Binding Mode Prediction of Strand Transfer HIV-1 Integrase Inhibitors Using Tn5 Transposase as a Plausible Surrogate Model for HIV-1 Integrase

Maria Letizia Barreca,* Laura De Luca, Nunzio Iraci, and Alba Chimirri

Dipartimento Farmaco-Chimico, Università di Messina, Viale Annunziata, 98168 Messina, Italy

Received March 21, 2006

The crystal structure of Tn5 transposase−DNA complex was used in docking experiments to predict binding modes of HIV-1 integrase strand transfer inhibitors (INSTIs). In fact, the identification of HIV-1 integrase inhibitors from an in vitro screen using Tn5 transposase as the target has been recently reported. Our results suggest the utility of this protein as a useful surrogate model for IN and also for in silico screening, in the search for new potential INSTIs.

Introduction

The HIV-1 integrase (IN) enzyme is an attractive target for the treatment of HIV infection.[1] It inserts a double stranded DNA copy of the viral RNA genome into the chromosomes of an infected cell through two separate catalytic steps: in the ''3′-processing" reaction, IN removes the terminal GT dinucleotides from each LTR (long-terminal repeat) 3′ end, while in the "strand transfer" reaction the two newly processed 3′- viral DNA ends are joined to the cellular DNA.[2] The integration reaction is completed by the removal of the two unpaired nucleotides at the 5′-end of the viral DNA and the repair of the single stranded gaps created between the viral and target DNA. Moreover, there are no cellular homologues to IN, and the reactions catalyzed by IN are unique.

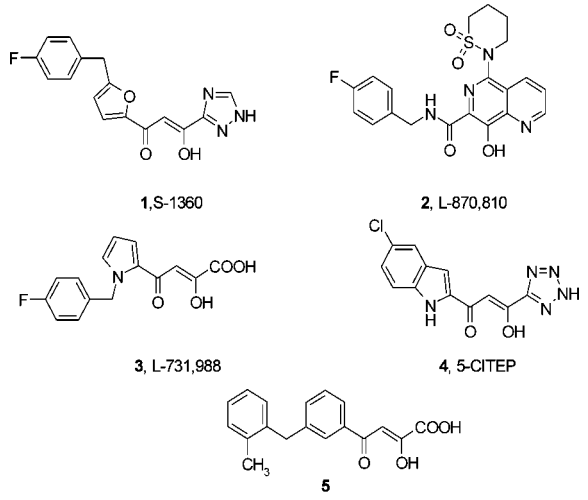
The full-length IN monomer is composed of three structurally and functionally distinct domains: the N-terminal domain, the central catalytic core domain, and the C-terminal domain.[2]

Moreover, a divalent cation-bound active site, involving Mg^{2+} or Mn^{2+}, is required for the integration reaction. The HIV-1 IN crystal structures available to date show a single binding site for Mg^{2+} or Mn^{2+}; on the other hand, both biochemical and structural studies offer a plausible two-metal model for the catalytic center of IN. In fact, because a second metal has been observed in an avian sarcoma virus (ASV) IN crystal structure,[3] and because of the two-metal structure for polynucleotide transferases,[4,5] it has been proposed that the second metal is carried into the active site once HIV-1 integrase binds its DNA substrate(s).[6,7]

Despite the unique role of IN in the viral replication process, drugs active against this enzyme have not as yet been approved by the FDA. The first two highly specific HIV-1 integrase inhibitors under clinical trials have been S-1360 and L-870,-810, which belong to a class of compounds known as strand-transfer inhibitors (diketo acids (DKAs) and DKA-like compounds), such as the other well-known IN inhibitors L-731,988, 5-CITEP, and **5** (Chart 1, **1**−**5**).[8−10]

These molecules have been shown to selectively inhibit the integrase strand transfer step[11] by sequestering the metal ions bound in the active site of the enzyme;[6,7] moreover, they bind to IN after the protein has formed a complex with target DNA.[12] Unfortunately, there is a lack of detailed structural information about the three-domain protein structure and the interaction between IN and its substrates, and consequently structure-based design of novel HIV-1 IN inhibitors is currently hampered.

Chart 1



However, it has been recently reported that the Tn5 transposase (Tnp) can be considered an excellent surrogate model for IN. In fact, Ason et al. screened a chemical library for inhibitors of Tn5 Tnp and identified six compounds that inhibit both Tn5 Tnp and HIV IN in vitro.[13] This interesting result can be explained considering that Tn5 Tnp and IN are both members of the superfamily of polynucleotidyl transferases, and there are many similarities between the catalytic mechanism and the active site architecture of these enzymes.[14] In particular, the two proteins share a high degree of structural similarity of the catalytic triad of acidic residues, known as DDE motif, which coordinate divalent metal ions required for catalysis.

The identification of IN inhibitors from a screen using Tn5 Tnp as the target suggested that the lack of structural data for the complex between IN and its DNA substrates could be overcome by the availability of crystal structure of Tn5 Tnp with DNA and two metal ions bound in the active site. Furthermore, looking at the six structures found to inhibit the catalytic activity of both IN and Tn5 Tnp, we have noticed that one of these compounds (Figure 1, **6**) showed all the chemical features recently highlighted by our 3D pharmacophore model for known IN inhibitors (Figure 1).[15]

For the above reasons and as continuation of our previous work, we have used the X-ray cocrystal structure of Tn5 Tnp−

* To whom correspondence should be addressed. Tel: (++39) 090-6766464; Fax: (++39) 090-355613; E-mail: barreca@pharma.unime.it.

10.1021/jm060323r CCC: $33.50
Published on Web 05/26/2006

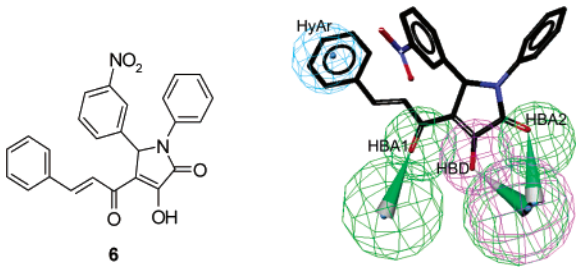


Figure 1. Compound **6** and its mapping into 3D pharmacophore model for INSTIs (HBD, hydrogen bond donor, purple; HBA, hydrogen bond acceptor, green; HyAr, hydrophobic aromatic region, cyan).

DNA−metal ternary complex as a model for our structure-based molecular modeling studies. In particular, automated docking experiments on the Tn5−DNA complex have been performed in an attempt to address the potential binding modes of integrase strand transfer inhibitors (INSTIs), and thus to obtain comparable information about the possible interactions between IN and its inhibitors.

Results and Discussion

Docking simulations of inhibitors **1**−**6** into the Tn5 Tnp−DNA complex were performed using the automated docking program GOLD 3.0[16,17] and the GOLD fitness function to rank the compounds on the basis of their ability to form favorable interactions with the protein active site. Due to the available experimental information on compound **6**, i.e. it is an inhibitor of both Tn5 Tnp and IN, this molecule was used as a reference for the analysis of the docking results of **1**−**5**.

The best docking poses of **1**−**6** were selected on the basis of the GOLD fitness score; for each molecule the bound conformation with the highest fitness score was chosen as the predicted ligand binding mode. Moreover, compounds **3**−**5** resulted in an almost identical docked pose and GOLD score in their neutral versus deprotonated forms.

The selected conformations of **1**−**6** revealed that the ligands bound to the same region of the protein (Figure 2) and were able to interact with the two divalent ions in the active site, in perfect agreement with the proposed mechanism of action for INSTIs, that is the functional sequestration of the critical metal cofactors. The docked ligand−target complexes generated by GOLD docking were analyzed by the Ligplot 4.22 program,[18] and a detailed list of all the interactions for a given ligand and the Tn5 Tnp/DNA complex are reported in the Supporting Information. Table 1 summarizes the fitness scores and the most important interactions between **1**−**6** and Tn5 Tnp−DNA complex, considering the residues included in a distance of 5.0 Å starting from the center of the ligand.

In their predicted bound conformation, compounds **1**−**6** show close contact with the three catalytic residues (i.e. D97, D188, and E326); moreover, all the ligands interact with the donor DNA, in accordance with biochemical experiments, which suggest an interaction between DKAs and the terminal unpaired dinucleotide present in the donor DNA.[19] The hydroxyl group of **1** makes a H-bonding interaction with T98, and the NH of the triazole ring establishes a hydrogen bond with the phosphodiester backbone of the 3′ terminal nucleotide of the DNA transferred strand (i.e. Gua 20). In its best docking pose, the hydroxyl group of **2** is hydrogen-bonded to T98, while one of the oxygens of the six-membered cyclic sulfonamide substituent of the inhibitor interacts via a hydrogen bond with R150. In the orientation that **3** adopts in the molecular docking calculations, the carboxylate group of the ligand is hydrogen-bonded to the K164. The enol oxygen and the keto group of **4** makes H-bonding interaction with the nucleotide Gua20 and T99, while the tetrazole motif is engaged in hydrogen bond interaction with T98. Compound **5** is also hydrogen-bonded to the macromolecule target; in particular, the keto-, the enol- and the carboxylate groups interact with T98, D188 and E190, respectively. Finally, since both enantiomers of **6** were considered for docking, the

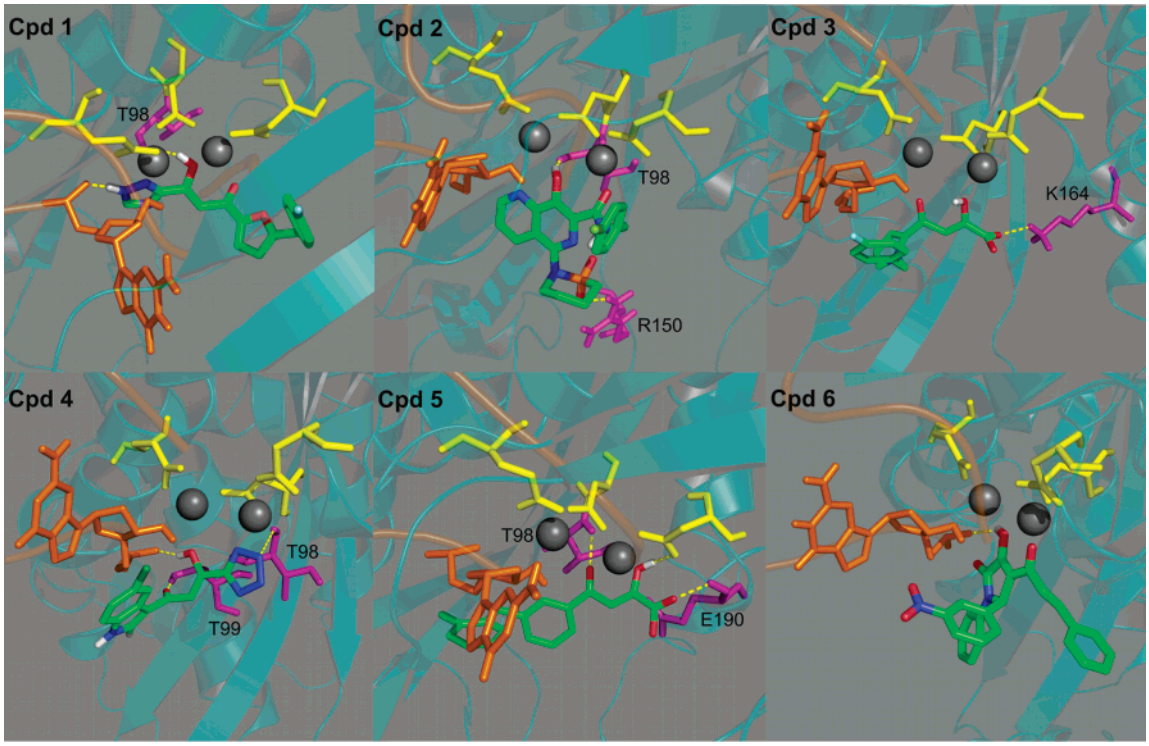


Figure 2. Proposed binding modes of compounds **1**−**6**. Catalytic residues (D97, D188, and E326) are shown in yellow with divalent metal ions as gray spheres. Hydrogen bonds (shown as dashed yellow lines) are formed between the ligands (colored by atom type) and the complex Tn5 Tnp−DNA, whose amino acids are in violet, while the terminal nucleotide of the 3′-viral DNA end (Gua 20) is in orange.

Table 1. Close Interatomic Contacts between the Ligands **1**−**6** and the Target, and Fitness Score Values

Tn5 Tnp[a]	IN[a]	**1**[b]	**2**[b]	**3**[b]	**4**[b]	**5**[b]	**6**[b]
D97	**D64**	X	X	X	X	X	X
T98	C65	**X**	**X**	X	**X**	**X**	X
T99	T66	X	X	X	**X**	X	X
S100	H67	X	X	X	X	X	X
W125	V72	X	X		X	X	X
W148	P90						X
R150			**X**				X
E161	T93						X
K164	E96	X		**X**		X	X
W165		X		X		X	
D188	**D116**	X	X	X		**X**	X
R189	N117	X	X			X	
E190	G118	X	X	X	X	**X**	X
R210		X					
K212		X	X				
E326	**E152**	X		X			X
H329	N155	X	X	X	X	X	X
K333				X		X	
C1		X	X	X	X	X	X
A19					X	X	
G20		**X**	X	X	**X**	X	**X**
GOLD fitness values[c]		61.15	59.36	60.54	52.81	52.54	51.29

[a] Selected Tn5 Tnp residues in close contact with the ligands (5.0 Å cutoff) and equivalent residues in IN. The active site residues are shown in bold; residues associated with resistance to IN inhibitors are underlined; C1 and A19 are DNA nucleotide bases, while G20 is the terminal nucleotide of the 3′-viral DNA end. [b] Residues that show close contacts and hydrogen bond interactions with the corresponding ligand are highlighted by a cross and a bold cross, respectively. [c] The pose with the highest GOLD score for each compound was considered as the best docking position.

top score result was obtained for the *(R)*-enantiomer (*R*-score: 51.29; *S*-score: 47.80), which exhibits an hydrogen-bond with Gua 20. In terms of VdW interactions; according to the six docking experiments, the following residues are always involved in favorable interactions: D97, T98, T99, S100, E190 and H329.

We later performed a structure-based sequence alignment of the Tn5 transposase core domain and the HIV-1 IN core domain (PDB code 1BIS), to look at the equivalent inhibitor binding pocket in the IN enzyme (Table 1). The alignment suggested that the INSTI binding site is formed by the three catalytic residues (i.e. D64, D116, and E152) and mainly by polar residues. It was worth noting that T99, W125, and H329 in Tn5 Tnp correspond to T66, V72, and N155 in IN, respectively; these residues have been associated with resistance to INSTIs,[9] indicating once again the plausibility of the proposed mode of interaction.

In particular, Hazuda et al. have recently suggested that, although **2** (the prototype of naphthyridine carboxamide inhibitors) and DKAs (i.e. **3** and **5**) bind to the same region in the IN active site and have a similar mechanism of action (i.e. the metal chelation), on the basis of their resistance profile they may show opposite orientations of their substituents (i.e the substituted-benzyl rings).[20] Our docking results are in perfect agreement with this hypothesis, as illustrated in Figure 2; in fact, even if compounds **2**, **3**, and **5** bind the divalent metals within the protein active site, the respective substituents of **3** and **5** point in the opposite direction of the compound **2** substituent (see also Supporting Information). In an attempt to assess the docking results with regard to our previously developed 3D pharmacophore model for IN inhibitors,[15] the inhibitors in their "predicted" conformation (i.e. bound to Tn5 Tnp) were mapped onto the Catalyst hypothesis using the "Fast fit" option; the selected poses were able to fit the four chemical features highlighted by the hypothesis (see Supporting Information), confirming the plausibility of the suggested GOLD binding modes.

Conclusion

In summary, our studies give structural insight into plausible binding mode for INSTIs that is consistent with the mechanism of action of this class of IN inhibitors. The results suggest that, in absence of complete information on IN/DNA interaction, alternative computational strategies for identifying new HIV-1 IN inhibitors may consist in using the coordinates of Tn5 Tnp−DNA complex as three-dimensional target for the purpose of structure-based discovery of IN inhibitors. For instance, we are going to use the above-mentioned complex for in silico screening of large databases of commercially available chemical compounds in order to identify new potential INSTIs.

Experimental Section

Mapping of Compound 6 into our 3D Pharmacophore Model for IN Strand Transfer Inhibitors. Compound **6** was sketched within Catalyst 4.10 software package[21] and minimized to its closest local energy minimum using a molecular mechanics approach. Poled conformations were generated for this molecule using the "Best" conformer generation option and an energy cutoff of 10 kcal/mol. The compound was later overlaid against our recently reported four-feature pharmacophore model for IN strand transfer inhibitors (INSTIs) by using the Best Fit option.

Automated Molecular Docking Experiments. Prior to docking, the ligand structures (compound **1**−**6**) were constructed using the 3D-Sketcher tool available in Maestro and processed using the Schrodinger LigPrep[22] utility, which produces a number of low-energy 3D structures from each molecule input with various ionization states, tautomers, stereochemistries, and ring conformations to be considered for the docking studies. For our study a pH range of 6−8 was set. The compounds were considered in their keto−enol tautomeric form, since it has been clearly established that these molecules mainly exist in this form in solution. Moreover, both neutral and ionic forms were generated for the carboxylic acid and tetrazole groups of compounds **3**−**5**.

Regarding the chirality of the asymmetric carbon atom in compound **6**, as no experimental data on the biologically relevant conformation of this molecule were available, both enantiomers were docked into our target.

The crystal structure of Tn5 transposase−DNA complex was retrieved from the RCSB Protein Data Bank (entry code 1MUS), and water and ethylene glycol molecules were discarded using Maestro.[22]

The element names and formal charges of the two metal ions (Mn^{2+}) in the protein active site were corrected. The FirstDiscovery Protein Preparation[22] procedure was used to obtain a reasonable starting structure for docking studies. This facility is designed to ensure chemical correctness and to optimize the protein structure for further analysis; the process adds hydrogens, neutralizes appropriate amino acid chains, and relieves steric clashes. In particular, it performs a series of restrained, partial minimizations on the cocrystallized structure; each of these employs a limited number of minimization steps and is not intended to minimize the system completely. In our study, the minimization (OPLS 2001 force field) was stopped after RMSD of the non-hydrogen atoms reached 0.30 Å, that is the specified limit by default.

Automated docking studies were then performed using CCDC's GOLD (Genetic Optimization for Ligand Docking) software version 3.0.[16,17] The algorithm has been previously validated and successfully tested on a data set of over 300 complexes extracted from the Protein DataBank.[23] The binding site was initially defined as all residues of the target (i.e. Transposase Tn5-DNA complex) within 10 Å from Mn atom number 8519, and later automated cavity detection was used.

GOLD score was chosen as fitness function and the standard default settings were used in all the calculations. For each of the 100 independent genetic algorithm runs, a default maximum of 100 000 genetic operations was performed, using the default operator weights and a population size of 100 chromosomes. Default

cutoff values of 2.5 Å for hydrogen bonds and 4.0 Å for VdW were employed. Results differing by less than 1.5 Å in ligand-all atom RMSD were clustered together.

Acknowledgment. Financial support for this research by MIUR (COFIN2004,Roma, Italy) and European TRIoH Consortium (LSHB-CT-2003-503480)) is gratefully acknowledged.

Supporting Information Available: Molecular modeling details are available free of charge via the Internet at http://pubs.acs.org.

References

(1) Anthony, N. J. HIV-1 integrase: a target for new AIDS chemotherapeutics. *Curr. Top. Med. Chem.* **2004**, *4*, 979−990.
(2) Chiu, T. K.; Davies, D. R. Structure and function of HIV-1 integrase. *Curr. Top. Med. Chem.* **2004**, *4*, 965−977.
(3) Bujacz, G.; Alexandratos, J.; Wlodawer, A.; Merkel, G.; Andrake, M. et al. Binding of different divalent cations to the active site of avian sarcoma virus integrase and their effects on enzymatic activity. *J. Biol. Chem.* **1997**, *272*, 18161−18168.
(4) Beese, L. S.; Steitz, T. A. Structural basis for the 3′-5′ exonuclease activity of Escherichia coli DNA polymerase I: a two metal ion mechanism. *EMBO J.* **1991**, *10*, 25−33.
(5) Yang, W.; Steitz, T. A. Recombining the structures of HIV integrase, RuvC and RNase H. *Structure* **1995**, *3*, 131−134.
(6) Grobler, J. A.; Stillmock, K.; Hu, B.; Witmer, M.; Felock, P. et al. Diketo acid inhibitor mechanism and HIV-1 integrase: implications for metal binding in the active site of phosphotransferase enzymes. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 6661−6666.
(7) Marchand, C.; Johnson, A. A.; Karki, R. G.; Pais, G. C.; Zhang, X. et al. Metal-dependent inhibition of HIV-1 integrase by beta-diketo acids and resistance of the soluble double-mutant (F185K/C280S). *Mol. Pharmacol.* **2003**, *64*, 600−609.
(8) Johnson, A. A.; Marchand, C.; Pommier, Y. HIV-1 integrase inhibitors: a decade of research and two drugs in clinical trial. *Curr. Top Med. Chem.* **2004**, *4*, 1059−1077.
(9) Pommier, Y.; Johnson, A. A.; Marchand, C. Integrase inhibitors to treat HIV/AIDS. *Nat. Rev. Drug Discovery* **2005**, *4*, 236−248.
(10) Cotelle, P. Patented HIV-1 Integrase Inhibitors (1998−2005). *Recent Patents on Anti-Infective Drug Discovery* **2006**, *1*, 1−15.
(11) Hazuda, D. J.; Felock, P.; Witmer, M.; Wolfe, A.; Stillmock, K. et al. Inhibitors of strand transfer that prevent integration and inhibit HIV-1 replication in cells. *Science* **2000**, *287*, 646−650.
(12) Espeseth, A. S.; Felock, P.; Wolfe, A.; Witmer, M.; Grobler, J. et al. HIV-1 integrase inhibitors that compete with the target DNA substrate define a unique strand transfer conformation for integrase. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 11244−11249.
(13) Ason, B.; Knauss, D. J.; Balke, A. M.; Merkel, G.; Skalka, A. M. et al. Targeting Tn5 transposase identifies human immunodeficiency virus type 1 inhibitors. *Antimicrob. Agents Chemother.* **2005**, *49*, 2035−2043.
(14) Rice, P. A.; Baker, T. A. Comparative architecture of transposase and integrase complexes. *Nat. Struct. Biol.* **2001**, *8*, 302−307.
(15) Barreca, M. L.; Ferro, S.; Rao, A.; De Luca, L.; Zappala, M. et al. Pharmacophore-based design of HIV-1 integrase strand-transfer inhibitors. *J. Med. Chem.* **2005**, *48*, 7084−7088.
(16) GOLD 3.0; Cambridge Crystallographic Data Centre: Cambridge CB2 1EZ, UK.
(17) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727−748.
(18) Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. LIGPLOT: A program to generate schematic diagrams of protein−ligand interactions. *Prot. Eng.* **1995**, *8*, 127−134.
(19) Marchand, C.; Zhang, X.; Pais, G. C.; Cowansage, K.; Neamati, N. et al. Structural determinants for HIV-1 integrase inhibition by beta-diketo acids. *J. Biol. Chem.* **2002**, *277*, 12596−12603.
(20) Hazuda, D. J.; Anthony, N. J.; Gomez, R. P.; Jolly, S. M.; Wai, J. S. et al. A naphthyridine carboxamide provides evidence for discordant resistance between mechanistically identical inhibitors of HIV-1 integrase. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 11233−11238.
(21) Catalyst 4.10; Accelrys Inc., San Diego, CA.
(22) Schrodinger, Inc., www.schrodinger.com.
(23) Nissink, J. W.; Murray, C.; Hartshorn, M.; Verdonk, M. L.; Cole, J. C. et al. A new test set for validating predictions of protein−ligand interaction. *Proteins* **2002**, *49*, 457−471.

JM060323R