



Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Small molecules targeting the interaction between HIV-1 integrase and LEDGF/p75 cofactor

Laura De Luca^{a,*}, Stefania Ferro^a, Rosaria Gitto^a, Maria Letizia Barreca^b, Stefano Agnello^a, Frauke Christ^c, Zeger Debyser^c, Alba Chimirri^a

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

^b Dipartimento di Chimica e Tecnologia del Farmaco, Università di Perugia Via del Liceo 1, I-06123 Perugia, Italy

^c Molecular Virology and Gene Therapy, Molecular Medicine Katholieke Universiteit Leuven and IRC KULAK Kapucijnenvoer 33, B-3000 Leuven, Flanders, Belgium

ARTICLE INFO

Article history:

Received 21 June 2010

Revised 24 August 2010

Accepted 27 August 2010

Available online xxxxx

Keywords:

HIV-1 integrase

LEDGF/p75

Protein–protein interactions

Molecular docking

GRID

ABSTRACT

The search of small molecules as protein–protein interaction inhibitors represents a new attractive strategy to develop anti-HIV-1 agents. We previously reported a computational study that led to the discovery of new inhibitors of the interaction between enzyme HIV-1 integrase (IN) and the nuclear protein lens epithelium growth factor LEDGF/p75.¹

Herein, we describe new findings about the binding site of LEDGF/p75 on IN employing a different computational approach. In this way further structural requirements, helpful to disrupt LEDGF/p75–IN binding, have been identified. The main result of this work was the exploration of a relevant hydrophobic region. So we planned the introduction of suitable and simple chemical modifications on our previously reported ‘hit’ and the new synthesized compounds were subjected to biological tests.

The results obtained demonstrate that the hydrophobic pocket could play a key role in improving inhibitory efficacy thus opening new suggestions to design active ligands.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

To date, 25 compounds for the treatment of HIV infections have been formally licensed. They belong to different categories: reverse transcriptase inhibitors (RTIs), protease inhibitors (PIs); fusion inhibitors (FIs), coreceptor inhibitors (CRIs), and integrase inhibitors (INIs).² However, even the combination of these drugs is not sufficient enough to stop the progression of AIDS partly due to the emergence of drug-resistant HIV-1 mutants as well as the severe side effects.^{3–5} Therefore, alternative strategies need to be considered and new targets need to be identified to combat the growing AIDS pandemic. Currently, small-molecule inhibitors of the direct protein–protein interactions (PPIs), that mediate important biological processes, are an emerging and interesting area in drug design. In particular, there is evidence that various proteins play auxiliary roles in the HIV-1 life cycle and that some host cellular cofactors specifically control the integration process of viral DNA. The disruption of these essential cofactor–IN protein–protein interactions could represent a new frontier to design and develop novel anti-HIV-1 agents for anti-AIDS therapy.⁶

Amongst these cofactors, the ubiquitous lens epithelium growth factor LEDGF/p75 seems to play a pivotal role in the transcriptional

regulation, as well as in cell survival and in prevention of apoptosis. In particular, cells lacking of LEDGF/p75 showed a severe defect in HIV-1 infection characterized by decreased levels of integrated viral DNA.^{7,8} This evidence pointed out the possibility that LEDGF/p75 could be a promising target to inhibit DNA integration of HIV-1 life cycle. The LEDGF/p75 acts through a tethering mechanism as a potent cofactor for HIV-1 integration with the N-terminal Pro-Trp-Trp-Pro (PWWP) and A/T-hook elements binding to chromatin, and a C-terminal integrase-binding domain (IBD) binding to IN.^{9,10}

It is well known that these two regions of LEDGF/p75 are crucial for cofactor activity.

The specific protein–protein interaction between HIV-1 IN catalytic core domain (CCD) and IBD has been well described by X-ray analysis (Protein Data Bank file 2B4J) as well as by site-directed mutation studies, and it is characterized by IBD residues Ile365, Asp366, Phe406, and Val408.¹¹ In a previous paper, we focused our interest on key residues Ile365–Asp366 and used a computational approach to describe the pharmacophoric requirements needed to disrupt the interaction between HIV-1 IN CCD and LEDGF/p75 IBD.¹ This study led to the identification of some small molecules mimicking the Ile–Asp dipeptide; the most active molecule was the 2-hydroxy-4-(4-hydroxy-1*H*-indol-3-yl)-4-oxobut-2-enoic acid (CHIBA-3003, **I**, Fig. 1) that inhibited the IN–LEDGF/p75 interaction showing IC₅₀ value of 35 μM in AlphaScreen assay.

We therefore considered that CHIBA-3003 was a promising ‘hit compound’ for the development of new and more potent inhibitors

* Corresponding author. Tel.: +39 0906766464; fax: +39 0906766402.

E-mail address: ldeluca@unime.it (L. De Luca).

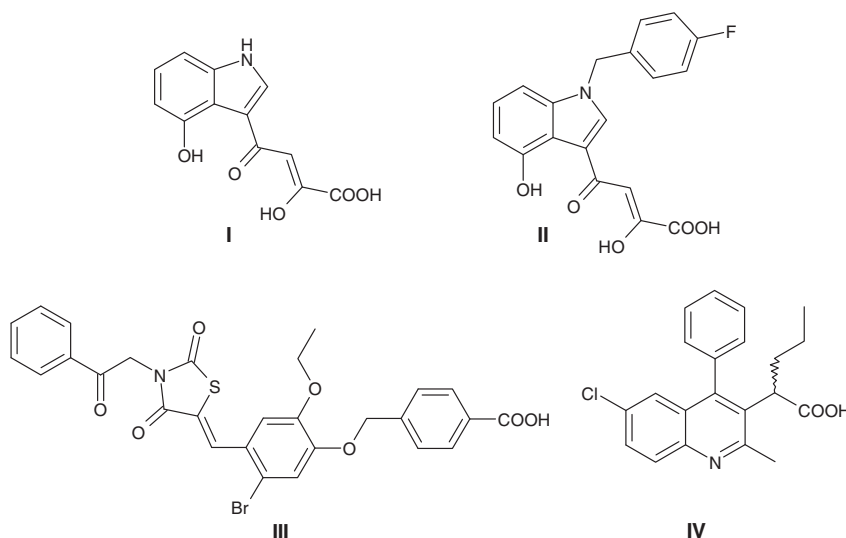


Figure 1. Chemical structures of CHIBA-3003 (I), CHIBA-3000 (II), D77 (III), and IV.

preventing IN-LEDGF/p75 interaction. Considering that very few inhibitors of this interaction are known and with the aim of achieving other information about the IN/LEDGF protein interaction and identify new important chemical frames improving the inhibitory potency of CHIBA-3003, in this study we have explored new binding pocket regions through GRID molecular interaction fields (MIFs) and synthesized several new small molecules targeting this PPI. Furthermore, docking experiments furnished new suggestions about the putative interactions within the binding site of these inhibitors, which hopefully could advance into novel anti-HIV drug targets.

2. Results and discussion

2.1. Rational design

It is well known that a typical protein–protein interaction occurs between the IBD of LEDGF/p75 and the integrase catalytic core domain (INCCD). The IBD of LEDGF/p75 forms a specific interaction with the binding pocket at the interface of two monomers of the INCCD. The IBD residues, involved in the interaction, are located in its interhelical loop regions and 4 relevant amino acid residues, Ile365, Asp366, Phe406, and Val408, are considered essential for the direct contact with INCCD residues.¹¹ The Ile365 and Asp366 hot-spot residues are contained on an octapeptide that fits into a HIV-IN binding pocket. Previous studies suggested that these two residues provide a binding site for small molecules as inhibitors of this protein–protein interaction. On this basis, we have recently disclosed that some small molecules containing indole nucleus could mimic the Ile365/Asp366 dipeptide thus inhibiting the interaction between IN and its cofactor LEDGF/p75.¹ The most active compound named CHIBA-3003 maps this binding site well, engaging interactions with some crucial residues of HIV-1 IN. In particular, we found, by docking experiments, that CHIBA-3003 (I) established H-bond interactions with Glu-170 and His-171 while the fused benzene ring of the indole system occupied an hydrophobic region formed by Leu102, Ala128 and Ala129.¹ Recently, other researchers identified, by docking experiments and site-directed mutagenesis, the benzoic acid derivative D77 as IN-LEDGF/p75 inhibitor (III, Fig. 1) and suggested that other hydrophobic regions could play a role in the IN/IBD interaction, in particular the area

corresponding to the IN residue Trp131.¹⁰ Furthermore it was highlighted that the importance of IBD aminoacids Phe406 and Val408 was correlated to their capability to contact and occlude solvent from the exposed IN residue Trp131.¹¹ In an effort to explore new regions of IN/LEDGF protein interaction, we studied the HIV-1 INCCD surface using molecular interaction fields (MIFs) calculated by the GRID program¹² and on the basis of structural information retrieved from IN/LEDGF crystal complex (PDB: 2 B4J).¹¹ The GRID program initially builds a grid over the molecular region of interest and then calculates the interaction energies between the protein and several molecular probes, at each grid point. The GRID probes incorporate different physicochemical properties and are able to mimic most of the atom types and small moieties commonly found in ligands. The set of energies calculated for a given probe constitutes the MIFs of the probe–target system. The probes selected for this study were the DRY probe representing steric and hydrophobic interactions; the N1 (amide nitrogen) probe to represent hydrogen-bond-donor groups; and the O (sp² carbonyl oxygen) probe to represent hydrogen-bond-acceptors. The results of the MIF calculations are shown in Figure 2.

We compared these MIFs with the previous docking results of our hit compound CHIBA-3003 in IN CCD. As shown in Figure 3A, CHIBA-3003 well mapped the most large areas described by N1 (magenta) and O (green) probes.

The CHIBA-3003 carboxylate group was able to interact with Glu-170 and His-171, as well as overlapping some areas described by N1 and O probes. We also observed that the most significant N1 area was filled by diketoacid portion. Moreover, the hydroxyl group at C-4 position of CHIBA-3003 formed a hydrogen-bond with the backbone carbonyl group of Gln168 whereas the benzene fused-ring of indole nucleus of CHIBA-3003 is placed into IN hydrophobic pocket well described by a DRY map. Nevertheless, we found that the most relevant area calculated by DRY probe (cyan and indicated by a red arrow in Fig. 3A) was vacant. This area corresponds to the hydrophobic region located near the residue Trp131 in the IN B chain. The analysis of these data prompted us to reconsider a class of CHIBA-3003 analogues containing hydrophobic moieties that could possibly occupy this large and relevant area described by the DRY probe.

In fact for the low activity of some derivatives and the lack in our pharmacophore model of this important hydrophobic area discovered by GRID we had previously supposed that the benzyl

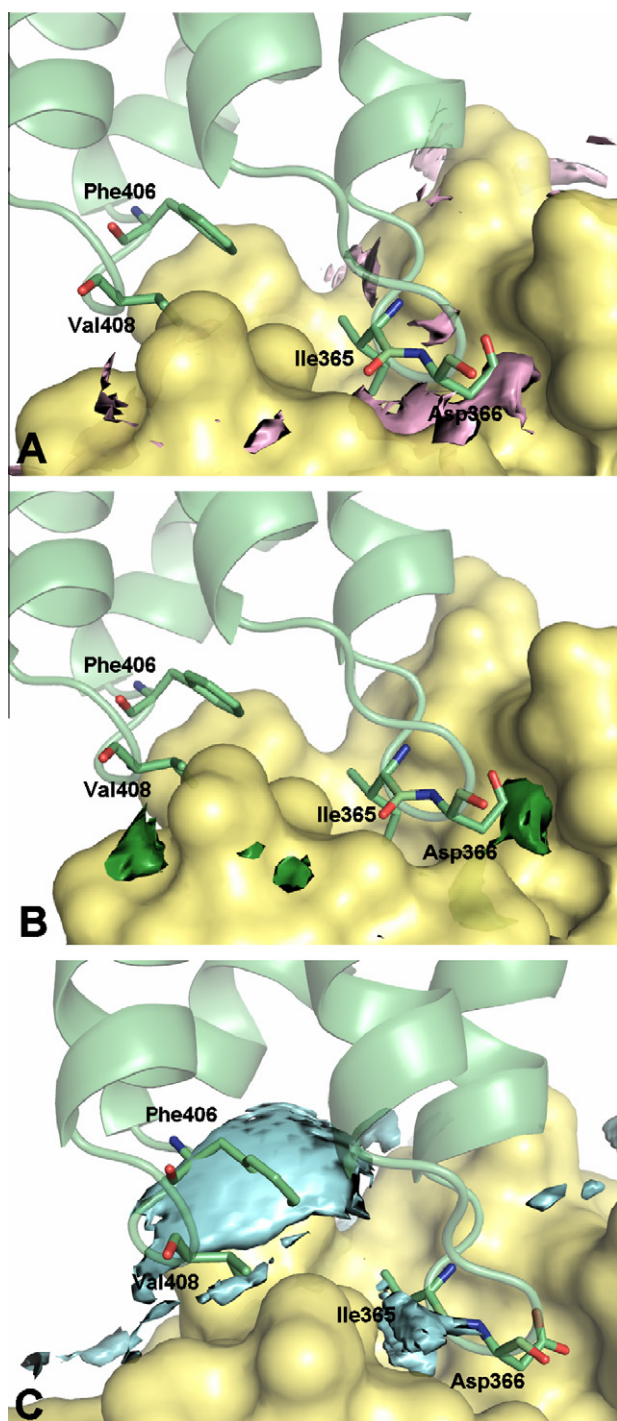


Figure 2. Molecular interaction fields MIFs calculated in the HIV-1 INCCD (surface, yellow) obtained by running GRID with selected probes. (a) N1 (−6 kcal/mol, magenta), (b) O (−6 kcal/mol, green), and (c) DRY (−0.5 kcal/mol, cyan) are viewed together with the IDB LEDGF/p75 portion (green cartoon). The IDB residues essential for the HIV-1 interaction (Ile365, Asp366, Phe406, Val408) are drawn in stick to visualize that the interaction fields correspond to actual atomic positions. This figure was prepared using the PyMOL program.¹⁹

group was not so important for the inhibition of PPIs.¹ Now, our docking experiments have allowed us to observe that the molecules benzylsubstituted such as CHIBA-3000, **II** (Figs. 1, 3B and Table 1) show a different binding mode when compared with N-unsubstituted analogue CHIBA-3003 (Fig. 3A) and have suggested the introduction of some hydrophobic groups (i.e., Me₂C₆H₃) suitable to improve the potency of CHIBA derivatives.

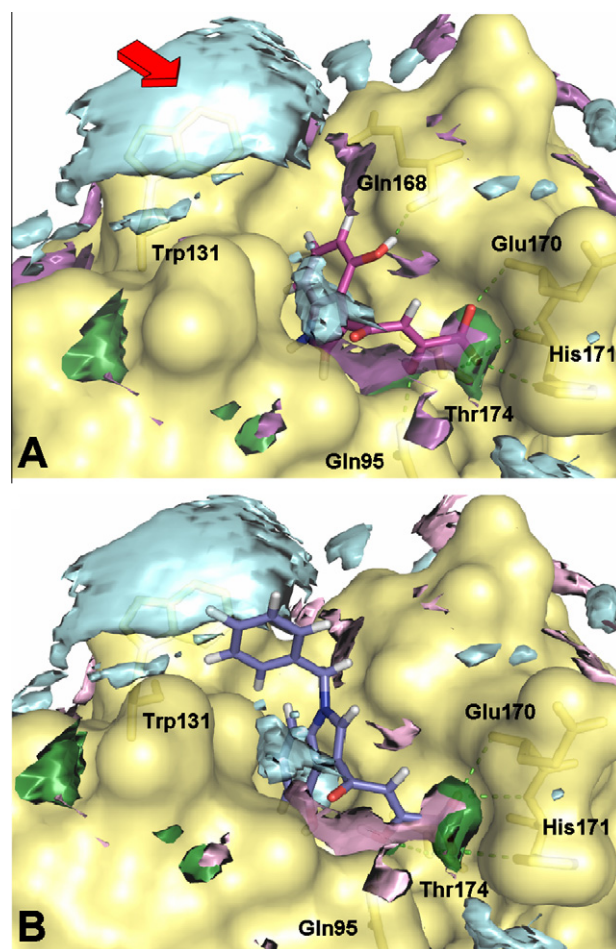


Figure 3. GRID MIFs calculated for the interaction between HIV-1 INCCD interface and the docking pose of CHIBA-3003 (I, Fig. 3A) and CHIBA-3000 (II, Fig. 3B). The magenta, green, and cyan contours correspond to the molecular interaction fields calculated for the selected probes (N1, O, and DRY). This figure was prepared using the PyMOL program.¹⁹

2.2. Chemistry and biological activity

In order to obtain more information on this hydrophobic portion, we synthesized and tested a small library of new *N*-benzyl-substituted indoles (**4a–h**, **5a–h**, and **6a–h**) in which we selected the mono- and dimethylbenzyl fragment as lipophilic moiety. The synthesis of the new compounds was accomplished following our previous reported procedures,¹⁰ employed to prepare indole derivatives (Scheme 1).

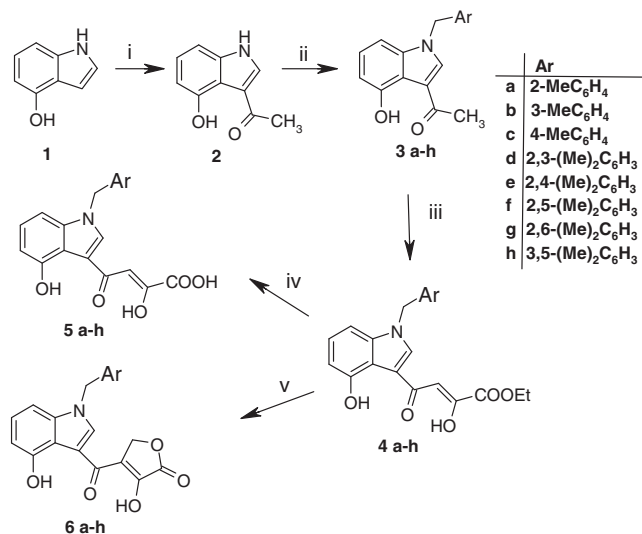
The experimental details are reported in Section 4. All synthesized compounds were tested in an AlphaScreen assay. This is a robust biochemical assay for identification of HIV integrase inhibitors, targeting the interaction between HIV-IN and the cellular cofactor LEDGF/p75 which has been developed for promoting the identification of novel inhibitors specific for this protein–protein interaction effectively.¹² The results of the screening are shown in Table 1. We tested the precursor ethyl esters (**4a–h**) as well as the corresponding diketoacid derivatives (**5a–h**); furthermore we evaluated compounds **6a–h** in which the diketoacid (DKA) portion was transformed into the hydroxyfuranone moiety which can be considered as a closed form of DKA function. All tested compounds produced inhibitory effects at 100 μM concentration with a percentage ranging from 16% to 96%. The analysis of available IC₅₀ values (see Table 1) pointed out that the ester moiety negatively affected the efficacy of this series of PP inhibitors

Table 1

Compound	% Inhibition at 100 μ M	IC ₅₀ ^a (μ M)
CHIBA-3003 (I)	71	35.0
CHIBA-3000	56	76
4a	24	•
4b	45	•
4c	60	170.7
4d	57	•
4e	58	•
4f	33	•
4g	47	•
4h	52	•
5a	20	•
5b	16	•
5c	86	68.0
5d	93	28.5
5e	55	•
5f	40	•
5g	72	•
5h	94	3.5
6a	34	•
6b	61	53.3
6c	81	8.7
6d	70	•
6e	96	11.7
6f	38	•
6g	70	7.5
6h	69	122.8

^a Concentration required to inhibit the HIV-1 IN-LEDGF/p75 interaction by 50%.

• Not determined.



Scheme 1. Reagents and conditions: (i) POCl₃, CH₃CON(CH₃)₂, 12 h, rt; (ii) appropriate methylbenzyl chloride or bromide, K₂CO₃, DMF, 10 min, 100 °C, 100 W; (iii) diethyl oxalate, dry CH₃ONa, THF, two separate steps under the same conditions: 2 min 50 °C, 250 W; (iv) NaOH 2 N, MeOH, 90 min, rt; (v) CH₂O, 1.5 h rt.

and that the most active molecules were the derivatives **5h**, **6c**, **6e**, and **6g**, active at low micromolar concentration. In particular the most potent derivative of the series (**5h**) was 10-fold more efficacious than CHIBA-3003 (3.5 μ M vs 35.0 μ M).

To investigate the recognition process we docked the synthesized compounds into the region of protein–protein interaction between HIV-1 IN and cofactor LEDGF/p75. All obtained complexes were subsequently optimized by a minimization procedure. Figure 4 shows a plausible binding mode of the most interesting derivative **5h** which displays the following interactions: (a) the carboxylate group was able to interact with Glu-170, His-171, and Thr174;

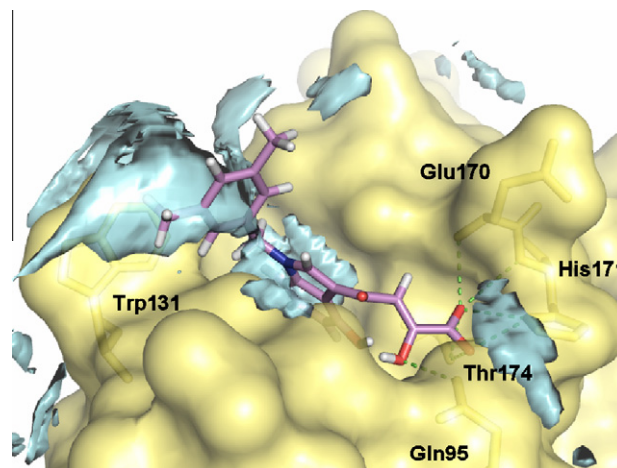


Figure 4. Docking pose of **5h** after minimization procedures and GRID MIF calculated for select probe DRY (cyan contours). This figure was prepared using the PyMOL program.¹⁹

(b) the hydroxyl group of the diketoacid moiety interacts with Gln95; (c) the benzene fused-ring of indole nucleus is placed into IN hydrophobic pocket; (d) the 3,5-dimethylbenzyl N-substituent occupies the most relevant area calculated by DRY probes.

The highest inhibitory efficacy of compound **5h** could be explained for its ability to well map the DRY area related to the crucial residue Trp131.

While our studies were in progress, the crystal structures of other small-molecule inhibitors active as PPIs in LEDGF/p75-binding pocket were solved.¹³ The most active reported compound was derivative **IV** (Fig. 1) that showed a biological activity of 1.37 μ M: the comparison of the docking result of our most active compound **5h** (3.5 μ M) and the X-ray position of compound **IV** is showed in Figure 5.

These two compounds show comparable activity and present a similar binding mode. In fact their carboxylate groups are in the same position and the indole nucleus of **5h** and the phenyl group of **IV** occupy the same hydrophobic area. A different orientation is observed between the dimethylphenyl moiety of **5h** and the benzene fused-ring containing a chlorine atom of **IV** which interacts with two different hydrophobic spaces. These observations could be particular useful for the optimization of these two important series of new integrase-LEDGF/p75 interaction inhibitors.

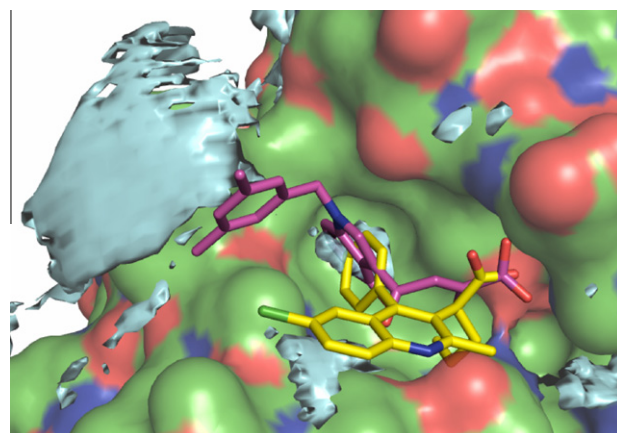


Figure 5. Docking pose of compound **5h** (magenta), compared to the crystallized position of **IV** (yellow) into the LEDGF/p75 binding pocket of IN-core and GRID MIF calculated for select probe DRY (cyan contours). This figure was prepared using the PyMOL program.¹⁹

3. Conclusion

In conclusion the importance of an additional lipophilic feature for IN-LEDGF/p75 inhibition has been suggested following the identification of a new hydrophobic region in IBD binding site. Novel small molecules active as PPIs have been designed, synthesized and tested. The obtained information could be helpful for the discovery of new inhibitors of this emerging and promising PPI target particularly important for anti-AIDS drug design.

4. Experimental section

4.1. Molecular modeling

Both the GRID analysis and docking were performed using IN structure retrieved by X-ray crystallographic of the dimeric catalytic core domain of HIV-1 IN complexed with LEDGF/p75 IBD deposited in the RCSB Protein Data Bank (entry code 2 B4J).¹¹

4.2. GRID calculations

The calculations were performed with version 22 of the GRID software.¹⁴ The GRID box dimensions were chosen to encompass all residues able to interact with IBD; and the NPLA (Number of Planes per Å for the grid box) was set to 2. The probes applied were DRY probes to describe the potential hydrophobic interactions, sp² carbonyl oxygen (O), and the neutral flat amino (N1) probes to evaluate the hydrogen-bond-donor and acceptor capacity of the target, respectively.

4.3. Molecular docking experiments

Hydrogen atoms were added to IN protein using the Biopolymer module in SYBYL 8.0.¹⁵ The structures of the ligands were constructed using standard bond lengths and angles from the SYBYL 8.0 fragment library and were fully optimized by the semiempirical quantum mechanical method AM1. Coordinates for the LEDGF dipetide Ile365–Asp366 were taken directly from the C chain of the PDB 2 B4J crystal structure and used as a control of program performance. Docking studies were performed using CCDC's GOLD (Genetic Optimization for Ligand Docking) software package version 3.1.1.¹⁶ ChemScore was chosen as a 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. All single bonds were treated as rotatable. Results differing by less than 1.5 Å in ligand-all atom RMSD, were clustered together. A 15.0 Å radius active site was drawn on the original position of the LEDGF/p75 IBD dipeptide I365–D366 and automated cavity detection was used.

4.4. Minimization process

All resulting docked complexes were subjected to energy minimization using 1000 steps of SD followed by 1000 steps of CG.

4.5. Chemistry

All microwave-assisted reactions were carried out in a CEM Focused Microwave Synthesis System, *Model Discover*, working at the power necessary for refluxing under atmospheric conditions (i.e., 250–300 W). Melting points were determined on a BUCHI Melting Point B-545 apparatus and have not been corrected. Elemental

analyses (C, H, N) were carried out on a Carlo Erba Model 1106 Elemental Analyzer and the results are within $\pm 0.4\%$ of the theoretical values. Merck silica gel 60 F₂₅₄ plates were used for analytical TLC; column chromatography was performed on Merck silica gel 60 (230–400 mesh) and Flash Chromatography (FC) on a Biotage SP1 EXP. ¹H NMR spectra were recorded in CDCl₃ with TMS as internal standard or DMSO-*d*₆ on a Varian Gemini-300 spectrometer. Chemical shifts are expressed in δ (ppm) and coupling constants (*J*) in hertz.

Compound **2** was prepared following a previously reported procedure and spectroscopic data are in accordance with the literature.¹

The synthesis of new compounds **3a–h**, **4a–h**, and **5a–h** was achieved using previously published procedures.¹

4.5.1. 3-Acetyl-1-(2-methyl-benzyl)-4-hydroxy-1H-indole (3a)

Compound **3a**: (87%) mp 130 °C dec; ¹H NMR (300 MHz, CDCl₃): δ = 2.34 (s, 3H, CH₃), 2.52 (s, 3H, CH₃), 5.46 (s, 2H, CH₂), 6.50–7.25 (m, 7H, ArH), 8.50 (s, 1H, ArH), 11.57 (s, 1H, OH). Anal. Calcd for C₁₈H₁₇FNO₂: C, 77.40; H, 6.13; N, 5.01. Found: C, 77.64; H, 6.42; N, 5.13.

4.5.2. 3-Acetyl-1-(3-methyl-benzyl)-4-hydroxy-1H-indole (3b)

Compound **3b**: (78%) mp 135–137 °C; ¹H NMR (300 MHz, CDCl₃): δ = 2.33 (s, 3H, CH₃), 2.53 (s, 3H, CH₃), 5.26 (s, 2H, CH₂), 6.71–7.23 (m, 7H, ArH), 7.66 (s, 1H, ArH), 11.53 (s, 1H, OH). Anal. Calcd for C₁₈H₁₇FNO₂: C, 77.40; H, 6.13; N, 5.01. Found: C, 77.22; H, 6.01; N, 5.32.

4.5.3. 3-Acetyl-1-(4-methyl-benzyl)-4-hydroxy-1H-indole (3c)

Compound **3c**: (73%) mp 129–131 °C; ¹H NMR (300 MHz, CDCl₃): δ = 2.35 (s, 3H, CH₃), 2.52 (s, 3H, CH₃), 5.25 (s, 2H, CH₂), 6.71–7.20 (m, 7H, ArH), 7.65 (s, 1H, ArH), 11.52 (s, 1H, OH). Anal. Calcd for C₁₈H₁₇FNO₂: C, 77.40; H, 6.13; N, 5.01. Found: C, 77.63; H, 6.41; N, 5.31.

4.5.4. 3-Acetyl-1-(2,3-dimethyl-benzyl)-4-hydroxy-1H-indole (3d)

Compound **3d**: (90%) mp 166–168 °C; ¹H NMR (300 MHz, CDCl₃): δ = 2.17 (s, 3H, CH₃), 2.34 (s, 3H, CH₃), 2.46 (s, 3H, CH₃), 5.25 (s, 2H, CH₂), 6.73–7.23 (m, 6H, ArH), 7.45 (s, 1H, ArH), 11.56 (s, 1H, OH). Anal. Calcd for C₁₉H₁₉FNO₂: C, 77.79; H, 6.53; N, 4.77. Found: C, 77.96; H, 6.72; N, 5.02.

4.5.5. 3-Acetyl-1-(2,4-dimethyl-benzyl)-4-hydroxy-1H-indole (3e)

Compound **3e**: (74%) mp 139–141 °C; ¹H NMR (300 MHz, CDCl₃): δ = 2.19 (s, 3H, CH₃), 2.32 (s, 3H, CH₃), 2.52 (s, 3H, CH₃), 5.21 (s, 2H, CH₂), 6.68–7.19 (m, 6H, ArH), 7.54 (s, 1H, ArH), 11.13 (s, 1H, OH). Anal. Calcd for C₁₉H₁₉FNO₂: C, 77.79; H, 6.53; N, 4.77. Found: C, 77.61; H, 6.69; N, 4.92.

4.5.6. 3-Acetyl-1-(2,5-dimethyl-benzyl)-4-hydroxy-1H-indole (3f)

Compound **3f**: (86%) mp 187–140 °C; ¹H NMR (300 MHz, CDCl₃): δ = 2.24 (s, 3H, CH₃), 2.27 (s, 3H, CH₃), 2.47 (s, 3H, CH₃), 5.19 (s, 2H, CH₂), 6.73–7.23 (m, 6H, ArH), 7.46 (s, 1H, ArH), 11.56 (s, 1H, OH). Anal. Calcd for C₁₉H₁₉FNO₂: C, 77.79; H, 6.53; N, 4.77. Found: C, 78.02; H, 6.38; N, 4.51.

4.5.7. 3-Acetyl-1-(2,6-dimethyl-benzyl)-4-hydroxy-1H-indole (3g)

Compound **3g**: (61%) mp 170–172 °C; ¹H NMR (300 MHz, CDCl₃): δ = 2.28 (s, 6H, CH₃), 2.35 (s, 3H, CH₃), 5.21 (s, 2H, CH₂), 6.76–7.31 (m, 7H, ArH), 11.60 (s, 1H, OH). Anal. Calcd for C₁₉H₁₉FNO₂: C, 77.79; H, 6.53; N, 4.77. Found: C, 77.62; H, 6.74; N, 4.91.

4.5.8. 3-Acetyl-1-(3,5-dimethyl-benzyl)-4-hydroxy-1H-indole (3h)

Compound **3h**: (89%) mp 167–169 °C; ¹H NMR (300 MHz, CDCl₃): δ = 2.29 (s, 6H, CH₃), 2.53 (s, 3H, CH₃), 5.21 (s, 2H, CH₂), 6.71–7.20 (m, 6H, ArH), 7.65 (s, 1H, ArH), 11.54 (s, 1H, OH). Anal. Calcd for C₁₉H₁₉NO₂: C, 77.79; H, 6.53; N, 4.77. Found: C, 77.53; H, 6.45; N, 4.96.

4.5.9. Ethyl 4-[1-(2-methyl-benzyl)-4-hydroxy-1H-indol-3-yl]-2-hydroxy-4-oxobut-2-enoate (4a)

Compound **4a**: (90%) mp 271 °C dec; ¹H NMR (300 MHz, DMSO-*d*₆): δ = 1.21 (t, *J* = 7.26, 3H, CH₃), 2.24 (s, 3H, CH₃), 4.08 (q, *J* = 7.26, 2H, CH₂), 5.25 (s, 2H, CH₂), 5.35–7.96 (m, 9H, ArH and CH), 14.33 (s, 1H, OH). Anal. Calcd for C₂₂H₂₁NO₅: C, 69.65; H, 5.58; N, 3.69. Found: C, 69.74; H, 5.71; N, 3.48.

4.5.10. Ethyl 4-[1-(3-methyl-benzyl)-4-hydroxy-1H-indol-3-yl]-2-hydroxy-4-oxobut-2-enoate (4b)

Compound **4b**: (88%) mp 268 °C dec; ¹H NMR (300 MHz, DMSO-*d*₆): δ = 1.21 (t, *J* = 7.26, 3H, CH₃), 2.24 (s, 3H, CH₃), 4.08 (q, *J* = 7.26, 2H, CH₂), 5.27 (s, 2H, CH₂), 5.32–7.92 (m, 9H, ArH and CH), 14.21 (s, 1H, OH). Anal. Calcd for C₂₂H₂₁NO₅: C, 69.65; H, 5.58; N, 3.69. Found: C, 69.32; H, 5.41; N, 3.73.

4.5.11. Ethyl 4-[1-(4-methyl-benzyl)-4-hydroxy-1H-indol-3-yl]-2-hydroxy-4-oxobut-2-enoate (4c)

Compound **4c**: (94%) mp 280 °C dec; ¹H NMR (300 MHz, DMSO-*d*₆): δ = 1.23 (t, *J* = 7.26, 3H, CH₃), 2.23 (s, 3H, CH₃), 4.09 (q, *J* = 7.26, 2H, CH₂), 5.26 (s, 2H, CH₂), 5.31–8.12 (m, 9H, ArH and CH), 14.54 (s, 1H, OH). Anal. Calcd for C₂₂H₂₁NO₅: C, 69.65; H, 5.58; N, 3.69. Found: C, 69.74; H, 5.47; N, 3.81.

4.5.12. Ethyl 4-[1-(2,3-dimethyl-benzyl)-4-hydroxy-1H-indol-3-yl]-2-hydroxy-4-oxobut-2-enoate (4d)

Compound **4d**: (82%) mp 205 °C dec; ¹H NMR (300 MHz, DMSO-*d*₆): δ = 1.22 (t, *J* = 7.26, 3H, CH₃), 2.22 (s, 6H, CH₃), 4.08 (q, *J* = 7.26, 2H, CH₂), 5.33 (s, 2H, CH₂), 5.38–7.99 (m, 8H, ArH and CH), 13.68 (s, 1H, OH). Anal. Calcd for C₂₃H₂₃NO₅: C, 70.22; H, 5.89; N, 3.56. Found: C, 70.39; H, 5.51; N, 3.86.

4.5.13. Ethyl 4-[1-(2,4-dimethyl-benzyl)-4-hydroxy-1H-indol-3-yl]-2-hydroxy-4-oxobut-2-enoate (4e)

Compound **4e**: (91%) mp 269 °C dec; ¹H NMR (300 MHz, DMSO-*d*₆): δ = 1.16 (t, *J* = 7.26, 3H, CH₃), 2.21 (s, 6H, CH₃), 4.08 (q, *J* = 7.26, 2H, CH₂), 5.27 (s, 2H, CH₂), 5.32–7.70 (m, 8H, ArH and CH), 14.56 (s, 1H, OH). Anal. Calcd for C₂₃H₂₃NO₅: C, 70.22; H, 5.89; N, 3.56. Found: C, 70.03; H, 5.71; N, 3.42.

4.5.14. Ethyl 4-[1-(2,5-dimethyl-benzyl)-4-hydroxy-1H-indol-3-yl]-2-hydroxy-4-oxobut-2-enoate (4f)

Compound **4f**: (83%) mp 298 °C dec; ¹H NMR (300 MHz, DMSO-*d*₆): δ = 1.16 (t, *J* = 7.26, 3H, CH₃), 2.11 (s, 3H, CH₃), 2.24 (s, 3H, CH₃), 4.09 (q, *J* = 7.26, 2H, CH₂), 5.28 (s, 2H, CH₂), 5.34–7.71 (m, 8H, ArH and CH), 14.56 (s, 1H, OH). Anal. Calcd for C₂₃H₂₃NO₅: C, 70.22; H, 5.89; N, 3.56. Found: C, 70.41; H, 5.62; N, 3.47.

4.5.15. Ethyl 4-[1-(2,6-dimethyl-benzyl)-4-hydroxy-1H-indol-3-yl]-2-hydroxy-4-oxobut-2-enoate (4g)

Compound **4g**: (70%) mp 199 °C; ¹H NMR (300 MHz, DMSO-*d*₆): δ = 1.19 (t, *J* = 7.26, 3H, CH₃), 2.20 (s, 6H, CH₃), 4.05 (q, *J* = 7.26, 2H, CH₂), 5.25 (s, 2H, CH₂), 5.28–7.20 (m, 8H, ArH and CH), 14.56 (s, 1H, OH). Anal. Calcd for C₂₃H₂₃NO₅: C, 70.22; H, 5.89; N, 3.56. Found: C, 70.41; H, 5.62; N, 3.47.

4.5.16. Ethyl 4-[1-(3,5-dimethyl-benzyl)-4-hydroxy-1H-indol-3-yl]-2-hydroxy-4-oxobut-2-enoate (4h)

Compound **4h**: (93%) mp 202 °C dec; ¹H NMR (300 MHz, DMSO-*d*₆): δ = 1.25 (t, *J* = 7.26, 3H, CH₃), 2.19 (s, 6H, CH₃), 4.11 (q, *J* = 7.26,

2H, CH₂), 5.22 (s, 2H, CH₂), 5.28–8.13 (m, 8H, ArH and CH), 14.63 (s, 1H, OH). Anal. Calcd for C₂₃H₂₃NO₅: C, 70.22; H, 5.89; N, 3.56. Found: C, 70.39; H, 5.64; N, 3.44.

4.5.17. 4-[1-(2-Methylbenzyl)-4-hydroxy-1H-indol-3-yl]-2-hydroxy-4-oxobut-2-enoic acid (5a)

Compound **5a**: (30%) mp 210 °C dec; ¹H NMR (300 MHz, DMSO-*d*₆): δ = 2.34 (s, 3H, CH₃), 5.49 (s, 2H, CH₂), 6.58–7.24 (m, 8H, ArH and CH), 8.91 (s, 1H, ArH), 11.30 (s, 1H, OH). Anal. Calcd for C₂₀H₁₇NO₅: C, 68.37; H, 4.88; N, 3.99. Found: C, 68.13; H, 4.59; N, 4.09.

4.5.18. 4-[1-(3-Methylbenzyl)-4-hydroxy-1H-indol-3-yl]-2-hydroxy-4-oxobut-2-enoic acid (5b)

Compound **5b**: (45%) mp 200 °C; ¹H NMR (300 MHz, DMSO-*d*₆): δ = 2.25 (s, 3H, CH₃), 5.42 (s, 2H, CH₂), 6.55–7.24 (m, 8H, ArH and CH), 9.07 (s, 1H, ArH), 11.25 (s, 1H, OH). Anal. Calcd for C₂₀H₁₇NO₅: C, 68.37; H, 4.88; N, 3.99. Found: C, 68.62; H, 4.63; N, 3.81.

4.5.19. 4-[1-(4-Methylbenzyl)-4-hydroxy-1H-indol-3-yl]-2-hydroxy-4-oxobut-2-enoic acid (5c)

Compound **5c**: (55%) mp 216–218 °C; ¹H NMR (300 MHz, DMSO-*d*₆): δ = 2.24 (s, 3H, CH₃), 5.40 (s, 2H, CH₂), 6.54–7.27 (m, 8H, ArH and CH), 9.06 (s, 1H, ArH), 11.25 (s, 1H, OH). Anal. Calcd for C₂₀H₁₇NO₅: C, 68.37; H, 4.88; N, 3.99. Found: C, 68.54; H, 4.99; N, 3.73.

4.5.20. 4-[1-(2,3-Dimethylbenzyl)-4-hydroxy-1H-indol-3-yl]-2-hydroxy-4-oxobut-2-enoic acid (5d)

Compound **5d**: (43%) mp 190 °C dec; ¹H NMR (300 MHz, DMSO-*d*₆): δ = 2.23 (s, 3H, CH₃), 2.26 (s, 3H, CH₃), 5.49 (s, 2H, CH₂), 6.47–7.11 (m, 7H, ArH and CH), 8.87 (s, 1H, ArH), 11.28 (s, 1H, OH). Anal. Calcd for C₂₁H₁₉NO₅: C, 69.03; H, 5.24; N, 3.83. Found: C, 69.32; H, 5.46; N, 3.68.

4.5.21. 4-[1-(2,4-Dimethylbenzyl)-4-hydroxy-1H-indol-3-yl]-2-hydroxy-4-oxobut-2-enoic acid (5e)

Compound **5e**: (56%) mp 235 °C dec; ¹H NMR (300 MHz, DMSO-*d*₆): δ = 2.20 (s, 3H, CH₃), 2.23 (s, 3H, CH₃), 5.29 (s, 2H, CH₂), 6.28–7.62 (m, 8H, ArH and CH), 10.14 (s, 1H, OH). Anal. Calcd for C₂₁H₁₉NO₅: C, 69.03; H, 5.24; N, 3.83. Found: C, 69.22; H, 5.56; N, 3.71.

4.5.22. 4-[1-(2,5-Dimethylbenzyl)-4-hydroxy-1H-indol-3-yl]-2-hydroxy-4-oxobut-2-enoic acid (5f)

Compound **5f**: (43%) mp 200 °C dec; ¹H NMR (300 MHz, DMSO-*d*₆): δ = 2.12 (s, 3H, CH₃), 2.26 (s, 3H, CH₃), 5.43 (s, 2H, CH₂), 6.56–7.12 (m, 7H, ArH and CH), 8.84 (s, 1H, ArH), 11.36 (s, 1H, OH). Anal. Calcd for C₂₁H₁₉NO₅: C, 69.03; H, 5.24; N, 3.83. Found: C, 69.18; H, 5.08; N, 3.98.

4.5.23. 4-[1-(2,6-Dimethylbenzyl)-4-hydroxy-1H-indol-3-yl]-2-hydroxy-4-oxobut-2-enoic acid (5g)

Compound **5g**: (43%) mp 199 °C dec; ¹H NMR (300 MHz, DMSO-*d*₆): δ = 2.21 (s, 6H, CH₃), 5.43 (s, 2H, CH₂), 6.58–7.21 (m, 7H, ArH and CH), 8.17 (s, 1H, ArH), 11.31 (s, 1H, OH). Anal. Calcd for C₂₁H₁₉NO₅: C, 69.03; H, 5.24; N, 3.83. Found: C, 68.92; H, 5.44; N, 3.62.

4.5.24. 4-[1-(3,5-Dimethylbenzyl)-4-hydroxy-1H-indol-3-yl]-2-hydroxy-4-oxobut-2-enoic acid (5h)

Compound **5h**: (40%) mp 170 °C dec; ¹H NMR (300 MHz, DMSO-*d*₆): δ = 2.20 (s, 6H, CH₃), 5.37 (s, 2H, CH₂), 6.55–7.22 (m, 7H, ArH and CH), 9.05 (s, 1H, ArH), 11.27 (s, 1H, OH). Anal. Calcd for C₂₁H₁₉NO₅: C, 69.03; H, 5.24; N, 3.83. Found: C, 68.89; H, 5.41; N, 3.65.

4.6. General procedure for the synthesis of 4-[1-(methylbenzyl)-4-hydroxy-1H-indol-3-oyl]-3-hydroxyfuran-2(5H)-ones (6a–h)^{17,18}

A solution of 40% aqueous formaldehyde in water (4 ml) was added to a mixture of **4a–h** (1 mmol) in diethyl ether (5 ml). The stirring was then continued until clear layers were formed (usually within 1–2 h). Sometimes, an additional 4 ml of water was added if the reaction was especially thick or when the solid appeared to react slowly. The clear, aqueous bottom layer was removed and the organic layer extracted twice with 5 ml of water. The combined aqueous extracts were cooled followed by acidification with 3 ml of concentrated hydrochloric acid. The corresponding furanone precipitate and the solution was then cooled overnight to ensure complete product formation. The resulting solid was collected, dried, and recrystallized from ethanol.

4.6.1. 4-[1-(2-Methylbenzyl)-4-hydroxy-1H-indol-3-oyl]-3-hydroxyfuran-2(5H)-one (6a)

Compound **6a**: (33%) mp 200 °C dec; ¹H NMR (300 MHz, DMSO-*d*₆): δ = 2.34 (s, 3H, CH₃), 4.96 (2H, CH₂), 5.43 (s, 2H, CH₂), 6.43–7.21 (m, 7H, ArH), 10.42 (s, 1H, ArH), 13.88 (s, 1H, OH). Anal. Calcd for C₂₁H₁₇NO₅: C, 69.41; H, 4.72; N, 3.85. Found: C, 69.57; H, 4.98; N, 3.71.

4.6.2. 4-[1-(3-Methylbenzyl)-4-hydroxy-1H-indol-3-oyl]-3-hydroxyfuran-2(5H)-one (6b)

Compound **6b**: (38%) mp 246–248 °C; ¹H NMR (300 MHz, DMSO-*d*₆): δ = 2.21 (s, 3H, CH₃), 4.89 (2H, CH₂), 5.35 (s, 2H, CH₂), 6.33–7.22 (m, 7H, ArH), 10.50 (s, 1H, ArH), 13.90 (s, 1H, OH). Anal. Calcd for C₂₁H₁₇NO₅: C, 69.41; H, 4.72; N, 3.85. Found: C, 69.21; H, 4.61; N, 3.98.

4.6.3. 4-[1-(4-Methylbenzyl)-4-hydroxy-1H-indol-3-oyl]-3-hydroxyfuran-2(5H)-one (6c)

Compound **6c**: (26%) mp 186 °C dec; ¹H NMR (300 MHz, DMSO-*d*₆): δ = 2.24 (s, 3H, CH₃), 5.00 (2H, CH₂), 5.38 (s, 2H, CH₂), 6.81–7.42 (m, 8H, ArH), 11.25 (s, 1H, OH). Anal. Calcd for C₂₁H₁₇NO₅: C, 69.41; H, 4.72; N, 3.85. Found: C, 69.33; H, 4.59; N, 3.72.

4.6.4. 4-[1-(2,3-Dimethylbenzyl)-4-hydroxy-1H-indol-3-oyl]-3-hydroxyfuran-2(5H)-one (6d)

Compound **6d**: (42%) mp 172 °C dec; ¹H NMR (300 MHz, DMSO-*d*₆): δ = 2.20 (s, 3H, CH₃), 2.25 (s, 3H, CH₃), 5.09 (2H, CH₂), 5.40 (s, 2H, CH₂), 6.48–7.20 (m, 6H, ArH), 8.69 (s, 1H, ArH), 11.39 (s, 1H, OH). Anal. Calcd for C₂₂H₁₉NO₅: C, 70.02; H, 5.07; N, 3.71. Found: C, 70.21; H, 5.23; N, 3.42.

4.6.5. 4-[1-(2,4-Dimethylbenzyl)-4-hydroxy-1H-indol-3-oyl]-3-hydroxyfuran-2(5H)-one (6e)

Compound **6e**: (42%) mp 184–186 °C; ¹H NMR (300 MHz, DMSO-*d*₆): δ = 2.23 (s, 3H, CH₃), 2.27 (s, 3H, CH₃), 5.07 (2H, CH₂), 5.42 (s, 2H, CH₂), 6.53–7.21 (m, 6H, ArH), 8.75 (s, 1H, ArH), 11.54 (s, 1H, OH). Anal. Calcd for C₂₂H₁₉NO₅: C, 70.02; H, 5.07; N, 3.71. Found: C, 70.38; H, 5.16; N, 3.59.

4.6.6. 4-[1-(2,5-Dimethylbenzyl)-4-hydroxy-1H-indol-3-oyl]-3-hydroxyfuran-2(5H)-one (6f)

Compound **6f**: (26%) mp 219–221 °C; ¹H NMR (300 MHz, DMSO-*d*₆): δ = 2.15 (s, 3H, CH₃), 2.26 (s, 3H, CH₃), 5.07 (2H, CH₂), 5.42 (s, 2H, CH₂), 6.54–7.19 (m, 6H, ArH), 8.71 (s, 1H, ArH), 11.47 (s, 1H, OH). Anal. Calcd for C₂₂H₁₉NO₅: C, 70.02; H, 5.07; N, 3.71. Found: C, 70.24; H, 5.19; N, 3.84.

4.6.7. 4-[1-(2,6-Dimethylbenzyl)-4-hydroxy-1H-indol-3-oyl]-3-hydroxyfuran-2(5H)-one (6g)

Compound **6g**: (20%) mp 235–237 °C; ¹H NMR (300 MHz, DMSO-*d*₆): δ = 2.24 (s, 6H, CH₃), 4.97 (2H, CH₂), 5.34 (s, 2H, CH₂),

6.58–7.22 (m, 6H, ArH), 8.18 (s, 1H, ArH), 11.61 (s, 1H, OH). Anal. Calcd for C₂₂H₁₉NO₅: C, 70.02; H, 5.07; N, 3.71. Found: C, 70.23; H, 5.29; N, 3.94.

4.6.8. 4-[1-(3,5-Dimethylbenzyl)-4-hydroxy-1H-indol-3-oyl]-3-hydroxyfuran-2(5H)-one (6h)

Compound **6h**: (31%) mp 202–204 °C; ¹H NMR (300 MHz, DMSO-*d*₆): δ = 2.21 (s, 6H, CH₃), 5.11 (2H, CH₂), 5.38 (s, 2H, CH₂), 6.51–7.23 (m, 6H, ArH), 8.96 (s, 1H, ArH), 11.58 (s, 1H, OH). Anal. Calcd for C₂₂H₁₉NO₅: C, 70.02; H, 5.07; N, 3.71. Found: C, 70.19; H, 5.19; N, 3.87.

4.7. LEDGF/p75-HIV-1 Integrase interaction screening (AlphaScreen technology)

The AlphaScreen assay was performed as described previously.¹² 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% (v/v) Tween-20 and 0.1% (w/v) bovine serum albumin. His₆-tagged integrase (300 nM final concentration) was incubated with the compounds at 4 °C for 30 min. The compounds were added in varying concentrations from 1 up to 100 μM. Afterward 100 nM of recombinant flag-LEDGF/p75 was added and incubation was extended by another hour at 4 °C. Subsequently, 5 μl of Ni-chelate-coated acceptor beads and 5 μl of anti-flag donor beads were added to a final concentration of 20 μg/ml 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.

Acknowledgment

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

References and notes

- De Luca, L.; Barreca, M. L.; Ferro, S.; Christ, F.; Iraci, N.; Gitto, R.; Monforte, A. M.; Debyser, Z.; Chimirri, A. *ChemMedChem* **2009**, *4*, 1311.
- De Clercq, E. *Med. Res. Rev.* **2009**.
- Barreca, M. L.; Iraci, N.; De Luca, L.; Chimirri, A. *ChemMedChem* **2009**, *4*, 1446.
- Nikolenko, G. N.; Delviks-Frankenberry, K. A.; Pathak, V. K. *J. Virol.* **2010**, *84*, 5238.
- Brogli, R.; Levy, Y.; Tiana, G. *Curr. Opin. Struct. Biol.* **2008**, *18*, 60.
- García-Rivera, J. A.; Bueno, M. T.; Morales, E.; Kugelman, J. R.; Rodriguez, D. F.; Llano, M. *J. Virol.* **2010**, *84*, 740.
- Busschots, K.; Voet, A.; De Maeyer, M.; Rain, J. C.; Emiliani, S.; Benarous, R.; Desender, L.; Debyser, Z.; Christ, F. *J. Mol. Biol.* **2007**, *365*, 1480.
- Greene, W. C.; Debyser, Z.; Ikeda, Y.; Freed, E. O.; Stephens, E.; Yonemoto, W.; Buckheit, R. W.; Este, J. A.; Cihlar, T. *Antiviral Res.* **2008**, *80*, 251.
- Hare, S.; Shun, M. C.; Gupta, S. S.; Valkov, E.; Engelman, A.; Cherepanov, P. *PLoS Pathog.* **2009**, *5*, e1000259.
- Du, L.; Zhao, Y.; Chen, J.; Yang, L.; Zheng, Y.; Tang, Y.; Shen, X.; Jiang, H. *Biochem. Biophys. Res. Commun.* **2008**, *375*, 139.
- Cherepanov, P.; Ambrosio, A. L.; Rahman, S.; Ellenberger, T.; Engelman, A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17308.
- Al-Mawsawi, L. Q.; Christ, F.; Dayam, R.; Debyser, Z.; Neamati, N. *FEBS Lett.* **2008**, *582*, 1425.
- Christ, F.; Voet, A.; Marchand, A.; Nicolet, S.; Desimmie, B. A.; Marchand, D.; Bardiot, D.; Van der Veken, N. J.; Van Remoortel, B.; Strelkov, S. V.; De Maeyer, M.; Chaltin, P.; Debyser, Z. *Nat. Chem. Biol.* **2010**, *6*, 442.
- Grid 22, Molecular Discovery Ltd, Pinner, Middlesex, UK.
- svbvl, Version 7.1; Tripos Inc.: 1699 South Hanley Road, Suite 303, St. Louis, MO 63144, 2005.
- Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.* **1997**, *267*, 727.
- Ferro, S.; Barreca, M. L.; De Luca, L.; Rao, A.; Monforte, A. M.; Debyser, Z.; Witvrouw, M.; Chimirri, A. *Arch. Pharm. (Weinheim)* **2007**, *340*, 292.
- Ferro, S.; De Grazia, S.; De Luca, L.; Barreca, M. L.; Debyser, Z.; Chimirri, A. *Heterocycles* **2009**, *78*, 947.
- The PyMOL Molecular Graphics System, W. L. DeLano, DeLano Scientific LLC, San Carlos, CA, USA, 2008; <http://www.pymol.org>.