Pharmacophore-Based Discovery of Small-Molecule Inhibitors of Protein–Protein Interactions between HIV-1 Integrase and Cellular Cofactor LEDGF/p75


The cellular protein lens epithelium-derived growth factor, or transcriptional coactivator p75 (LEDGF/p75), plays a crucial role in HIV integration. The protein–protein interactions (PPIs) between HIV-1 integrase (IN) and its cellular cofactor LEDGF/p75 may therefore serve as targets for the development of new anti-HIV drugs. In this work, a structure-based pharmacophore model for potential small-molecule inhibitors of HIV-1 IN–LEDGF/p75 interaction was developed using the LigandScout software. The 3D model obtained was used for virtual screening of our in-house chemical database, CHIME, leading to the identification of compound CHIBA-3002 as an interesting hit for further optimization. The rational design, synthesis and biological evaluation of four derivatives were then carried out. Our studies resulted in the discovery of a new and more potent small molecule (7, CHIBA-3003) that is able to interfere with the HIV-1 IN–LEDGF/p75 interaction at micromolar concentration, representing one of the first compounds to show activity against these specific PPIs. Docking simulations were subsequently performed in order to investigate the possible binding mode of our new lead compound to HIV-1 IN. This study is a valid starting point for the identification of anti-HIV agents with a different mechanism of action from currently available antiviral drugs.

Introduction

Although anti-HIV drugs typically target viral proteins, the manipulation of specific protein–protein interactions (PPIs) involved in the HIV life cycle could potentially result in potent drugs that lack cellular toxicity. With this in mind, the objective of our research is hit identification and lead development of small-molecule inhibitors that are able to disrupt specific protein–protein interactions (PPIs). In particular, the interaction between HIV-1 integrase (IN) and the cellular protein lens epithelium-derived growth factor or transcriptional coactivator p75 (LEDGF/p75) has increasingly gained attention as a valuable target for a novel antiviral strategy.[1–4]

The HIV-1 IN enzyme catalyzes an essential step in the retroviral life cycle—insertion of viral DNA into the genome of the host cell through a multistep process. In the "3'-processing" step, HIV-1 IN removes a dinucleotide from each 3' end of viral cDNA. These two newly processed 3'-viral DNA ends are then inserted into the host cell DNA during the "strand transfer" reaction.

Sequence alignments, mutagenesis and proteolytic digestion suggested that HIV-1 IN is comprised of three structurally and functionally distinct domains: an N-terminal zinc-binding HHCC domain, a central catalytic core domain and a C-terminal DNA-binding domain.[5]

All three domains bind DNA, and each separate domain forms a homodimer in solution. Even if all three domains are required for full catalytic activity, site-directed mutagenesis experiments have shown that the central core domain is sufficient to promote a reverse integration reaction in vitro, known as disintegration, indicating that this region contains the enzymatic catalytic center.

Although HIV-1 IN plays a key role in integration, the process takes place in a more complex environment in vivo. Before integration, viral cDNA becomes associated with a number of viral and cellular proteins to form a large nucleoprotein assembly called a preintegration complex. LEDGF/p75 is a cellular protein that has recently been identified and validated as a novel cellular cofactor of HIV integration and replication.[6,7] LEDGF/p75 binds HIV-1 IN via a small (~80 residues) IN-binding domain (IBD, amino acids 347–429) within its C-terminal region. IBD is both necessary and sufficient for interaction with HIV-1 IN. The crystal structure of the dimeric catalytic core

[a] Dr. L. De Luca, Dr. S. Ferro, Prof. R. Gitto, Prof. A. M. Monforte, Prof. A. Chimirri
Department of Medicinal Chemistry, University of Messina
Viale Annunziata, 98168 Messina (Italy)
Fax: (+39) 090-355613
E-mail: ideluca@pharma.unime.it

[b] Dr. M. L. Barreca, Dr. N. Iraci
Dipartimento di Chimica e Tecnologia del Farmaco
Università degli Studi di Perugia, Via del Liceo 1, 06123 Perugia (Italy)
Fax: (+39) 075-5855115
E-mail: mbarreca@unipg.it

[c] Dr. F. Christ, Prof. Z. Debyser
Molecular Virology and Gene Therapy Molecular Medicine
Katholieke Universiteit Leuven and IRC KULAK
Kapucijnenvoer 33, 3000 Leuven, Flanders (Belgium)
Fax: (+32) 16-336-336
E-mail: zeger.debyser@med.kuleuven.be
domain of HIV-1 IN complexed with the LEDGF/p75 IBD has recently been reported (PDB code 2B4J).[8]

The aim of our study was thus to use these structural data as the starting point for a pharmacophore modeling approach to identify novel compounds able to disrupt HIV-1 IN–LEDGF/p75 interactions. The computational approaches presented herein could also be useful in general PPI inhibition strategies.

Results and Discussion

The cellular cofactor LEDGF/p75 plays a unique role during HIV integration and the identification of small molecules that interfere with HIV-1 IN–LEDGF/p75 interactions and block HIV replication might prove an enormous driving force in the field of antiretroviral therapeutics.[3]

The most critical interacting residues of the IBD are Ile365, Asp366 and Phe406; there have been recent reports that mutation of these LEDGF/p75 residues destroyed interaction with HIV-1 IN.[9] Therefore, we focused our interest on the LEDGF hotspot residues Ile365 and Asp366, supposing that a small molecule able to mimic this IBD dipeptide might inhibit IN–LEDGF/p75 recognition. This idea was also supported by a recent report showing that a LEDGF/p75-derived peptide (amino acids 355–377), which contains the two HIV-1 IN interacting residues of interest to us (Ile365 and Asp366), could disrupt the IN–LEDGF/p75 interaction.[9]

In the first step of our rational approach, the X-ray crystal structure of HIV-1 IN in complex with LEDGF/p75 IBD was used to generate a structure-based pharmacophore model. The assumption that this computational methodology could work well in a PPI inhibitor discovery process was partly encouraged by the successful application of the same approach in the identification of new non-nucleoside reverse transcriptase inhibitors.[10]

The analysis of protein–protein contacts by LigandScout enabled us to highlight the chemical features mainly involved in the interactions between LEDGF residues Ile365–Asp366 and IN, as described by Cherepanov et al.[8] The hypothesis obtained contained 14 features: one H-bond donor, two H-bond acceptors, two hydrophobic groups and nine excluded volumes (Figure 1). In particular, the H-bond donor reflected one of the two NH groups of Ile365–Asp366 that interacts with the carbonyl oxygen of Gln168. The two H-bond acceptor features were occupied by the Asp366 carboxylate oxygen atoms, which made a bidentate H bond to the main-chain amides of IN residues Glu170 and His171 in chain A. Lastly, the two hydrophobic spheres were occupied by the Ile365 side chain, which projects into a hydrophobic pocket formed by IN chain B residues Leu102, Ala128, Ala129 and Trp132 and chain A residues Thr174 and Met178. The nine excluded volumes reflected potential steric restriction and corresponded to the positions that are sterically claimed by the macromolecular environment surrounding the LEDGF dipeptide (residues Ala128, Ala129, Gln168, Trp132, Glu170, His171, Thr174 and Met178).

The 3D pharmacophore model obtained was then used as a query in a virtual screening approach to filter the CHIME database, which consists of 3055 small molecules synthesized and/or published by Chimirri and co-workers. The “fast flexible search” option in Catalyst 4.10 was selected. The search returned ~936 compounds, which contained the specified 3D location of chemical functions. A subset of 234 structures was created by considering only compounds with a Catalyst fitness score equal to or greater than 2.00. The best fitness values were obtained for the series of benzylindoles previously reported by our group as strand-transfer IN inhibitors.[11] In particular, the best fitness value (Fit = 2.8) was obtained for compound CHIBA-3002 (shown).

As a preliminary assessment, we tested CHIBA-3002 in the AlphaScreen assay and found that it inhibited the IN–LEDGF/p75 interaction by up to 46% at a concentration of 100 μM. However, the mapping of this molecule into the obtained pharmacophore model lacked the feature needed to form the H-bond interaction with IN residue Gln168. Moreover, we noticed that the CHIBA-3002 phenyl group was very close to a region considered off limits to the ligands. Based on these observations, we reasoned that suitable chemical modifications of CHIBA-3002 might improve the fitness to the pharmacophore hypothesis, leading to improved small-molecule inhibitors of the IN–LEDGF/p75 interaction.
We rationally design and synthesized some CHIBA-3002 analogues (4–7, Scheme 1). The 2-hydroxy-4-(4-hydroxy-1H-indol-3-yl)-4-oxobut-2-enoic acids 5 and 7 (CHIBA-3003) were prepared following a procedure similar to that employed for hit compound CHIBA-3002. 

The commercially available 4-hydroxy-1H-indole (1) was acylated under Vilsmeier–Haack reaction conditions, using N,N-dimethylacetamide and phosphoryl chloride, to give compound 2. Ethyl esters 4 and 6 were synthesized from this key intermediate using microwave-assisted organic synthesis (MAOS). The 4-fluorobenzyl derivative 5 and unsubstituted analogue 7 were subsequently formed by base-catalyzed ester hydrolysis of 4 and 6, respectively.

As a first step, we decided to introduce a potential H-bond functionality, absent in CHIBA-3002, by inserting a hydroxy group at position C-4 of the indole system (compound 5). The alignment of compound 5 to the 3D model provided a higher fit value compared to CHIBA-3002 (3.6 vs 2.8), since, as hypothesized above, derivative 5 was able to fulfill the pharmacophore hypothesis. Derivative 7 was designed, incorporating the hydroxy group at position C-4 of the indole but lacking the benzyl moiety. The fit value of this analogue (Fit = 4.1) surpassed those of both compounds 5 and 7. This improvement could be due to the absence of steric hindrance caused by the benzyl group, and consequently enabling a better positioning of the molecule in the pharmacophore model (Figure 2).

The synthesized compounds 5 and 7 were subsequently tested in the AlphaScreen assay to evaluate their ability to disrupt the HIV-1 IN interaction with LEDGF/p75 IBD, and proved to be inhibitors of that specific PPI. In agreement with our hypothesis, compound 5 showed a higher potency than CHIBA-3002, while compound 7 proved to be the most active molecule (Table 1). Ester derivatives 4 and 6 were also tested in the assay. Both esters showed lower activity than the corresponding acids. This result could be due to the steric hindrance of the alkyl group, preventing the molecule from adopting a suitable conformation so as to form H bonds between the ligand oxygen atoms and Glu 170 and His 171.

As a final step, the possible binding mode of the most active derivative 7 against HIV-1 IN was investigated by performing automated docking studies using the software GOLD. Firstly, test docking calculations using the LEDGF dipeptide Ile365–Asp366 were carried out with the aim of comparing experimental and predicted binding modes and validating the docking protocol. The best

**Table 1.** Inhibitory activities of CHIBA-3002 and compounds 4–7.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Inhibition [%]</th>
<th>IC₅₀ [µM]</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHIBA-3002</td>
<td>46</td>
<td>–</td>
</tr>
<tr>
<td>4</td>
<td>NA</td>
<td>–</td>
</tr>
<tr>
<td>5</td>
<td>56</td>
<td>76</td>
</tr>
<tr>
<td>6</td>
<td>37</td>
<td>–</td>
</tr>
<tr>
<td>7 (CHIBA-3003)</td>
<td>71</td>
<td>35</td>
</tr>
</tbody>
</table>

[a] Measured at 100 µM. [b] Concentration required to inhibit the HIV-1 IN–LEDGF/p75 interaction by 50%. NA = not active.

![ChemMedChem 0000, 00, 1–7 © 2009 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim www.chemmedchem.org](https://www.chemmedchem.org)
docking pose of the 19 clusters found for the LEDGF dipeptide agreed well with the experimental binding mode with a root-mean square deviation (RMSD) of 0.82.

The best docked conformation of compound 7 is shown in Figure 3a, together with the experimental position of the LEDGF residues Ile 365–Asp 366. Compound 7 interacted with the HIV-1 IN residues in a similar fashion to that of the LEDGF dipeptide. Indeed, in agreement with the mapping of the designed molecule to the 3D pharmacophore model, the carboxylate group interacted with the main chains of Glu 170 and His 171 by H-bond interactions, while the hydrophobic pocket formed by IN chain B residues was occupied by the fused ben-

zene ring of the indole system. Moreover, compound 7 formed a H bond with the backbone carbonyl group of IN amino acid Gln 168 as expected, due to the presence of the hydroxy group at C-4 position. (Figure 3b) The docking results confirmed that CHIBA-3003 (7) could inhibit the IN–LEDGF/p75 interaction by mimicking some specific PPIs between the viral enzyme and its cellular cofactor.

Recently, a benzoic acid derivative was reported as an inhibitor of the IN–LEDGF/p75 interaction,[2] although the proposed IN-binding site seems to be different from the one proposed in this manuscript.

Conclusions

This study identified small molecules able to inhibit IN–LEDGF/p75 interactions, through pharmacophore modeling, virtual library screening and rational drug design.

In particular, compound 7 could prove to be a valid starting point for the discovery of new and more potent derivatives able to disrupt PPIs between IN and its intracellular cofactor LEDGF/p75, potentially blocking HIV replication through a mechanism of action that differs to those of antiviral agents used currently. Furthermore, some derivatives were also active in the IN inhibition test suggesting a possible dual mode of action may be responsible for their potency. This could be particularly important considering the need for combined therapy targeting the different stages of the HIV life cycle to effectively suppress infection.

Further studies are in progress to improve the affinity of the inhibitors for the protein–protein interface, and to obtain anti-HIV agents active against multiple targets.

Experimental Section

Pharmacophore modeling and in silico screening

The software LigandScout[14] was used for the detection and interpretation of crucial interaction patterns between IN and specific cofactor LEDGF/p75 residues. LigandScout allows the automatic construction and visualization of 3D pharmacophores from structural data of macromolecule–ligand complexes. Detected chemical features include H-bond donors and acceptors as direct-ed vectors, positive and negative ionizable regions, as well as lipophilic areas represented by spheres. In order to increase selectivity, the LigandScout model also includes spatial information about potentially inaccessible areas for any ligand, thus taking into account any potential steric restrictions. In particular, excluded volume spheres placed in positions that are sterically forbidden are automatically added to the generated pharmacophore model.

We have used the crystal structure of the dimeric catalytic core domain of HIV-1 IN complexed with the recently reported LEDGF/p75 IN-binding domain (IBD)
Molecular Docking experiments

The crystal structure of the dimeric catalytic core domain of HIV-1 IN complexed with the LEDGF/p75 IBD was retrieved from the RCSB Protein Data Bank (entry code 2B4J) and used as our docking simulation target. The LEDGF structure was removed and hydrogen atoms were added to the IN protein using the Biopolymer module in Sybyl 8.0. The structures of the ligands were constructed using standard bond lengths and angles from the sp3 format for the Catalyst software package version 4.10 giving the title compound as a white solid (92%): mp: 264–266 °C; 1H NMR: δ = 1.35 (t, J = 7.1 Hz, 3H, CH3), 4.08 (q, J = 5.45 (s, 2H, CH2), 6.56–9.07 (m, 9 H, ArH and CH), 14.64 ppm (s, 1 H, OH). Anal. calcd for C10H9NO2; C, 68.56; H, 5.18; N, 8.00. Found: C, 68.72; H, 5.34; N, 7.51.

3-Acetyl-1-(4-fluorobenzyl)-4-hydroxy-1H-indole (3): 3-Acetyl-4-hydroxy-1H-indole (2) (159 mg, 0.001 mol) was dissolved in DMF (2 mL) and K2CO3 (138 mg, 0.001 mol) was added. The mixture was stirred for 2 min and then 4-fluorobenzyl bromide (568 mg, 0.003 mol) was added dropwise. The resulting solution was placed in a quartz tube (Ø 2 cm) and irradiated in a microwave oven (100 W) at 100 °C for 10 min with stirring. A saturated NaHCO3 solution was added and the mixture was extracted with EtOAc. The combined organic extracts were dried (Na2SO4), filtered and concentrated in vacuo. The residue was powdered with treatment with Et2O and crystallized from CH2Cl2 to give the title compound as a white solid (77%): mp: 138–140 °C; 1H NMR: δ = 2.53 (s, 3H, CH3), 5.28 (s, 2H, CH2), 6.72–7.18 (m, 7H, ArH), 7.66 (s, 1H, ArH), 11.50 ppm (s, 1H, OH). Anal. calcd for C9H8F3NO; C, 65.79; H, 4.73; N, 3.65. Found: C, 65.93; H, 4.62; N, 3.84.

Ethy1 4-[1-(4-fluorobenzyl)-4-hydroxy-1H-indol-3-yl]-2-hydroxy-4-oxo-2-enoic acid (5): Prepared following a previously reported procedure giving the title compound as a white solid (90%): mp: 223–225 °C; 1H NMR: δ = 5.45 (s, 2H, CH2), 6.56–9.07 (m, 9H, ArH and CH), 11.26 ppm (s, 1H, OH). Anal. calcd for C10H9NO2; C, 64.23; H, 3.97; N, 3.96. Found: C, 64.48; H, 4.22; N, 3.78.

2-Hydroxy-4-[1-(1H-indol-3-yl)-4-oxobut-2-enoic acid (7): Prepared following a previously reported procedure giving the title compound as a white solid (82%): mp: 208–210 °C; 1H NMR: δ = 6.54 (d, 1H, J = 8.11 Hz, ArH), 6.91 (d, 1H, J = 8.11 Hz, ArH), 7.11 (t, 1H, J = 8.11 Hz, ArH), 8.77 (s, 1H, ArH), 11.17 (s, 1H, OH). Anal. calcd for C10H10NO4; C, 58.30; H, 3.67; N, 5.67. Found: C, 58.46; H, 3.82; N, 5.44.

HIV-1 IN−LEDGF/p75 interaction screening

The AlphaScreen assay was performed as described previously. Reactions were performed in 25 μL final volume in 384-well Optiwell microtiter plates (Perkin−Elmer). The reaction buffer contained 25 mM Tris-HCl (pH 7.4), 150 mM NaCl, 1 mM MgCl2, 0.01% (w/v) Tween-20 and 0.1% (w/v) bovine serum albumin. His6-tagged inte-
grase (300 nM final concentration) was incubated with the compounds at 4°C for 30 min. The compounds were added in varying concentrations (1–100 μM). Afterwards 100 nM of recombinant flag-LEDGF/p75 was added and incubation was extended by 1 h at 4°C. Subsequently, 5 μL of Ni-chelate-coated acceptor beads and 5 μL of antiflag donor beads were added to a final concentration of 20 μg mL⁻¹ of both beads. Proteins and beads were incubated at 30°C for 1 h in order to allow association to occur. Exposure of the reaction to direct light was prevented as much as possible and the emission of light from the acceptor beads was measured in the EnVision plate reader (Perkin–Elmer, Benelux) and analyzed using the EnVision manager software.

Acknowledgements

This work was supported by the European Commission (HEALTH-F3–2008–201032) (THINC project).

Keywords: antiviral agents · inhibitors · HIV-1 integrase · protein–protein interactions · molecular modeling


Received: February 20, 2009
Revised: June 4, 2009
Published online on 2009-09-16
Anti-HIV agents: The identification of potential antiviral agents discovered using crystal data pharmacophore modeling, virtual library screening and rational drug design is described. These compounds are one of the first examples of small molecules able to inhibit HIV integrase–LEDGF/p75 interactions.