lentiviral vectors. *Mol Ther*, **2006**, *13* (6), 1121-1132.

[68] Saeh, J. C.; Lyne, P. D.; Takasaki, B. K.; Cosgrove, D. A. Lead hopping using SVM and 3D pharmacophore fingerprints. *J Chem Inf Model*, **2005**, *45* (4), 1122-1133.

[69] Basak, S. C.; Natarajan, R.; Mills, D.; Hawkins, D. M.; Kraker, J. J. Quantitative structure-activity relationship modeling of juvenile hormone mimetic compounds for Culex pipiens larvae, with a discussion of descriptor-thinning methods. *J Chem Inf Model*, **2006**, *46* (1), 65-77.

[70] Singh, J.; Chuaqui, C. E.; Boriack-Sjodin, P. A.; Lee, W. C.; Pontz, T.; Corbley, M. J.; Cheung, H. K.; Arduini, R. M.; Mead, J. N.; Newman, M. N.; Papadatos, J. L.; Bowes, S.; Josiah, S.; Ling, L. E. Successful shape-based virtual screening: The discovery of a potent inhibitor of the type I TGF beta receptor kinase (T beta RI). *Bioorg Med Chem Lett*, **2003**, *13* (24), 4355-4359.

[71] Komano, J.; Futahashi, Y.; Urano, E.; Miyauchi, K.; Murakami, T.; Matsuda, Z.; Yamamoto, N. The interaction of HIV-1 with the host factors. *Jpn J Infect Dis*, **2005**, *58* (3), 125-130.

[72] Bhattacharya, S.; Osman, H. Novel targets for anti-retroviral therapy. *J Infection*, **2009**, *59* (6), 377-386.

[73] Gupta, P.; Roy, N.; Garg, P. Docking-based 3D-QSAR study of HIV-1 integrase inhibitors. *Eur J Med Chem*, **2009**, *44* (11), 4276-87.

[74] Leonard, J. T.; Roy, K. Exploring molecular shape analysis of styrylquinoline derivatives as HIV-1 integrase inhibitors. *Eur J Med Chem*, **2008**, *43* (1), 81-92.

[75] de Melo, E. B.; Ferreira, M. M. C. Multivariate QSAR study of 4,5-dihydroxypyrimidine carboxamides as HIV-1 integrase inhibitors. *Eur J Med Chem*, **2009**, *44* (9), 3577-3583.

[76] Dhaked, D. K.; Verma, J.; Saran, A.; Coutinho, E. C. Exploring the binding of HIV-1 integrase inhibitors by comparative residue interaction analysis (CoRIA). *J Mol Model*, **2009**, *15* (3), 233-245.

[77] Liu, Q.; Che, D. S.; Huang, Q.; Cao, Z. W.; Zhu, R. X. Multi-target QSAR Study in the Analysis and Design of HIV-1 Inhibitors. *Chinese J Chem*, **2010**, *28* (9), 1587-1592.

[78] Cheng, Z.; Zhang, Y.; Fu, W. QSAR study of carboxylic acid derivatives as HIV-1 Integrase inhibitors. *Eur J Med Chem*, **2010**, *45* (9), 3970-80.

[79] Ferro, S.; De Luca, L.; Barreca, M. L.; Iraci, N.; De Grazia, S.; Christ, F.; Witvrouw, M.; Debyser, Z.; Chimirri, A. Docking Studies on a New Human Immodeficiency Virus Integrase-Mg-DNA Complex: Phenyl Ring Exploration and Synthesis of 1H-Benzylindole Derivatives through Fluorine Substitutions. *J Med Chem*, **2009**, *52* (2), 569-573.

[80] Tintori, C.; Manetti, F.; Botta, M. Pharmacophoric models and 3D QSAR studies of the adenosine receptor ligands. *Curr Top Med Chem*, **2010**, *10* (10), 1019-35.

[81] Leach, A. R.; Gillet, V. J.; Lewis, R. A.; Taylor, R. Three-dimensional pharmacophore methods in drug discovery. *J Med Chem*, **2010**, *53* (2), 539-58.

[82] Patel, Y.; Gillet, V. J.; Bravi, G.; Leach, A. R. A comparison of the pharmacophore identification programs: Catalyst, DISCO and GASP. *J Comput Aid Mol Des*, **2002**, *16* (8-9), 653-681.

[83] Barreca, M. L.; Ferro, S.; Rao, A.; De Luca, L.; Zappala, M.; Monforte, A. M.; Debyser, Z.; Witvrouw, M.; Chimirri, A. Pharmacophore-based design of HIV-1 integrase strand-transfer inhibitors. *J Med Chem*, **2005**, *48* (22), 7084-7088.

[84] Barreca, M. L.; Rao, A.; De Luca, L.; Zappala, M.; Gurnari, C.; Monforte, P.; De Clercq, E.; Van Maele, B.; Debyser, Z.; Witvrouw, M.; Briggs, J. M.; Chimirri, A. Efficient 3D database screening for novel HIV-1IN inhibitors. *J Chem Inf Comp Sci*, **2004**, *44* (4), 1450-1455.

[85] Deng, J. X.; Kelley, J. A.; Barchi, J. J.; Sanchez, T.; Dayam, R.; Pommier, Y.; Neamati, N. Mining the NCI antiviral compounds for HIV-1 integrase inhibitors. *Bioorgan Med Chem*, **2006**, *14* (11), 3785-3792.

[86] Deng, J. X.; Sanchez, T.; Al-Mawsawi, L. Q.; Dayam, R.; Yunes, R. A.; Garofalo, A.; Bolger, M. B.; Neamati, N. Discovery of

structurally diverse HIV-1 integrase inhibitors based on a chalcone pharmacophore. *Bioorgan Med Chem*, **2007**, *15* (14), 4985-5002.

[87] Dayam, R.; Al-Mawsawi, L. Q.; Zawahir, Z.; Witvrouw, M.; Debyser, Z.; Neamati, N. Quinolone 3-carboxylic acid pharmacophore: Design of second generation HIV-1 integrase inhibitors. *J Med Chem*, **2008**, *51* (5), 1136-1144.

[88] Zhang, X. Y.; Liu, B.; He, H. Q.; Yang, D.; Wang, C. X. Human Immunodeficiency Virus Integrase Pharmacophore Model Derived from Diketoacids Inhibitors. *Acta Phys-Chim Sin*, **2009**, *25* (5), 817-824.

[89] Deng, J. X.; Sanchez, T.; Neamati, N.; Briggs, J. M. Dynamic pharmacophore model optimization: Identification of novel HIV-1 integrase inhibitors. *J Med Chem*, **2006**, *49* (5), 1684-1692.

[90] Deng, J. X.; Lee, K. W.; Sanchez, T.; Cui, M.; Neamati, N.; Briggs, J. M. Dynamic receptor-based pharmacophore model development and its application in designing novel HIT-1 integrase inhibitors. *J Med Chem*, **2005**, *48* (5), 1496-1505.

[91] Tintori, C.; Corradi, V.; Magnani, M.; Manetti, F.; Botta, M. Targets Looking for Drugs: A Multistep Computational Protocol for the Development of Structure-Based Pharmacophores and Their Applications for Hit Discovery. *J Chem Inf Model*, **2008**, *48* (11), 2166-2179.

[92] Klebe, G. Virtual ligand screening: strategies, perspectives and limitations. *Drug Discov Today*, **2006**, *11* (13-14), 580-594.

[93] Tuccinardi, T. Docking-Based Virtual Screening: Recent Developments. *Comb Chem High T Scr*, **2009**, *12* (3), 303-314.

[94] Sun, H. M. Pharmacophore-based virtual screening. *Curr Med Chem*, **2008**, *15* (10), 1018-1024.

[95] Villar, H. O.; Hansen, M. R. Design of chemical libraries for screening. *Expert Opin Drug Dis*, **2009**, *4* (12), 1215-1220.

[96] Good, A. C.; Lewis, R. A. New methodology for profiling combinatorial libraries and screening sets: Cleaning up the design process with HARPick. *J Med Chem*, **1997**, *40* (24), 3926-3936.

[97] Gillet, V. J.; Khatib, W.; Willett, P.; Fleming, P. J.; Green, D. V. S. Combinatorial library design using a multiobjective genetic algorithm. *J Chem Inf Comp Sci*, **2002**, *42* (2), 375-385.

[98] Schneider, G.; Lee, M. L.; Stahl, M.; Schneider, P. De novo design of molecular architectures by evolutionary assembly of drug-derived building blocks. *J Comput Aid Mol Des*, **2000**, *14* (5), 487-494.

[99] Amaro, R. E.; Li, W. W. Emerging Methods for Ensemble-Based Virtual Screening. *Curr Top Med Chem*, **2010**, *10* (1), 3-13.

[100] Tintori, C.; Manetti, F.; Veljkovic, N.; Perovic, V.; Vercammen, J.; Hayes, S.; Massa, S.; Witvrouw, M.; Debyser, Z.; Veljkovic, V.; Botta, M. Novel virtual screening protocol based on the combined use of molecular modeling and electron-ion interaction potential techniques to design HIV-1 integrase inhibitors. *J Chem Inf Model*, **2007**, *47* (4), 1536-1544.

[101] Mugnaini, C.; Rajamaki, S.; Tintori, C.; Corelli, F.; Massa, S.; Witvrouw, M.; Debyser, Z.; Veljkovic, V.; Botta, M. Toward novel HIV-1 integrase binding inhibitors: Molecular modeling, synthesis, and biological studies. *Bioorg Med Chem Lett*, **2007**, *17* (19), 5370-5373.

[102] Liao, C. Z.; Karki, R. G.; Marchand, C.; Pommier, Y.; Nicklaus, M. C. Virtual screening application of a model of full-length HIV-1 integrase complexed with viral DNA. *Bioorg Med Chem Lett*, **2007**, *17* (19), 5361-5365.

[103] Nishibata, Y.; Itai, A. Automatic Creation of Drug Candidate Structures Based on Receptor Structure - Starting Point for Artificial Lead Generation. *Tetrahedron*, **1991**, *47* (43), 8985-8990.

[104] Bohm, H. J. The Computer-Program Ludi - a New Method for the Denovo Design of Enzyme-Inhibitors. *J Comput Aid Mol Des*, **1992**, *6* (1), 61-78.

[105] Payne, A. W. R.; Glen, R. C. Molecular Recognition Using a Binary Genetic Search Algorithm. *J Mol Graphics*, **1993**, *11* (2), 74.

[106] Gillet, V. J.; Myatt, G.; Zsoldos, Z.; Johnson, A. P. SPROUT, HIPPO and CAESA: Tools for de novo structure generation and estimation of synthetic accessibility. *Perspect Drug Discov*, **1995**, *3*, 34-50.

[107] Eisen, M. B.; Wiley, D. C.; Karplus, M.; Hubbard, R. E. Hook - a Program for Finding Novel Molecular Architectures That Satisfy the Chemical and Steric Requirements of a Macromolecule Binding-Site. *Proteins-Structure Function and Genetics*, **1994**, *19* (3), 199-221.

[108] Murray, C. W.; Clark, D. E.; Byrne, D. G. Pro-Ligand - an Approach to De-Novo Molecular Design .6. Flexible Fitting in the Design of Peptides. *J Comput Aid Mol Des*, **1995**, *9* (5), 381-395.

[109] Schneider, G.; Fechner, U. Computer-based de novo design of drug-like molecules. *Nat Rev Drug Discov*, **2005**, *4* (8), 649-663.

[110] Schneider, G.; Hartenfeller, M.; Reutlinger, M.; Tanrikulu, Y.; Proschak, E.; Schneider, P. Voyages to the (un)known: adaptive design of bioactive compounds. *Trends Biotechnol*, **2009**, *27* (1), 18-26.

[111] Huang, Q.; Li, L. L.; Yang, S. Y. PhDD: A new pharmacophore-based de novo design method of drug-like molecules combined with assessment of synthetic accessibility. *J Mol Graph Model*, **2010**, *28* (8), 775-787.

[112] Dayam, R.; Sanchez, T.; Clement, O.; Shoemaker, R.; Sei, S.; Neamati, N. beta-Diketo acid pharmacophore hypotesis. 1. Discovery of a novel class of HIV-1 integrase inhibitors. *J Med Chem*, **2005**, *48* (1), 111-120.

Received: November 05, 2011 Revised: March 30, 2012 Accepted: April 02, 2012