

1,2,3,4-tetrahydroisoquinolines as inhibitors of HIV-1 integrase and human LEDGF/p75 interaction

Anu George^a, Alavala Gopi Krishna Reddy^b, Gedu Satyanarayana^b, Nidhanapati K Raghavendra^{a*}

^aDepartment of Biotechnology, ^bDepartment of Chemistry, Indian Institute of Technology Hyderabad, Kandi Sangareddy-502285, Telangana, India.

* Correspondence: E-mail: raghunk@iith.ac.in; Tel: +91-40-2301-6054; Fax: +91-40-2301-6032.

Key words: HIV-1 integrase; LEDGF/p75; interaction; isoquinoline; inhibition

Short Title: Isoquinoline based inhibitor of LEDGF/p75-IN

1,2,3,4-tetrahydroisoquinolines as inhibitors of HIV-1 integrase and human LEDGF/p75 interaction

Abstract

Alkaloids are a class of organic compounds with a wide range of biological properties, including anti-HIV activity. The 1,2,3,4-tetrahydroisoquinoline is a ubiquitous structural motif of many alkaloids. Using a short and an efficient route for synthesis, a series of 1,2,3,4-tetrahydroisoquinolines/isoquinolines was developed. These compounds have been analysed for their ability to inhibit an important interaction between HIV-1 integrase enzyme (IN) and human LEDGF/p75 protein (p75) which assists in the viral integration into the active genes. A lead compound **6d** is found to inhibit the LEDGF/p75-IN interaction *in vitro* with an IC₅₀ of ~10 μM. Molecular docking analysis of the isoquinoline **6d** reveals its interactions with the LEDGF/p75-binding residues of IN. Based on an order of addition experiment, the binding of **6d** or LEDGF/p75 to IN is shown to be mutually exclusive. Also, the activity of **6d** *in vitro* is found to be unaffected by the presence of a non-specific DNA. As reported earlier for the inhibitors of LEDGF/p75-IN interaction, **6d** exhibits a potent inhibition of both the early and late stages of HIV-1 replication. Compound **6d** differing from the known inhibitors in the chemical moieties and interactions with

CCD could potentially be explored further for developing small molecule inhibitors of LEDGF/p75-IN interaction having a higher potency.

Introduction

HIV-1 integrase (IN) catalyses the integration of a viral DNA (vDNA) into the host genomic DNA (gDNA) during the early stage of replication.^[1-4] Although IN can exist as a monomer, dimer, tetramer and a higher oligomer in the solution, it is the tetramer of IN that binds to the LTR (long terminal repeat) termini of the vDNA, forming an uncleaved intasome (UI). The UI catalyses a 3' end processing reaction resulting in the formation of cleaved intasome (CI) in the cytoplasm. The CI captures the gDNA to form a transient target capture complex (TCC) in the nucleus. The IN in the TCC catalyses the strand transfer reaction, covalently linking the 3' ends of the vDNA to the gDNA and forming a strand transfer complex (STC). HIV-1 preferentially integrates into the actively transcribing genes of the gDNA.^[5,6] An important co-factor of IN that is required for vDNA integration into the active genes is the human transcriptional co-activator LEDGF/p75 (lens epithelium-derived growth factor).^[7-11] Amino acid residues 347- 429 at the C-terminus of LEDGF/p75 termed as the integrase binding domain (IBD)^[12-14] binds to IN in the PIC (pre-integration complex), while the PWWP domain at the N-terminus interacts with the chromatin,^[15-18] making LEDGF/p75 as a bimodal tether for integration of HIV-1 into the active genes^[19,20]. Co-crystal of IBD and CCD (catalytic core domain) of IN reveals that two interhelical loops of IBD, comprised of the residues 364-368 and 405-408, respectively, interact with the LEDGF-binding site at the dimer interface of CCD of IN.^[13,21,22] The hydrophobic interaction between I365 of IBD and the hydrophobic pocket formed by L102, A128, A129, W132 of one monomer and T174, M178 of the other monomer of IN dimer, is one of the three critical interactions in the LEDGF/p75-IN complex. The hydrophobic interaction between W131 of IN and F406, V408 of IBD forms the second critical interaction. A bidentate H-bond is formed between the side chain carboxyl of D366 of IBD and the main chain amides of E170, H171 of IN. A D366N mutation in IBD abrogates the formation of a LEDGF/p75-IN complex, highlighting the importance of the third critical interaction. These critical interactions in the LEDGF/p75-IN complex are targeted by small molecule inhibitors for discovering anti-HIV drugs.

A halogen containing quinoline molecule, CX04328, was developed as the first LEDGIN (inhibitor of LEDGF/p75-IN interaction) and demonstrated to be effective at micro-molar concentration in inhibiting the interaction of LEDGF/p75 and IN *in vitro*, as well as, HIV-1 infection in a mammalian cell line.^[23] Later studies have demonstrated that LEDGINs such as BI-1001, GS-A, GS-B, CX14442, GSK1264 are allosteric IN inhibitors (ALLINIs) as they allosterically stabilize an inactive oligomeric form of IN and interfere with IN-vDNA assembly, while preformed IN-vDNA complexes are refractory.^[24-30] The ALLINIs present during the late phase of HIV-1 replication induce a premature multimerization of IN (a LEDGF/p75-independent effect) resulting in the production of less infectious virions that are defective for reverse transcription, nuclear import of PIC and integration.^[31-33] The LEDGINs such as GS-A and GS-B also qualify as non-catalytic site IN inhibitors (NCINIs) as they interfere with IN-vDNA assembly.^[24, 27] The LEDGF/p75-independent effects of these molecules are found to contribute to their higher potency for inhibiting HIV-1 replication.

In this study, starting with a series of 1,2,3,4-tetrahydroisoquinolines, a non-halogenated isoquinoline compound **6d** has been identified as a lead molecule for inhibiting LEDGF/p75-IN interaction. The chosen tetrahydroisoquinoline moiety is present in a number of isoquinoline based alkaloid natural products having antitumor, anti-microbial, anti-inflammatory, anti-analgesic and anti-HIV-1 activity.^[34-36] A short and efficient route for the synthesis of substituted 1,2,3,4-

tetrahydroisoquinolines was developed in an earlier study and shown to yield pure compounds suitable for biological evaluation.^[37,38] Following an initial screening of the series of tetrahydroisoquinolines, a lead compound **6d** with an IC₅₀ of ~10 μM for inhibiting *in vitro* interaction between recombinant LEDGF/p75 and IN was obtained. Compared to the LEDGF/p75-IN inhibitors reported earlier,^[23-30] **6d** differs in: having an isoquinoline rather than a quinoline core, lacking a halogen or sulphur atom and having an ester rather than a carboxyl moiety (supplementary Figure 1). As seen in the three dimensional structures of known LEDGINs bound to the CCD dimer,^[23,24] molecular docking analysis shows **6d** interacting with the LEDGF/p75 binding site of a CCD dimer, although the type of interactions is different. Both IN and LEDGF/p75 can bind to DNA in a sequence independent manner,^[39-43] however, the presence of a non-specific DNA did not alter the inhibitory activity of **6d** *in vitro*. Also, the pre-formed LEDGF/p75-IN complexes could not be disrupted by **6d** *in vitro*, indicating a requirement for **6d** to access the LEDGF/p75-binding site on IN for having an inhibitory effect. Interaction of LEDGF/p75 and IN in the cell lysates, as well as, the early and late stages of HIV-1 replication is efficiently inhibited by **6d**. Taken together, the unique features of **6d** described here could potentially be used for developing inhibitors of LEDGF/p75-IN interaction having a higher potency.

Materials and Methods

Synthesis of 1,2,3,4-tetrahydroisoquinoline compounds

The synthesis of 1,2,3,4-tetrahydroisoquinoline compounds **1a-1m** (supplementary figure 2) was performed as described earlier.^[37,38] The reduced ester moiety of **1a** (labelled **4a**) and rearranged aromatization of compounds **1a**, **1d** (labelled **6a**, **6d**) were synthesized as described in the supplementary material.

In vitro pull-down assay

The pull-down assay using a purified C-terminus His-tagged IN (His-IN) and a N-terminus GST tagged IBD (Integrase Binding Domain, GST-IBD) or a N-terminus GST tagged full length LEDGF/p75 (GST-p75) was performed as described earlier.^[44] Briefly, in a 150 μl of phosphate pull-down buffer (PPD buffer) having 50 mM Phosphate buffer (pH 6.5), 150 mM NaCl, 2 mM MgCl₂, 25 mM Imidazole and 0.1% NP-40, with or without indicated amount of a given compound, a 3 μg of His-IN was added and incubated at 37°C for 10 min. This was followed by addition of a 3 μg of GST-IBD or GST-p75 and incubation was continued at 37°C for another 10 min. The formed complexes were pull-down by binding to the Ni sepharose 6 Fast flow beads (GE Healthcare) for 30 min at room temperature. The bound complexes were eluted in the PPD buffer containing 400 mM Imidazole and resolved on a 10% SDS-PAGE gel. Intensity of the coomassie blue stained protein bands was quantified using an Image Lab Software 4.0 (Bio-Rad). Stock solution of all the compounds were prepared in DMSO. The required dilutions of the compounds were prepared in the PPD buffer just prior to the assay.

Pull-down of ectopically expressed proteins from a mammalian cell lysate

Pull-down of a Flag-tagged IN (Flag-IN) and a HA-tagged LEDGF/p75 (HA-p75) over-expressed in the HEK293T cell line was performed as described earlier,^[45,46] with slight modifications. Cells over-expressing either a Flag-IN or HA-p75, were lysed in a 300 μl of the lysis buffer (500 mM NaCl, 50 mM HEPES (pH 7.8), 0.5% Triton-X-100, 5% glycerol, 2 mM MgCl₂), to this 1.2 ml of a binding buffer (50 mM HEPES (pH7.8), 5% glycerol, 2 mM MgCl₂) with or without indicated amount of **6d** was added. The mixture was nutated at 4°C for 4 h to allow binding of the compound. This was followed by the addition of a cell lysate containing the other protein partner and the nutation was continued at 4°C for 12 h, for the complex to form. Anti-Flag beads (Clontech) were

added to the mixed lysates and nutated at 4°C for additional 2 h. Beads were washed with buffer containing 150 mM NaCl, 50 mM HEPES (pH7.8), 0.1% Triton-X-100, 0.1% NP-40, 5% glycerol and 2 mM MgCl₂ to remove any non-specific proteins. The proteins bound to the beads were eluted in a SDS-PAGE loading buffer and resolved on a 10% SDS-PAGE gel. The Flag-IN and HA-p75 bands were detected on a Western blot using anti-Flag (Clontech) and anti-HA (Thermo Scientific) antibodies, respectively.

Infection of pseudo-typed HIV-1

The production of NLX-Luc-VSVg virus, setting up of infection and measuring of luciferase activity at 48 h post-infection, in the HEK293T cells, were performed as described earlier,^[47] with slight modifications. The growth media on HEK293T cells containing the virus was collected at 48 h post-transfection and stored at -80°C till it is used for infecting the cells. For analysing the effect of **6d** on the late stage of replication, virus was generated by treating the HEK293T cells with **6d** for the duration of transfection. In experiments analysing inhibition of the early stage of replication, cells were treated with **6d** for 2 h prior to the addition of virus. To set up an infection using 70 µl of the virus, 0.08 – 0.12 million cells were plated per well of a 24-well plate at least 12 h in advance. At 48 h post-infection, cells were harvested and the cell lysates was prepared using a 1X passive lysis buffer (Promega, cat. No. E1941). The luciferase activity in the lysates was measured using a D-luciferin sodium salt (Invitrogen, cat no. L2912). The luciferase activity was normalized to the concentration of the protein determined using a BCA reagent (Pierce, cat. No. 23221). The normalized luciferase activity in untreated infected cells was taken as 100% and that in **6d** treated cells was determined accordingly. To measure the viability of cells treated with **6d**, addition of virus was omitted in the above infection protocol and at 48 h post-treatment with **6d**, a trypan blue exclusion method was used to determine the percent of live cells. The number of live cells in untreated wells was set as 100% and that in **6d** treated cells was determined accordingly.

Molecular docking analysis

Molecular docking of **6d** to CCD dimer of IN was performed using a Discovery Studio version 4.0 software (Accelrys, USA). The three dimensional structure of CCD of IN (PDB accession 1BL3)^[48] was used as a receptor and **6d** was used as a ligand, in the LibDock protocol. The known LEDGF/p75 interacting residues of IN^[12,13,21]: L102, A128, A129, W131, W132, E170, H171, T174 and M178 were used to define the interaction sphere for docking. The best docked pose was ranked based on the LibDock score and binding energy and was used for determining the interactions between **6d** and the CCD dimer.

Statistical analysis

Intensity of the bands corresponding to IBD or LEDGF/p75 over IN (ratio of IBD or LEDGF/p75 to IN) was calculated to obtain the extent of complex formation in the pull-down assays. The IBD:IN ratio in the sample without compound was set to 100% binding and in those having compound was calculated accordingly. The concentration of a compound giving a 50% inhibition of the complex formation is determined to be its IC₅₀. Data from at least three independent experiments has been used to obtain the standard error.

Results and Discussion

Tetrahydroisoquinolines can inhibit interaction between IBD and IN *in vitro*

The first series of 1,2,3,4-tetrahydroisoquinoline compounds **1a-1m** (supplementary Fig. 2) were tested *in vitro* for their ability to inhibit the interaction between His-IN and GST-IBD in the pull-down assay. A serial double dilution of the compounds with a starting concentration of 24 mM was used in the initial screening and the *in vitro* assay was performed as described in the methods.

The IC₅₀ of the compounds **1a-1m** for inhibiting the formation of IBD-IN complex is shown in Table 1. Of all the compounds tested in this series, **1a** is found to be the most potent as it has the least IC₅₀ of ~80 μM (Fig. 1A). Although **1h** has the second best IC₅₀, **1d** that differs from **1h** only in the presence of an ethyl rather than a *tert*-butyl group on the ester moiety, did not show any inhibition even at the highest concentration tested in the *in vitro* assay.

Increasing the potency of compounds by rearranged aromatization

With the aim of increasing the potency of the compounds in the first series, either a reducing or an oxidizing condition was used as an additional step to the synthesis scheme and those compounds with an appreciable purity were tested *in vitro*. Reducing of the ester group on the best compound **1a** (IC₅₀ ~80 μM) resulted in an alcohol containing derivative labelled as **4a** (scheme 2, supplementary material), that has an IC₅₀ greater than 24 mM, emphasizing the importance of an ester moiety for activity of the compound. Subjecting **1a** to an oxidative condition (scheme 3, supplementary material) resulted in a rearranged aromatization and significantly enhanced the potency of the derivative labelled as **6a**. Compound **6a** showed the inhibition of IBD-IN complex formation with an IC₅₀ of 20.2±8.1 μM (Fig. 1B). Interestingly, a similarly rearranged aromatization of **1d** (that lacks any inhibitory activity, Table 1) resulted in a lead molecule **6d** (ethyl-1-benzyl-7-methoxyisoquinoline-4-carboxylate) that is characterized further in this study. Compound **6d** inhibited the interaction between IBD and IN with an IC₅₀ of 10.4±3.8 μM (Fig. 1C) and the interaction between IN and full-length LEDGF/p75 with an IC₅₀ of 11.1±4.4 μM (Supplementary Fig 3). Thus, compared to **1d**, a benzyl group at the 1st position of an unsaturated bicyclic system dramatically enhanced the ability of **6d** for inhibiting the interaction between IBD and IN.

Molecular docking shows interaction of 6d with the critical residues of CCD

In order to gain insight into how **6d** interacts with IN, molecular docking of **6d** to CCD dimer of IN was performed as described in the methods. The interactions between **6d** and CCD are summarized in table 2. The results of docking analysis indicate that **6d** inhibits binding of IBD by sterically blocking access to the LEDGF/p75 binding site on IN (Fig. 1D), similar to the LEDGINs reported earlier. Co-crystal structures of the known LEDGINs with CCD^[23-26,30] reveal H-bonds between the carboxyl moiety of the inhibitor and the backbone amides of E170, H171 and the side chain hydroxyl group of T174, while **6d** has pi-donor H-bonds between the isoquinoline core and the side chain hydroxyl group of T174. In addition, methoxy and ester moieties of **6d** make carbon H-bond with the backbone amides of Q95 and A128, respectively. The quinoline core of known LEDGINs has a hydrophobic interaction with one face of the hydrophobic pocket of CCD dimer comprised of residues Q95, T124, T125, A128, while the isoquinoline core of **6d** has a hydrophobic interaction with A98, A169 and T174. The ethyl group (in the ester moiety) of **6d** is surrounded by the hydrophobic side chains of L102, W132 and M178. A pi-anion charge interaction is present between the benzyl group of **6d** and the side chain carboxyl group of E170. Overall, the binding of **6d** to CCD dimer while being similar to that of known LEDGINs, seems to offer sufficient differences in the type of interactions at the LEDGF/p75 binding site that could potentially be explored further. It has been shown earlier that a virus resistant to LEDGINs carries one or more of the following mutations: Y99H, L102F, T124N, A128T, A129T, H171T, T174I in CCD, with A128T being the most common resistance mutation.^[32,49,50] Owing to the overlap in the residues of CCD interacting with **6d** and known LEDGINs, similar mutations in IN can be expected to result in an emergence of resistance to **6d**.

Binding of IBD, 6d to IN is mutually exclusive

To gain a further insight into the details of inhibition by **6d**, IBD-IN or LEDGF/p75-IN complexes were formed in the absence of **6d** and in the absence or presence of a non-specific DNA and activity of **6d** on such complexes was analysed. First, displacing of IBD in the complex by **6d** would require accessing the same interacting residues on CCD, which might be feasible if **6d** can also bind at a secondary site on IN and induce conformational changes in CCD that reduce the affinity for IBD. Second, since both IN and LEDGF/p75 are non-specific DNA binding proteins,^[39-43] it is possible that a DNA bound protein has an altered affinity for **6d** or DNA by itself binds to **6d**, consequently influencing the activity of **6d**. Following the formation of a IBD-IN complex at 37°C for 10 min, 150 µM of **6d** (15 times the IC₅₀) was added and pull-down assay was performed as described in the methods. As can be seen in Fig. 2A, **6d** did not disrupt a pre-formed complex (lane 3) and the extent of pre-formed complex remained similar to that formed in the absence of **6d** (lane 1), while the interaction of IBD with a **6d** bound IN is completely inhibited (lane 2). These results indicate that the binding of IBD, **6d** to IN might be mutually exclusive due to an overlap at the site of interaction. Although the possibility of **6d** binding at a secondary site on IN is not ruled out, it nonetheless does not seem to alter the affinity for IBD. Incubating IN with a non-specific DNA (1 µg of plasmid pUC18) at 37°C for 10 min, prior to addition of **6d**, did not affect the ability of compound to inhibit IBD-IN interaction (lane 4). Formation of a IBD-IN complex in the presence of a non-specific DNA (lane 5) or addition of a non-specific DNA to a pre-formed IBD-IN complex (lane 6), prior to the addition of **6d**, did not result in the disruption of the complex. Similar results were obtained when IBD was replaced with a full length LEDGF/p75 (GST-p75) (Fig. 2B).

Reversing the order of addition of protein partners by incubating IBD with **6d**, prior to the addition of IN (Fig 2A, lane 7) or a non-specific DNA bound IN (lane 8), did not result in any inhibition by **6d**. In this case, it may be speculated that binding of **6d** to IBD at a site(s) other than the inter helical loops involved in the interaction with IN^[13,21,22] decreases the effective concentration of **6d** that is required to inhibit complex formation. Although full length LEDGF/p75 binds to a non-specific DNA,^[42,43] addition of **6d** to a DNA bound LEDGF/p75 prior to that of IN (Fig 2B, lane 9) or vice versa (lane 10), did not result in an inhibition of the LEDGF/p75-IN complex formation. These results indicate that neither the presence of a non-specific DNA nor the binding of DNA by any of the protein partners, at any stage of the complex formation, interferes with the inhibitory activity of **6d** *in vitro*.

Inhibition of interaction in the presence of cellular proteins

Prior to testing the ability of **6d** to inhibit HIV-1 infection, the inhibitory activity of **6d** on an interaction between the ectopically expressed proteins in a mammalian cell lysate was analysed. A lysate of HEK293T cells overexpressing either a Flag-IN or a HA-p75 was treated with 150 µM of **6d** prior to the addition of a lysate containing the other protein partner and the formation of LEDGF/p75-IN complex was analysed as described in the methods. Treating the lysate having Flag-IN with **6d**, prior to mixing with the lysate containing HA-p75 resulted in an 81±10% reduction in the amount of LEDGF/p75-IN complex formed (Fig. 3A, lane 1), as compared to that formed in the absence of **6d** (lane 3). Treating the lysate having HA-p75 with **6d**, prior to mixing with the lysate having Flag-IN resulted in a 71±12% reduction in the amount of LEDGF/p75-IN complex formed (lane 2). The inhibition of complex formation observed in Fig 3A, lane 2 (pre-treating cell lysate containing LEDGF/p75 with **6d**) is in contrast to a lack of inhibition shown in Fig. 2B, lanes 7-9 (pre-treating purified recombinant LEDGF/p75 with **6d**) and can be explained if the speculated secondary site on LEDGF/p75 for binding of **6d** is masked by other cellular proteins in the lysate and hence the effective concentration of **6d** remains unaltered. The results

shown in figures 2 and 3 indicate that **6d** can inhibit the interaction between IN and LEDGF/p75 inside the cell and can be expected to inhibit HIV-1 infection.

Inhibition of HIV-1 infection by 6d

LEDGINs are known to affect the early stage of HIV-1 replication by inhibiting assembly of IN-vDNA complex, as well as, the interaction between LEDGF/p75 and IN at the integration step and also affect the late stage of replication by stabilizing the aberrant multimerization of IN in the newly formed virions that consequently become less infectious.^[23-33] ALLINIs inhibit multiple steps of viral replication such as Prior to analysing the effect of **6d** on HIV-1 infection, the viability of HEK293T cells treated with 0, 12.5, 25, 50, 100, 150 and 200 μM of **6d** was measured as described in the methods. No cytotoxicity was observed up to 50 μM of **6d** (Fig. 3B, filled squares) however, a 30 ± 5 , 55 ± 7 and $63\pm 3\%$ drop in live cells was noticed in the samples treated with 100, 150 and 200 μM , respectively, giving a CC_{50} of ~ 140 μM . For analysing the effect of **6d** on the early stage of replication, the cells were treated with 0, 12.5, 25 and 50 μM of **6d** prior to the infection. The normalized luciferase activity in the cell lysate was used to determine the levels of infection. As shown in Fig. 3B, a 0, 19 ± 8 and $63\pm 7\%$ inhibition of infection was observed with 12.5, 25 and 50 μM concentration of **6d** (Fig. 3B, filled circles). Using the linear range of the inhibition curve (Fig 3B), an IC_{50} of ~ 42.5 μM for **6d** has been obtained. Based on these results, **6d** is determined to have a selectivity index ($\text{CC}_{50}/\text{IC}_{50}$) of 3.3. Infectivity of virus generated in the presence of 0, 12.5, 25, 50 μM of **6d** was measured to assess the inhibition of the late stage of replication. A noticeable increase in the inhibitory potency of **6d** could be seen as a 33 ± 2 , 41 ± 8 and $75\pm 7\%$ decrease in the infectivity of the virus generated in the presence of 12.5, 25 and 50 μM of **6d** was observed (Fig. 3B, filled triangles). An IC_{50} of ~ 31 μM , with a selectivity index of 4.5, for inhibition of late stage of replication has been obtained. At concentrations above 50 μM , the cytotoxic effects of **6d** interfered with the analysis of its inhibitory effects. These results indicate that, similar to the known LEDGINs,^[23-30] **6d** could also be affecting both the LEDGF/p75-dependent integration rather than steps such as reverse transcription, nuclear import of PIC, during early stage of replication, while the virus generated in the presence of **6d** could be less infectious due to an effect on the maturation stage, arising from the aberrant multimerization of IN (a LEDGF/p75-independent effect).

Conclusion

In conclusion, compound **6d** (ethyl-1-benzyl-7-methoxyisoquinoline-4-carboxylate) that efficiently inhibits the interaction between LEDGF/p75 and IN *in vitro*, as well as, HIV-1 infection in a cell line, has been identified as a lead isoquinoline molecule in this study. The order of addition experiment indicates an overlap in the binding site for IBD, **6d** on CCD of IN. Similar to the known LEDGINs,^[23-33] **6d** shows an effect on the early and late stages of HIV-1 replication. In addition to the chemical differences between **6d** and already known LEDGINs (supplementary figure 1), molecular docking indicates a possibility for sufficient differences in the type of interactions between **6d** and the residues of CCD (Table 2). Of interest, is the ethyl group in the ester moiety of **6d** which fits into the hydrophobic pocket of CCD that is occupied by the side chain of one of the critical residue I365 of LEDGF/p75. In a recent study, we have shown that M178 of CCD might probably be the first residue from the hydrophobic pocket of IN to interact with I365 of LEDGF/p75.^[44] Since the mutation of M178 is shown to result in a catalytically inactive IN,^[51] as well as, result in an abrogation of its interaction with LEDGF/p75,^[44] the features of **6d** imparting interaction with M178 could potentially be exploited further for developing a small molecule having higher potency and variation from the known LEDGINs.

Acknowledgements

The authors are grateful to Alan N Engelman (Harvard Medical School) for sharing the plasmids expressing IN, IBD, p75 and pNLX.Luc.R⁻, pCG-VSVG for virus production and to Anand K Kondapi (University of Hyderabad) for sharing cell lines. AG acknowledges UGC, India, for PhD research fellowship. This work was supported by ICMR-DBT (India) collaborative effort on HIV/AIDS and Microbicides program (grant no. BT/PR4843/MED/29/372/2012) to NKR and GS.

Declaration of Interest

The authors report no conflict of interest.

Abbreviations

IN: integrase; IBD: integrase binding domain; LEDGF/p75: Lens epithelium derived growth factor; CCD: catalytic core domain

References

- [1] L. Krishnan, A. Engelman. *J. Biol. Chem.* 2012, 287, 40858.
- [2] M. Li, R. Craigie. *Methods.* 2009, 47, 237.
- [3] P. Cherepanov, G.N. Maertens, S. Hare. *Curr. Opin. Struc. Biol.* 2011, 21, 249.
- [4] S. Hare, G.N. Maertens, P. Cherepanov. *The EMBO J.* 2012, 31, 3020.
- [5] A. R. Schröder, P. Shinn, H. Chen, C. et al. *Cell.* 2002, 110, 521.
- [6] G. P. Wang, A. Ciuffi, J. Leipzig, et al. *Genome Research.* 2007, 17, 1186.
- [7] A. Ciuffi, M. Llano, E. Poeschla, et al. *Nature Medicine.* 2005, 11, 1287.
- [8] M. C. Shun, N. K. Raghavendra, N. Vandegraaff, et al. *Genes & Devel.* 2007, 21, 1767.
- [9] P. K. Singh, M. R. Plumb, A. L. Ferris, et al. *Genes & Devel.* 2015, 29, 2287.
- [10] G. A. Sowd, E. Serrao, H. Wang, et al. *Proc. Nat. Acad. Sci. (U.S.A).* 2016, 113, E1054.
- [11] J. De Rijck, K. Bartholomeeusen, H. Ceulemans, et al. *Nucl. Acids Res.* 2010, 38, 6135.
- [12] P. Cherepanov, Z. Y. Sun, S. Rahman, et al. *Nat. Struc. & Mol. Biol.* 2005, 12, 526.
- [13] P. Cherepanov, A. L. Ambrosio, S. Rahman, et al. *Proc. Nat. Acad. Sci. (U.S.A).* 2005, 102, 17308.
- [14] P. Cherepanov, E. Devroe, P. A. Silver, et al. *J. Biol. Chem.* 2004, 279, 48883.
- [15] M. Llano, M. Vanegas, N. Hutchins, et al. *J. Mol. Biol.* 2006, 360, 760.
- [16] M. C. Shun, Y. Botbol, X. Li, et al. *J. Virol.* 2008, 82, 11555.
- [17] R. Gijsbers, S. Vets, J. De Rijck, et al. *J. Biol. Chem.* 2011, 286, 41812.
- [18] J. O. Eidahl, B.L. Crowe, J. A. North, et al. *Nucl. Acids Res.* 2013, 41, 3924.
- [19] E. M. Poeschla. *Cell. & Mol. Life Scien.* 2008, 65, 1403.
- [20] A. Engelman, P. Cherepanov. *PLoS Patho.* 2008, 4, e1000046.
- [21] S. Hare, M. C. Shun, S. S. Gupta, et al. *PLoS Patho.* 2009, 5, e1000259.
- [22] S. Rahman, R. Lu, N. Vandegraaff, et al. *Virology.* 2007, 357, 79.
- [23] F. Christ, A. Voet, A. Marchand, et al. *Nat. Chem. Biol.* 2010, 6, 442.
- [24] J. J. Kessl, N. Jena, Y. Koh, et al. *J. Biol. Chem.* 2012, 287, 16801.
- [25] K. A. Jurado, H. Wang, A. Slaughter, et al. *Proc. Nat. Acad. Sci. (U.S.A).* 2013, 110, 8690.
- [26] M. Tsiang, G. S. Jones, A. Niedziela-Majka, et al. *J. Biol. Chem.* 2012, 287, 21189.
- [27] M. Balakrishnan, S. R. Yant, L. Tsai, et al. *PloS One.* 2013, 8, e74163.
- [28] F. Christ, S. Shaw, J. Demeulemeester, et al. *Anti. Agents & Chemo.* 2012, 56, 4365.
- [29] B. A. Desimmie, R. Schrijvers, J. Demeulemeester, et al. *Retrovirology.* 2013, 10, 57.
- [30] K. Gupta, T. Brady, B. M. Dyer, et al. *J. Biol. Chem.* 2014, 289, 20477.
- [31] Z. Debyser, F. Christ, J. De Rijck, et al. *Trends Biochem. Scien.* 2015, 40, 108.

- [32] K. A. Jurado, A. Engelman. *Exp. Rev. Mol. Med.* 2013, 15, e14.
- [33] H. Wang, K.A. Jurado, X. Wu, et al. *Nucleic Acids Res.* 2012, 40, 11518.
- [34] Y. Kashiwada, A. Aoshima, Y. Ikeshiro, et al. *Bio. & Med. Chem.* 2005, 13, 443.
- [35] T. M. Kutchan. *The Plant Cell.* 1995, 7, 1059.
- [36] A. Diamond, I. Desgagné-Penix. *Plant Biotech. J.* 2016, 14,1319.
- [37] A. G. K. Reddy, G. Satyanarayana. *Tetrahedron.* 2012, 68, 8003.
- [38] A. G. K. Reddy, J. Krishna, G. Satyanarayana. *Synlett.* 2011, 2011, 1756.
- [39] R. A. Puras Lutzke, C. Vink, R. H. Plasterk. *Nucl. Acids Res.* 1994, 22, 4125.
- [40] A. Engelman, A. B. Hickman, R. Craigie. *J. Virol.* 1994, 68, 5911.
- [41] D. Esposito, R. Craigie. *The EMBO J.* 1998, 17, 5832.
- [42] F. Turlure, G. Maertens, S. Rahman, et al. *Nucl. Acids Res.* 2006, 34, 1653.
- [43] K. M. Tsutsui, K. Sano, O. Hosoya, et al. *Nucl. Acids Res.* 2011, 39, 5067.
- [44] A. George, N. K. Raghavendra. *Biochem. & Biophys. Res. Comm.* 2017, 490, 271.
- [45] P. Cherepanov, G. Maertens, P. Proost, *J. Biol. Chem.* 2003, 278, 372.
- [46] G. N. Maertens, P. Cherepanov, A. Engelman. *J. Cell Scien.* 2006, 119, 2563.
- [47] J. DeBoer, C. J. Madson, M. Belshan. *Virology.* 2016, 489, 282.
- [48] S. Maignan, J. P. Guilloteau, Q. Zhou-Liu, et al. *J. Mol. Biol.* 1998, 282, 359.
- [49] A. Engelman, J. J. Kessl, M. Kvaratskhelia. *Curr. Opin. Chem. Biol.* 2013, 17, 339.
- [50] F. Christ, Z. Debyser. *Virology.* 2013, 435, 102.
- [51] L. Q. Al-Mawsawi, A. Hombrouck, R. Dayam, et al. *Virology.* 2008, 377, 355.

Figure Legends

Figure 1. *In vitro* pull-down analysis. The figure shows pull-down assays performed using recombinant IN and IBD in the presence of the compounds. Panels A, B and C show assay performed using compounds **1a**, **6a** and **6d**, respectively. The concentration (mM or μ M) of the compound in the assay samples is indicated at the top of the respective lanes. The numbers below each lane represent the percent inhibition of formation of IBD-IN complex by the compound, calculated with respect to the assay having no compound. Bands corresponding to IBD and IN are indicated. The picture in panel C is assembled from two separate gels. Panel D shows the interactions between **6d** and the residues of CCD of IN in the docked complex of **6d**-CCD. The amino acid residues are indicated by solid black lines, the structure of **6d** is shown with gray lines and the dashed lines represent the interactions between **6d** and CCD.

Figure 2. Pre-formed complexes are not disrupted by compound **6d**. Panels A and B show the formation of IN-IBD and IN-p75 complexes respectively. The order of addition of the components in the pull-down assay is indicated at the top of respective lane. The symbol '>' indicates sequential addition and '+' indicates incubation of components together prior to the addition of the next component in the assay. Letter D denotes inclusion of a non-specific DNA to the assay. Bands corresponding to IBD, p75 and IN are indicated to the right. Lane numbers are indicated below the gel picture.

Figure 3. Compound **6d** inhibits LEDGF/p75-IN interaction in the mammalian cell. Panel A shows the Western blot analysis of the complex pull-down from the cell lysates using anti-Flag beads. The 'IN+6d' in lane 1 represents treatment of the cell lysate having Flag-IN with **6d**, prior to addition of lysate containing HA-p75. The 'p75+6d' in lane 2 represents treatment of the cell lysate having HA-p75 with **6d**, prior to addition of lysate containing Flag-IN. The '-6d' in lane 3 represents the LEDGF/p75-IN interaction obtained by mixing lysates expressing either one of the

protein partners without **6d** treatment. Top part shows blotting of the pull-down complex using anti-HA antibodies for analysing HA-p75. Bottom part shows the blotting of pull-down complex using anti-Flag antibodies for analysing Flag-IN. Panel B shows the inhibition of HIV-1 infection by **6d**. The concentration of **6d** in μM (0, 12.5, 25 and 50) used in the analysis is indicated on the X-axis. Dotted line (filled square marks) indicates the percent of live cells in samples treated with different concentrations of **6d**. Dashed dot line (filled circle marks) and dashed line (filled triangle marks) represent the percent inhibition of the early and late stages of HIV-1 replication by **6d**, respectively.

Table Legends

Table 1. IC_{50} of the tetrahydroisoquinolines for inhibition of the interaction between IN and IBD. ‘N.I.’ indicates no inhibition with 24 mM concentration of compound; >24 mM indicates IC_{50} in range of 25-34 mM and >>24 mM indicates IC_{50} in range of 35-60 mM; ‘N.T.’ indicates compound not tested in the assay due to insolubility of the compound.

Table 2. Interactions between CCD and compound **6d**. Compound is indicated as molecule 1 and the two chains of CCD dimer are denoted as A and B. Positions of H, C and O atoms on compound and those of amino acid residues in IN are mentioned against each interaction.

Figure 1. *In vitro* pull-down analysis.

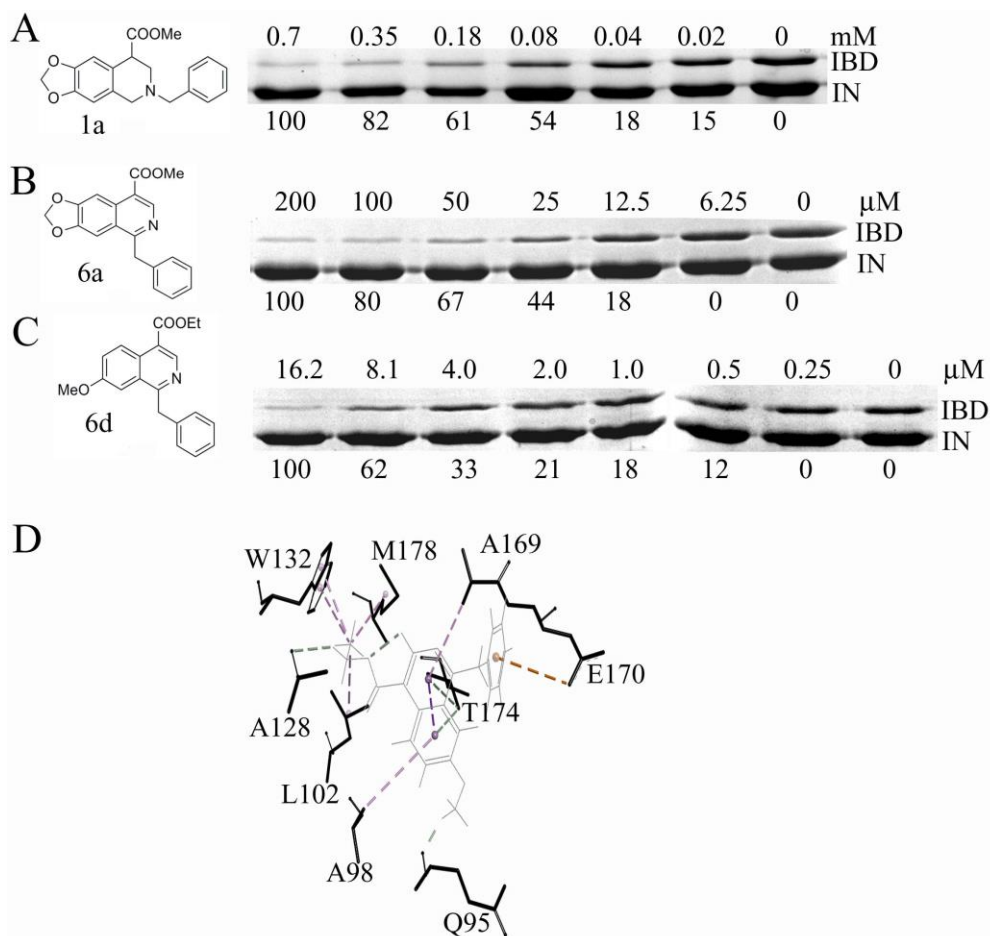


Figure 2. Pre-formed complexes are not disrupted by compound **6d**.

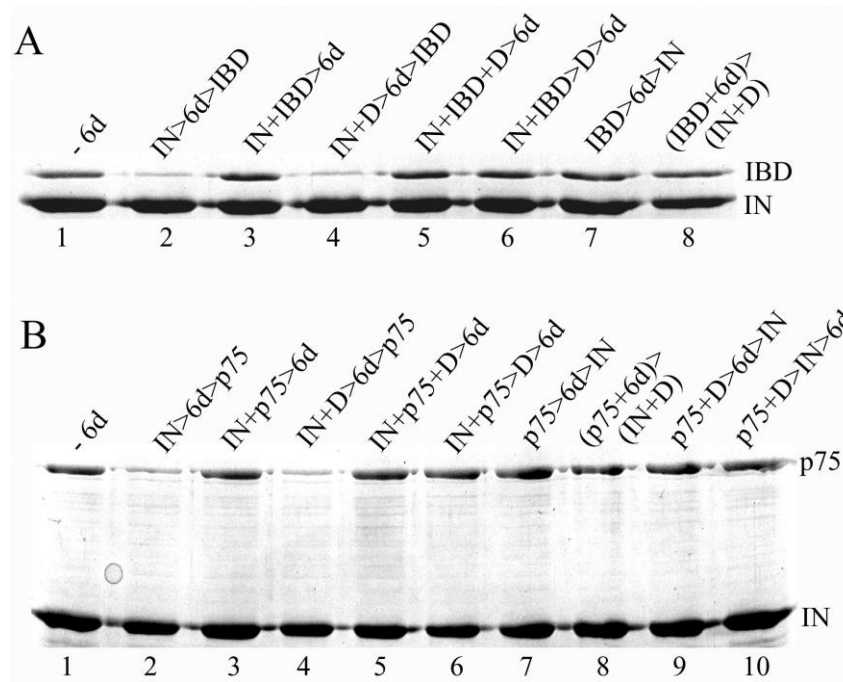


Figure 3. Compound **6d** inhibits LEDGF/p75-IN interaction in cell.

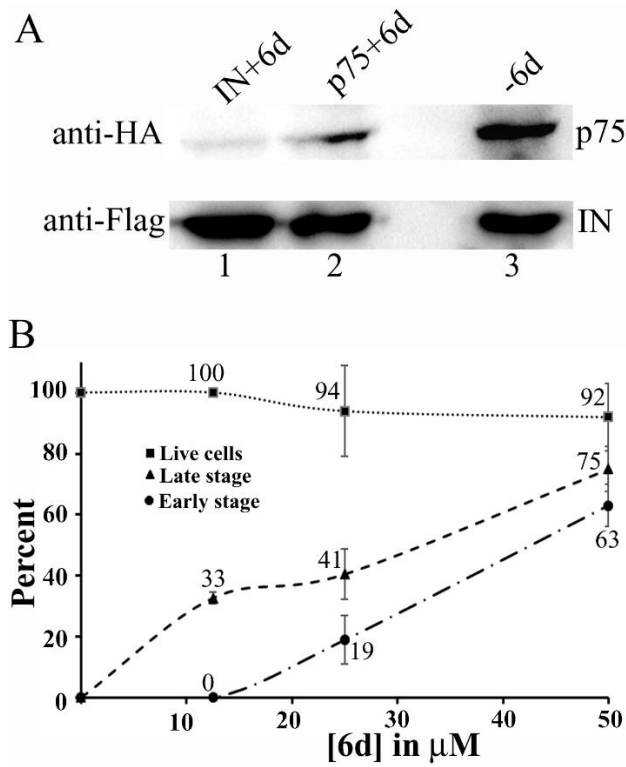


Table 1. IC₅₀ of 1,2,3,4-tetrahydroisoquinolines.

Compound	IC₅₀ (mM)	Compound	IC₅₀ (mM)
1a	0.081±0.01	1e	>> 24 mM
1h	1.63±0.02	1i	> 24 mM
1k	2.23±0.15	1j	> 24 mM
1b	7.07±0.62	1m	>> 24 mM
1f	10.03±2.11	1d	N.I
1l	12.22±1.4	1g	N.T
1c	>> 24 mM		

Table 2. Interactions between CCD and compound **6d**.

CCD - 6d complex	Type of Interaction
Molecule 1:H29 - B:GLN95:O	Carbon Hydrogen Bond
Molecule 1:H39 - B:ALA128:O	Carbon Hydrogen Bond
A:GLU170:OE1 - Molecule 1	Pi-Anion
A:THR174:OG1 - Molecule 1	Pi-Donor Hydrogen Bond
A:THR174:OG1 - Molecule 1	Pi-Donor Hydrogen Bond
A:THR174:CG2 - Molecule 1	Pi-Sigma
A:THR174:CG2 - Molecule 1	Pi-Sigma
Molecule 1:C24 - A:MET178	Alkyl
Molecule 1:C24 - B:LEU102	Alkyl
B:TRP132 - Molecule 1:C24	Pi-Alkyl
B:TRP132 - Molecule 1:C24	Pi-Alkyl
Molecule 1 - B:ALA98	Pi-Alkyl
Molecule 1 - A:ALA169	Pi-Alkyl