RSC Advances

PAPER

View Article Online
View Journal | View Issue

Cite this: *RSC Adv.*, 2024, **14**, 9020

Design, synthesis and biological evaluation of indole-2-carboxylic acid derivatives as novel HIV-1 integrase strand transfer inhibitors†

Rong-Hong Zhang,‡[ab] Guo-Qi Chen,‡[ac] Weilin Wang,[e] Yu-Chan Wang,[c] Wen-Li Zhang,[c] Ting Chen,[c] Qian-Qian Xiong,[c] Yong-Long Zhao,[c] Shang-Gao Liao,[c] Yong-Jun Li,[a] Guo-Yi Yan*[d] and Meng Zhou *[ac]

Received 6th December 2023
Accepted 8th March 2024

DOI: 10.1039/d3ra08320a

rsc.li/rsc-advances

Integrase plays an important role in the life cycle of HIV-1, and integrase strand transfer inhibitors (INSTIs) can effectively impair the viral replication. However, drug resistance mutations have been confirmed to decrease the efficacy of INSTI during the antiviral therapy. Herein, indole-2-carboxylic acid (**1**) was found to inhibit the strand transfer of integrase, and the indole nucleus of compound **1** was observed to chelate with two Mg^{2+} ions within the active site of integrase. Through optimization of compound **1**, a series of indole-2-carboxylic acid derivatives were designed and synthesized, and compound **17a** was proved to markedly inhibit the effect of integrase, with IC_{50} value of 3.11 μM. Binding mode analysis of **17a** demonstrated that the introduced C6 halogenated benzene ring could effectively bind with the viral DNA (dC20) through π–π stacking interaction. These results indicated that indole-2-carboxylic acid is a promising scaffold for the development of integrase inhibitors.

Introduction

Acquired immune deficiency syndrome (AIDS), which reduces the effectiveness of the immune system of the body, is a potentially life-threatening illness caused by the infection of human immunodeficiency virus (HIV), making this pandemic a health problem all over the world.[1–3] Highly active antiretroviral therapy (HAART) has improved the survival of patients with AIDS, which is usually comprised of three or more antiretroviral drugs including reverse transcriptase, protease, and integrase inhibitors.[4–6] However, due to the high variability of HIV-1, drug resistant strains emerged rapidly during the clinical use of antiviral agents, leading to the decline in efficacy of the HAART treatment.[7–9]

HIV-1 integrase plays a critical role in the viral replication, which inserts a double stranded viral DNA copy of the viral RNA genome into the host genome.[10] Generally, the insertion is achieved through a multistep process, mainly including two separate reactions, 3′-processing and strand transfer.[11,12] Most HIV-1 integrase inhibitors were reported to target the strand transfer step, which involves a concerted nucleophilic attack by two reactive 3′-OH ends of the viral DNA to the host chromosomal DNA.[12,13] During the integration reaction, no energy source is needed, and only divalent cations are required under physiological conditions, particularly two Mg^{2+} ions,[14] which are coordinated by three catalytic carboxylate residues (Asp64, Asp116, and Glu152, called DDE motif) within the enzyme active site.[15] Unlike transcriptase and protease, there is no host cellular counterpart for HIV-1 integrase, and specific integrase inhibitors would not interfere with cellular functions.[16] In addition, the integrase uses the same active site (DDE motif) for both the 3′-processing and DNA strand transfer steps, and inhibitors targeting integrase could benefit from a potentially high genetic barrier to resistance selection.[12] Therefore, extensive efforts have been made to identify effective integrase inhibitors with diverse chemical features.[17,18]

After two decades of attempts, it is believed that a metal binding group (MBG) interacting with two Mg^{2+} cofactors and a pendent aromatic or heteroaromatic hydrophobic moiety close to the MBG are essential for the development of integrase inhibitors.[19–23] Until now, five integrase antagonists have been

[a]*State Key Laboratory of Functions and Applications of Medicinal Plants, Engineering Research Center for the Development and Application of Ethnic Medicine and TCM (Ministry of Education), Guizhou Medical University, Guiyang, 550004, P. R. China. E-mail: gmu_mengzhou@163.com*

[b]*Center for Tissue Engineering and Stem Cell Research, Key Laboratory of Regenerative Medicine of Guizhou Province, School of Basic Medical Sciences, Guizhou Medical University, Guiyang, 550004, P. R. China*

[c]*School of Pharmacy, Guizhou Medical University, Guian New District, Guizhou, 550025, P. R. China*

[d]*School of Pharmacy, Xinxiang University, Xinxiang, 453000, P. R. China*

[e]*State Key Laboratory of Biotherapy, Collaborative Innovation of Biotherapy and Cancer Center, West China Hospital of Sichuan University, Chengdu 610041, Sichuan, China*

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d3ra08320a

‡ These authors contributed equally to this work.

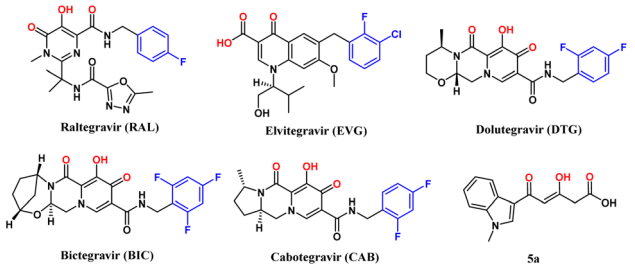


Fig. 1 Structure of HIV-1 integrase inhibitors. The metal chelating heteroatoms and halogenated phenyl groups interacting with dC20 are colored in red and blue, respectively.

approved for clinical use: Raltegravir (RAL, Fig. 1),[24] Elvitegravir (ELV),[25] Dolutegravir (DTG),[26] Bictegravir (BIC),[27] and Cabotegravir (CAB).[28] All these drugs effectively inhibit the DNA strand transfer catalyzed by integrase and are called integrase strand transfer inhibitors (INSTIs).[29,30] However, drug resistance mutations inevitably emerged during the antiviral therapy, which impaired the susceptibility of the virus to the INSTIs.[27,31–33] Thus, development of novel integrase inhibitors has been always an important research topic for the treatment of AIDS.

Indole alkaloids represent a well-known natural product, and majority of these alkaloids exhibit a variety of biological effect, such as antiviral, antitumor, antibacterial, antifungal, and antiplasmodial activities.[34–37] In particular, indole derivatives have been considered as one class of promising HIV-1 inhibitors targeting reverse transcriptase, protease, and CCR5.[38] Moreover, indole-β-diketo acid **5a** was reported to inhibit the strand transfer process of integration,[39] and indole derivative was proved to be a allosteric inhibitor of HIV-1 integrase.[40] However, none of these compounds entered the clinical studies because of low selectivity or potency.

Since the report of the cryo-electron microscopy (cryo-EM) of integrase, the mode of action of INSTIs has been well analyzed, including a chelating core with two Mg^{2+} and a π-stacking interaction with 3′ terminal adenosine of processed vDNA (dC20).[15,41] In this study, indole-2-carboxylic acid-based integrase inhibitors were designed and synthesized, and structural optimizations were further carried out to improve the metal–chelating and π-stacking interactions with two Mg^{2+} and vDNA dC20, respectively. Through these efforts, potential INSTIs were developed, and the probably binding modes with HIV-1 integrase were analyzed as well.

Results and discussion

Discovery strategy

Through analysing the structure of the reported integrase inhibitors, it can be found that a carboxyl group is usually used to chelate with the metal ions in the active site of integrase.[42,43] However, a bis-bidentate chelation with two Mg^{2+} ions is hard to establish for a single carboxyl group, and additional heteroatoms (*e.g.*, O, N) are required for a stable metal–chelating interaction. Considering the potential of indole scaffold in the development of INSTIs, herein, a carboxyl group was introduced to C2 of indole structure to

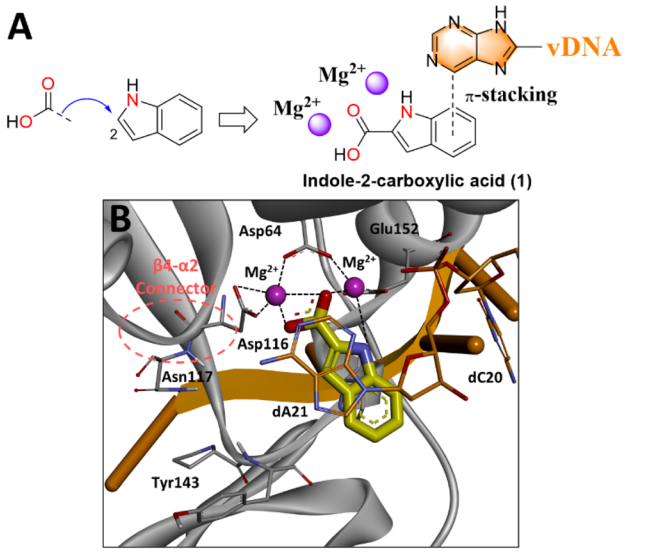


Fig. 2 (A) Design strategy of indole-2-carboxylic acid (**1**) as a potential HIV-1 integrase inhibitor. (B) Putative binding modes of compound **1** with HIV-1 integrase (PDB ID: 6PUY) is depicted in the right panel. HIV-1 integrase is displayed in gray, the 3′ end of the viral DNA (dA21 and dC20) is shown in stick representation in orange, and the chelate bond is represented with dashed line in black.

enhance the metal chelation (Fig. 2A). Meanwhile, the aromatic moiety of indole-2-carboxylic acid might interact with the viral DNA in the active region as well.

To verify this hypothesis, indole-2-carboxylic acid (**1**) was docked into the crystal structure of HIV-1 integrase (PDB ID: 6PUY), and the location of the original substrate (**4d**) was designated as the docking site.[15] The result showed that the indole nitrogen and 2-carboxyl group were found to chelate two metals within the active site, which was consistent with the former hypothesis (Fig. 2B). In addition, a π-stacking interaction was observed between the indole core and dA21. To confirm the antiviral potency of compound **1**, the integrase strand transfer inhibitory activity was then investigated using HIV-1 integrase assay kit.[44,45] The result proved that compound **1** indeed inhibited the integration of vDNA with IC_{50} of 32.37 μM (Table 1), indicating the potential of indole-2-carboxylic acid as a HIV-1 integrase inhibitor. Thus, to improve the integrase inhibitory effect, structural optimizations were further performed on indole-2-carboxylic acid scaffold.

Structural optimizations on indole-2-carboxylic acid

The crystal structures of HIV-1 integrase have revealed that a hydrophobic cavity is observed adjacent to the β4–α2 connector (Fig. 1),[18,41] which might be the key to increase the antiviral effect of compound **1** derivatives. Although the reported INSTIs (*e.g.* DTG and BIC) were designed to extend to the cavity, substantial free space was still observed around the pocket, and improving the interaction with the hydrophobic chamber might increase the antiviral effect of compound **1**. Therefore, to fill the space of the hydrophobic cavity, a bulky hydrophobic pharmacophore was proposed to be introduced at

Table 1 HIV-1 integrase inhibitory effect and cytotoxicity of indole-2-carboxylic acid derivatives

Compd	IC_{50}[a] (μM)	CC_{50}[b] (μM)	Compd	IC_{50} (μM)	CC_{50} (μM)
1	32.37 ± 4.51	>80	**16d**	>40	>80
4a	10.06 ± 0.56	>80	**16e**	>40	>80
4b	10.18 ± 0.67	>80	**16f**	>40	>80
4c	>40	>80	**16g**	>40	>80
4d	15.64 ± 1.03	>80	**16h**	8.68 ± 0.77	>80
4e	15.70 ± 1.06	>80	**16i**	14.65 ± 1.12	>80
4f	>40	>80	**16j**	9.67 ± 0.89	>80
9a	>40	54.31 ± 4.79	**16k**	>40	>80
9b	>40	>80	**17a**	3.11 ± 0.18	>80
10a	10.55 ± 0.89	65.21 ± 5.22	**17b**	6.67 ± 0.51	>80
13a	26.03 ± 1.27	>80	**17c**	12.28 ± 1.06	>80
13b	11.67 ± 1.08	>80	**17e**	7.08 ± 0.63	>80
16a	>40	>80	**17f**	14.10 ± 0.78	>80
16b	>40	>80	**17g**	15.56 ± 1.09	>80
16c	>40	>80	**RAL**	0.06 ± 0.04	>80

[a] Concentration required to inhibit the *in vitro* strand-transfer step of integration by 50%. [b] Concentration that inhibits the proliferation of MT-4 cells by 50%. All data values are average for at least three independent experiments.

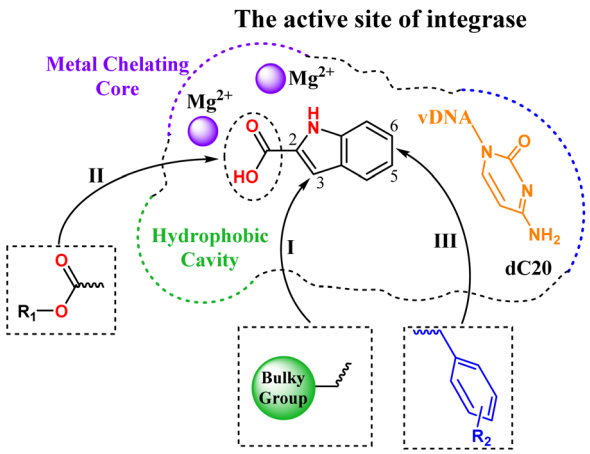


Fig. 3 Optimization strategies of compound **1**.

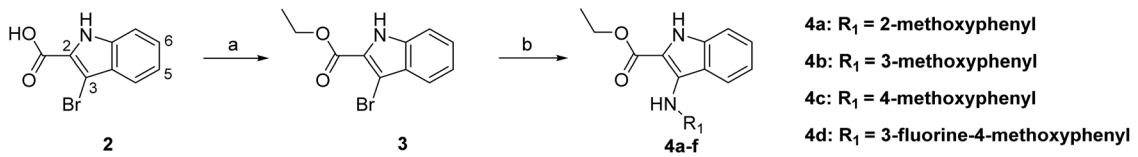
C3 of the indole structure (Fig. 3 strategy I). Meanwhile, modifications on C2 carboxyl acid would also be performed to evaluate the chelation with the metal ions of the intasome core (Fig. 3 strategy II). In addition, to improve the π–π stacking interaction with the viral DNA (dC20) as the reported integrase inhibitors,[46] a halogenated phenyl would then be added to the C5 or C6 position of indole scaffold (Fig. 3 strategy III).

Synthesis route of compound 1 derivatives

To facilitate the modifications on C3 position of indole core, 3-bromoindole-2-carboxylic acid (**2**) was used as the initial material, and the synthetic route was described in Scheme 1. The carboxyl group of **2** was esterificated to give compound **3** in the presence of concentrated sulfuric acid. Substituted anilines were then introduced at C3 position through the Buchwald–Hartwig reaction catalyzed by palladium acetate to offer compounds **4a–4f**.

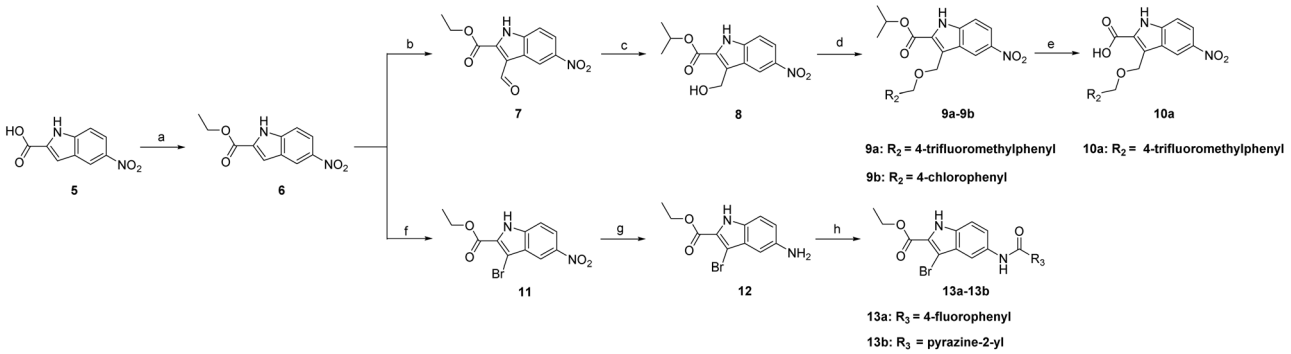
To explore the effect of C5 position of indole scaffold, 5-nitroindole-2-carboxylic acid (**5**) was selected as the material, which was esterificated to give compound **6** (Scheme 2). Due to the electron absorption effect of C2 carbonyl group, a formyl group was easily added at C3 position through the Vilsmeier–Haack reaction to provide compound **7** in a high yield (95%), which was then reduced to hydroxymethyl (**8**) by using isopropanolic aluminium *via* the Meerwein–Ponndorf–Verley reaction. Meanwhile, an ester-exchange product (**8**) was detected under this condition. Sequentially, compound **8** was condensed with 4-(trifluoromethyl)benzyl alcohol and 4-chlorobenzyl alcohol to give compounds **9a** and **9b** under alkaline conditions respectively, and compound **9a** was further hydrolyzed to provide **10a**.

On the other hand, a bromine atom was added to C3 of compound **6** by using *N*-bromo-succinimide (NBS) to offer

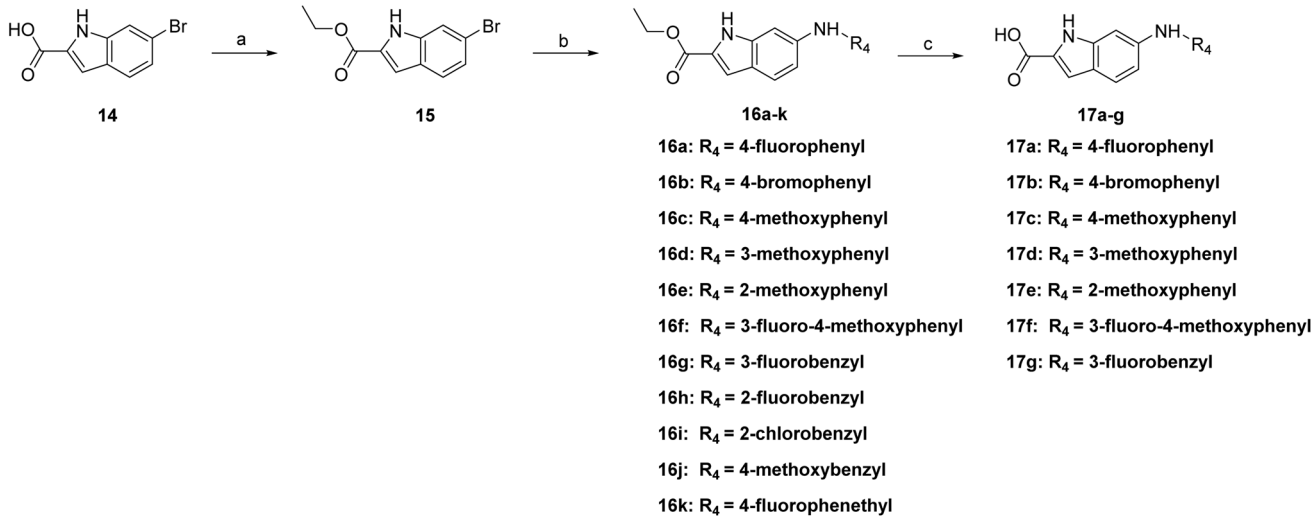


4a: R_1 = 2-methoxyphenyl
4b: R_1 = 3-methoxyphenyl
4c: R_1 = 4-methoxyphenyl
4d: R_1 = 3-fluorine-4-methoxyphenyl
4e: R_1 = 4-fluorinephenyl
4f: R_1 = 2-chlorinephenyl

Scheme 1 Synthetic route of compounds **4a–4f**. Reagents and conditions: (a) Con. H_2SO_4, EtOH, 80 °C, 2 h, 85%; (b) substituted anilines, $Pd(OAc)_2$, CS_2CO_3, Xphos, 1,4-dioxane, 110 °C, 2–4 h, 60–86%.



Scheme 2 Synthetic route of compounds **9a–9b**, **10a**, and **13a–13b**. Reagents and conditions: (a) Con. H_2SO_4, EtOH, 80 °C, 2 h, 65%; (b) $POCl_3$, DMF, rt-50 °C, 4 h, 95%; (c) $Al(O\text{-}i\text{-}Pr)_3$, i-PrOH, 60 °C, 5 h, 80%; (d) 4-(trifluoromethyl)benzyl alcohol or 4-chlorobenzyl alcohol, K_2CO_3, DMF, rt, 3–5 h, 79–86%; (e) NaOH, MeOH, H_2O, 80 °C, 2 h, 43%; (f) NBS, DMF, rt 4 h, 85%; (g) $SnCl_2 \cdot 2H_2O$, EtOAc, 80 °C, 3 h, 48%; (h) 4-fluorobenzoic acid or pyrazine-2-carboxylic acid, DMAP, EDAC, THF, rt, 3–5 h, 78–82%.



Scheme 3 Synthetic route of compounds **16a–16k** and **17a–17g**. Reagents and conditions: (a) Con. H_2SO_4, EtOH, 80 °C, 2 h, 82%; (b) substituted anilines, benzylamines or phenethylamine, $Pd(OAc)_2$, Cs_2CO_3, Xphos, 1,4-dioxane, 110 °C, 2–4 h, 63–82%; (c) NaOH, MeOH, H_2O, 80 °C, 1–3 h, 35–50%.

intermediate **11** (Scheme 2), the nitro of which was then reduced to an amino group (**12**) to facilitate the introduction of aromatic carboxylic acids. Subsequently, an amide condensation reaction was proceeded between compound **12** and 4-fluorobenzoic acid or pyrazine-2-carboxylic acid to afford **13a** or **13b** in the presence of condensing agent *N*-(3-dimethylaminopropyl)-*N*′-ethylcarbodiimide hydrochloride (EDAC).

To performed the structural optimizations at C6 position of indole scaffold, 6-bromoindole-2-carboxylic acid (**14**) was chosen as the initial material, and the synthetic route was described in Scheme 3. The carboxyl group of **14** was esterificated to give compound **15** as the former. Then, substituted anilines, benzylamines, and phenethylamine were introduced at C6 *via* the Buchwald–Hartwig reaction catalyzed by palladium acetate to provide compounds **16a–16k**. Among these molecules, carboxylic acid esters **16a–16h** were further hydrolyzed to yield **17a–17k**, respectively.

Biological activity evaluation and molecular docking study

In order to evaluate the effect of the hydrophobic cavity near the integrase active site, six substituted anilines were introduced to C3 position of compound **1** (**4a–4f**) to improve the interaction with this hydrophobic site (Scheme 1). The inhibitory activity of these derivatives against integrase strand transfer was then evaluated (Table 1). The results showed that four derivatives (**4a**, **4b**, **4d**, and **4e**) exhibited better antiviral activity than the parent compound **1**, with IC_{50} values of 10.06–15.70 μM, indicating that the optimizations on C3 of compound **1** indeed enhanced the inhibitory activity. Among them, the derivatives with 2-methoxyphenyl (**4a**) or 3-methoxyphenyl (**4b**) at C3 of indole core exhibited the best activity, improving the inhibitory effect about 3-folds compared with compound **1**.

The binding modes of compounds **4a** and **4b** with integrase were then investigated by using AutoDock Vina. The conformation of **4a** revealed that the introduced 2-methoxyphenyl

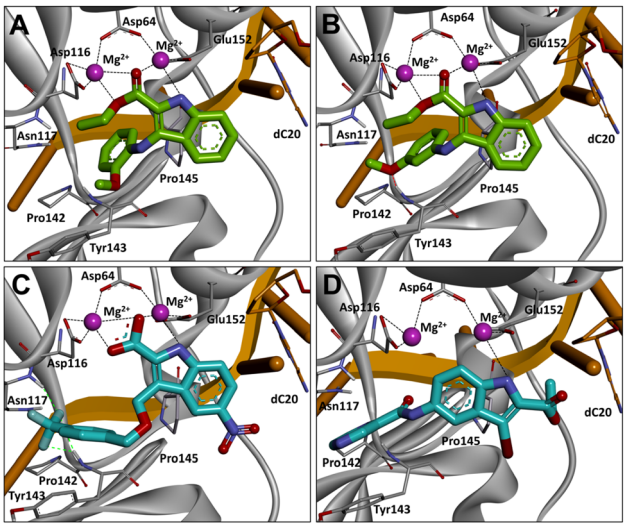


Fig. 4 Putative binding modes of compounds **4a** (A), **4b** (B), **10a** (C), and **13b** (D) with HIV-1 integrase. Compounds **4a** and **4b** are showed in green carbon, and compounds **10a** and **13b** are showed in cyan carbon. dA21 and all water molecules are omitted for clarity.

extended to the hydrophobic region and interacted with Pro145 (Fig. 4A). For compound **4b**, a similar binding mode was found, and a hydrogen bond was formed with Tyr143 (Fig. 4B). The above binding mode analysis demonstrating that the introduced C3 groups could interact with the hydrophobic cavity and thus improve the activity, which was consistent with the previous design.

Ligand efficiency is a widely used metric for selecting favorable fragments in drug discovery,[47] and the efficiency of the introduced fragments was then evaluated using this parameter. As the initial fragment, compound **1** bound with integrase with a binding free energy of −9.49 kcal mol^{-1} (Table 2) and a high ligand efficiency of 0.79. Probably because the low molecular complexity of this fragment allows target interaction potential to be explored more efficiently and accurately.[48–50] For compounds **4a** and **4b**, although the introduction of methoxyphenyl increased the binding free energies (−8.39 and 9.25 kcal mol^{-1}) and decreased the ligand efficiency (0.36 and 0.40 respectively), the strand transfer inhibitory activity was improved because of the interaction with the hydrophobic cavity of integrase. However, due to the short chain of the C3 substituents, the interaction with the hydrophobic pocket still needed to be improved, and further modifications were then carried out to lengthen the linker of C3 substituents.

Table 2 Binding free energy and ligand efficiency of indole-2-carboxylic acid derivatives

Compd	ΔG^a (kcal mol^{-1})	N^b	Ligand efficiency
1	−9.49	12	0.79
4a	−8.39	23	0.36
4b	−9.25	23	0.40
10a	−10.37	28	0.37
13b	−8.09	24	0.34
16a	−6.82	22	0.31
17a	−17.46	20	0.87
17g	−16.3	21	0.78

[a] ΔG is the binding free energy. [b] N is the number of heavy atoms.

To enhance the interaction with the backbone of residues that forms the hydrophobic pocket, the linker of C3 substituent was lengthened, and two substituted benzyloxymethyls were added to the indole core (**9a** and **9b**). Whereas, none of these compounds inhibited the strand transfer of integrase (Table 1), probably because the esterified C2 carboxyl impaired the chelation with the metal ions of the intasome core. Thus, the isopropyl ester of **9a** was further hydrolyzed to give **10a**, and encouragingly, the inhibitory effect was restored, with IC_{50} of 10.55 μM. To analyze the above optimizations on C3, compound **10a** was docked into the active site of integrase (Table 2). Compared with compound **1**, the binding free energy was decreased to 10.37 kcal mol^{-1}. Binding modes analysis showed that a strong chelating core was observed in the free carboxyl group, which was similar with compound **1** (Fig. 4C). Hydrophobic interactions were found between **10a** and Pro142, and three hydrogen bonds were formed with Tyr143 and Asn117, indicating that the lengthened C3 group could significantly enhance the adhesion with the hydrophobic chamber.

Since most of the reported INSTIs contain halogenated phenyl groups near the chelating core,[46] the effect of C5 position on the indole structure was then evaluated. According to the preliminary design, a halogenated benzene (**13a**) or pyrazine (**13b**) was introduced at the 5-position of compound **1** *via* an amide bond. The biological results revealed that these two compounds exhibited better strand transfer inhibitory effect than the parent **1** (Table 1), and the IC_{50} of **13b** reached to 11.67 μM. However, compared with compounds **4a**, **4b** and **10a**, no improvement was observed in the derivatives **13a** and **13b**. Binding mode analysis revealed that, unlike the previous compounds, **13b** bound with the integrase using a quite different conformation (Fig. 4D). Probably because of the rigid planar nature of C5 benzamide group, this part was hard to insert into the pocket formed by the viral DNA through the rotation of chemical bond. Instead, C5 side chain was located into the hydrophobic cavity near the β4–α2 connector, and obvious an apparent π–π stacking interaction was formed with Tyr143. Meanwhile, due to the conversion of the binding conformation, the chelation of the indole core was disappeared, and the binding interaction and ligand efficiency were both decreased (Table 2). Therefore, to improve the π-stacking interaction with the viral DNA (dC20), a flexible side chain might be more suitable for the antiviral activity.

According to the above analysis, flexible substituted anilines, benzylamines and phenethylamine were added to C6 of indole scaffold (**16a–16f**, **16g–16j**, and **16k**). Nevertheless, no integrase inhibitory activity was observed for most of the derivatives (Table 1), probably due to the esterified carboxyl groups. However, three compounds (**16h**, **16i** and **16j**) were found effectively inhibited the strand transfer of integrase, with IC_{50} values of 8.68, 14.65 and 9.67 μM respectively. Further hydrolysis of the carboxylate of **16a–16g** was carried out to provide **17a–17g**, which resulted in obvious enhancements in activity,

with IC_{50}s of 3.11–15.56 μM. Among them, compound **17a** exhibited the best integrase inhibitory effect (IC_{50} = 3.11 μM), increasing the activity about 10.4-folds compared with compound **1**. Whereas compound **17g**, which contains a benzylamine at C6 of indole scaffold, only showed a mild inhibitory activity (IC_{50} = 15.56 μM), suggesting that a longer linker would weaken the activity.

To better understand the effect of C6 substituent groups on biological activity, three compounds (**16a**, **17a**, and **17g**) were selected to study the binding modes with HIV-1 integrase (Fig. 5). Similar with the parent compound **1**, three central electronegative heteroatoms of **17a** were observed to chelate with two Mg^{2+} cofactors, suggesting a strong metal–chelating cluster between the indole nucleus and the integrase (Fig. 5A). Meanwhile, a π–π stacking interaction with dC20 was also found in C6 halogenated benzene ring, and the introduced fragment markedly improved the binding free energy and ligand efficiency, with the values of −17.46 kcal mol^{-1} and 0.87 respectively (Table 2). These results demonstrated that the introduced C6 halogenated benzene indeed increased the adhesion with the integrase and, as a result, enhanced the HIV-1 integrase inhibitory activity. Although a similar binding mode was observed for compound **16a** (Fig. 5B), the steric hindrance effect of esterified carboxylic acid markedly impaired the chelation with the metal ions and decreased the binding interaction (−6.82 kcal mol^{-1}) and ligand efficiency (0.31), leading to the decrease in activity. For compound **17g**, which contains a longer chain at C6 of indole scaffold, the chelating core was observed in this molecule as well (Fig. 5C), whereas π–π stacking interaction was not found. Instead, C6 side chain extended to Tyr143 and formed a halogen bond with this residue, resulting in a increased binding free energy (−16.3 kcal mol^{-1}) and a lower ligand efficiency (0.78) compared with **17a** (Table 2). The above binding mode analysis revealed that a flexible short chain at C6 of indole scaffold was appropriate to improve the inhibitory effect against integrase strand transfer.

On the other hand, the toxicity of the synthesized compounds to the MT-4 human T-cell line was also assessed (Table 1). The results revealed that no toxicity was observed for most of the compounds at 80 μM except **9a** and **10a**, and the CC_{50} values of these two compounds were both higher than 54 μM, indicating lower toxicities.

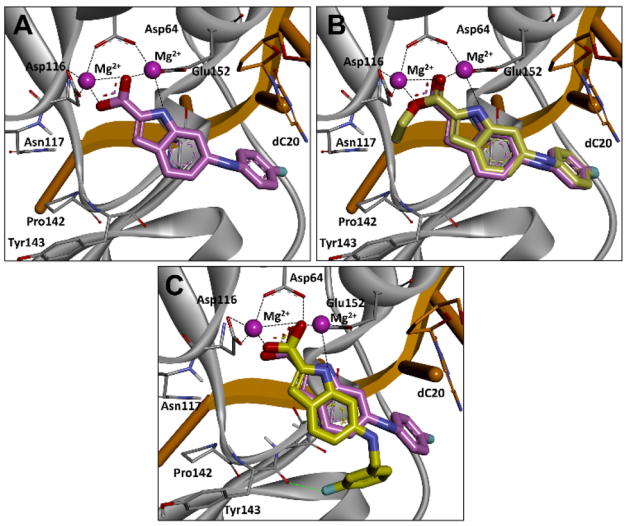


Fig. 5 Putative binding patterns of compound **17a** with HIV-1 integrase (A); superimposed conformations of **17a** and **16a** (B); superimposed conformations of **17a** and **17g** (C). Compounds **17a** is display in pink carbon, **16a** and **17g** are showed in yellow carbon, and the halogen bond is represented with dashed line in green.

Experimental section

Chemistry

All analytical grade reagents and chemicals were purchased from commercial suppliers. Column chromatography was performed on silica gel (Qingdao, 300–400 mesh) using the indicated eluents. ^{1}H NMR and ^{13}C NMR spectra were recorded on JEOL spectrometer (400 MHz) using $CDCl_3$, CD_3OD or DMSO-d_6 as solvents with the internal standard of tetramethylsilane. HRMS was recorded on a Thermo Scientific Q Exactive Plus Orbitrap LC-MS/MS. All prepared compounds were purified to >96% purity as determined by HPLC (Thermo Scientific Dionex Ultimate 3000, USA) analysis using the following methods. Purity analysis of final compounds was performed through a SuperLu C_{18} (particle size = 5 μm, pore size = 4.6 nm, dimensions = 250 mm) column. The injection volume was 10 μL, the mobile phase was consisted of methyl alcohol and water (80 : 20) with a flow rate of 1.0 mL min^{-1}, and each analysis lasted for 20 min. The detection wavelength was 280 nm. The retention time of each compound ($R_{T,HPLC}$) was displayed in the analytical data of the respective compounds. All melting points were obtained on a WRS-2 microcomputer melting point apparatus.

General procedure for the synthesis of compounds 3–17

Synthesis of ethyl 3-bromo-1*H*-indole-2-carboxylate (3). 3-Bromo-1*H*-indole-2-carboxylic acid (100 mg, 0.42 mmol) was dissolved in anhydrous ethanol (10 mL), and concentrated sulphuric acid (20 mg, 0.21 mmol) was added dropwise to the solution. The mixture was stirred at 80 °C for 2 h (monitored by TLC). The reaction was quenched with a saturated aqueous solution of sodium bicarbonate and extracted three times with ethyl acetate (30 mL × 3). The combined organic phase was dried over anhydrous sodium sulfate before vacuum suction filtration. The crude product was chromatographed on silica gel (1 : 5 v/v ethyl acetate/petroleum ether) to afford a white solid (91 mg, 85%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.22 (s, 1H), 7.68 (d, J = 7.2 Hz, 1H), 7.39 (q, J = 8.0 Hz, 2H), 7.26–7.21 (m, J = 13.3, 1H), 4.47 (q, J = 7.2 Hz, 2H), 1.46 (t, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 161.18, 135.41, 128.05, 126.68, 124.15, 121.56, 121.41, 112.13, 98.40, 61.63, 14.45. ESI-HRMS (m/z): calcd $C_{11}H_{10}BrNO_2$ for [M − H]−, 265.9822, 267.9801; found, 265.9821, 267.9796. $R_{T,HPLC}$ = 7.15 min, purity >99%. The same procedure was also followed for the synthesis of **6** and **15**.

Synthesis of ethyl 3-((2-methoxyphenyl)amino)-1*H*-indole-2-carboxylate (4a). 2-Methoxyaniline 4-bromoaniline (46 mg, 0.37 mmol), palladium(II) acetate (16 mg, 0.07 mmol), cesium carbonate (76 mg, 0.56 mmol), and 2-(dicyclohexylphosphino)-2′,4′,6′-tri-i-propyl-1,1′-biphenyl (71 mg, 0.15 mmol) were added to a solution of **6** (100 mg, 0.37 mmol) in 1,4-dioxane (10 mL) under the protection of nitrogen. The mixture was stirred at 110 °C for 2 h and monitored by TLC. Upon completion, the mixture was concentrated under reduced pressure to provide the crude product and purified by silica gel chromatography from (1 : 5 v/v ethyl acetate/petroleum ether) to afford a white solid (93 mg, 81%). mp 131.4–135.6 °C. ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.49 (s, 1H), 7.47 (d, J = 8.0 Hz, 2H), 7.38 (d, J = 8.0 Hz, 1H), 7.31 (t, J = 7.2 Hz, 1H), 7.01 (t, J = 8.0 Hz, 2H), 6.83–6.79 (m, 3H), 4.35 (q, J = 6.8 Hz, 2H), 3.92 (s, 3H), 1.34 (t, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, DMSO-d_6) δ 162.45, 148.27, 136.48, 134.07, 127.49, 126.00, 122.24, 120.93, 120.90, 119.90, 119.28, 114.80, 113.88, 113.57, 110.99, 60.72, 56.03, 14.74. ESI-HRMS (m/z): calcd $C_{18}H_{18}N_2O_3$ for $[M + H]^+$, 311.1396; found, 311.1402. $R_{T,HPLC}$ = 9.35 min, purity >97%. The same procedure was also followed for the synthesis of **4b–4f**, and **16a–16k**.

Ethyl 3-((3-methoxyphenyl)amino)-1*H*-indole-2-carboxylate (4b). White solid, 69% yield. mp 130.3–134.7 °C, ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.44 (s, 1H), 7.75 (s, 1H), 7.44 (d, J = 8.4 Hz, 1H), 7.36 (d, J = 8.2 Hz, 1H), 7.30–7.26 (m, 1H), 7.07–6.97 (m, 2H) 6.45 (d, J = 7.6 Hz, 2H), 6.34 (d, J = 6.4 Hz, 1H), 4.30 (q, J = 6.4 Hz, 2H), 3.64 (s, 3H), 1.28 (t, J = 6.8 Hz, 3H). ^{13}C NMR (100 MHz, DMSO-d_6) δ 162.05, 160.55, 147.27, 136.24, 130.01, 126.49, 125.86, 121.99, 121.81, 119.38, 116.49, 113.42, 108.54, 104.63, 101.49, 60.55, 55.21, 14.77. ESI-HRMS (m/z): calcd $C_{18}H_{18}N_2O_3$ for $[M + H]^+$, 311.1396; found, 311.1402. $R_{T,HPLC}$ = 6.65 min, purity >99%.

Ethyl 3-((4-methoxyphenyl)amino)-1*H*-indole-2-carboxylate (4c). White solid, 75% yield. mp 128.4–131.1 °C, ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.21 (s, 1H), 7.53 (s, 1H), 7.39 (d, J = 8.4 Hz, 1H), 7.26–7.21 (m, 1H), 7.16 (d, J = 8.4 Hz 1H), 6.93–6.87 (m, 3H), 6.82 (d, J = 8.8 Hz, 2H), 4.31 (q, J = 6.8 Hz, 2H), 3.70 (s, 3H), 1.31 (t, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, DMSO-d_6) δ 162.46, 154.07, 138.30, 136.65, 129.56, 125.95, 122.07, 120.44, 119.43, 118.84, 114.69, 113.77, 113.33, 60.39, 55.72, 14.91. ESI-HRMS (m/z): calcd $C_{18}H_{18}N_2O_3$ for $[M + H]^+$, 311.1396; found, 311.1402. $R_{T,HPLC}$ = 7.92 min, purity >99%.

Ethyl 3-((3-fluoro-4-methoxyphenyl)amino)-1*H*-indole-2-carboxylate (4d). White solid, 86% yield. mp 145.6–148.9 °C, ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.40 (s, 1H), 7.70 (s, 1H), 7.43 (d, J = 8.4 Hz, 1H), 7.27 (t, J = 8.0 Hz, 2H), 7.02–6.95 (m, 2H), 6.75 (d, J = 13.6 Hz, 1H), 6.62 (d, J = 8.8 Hz, 1H), 4.30 (q, J = 7.2 Hz, 2H), 3.76 (s, 3H), 1.29 (t, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, DMSO-d_6) δ 162.06, 152.41 (J_{CF} = 240.6 Hz), 151.20, 140.35 (J_{CF} = 48.1 Hz), 136.35, 127.14, 125.93, 121.61, 121.41, 119.35, 115.80, 115.40, 113.43, 111.92, 105.40 (J_{CF} = 21.2 Hz), 60.52, 57.02, 14.80. ^{19}F NMR (376 MHz, DMSO-d_6): δ −134.11 (t, J = 11.5 Hz). ESI-HRMS (m/z): calcd $C_{18}H_{17}FN_2O_3$ for $[M + H]^+$, 329.1301; found, 329.1309. $R_{T,HPLC}$ = 7.04 min, purity >99%.

Ethyl 3-((4-fluorophenyl)amino)-1*H*-indole-2-carboxylate (4e). White solid, 80% yield. mp 138.1–141.0 °C, ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.41 (s, 1H), 7.73 (s, 1H), 7.43 (d, J = 8.8, 1H), 7.26 (d, J = 7.6, 1H), 7.03–6.95 (m, 3H), 6.88–6.85 (m, 2H), 4.29 (q, J = 6.8 Hz, 2H), 1.27 (t, J = 6.8 Hz, 3H). ^{13}C NMR (100 MHz, DMSO-d_6) δ 137.33, 132.86 (J_{CF} = 186.4 Hz), 121.56, 116.75, 109.38, 108.40, 104.97, 103.17, 101.55 (J_{CF} = 6.0 Hz), 100.53, 100.25 (J_{CF} = 18.2 Hz), 100.16, 98.42, 56.09, 19.50. ^{19}F NMR (376 MHz, DMSO-d_6): δ −125.71 to −125.64 (m). ESI-HRMS (m/z): calcd $C_{17}H_{15}FN_2O_2$ for $[M + H]^+$, 299.1196; found, 299.1202. $R_{T,HPLC}$ = 4.14 min, purity >99%.

Ethyl 3-((2-chlorophenyl)amino)-1*H*-indole-2-carboxylate (4f). White solid, 60% yield. mp 141.2–143.4 °C, ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.68 (s, 1H), 7.51–7.43 (m, 3H), 7.33–7.29 (m, 2H), 7.12 (t, J = 8.0 Hz, 2H), 7.05–7.02 (m, 1H), 6.84–6.78 (m, 2H), 4.32 (q, J = 6.8 Hz, 2H), 1.30 (t, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, DMSO-d_6) δ 162.21, 141.68, 136.21, 129.84, 128.27, 126.03, 125.32, 121.58, 121.53, 120.44, 120.36, 119.88, 116.78, 115.45, 113.65, 60.93, 14.69. ESI-HRMS (m/z): calcd $C_{17}H_{15}ClN_2O_2$ for $[M + H]^+$, 315.0900; found, 315.0907. $R_{T,HPLC}$ = 6.71 min, purity >99%.

Ethyl 5-nitro-1*H*-indole-2-carboxylate (6). White solid, 65% yield. ^{1}H NMR (400 MHz, DMSO-d_6) δ 12.63 (s, 1H), 8.73 (s, 1H), 8.14 (d, J = 8.8 Hz, 1H), 7.61 (d, J = 9.2 Hz, 1H), 7.44 (s, 1H), 4.38 (q, J = 7.2 Hz, 2H), 1.36 (t, J = 6.8 Hz, 3H) ^{13}C NMR (100 MHz, DMSO-d_6) δ 160.61, 141.46, 139.98, 130.83, 125.90, 119.65, 119.37, 113.18, 110.09, 61.00, 14.19. ESI-HRMS (m/z): calcd $C_{11}H_{10}N_2O_4$ for $[M - H]^-$, 233.0641; found, 233.0559. $R_{T,HPLC}$ = 12.43 min, purity >96%.

Synthesis of ethyl 3-formyl-5-nitro-1*H*-indole-2-carboxylate (7). Phosphorus oxychloride (572 mg, 3.73 mmol) was added dropwise to a solution of compound **6** (87 mg, 0.37 mmol) in 10 mL of DMF. The mixture was stirred at room temperature for 2 h and heated under reflux for another 2 h (monitored by TLC). The solution was cooled to room temperature, brought to pH 8 by adding anhydrous sodium carbonate, extracted with ethyl acetate (20 mL × 3), and concentrated under reduced pressure to afford the crude product. The residue was applied to silica gel column chromatography (1 : 4 v/v ethyl acetate/petroleum ether) to give a white solid (92 mg, 95%). ESI-HRMS (m/z): calcd $C_{12}H_{10}N_2O_5$ for $[M + H]^+$, 263.0656; found, 263.0673.

Synthesis of isopropyl 3-(hydroxymethyl)-5-nitro-1*H*-indole-2-carboxylate (8). The intermediate **7** (89 mg, 0.34 mmol) and aluminium isopropoxide (207 mg, 1.01 mmol) were dissolved in isopropanol (15 mL) and stirred at 60 °C for 5 h (monitored by TLC). After removal of isopropanol using a rotary evaporator, the reaction mixture was applied to silica gel column chromatography (1 : 5 v/v ethyl acetate/petroleum ether) to provide a white solid (76 mg, 80%). ^{1}H NMR (400 MHz, DMSO-d_6) δ 12.18 (s, 1H), 8.92 (s, 1H), 8.11 (d, J = 9.2 Hz, 1H), 7.58 (d, J = 9.2 Hz, 1H), 5.24–5.17 (m, 2H), 5.08 (s, 2H), 1.38 (d, J = 6.4 Hz, 6H). ^{13}C NMR (100 MHz, DMSO-d_6) δ 159.91, 140.33, 138.37, 125.47, 125.45, 124.69, 118.99, 118.76, 112.35, 68.03, 54.35, 21.06. ESI-HRMS (m/z): calcd $C_{13}H_{14}N_2O_5$ for $[M - H]^-$, 277.0903; found, 277.0830. $R_{T,HPLC}$ = 9.08 min, purity >99%.

Synthesis of isopropyl 5-nitro-3-((4-(trifluoromethyl) phenethoxy) methyl)-1*H*-indole-2-carboxylate (9a). Compound **8** (89 mg, 0.32 mmol), 4-(trifluoromethyl)benzyl bromide (76 mg, 0.32 mmol), and potassium carbonate (133 mg, 0.96

mmol) were dissolved in 10 mL of DMF. The solution was stirred at room temperature for 5 h (monitored by TLC). After the addition of 5 mL of water, the mixture was extracted with ethyl acetate (30 mL × 3), concentrated under reduced pressure, and applied to silica gel column chromatography (1 : 5 v/v ethyl acetate/petroleum ether) to afford a white solid (124 mg, 86%). mp 124.4–127.3 °C, ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.03 (s, 1H), 8.16 (d, J = 9.2 Hz, 1H), 7.78 (d, J = 9.6 Hz, 1H), 7.66 (d, J = 7.6 Hz, 2H), 7.19 (d, J = 7.6 Hz, 2H), 5.93 (s, 2H), 5.08 (s, 3H), 1.23 (d, J = 6.0 Hz, 6H). ^{13}C NMR (100 MHz, DMSO-d_6) δ 160.95, 143.23, 142.13, 141.05, 128.64, 128.32, 127.74, 127.39, 127.28, 126.03 (J_{CF} = 7.5 Hz), 125.97, 125.76, 123.33, 120.47, 120.31, 112.28, 69.72, 56.15, 48.38, 21.88. ^{19}F NMR (376 MHz, DMSO-d_6): δ −60.89 (s). ESI-HRMS (m/z): calcd $C_{21}H_{19}F_3N_2O_5$ for $[M + Na]^+$, 459.1144; found, 459.1129. $R_{T,HPLC}$ = 8.94 min, purity >98%. The same procedure was also followed for the synthesis of **9b**.

Isopropyl 3-((4-chlorophenethoxy)methyl)-5-nitro-1*H*-indole-2-carboxylate (9b). White solid, 79% yield. mp 138.2–141.4 °C, ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.01 (s, 1H), 8.15 (d, J = 9.2 Hz, 1H), 7.77 (d, J = 9.2 Hz, 1H), 7.35 (d, J = 7.6 Hz, 2H), 7.03 (d, J = 7.6 Hz, 2H), 5.81 (s, 2H), 5.27 (t, J = 5.6 Hz, 1H), 5.07 (d, J = 4.8 Hz, 2H), 1.27 (d, J = 6.0 Hz, 6H). ^{13}C NMR (100 MHz, DMSO-d_6) δ 159.84, 140.87, 139.78, 136.16, 131.26, 127.87, 127.35, 126.54, 126.10, 124.57, 119.17, 119.08, 111.11, 68.55, 54.96, 46.81, 20.77. ESI-HRMS (m/z): calcd $C_{20}H_{19}ClN_2O_5$ for $[M + Na]^+$, 425.0880; found, 425.0865, $R_{T,HPLC}$ = 9.72 min, purity >98%.

Synthesis of 5-nitro-3-(((4-(trifluoromethyl)benzyl)oxy)methyl)-1*H*-indole-2-carboxylic acid (10a). Sodium hydroxide (56 mg, 1.39 mmol) was added to a solution of **9a** (122 mg, 0.28 mmol) in a mixture of methanol and water (16 mL, 3 : 1 methanol/water), and the reaction was stirred at 80 °C for 1.5 h. The solution was then cooled to room temperature, brought to a pH of 6 by adding dilute hydrochloric acid, extracted with ethyl acetate (15 mL × 3), and concentrated under reduced pressure to provide the crude product. The residue was purified by silica gel chromatography (1 : 2 v/v ethyl acetate/petroleum ether) to afford a light-yellow solid (47 mg, 43%). mp 182.6–186.3 °C, ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.20 (s, 1H), 6.52 (d, J = 6.0 Hz, 1H), 6.19 (d, J = 7.2 Hz, 1H), 6.12 (d, J = 6.4 Hz, 2H), 5.76 (d, J = 6.4 Hz, 2H), 4.78 (s, 2H), 4.08 (s, 2H). ^{13}C NMR (100 MHz, DMSO-d_6) δ 163.06, 143.37, 142.03, 141.00, 128.40, 127.39, 127.13, 126.03, 125.95 (J_{CF} = 7.3 Hz), 120.26, 120.19, 112.23, 56.10, 48.14. ^{19}F NMR (376 MHz, DMSO-d_6): δ −60.88 (s). ESI-HRMS (m/z): calcd $C_{18}H_{13}F_3N_2O_5$ for $[M - H]^-$, 393.0777; found, 393.0774. $R_{T,HPLC}$ = 9.06 min, purity >98%. The same procedure was also followed for the synthesis of **17a–17g**.

Synthesis of ethyl 3-bromo-5-nitro-1*H*-indole-2-carboxylate (11). Compound **9** (100 mg, 0.43 mmol) and NBS (75 mg, 0.43 mmol) were dissolved in 10 mL of DMF, and the mixture was stirred at room temperature and monitored by TLC. 4 h later, 10 mL water was added into the reaction and the solution was extracted with ethyl acetate (30 mL × 3). The combined organic phase was dried over anhydrous sodium sulfate before vacuum suction filtration. The crude product was chromatographed on silica gel (1 : 5 v/v ethyl acetate/petroleum ether) to give a light yellow solid (114 mg, 85%). ^{1}H NMR (400 MHz, DMSO-d_6) δ 12.93 (s, 1H), 8.41 (s, 1H), 8.19 (d, J = 9.2 Hz, 1H), 7.65 (d, J = 9.2 Hz, 1H), 4.42 (q, J = 7.2 Hz, 2H), 1.39 (t, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, DMSO-d_6) δ 160.02, 142.56, 139.01, 128.07, 126.63, 121.12, 117.85, 114.65, 98.60, 61.95, 14.67. ESI-HRMS (m/z): calcd $C_{11}H_9N_2O_4$ for $[M - H]^-$, 310.9668, 312.9647; found, 310.9677, 312.9653. $R_{T,HPLC}$ = 9.97 min, purity >99%.

Synthesis of ethyl 5-amino-3-bromo-1*H*-indole-2-carboxylate (12). To a solution of compound **11** (100 mg, 0.32 mmol) in ethyl acetate (20 mL) was added $SnCl_2 \cdot 2H_2O$ (302 mg, 1.28 mmol). The resulting mixture was heated under reflux for 3 h (monitored by TLC). Then the solution was cooled to room temperature, brought to pH 7–8 by adding anhydrous sodium carbonate, and concentrated under reduced pressure to afford the crude product. The residue was applied to silica gel column chromatography (1 : 4 v/v ethyl acetate/petroleum ether) to afford a light yellow solid (43 mg, 48%). ^{1}H NMR (400 MHz, CD_3OD) δ 7.25 (d, J = 8.4 Hz, 1H), 6.91–6.89 (m, 2H), 4.39 (q, J = 7.2 Hz, 2H), 1.41 (t, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, CD_3OD) δ 160.98, 141.26, 131.13, 128.33, 123.66, 118.52, 112.85, 103.75, 95.25, 60.59, 13.38. ESI-HRMS (m/z): calcd $C_{11}H_{11}BrN_2O_2$ for $[M + H]^+$, 283.0082, 285.0062; found, 283.0091, 285.0070. $R_{T,HPLC}$ = 8.08 min, purity >99%.

Synthesis of ethyl 3-bromo-5-(4-fluorobenzamido)-1*H*-indole-2-carboxylate (13a). Compound **12** (100 mg, 0.35 mmol), 4-fluorobenzoic acid, DMAP, and EDAC were dissolved in 15 mL of THF, and the mixture was stirred at room temperature for 3 h (monitored by TLC). Upon completion, the reaction was concentrated under reduced pressure, and 10 mL water was added into the mixture, and the solution was then extracted three times with ethyl acetate (30 mL × 3). The combined organic phase was dried over anhydrous sodium sulfate before vacuum suction filtration. The crude product was chromatographed on silica gel (1 : 4 v/v ethyl acetate/petroleum ether) to give a white solid (111 mg, 78%). mp 196.3–199.6 °C, ^{1}H NMR (400 MHz, DMSO-d_6) δ 12.24 (s, 1H), 10.35 (s, 1H), 8.15 (s, 1H), 8.08 (dd, J = 8.8, 5.6 Hz, 2H), 7.69 (d, J = 9.2 Hz, 1H), 7.49 (d, J = 8.8 Hz, 1H), 7.39 (t, J = 8.8 Hz, 2H), 4.39 (q, J = 7.2 Hz, 2H), 1.38 (t, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, DMSO-d_6) δ 164.5 (J_{CF} = 247.3 Hz), 164.28, 160.09, 132.98 (J_{CF} = 32.6 Hz), 131.46, 130.39, 130.30, 126.71, 124.38, 121.09, 115.35 (J_{CF} = 22.4 Hz), 113.17, 110.70, 95.88, 60.83, 14.27. ^{19}F NMR (376 MHz, DMSO-d_6): δ −108.97 (br s). ESI-HRMS (m/z): calcd $C_{18}H_{14}BrFN_2O_3$ for $[M + H]^+$, 405.0250, 407.0230; found, 405.0246, 407.0224. $R_{T,HPLC}$ = 5.37 min, purity >95%. The same procedure was also followed for the synthesis of **13b**.

Ethyl 3-bromo-5-(pyrazine-2-carboxamido)-1*H*-indole-2-carboxylat (13b). White solid, 82% yield. mp 187.6–191.7 °C, ^{1}H NMR (400 MHz, DMSO-d_6) δ 12.26 (s, 1H), 10.82 (s, 1H), 9.33 (s, 1H), 8.94 (d, J = 2.8 Hz, 1H), 8.84–8.83 (m, 1H), 8.32 (s, 1H), 7.79 (dd, J = 9.2, 2 Hz, 1H), 7.50 (d, J = 8.8 Hz, 1H), 4.39 (q, J = 6.8 Hz, 2H), 1.39 (t, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, DMSO-d_6) δ 162.08, 160.55, 148.13, 145.74, 144.52, 143.76, 133.53, 132.76, 127.19, 125.01, 121.61, 113.76, 111.48, 96.50, 61.35, 14.76. ESI-HRMS (m/z): calcd $C_{16}H_{14}BrN_4O_3$ for $[M + Na]^+$, 411.0069, 413.0048; found, 411.0081, 413.0060. $R_{T,HPLC}$ = 14.65 min, purity >98%.

Ethyl 6-bromo-1*H*-indole-2-carboxylate (15). White solid, 82% yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.19 (s, 1H), 7.60 (s, 1H), 7.54 (d, J = 8.4 Hz, 1H), 7.24 (s, 1H), 7.19 (s, 1H), 4.43 (q, J = 7.2 Hz, 2H), 1.43 (t, J = 6.8 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 161.81, 137.38, 128.07, 126.27, 124.37, 123.81, 119.07, 114.77, 108.62, 61.29, 14.38. ESI-HRMS (m/z): calcd $C_{11}H_{10}BrNO_2$ for $[M - H]^-$, 265.9817, 267.9796; found, 265.9822, 267.9798. $R_{T,HPLC}$ = 12.51 min, purity >99%.

Ethyl 6-((4-fluorophenyl)amino)-1*H*-indole-2-carboxylate (16a). White solid, 82% yield. mp 153.6–156.2 °C, ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.94 (s, 1H), 7.53 (d, J = 8.8 Hz, 1H), 7.15 (s, 1H), 7.11–7.08 (m, 2H), 7.01 (s, 1H), 7.00–6.97 (m, 2H), 6.80 (d, J = 8.4, 1H), 4.36 (q, J = 6.8 Hz, 2H), 1.38 (t, J = 7.2 Hz 3H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 162.11, 158.26 (J_{CF} = 275.6 Hz), 142.41, 138.86, 138.17, 126.28, 123.52, 122.20, 120.98, 120.91, 116.11 (J_{CF} = 22.4 Hz), 115.89, 114.06, 109.11, 97.24, 60.83, 14.40. ^{19}F NMR (376 MHz, DMSO-d_6): δ −123.78 to −123.71 (m). ESI-HRMS (m/z): calcd $C_{17}H_{15}FN_2O_2$ for $[M + H]^+$, 299.1196; found, 299.1184. $R_{T,HPLC}$ = 6.64 min, purity >99%.

Ethyl 6-((4-bromophenyl)amino)-1*H*-indole-2-carboxylate (16b). White solid, 75% yield. mp 178.5–182.1 °C, ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.98 (s, 1H), 7.53 (d, J = 9.2 Hz, 1H), 7.14 (s, 1H), 7.10–7.06 (m, 2H), 7.00–6.96 (m 3H), 6.80 (d, J = 8.4 Hz, 1H), 4.36 (q, J = 7.2 Hz, 2H), 1.38 (t, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 162.14, 159.30, 156.91, 142.49, 138.88, 138.21, 126.23, 123.51, 122.13, 120.94, 120.86, 116.09, 115.87, 114.03, 109.11, 97.12, 60.83, 14.40. ESI-HRMS (m/z): calcd $C_{17}H_{15}BrN_2O_2$ for $[M - H]^-$, 357.0244, 359.0223; found, 357.0245, 359.0224. $R_{T,HPLC}$ = 13.02 min, purity >99%.

Ethyl 6-((4-methoxyphenyl)amino)-1*H*-indole-2-carboxylate (16c). White solid, 80% yield. mp 159.1–162.3 °C, ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.73 (s, 1H), 7.50 (d, J = 8.8 Hz, 1H), 7.12 (d, J = 8.0 Hz, 3H), 6.88 (d, J = 8.8 Hz, 3H), 6.76 (d, J = 10 Hz, 1H), 4.36 (q, J = 7.2 Hz, 2H), 3.81 (s, 3H), 1.38 (t, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, DMSO-d_6) δ 161.28, 153.83, 143.28, 139.13, 136.40, 125.15, 122.66, 120.52, 120.29, 114.56, 113.09, 108.46, 94.89, 59.95, 55.25, 14.39. ESI-HRMS (m/z): calcd $C_{18}H_{18}N_2O_3$ for $[M + H]^+$, 311.1396; found, 311.1397. $R_{T,HPLC}$ = 5.70 min, purity >99%.

Ethyl 6-((3-methoxyphenyl)amino)-1*H*-indole-2-carboxylate (16d). White solid, 78% yield. mp 155.9–157.3 °C, ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.81 (s, 1H), 7.56 (d, J = 8.4 Hz, 1H), 7.19 (s, 1H), 7.17–7.16 (m, 2H), 6.90 (dd, J = 8.8, 2 Hz, 1H), 6.72 (s, 1H), 6.71 (d, J = 7.6 Hz, 1H), 6.51 (d, J = 6.4 Hz, 1H), 4.39 (q, J = 6.8 Hz, 2H), 3.78 (s, 3H), 1.40 (t, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 162.06, 160.77, 144.46, 141.03, 138.00, 130.27, 126.68, 123.53, 122.83, 115.22, 110.65, 109.12, 106.58, 103.79, 99.35, 60.96, 55.34, 14.53. ESI-HRMS (m/z): calcd $C_{18}H_{18}N_2O_3$ for $[M + H]^+$, 311.1396; found, 311.1385. $R_{T,HPLC}$ = 6.22 min, purity >99%.

Ethyl 6-((2-methoxyphenyl)amino)-1*H*-indole-2-carboxylate (16e). White solid, 81% yield. mp 146.9–149.2 °C, ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.89 (s, 1H), 7.56 (d, J = 8.4 Hz, 1H), 7.39–7.36 (m, 1H), 7.21 (s, 1H), 7.16 (s, 1H), 6.95 (d, J = 2 Hz, 1H), 6.93 (d, J = 2 Hz, 1H), 6.91–6.88 (m, 2H), 4.38 (q, J = 7.2 Hz, 2H), 3.90 (s, 3H), 1.40 (t, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 162.08, 148.35, 141.03, 138.07, 132.80, 126.33, 123.29, 122.49, 120.77, 120.12, 115.49, 115.04, 110.49, 109.07, 98.89, 60.81, 55.59, 14.43. ESI-HRMS (m/z): calcd $C_{18}H_{18}N_2O_3$ for $[M + H]^+$, 311.1396; found, 311.1381. $R_{T,HPLC}$ = 8.32 min, purity >99%.

Ethyl 6-((3-fluoro-4-methoxyphenyl)amino)-1*H*-indole-2-carboxylate (16f). White solid, 77% yield. mp 163.8–167.4 °C, ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.92 (s, 1H), 7.53 (d, J = 8.8 Hz, 1H), 7.15 (s, 1H), 6.97 (d, J = 2.8 Hz, 1H), 6.94–6.93 (m, 1H), 6.89 (d, J = 9.2 Hz, 1H), 6.84–6.81 (m, 1H), 6.80 (d, J = 2.0 Hz, 1H), 6.78 (d, J = 2.0 Hz, 1H), 4.37 (q, J = 7.2 Hz, 2H), 3.88 (s, 3H), 1.39 (t, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 162.19, 152.88 (J_{CF} = 244.2 Hz), 142.80, 142.69, 142.37, 138.26, 136.86 (J_{CF} = 8.8 Hz), 126.37, 123.63, 122.26, 115.19, 114.72, 114.10, 109.21, 108.41 (J_{CF} = 20.7 Hz), 97.23, 60.93, 56.97, 14.51. ^{19}F NMR (376 MHz, DMSO-d_6): δ −133.55 to −133.49 (m). ESI-HRMS (m/z): calcd $C_{18}H_{17}FN_2O_3$ for $[M + H]^+$, 329.1301; found, 329.1298. $R_{T,HPLC}$ = 5.76 min, purity >99%.

Ethyl 6-((3-fluorobenzyl)amino)-1*H*-indole-2-carboxylate (16g). White solid, 67% yield. mp 150.0–153.6 °C, ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.60 (s, 1H), 7.44 (d, J = 8.8 Hz, 1H), 7.33–7.27 (m, 1H), 7.16 (d, J = 7.6 Hz, 1H), 7.11–7.09 (m, 2H), 6.98–6.93 (m, 1H), 6.58 (dd, J = 8.8, 2.0 Hz, 1H), 6.44 (s, 1H), 4.38 (s, 2H), 4.35 (q, J = 7.2 Hz, 2H), 1.38 (t, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 162.22 (J_{CF} = 238.5 Hz), 146.47, 141.92, 138.75, 130.22, 125.29, 123.43, 122.88, 120.44, 114.38 (J_{CF} = 2.2 Hz), 114.17, 112.06, 109.44, 92.35, 60.67, 48.20, 14.53. ^{19}F NMR (376 MHz, DMSO-d_6): δ −113.62 to −113.55 (m). ESI-HRMS (m/z): calcd $C_{18}H_{17}FN_2O_2$ for $[M + H]^+$, 313.1352; found, 313.1344. $R_{T,HPLC}$ = 5.91 min, purity >99%.

Ethyl 6-((2-fluorobenzyl)amino)-1*H*-indole-2-carboxylate (16h). White solid, 63% yield. mp 155.7–158.1 °C, ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.69 (s, 1H), 7.44 (d, J = 8.4 Hz, 1H), 7.39 (t, J = 6.0 Hz, 1H), 7.24–7.22 (m, 1H), 7.10–7.05 (m, 3H), 6.58 (dd, J = 8.4, 2.0 Hz, 1H), 6.48 (s, 1H), 4.44 (s, 2H), 4.35 (q, J = 6.8 Hz, 2H), 1.37 (t, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 162.47 (J_{CF} = 241.5 Hz), 146.66, 138.87, 129.45 (J_{CF} = 4.1 Hz), 129.01, 128.93, 126.09, 125.95, 125.17, 124.35, 123.38, 120.31, 115.46 (J_{CF} = 21.2 Hz), 115.36, 112.11, 109.46, 92.10, 60.68, 42.14, 14.54. ^{19}F NMR (376 MHz, DMSO-d_6): δ −133.62 to −133.55 (m). ESI-HRMS (m/z): calcd $C_{18}H_{17}FN_2O_2$ for $[M + H]^+$, 313.1352; found, 313.1341. $R_{T,HPLC}$ = 6.01 min, purity >99%.

Ethyl 6-((2-chlorobenzyl)amino)-1*H*-indole-2-carboxylate (16i). White solid, 65% yield. mp 173.3–176.8 °C, ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.59 (s, 1H), 7.46–7.38 (m, 3H), 7.22–7.20 (m, 2H), 7.10 (s, 1H), 6.61 (dd, J = 8.8, 2.0 Hz, 1H), 6.44 (s, 1H), 4.49 (s, 2H), 4.35 (q, J = 7.2 Hz, 2H), 1.37 (t, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 161.98, 146.33, 138.70, 136.15, 133.29, 129.59, 129.01, 128.52, 127.01, 125.20, 123.34, 120.36, 112.03, 109.36, 92.31, 60.59, 46.24, 14.47. ESI-HRMS (m/z): calcd $C_{18}H_{17}ClN_2O_2$ for $[M + H]^+$, 329.1057; found, 329.1047. $R_{T,HPLC}$ = 7.83 min, purity >98%.

Ethyl 6-((4-methoxybenzyl)amino)-1*H*-indole-2-carboxylate (16j). White solid, 66% yield. mp 125.6–127.1 °C, ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.68 (s, 1H), 7.43 (d, J = 8.4 Hz, 1H), 7.30 (d, J = 8.8 Hz, 2H), 7.11 (s, 1H), 6.87 (d, J = 8.8 Hz, 2H), 6.57 (dd, J = 8.4, 2.0 Hz, 1H), 6.49 (s, 1H), 4.35 (q, J = 6.8 Hz, 2H), 4.28 (s, 2H), 3.79 (s, 3H), 1.38 (t, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$)

δ 162.06, 158.92, 146.65, 138.80, 130.78, 128.89, 125.06, 123.23, 120.22, 114.05, 112.13, 109.39, 92.20, 60.56, 55.30, 48.20, 14.46. ESI-HRMS (*m*/*z*): calcd $C_{19}H_{20}N_2O_3$ for $[M + H]^+$, 325.1552; found, 325.1445. $R_{T,HPLC}$ = 8.23 min, purity >99%.

Ethyl 6-((4-fluorophenethyl)amino)-1*H*-indole-2-carboxylate (16k). White solid, 63% yield. mp 182.4–185.7 °C, ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.68 (s, 1H), 7.44 (d, *J* = 8.4 Hz, 1H), 7.20–7.16 (m, 2H), 7.11 (s, 1H), 7.00 (t, *J* = 8.8 Hz, 2H), 6.59–6.53 (m, 2H), 4.37 (q, *J* = 7.2 Hz, 2H), 3.42 (t, *J* = 7.2 Hz, 2H), 2.95 (t, *J* = 6.8 Hz, 2H), 1.39 (t, *J* = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 162.87, 162.04, 160.43, 138.71, 134.63, 130.21, 130.13, 125.30, 123.40, 115.55, 115.34, 112.27, 109.34, 60.64, 45.77, 34.24, 14.46. ESI-HRMS (*m*/*z*): calcd $C_{19}H_{19}FN_2O_2$ for $[M + H]^+$, 327.1509; found, 327.1483. $R_{T,HPLC}$ = 5.27 min, purity >98%.

6-((4-Fluorophenyl)amino)-1*H*-indole-2-carboxylic acid (17a). Yellow solid, 47% yield. mp 158.5–161.2 °C, ^{1}H NMR (400 MHz, CD_3OD) δ 12.65 (s, 1H), 11.34 (s, 1H), 8.17 (s, 1H), 7.50 (d, *J* = 8.8 Hz, 1H), 7.12–7.09 (m, 5H), 7.00 (s, 1H), 6.83 (d, *J* = 8.8 Hz, 1H). ^{13}C NMR (100 MHz, DMSO-d_6) δ 163.17, 156.72 (J_{CF} = 234.7 Hz), 155.54, 141.94, 140.71, 139.14, 127.16, 123.10, 121.65, 119.12, 116.14, (J_{CF} = 22.1 Hz), 113.97, 108.43, 97.57. ^{19}F NMR (376 MHz, DMSO-d_6): δ −124.03 to −123.96 (m). ESI-HRMS (*m*/*z*): calcd $C_{15}H_{11}FN_2O_2$ for $[M + H]^+$, 271.0883; found, 271.0892. $R_{T,HPLC}$ = 6.78 min, purity >97%. The same procedure was also followed for the synthesis of **17b–17g**.

6-((4-Bromophenyl)amino)-1*H*-indole-2-carboxylic acid (17b). Yellow solid, 45% yield. mp 153.4–157.9 °C, ^{1}H NMR (400 MHz, CD_3OD) δ 7.50 (d, *J* = 8.8 Hz, 1H), 7.31 (s, 1H), 7.29 (s, 1H), 7.17 (s, 1H), 7.08 (s, 1H), 7.04 (s, 1H), 7.02 (s, 1H), 6.85 (dd, *J* = 8.8, 2.0 Hz, 1H). ^{13}C NMR (100 MHz, CD_3OD) δ 165.14, 145.08, 142.26, 140.16, 132.93, 130.12, 128.00, 123.88, 123.81, 121.01, 119.34, 118.37, 115.84, 111.92, 109.90, 99.98. ESI-HRMS (*m*/*z*): calcd $C_{15}H_{11}BrN_2O_2$ for $[M - H]^-$, 328.9931, 330.9910; found, 328.9930, 330.9907. $R_{T,HPLC}$ = 5.14 min, purity >99%.

6-((4-Methoxyphenyl)amino)-1*H*-indole-2-carboxylic acid (17c). Yellow solid, 50% yield. mp 156.1–160.8 °C, ^{1}H NMR (400 MHz, CD_3OD) δ 7.42 (d, *J* = 8.8 Hz, 1H), 7.10 (d, *J* = 8.0 Hz, 1H), 7.05 (s, 1H), 6.97 (s, 1H), 6.85 (d, *J* = 9.2 Hz, 2H), 6.76 (d, *J* = 8.4 Hz, 1H), 3.76 (s, 3H). ^{13}C NMR (100 MHz, CD_3OD) δ 163.94, 154.75, 144.00, 139.28, 136.91, 125.79, 122.39, 121.26, 120.95, 120.29, 117.52, 114.21, 113.25, 108.82, 95.20, 54.65. ESI-HRMS (*m*/*z*): calcd $C_{16}H_{14}N_2O_3$ for $[M + H]^+$, 283.1083; found, 283.1079. $R_{T,HPLC}$ = 8.78 min, purity >99%.

6-((3-Methoxyphenyl)amino)-1*H*-indole-2-carboxylic acid (17d). Yellow solid, 42% yield. mp 176.7–178.3 °C, ^{1}H NMR (400 MHz, CD_3OD) δ 7.49 (d, *J* = 8.8 Hz, 1H), 7.20 (s, 1H), 7.12 (d, *J* = 7.6 Hz, 1H), 7.09 (s, 1H), 6.87 (d, *J* = 8.8 Hz, 1H), 6.71 (d, *J* = 8.4 Hz, 2H), 6.41 (d, *J* = 9.2 Hz, 1H), 3.75 (s, 3H). ^{13}C NMR (100 MHz, CD_3OD) δ 165.21, 162.10, 146.88, 142.92, 140.28, 130.84, 127.75, 123.73, 123.50, 115.91, 110.81, 109.98, 106.36, 103.74, 99.64, 55.51. ESI-HRMS (*m*/*z*): calcd $C_{16}H_{14}N_2O_3$ for $[M + H]^+$, 283.1083; found, 283.1075. $R_{T,HPLC}$ = 7.27 min, purity >98%.

6-((2-Methoxyphenyl)amino)-1*H*-indole-2-carboxylic acid (17e). Yellow solid, 40% yield. mp 177.3–179.4 °C, ^{1}H NMR (400 MHz, DMSO-d_6) δ 12.60 (s, 1H), 11.26 (s, 1H), 7.41 (dd, *J* = 8.8, 4.4 Hz, 1H), 7.31 (d, *J* = 3.6 Hz, 1H), 7.21–7.18 (m, 1H), 7.02 (s, 1H), 6.96–6.93 (m, 2H), 6.87–6.81 (m, 3H), 3.77 (s, 3H). ^{13}C NMR (100 MHz, DMSO-d_6) δ 162.23, 149.05, 140.85, 138.12, 132.12, 126.12, 121.76, 120.66, 120.14, 120.07, 116.32, 113.77, 110.96, 107.43, 97.39, 54.95. ESI-HRMS (*m*/*z*): calcd $C_{16}H_{14}N_2O_3$ for $[M + H]^+$, 283.1083; found, 283.1088. $R_{T,HPLC}$ = 8.73 min, purity >96%.

6-((3-Fluoro-4-methoxyphenyl)amino)-1*H*-indole-2-carboxylic acid (17f). Yellow solid, 45% yield. mp 170.9–173.2 °C, ^{1}H NMR (400 MHz, CD_3OD) δ 7.47 (d, *J* = 8.4 Hz, 1H), 7.17 (s, 1H), 7.06 (d, *J* = 9.6 Hz, 3H), 6.96 (d, *J* = 8.8 Hz, 1H), 6.79 (d, *J* = 8.0 Hz, 1H), 3.82 (s, 3H). ^{13}C NMR (100 MHz, CD_3OD) δ 163.89, 152.95 (J_{CF} = 242.4 Hz), 149.47, 142.58, 139.07, 138.24, 122.57, 121.85, 120.18, 117.76, 113.54 (J_{CF} = 8.2 Hz), 108.73, 96.59, 55.77. ^{19}F NMR (376 MHz, DMSO-d_6): δ −133.56 to −133.50 (m). ESI-HRMS (*m*/*z*): calcd $C_{16}H_{13}FN_2O_3$ for $[M + H]^+$, 301.1088; found, 301.1418. $R_{T,HPLC}$ = 9.32 min, purity >96%.

6-((3-Fluorobenzyl)amino)-1*H*-indole-2-carboxylic acid (17g). Yellow solid, 35% yield. mp 166.7–170.3 °C; ^{1}H NMR (400 MHz, CD_3OD) δ 7.35–7.28 (m, 2H), 7.21 (d, *J* = 8.0 Hz, 1H), 7.13 (d, *J* = 10 Hz, 1H), 7.00 (s, 1H), 6.93 (t, *J* = 8.8 Hz, 1H), 6.60 (dd, *J* = 8.8, 2.0 Hz, 1H), 6.43 (s, 1H), 4.37 (s, 2H). ^{13}C NMR (100 MHz, CD_3OD) δ 163.21 (J_{CF} = 243.0 Hz), 161.99, 147.10, 143.43, 143.36, 139.65, 129.81, 129.73, 125.07, 122.67, 122.64, 122.27, 119.90, 113.51 (J_{CF} = 2.2 Hz), 113.04 (J_{CF} = 2.1 Hz), 112.93, 111.80, 108.90, 91.71. ^{19}F NMR (376 MHz, DMSO-d_6): δ −113.57 to −113.64 (m). ESI-HRMS (*m*/*z*): calcd $C_{16}H_{13}FN_2O_2$ for $[M + H]^+$, 285.1039; found, 285.1022. $R_{T,HPLC}$ = 7.13 min, purity >99%.

Biochemistry

Cell lines, and culture conditions. All the reagents and chemicals were purchased from commercial sources. HIV-1 Integrase Assay Kit was purchased from Abnova. MT-4 cells were cultured in RPMI-1640 media and equipped with 10% FBS (Gibco, Australia), 100 units per mL penicillin and 100 units per mL streptomycin at 37 °C in a humidified atmosphere (5% CO_2 and 95% air).

Strand transfer inhibition assay. The strand transfer assay was performed using a HIV-1 integrase assay kit (Abnova, Taipei, China) according to the manufacturer's instructions. Each sample was dissolved in DMSO and diluted to final desired concentration (80, 40, 20, 10, 5 μM) with reaction buffer. Briefly, 100 μL of DS DNA solution was added to each well and incubated for 30 min at 37 °C, and the liquid was removed from the wells and washed 3 times with wash buffer. Then, 200 μL of blocking buffer was added to each well and incubated for 30 min in a 37 °C incubator. Followed by the aspiration of the liquid, each well was washed 3 times with reaction buffer, and 100 μL integrase enzyme solution was added to each well and incubated under similar conditions. Sequentially, each well was added 50 μL reaction buffer containing different concentrations of samples and incubated for 5 min at room temperature. To each well was added 50 μL of TS DNA solution, and the reactions were mixed gently and incubated for 30 min at 37 °C. Then, each well was washed 3 times with wash solution, incubated with 100 μL HRP antibody, and incubated for 30 min at 37 °C. After washing the plate under the same conditions, 100 μL TMB

peroxidase substrate solution was added to each well and incubated at room temperature for 10 min. To terminate the reaction, 100 μL TMB stop solution was directly added to each well and the absorbance of the sample was determined using a plate reader (B20-TEK ELX800UV, USA) at 450 nm.

Cytotoxicity assay. The cytotoxicity of the title compounds was tested through MTT analysis on MT-4 cells. Cells were seeded in 96-well plates with a density of 50 000 cells per well and coincubated with serial diluted compounds (2.5, 5, 10, 20, 40, and 80 μM) at 37 °C for five days. MTT (Sigma) was then added to each well with a final concentration of 0.5 mg mL^{-1} and incubated for 2 h. Subsequently, the medium was decanted, and 150 μL of DMSO was added to each well. The absorbance was obtained by a microplate reader at 570 nm, and all measurements were conducted three times under the same condition. 50% cytotoxicity concentration (CC50) values were calculated by Graphpad Prism 8.

Molecular docking. Molecular docking of 168 compounds with drug-like property was conducted by AutoDock Vina to predict the binding affinity as well as binding modes of the compounds with integrase. The binding site of ligand (**4d**) was used as the binding site with the central coordinate of *X* (141.136), *Y* (159.768), and *Z* directions (179.84). All the rotatable bonds of ligands were allowed to rotate freely, and 100 conformations were generated for each molecule. After docking, the best conformation of each ligand was selected and the corresponding score of the ligand was ranked to find the compounds with the highest score. The top 25 compounds with the highest scores were selected for PAINS remover and manual analysis.

Conclusion

In the present work, indole-2-carboxylic acid (**1**) was proved to inhibit the strand transfer of HIV-1 integrase, and a bis-bidentate chelation with two Mg^{2+} ions was observed for this molecule. Through the optimizations on C2, C3, C5 and C6 of compound **1**, a series of indole-2-carboxylic acid derivatives were designed and synthesized. Biological assay revealed that the introduction of substituted benzene at C3 (**10a**) extended to the hydrophobic cavity near the integrase active site and interacted with Pro142 and Tyr143, and thus increased the inhibitory effect. The structural modifications on C6 of indole core showed that C6 halogenated benzene could bind with the viral DNA dC20 *via* π–π stacking interaction. Through these optimizations, compound **17a** was finally developed, which effectively inhibited the strand transfer of integrase with IC_{50} of 3.11 μM, demonstrating the potential of this molecule as a novel HIV-1 INSTI.

Author contributions

Conceptualization, Zhou, M. and Yan, G.-Y.; methodology, Chen, G.-Q.; software, Chen, T.; validation, Zhang, R.-H., Wang, Y.-C., and Xiong, Q.-Q.; biological tests, Zhang, R.-H. and Zhang, W.-L.; investigation, Zhao, Y.-L.; resources, Li, Y.-J.; data curation, Liao, S.-G.; writing—original draft preparation, Zhang, R.-H. writing—review and editing, Zhou, M. All authors have read and agreed to the published version of the manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by the Foundation of Guizhou Provincial Department of Science and Technology (qiankehejichu [2020]1Z008, qiankehejichu-ZK[2021]yiban551, qiankehezhicheng [2021]yiban425, YQK[2023]031, qiankehejichu-ZK [2024]zhongdian038), the Project of Guizhou Medical University (21NSFCP44, and 20NSP076), and the Project of Guizhou Provincial Department of Education (qianjiaoji[2022]081, [2023] 035).

References

1 K. Yoshimura, *J. Infect. Chemother.*, 2017, **23**, 12–16.
2 V. Simon, D. D. Ho and Q. Abdool Karim, *Lancet*, 2006, **368**, 489–504.
3 R. J. Pomerantz and D. L. Horn, *Nat. Med.*, 2003, **9**, 867–873.
4 S. Viswanathan, A. C. Justice, G. C. Alexander, T. T. Brown, N. R. Gandhi, I. R. McNicholl, D. Rimland, M. C. Rodriguez-Barradas and L. P. Jacobson, *J. Acquired Immune Defic. Syndr.*, 2015, **69**, 493–498.
5 Tazeem, X. Han, Q. Zhou, J. Wei, P. Tien, G. Yang, S. Wu and C. Dong, *RSC Adv.*, 2016, **6**, 95177–95188.
6 Y. Wan, Y. Tian, W. Wang, S. Gu, X. Ju and G. Liu, *RSC Adv.*, 2018, **8**, 40529–40543.
7 N. Bulteel, L. Bansi-Matharu, D. Churchill, D. Dunn, D. Bibby, T. Hill, C. Sabin and M. Nelson, *J. Infect.*, 2014, **68**, 77–84.
8 R. S. Hogg, D. R. Bangsberg, V. D. Lima, C. Alexander, S. Bonner, B. Yip, E. Wood, W. W. Y. Dong, J. S. G. Montaner and P. R. Harrigan, *PLoS Med.*, 2006, **3**, e356.
9 D. Kumar, P. Sharma, Shabu, R. Kaur, M. M. M. Lobe, G. K. Gupta and F. Ntie-Kang, *RSC Adv.*, 2021, **11**, 17936–17964.
10 P. Lesbats, A. N. Engelman and P. Cherepanov, *Chem. Rev.*, 2016, **116**, 12730–12757.
11 M. Su, J. Tan and C.-Y. Lin, *Drug Discovery Today*, 2015, **20**, 1337–1348.
12 K. T. Chiu and R. D. Davies, *Curr. Top. Med. Chem.*, 2004, **4**, 965–977.
13 Y. Pommier, A. A. Johnson and C. Marchand, *Nat. Rev. Drug Discovery*, 2005, **4**, 236–248.
14 O. Delelis, K. Carayon, A. Saïb, E. Deprez and J.-F. Mouscadet, *Retrovirology*, 2008, **5**, 114.
15 D. O. Passos, M. Li, I. K. Jóźwik, X. Z. Zhao, D. Santos-Martins, R. Yang, S. J. Smith, Y. Jeon, S. Forli, S. H. Hughes, T. R. Burke, R. Craigie and D. Lyumkis, *Science*, 2020, **367**, 810–814.

16 D. J. Hazuda, *Curr. Opin. HIV AIDS*, 2012, **7**, 383–389.
17 M. S. A. Gill, S. S. Hassan and N. Ahemad, *Eur. J. Med. Chem.*, 2019, **179**, 423–448.
18 X. Z. Zhao, S. J. Smith, D. P. Maskell, M. Métifiot, V. E. Pye, K. Fesen, C. Marchand, Y. Pommier, P. Cherepanov, S. H. Hughes and T. R. Burke Jr, *J. Med. Chem.*, 2017, **60**, 7315–7332.
19 Y. Wang, S.-X. Gu, Q. He and R. Fan, *Eur. J. Med. Chem.*, 2021, **225**, 113787.
20 V. Nair and G. Chi, *Rev. Med. Virol.*, 2007, **17**, 277–295.
21 B. A. Johns and A. C. Svolto, *Expert Opin. Ther. Pat.*, 2008, **18**, 1225–1237.
22 Q.-Q. He, X. Zhang, L.-M. Yang, Y.-T. Zheng and F. Chen, *J. Enzyme Inhib. Med. Chem.*, 2013, **28**, 671–676.
23 M. M. F. Ismail and M. S. Ayoup, *RSC Adv.*, 2022, **12**, 31032–31045.
24 V. Summa, A. Petrocchi, F. Bonelli, B. Crescenzi, M. Donghi, M. Ferrara, F. Fiore, C. Gardelli, O. Gonzalez Paz, D. J. Hazuda, P. Jones, O. Kinzel, R. Laufer, E. Monteagudo, E. Muraglia, E. Nizi, F. Orvieto, P. Pace, G. Pescatore, R. Scarpelli, K. Stillmock, M. V. Witmer and M. Rowley, *J. Med. Chem.*, 2008, **51**, 5843–5855.
25 K. Shimura and E. N. Kodama, *Antiviral Chem. Chemother.*, 2009, **20**, 79–85.
26 D. E. Dow and J. A. Bartlett, *Infect. Dis. Ther.*, 2014, **3**, 83–102.
27 A. D. Rowan-Nash, B. J. Korry, E. Mylonakis and P. Belenky, *Microbiol. Mol. Biol. Rev.*, 2019, **83**, e00044.
28 S. Swindells, J.-F. Andrade-Villanueva, G. J. Richmond, G. Rizzardini, A. Baumgarten, M. Masiá, G. Latiff, V. Pokrovsky, F. Bredeek, G. Smith, P. Cahn, Y.-S. Kim, S. L. Ford, C. L. Talarico, P. Patel, V. Chounta, H. Crauwels, W. Parys, S. Vanveggel, J. Mrus, J. Huang, C. M. Harrington, K. J. Hudson, D. A. Margolis, K. Y. Smith, P. E. Williams and W. R. Spreen, *N. Engl. J. Med.*, 2020, **382**, 1112–1123.
29 M. Métifiot, B. C. Johnson, E. Kiselev, L. Marler, X. Z. Zhao, T. R. Burke Jr, C. Marchand, S. H. Hughes and Y. Pommier, *Nucleic Acids Res.*, 2016, **44**, 6896–6906.
30 S. Hare, G. N. Maertens and P. Cherepanov, *EMBO J.*, 2012, **31**, 3020–3028.
31 M. A. Wainberg, G. J. Zaharatos and B. G. Brenner, *N. Engl. J. Med.*, 2011, **365**, 637–646.
32 C. B. Hurt, J. Sebastian, C. B. Hicks and J. J. Eron, *Clin. Infect. Dis.*, 2014, **58**, 423–431.
33 S. J. Smith, X. Z. Zhao, T. R. Burke and S. H. Hughes, *Retrovirology*, 2018, **15**, 37.
34 Q. Chen, C. Wu, J. Zhu, E. Li and Z. Xu, *Curr. Top. Med. Chem.*, 2022, **22**, 993–1008.
35 A. K. Sinha, D. Equbal, S. K. Rastogi, S. Kumar and R. Kumar, *Asian J. Org. Chem.*, 2022, **11**, 31–42.
36 X. Han, H. Wu, W. Wang, C. Dong, P. Tien, S. Wu and H.-B. Zhou, *Org. Biomol. Chem.*, 2014, **12**, 8308–8317.
37 M. Palomba, L. Rossi, L. Sancineto, E. Tramontano, A. Corona, L. Bagnoli, C. Santi, C. Pannecouque, O. Tabarrini and F. Marini, *Org. Biomol. Chem.*, 2016, **14**, 2015–2024.
38 H. Xu and M. Lv, *Curr. Pharm. Des.*, 2009, **15**, 2120–2148.
39 M. Sechi, M. Derudas, R. Dallocchio, A. Dessì, A. Bacchi, L. Sannia, F. Carta, M. Palomba, O. Ragab, C. Chan, R. Shoemaker, S. Sei, R. Dayam and N. Neamati, *J. Med. Chem.*, 2004, **47**, 5298–5310.
40 P. A. Patel, N. Kvaratskhelia, Y. Mansour, J. Antwi, L. Feng, P. Koneru, M. J. Kobe, N. Jena, G. Shi, M. S. Mohamed, C. Li, J. J. Kessl and J. R. Fuchs, *Bioorg. Med. Chem. Lett.*, 2016, **26**, 4748–4752.
41 N. J. Cook, W. Li, D. Berta, M. Badaoui, A. Ballandras-Colas, A. Nans, A. Kotecha, E. Rosta, A. N. Engelman and P. Cherepanov, *Science*, 2020, **367**, 806–810.
42 H. Shinkai, in *Successful Drug Discovery*, 2015, pp. 113–128.
43 K. K. Pandey and D. P. Grandgenett, *Retrovirology*, 2008, **2**, 11–16.
44 P. Kapewangolo, T. Tawha, T. Nawinda, M. Knott and R. Hans, *S. Afr. J. Bot.*, 2016, **106**, 140–143.
45 F. Jiang, W. Chen, K. Yi, Z. Wu, Y. Si, W. Han and Y. Zhao, *Clin. Immunol.*, 2010, **137**, 347–356.
46 C. Zhang, Q. Xie, C. C. Wan, Z. Jin and C. Hu, *Curr. Med. Chem.*, 2021, **28**, 4910–4934.
47 C. W. Murray, M. G. Carr, O. Callaghan, G. Chessari, M. Congreve, S. Cowan, J. E. Coyle, R. Downham, E. Figueroa, M. Frederickson, B. Graham, R. McMenamin, M. A. O'Brien, S. Patel, T. R. Phillips, G. Williams, A. J. Woodhead and A. J. A. Woolford, *J. Med. Chem.*, 2010, **53**, 5942–5955.
48 N. Blomberg, D. A. Cosgrove, P. W. Kenny and K. Kolmodin, *J. Comput.-Aided Mol. Des.*, 2009, **23**, 513–525.
49 M. M. Hann, A. R. Leach and G. Harper, *J. Chem. Inf. Comput. Sci.*, 2001, **41**, 856–864.
50 M. A. McCoy and D. F. Wyss, *J. Am. Chem. Soc.*, 2002, **124**, 11758–11763.