Computational drug design strategies applied to the modelling of human immunodeficiency virus-1 reverse transcriptase inhibitors

Lucianna Helene Santos[1], Rafaela Salgado Ferreira[2], Ernesto Raúl Caffarena[1]/+

[1]Fundação Oswaldo Cruz, Programa de Computação Científica, Rio de Janeiro, RJ, Brasil
[2]Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas, Belo Horizonte, MG, Brasil

Reverse transcriptase (RT) is a multifunctional enzyme in the human immunodeficiency virus (HIV)-1 life cycle and represents a primary target for drug discovery efforts against HIV-1 infection. Two classes of RT inhibitors, the nucleoside RT inhibitors (NRTIs) and the nonnucleoside transcriptase inhibitors are prominently used in the highly active antiretroviral therapy in combination with other anti-HIV drugs. However, the rapid emergence of drug-resistant viral strains has limited the successful rate of the anti-HIV agents. Computational methods are a significant part of the drug design process and indispensable to study drug resistance. In this review, recent advances in computer-aided drug design for the rational design of new compounds against HIV-1 RT using methods such as molecular docking, molecular dynamics, free energy calculations, quantitative structure-activity relationships, pharmacophore modelling and absorption, distribution, metabolism, excretion and toxicity prediction are discussed. Successful applications of these methodologies are also highlighted.

Key words: HIV-1 - computer-aided drug design - reverse transcriptase inhibitors - molecular modelling

Established in 1983 as the causative agent of the acquired immune deficiency syndrome (AIDS) (Barre-Sinoussi et al. 1983), the human immunodeficiency virus (HIV) remains a worldwide health care issue. HIV has two known variants: HIV-1, which causes HIV infections worldwide, and HIV-2, mostly confined to West Africa (Reeves & Doms 2002). Thirty years of research and technological innovation have allowed validation of several steps of the HIV life cycle as intervention points for antiretroviral therapies. The highly active antiretroviral therapy (ART) is the standard treatment for HIV-infected patients and consists of the combination of three or more HIV drugs to reach maximal virological response and reduce the potential development of antiviral resistance (Asahchop et al. 2012). Currently, 26 antiretroviral drugs have been approved by the United States Food and Drug Administration (FDA) (FDA 2014).

Although the currently available ART proved that HIV infection is treatable, some challenges remain (Broder 2010). One important factor is the constant occurrence of new infections in many parts of the world. According to the Joint United Nations Programme on HIV/AIDS, approximately 35 million people were living with HIV and an estimated 2.3 million new HIV infections happened globally in 2012 (UNAIDS 2013). The life-long treatment brings another challenge. It can lead to long-term cardiac and metabolic complications such as dyslipidemias, insulin resistance, lipodystrophy, heart diseases and other related disorders (Filardi et al. 2008, Silverberg et al. 2009). Also, treatment can be impaired by the development of drug resistance strains when viral suppression is not maintained (Scarth et al. 2011). A vast number of viruses are produced daily in an infected individual and genetic variation within individuals has contributed to the emergence of diverse HIV-1 subtypes, complicating extensively the development of active drugs (Sarafianos et al. 2004). Therefore, current antiretroviral research efforts have been aiming at refining present therapies and discovering new drugs with lower toxicity and favourable resistance profile (Ghosh et al. 2008, 2011, Maga et al. 2010, Quashie et al. 2012, Cao et al. 2014, Michailidis et al. 2014).

Presently, computational methods are an important part of the drug design process and this kind of modelling is often denoted as computer-aided drug design (CADD). Computational methods can offer detailed information about the interaction between compounds and targets, increasing the efficiency and lowering the cost of research in several stages of drug discovery (Kirchmair et al. 2011). Choosing the most appropriate computational technique to apply when planning novel drugs depends on the understanding of the target of interest (Jorgensen 2004). So far, various computational methods have been employed to the development of anti-viral drugs [reviewed by Kirchmair et al. (2011) and Wlodawer (2002)]. It is noteworthy that some approved drugs for the treatment of an assortment of diseases owe their discovery in part to CADD methods [recently reviewed by Sliwoski et al. (2014)]. This group includes anti-HIV drugs such as protease inhibitors saquinavir (Invirase®), ritonavir (Norvir®) and indinavir (Crixivan®), integrase inhibitor raltegravir (Isentress®), reverse transcriptase (RT) inhibitor rilpivirine (RPV) (Edurant®) and fusion inhibitor enfuvirtide (Fuzeon®).
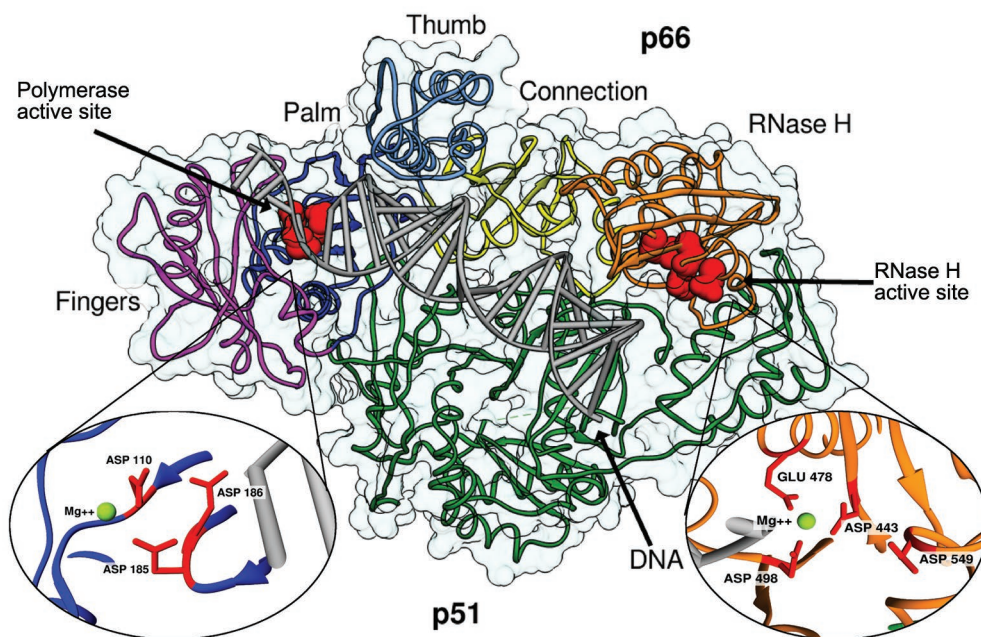
Fig. 1: structure of human immunodeficiency virus-1 reverse transcriptase in complex with DNA [Protein Data Bank code: 1T05 (Tuske et al. 2004)]. The two domains are the p66 (coloured) and the p51 (green). The polymerase domain displays a highly conserved structure that resembles the shape of the human right hand, consisting of fingers domain (magenta), palm domain (blue), thumb domain (light blue). The p66 subunit also includes the connection domain (yellow) and ribonuclease H (RNase H) domain (orange). The polymerase active site is located in the canter of palm, fingers and thumb subdomains. The three catalytic aspartic acid residues (110, 185 and 186), shown in red, are located in the palm subdomain and bind the cofactor divalent ion (Mg^{2+}). The RNase H domain is situated at the p66 C-terminus, approximately 60 Å from polymerase active site. The RNase H active site contains a DDE motif comprising the carboxylates residues ASP443, GLU478, ASP498 and ASP549 that can coordinate a divalent Mg^{2+} ion.

The goal of the present review is to give an overview of CADD methods, the challenges involved and current innovations when modelling one of the HIV-1 enzymes: the RT.

HIV-1 RT enzyme and inhibitors

The HIV-1 enzyme RT is a primary target for antiretroviral drugs. Today, 13 inhibitors act against it, including the very first drug used in HIV treatment, the nucleoside RT inhibitor (NRTI) zidovudine (AZT) (Retrovir®) (Esposito et al. 2012). RT is the enzyme that converts the single-stranded RNA viral genome into a double-stranded DNA (dsDNA) provirus, which is afterwards imported into the cell nucleus to be integrated into the host chromosome with the help of integrase (Esposito et al. 2012), another HIV enzyme. Other crucial activities of the retrotranscription process can be attributed to this highly dynamic enzyme: an endonucleolytic ribonuclease H (RNase H) activity and strand transfer (Liu et al. 2008). RT is a heterodimer (Fig. 1) composed of two subunits of 560 and 440 amino acid (aa) residues, referred to as p66 and p51, respectively (Menendez-Arias 2013). These subunits share almost the same aa sequences. However, p51 lacks the catalytic activity and the RNase H domain, performing a structural role (Kohlstaedt et al. 1992). Unlike p51, p66 has a more flexible structure and contains the polymerase and RNase H active sites (Kohlstaedt et al. 1992). Although, all the commercially available RT-targeting drugs affect the polymerase activity inhibiting its

function, some RNase H inhibitors have recently been designed and studied (Tramontano & Di Santo 2010, Distinto et al. 2013) (Steitz 1999, Tuske et al. 2004).

The two main classes of RTIs include NRTIs and nonnucleoside transcriptase inhibitors RTIs (NNRTIs). The NRTIs are composed of modified nucleosides that mimic and compete with natural substrates for binding and incorporation at the polymerase site (Fig. 2B) (De Clercq 2010). They act as chain terminators due to the lack of a 3'-OH group on their sugar moiety. Similarly to their natural counterparts, the NRTIs need to be converted in 5'-triphosphate nucleotides by host-cell kinases to compete with the analogous deoxynucleotide-triphosphates and consequently be incorporated into the growing DNA strand (Esposito et al. 2012). The current clinically available NRTIs are structurally similar to pyrimidine and purine analogues, including thymidine analogues AZT and stavudine (Zerit®); together with cytidine analogues zalcitabine (Hivid®), lamivudine (Epivir®) and emtricitabine (Emtriva®). Purine analogues include the inosine analogue didanosine (Videx®) along with the carbocyclic nucleoside analogue abacavir (Ziagen®), a guanine analogue when in its active form (Fig. 3) (Mehellou & De Clercq 2010).

In the NRTI class, there are RTIs that already have a phosphate group incorporated into their structure. Also known as nucleotide RTIs, such as tenofovir (TFV) (Fig. 3), formulated as TFV disoproxil fumarate (TDF) (Viread®), they require only two phosphorylation steps to achieve

their active triphosphate derivatives (Squires 2001). However, their mode of action is the same as for the NRTIs.

The NNRTIs are allosteric inhibitors of DNA polymerisation. These compounds bind in a noncompetitive manner to a hydrophobic pocket (Fig. 2A) located approximately 10 Å away from the polymerase active site, causing conformation changes that impair DNA synthesis (Squires 2001). During the DNA synthesis, the RT fits a "closed" conformation bringing the fingers and thumb subdomains closer to the palm one and allowing the binding of nucleic acids. The presence of an NNRTI leads to an open conformation that restricts the thumb to a hyperextension position, which prevents the polymerisation (de Bethune 2010, Das et al. 2012). The currently approved NNRTIs are nevirapine (NVP) (Viramune®), efavirenz (Sustiva®), delavirdine (DLV) (Rescriptor®), etravirine (ETR) (Intelence®) and RPV (Fig. 4).

Despite their popularity and the number of drugs already approved for this class, most RTIs have their antiviral potency limited by several factors such as mutations in the binding site, drug-drug harmful interactions, toxicity and long-term complications (Ho & Hitchcock 1989, Waters et al. 2007, Johnson et al. 2008, Cihlar & Ray 2010). Consequently, new inhibitors are being sought out and, supported by the available knowledge of the RT structure and its known inhibitors, the field of drug design has been adequately applied to study and optimise lead compounds. RT has been the focus of extensive research, including several structural biology studies that resulted in the determination of numerous crystallographic structures. Currently, over 100 RT crystal structures are available in the Research Collaboratory for Structural Bioinformatics Protein Data Bank (PDB) repository (Berman et al. 2000). The available RT crystal structures provide insights into the conformational flexibility of the protein, including the conformational changes induced by inhibitor and DNA binding (Titmuss et al. 1999). For instance, the formation of the nonnucleoside inhibitor-binding pocket (NNIBP) is in-

duced by the presence of an NNRTI, i.e., it only exists in RT structures complexed with this kind of inhibitors. The "open" and "closed" conformations can be found in crystal structures with bound and unbound DNA, respectively. The RT structures are alike, presenting some structural changes mainly in the binding pockets. Com-
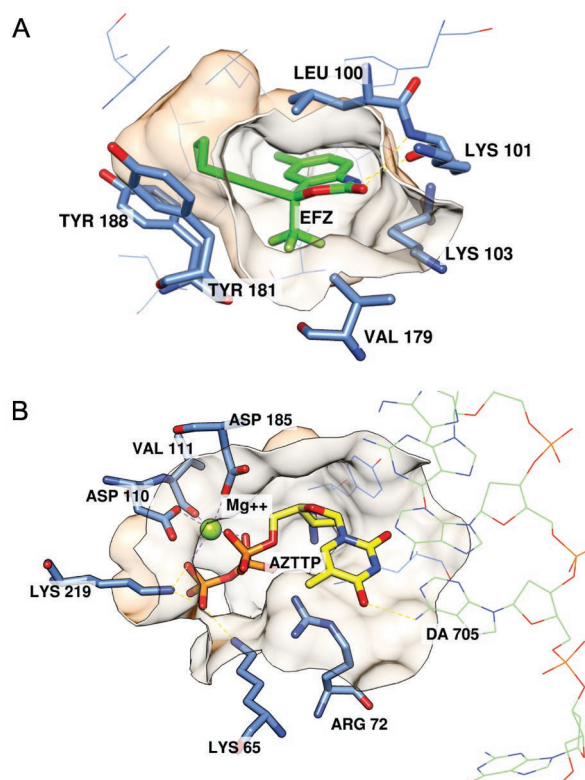

Fig. 2A: efavirenz (EFZ) (green) within the nonnucleoside reverse transcriptase inhibitor (NRTI) allosteric binding site [Protein Data Bank (PDB) code: 1FK9 (Ren et al. 2000b)]: B: zidovudine (AZT) (yellow) within the NRTI binding site [PDB code: 3V4I (Das et al. 2012)].
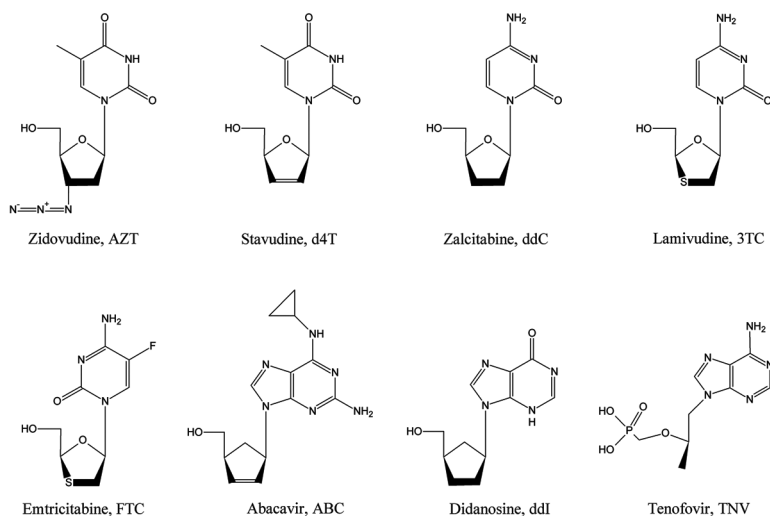

Fig. 3: chemical structures of eight approved nucleoside and nucleotide reverse transcriptase inhibitors.
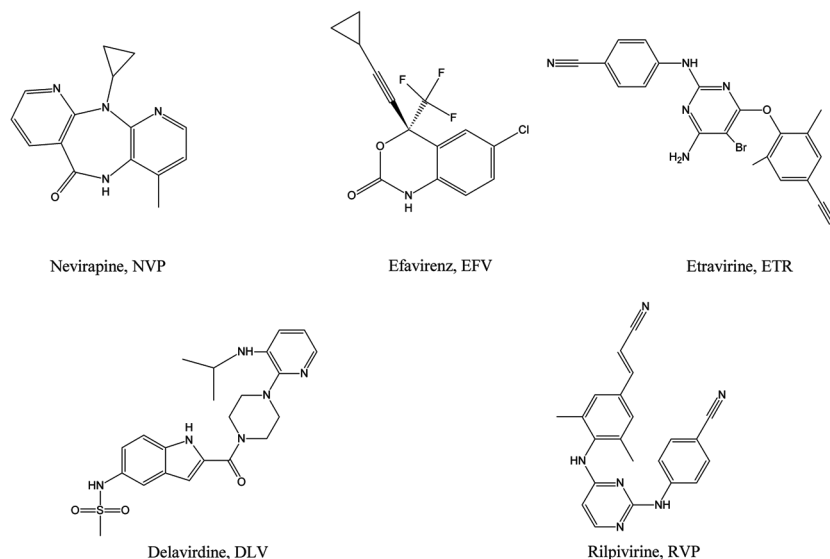
Fig. 4: chemical structures of five approved nonnucleoside reverse transcriptase inhibitors.

monly, when combined with computational methods, crystallographic structures provide molecular insights into drug-target interactions and the mechanisms that set different drug responses. Computational studies, frequently applied in CADD, such as molecular docking, molecular dynamics (MD), free energy calculations, quantitative structure-activity relationships (QSARs), pharmacophore modelling and absorption, distribution, metabolism, excretion and toxicity (ADMET) have been performed using the RT and its inhibitors as targets. A successful example of the multidisciplinary effort in drug discovery, when modelling RTIs, is the 2011 FDA-approved NNRTI RPV. RPV was developed by combining chemical synthesis with broad antiviral screening; bioavailability and safety assessments in animals and molecular modelling, including analysis of three-dimensional (3D) structures and ligand-target relationships by molecular docking (Janssen et al. 2005).

Molecular docking

A molecular docking study can provide a better understanding of the interactions between a protein and a ligand. Such applications of this method in finding lead compounds are described in details by Shoichet et al. (2002), Kroemer (2007) and Cavasotto and Orry (2007). Docking begins with sampling ligands orientations and conformations within the target binding site [for reviews see Taylor et al. (2002), Moitessier et al. (2008), Meng et al. (2011) and Yuriev and Ramsland (2013)]. Afterwards, the best poses for each ligand are determined and the compounds are ranked according to a scoring function (Lahti et al. 2012). One of the earliest docking methods was constructed based on the lock-and-key theory of ligand-protein binding, where both the protein and ligand structures are treated as rigid bodies (Kuntz et al. 1982). Currently, the most popular docking programs address the ligand flexibility when binding to rigid targets, such

as AutoDock (Goodsell et al. 1996), DOCK (Ewing et al. 2001), FlexX (Kramer et al. 1999), Glide (Friesner et al. 2004, Halgren et al. 2004), GOLD (Verdonk et al. 2003), Molegro Virtual Docker (Thomsen & Christensen 2006), AutoDock Vina (Trott & Olson 2010) and Surflex (Jain 2003, 2007, Spitzer & Jain 2012), to name a few.

The most explored RTIs are the NNRTIs, with a large number of chemical and structurally diverse compounds identified as genuine inhibitors that suppress HIV-1 replication (De Clercq 2009, de Bethune 2010). Although diverse, all compounds bind in the NNRTI binding pocket in similar conformation and manner (Zhan et al. 2013). The NNRTI binding pocket consists of hydrophobic residues with significant aromatic character (Y181, Y188, F227, W229, Y232 and Y318 of p66) and hydrophilic residues (K101, K103, S105, D192, E224 and H235 of p66 and E138 of p51) (Sluis-Cremer et al. 2004). The solvent accessible entrance is formed by the residues L100, K101, K103, V179, Y181 and E138 (Fig. 2A). However, this open state of the binding pocket is only noticeable when the structure is co-crystallised with NNRTIs, mainly due to significant torsional shifts of the Y181 and Y188 residues to accommodate the ligand (Hsiou et al. 1996). In the absence of a ligand, the binding pocket is blocked since the side chains of Y181 and Y188 are situated at the hydrophobic core, representing a closed state of the pocket. This inherent flexibility of the binding pocket provides a challenge to molecular docking. Previous docking studies showed that the difference in geometries can affect the accuracy of ligand binding energies when docking other NNRTIs into the inhibitor binding pocket (Smith et al. 1995, Titmuss et al. 1999). Numerous studies (Titmuss et al. 1999, Zhou et al. 2002, Ragno et al. 2005, Sherman et al. 2006, Ivetac & McCammon 2011) have reported the employment of molecular docking, by itself or in combination with other molecular modelling techniques, upon targeting the binding pocket with different approaches to ligand and receptor flexibility.

The search for novel anti-HIV inhibitors is also extended to natural products [reviewed by Asres et al. (2005) and Vo and Kim (2010)]. Historically, natural products have been a prolific source for lead drugs and continue to provide structural templates for drug discovery, since the majority of all market drugs have their origin *in nature* (Chin et al. 2006, Newman & Cragg 2012). However, only a few of anti-HIV natural products that have been reported to exhibit inhibition activities have reached clinical trial and so far none of them is commercially available (Asres et al. 2005). Recently, computer-aided approaches have found room in natural product research (Rollinger et al. 2006a, b, 2008) and some studies had RT as their target (Sangma et al. 2005, Ehrman et al. 2007, Seal et al. 2011, Ashok et al. 2015).

In an early work (Currens et al. 1996), a natural product extracted from the tropical rainforest tree *Calophyllum lanigerum*, calanolide A, showed promising results as an NNRTI. However, this natural product is difficult to purify from its natural source in a sufficient amount for clinical use and its low therapeutic index contributed to the delay of its clinical development. Lu et al. (2012) investigated a calanolide A analogue, 10-chloromethyl-11-demethyl-12-oxo-calanolide A (F18), using experimental and docking studies. F18 was chosen since it showed high potency against wild-type (WT) HIV-1 [half maximal effective concentration (EC_{50}) = 7.4 nM] in a TZM-bl cell based assay. Docking studies were conducted using AutoDock 4.2 and the structures of F18 and NVP (control) were docked into three different RT crystal structures: WT [PDB code: 1VRT (Ren et al. 1995)], L100I mutant [PDB code: 1S1U (Ren et al. 2004)] and Y181C mutant [PDB code: 1JLB (Ren et al. 2001)]. The results showed that F18 had a rigid structure and restricted binding when compared to NVP. In the WT structure, no meaningful interactions between F18 and the binding pocket residues were predicted. On the other hand, an aromatic interaction between Y188 and NVP was observed, indicating that the WT structure is more sensitive to NVP than to F18. With the L100I structure, the binding pocket was altered and there was less hydrophobic interaction with F18 and thus, L100I mutation conferred moderate resistance to F18. However, Y181C structure was favourable to F18, since the change of tyrosine to cysteine permitted more spatial flexibility to the compound and increased antiviral activity. The docking analysis was later correlated with cell-based assays and both results indicated that F18 might bind to a distinct motif on the RT from that of NVP, which can be the cause of its drug resistance profile.

A molecular docking study by Allen et al. (2015) evaluated the latest version of the program DOCK with the SB2012 test set [expanded from the SB2010 (Mukherjee et al. 2010)] composed of a diverse range of receptors, including 21 RT-NNRTI crystal structures. All receptors were structurally aligned to facilitate docking all the ligands into all of them. Therefore, docking statistics was based on the crystallographic ligand and its pose prediction the root-mean-square deviation (RMSD) when docked into its native receptor (redocking) or a nonnative structure from the same drug-target family (cross-docking). Also, if any particular pose comparison achieved either an RMSD over 2.0 Å or a positive score, it was considered a nonviable reference and the pairing was not included in the docking statistics. Overall, the success rates for RT structures were 71.4%, whereas the success rate of redocking by itself was 95.2%. Scoring and sampling failure rates were 18.3% and 10.4%, respectively. However, there was a high incidence of nonviable pairings, 220 nonviable out of 441 pairings, which could be related to the conformation of the binding site and the presence six mutations (L100I, K101E, K103, E138K, Y181C and Y188C) known to confer resistance to NNRTIs in the set of structures. The strategy of starting the docking process from 3D structures of RT-NNRTI complexes is a very delicate one, not only due to the intrinsic flexibility of the RT, but also due to the allosteric binding pocket conformational changes to accommodate the NNRTIs (Tronchet & Seman 2003). In general, proteins go through conformational changes when performing their functions and the molecular recognition between a protein and small molecules involves structural flexibility (Ivetac & McCammon 2011). Recently, more advanced methods have introduced protein flexibility and its influence on ligand recognition, supported by the exponential growth in computer processing and disk capacity (Carlson & McCammon 2000, Cavasotto & Singh 2008).

An attempt to account for a small amount of plasticity of the receptor is to use soft scoring functions, capable of tolerating some overlapping between the ligand and the protein, but still maintaining a rigid receptor (Jiang & Kim 1991, Claussen et al. 2001). This implementation, which is known as soft docking, is computationally efficient since only the scoring parameters need to be changed whereas everything else remains unaltered when compared to rigid docking (B-Rao et al. 2009). Although, this option has been pursued due to its computational simplicity, it can introduce false-positives if the tolerance is set too high (Ivetac & McCammon 2011). Other strategies to incorporate receptor flexibility involve sampling of side-chain conformers within the binding pocket, with the use of a library of rotamers and the use of an ensemble of receptor structures (Cavasotto & Orry 2007).

Glide is one of the programs that uses soft docking receptors by scaling the van der Waals radii (Elokely & Doerksen 2013). In their study, Bahare and Ganguly (2014) evaluated the accuracy of their docking procedure consisting of the docking of the NNRTI TNK 651, extracted from a X-ray crystallographic RT structure [PDB code: 1RT2 (Hopkins et al. 1996)], by means of two different programs: Glide and FlexX. The first employs a hybrid approach that combines one or more docking algorithms in the generation of the ligand poses (Moitessier et al. 2008). The second is based on incremental construction, where the ligand is built dynamically in the active site, frequently counting on libraries of favoured conformations (Moitessier et al. 2008). The RMSD values between the docking prediction and the experimental conformation of TNK 651 were 0.370 Å and 1.254 Å, with Glide and FlexX, respectively.

Recently, Fraczek et al. (2013) assessed the ability of docking programs and their scoring functions to predict the relative biological activity of triazole NNRTIs. In total 111 known 1,2,4-triazole and 76 other azole type NN-RTI were submitted to different docking protocols that involved softened van der Waals potentials (FlexX, Molegro Virtual Docker and Glide XP and SP), ligand flexibility (AutoDock Vina) and receptor flexibility employing the Induced Fit Docking (IFD) method (Sherman et al. 2006). The IFD method combines an iterative procedure to obtain initial poses allowing flexibility into rigid receptors, followed by a technique for modelling receptor conformational changes, present in the refinement module of Prime (Jacobson et al. 2002, 2004) program that explores flexibility. However, while the method allows efficient small backbone movements, it is inappropriate to more severe conformational changes due to an increase in complexity and computational cost (Ivetac & McCammon 2011). The RT structures 2RKI (Kirschberg et al. 2008), in complex with a triazole and 3DLG (Ren et al. 2008), in complex with a benzophenone, were used in the docking procedures. Since the core structure of the compounds is similar, the triazole binding mode was assumed as the reference pose. For 2RKI, all programs showed good predictions of ligand orientation in the binding site when compared to the reference pose, Glide SP (97.3%), Glide XP (59.5%), IFD (97.3%), Auto-Dock Vina (94.6%), FlexX (65.8%) and Molegro Virtual Docker (61.3%). The predictions for 3DLG were lower to most of them, Glide SP (82.9%), Glide XP (67.6%), IFD (88.3%), AutoDock Vina (75.7%), FlexX (36%) and Molegro Virtual Docker (71.2%). However, none of the scoring functions reached a perfect ranking of the compounds according to their activities. Glide XP achieved the highest correlation, Spearman's ρ of 0.7, which corresponds to around 0.75 probability of identifying the most active compound from two compounds. The outcomes from this study suggested that different docking methods can provide good binding mode predictions, yet results should not rely only on docking scores when trying to rank active compounds with different potencies.

Another approach considers a discrete number of receptor conformations (obtained either experimentally or by computational means) to represent the flexibility instead of making the protein flexible throughout the docking process (Knegtel et al. 1997, Osterberg et al. 2002, Huang & Zou 2007). This procedure is known as ensemble docking. Structure ensembles can diverge in their sidechain, loops and domain orientations. A study by Meleddu et al. (2014) performed ensemble docking experiments in an attempt to predict the binding mode of a compound from a series of dual inhibitors, a single molecule that is able to inhibit two enzymes activities, of RT-associated functions. Since the most promising compound showed activity in vitro against both the RNA-dependent DNA polymerase (RDDP) and RNase H of RT [half maximal inhibitory concentration (IC_{50}) of 6 ± 2 μM and 4 ± 1 μM, respectively], docking was performed into six NNRTI bound structures [PDB codes: 1VRT, 2ZD1 (Das et al. 2008), 1EP4 (Ren et al. 2000c), 3QO9 (Das et al. 2011), 1RTI (Ren et al. 1995) and 1TV6

(Pata et al. 2004)] and one RNase H inhibitor bound structure [PDB codes: 3LP2 (Su et al. 2010)], using the QM-Polarized Ligand docking protocol. The 1TV6 structure was also considered for RNase H docking experiments, using the whole domain. Post-docking procedures based on energy minimisation and binding free energy [molecular mechanics with generalised Born and surface area solvation (MM-GBSA) calculations] were also performed. The best ensemble score was obtained in the WT structure 1RTI (GScore of -11.04 kcal/mol), however the best free energy of binding (-48.0 kcal/mol) when comparing MM-GBSA values were obtained in the mutated Y181C NNIBP of the 1TV6 structure. Poses in the RNase H binding sites achieved worse ensemble scores (> -7.50 kcal/mol) and free energy of binding than those in the NNRTI binding pocket (> -38.60 kcal/mol). Biochemical and modelling studies combined suggested that polymerase inhibition was due to the compound binding into the NNRTI pocket, where the RDDP activity was retained in all RT strains. Whereas, binding into an allosteric site close to RNase H catalytic residues might be responsible for RNase H inhibitory activity, since a single-point mutation inserted in this site decreased the inhibition of the RNase function by the compound. Therefore, the compound might behave as a dual-site dual-function inhibitor.

As significant as docking methods are in drug discovery, the search for potential drug candidates often initially requires screening libraries of available compounds to identify novel hits. This computational approach, referred to as virtual screening (VS) is an important drug discovery tool, which allows identification of lead compounds among large databases, thanks to its ability to discriminate between true and false-positives (Cummings et al. 2005). Several VS approaches have been described, among which the most common one uses molecular docking as a faster and more cost-effective alternative than experimental high-throughput screening. VS aims to reduce a vast virtual library of approximately 10^5-10^6 chemical compounds, to a more manageable number for experimental screening against biological targets and further synthesis of analogues, which could lead to potential drug candidates.

Herschhorn and Hizi (2008) conducted a VS study to identify novel NNRTIs from a commercially available library of 46,000 compounds (Tripos Leadquest3) against two RT crystal structures [PDB codes: 1FK9 (Ren et al. 2000b) and 1DTQ (Ren et al. 2000a)]. The library of "druglike" (Lipinski et al. 2001) compounds was docked into the two structures in parallel using Surflex, in a way that the difference in the results for both structures could be accounted. The molecules were ranked by their score according to the average of their top ten conformations. Compounds with exceptionally high scores or high ratio of docking score to the number of rotational bonds were all included in the list of potential compounds. Overall, 740 out of the 46,000 were selected, purchased and submitted to a primarily experimental test for inhibiting in vitro RDDP of recombinant HIV-1 RT. Only 71 of the selected compounds inhibited more than 84% of RT-associated RDDP at the tested concentration (50 μg/mL). A total of

17 novel compounds were later chosen to further experimental evaluations according to their high RT inhibition at nanomolar concentrations and structural diversity. Some of the molecules shared similar elements with the phenylethylthiazolylthiourea (PETT) series (Ahgren et al. 1995, Ren et al. 2000a), which are known NNRTIs, however instead of the original PETT pyridine rings, other chemical structures such as phenyl, furan or cyclohexane rings were found. The original inhibitor in the selected structure 1DTQ is a PETT derivative, showing that the docking process could retrieve compounds that resemble the native inhibitor found in the crystallographic structure.

An interesting report displays an example in which a new class of inhibitors was identified from VS; despite the fact that the compounds initially evaluated were false-positives. After reporting failure to yield active NNRTIs from their top-scoring compounds, Barreiro et al. (2007b) still pursued one of the scaffolds. VS was performed in a library of 70,000 compounds (Maybridge Library) using first a chemical similarity search, considering known NNRTIs as reference structures and then the subsequent library (2,000 molecules) was docked into a single RT structure [PDB code: 1RT4 (Ren et al. 1998)] using Glide 3.5. The top 100 scored compounds were later submitted to MD simulations to estimate the free energy of binding by means of the MM-GBSA method (Kollman et al. 2000), as well as to evaluate the change in free energy of hydration using the GBSA (Still et al. 1990). Finally, four top-scoring compounds were subjected to experimental evaluations and showed no ability to inhibit HIV replication. Nevertheless, the top-scoring compound, a confirmed false-positive of the VS procedure, was assessed by computational analysis and modifications were made in its structure, removing or adding functional groups to create analogues. These last sets of compounds provided more favourable results than the original one and some of them were reported to be potent anti-HIV agents (lowest IC_{50} = 0.31 μM). This study demonstrates the importance of chemical insights and that even compounds that do not inhibit an enzyme with detectable activity may provide a scaffold to find new inhibitors.

In a recent study, Chander et al. (2015) performed a VS study to identify novel NNRTIs that could potentially act against WT and drug resistance RT strains. Firstly, a screening of 30,000 molecules, extracted from the Maybridge Library and filtered by the Lipinski Rule of Five (Lipinski et al. 2001), were performed by the Glide high-throughput VS module against a WT RT structure [PDB code: 4G1Q (Kuroda et al. 2013)]. Afterwards, compounds in the top 10% were retrieved and submitted to docking into the WT structure using the Glide SP module. This procedure was once again performed with the Glide XP module. After all, the top 30 scored hits were subjected to another round of docking into the RT mutant strains K103N [PDB code: 3TAM (Gomez et al. 2011)] and K103N/Y181C [PDB code: 4I2Q (Johnson et al. 2012)]. Out of 30 compounds, around nine exhibited good binding modes and hydrophobic interaction with binding site residues Y181, Y188, F227, W229 and Y318 toward all three RT strains. Hydrogen bonding interaction with the residue K101 was also presented for the majority of the hit compounds. Although no experimental studies were conducted, all nine compounds had favourable predictions for ADMET properties.

In the next section, we discuss works where MD simulations alone or in combination with other methods were applied to RT systems.

MD

MD is a powerful and extensively used method to gather information on the dynamical properties and processes of proteins and other biological macromolecules, also time-dependent and thermodynamical information (Adcock & McCammon 2006). It is a commonly employed tool in a vast number of fields such as structural biochemistry, biophysics, molecular biology and pharmaceutical industry (Galeazzi 2009). MD simulations have a broad range of usage. For instance, they are extensively employed to refine experimental or model-derived protein structures, to inspect the strength and stability protein-ligand complexes resulting from a docking study, to aid drug discovery and many more (Lahti et al. 2012).

In MD simulations, physical movements of atoms and molecules are portrayed over time, usually over tens to hundreds of nanoseconds (ns) reaching up to milliseconds, a feat provided by iterative calculations of the forces present that act on the system (a complex of protein, ligand, solvent and often a lipid bilayer) and the consequential movements (Adcock & McCammon 2006). A successful MD simulation depends on the choice of a suitable energy function for describing the inter and intramolecular interactions (Galeazzi 2009). Forces between atoms and the potential energy of the system are described by the force fields, well-parameterised functions obtained from experimental or quantum mechanical studies. Widely applied force fields included several versions from OPLS-AA (Jorgensen et al. 1996), CHARMM (MacKerell et al. 1998), AMBER (Cornell et al. 1995) and GROMOS (Oostenbrink et al. 2004). Common MD softwares are GROMACS (Van Der Spoel et al. 2005), AMBER (Pearlman et al. 1995) and NAMD (Kale et al. 1999).

RT flexibility is essential for the polymerisation and RNase H activities, in addition to inhibition of enzymatic activity. Madrid et al. (2001) conducted a study to analyse flexibility for two RT systems, bound [PDB code: 2HMI (Ding et al. 1998)] and unbound [PDB code: 1DLO (Hsiou et al. 1996)] to dsDNA, by means of MD simulations. MD simulations of 125 ps, with an integration step of 1 fs, were performed using AMBER keeping DNA and protein unrestrained in solution. From the simulations analysis, it was concluded that the RT flexibility depends on its ligation state. The complex RT/dsDNA showed more flexible regions than the unbound RT, particularly in the fingers and thumb p66 subdomains. This outcome was consistent with the conformation changes found in crystallographic structures and biochemical data. Although the simulation times to the systems were very short, probably due to hardware limitations at that time, these simulations showed that it is possible to complement the RT information available from existing crystal structures by means of MD.

A few years later, Ivetac and McCammon (2009) published an impressive paper focused on the inhibition mechanism of NNRTIs, using structures from crystallographic and MD data, through multicopy MD simulations (cumulative total simulation time of 360 ns). Principal components analysis (PCA) were employed to interpret the dynamics from both a crystallographic ensemble of 13 RT structures (1 apo, 2 substrate-bound and 10 NNRTI-bound) and a MD ensemble from three simulated systems (RT with NNRTI binding pocket closed, open and bound to NVP). PCA has been performed previously on other proteins for which substantial crystallographic data exists (van Aalten et al. 1997, Gorfe et al. 2008). Comparison of the systems showed similar movements, characterised by opening/closing of the fingers and thumb subdomains, between NNRTI-free simulations and crystallographic ensemble and quite distinct of those of the NNRTI bound simulations. The fingers and thumb subdomains in the NNRTI bound simulation made movements roughly orthogonal to those presented by the other simulated systems. This difference might demonstrate that the effect of an NNRTI is to constrain the motion between these subdomains. Consequently, NNRTIs may act as "molecular wedges" sterically blocking the full range of the subdomain movements since the NNRTI binding pocket is located proximally of their hinge points. The time scale of the simulations and the multicopy approach, chosen in this work, helped enhance the sampling of the dominant motions found in the ensembles.

Both studies displayed the use of MD simulations to understand RT flexibility. However, MD simulations and docking are usually employed as complementarily methods. While docking techniques allows for a vast exploration of ligand conformation and screening of large libraries in a short time, MD simulations can be employed to optimise conformations of the receptor-ligand complex, explore other receptor conformations and achieve accurate binding free energy predictions (Alonso et al. 2006).

Methods to consider receptor flexibility are also employed in VS, commonly using conformation ensembles. The idea is that screening against several structures might increase the chances of finding the right receptor conformation that accommodate ligand sampling. In a recent paper, Ivetac et al. (2014) described a VS approach using ensembles of experimental and theoretical RT structures to identify novel NNRTIs. A screening library of 2,864 compounds from the National Cancer Institute (NCI) was collected combining compounds from the NCI Diversity Set II (NCIDS-2), a general chemical diversity subset and molecules similar to a set of six known NNRTIs filtered from the NCI repository. The relatively small library was chosen due to the computational demands of the subsequent steps of docking and multistructure docking. An ensemble of diverse RT crystal structures complexed with different NNRTIs was selected to guarantee variation in the conformation of the NNRTI binding site. The screening library was docked, using Glide, into each of the 10 RT structures [PDB codes: 1VRT, 1RT1 (Hopkins et al. 1996), 1HNV (Ding et al. 1995), 1FK9, 1RTH (Ren et al. 1995), 1VRU (Ren et al. 1995), 1EP4, 1BQM (Hsiou et al. 1998), 1KLM (Esnouf et al. 1997) and 2ZD1] and a score

was calculated according to the average binding energy of each compound across the ensemble. This score was then used to rank the screening library, favouring molecules that could bind to diverse conformations of the pocket, as opposed to only binding favourably to a particular conformation. MD simulations of 30 ns were performed using the 1VRT structure in complex with four different NNRTIs (each solvated system had approximately 160,000 atoms). Snapshots of the simulations with similar conformations were clustered to ensure representative structures, yielding 30 clusters for each system. From the previous step, only 150 compounds were kept to take part in this secondary screening against all the 120 theoretical conformations. The compounds were ranked and re-scored by taking the mean rank of each compound throughout all simulation systems. Finally, 16 compounds were experimentally tested for inhibition of HIV infection and two of them showed potential inhibition of RT polymerase activity (with potency similar to the positive control NVP). Although successful, this is a very expensive and time-consuming approach. The use of the accurate scoring function is of great importance since a broad range of ligands will be able to fit in some of these more relaxed receptor conformations (Alonso et al. 2006).

This last study is an exceptional example of combining docking methods to MD simulations in a successful way. Given that the use of structures from MD simulations has been successfully employed in docking methods, it has also been investigated the benefits of using multiple of such structures, obtained from a crystallographic one. Nichols et al. (2011) presented an analysis of the usage of structures from MD simulations, with respect to the experimentally determined ones, to improve the predictive power in VS. Two proteins structures were selected, being the HIV-1 RT one of them. Their work consisted in the simulation of two bound systems (PDB code: 1VRT bound with α-APA and UC-781, both NNRTIs) as well as two unbound systems, one with the NNIBP in its open state (PDB code: 1VRT with the NNRTI extracted) and the other with the NNIBP in its closed state (PDB code: 1DLO). All simulations were performed using the GROMACS software along with the GROMOS 53A6 force field (Oostenbrink et al. 2004) and four independent 30 ns trajectories were generated and MD snapshots were extracted. A screening library, consisted of 20 diverse known RTIs combined with a set of 1,323 decoys (compounds from the NCIDS-2) as RT ligands, was prepared. The screening library was docked in each structure, 2,500 MD snapshots for each of the four systems as well as 15 RT X-ray structures (10 diverse NNRTI-bound and 5 NNRTI-free states) using Glide. The predictive power of VS was tested by using the receiver operating characteristic curves (Triballeau et al. 2005), a classification model to establish the probability of ranking active compounds over inactive ones (decoys). This analysis was performed for all different receptor conformations and the results were compared with the ones obtained when employing the same VS approach to the crystallographic structures. In all systems, the maximum MD area under the curve (AUC) value (bound AUC = 0.96 and unbound AUC = 0.77) surpass-

es the maximum X-ray AUC value (bound AUC = 0.93 and unbound AUC = 0.49). By contrast, when considering mean AUC values, on average bound MD snapshots (<AUC> = 0.76) were less predictive than bound X-ray structures (<AUC> = 0.81), indicating that some MD originated structures had inferior predictive power. However, mean values from unbound systems are equivalent (<AUC> = 0.44 for both). Overall, the advantage of using MD snapshots in VS may depend on the enrichment that a reduced number of MD generated structures provide, rather than the whole configurational ensemble. However, a more accurate method to select the best MD created structure is needed to identify the best conformations and to reduce, time-wise, MD sampling to a more efficient simulation time scale.

In later stages of lead refinement, other calculations might be needed to estimate the relative or absolute free energy of the final complexes. Some of these methods are discussed next.

Free energy calculations

Free energy calculations methodologies are currently employed in several research areas including solvation thermodynamics, molecular recognition and protein folding (Hansen & van Gunsteren 2014). Reviews and applications of free energy calculations in drug design have been described by different authors (Deng & Roux 2009, Michel et al. 2010, Hansen & van Gunsteren 2014). The most rigorous methods to compute relative free energy are free energy perturbation (FEP) and thermodynamic integration (TI) (Pohorille et al. 2010). Using computer aided statistical mechanics, this methods calculate binding free energies of small molecules to a protein through MD or Monte Carlo (MC) simulations (Deng & Roux 2009). For receptor-ligand affinities, perturbations are made to transform one ligand into another using a thermodynamic cycle (Fig. 5). These transformations comprise a coupling parameter that smoothly mutates one molecule to the other. The difference in free energies of binding, from the initial ligand to the final one is calculated by $\Delta\Delta G_B = \Delta G_X - \Delta G_Y = \Delta G_S - \Delta G_C$. To calculate the free energy differences, two transformation systems need to be prepared: one for the unbound ligands in solution (ΔG_S) and the other complexed to the receptor (ΔG_C). These methods can be used to determine the relative free energy, as the free energy is a state function that can be calculated by any reversible path between the initial and final states. Despite their accuracy, these methods are computationally expensive and with slow convergence.

Zeevaart et al. (2008), following the early success of searching and optimising of a top scoring compound (1) from a screening library [described Barreiro et al. (2007a)], reported a series of FEP guided simulations with analogues of the modified compound (2), with potencies in the 10-20 nM range. To predict relative free energies of binding, the calculations were carried out in the context of FEP/MC statistical mechanics simulations. These calculations were performed using the thermodynamic cycle theory, to interconvert two ligands unbound in water and bound to the protein. The systems were cal-
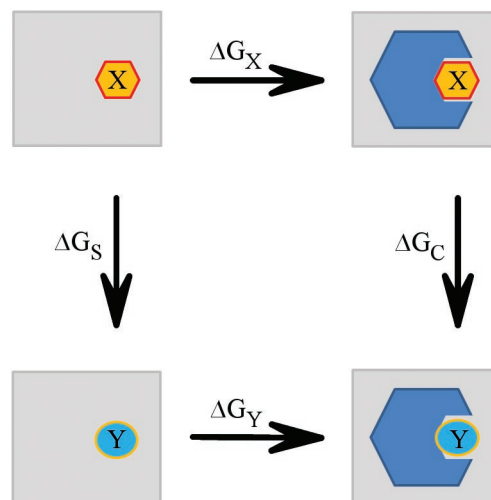

Fig. 5: thermodynamic cycle for relative free energies of binding. The receptor is in blue and X and Y are two ligands.

culated using dual-topology sampling with 14 windows or simple topology with 11 windows. In FEP calculations, a window refers to a simulation at one point along the mutation coordinate λ, which interconverts two ligands as λ goes from 0-1; the free energy changes are computed for each window, corresponding to a forward and backward increment (the space between windows $\Delta\lambda$) (Lu et al. 2004). When dual-topology is chosen, the system is prepared in a way that the two complete versions (initial state and final state) of the changing group coexist at every λ (Pearlman 1994). First, a so-called chlorine scan was performed, in which FEP calculations were used to transform each hydrogen individually in the phenyl rings into chlorine, resulting in 10 structures to be converted into compound 2. These FEP results indicated the most promising places for chlorine atoms were at positions 3, 4, 2' and 6' (Fig. 6B). Further optimisation guided the substitution at position 4, resulted in compounds with activity (EC_{50}) of 820 nM (3), 310 nM (4) and 130 nM (5) (Fig. 6C). Other FEP scans and ring modifications were made producing compounds with activity (EC_{50}) of 22 nM (6), 13 nM (7) and 6 nM (8) (Fig. 6C), the last two found in a later study (Leung et al. 2010).

The same FEP guided optimisation approach was used to improve the performance of a compound, discovered by VS using multiple proteins (Nichols et al. 2009), which showed activity against both WT and Y181C HIV-1 strains. The work by Bollini et al. (2011) started with compound 9, that presented anti-HIV activity EC_{50} of 5 μM and with the aid of FEP/MC outcomes, it was possible to yield a very potent compound (10), EC_{50} values of 55 pM, 42 nM and 220 nM against the WT, the Y181C and K103N/Y181C strains, respectively (Fig. 6D). Further optimisation of compound 10 produced compounds with EC_{50} values of 0.4 nM for the WT and 10 nM for the K103N/Y181C strain (Lee et al. 2013).

Currently developed approaches such as the linear interaction energy method (Aqvist & Marelius 2001) and the so-called MM-Poisson-Boltzmann (PB)/GBSA
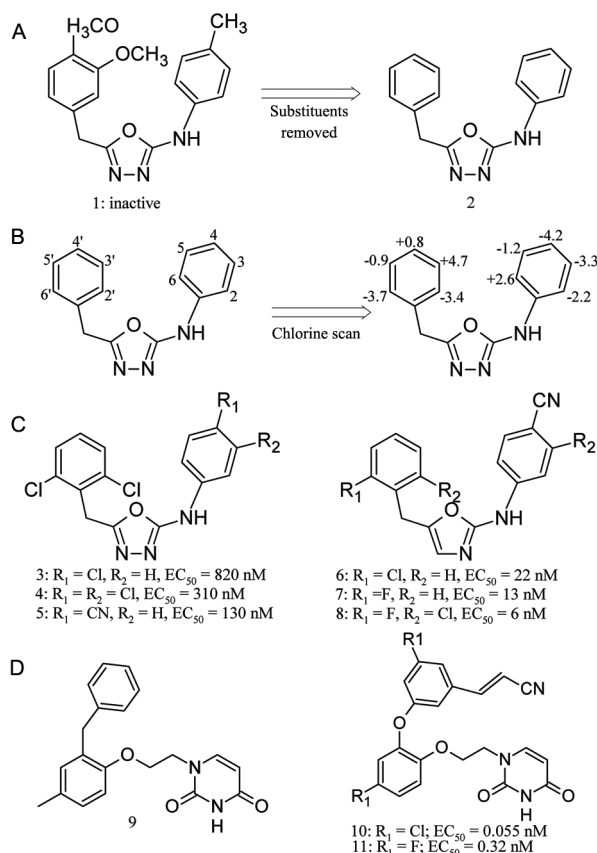
Fig. 6A: compound 1 and compound 2; B: chlorine scan results performed in compound 2; C, D: resulting compounds from free energy perturbation guided optimisation; EC_{50}: half maximal effective concentration.

(MM-PB/GBSA) method (Kollman et al. 2000), both MD-based, provide relatively good free energy predictions at a reasonable cost. Liu et al. (2014) applied the MM-GBSA method to investigate the binding affinities of NVP and two novel NVP analogues when in complex to WT RT (PDB code: 1VRT) and its mutants K103N [PDB code: 1FKP (Ren et al. 2000b)] and Y181C (PDB code: 1JLB). Based on He et al. (2005), whose study showed that NVP binds to the RT through several weak hydrogen bonds, the NVP analogues were constructed in a way that potential strong hydrogen bonds could binding to aas H235 and Y318 in the NNIBP. Alterations were also made to avoid repulsion found in NVP between a carbon atom and the side chain sulfur atom of the mutated C181. Docking was performed to predict the binding mode of the new compounds in the selected structures using AutoDock 4.2 and followed by 20 ns MD simulations for each system with AMBER 11, where snapshots were extracted to obtain ensemble-average binding free energies with the MM-GBSA method. The relative binding free energies of the two NVP analogues (-13.20 and -12.29 kcal/mol) were less favourable than that of NVP (-14.75 kcal/mol) when bound to the WT RT. However, the analogues affinity were more promising to the mutations K103N (-15.57 and -14.76 kcal/mol) and Y181C (-15.50 and -16.32 kcal/mol) than to the NVP affinity

when bound to the mutant structures (-12.14 and -10.39 kcal/mol, K103N and Y181C, respectively) and WT. This study showed that these calculations might be helpful to filter and overall assist in the design of new RT drug candidates. Although less computational demanding than FEP or TI, MM-PB/GBSA may not be accurate for the prediction of the entropic component of the free energy and errors might be produced in flexible systems since the internal energy of ligand and receptor upon complex formation are ignored (Alonso et al. 2006).

QSAR and pharmacophore studies

QSAR is an effort to associate structural or property descriptors of molecules with biological activities quantitatively (Vaidya et al. 2014). The structure-activity relationships are described in terms of physicochemical parameters such as constitutional, fragment constant, thermodynamic, conformational, hydrophobicity, topology, electronic properties, steric effects, hydrogen bond donors, hydrogen bond acceptors, among others (Kubinyi & Sadowski 1999). These descriptors can be determined empirically or by computational methods. Computational QSAR studies are often used to filter virtually large compounds libraries, to eliminate the molecules with predicted toxic or poor pharmacokinetic properties early on and to narrow the libraries to drug-like or lead-like compounds (Dudek et al. 2006). Some QSAR studies applied to RT have been reported (Gaudio & Montanari 2002, Gayen et al. 2004, Guimarães et al. 2014, Dong & Ren 2015, Nazar et al. 2015).

Recently, Tarasova et al. (2015) published a study addressing the use of data from publicly and commercially available databases to produce accurate and predictive QSAR models using RT as the case study. Two databases, Thomson Reuters' Integrity and ChEMBL (Bento et al. 2014), were chosen to collect all RTIs assayed against both WT and mutants RT. Several methods for the creation of modelling sets from the chosen databases were proposed and their accuracy investigated. The program GUSAR was used to build the QSAR models. From the Integrity database, when the compilation of modelling sets were according to their assay data (i.e., associated with just one material and method for testing), it yielded high-performance QSAR models for all RT forms. While ChEMBL database, compounds derived from individual scientific publications provided more consistent and higher quality QSAR models than the other methods employed in the same database. Although, some of the methods worked within the databases it did not work across them in a mix-and-match QSAR model approach. The lack of unified and standardised descriptors between Integrity and ChEMBL revealed to be a problem to data aggregation.

Li et al. (2008) published a paper that combined 3D-QSAR with MD simulations. The 3D-QSAR methodology consists of obtaining compound descriptors from an experimentally determined ligand and aligning conformers of the chosen dataset in space (Dudek et al. 2006). A series of diaryltriazine analogues, a category of NNRTIs, were extracted from the literature and separated in two data sets (8A-G and 9A-R) based on their structure and activity (IC_{50}). The steric and electrostatic

interactions were calculated by the comparative molecular field analysis (MFA) method and compounds with low, moderate and high activity were selected. Five physicochemical properties related to steric, electrostatic, hydrophobic, hydrogen bond donor and hydrogen bond acceptor parameters were evaluated by the comparative molecular similarity analysis method. The two most active molecules, one from each data set, 8E and 9H (pIC$_{50}$ of 1.21 and 2.04, respectively), were further analysed by 2.0 ns MD simulations. The simulations showed more hydrophobic contacts and hydrogen bonds for 9H, which might make it more active and stable than 8E. However, superposition of these two compounds showed similar binding modes with the RT, indicating that a conventional 3D-QSAR model of these two types of RTI could be constructed.

Another method frequently used in drug discovery is pharmacophore modelling. In it, the steric and electronic features a query molecule possesses, essential for receptor-ligand interaction, is analysed (Zheng et al. 2013). The resulting model can be determined either based on ligand information, by superposing a set of active molecules and selecting essential common features for their bioactivity or based on structural information. The method helps in the search for possible interaction points identified between receptor and ligands (Yang 2010). Some pharmacophore studies have been reported for RTIs (Keller et al. 2003, Balaji et al. 2004, Distinto et al. 2012).

Vadivelan et al. (2011) developed a work where potential anti-HIV lead compounds could be generated by analogue based design studies using 3D-QSAR and pharmacophore models. A training set of 36 molecules was used to develop the best model. The MFA model was generated based on the feature query of the biologically active conformation of the most active compound (pIC$_{50}$ of 8.57). The best pharmacophore model was composed of three characteristics: one hydrogen bond acceptor, one hydrophobic aliphatic and one aromatic ring. Both models were validated and their predictive ability evaluated by knowledge-based screening. A total of 10,000 molecules were generated based on the knowledge of the binding interaction of ligands to the RT and also the common features necessary for the molecule biological activity. Cross validation was made with both models and the results suggested that these techniques yield almost the same results. However, the screening produced some false-positives and a few false-negatives. Therefore, docking studies were performed on a RT structure [PDB code: 2VG5 (Spallarossa et al. 2008)] using Glide in an attempt to produce reliable true positives and negatives. The combined approach developed in their study showed a possible way to assess critically the identification and optimisation of lead compounds through better understanding of protein and ligand features.

ADMET studies

ADMET studies are commonly applied in drug discovery to optimise leads compounds into drug candidates (Selick et al. 2002). Experimental ADMET investigations allow classifying based on characteristics such as the ability to cross physiological barriers, group reactivity, metabolism and so on (Oprea et al. 2001, Selick et al. 2002,

Kubinyi 2003). *In silico* computations can be carried out to analyse the drug-likeness of a compound prior to its synthesis (Beresford et al. 2004). A series of filtering rules are defined to compute what are called descriptors that classify the compounds and to predict their ADMET properties (Lagorce et al. 2008). While these descriptors are not accurate enough to replace in vivo or in vitro methods, they can help point out physicochemical properties and lead to the optimisation of them (Gleeson et al. 2011).

An early work from Sengupta et al. (2007) analysed 15 DLV analogues for their potential to be used as drug candidates. Their approach consisted of docking the compounds to determine an initial binding mode of the ligand with the receptor. Then, free energy calculations with MM-GBSA were performed. Finally, ADME properties were estimated by Qikprop (Duffy & Jorgensen 2000). The program predicted 44 properties consisting of principal descriptors and physiochemical properties such as log P (Octanol/Water), log P Madin-Darby canine kidney (MDCK) (predicted apparent MDCK cell permeability) and log K$_p$ (skin permeability). Violations of the Lipinski's rule of five were also considered. From this analysis, 15 out of the 16 compounds showed acceptable values for all the properties analysed. Based on the overall examination, three analogues showed potential as a leads to be used for drug development. These three compounds exhibited efficient binding in the active site, showing ideal pIC$_{50}$ (~7.0) values and passed the rule of five. This work demonstrated the use of ADME properties as a tool to aggregate value to suitable candidates for drug development.

Pirhadi and Ghasemi (2012) used a combination of pharmacophore model for NNRTIs, docking and ADME studies in the search for novel compounds. Firstly, a set of 219 compounds comprising diverse structures was obtained. Based on these compounds, quantitative pharmacophore models were developed to identify critical features among NNRTIs. The best pharmacophore model took into account four descriptors, including two hydrogen bond acceptors, one hydrophobic and one aromatic feature, in agreement with previously reported pharmacophore models. The model was used as a 3D VS query for recovering novel and potent candidates from ZINC (Irwin & Shoichet 2005), resulting in 8,631 hits from this first screening. Next, this set was filtered based on pharmacokinetic properties (Lipinski's rule of 5) and the 6,229 molecules that remained were then docked into the NNRTI binding pocket of the RT structure [PDB code: 3DLG (Ren et al. 2008)]. Seven compounds were retrieved and submitted for ADME prediction studies. Nearly all the structures presented acceptable values for the ADME properties analysed, such as log K$_p$, apparent Caco-2 and MDCK permeability, log BB (predicted brain/blood partition coefficient), aqueous solubility (log S), maximum of transdermal transport rate (J$_m$), human oral absorption in the gastrointestinal tract, log K$_{hsa}$ for serum protein binding and log P. No experimental results were reported in the paper. However, their approach seemed to favour high potency compounds since three of the compounds are available in the ChEMBL database with varied but high reported potency, yet none of the potency reported was against the RT.

Concluding remarks

In the last two decades, substantial advances have been made in development of novel antiretroviral drugs. The newest FDA approved drugs, ETR (2008, NNRTI), RPV (2011, NNRTI), dolutegravir (2013, integrase inhibitor) and elvitegravir (2014, integrase inhibitor) indicate recent research efforts to the current antiretroviral drug classes. However, the emergence of drug-resistance strains call for not only new classes of anti-HIV drugs with lower toxicity and favourable resistance profile, but also innovative drug discovery strategies for antiretroviral treatment. For instance, a few compounds targeting the existing classes are in advanced stages of development: TFV alafenamide fumarate is a pro-drug of TFV, currently in Phase 3 of clinical trials, which seems to have less renal and bone toxicity than its precursor (Sax et al. 2014); the NNRTI doravirine (MK-1439), currently in Phase 2, exhibits activity against resistant viral strains (Gatell et al. 2014) and an integrase inhibitor currently in Phase 2, GSK1265744 (an experimental analogue of dolutegravir), is being established in a long-acting preparation (Spreen et al. 2013). In addition, novel explored alternatives to prevention and progression, such as microbicides (Buckheit et al. 2010), antiretroviral prophylaxis (Karim & Karim 2012), CD4-mimetic compounds (Gardner et al. 2015, Richard et al. 2015) and broadly neutralising HIV-specific antibodies (Diskin et al. 2011, Moir et al. 2011, McCoy & Weiss 2013), show potential for reducing HIV-1 transmission rates. Investigational drugs such as the CD4 attachment inhibitor BMS-663068 completed Phase 1 testing (Nettles et al. 2012) and ceniciviroc, a novel CCR5/CCR2 antagonist currently in Phase 3 (Klibanov et al. 2010), that suggests both antiretroviral activity and potential for an antiinflammatory effect.

Although HIV-1 RT is an extremely validated target, which has been widely studied, the discovery of novel allosteric sites and alternative mechanisms to this enzyme provide insights to develop new therapeutic classes of inhibitors (Kang et al. 2014). Consequently, inhibitors with distinct mechanisms have been exploited and can be found in the literature, comprising, RT-directed mutagenic inducers (Smith et al. 2005), nucleotide-competing RTIs (Maga et al. 2010), RT-associated RNase H function inhibitors (Yu et al. 2008), primer/template-competing RTIs (Wang et al. 2004), dual inhibitors of the RT associated polymerase and RNase H activities (Esposito et al. 2011) and NNRTIs with not-conventionally-binding modes or alternative mechanisms (Pata et al. 2004, Cullen et al. 2009, Zhan et al. 2010, Das et al. 2011). However, the majority of these inhibitors have not been explored by means of computational methods.

The application of computational methods is of great importance to drug discovery nowadays and it is principally beneficial to investigate drug resistance development. Methods such as molecular docking, MD, free energy calculations, QSAR, pharmacophore modelling and AD-MET are broadly applied in anti-HIV drug development. The focus of this review was the HIV-1 RT; however, the approaches discussed are also in use when targeting other HIV proteins. The extensive research targeting the RT throughout the years has benefited from the employment of computational methods, extracting information from the currently available compounds and crystallographic structures to generate many successful stories in inhibitor discovery and optimisation. The computational methods employed provide beneficial results that can expand and guide the drug discovery process in all stages.

REFERENCES

Adcock SA, McCammon JA 2006. Molecular dynamics: survey of methods for simulating the activity of proteins. *Chem Rev 106*: 1589-1615.

Ahgren C, Backro K, Bell FW, Cantrell AS, Clemens M, Colacino JM, Deeter JB, Engelhardt JA, Hogberg M, Jaskunas SR, Johansson NG, Jordan CL, Kasher JS, Kinnick MD, Lind P, Lopez C, Morin JM, Muesing MA, Noreen R, Oberg B, Paget CJ, Palkowitz JA, Parrish CA, Pranc P, Rippy MK, Rydergard C, Sahlberg C, Swanson S, Ternansky RJ, Unge T, Vasileff RT, Vrang L, West SJ, Zhang H, Xhou XX 1995. The PETT series, a new class of potent nonnucleoside inhibitors of human-immunodeficiency-virus type-1 reverse-transcriptase. *Antimicrob Agents Chemother 39*: 1329-1335.

Allen WJ, Balius TE, Mukherjee S, Brozell SR, Moustakas DT, Lang PT, Case DA, Kuntz ID, Rizzo RC 2015. DOCK 6: impact of new features and current docking performance. *J Comput Chem 36*: 1132-1156.

Alonso H, Bliznyuk AA, Gready JE 2006. Combining docking and molecular dynamic simulations in drug design. *Med Res Rev 26*: 531-568.

Aqvist J, Marelius J 2001. The linear interaction energy method for predicting ligand binding free energies. *Comb Chem High Throughput Screen 4*: 613-626.

Asahchop EL, Wainberg MA, Sloan RD, Tremblay CL 2012. Antiviral drug resistance and the need for development of new HIV-1 reverse transcriptase inhibitors. *Antimicrob Agents Chemother 56*: 5000-5008.

Ashok P, Sharma H, Lathiya H, Chander S, Murugesan S 2015. In-silico design and study of novel piperazinyl beta-carbolines as inhibitor of HIV-1 reverse transcriptase. *Med Chem Res 24*: 513-522.

Asres K, Seyoum A, Veeresham C, Bucar F, Gibbons S 2005. Naturally derived anti-HIV agents. *Phytother Res 19*: 557-581.

Bahare RS, Ganguly S 2014. Comparative molecular docking studies of novel 3, 5-disubstituted thiazolidinedione analogs as HIV-1-RT inhibitors. *Med Chem Res 23*: 1300-1308.

Balaji S, Karthikeyan C, Moorthy NSHN, Trivedi P 2004. QSAR modelling of HIV-1 reverse transcriptase inhibition by benzoxazinones using a combination of P_VSA and pharmacophore feature descriptors. *Bioorg Med Chem Lett 14*: 6089-6094.

Barreiro G, Guimarães CR, Tubert-Brohman I, Lyons TM, Tirado-Rives J, Jorgensen WL 2007a. Search for non-nucleoside inhibitors of HIV-1 reverse transcriptase using chemical similarity, molecular docking and MM-GB/SA scoring. *J Chem Inf Model 47*: 2416-2428.

Barreiro G, Kim JT, Guimarães CR, Bailey CM, Domaoal RA, Wang L, Anderson KS, Jorgensen WL 2007b. From docking false-positive to active anti-HIV agent. *J Med Chem 50*: 5324-5329.

Barre-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, Dauguet C, Axler-Blin C, Vezinet-Brun F, Rouzioux C, Rozenbaum W, Montagnier L 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science 220*: 868-871.

Bento AP, Gaulton A, Hersey A, Bellis LJ, Chambers J, Davies M, Krüger FA, Light Y, Mak L, McGlinchey S 2014. The ChEMBL bioactivity database: an update. *Nucleic Acids Res 42*: D1083-D1090.

Beresford AP, Segall M, Tarbit MH 2004. In silico prediction of ADME properties: are we making progress? *Curr Opin Drug Discov Devel 7*: 36-42.

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE 2000. The Protein Data Bank. *Nucleic Acids Res 28*: 235-242.

Bollini M, Domaoal RA, Thakur VV, Gallardo-Macias R, Spasov KA, Anderson KS, Jorgensen WL 2011. Computationally-guided optimization of a docking hit to yield catechol diethers as potent anti-HIV agents. *J Med Chem 54*: 8582-8591.

B-Rao C, Subramanian J, Sharma SD 2009. Managing protein flexibility in docking and its applications. *Drug Discov Today 14*: 394-400.

Broder S 2010. The development of antiretroviral therapy and its impact on the HIV-1/AIDS pandemic. *Antiviral Res 85*: 1-18.

Buckheit RW, Watson KM, Morrow KM, Ham AS 2010. Development of topical microbicides to prevent the sexual transmission of HIV. *Antiviral Res 85*: 142-158.

Cao L, Song W, De Clercq E, Zhan P, Liu X 2014. Recent progress in the research of small molecule HIV-1 RNase H inhibitors. *Curr Med Chem 21*: 1956-1967.

Carlson HA, McCammon JA 2000. Accommodating protein flexibility in computational drug design. *Mol Pharmacol 57*: 213-218.

Cavasotto CN, Orry AJW 2007. Ligand docking and structure-based virtual screening in drug discovery. *Curr Top Med Chem 7*: 1006-1014.

Cavasotto N, Singh CN 2008. Docking and high throughput docking: successes and the challenge of protein flexibility. *Curr Comput Aided Drug Des 4*: 221-234.

Chander S, Penta A, Murugesan S 2015. Structure-based virtual screening and docking studies for the identification of novel inhibitors against wild and drug resistance strains of HIV-1 RT. *Med Chem Res 24*: 1869-1883.

Chin YW, Balunas MJ, Chai HB, Kinghorn AD 2006. Drug discovery from natural sources. *AAPS J 8*: E239-E253.

Cihlar T, Ray AS 2010. Nucleoside and nucleotide HIV reverse transcriptase inhibitors: 25 years after zidovudine. *Antiviral Res 85*: 39-58.

Claussen H, Buning C, Rarey M, Lengauer T 2001. FlexE: efficient molecular docking considering protein structure variations. *J Mol Biol 308*: 377-395.

Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA 1995. A 2nd generation force-field for the simulation of proteins, nucleic-acids and organic-molecules. *J Am Chem Soc 117*: 5179-5197.

Cullen MD, Ho WC, Bauman JD, Das K, Arnold E, Hartman TL, Watson KM, Buckheit RW, Pannecouque C, De Clercq E, Cushman M 2009. Crystallographic study of a novel subnanomolar inhibitor provides insight on the binding interactions of alkenyl-diarylmethanes with human immunodeficiency virus-1 reverse transcriptase. *J Med Chem 52*: 6467-6473.

Cummings MD, Des Jarlais RL, Gibbs AC, Mohan V, Jaeger EP 2005. Comparison of automated docking programs as virtual screening tools. *J Med Chem 48*: 962-976.

Currens MJ, Gulakowski RJ, Mariner JM, Moran RA, Buckheit RW, Gustafson KR, McMahon JB, Boyd MR 1996. Antiviral activity and mechanism of action of calanolide A against the human immunodeficiency virus type-1. *J Pharmacol Exp Ther 279*: 645-651.

Das K, Bauman JD, Clark AD, Frenkel YV, Lewi PJ, Shatkin AJ, Hughes SH, Arnold E 2008. High-resolution structures of HIV-1 reverse transcriptase/TMC278 complexes: strategic flexibility explains potency against resistance mutations. *Proc Natl Acad Sci USA 105*: 1466-1471.

Das K, Bauman JD, Rim AS, Dharia C, Clark AD, Camarasa MJ, Balzarini J, Arnold E 2011. Crystal structure of tert-butyldimethylsilyl-spiroaminooxathioledioxide-thymine (TSAO-T) in complex with HIV-1 reverse transcriptase (RT) redefines the elastic limits of the non-nucleoside inhibitor-binding pocket. *J Med Chem 54*: 2727-2737.

Das K, Martinez SE, Bauman JD, Arnold E 2012. HIV-1 reverse transcriptase complex with DNA and nevirapine reveals non-nucleoside inhibition mechanism. *Nat Struct Mol Biol 19*: 253-259.

de Bethune MP 2010. Non-nucleoside reverse transcriptase inhibitors (NNRTIs), their discovery, development and use in the treatment of HIV-1 infection: a review of the last 20 years (1989-2009). *Antiviral Res 85*: 75-90.

De Clercq E 2009. Another ten stories in antiviral drug discovery (Part C): "old" and "new" antivirals, strategies and perspectives. *Med Res Rev 29*: 611-645.

De Clercq E 2010. Antiretroviral drugs. *Curr Opin Pharmacol 10*: 507-515.

Deng YQ, Roux B 2009. Computations of standard binding free energies with molecular dynamics simulations. *J Phys Chem B 113*: 2234-2246.

Ding J, Das K, Hsiou Y, Sarafianos SG, Clark AD, Jacobo-Molina A, Tantillo C, Hughes SH, Arnold E 1998. Structure and functional implications of the polymerase active site region in a complex of HIV-1 RT with a double-stranded DNA template-primer and an antibody Fab fragment at 2.8 A resolution. *J Mol Biol 284*: 1095-1111.

Ding J, Das K, Moereels H, Koymans L, Andries K, Janssen PA, Hughes SH, Arnold E 1995. Structure of HIV-1 RT/TIBO R 86183 complex reveals similarity in the binding of diverse non-nucleoside inhibitors. *Nat Struct Biol 2*: 407-415.

Diskin R, Scheid JF, Marcovecchio PM, West Jr AP, Klein F, Gao H, Gnanapragasam PNP, Abadir A, Seaman MS, Nussenzweig MC, Bjorkman PJ 2011. Increasing the potency and breadth of an HIV antibody by using structure-based rational design. *Science 334*: 1289-1293.

Distinto S, Esposito F, Kirchmair J, Cardia MC, Gaspari M, Maccioni E, Alcaro S, Markt P, Wolber G, Zinzula L, Tramontano E 2012. Identification of HIV-1 reverse transcriptase dual inhibitors by a combined shape, 2D-fingerprint and pharmacophore-based virtual screening approach. *Eur J Med Chem 50*: 216-229.

Distinto S, Maccioni E, Meleddu R, Corona A, Alcaro S, Tramontano E 2013. Molecular aspects of the RT/drug interactions. Perspective of dual inhibitors. *Curr Pharm Des 19*: 1850-1859.

Dong M, Ren Y 2015. Molecular modeling studies of dihydro-alkyloxy-benzyl-oxopyrimidines (DABOs) as non-nucleoside inhibitors of HIV-1 reverse transcriptase using 3D-QSAR, Topomer CoMFA and molecular docking simulations. *RSC Adv 5*: 13754-13761.

Dudek AZ, Arodz T, Galvez J 2006. Computational methods in developing quantitative structure-activity relationships (QSAR): a review. *Comb Chem High Throughput Screen 9*: 213-228.

Duffy EM, Jorgensen WL 2000. Prediction of properties from simulations: free energies of solvation in hexadecane, octanol and water. *J Am Chem Soc 122*: 2878-2888.

Ehrman TM, Barlow DJ, Hylands PJ 2007. Phytochemical databases of Chinese herbal constituents and bioactive plant compounds with known target specificities. *J Chem Inf Model 47*: 254-263.

Elokely KM, Doerksen RJ 2013. Docking challenge: protein sampling and molecular docking performance. *J Chem Inf Model 53*: 1934-1945.

Esnouf RM, Ren J, Hopkins AL, Ross CK, Jones EY, Stammers DK, Stuart DI 1997. Unique features in the structure of the complex between HIV-1 reverse transcriptase and the bis(heteroaryl)piperazine (BHAP) U-90152 explain resistance mutations for this nonnucleoside inhibitor. *Proc Natl Acad Sci USA 94*: 3984-3989.

Esposito F, Corona A, Tramontano E 2012. HIV-1 reverse transcriptase still remains a new drug target: structure, function, classical inhibitors and new inhibitors with innovative mechanisms of actions. *Mol Biol Int 2012*: 586401.

Esposito F, Kharlamova T, Distinto S, Zinzula L, Cheng YC, Dutschman G, Floris G, Markt P, Corona A, Tramontano E 2011. Alizarine derivatives as new dual inhibitors of the HIV-1 reverse transcriptase-associated DNA polymerase and RNase H activities effective also on the RNase H activity of non-nucleoside resistant reverse transcriptases. *FEBS J 278*: 1444-1457.

Ewing TJ, Makino S, Skillman AG, Kuntz ID 2001. DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. *J Comput Aided Mol Des 15*: 411-428.

FDA - United States Food and Drug Administration 2014. Antiretroviral drugs used in the treatment of HIV infection. Available from: fda.gov/forpatients/illness/hivaids/ucm118915.htm.

Filardi PP, Paolillo S, Marciano C, Iorio A, Losco T, Marsico F, Scala O, Ruggiero D, Ferraro S, Chiariello M 2008. Cardiovascular effects of antiretroviral drugs: clinical review. *Cardiovasc Hematol Disord Drug Targets 8*: 238-244.

Fraczek T, Siwek A, Paneth P 2013. Assessing molecular focking tools for relative biological activity prediction: a case study of triazole HIV-1 NNRTIs. *J Chem Inf Model 53*: 3326-3342.

Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS 2004. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem 47*: 1739-1749.

Galeazzi R 2009. Molecular dynamics as a tool in rational drug design: current status and some major applications. *Curr Comput Aided Drug Des 5*: 225-240.

Gardner MR, Kattenhorn LM, Kondur HR, von Schaewen M, Dorfman T, Chiang JJ, Haworth KG, Decker JM, Alpert MD, Bailey CC, Neale Jr ES, Fellinger CH, Joshi VR, Fuchs SP, Martinez-Navio JM, Quinlan BD, Yao AY, Mouquet H, Gorman J, Zhang B, Poignard P, Nussenzweig MC, Burton DR, Kwong PD, Piatak Jr M, Lifson JD, Gao G, Desrosiers RC, Evans DT, Hahn BH, Ploss A, Cannon PM, Seaman MS, Farzan M 2015. AAV-expressed eCD4-Ig provides durable protection from multiple SHIV challenges. *Nature 519*: 87-91.

Gatell JM, Morales-Ramirez JO, Hagins DP, Thompson M, Keikawus A, Hoffmann C, Rugina S, Osiyemi O, Escoriu S, Dretler R 2014. Forty-eight-week efficacy and safety and early CNS tolerability of doravirine (MK-1439), a novel NNRTI, with TDF/FTC in ART-naive HIV-positive patients. *J Int AIDS Soc 17* (Suppl. 3): 19532.

Gaudio AC, Montanari CA 2002. HEPT derivatives as non-nucleoside inhibitors of HIV-1 reverse transcriptase: QSAR studies agree with the crystal structures. *J Comput Aided Mol Des 16*: 287-295.

Gayen S, Debnath B, Samanta S, Jha T 2004. QSAR study on some anti-HIV HEPT analogues using physicochemical and topological parameters. *Bioorg Med Chem 12*: 1493-1503.

Ghosh AK, Chapsal BD, Parham GL, Steffey M, Agniswamy J, Wang YF, Amano M, Weber IT, Mitsuya H 2011. Design of HIV-1 protease inhibitors with C3-substituted hexahydrocyclopentafuranyl urethanes as P2-ligands: synthesis, biological evaluation and protein-ligand X-ray crystal structure. *J Med Chem 54*: 5890-5901.

Ghosh AK, Chapsal BD, Weber IT, Mitsuya H 2008. Design of HIV protease inhibitors targeting protein backbone: an effective strategy for combating drug resistance. *Acc Chem Res 41*: 78-86.

Gleeson MP, Hersey A, Hannongbua S 2011. In-silico ADME models: a general assessment of their utility in drug discovery applications. *Curr Top Med Chem 11*: 358-381.

Gomez R, Jolly S, Williams T, Tucker T, Tynebor R, Vacca J, McGaughey G, Lai MT, Felock P, Munshi V, DeStefano D, Touch S, Miller M, Yan Y, Sanchez R, Liang Y, Paton B, Wan BL, Anthony N 2011. Design and synthesis of pyridone inhibitors of non-nucleoside reverse transcriptase. *Bioorg Med Chem Lett 21*: 7344-7350.

Goodsell DS, Morris GM, Olson AJ 1996. Automated docking of flexible ligands: applications of AutoDock. *J Mol Recognit 9*: 1-5.

Gorfe AA, Grant BJ, McCammon JA 2008. Mapping the nucleotide and isoform-dependent structural and dynamical features of Ras proteins. *Structure 16*: 885-896.

Guimarães MC, Silva DG, da Mota EG, da Cunha EF, Freitas MP 2014. Computer-assisted design of dual-target anti-HIV-1 compounds. *Med Chem Res 23*: 1548-1558.

Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, Banks JL 2004. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J Med Chem 47*: 1750-1759.

Hansen N, van Gunsteren WF 2014. Practical aspects of free-energy calculations: a review. *J Chem Theory Comput 10*: 2632-2647.

He X, Mei Y, Xiang Y, Zhang DW, Zhang JZH 2005. Quantum computational analysis for drug resistance of HIV-1 reverse transcriptase to nevirapine through point mutations. *Proteins 61*: 423-432.

Herschhorn A, Hizi A 2008. Virtual screening, identification and biochemical characterization of novel inhibitors of the reverse transcriptase of human immunodeficiency virus type-1. *J Med Chem 51*: 5702-5713.

Ho HT, Hitchcock MJ 1989. Cellular pharmacology of 2',3'-dideoxy-2',3'-didehydrothymidine, a nucleoside analog active against human immunodeficiency virus. *Antimicrob Agents Chemother 33*: 844-849.

Hopkins AL, Ren J, Esnouf RM, Willcox BE, Jones EY, Ross C, Miyasaka T, Walker RT, Tanaka H, Stammers DK, Stuart DI 1996. Complexes of HIV-1 reverse transcriptase with inhibitors of the HEPT series reveal conformational changes relevant to the design of potent non-nucleoside inhibitors. *J Med Chem 39*: 1589-1600.

Hsiou Y, Das K, Ding J, Clark AD, Kleim JP, Rosner M, Winkler I, Riess G, Hughes SH, Arnold E 1998. Structures of Tyr188Leu mutant and wild-type HIV-1 reverse transcriptase complexed with the non-nucleoside inhibitor HBY 097: inhibitor flexibility is a useful design feature for reducing drug resistance. *J Mol Biol 284*: 313-323.

Hsiou Y, Ding J, Das K, Clark AD, Hughes SH, Arnold E 1996. Structure of unliganded HIV-1 reverse transcriptase at 2.7 A resolution: implications of conformational changes for polymerization and inhibition mechanisms. *Structure 4*: 853-860.

Huang SY, Zou XQ 2007. Ensemble docking of multiple protein structures: considering protein structural variations in molecular docking. *Proteins 66*: 399-421.

Irwin JJ, Shoichet BK 2005. ZINC - a free database of commercially available compounds for virtual screening. *J Chem Inf Model 45*: 177-182.

Ivetac A, McCammon JA 2009. Elucidating the inhibition mechanism of HIV-1 non-nucleoside reverse transcriptase inhibitors through multicopy molecular dynamics simulations. *J Mol Biol 388*: 644-658.

Ivetac A, McCammon JA 2011. Molecular recognition in the case of flexible targets. *Curr Pharm Des 17*: 1663-1671.

Ivetac A, Swift SE, Boyer PL, Diaz A, Naughton J, Young JA, Hughes SH, McCammon JA 2014. Discovery of novel inhibitors of HIV-1 reverse transcriptase through virtual screening of experimental and theoretical ensembles. *Chem Biol Drug Des 83*: 521-531.

Jacobson MP, Friesner RA, Xiang Z, Honig B 2002. On the role of the crystal environment in determining protein side-chain conformations. *J Mol Biol 320*: 597-608.

Jacobson MP, Pincus DL, Rapp CS, Day TJ, Honig B, Shaw DE, Friesner RA 2004. A hierarchical approach to all-atom protein loop prediction. *Proteins 55*: 351-367.

Jain AN 2003. Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. *J Med Chem 46*: 499-511.

Jain AN 2007. Surflex-Dock 2.1: robust performance from ligand energetic modeling, ring flexibility and knowledge-based search. *J Comput Aided Mol Des 21*: 281-306.

Janssen PA, Lewi PJ, Arnold E, Daeyaert F, de Jonge M, Heeres J, Koymans L, Vinkers M, Guillemont J, Pasquier E, Kukla M, Ludovici D, Andries K, de Bethune MP, Pauwels R, Das K, Clark Jr AD, Frenkel YV, Hughes SH, Medaer B, De Knaep F, Bohets H, De Clerck F, Lampo A, Williams P, Stoffels P 2005. In search of a novel anti-HIV drug: multidisciplinary coordination in the discovery of 4-[[4-[[4-[(1E)-2-cyanoethenyl]-2,6-dimethylphenyl] amino]-2- pyrimidinyl]amino]benzonitrile (R278474, rilpivirine). *J Med Chem 48*: 1901-1909.

Jiang F, Kim SH 1991. "Soft docking": matching of molecular surface cubes. *J Mol Biol 219*: 79-102.

Johnson BC, Pauly GT, Rai G, Patel D, Bauman JD, Baker HL, Das K, Schneider JP, Maloney DJ, Arnold E, Thomas CJ, Hughes SH 2012. A comparison of the ability of rilpivirine (TMC278) and selected analogues to inhibit clinically relevant HIV-1 reverse transcriptase mutants. *Retrovirology 9*: 99.

Johnson JA, Li JF, Wei X, Lipscomb J, Irlbeck D, Craig C, Smith A, Bennett DE, Monsour M, Sandstrom P, Lanier ER, Heneine W 2008. Minority HIV-1 drug resistance mutations are present in antiretroviral treatment-naïve populations and associate with reduced treatment efficacy. *PLoS Med 5*: e158.

Jorgensen WL 2004. The many roles of computation in drug discovery. *Science 303*: 1813-1818.

Jorgensen WL, Maxwell DS, Tirado-Rives J 1996. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc 118*: 11225-11236.

Kale L, Skeel R, Bhandarkar M, Brunner R, Gursoy A, Krawetz N, Phillips J, Shinozaki A, Varadarajan K, Schulten K 1999. NAMD2: greater scalability for parallel molecular dynamics. *J Comput Phys 151*: 283-312.

Kang D, Song Y, Chen W, Zhan P, Liu X 2014. "Old Dogs with New Tricks": exploiting alternative mechanisms of action and new drug design strategies for clinically validated HIV targets. *Mol Biosyst 10*: 1998-2022.

Karim SS, Karim QA 2012. Antiretroviral prophylaxis for HIV prevention reaches a key milestone. *Lancet 379*: 2047-2048.

Keller PA, Birch C, Leach SP, Tyssen D, Griffith R 2003. Novel pharmacophore-based methods reveal gossypol as a reverse transcriptase inhibitor. *J Mol Graph Model 21*: 365-373.

Kirchmair J, Distinto S, Liedl KR, Markt P, Rollinger JM, Schuster D, Spitzer GM, Wolber G 2011. Development of anti-viral agents using molecular modeling and virtual screening techniques. *Infect Disord Drug Targets 11*: 64-93.

Kirschberg TA, Balakrishnan M, Huang W, Hluhanich R, Kutty N, Liclican AC, McColl DJ, Squires NH, Lansdon EB 2008. Triazole derivatives as non-nucleoside inhibitors of HIV-1 reverse transcriptase - structure-activity relationships and crystallographic analysis. *Bioorg Med Chem Lett 18*: 1131-1134.

Klibanov OM, Williams SH, Iler CA 2010. Cenicriviroc, an orally active CCR5 antagonist for the potential treatment of HIV infection. *Curr Opin Investig Drugs 11*: 940-950.

Knegtel RM, Kuntz ID, Oshiro CM 1997. Molecular docking to ensembles of protein structures. *J Mol Biol 266*: 424-440.

Kohlstaedt LA, Wang J, Friedman JM, Rice PA, Steitz TA 1992. Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science 256*: 1783-1790.

Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA, Cheatham TE 2000. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc Chem Res 33*: 889-897.

Kramer B, Rarey M, Lengauer T 1999. Evaluation of the FLEXX incremental construction algorithm for protein-ligand docking. *Proteins 37*: 228-241.

Kroemer RT 2007. Structure-based drug design: docking and scoring. *Curr Protein Pept Sci 8*: 312-328.

Kubinyi H 2003. Drug research: myths, hype and reality. *Nat Rev Drug Discov 2*: 665-668.

Kubinyi H, Sadowski J 1999. Qsar, 3d Qsar and Beyond. *Abstr Pap Am Chem Soc 217*: U667-U667.

Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE 1982. A geometric approach to macromolecule-ligand interactions. *J Mol Biol 161*: 269-288.

Kuroda DG, Bauman JD, Challa JR, Patel D, Troxler T, Das K, Arnold E, Hochstrasser RM 2013. Snapshot of the equilibrium dynamics of a drug bound to HIV-1 reverse transcriptase. *Nat Chem 5*: 174-181.

Lagorce D, Sperandio O, Galons H, Miteva MA, Villoutreix BO 2008. FAF-Drugs2: free ADME/tox filtering tool to assist drug discovery and chemical biology projects. *BMC Bioinformatics 9*: 396.

Lahti JL, Tang GW, Capriotti E, Liu TY, Altman RB 2012. Bioinformatics and variability in drug response: a protein structural perspective. *J R Soc Interface 9*: 1409-1437.

Lee WG, Gallardo-Macias R, Frey KM, Spasov KA, Bollini M, Anderson KS, Jorgensen WL 2013. Picomolar inhibitors of HIV reverse transcriptase featuring bicyclic replacement of a cyanovinylphenyl group. *J Am Chem Soc 135*: 16705-16713.

Leung CS, Zeevaart JG, Domaoal RA, Bollini M, Thakur VV, Spasov KA, Anderson KS, Jorgensen WL 2010. Eastern extension of azoles as non-nucleoside inhibitors of HIV-1 reverse transcriptase; cyano group alternatives. *Bioorg Med Chem Lett 20*: 2485-2488.

Li Z, Han J, Chen HF 2008. Revealing interaction mode between HIV-1 reverse transcriptase and diaryltriazine analog inhibitor. *Chem Biol Drug Des 72*: 350-359.

Lipinski CA, Lombardo F, Dominy BW, Feeney PJ 2001. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev 46*: 3-26.

Liu J, He X, Zhang JZ 2014. Novel theoretically designed HIV-1 nonnucleoside reverse transcriptase inhibitors derived from nevirapine. *J Mol Model 20*: 1-11.

Liu S, Abbondanzieri EA, Rausch JW, Le Grice SFJ, Zhuang X 2008. Slide into action: dynamic shuttling of HIV reverse transcriptase on nucleic acid substrates. *Science 322*: 1092-1097.

Lu N, Kofke DA, Woolf TB 2004. Improving the efficiency and reliability of free energy perturbation calculations using overlap sampling methods. *J Comput Chem 25*: 28-39.

Lu X, Liu L, Zhang X, Lau TC, Tsui SK, Kang Y, Zheng P, Zheng B, Liu G, Chen Z 2012. F18, a novel small-molecule nonnucleoside reverse transcriptase inhibitor, inhibits HIV-1 replication using distinct binding motifs as demonstrated by resistance selection and docking analysis. *Antimicrob Agents Chemother 56*: 341-351.

MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D, Karplus M 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B 102*: 3586-3616.

Madrid M, Lukin JA, Madura JD, Ding J, Arnold E 2001. Molecular dynamics of HIV-1 reverse transcriptase indicates increased flexibility upon DNA binding. *Proteins 45*: 176-182.

Maga G, Radi M, Gerard MA, Botta M, Ennifar E 2010. HIV-1 RT inhibitors with a novel mechanism of action: NNRTIs that compete with the nucleotide substrate. *Viruses 2*: 880-899.

McCoy LE, Weiss RA 2013. Neutralizing antibodies to HIV-1 induced by immunization. *J Exp Med 210*: 209-223.

Mehellou Y, De Clercq E 2010. Twenty-six years of anti-HIV drug discovery: where do we stand and where do we go? *J Med Chem 53*: 521-538.

Meleddu R, Cannas V, Distinto S, Sarais G, Del Vecchio C, Esposito F, Bianco G, Corona A, Cottiglia F, Alcaro S, Parolin C, Artese A, Scalise D, Fresta M, Arridu A, Ortuso F, Maccioni E, Tramontano E 2014. Design, synthesis and biological evaluation of 1,3-diarylpropenones as dual inhibitors of HIV-1 reverse transcriptase. *ChemMedChem 9*: 1869-1879.

Menendez-Arias L 2013. Molecular basis of human immunodeficiency virus type 1 drug resistance: overview and recent developments. *Antiviral Res 98*: 93-120.

Meng XY, Zhang HX, Mezei M, Cui M 2011. Molecular docking: a powerful approach for structure-based drug discovery. *Curr Comput Aided Drug Des 7*: 146-157.

Michailidis E, Huber AD, Ryan EM, Ong YT, Leslie MD, Matzek KB, Singh K, Marchand B, Hagedorn AN, Kirby KA, Rohan LC, Kodama EN, Mitsuya H, Parniak MA, Sarafianos SG 2014. 4'-Ethynyl-2-fluoro-2'-deoxyadenosine (EFdA) inhibits HIV-1 reverse transcriptase with multiple mechanisms. *J Biol Chem 289*: 24533-24548.

Michel J, Foloppe N, Essex JW 2010. Rigorous free energy calculations in structure-based drug design. *Mol Inform 29*: 570-578.

Moir S, Malaspina A, Fauci AS 2011. Prospects for an HIV vaccine: leading B cells down the right path. *Nat Struct Mol Biol 18*: 1317-1321.

Moitessier N, Englebienne P, Lee D, Lawandi J, Corbeil CR 2008. Towards the development of universal, fast and highly accurate docking/scoring methods: a long way to go. *Br J Pharmacol 153* (Suppl. 1): 7-26.

Mukherjee S, Balius TE, Rizzo RC 2010. Docking validation resources: protein family and ligand flexibility experiments. *J Chem Inf Model 50*: 1986-2000.

Nazar MF, Abdullah MI, Badshah A, Mahmood A, Rana UA, Khan SU-D 2015. Synthesis, structure-activity relationship and molecular docking of cyclohexenone based analogous as potent non-nucleoside reverse-transcriptase inhibitors. *J Mol Struct 1086*: 8-16.

Nettles RE, Schürmann D, Zhu L, Stonier M, Huang S-P, Chang I, Chien C, Krystal M, Wind-Rotolo M, Ray N 2012. Pharmacodynamics, safety and pharmacokinetics of BMS-663068, an oral HIV-1 attachment inhibitor in HIV-1-infected subjects. *J Infect Dis 206*: 1002-1011.

Newman DJ, Cragg GM 2012. Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J Nat Prod 75*: 311-335.

Nichols SE, Baron R, Ivetac A, McCammon JA 2011. Predictive power of molecular dynamics receptor structures in virtual screening. *J Chem Inf Model 51*: 1439-1446.

Nichols SE, Domaoal RA, Thakur VV, Tirado-Rives J, Anderson KS, Jorgensen WL 2009. Discovery of wild-type and Y181C mutant non-nucleoside HIV-1 reverse transcriptase inhibitors using virtual screening with multiple protein structures. *J Chem Inf Model 49*: 1272-1279.

Oostenbrink C, Villa A, Mark AE, van Gunsteren WF 2004. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J Comput Chem 25*: 1656-1676.

Oprea TI, Davis AM, Teague SJ, Leeson PD 2001. Is there a difference between leads and drugs? A historical perspective. *J Chem Inf Comput Sci 41*: 1308-1315.

Osterberg F, Morris GM, Sanner MF, Olson AJ, Goodsell DS 2002. Automated docking to multiple target structures: incorporation of protein mobility and structural water heterogeneity in AutoDock. *Proteins 46*: 34-40.

Pata JD, Stirtan WG, Goldstein SW, Steitz TA 2004. Structure of HIV-1 reverse transcriptase bound to an inhibitor active against mutant reverse transcriptases resistant to other nonnucleoside inhibitors. *Proc Natl Acad Sci USA 101*: 10548-10553.

Pearlman DA 1994. A comparison of alternative approaches to free-energy calculations. *J Phys Chem 98*: 1487-1493.

Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE, DeBolt S, Ferguson D, Seibel G, Kollman P 1995. AMBER, a package of computer-programs for applying molecular mechanics, normal-mode analysis, molecular-dynamics and free-energy calculations to simulate the structural and energetic properties of molecules. *Comput Phys Commun 91*: 1-41.

Pirhadi S, Ghasemi JB 2012. Pharmacophore identification, molecular docking, virtual screening and in silico ADME studies of non-nucleoside reverse transcriptase inhibitors. *Mol Inf 31*: 856-866.

Pohorille A, Jarzynski C, Chipot C 2010. Good practices in free-energy calculations. *J Phys Chem B 114*: 10235-10253.

Quashie PK, Sloan RD, Wainberg MA 2012. Novel therapeutic strategies targeting HIV integrase. *BMC Med 10*: 34.

Ragno R, Frasca S, Manetti F, Brizzi A, Massa S 2005. HIV-reverse transcriptase inhibition: inclusion of ligand-induced fit by cross-docking studies. *J Med Chem 48*: 200-212.

Reeves JD, Doms RW 2002. Human immunodeficiency virus type 2. *J Gen Virol 83*: 1253-1265.

Ren J, Chamberlain PP, Stamp A, Short SA, Weaver KL, Romines KR, Hazen R, Freeman A, Ferris RG, Andrews CW, Boone L, Chan JH, Stammers DK 2008. Structural basis for the improved drug resistance profile of new generation benzophenone non-nucleoside HIV-1 reverse transcriptase inhibitors. *J Med Chem 51*: 5000-5008.

Ren J, Diprose J, Warren J, Esnouf RM, Bird LE, Ikemizu S, Slater M, Milton J, Balzarini J, Stuart DI, Stammers DK 2000a. Phenylethylthiazolylthiourea (PETT) non-nucleoside inhibitors of HIV-1 and HIV-2 reverse transcriptases. Structural and biochemical analyses. *J Biol Chem 275*: 5633-5639.

Ren J, Esnouf R, Garman E, Somers D, Ross C, Kirby I, Keeling J, Darby G, Jones Y, Stuart D 1995. High resolution structures of HIV-1 RT from four RT-inhibitor complexes. *Nat Struct Biol 2*: 293-302.

Ren J, Esnouf RM, Hopkins AL, Warren J, Balzarini J, Stuart DI, Stammers DK 1998. Crystal structures of HIV-1 reverse transcriptase in complex with carboxanilide derivatives. *Biochemistry 37*: 14394-14403.

Ren J, Milton J, Weaver KL, Short SA, Stuart DI, Stammers DK 2000b. Structural basis for the resilience of efavirenz (DMP-266) to drug resistance mutations in HIV-1 reverse transcriptase. *Structure 8*: 1089-1094.

Ren J, Nichols C, Bird L, Chamberlain P, Weaver K, Short S, Stuart DI, Stammers DK 2001. Structural mechanisms of drug resistance for mutations at codons 181 and 188 in HIV-1 reverse transcriptase and the improved resilience of second generation non-nucleoside inhibitors. *J Mol Biol 312*: 795-805.

Ren J, Nichols C, Bird LE, Fujiwara T, Sugimoto H, Stuart DI, Stammers DK 2000c. Binding of the second generation non-nucleoside inhibitor S-1153 to HIV-1 reverse transcriptase involves extensive main chain hydrogen bonding. *J Biol Chem 275*: 14316-14320.

Ren J, Nichols CE, Chamberlain PP, Weaver KL, Short SA, Stammers DK 2004. Crystal structures of HIV-1 reverse transcriptases mutated at codons 100, 106 and 108 and mechanisms of resistance to non-nucleoside inhibitors. *J Mol Biol 336*: 569-578.

Richard J, Veillette M, Brassard N, Iyer SS, Roger M, Martin L, Pazgier M, Schon A, Freire E, Routy JP, Smith AB, Park J, Jones DM, Courter JR, Melillo BN, Kaufmann DE, Hahn BH, Permar SR, Haynes BF, Madani N, Sodroski JG, Finzi A 2015. CD4 mimetics sensitize HIV-1-infected cells to ADCC. *Proc Natl Acad Sci USA 112*: E2687-E2694.

Rollinger JM, Langer T, Stuppner H 2006a. Integrated in silico tools for exploiting the natural products' bioactivity. *Planta Med 72*: 671-678.

Rollinger JM, Langer T, Stuppner H 2006b. Strategies for efficient lead structure discovery from natural products. *Curr Med Chem 13*: 1491-1507.

Rollinger JM, Stuppner H, Langer T 2008. Virtual screening for the discovery of bioactive natural products. *Prog Drug Res 65*: 211-249.

Sangma C, Chuakheaw D, Jongkon N, Saenbandit K, Nunrium P, Uthayopas P, Hannongbua S 2005. Virtual screening for anti-HIV-1 RT and anti-HIV-1 PR inhibitors from the Thai medicinal plants database: a combined docking with neural networks approach. *Comb Chem High Throughput Screen 8*: 417-429.

Sarafianos SG, Das K, Hughes SH, Arnold E 2004. Taking aim at a moving target: designing drugs to inhibit drug-resistant HIV-1 reverse transcriptases. *Curr Opin Struct Biol 14*: 716-730.

Sax PE, Zolopa A, Brar I, Elion R, Ortiz R, Post F, Wang H, Callebaut C, Martin H, Fordyce MW, McCallister S 2014. Tenofovir alafenamide vs. tenofovir disoproxil fumarate in single tablet regimens for initial HIV-1 therapy: a randomized Phase 2 study. *J Acquir Immune Defic Syndr 67*: 52-58.

Scarth BJ, Ehteshami M, Beilhartz GL, Gotte M 2011. HIV-1 reverse transcriptase inhibitors: beyond classic nucleosides and non-nucleosides. *Future Virol 6*: 581-598.

Seal A, Aykkal R, Babu RO, Ghosh M 2011. Docking study of HIV-1 reverse transcriptase with phytochemicals. *Bioinformation 5*: 430-439.

Selick HE, Beresford AP, Tarbit MH 2002. The emerging importance of predictive ADME simulation in drug discovery. *Drug Discov Today 7*: 109-116.

Sengupta D, Verma D, Naik PK 2007. Docking mode of delvardine and its analogues into the p66 domain of HIV-1 reverse transcriptase: screening using molecular mechanics-generalized born/surface area and absorption, distribution, metabolism and excretion properties. *J Biosci 32*: 1307-1316.

Sherman W, Day T, Jacobson MP, Friesner RA, Farid R 2006. Novel procedure for modeling ligand/receptor induced fit effects. *J Med Chem 49*: 534-553.

Shoichet BK, McGovern SL, Wei BQ, Irwin JJ 2002. Lead discovery using molecular docking. *Curr Opin Chem Biol 6*: 439-446.

Silverberg MJ, Leyden W, Hurley L, Go AS, Quesenberry CP, Klein D, Horberg MA 2009. Response to newly prescribed lipid-lowering therapy in patients with and without HIV infection. *Ann Intern Med 150*: 301-313.

Sliwoski G, Kothiwale S, Meiler J, Lowe EW 2014. Computational methods in drug discovery. *Pharmacol Rev 66*: 334-395.

Sluis-Cremer N, Temiz NA, Bahar I 2004. Conformational changes in HIV-1 reverse transcriptase induced by nonnucleoside reverse transcriptase inhibitor binding. *Curr HIV Res 2*: 323-332.

Smith MBK, Hughes SH, Boyer PL, Michejda CJ, Rouzer CA, Taneyhill LA, Smith RH, Smith NA, Janssen PA, Moereels H 1995. Molecular modeling studies of HIV-1 reverse transcriptase non-nucleoside inhibitors: total energy of complexation as a predictor of drug placement and activity. *Protein Sci 4*: 2203-2222.

Smith RA, Loeb LA, Preston BD 2005. Lethal mutagenesis of HIV. *Virus Res 107*: 215-228.

Spallarossa A, Cesarini S, Ranise A, Ponassi M, Unge T, Bolognesi M 2008. Crystal structures of HIV-1 reverse transcriptase complexes with thiocarbamate non-nucleoside inhibitors. *Biochem Biophys Res Commun 365*: 764-770.

Spitzer R, Jain AN 2012. Surflex-Dock: docking benchmarks and real-world application. *J Comput Aided Mol Des 26*: 687-699.

Spreen W, Williams P, Margolis D, Ford S, Crauwels H, Lou Y, Gould E, Stevens M, Piscitelli S 2013. First study of repeat dose co-administration of GSK1265744 and TMC278 long-acting parenteral nanosuspensions: pharmacokinetics, safety and tolerability in healthy adults. 7th International AIDS Society Conference on HIV Pathogenesis, Treatment and Prevention, 30 June-3 July 2013, Kuala Lumpur, Malaysia, p. 2-12.

Squires KE 2001. An introduction to nucleoside and nucleotide analogues. *Antivir Ther 6* (Suppl. 3): 1-14.

Steitz TA 1999. DNA polymerases: structural diversity and common mechanisms. *J Biol Chem 274*: 17395-17398.

Still WC, Tempczyk A, Hawley RC, Hendrickson T 1990. A general treatment of solvation for molecular mechanics. *J Am Chem Soc 112*: 6127.

Su H-P, Yan Y, Prasad GS, Smith RF, Daniels CL, Abeywickrema PD, Reid JC, Loughran HM, Kornienko M, Sharma S 2010. Structural basis for the inhibition of RNase H activity of HIV-1 reverse transcriptase by RNase H active site-directed inhibitors. *J Virol 84*: 7625-7633.

Tarasova OA, Urusova AF, Filimonov DA, Nicklaus MC, Zakharov AV, Poroikov VV 2015. QSAR modeling using large-scale databases: case study for HIV-1 reverse transcriptase inhibitors. *J Chem Inf Model 55*: 1388-1399.

Taylor RD, Jewsbury PJ, Essex JW 2002. A review of protein-small molecule docking methods. *J Comput Aided Mol Des 16*: 151-166.

Thomsen R, Christensen MH 2006. MolDock: a new technique for high-accuracy molecular docking. *J Med Chem 49*: 3315-3321.

Titmuss SJ, Keller PA, Griffith R 1999. Docking experiments in the flexible non-nucleoside inhibitor binding pocket of HIV-1 reverse transcriptase. *Bioorg Med Chem 7*: 1163-1170.

Tramontano E, Di Santo R 2010. HIV-1 RT-associated RNase H function inhibitors: recent advances in drug development. *Curr Med Chem 17*: 2837-2853.

Triballeau N, Acher F, Brabet I, Pin JP, Bertrand HO 2005. Virtual screening workflow development guided by the "receiver operating characteristic" curve approach. Application to high-throughput docking on metabotropic glutamate receptor subtype 4. *J Med Chem 48*: 2534-2547.

Tronchet JM, Seman M 2003. Nonnucleoside inhibitors of HIV-1 reverse transcriptase: from the biology of reverse transcription to molecular design. *Curr Top Med Chem 3*: 1496-1511.

Trott O, Olson AJ 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J Comput Chem 31*: 455-461.

Tuske S, Sarafianos SG, Clark AD, Ding J, Naeger LK, White KL, Miller MD, Gibbs CS, Boyer PL, Clark P, Wang G, Gaffney BL, Jones RA, Jerina DM, Hughes SH, Arnold E 2004. Structures of HIV-1 RT-DNA complexes before and after incorporation of the anti-AIDS drug tenofovir. *Nat Struct Mol Biol 11*: 469-474.

UNAIDS - Joint United Nations Programme on HIV/AIDs 2013. Global report. UNAIDS report on the global AIDS epidemic 2013. Available from: unaids.org/sites/default/files/media_asset/UNAIDS_Global_Report_2013_en_1.pdf.

Vadivelan S, Deeksha TN, Arun S, Machiraju PK, Gundla R, Sinha BN, Jagarlapudi SA 2011. Virtual screening studies on HIV-1 reverse transcriptase inhibitors to design potent leads. *Eur J Med Chem 46*: 851-859.

Vaidya A, Jain S, Jain S, Jain AK, Agrawal RK 2014. Quantitative structure-activity relationships: a novel approach of drug design and discovery. *J Pharm Sci Pharmacol 1*: 219-232.

van Aalten DM, Conn DA, de Groot BL, Berendsen HJ, Findlay JB, Amadei A 1997. Protein dynamics derived from clusters of crystal structures. *Biophys J 73*: 2891-2896.

Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ 2005. GROMACS: fast, flexible and free. *J Comput Chem 26*: 1701-1718.

Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD 2003. Improved protein-ligand docking using GOLD. *Proteins 52*: 609-623.

Vo TS, Kim SK 2010. Potential anti-HIV agents from marine resources: an overview. *Mar Drugs 8*: 2871-2892.

Wang LZ, Kenyon GL, Johnson KA 2004. Novel mechanism of inhibition of HIV-1 reverse transcriptase by a new non-nucleoside analog, KM-1. *J Biol Chem 279*: 38424-38432.

Waters LJ, Moyle G, Bonora S, D'Avolio A, Else L, Mandalia S, Pozniak A, Nelson M, Gazzard B, Back D, Boffito M 2007. Abacavir plasma pharmacokinetics in the absence and presence of atazanavir/ritonavir or lopinavir/ritonavir and *vice versa* in HIV-infected patients. *Antivir Ther 12*: 825-830.

Wlodawer A 2002. Rational approach to AIDS drug design through structural biology. *Annu Rev Med 53*: 595-614.

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

Yu F, Liu X, Zhan P, De Clercq E 2008. Recent advances in the research of HIV-1 RNase H inhibitors. *Mini Rev Med Chem 8*: 1243-1251.

Yuriev E, Ramsland PA 2013. Latest developments in molecular docking: 2010-2011 in review. *J Mol Recognit 26*: 215-239.

Zeevaart JG, Wang LG, Thakur VV, Leung CS, Tirado-Rives J, Bailey CM, Domaoal RA, Anderson KS, Jorgensen WL 2008. Optimization of azoles as anti-human immunodeficiency virus agents guided by free-energy calculations. *J Am Chem Soc 130*: 9492-9499.

Zhan P, Chen X, Li D, Fang Z, Clercq E, Liu X 2013. HIV-1 NNRTIs: structural diversity, pharmacophore similarity and impliations for drug design. *Med Res Rev 33*: E1-E72.

Zhan P, Li Z, Liu X 2010. Cosalane and its analogues: a unique class of anti-HIV agents. *Mini Rev Med Chem 10*: 966-976.

Zheng MY, Liu X, Xu Y, Li HL, Luo C, Jiang HL 2013. Computational methods for drug design and discovery: focus on China. *Trends Pharmacol Sci 34*: 549-559.

Zhou Z, Madrid M, Madura JD 2002. Docking of non-nucleoside inhibitors: neotripteriiordin and its derivatives to HIV-1 reverse transcriptase. *Proteins 49*: 529-542.