Supporting Information for

Original article

Discovery of novel sulfonamide substituted indolylarylsulfones as potent HIV-1 inhibitors with better safety profiles

Shenghua Gao^{a,b,d}, Letian Song^a, Yusen Cheng^a, Fabao Zhao^a, Dongwei Kang^{a,b}, Shu Song^a, Mianling Yang^a, Bing Ye^a, Wei Zhao^e, Yajie Tang^e, Erik De Clercq^c, Christophe Pannecouque^{c,*}, Peng Zhan^{a,b,*}, Xinyong Liu^{a,b,*}

^aDepartment of Medicinal Chemistry, Key Laboratory of Chemical Biology, Ministry of Education, School of Pharmaceutical Sciences, Shandong University, Ji'nan 250012, China

^bChina-Belgium Collaborative Research Center for Innovative Antiviral Drugs of Shandong Province, Ji'nan 250012, China

^cRega Institute for Medical Research, K.U.Leuven, Minderbroedersstraat 10, Leuven B-3000, Belgium

^dShenzhen Research Institute of Shandong University, A301 Virtual University Park in South District of Shenzhen, Shenzhen 518057, China.

^eState Key Laboratory of Microbial Technology, Shandong University, Qingdao 266237, China

Received 26 October 2022; received in revised form 17 December 2022; accepted 26 December 2022

*Corresponding authors.

E-mail addresses: christophe.pannecouque@kuleuven.be (Christophe Pannecouque); zhanpeng1982@sdu.edu.cn (Peng Zhan); xinyongl@sdu.edu.cn (Xinyong Liu.).

Table of Contents

Methods	S1
<i>In vivo</i> pharmacokinetics study	
Acute and subacute toxicity assay procedure	
Molecular docking simulation	
Table S1. RT inhibitory activity of target compounds	
¹ H NMR and ¹³ C NMR spectra of final compounds	
HRMS spectra of representative final compounds	
HPLC trace of representative final compounds	

Animal husbandry conditions

Animals are group housed during acclimation and study. The animal room environment is controlled (target conditions: temperature 20 °C to 25 °C, relative humidity 40% to 70%, 12 hours' artificial light and 12 hours' dark). In the acute toxicity experiment, subacute toxicity experiment, and PK study, All the animals will be fasted for 12 hours before dosing. The age of Sprague-Dawley rats in PK study is 6-8 weeks, and the body weights range from 180 - 220 g. The age of Kunming mice in acute/subacute toxicity study is 3-4 weeks, and the body weights are about 20 g. *In vivo* pharmacokinetics study

The R₁₀L₄ (14.56 mg) was dissolved in a mixture of 0.713 mL DMSO, 1.427 mL polyethylene glycol-15 hydroxystearate (Solutol) and 12.128 mL normal saline for a 1.00 mg/mL solution. Six male Sprague-Dawley rats were randomly divided into two groups to receive intravenous $(2 \text{ mg} \cdot \text{kg}^{-1})$ or oral administration $(10 \text{ mg} \cdot \text{kg}^{-1})$ of a test drug. Blood samples of the intravenous group were collected from the sinus jugularis into heparinized centrifugation tubes at 5 min, 15 min, 30 min, 1 h, 2 h, 4 h, 8 h, and 24 h after dosing, and blood samples of the oral administration group were collected at 15 min, 30 min, 1 h, 2 h, 4 h, 8 h, and 24 h after dosing (250 µL of blood each time). The blood samples were centrifuged at 6000 rpm for 3 min to separate plasma, which was stored at -20 °C for next steps of analysis. Upon LC-MS analysis, plasma samples were thawed, shaken for 30 seconds, and centrifuged at 4000 rpm for 0.5 min under 4°C. The internal standard solution contains 50 µg/L tolbutamide in a 1:1 mixture of MeOH and ACN. Then 40 µL of plasma, calibration standard, quality control and dilution quality control, blank sample were added to the 96-well plate respectively. Each sample was quenched with 200 μ L of internal standard respectively (blank sample was quenched with 200 μ L of ACN:MeOH = 1:1), and then the mixture was vortex-mixed for 5 min and centrifuged for 5 min at 14000 rpm, 4 °C. Subsequently, 100 µL supernatant was transferred to another clean 96-well plate and diluted with 100 µL of ACN, vortex-mixed for 5 min and centrifuged for 5 min at 800 rpm, 4 °C, then injected for LC-MS/MS analysis. All samples were quantified with an LC-MS/MS-AR Triple Quad 6500+ (SCIEX, USA) and the liquid phase system was ExionLC. The mobile phase was 0.1% formic acid-water/ACN with gradient elute at a

flow rate of 0.8 mL/min (total time 2 min), and the test wavelength was 225 nm. All blood samples were centrifuged in an Eppendorf 5424R centrifuge and quantified with an LC-MS/MS-AR Triple Quad 6500+ (SCIEX, USA). Pharmacokinetic parameters were calculated by WinNolin 8.2 software.

Acute Toxicity Experiment

A group of Kunming mice (three male and three female) was supplied by the Animal Experimental Center of Shandong University. The mice were fasted for 12 h; then, a suspension of $\mathbf{R}_{10}\mathbf{L}_4$ in 0.5% CMC-Na and 3% DMSO at the concentration of 100 mg·mL⁻¹ was administered intragastrically to provide a dose of 800 mg·kg⁻¹. Experimental group is consisted of six mice (three males and three females).

Subacute Toxicity Experiment

Another batch of three male and three female Kunming mice was randomly divided into four groups (n = 3): male test group, female test group, male vehicle group, and female vehicle group. All mice were deprived of feed for 12 h and then the mice in the test groups were given 50 mg· kg⁻¹ p.o. of **R**₁₀L₄ once every other day for 14 days (D0, D2, D4, D6, D8, D10, and D12), while the mice in control groups received the same volume of blank solution. The mice were weighed before each dosing. All the mice were dissected after euthanasia at D14, and the heart, liver, spleen, lung, and kidney were extracted. These organs were sliced and examined by HE staining.

Molecular docking simulation

The PDB files (6C0N, 6C0L, 2OPQ) are downloaded from Protein Database Bank (rcsb.org). Protein preparation, ligand preparation and docking studies were performed by Schrödinger platform (Maestro version 13.2.128). Protein models are preprocessed, optimized, and minimized under default settings of Protein Preparation Workflow. Waters, ions and unrelated small molecules are deleted manually. Then ligand molecules in NNIBP of each model were used to locate the center of receptor grid boxes, and automatically removed in the established receptor grids. The structure of $\mathbf{R}_{10}\mathbf{L}_4$ and IAS-0 were prepared by LigPrep under default settings. Ligand docking was operated by Glide XP (extra precision) module, under default settings and unadjusted van der Wall radii scaling. Upon completion, docking scores and interactions were shown in Schrödinger. 3D files are exported and visualized in Pymol (Version 2.5.2).

		-	
Compound	$IC_{50}\pm SD(\mu M)^a$	Compound	$IC_{50}\pm SD\left(\mu M\right)$
R_1L_2	0.075 ± 0.005	R_1L_3	0.160 ± 0.004
R_1L_4	0.140 ± 0.007	R_1L_5	0.130 ± 0.007
R_2L_2	0.094 ± 0.002	R_2L_3	0.130 ± 0.002
R_2L_4	0.130 ± 0.003	R_2L_5	0.160 ± 0.000
R_3L_2	0.100 ± 0.007	R_3L_3	0.120 ± 0.005
R ₃ L ₄	0.150 ± 0.013	R_3L_5	0.160 ± 0.016
R_4L_2	0.066 ± 0.018	R_4L_3	0.100 ± 0.005
R4L4	0.150 ± 0.009	R_4L_5	0.150 ± 0.012
R_5L_2	0.110 ± 0.032	R_5L_3	0.100 ± 0.003
R_5L_4	0.120 ± 0.003	R_5L_5	0.160 ± 0.002
R_6L_2	0.120 ± 0.011	R_6L_3	0.160 ± 0.016
R_6L_4	0.150 ± 0.015	R_6L_5	0.130 ± 0.009
R_7L_2	0.055 ± 0.002	R ₇ L ₃	0.078 ± 0.013
R_7L_4	0.094 ± 0.003	R_7L_5	0.120 ± 0.011
R_8L_2	0.140 ± 0.015	R_8L_3	0.190 ± 0.035
R_8L_4	0.190 ± 0.017	R_8L_5	0.220 ± 0.017
R ₉ L ₂	0.055 ± 0.006	R ₉ L ₃	0.084 ± 0.011
R9L4	0.130 ± 0.014	R_9L_5	0.120 ± 0.015
$R_{10}L_2$	0.060 ± 0.009	R ₁₀ L ₃	0.041 ± 0.009
$R_{10}L_{4}$	0.077 ± 0.009	$R_{10}L_5$	0.095 ± 0.012
$R_{11}L_2$	0.083 ± 0.007	$R_{11}L_3$	0.076 ± 0.019
R ₁₁ L ₄	0.120 ± 0.015	$R_{12}L_2$	0.083 ± 0.011
$R_{12}L_3$	0.130 ± 0.016	$R_{12}L_4$	0.150 ± 0.017
$R_{13}L_2$	0.057 ± 0.012	$R_{14}L_2$	0.180 ± 0.000
IAS-0	0.018 ± 0.002	EFV ^b	0.030 ± 0.000

Table S1. RT inhibitory activity of target compounds.

^aIC₅₀: Compound concentration required to inhibit 50% incorporation of biotin-labeled UTP by HIV-1 (WT) RT. ^bAll data were obtained from one batch with the same method.

¹H NMR and ¹³C NMR spectra of final compounds

Feb23-2021-R1L2.1.fid -12.95 1,7,69 1,7,69 1,7,61 1,7,61 1,7,61 1,7,61 1,7,61 1,7,61 1,7,61 1,7,61 1,7,61 1,7,61 1,7,61 1,7,73 1,2,23 1, -9.02 -9.00 -8.99 -7.94 -7.93 -5500 -5000 -4500 , |||| ſ ſ -4000 -3500 -3000 -2500 -2000 -1500 -1000 -500 -0 2.14 ± 2.09 ± 3.32 € 0.98⊣ 1.05 H 1.00 2.11 1.12 3.23 1.12 3.22 1.01 --500 9 16 13 15 14 12 'n 10 6 3 0 8 fl (ppm) 5 4 2

¹H NMR (600 MHz, DMSO- d_6) of $\mathbf{R}_1\mathbf{L}_2$

¹³C NMR (150 MHz, DMSO- d_6) of R_1L_2



¹H NMR (600 MHz, DMSO- d_6) of R_1L_3



¹³C NMR (150 MHz, DMSO- d_6) of R₁L₃



¹H NMR (600 MHz, DMSO- d_6) of R₁L₄



¹³C NMR (150 MHz, DMSO-*d*₆) of **R**₁L₄



¹H NMR (600 MHz, DMSO- d_6) of R_1L_5



¹³C NMR (150 MHz, DMSO- d_6) of R₁L₅



¹H NMR (600 MHz, DMSO- d_6) of R_2L_2



¹³C NMR (150 MHz, DMSO- d_6) of R_2L_2







¹³C NMR (150 MHz, DMSO- d_6) of R_2L_3





¹H NMR (600 MHz, DMSO- d_6) of R_2L_4

¹³C NMR (150 MHz, DMSO-*d*₆) of **R**₂L₄



¹H NMR (600 MHz, DMSO- d_6) of R_2L_5



¹³C NMR (150 MHz, DMSO-*d*₆) of **R**₂**L**₅





¹H NMR (600 MHz, DMSO- d_6) of $\mathbf{R}_3\mathbf{L}_2$

¹³C NMR (150 MHz, DMSO- d_6) of R₃L₂





¹H NMR (600 MHz, DMSO- d_6) of **R**₃L₃

¹³C NMR (150 MHz, DMSO- d_6) of R₃L₃







¹³C NMR (150 MHz, DMSO- d_6) of R₃L₄





¹H NMR (600 MHz, DMSO- d_6) of **R**₃L₅

¹³C NMR (150 MHz, DMSO- d_6) of R₃L₅





¹H NMR (600 MHz, DMSO- d_6) of R_4L_2

¹³C NMR (150 MHz, DMSO- d_6) of R₄L₂





¹H NMR (600 MHz, DMSO- d_6) of R₄L₃

¹³C NMR (150 MHz, DMSO- d_6) of R₄L₃





¹H NMR (600 MHz, DMSO- d_6) of R₄L₄

¹³C NMR (150 MHz, DMSO- d_6) of R₄L₄



¹H NMR (600 MHz, DMSO- d_6) of R₄L₅



¹³C NMR (150 MHz, DMSO-*d*₆) of **R**₄L₅







¹³C NMR (150 MHz, DMSO- d_6) of R₅L₂





¹H NMR (600 MHz, DMSO- d_6) of $\mathbf{R}_5\mathbf{L}_3$

¹³C NMR (150 MHz, DMSO-*d*₆) of **R**₅L₃





¹H NMR (600 MHz, DMSO- d_6) of $\mathbf{R}_5\mathbf{L}_4$

¹³C NMR (150 MHz, DMSO-*d*₆) of **R**₅L₄





¹H NMR (600 MHz, DMSO- d_6) of R₅L₅

¹³C NMR (150 MHz, DMSO- d_6) of R₅L₅





¹H NMR (600 MHz, DMSO- d_6) of R₆L₂

¹³C NMR (150 MHz, DMSO- d_6) of R₆L₂





¹H NMR (600 MHz, DMSO- d_6) of R_6L_3

¹³C NMR (150 MHz, DMSO- d_6) of R₆L₃





-3500

-2500

-2000

-1000

-500

2.09-≖

5.79-

4.27-I

¹H NMR (600 MHz, DMSO- d_6) of R_6L_4

¹³C NMR (150 MHz, DMSO- d_6) of R₆L₄

-16.0



0.94 3.09 1.92 0.99

fl (ppm)

1.00-



¹H NMR (600 MHz, DMSO- d_6) of R_6L_5

¹³C NMR (150 MHz, DMSO- d_6) of R₆L₅



¹H NMR (600 MHz, DMSO- d_6) of $\mathbf{R}_7 \mathbf{L}_2$



$^{13}\mathrm{C}$ NMR (150 MHz, DMSO- $d_6)$ of R_7L_2



¹H NMR (600 MHz, DMSO- d_6) of R_7L_3



¹³C NMR (150 MHz, DMSO-*d*₆) of **R**₇**L**₃





¹H NMR (600 MHz, DMSO- d_6) of $\mathbf{R}_7\mathbf{L}_4$

¹³C NMR (150 MHz, DMSO-*d*₆) of **R**₇L₄



¹H NMR (600 MHz, DMSO- d_6) of R_7L_5



$^{13}\mathrm{C}$ NMR (150 MHz, DMSO- $d_6)$ of R₇L₅







¹³C NMR (150 MHz, DMSO- d_6) of **R**₈L₂





¹H NMR (600 MHz, DMSO- d_6) of R₈L₃

¹³C NMR (150 MHz, DMSO- d_6) of R₈L₃




¹H NMR (600 MHz, DMSO- d_6) of **R**₈L₄

¹³C NMR (150 MHz, DMSO-*d*₆) of **R**₈L₄





¹H NMR (600 MHz, DMSO- d_6) of **R**₈L₅

¹³C NMR (150 MHz, DMSO- d_6) of R₈L₅





¹H NMR (600 MHz, DMSO- d_6) of **R**₉L₂

¹³C NMR (150 MHz, DMSO- d_6) of **R**₉L₂





¹H NMR (600 MHz, DMSO- d_6) of R₉L₃

¹³C NMR (150 MHz, DMSO- d_6) of **R**₉L₃





¹H NMR (600 MHz, DMSO- d_6) of **R**₉L₄

¹³C NMR (150 MHz, DMSO-*d*₆) of **R**₉L₄



¹H NMR (600 MHz, DMSO- d_6) of **R**₉L₅



¹³C NMR (150 MHz, DMSO-*d*₆) of **R**₉L₅





¹H NMR (600 MHz, DMSO- d_6) of R₁₀L₂

¹³C NMR (150 MHz, DMSO- d_6) of R₁₀L₂





¹H NMR (600 MHz, DMSO- d_6) of R₁₀L₃

¹³C NMR (150 MHz, DMSO- d_6) of R₁₀L₃







¹³C NMR (150 MHz, DMSO- d_6) of R₁₀L₄



¹H NMR (600 MHz, DMSO- d_6) of R₁₀L₅



¹³C NMR (150 MHz, DMSO- d_6) of R₁₀L₅





¹H NMR (600 MHz, DMSO- d_6) of R₁₁L₂

¹³C NMR (150 MHz, DMSO-*d*₆) of **R**₁₁L₂





¹H NMR (600 MHz, DMSO- d_6) of R₁₁L₃

¹³C NMR (150 MHz, DMSO-*d*₆) of **R**₁₁L₃







¹³C NMR (150 MHz, DMSO- d_6) of R₁₁L₃





¹H NMR (600 MHz, DMSO- d_6) of R₁₂L₂

¹³C NMR (150 MHz, DMSO- d_6) of $R_{12}L_2$





¹H NMR (600 MHz, DMSO- d_6) of R₁₂L₃

13 C NMR (150 MHz, DMSO- d_6) of R₁₂L₃



¹H NMR (600 MHz, DMSO- d_6) of R₁₂L₄



¹³C NMR (150 MHz, DMSO- d_6) of R₁₂L₄







¹³C NMR (150 MHz, DMSO- d_6) of R₁₃L₂





¹H NMR (600 MHz, DMSO- d_6) of R₁₄L₂

¹³C NMR (150 MHz, DMSO-*d*₆) of **R**₁₄L₂



HRMS spectra of representative final compounds HRMS spectrum of R₁L₂



HRMS spectrum of R₁L₃



HRMS spectrum of R₁L₄



HRMS spectrum of R₁L₅



HRMS spectrum of R₂L₂



HRMS spectrum of R₂L₃



HRMS spectrum of R₂L₄



HRMS spectrum of R₂L₅



HRMS spectrum of R₃L₂



HRMS spectrum of R₃L₃



HRMS spectrum of R₃L₄



HRMS spectrum of R₃L₅



HRMS spectrum of R₄L₂



HRMS spectrum of R₄L₃



HRMS spectrum of R₄L₄



HRMS spectrum of R₄L₅



HRMS spectrum of R₅L₂



HRMS spectrum of R₅L₃



HRMS spectrum of R₅L₄



HRMS spectrum of R₅L₅



HRMS spectra of R₆L₂





HRMS spectra of R₆L₃



HRMS spectra of R₆L₄





HRMS spectra of R₆L₅





HRMS spectrum of R₇L₂



HRMS spectrum of R7L3



HRMS spectrum of R7L4



HRMS spectrum of R7L5



HRMS spectrum of R₈L₂



HRMS spectrum of R₈L₃



HRMS spectrum of R₈L₄



HRMS spectrum of R₈L₅



HRMS spectrum of R₉L₂



HRMS spectrum of R₉L₃



HRMS spectrum of R₉L₄



HRMS spectrum of R₉L₅



HRMS spectrum of R₁₀L₂



HRMS spectrum of R₁₀L₃



HRMS spectrum of R₁₀L₄



HRMS spectrum of R₁₀L₅



HRMS spectrum of R₁₁L₂



HRMS spectrum of R₁₁L₃



HRMS spectrum of R₁₁L₄



HRMS spectrum of R₁₂L₂



HRMS spectrum of R₁₂L₃



HRMS spectrum of R₁₂L₄



HRMS spectrum of R₁₃L₂



HRMS spectrum of R₁₄L₂


HPLC trace of representative final compounds



HPLC trace of R₂L₂

HPLC trace of R₂L₄



HPLC trace of R₂L₅



HPLC trace of R₃L₂



HPLC trace of R₃L₄













HPLC trace of **R**₄L₄

HPLC trace of R₅L₂



HPLC trace of R₅L₃



HPLC trace of R₅L₅



HPLC trace of **R**₆L₂



HPLC trace of R₆L₃







HPLC trace of R₇L₂



HPLC trace of **R**₇**L**₄



HPLC trace of R₈L₂



HPLC trace of **R**₈L₃



HPLC trace of R₈L₄



HPLC trace of R₈L₅



HPLC trace of R₉L₂







HPLC trace of R₉L₅



HPLC trace of **R**₁₀L₂



HPLC trace of R₁₀L₃



HPLC trace of R₁₀L₄



Peak#	Ret. time	Square	Height	Square%	Height%
1	1.624	15189	902	0.351	0.243
2	2.165	8092	684	0.187	0.184
3	2.783	4299117	369902	99.431	99.527
4	3.914	1320	173	0.031	0.046
Total		4323718	371661	100.000	100.000

HPLC trace of R₁₀L₅



mV 2.824 750 500-250 3.301 *2.474 *4.006 *1.838 2.151 1.600 0 2 3 4 5 6 ò 1 min Square% Height% Peak# Ret. time Square Height 0. 196 0. 030 0. 163 0. 038 99. 037 0. 451 0. 028 0. 058 1592 479 931 474 829578 18878 2908 15689 1.600 0.190 $\begin{array}{r}
 1.600 \\
 1.838 \\
 2.151 \\
 2.474 \\
 2.824 \\
 3.301 \\
 4.006 \\
 4.165 \\
\end{array}$ 0.057 23 15689 3630 9547890 43442 2725 5580 9640741 0.056 98.831 0.644 4 5 6 7 8 5403 336 597 0.040 4.165 9640741 839391 Total 100.000 100.000

HPLC trace of $R_{11}L_2$





HPLC trace of R₁₁L₄



HPLC trace of $R_{12}L_2$



HPLC trace of $R_{12}L_4$



HPLC trace of $R_{13}L_2$

