Supplementary Information for

A novel G protein-biased and subtype selective agonist for a G protein-coupled receptor discovered from screening herbal extracts

Bingjie Zhang[1‡], Simeng Zhao[1‡], Dehua Yang[2], Yiran Wu[1], Ye Xin[1], Haijie Cao[1], Xi-Ping Huang[5], Xiaoqing Cai[2], Wen Sun[2,3], Na Ye[4], Yueming Xu[1], Yao Peng[1], Suwen Zhao[1,6], Zhi-Jie Liu[1,6], Guisheng Zhong[1,6*], Ming-Wei Wang[2,6,7*], Wenqing Shui[1,6*]

[1]iHuman Institute, ShanghaiTech University, Shanghai 201210, China

[2]The National Center for Drug Screening and the CAS Key Laboratory of Receptor Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China

[3]University of Chinese Academy of Sciences, Beijing 100049, China

[4]Jiangsu Key Laboratory of Neuropsychiatric Diseases and College of Pharmaceutical Sciences, Soochow University, Suzhou, Jiangsu 215123, China

[5]Department of Pharmacology, School of Medicine, University of North Carolina , Chapel Hill, NC 27599, USA

[6]School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China

[7]School of Pharmacy, Fudan University, Shanghai 201203, China

‡Equal contribution

Contents

Supplementary Table 3. Identification of eight aporphine analogues from *Stephania tetrandra* (ST) as putative 5-HT_{2c} ligands.

Each compound was identified from screening the total extract (ST-total) or a specific fraction of the extract (ST-F1/F2/F3) as indicated in "Screening source".

Compound	Name	Substituent	Formula	Accurate mass	Mass error (ppm)	RT (min) In extract	RT (min) In standard	Screening source
1857	Asimilobine	R_1 = OH R_2 = OCH_3 R_3 = H R_4 = R_5 = R_6 = H	C17H17N1O2	267.1259	1.5	6.59	6.63	ST-F3
15781	Nornuciferine	R_1 = OCH_3 R_2 = OCH_3 R_3 = H R_4 = R_5 = R_6 = H	C18H19N1O2	281.1416	0.9	9.36	9.44	ST-F3
14148	N-Methylasimilobine	R_1 = OH R_2 = OCH_3 R_3 = CH_3 R_4 = R_5 = R_6 = H	C18H19N1O2	281.1416	0.2	5.01	4.99	ST-total
15856	Nuciferine	R_1 = OCH_3 R_2 = OCH_3 R_3 = CH_3 R_4 = R_5 = R_6 = H	C19H21N1O2	295.1572	-0.9	9.59	9.56	ST-F3
8513	Glaucine	R_1 = OCH_3 R_2 = OCH_3 R_3 = CH_3 R_4 = H R_5 = R_6 = OCH_3	C21H25N1O4	355.1784	-0.2	6.42	ND	ST-total
9454	Hernovine	R_1 = OH R_2 = OCH_3 R_3 = H R_4 = OCH_3 R_5 = OH R_6 = H	C18H19N1O4	313.1314	-0.9	4.14	ND	ST-F2
12573	Laurolitsine	R_1 = OH R_2 = OCH_3 R_3 = H R_4 = H R_5 = OCH_3 R_6 = OH	C18H19N1O4	313.1314	2.4	0.6	ND	ST-F1
12915	Lirinidine	R_1 = OCH_3 R_2 = OH R_3 = CH_3 R_4 = R_5 = R_6 = H	C18H19N1O2	281.1416	-5.2	7.71	ND	ST-total, ST-F1

RT, retention time. ND indicates the standard is not available.

Supplementary Table 4. K_i values of four aporphines and lorcaserin at 5-$HT_{2A/2B/2C}$ receptors.

Compound	h5-HT_{2A} [^{3}H]-Ketanserin K_i nM (pK_i ± SEM)	h5-HT_{2B} [^{3}H]-LSD K_i nM (pK_i ± SEM)	h5-HT_{2C} [^{3}H]-Mesulergine K_i nM (pK_i ± SEM)
1857	375.9 (6.43 ± 0.08)	74.2 (7.13 ± 0.05)	193.9 (6.71 ± 0.04)
15781	925.0 (6.03 ± 0.06)	263.5 (6.58 ± 0.06)	446.2 (6.35 ± 0.05)
14148	60.0 (7.22 ± 0.04)	3.38 (8.47 ± 0.03)	20.9 (7.68 ± 0.03)
15856	149.7 (6.83 ± 0.05)	14.3 (7.85 ± 0.03)	22.2 (7.65 ± 0.04)
Lorcaserin	765.1 (6.11 ± 0.06)	140.2 (6.85 ± 0.05)	123.1 (6.91 ± 0.05)

Data represent K_i and pK_i from three independent experiments.

Supplementary Table 5. IC_{50}/EC_{50} values of 1857 and lorcaserin at wild-type and mutant 5-HT_{2C} receptors.

	1857			Lorcaserin		
	Radiolabeled ligand binding assay (pIC_{50} ± SEM)	G_q calcium flux assay (pEC_{50} ± SEM)	β-arrestin2 recruitment (pEC_{50} ± SEM)	Radiolabeled ligand binding assay (pIC_{50} ± SEM)	G_q calcium flux assay (pEC_{50} ± SEM)	β-arrestin2 recruitment (pEC_{50} ± SEM)
WT	5.80 ± 0.07	6.10 ± 0.05	ND	5.98 ± 0.08	8.41 ± 0.03	6.94 ± 0.07
V135L$^{3.33}$	6.24 ± 0.05	6.89 ± 0.05	ND	6.10 ± 0.09	8.28 ± 0.04	ND
A222I$^{5.46}$	6.95 ± 0.04	ND	ND	6.84 ± 0.06	8.47 ± 0.03	ND
F327L$^{6.52}$	5.52 ± 0.08	5.41 ± 0.11	NA	6.58 ± 0.07	8.00 ± 0.04	NA
V354A$^{7.39}$	6.54 ± 0.11	8.19 ± 0.04	5.64 ± 0.07	6.98 ± 0.16	8.38 ± 0.03	6.61 ± 0.07
V354F$^{7.39}$	5.80 ± 0.10	7.27 ± 0.03	ND	5.73 ± 0.12	7.34 ± 0.03	ND

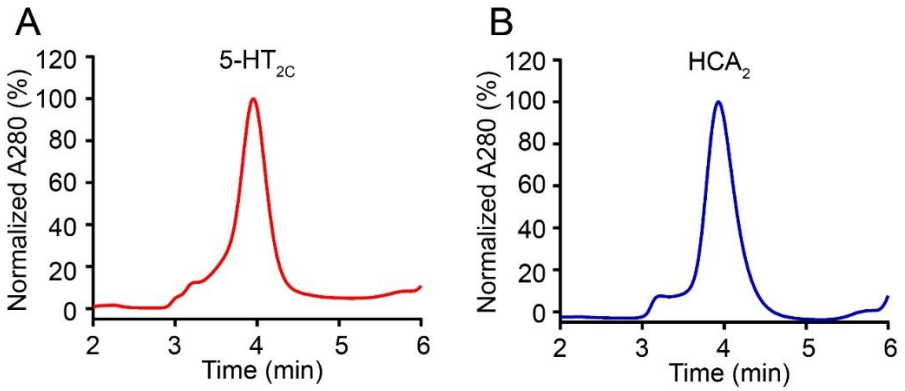
IC_{50} values were estimated from competitive inhibition of radiolabeled ligand binding using a three-parameter logistic equation [log (inhibitor versus response)] in Prism 7 (GraphPad). ND indicates no detectable activity and NA indicates this mutant was not assayed.

Supplementary Table 6. Blood-brain barrier penetrance of 1857 in mice.

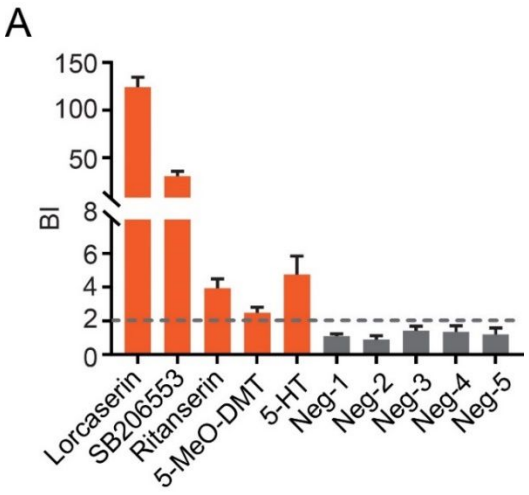
Sample	Plasma	Brain	Ratio (C_{brain}/C_{plasma})
	Conc. (ng/mL)	Conc. (ng/g)	
1857-30 min	659.7 ± 176.7	11060.3 ± 2774.9	16.8
1857-240 min	36.7 ± 3.4	236.0 ± 33.1	6.4

1857 brain and plasma pharmacokinetics were evaluated after a single intravenous dose of 10 mg/kg in male C57/6BJ mice (n=3 for each time point). After 30 or 240 minutes (min), mice were anesthetized to obtain whole blood and brain samples. Plasma was prepared from sodium heparin-treated whole blood and separated by centrifugation. Plasma and brain samples were frozen and stored at -80°C until analysis. Compound concentration (Conc.) was determined using single reaction monitoring (SRM)-based MS analysis. Compound with a brain/plasma ratio of greater than 1 is considered to be able to cross the blood-brain barrier freely. Data presented are means ± SEM.

The SRM method is as follows: A Waters Acquity I-Class UPLC coupled with a Xevo TQ-S mass spectrometer was used for determination of each compound concentration. LC separation was performed on a Waters Acquity BEH C18 (2.1×50 mm, 1.7 μm) column with the gradient being 0-1.2 min, B at 5%; 1.2-1.5 min, B at 70-95%; then re-equilibrating for 5 min. Mobile phase was water/0.1% formic acid (A solvent) and acetonitrile/0.1% formic acid (B solvent). MS instrument operated in an MRM mode with positive electrospray ionization to monitor analyte transitions of m/z 268.068 > 190.976 and m/z 282.068 > 265.07 for measurement of compound 1857 and 15781, respectively.

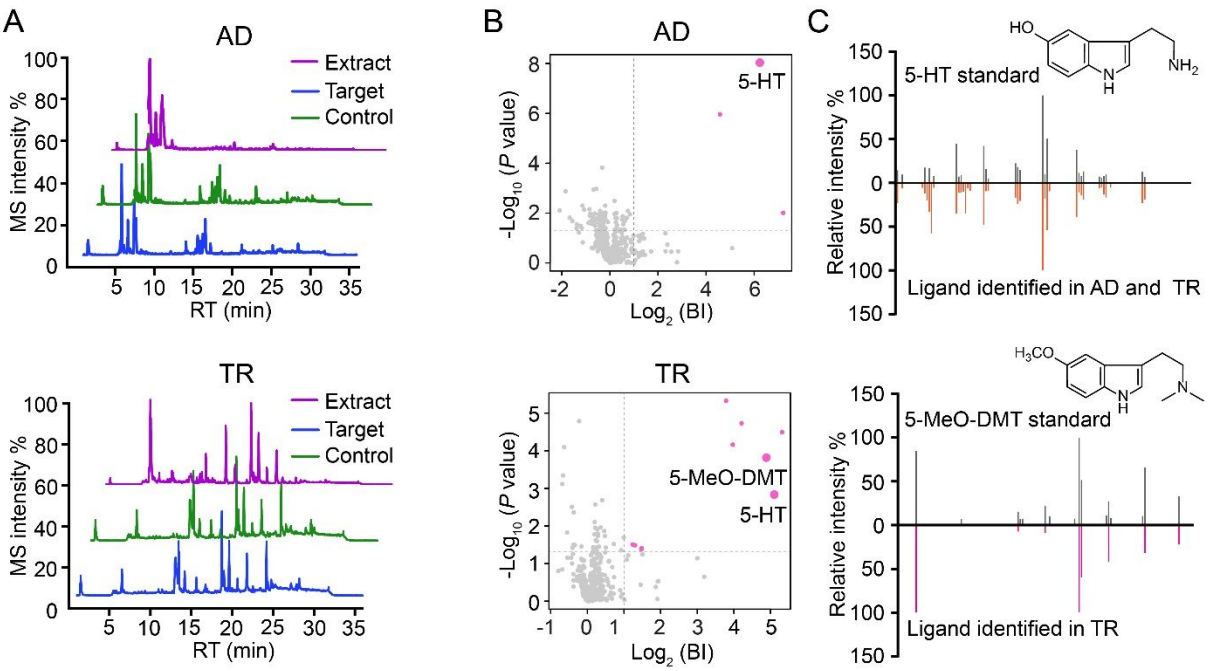


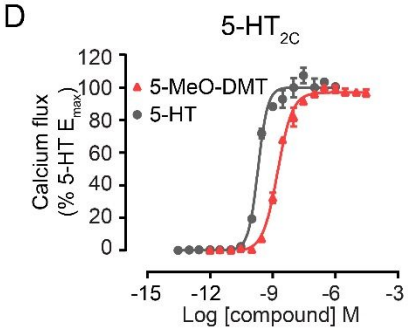
Supplementary Figure 1. SEC analysis of purified 5-HT_{2C} and HCA_2 proteins.



Ligand	pK_i
Lorcaserin	7.82
SB206553	8.00
Ritanserin	8.74
5-MeO-DMT	7.38
5-HT	8.16

Supplementary Figure 2. Affinity MS screening of a simple mixture of five known 5-HT_{2C} ligands (red bars) and five unrelated compounds (grey bars). Binding index (BI) refers to the ratio of MS intensity of each compound detected in the 5-HT_{2C} target *vs.* control. Data were obtained from three independent experiments. Error bars represent SEM. Binding affinity data for known ligands were obtained from ChEMBL database.



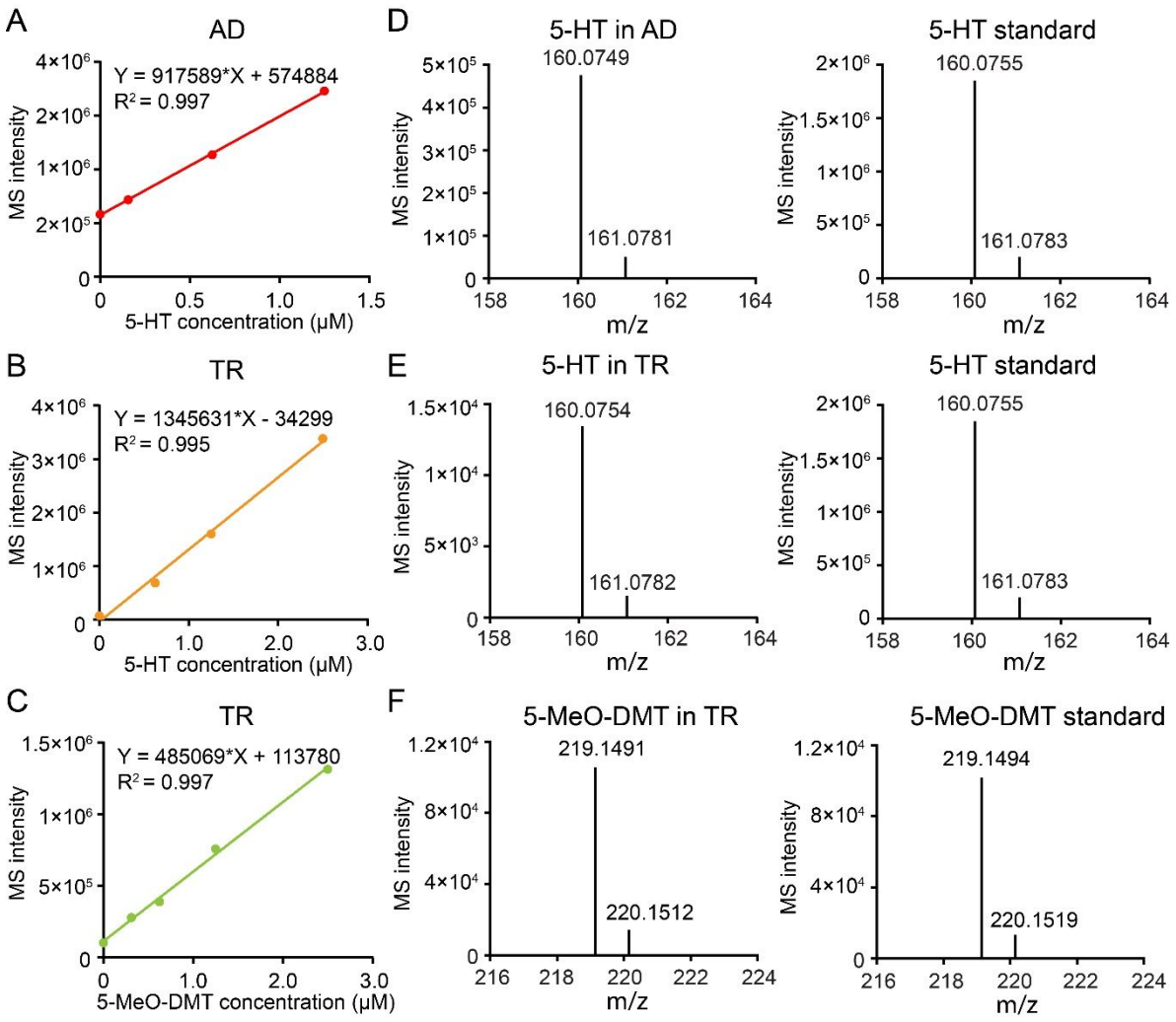


E

Ligand	Extract	Concentration in the extract (*a*)	EC_{50} of pure agonist (*b*)	EC_{50} of extract (*c*)	Ligand contribution (*d**)
5-HT	AD	0.22%	0.2 nM	94.8 nM	100%
5-HT	TR	0.02%	0.2 nM	1.4 μM	100%
5-MeO-DMT	TR	0.26%	5.4 nM	1.4 μM	67%

**d* = [(*a*×*c*)/*b*]%
Ligand contribution is defined as 100% if calculated *d* > 100%.

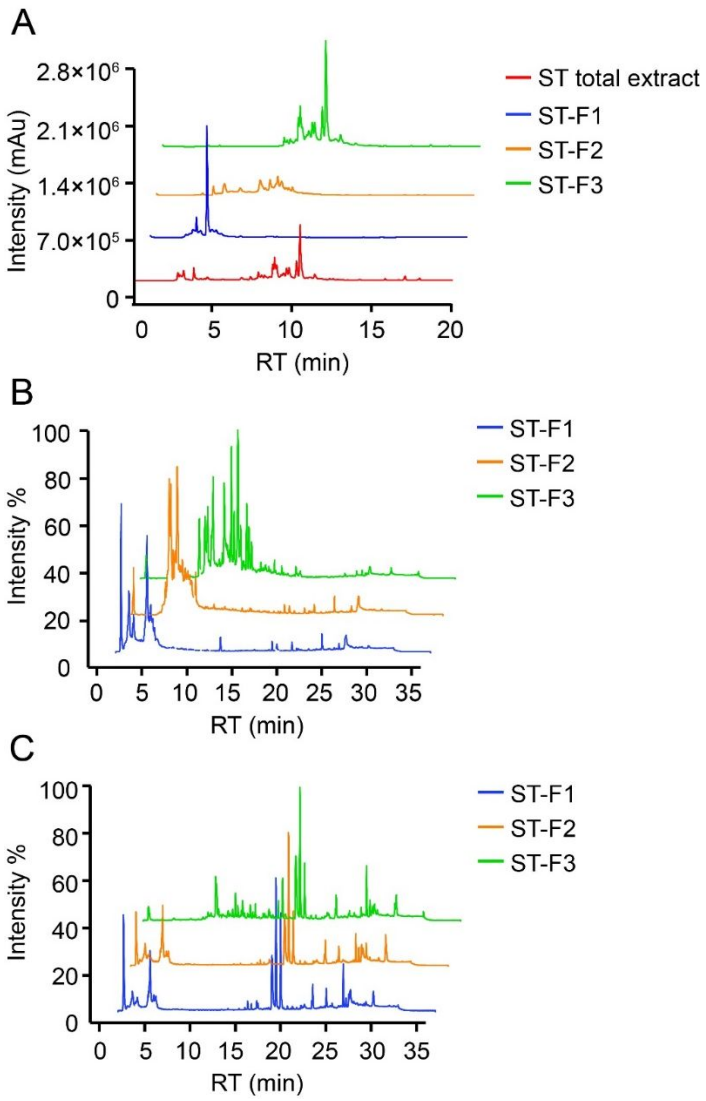
Supplementary Figure 3. Identification of known 5-HT_{2C} agonists by screening crude extracts of *Aristolocxsshia debilis* (AD) and *Tetradium ruticarpum* (TR). (A) Representative LC-MS chromatograms of crude extract, 5-HT_{2C} target and control for AD (top) or TR (bottom). (B) Initial hits (pink dots) from screening crude extract of AD (top) or TR (bottom) by affinity MS combined with metabolomics. 5-HT and 5-MeO-DMT identified are marked by larger dots. BI, binding index. (C) Structural validation of 5-HT and 5-MeO-DMT by MSMS analysis of identified ligands and the standard. (D) 5-HT_{2C} mediated calcium mobilization elicited by 5-HT (EC_{50} = 0.2 nM) and 5-MeO-DMT (EC_{50} = 5.4 nM). (E) Calculated contribution of 5-HT and 5-MeO-DMT to the total 5-HT_{2C} activity manifested by AD and TR extracts.



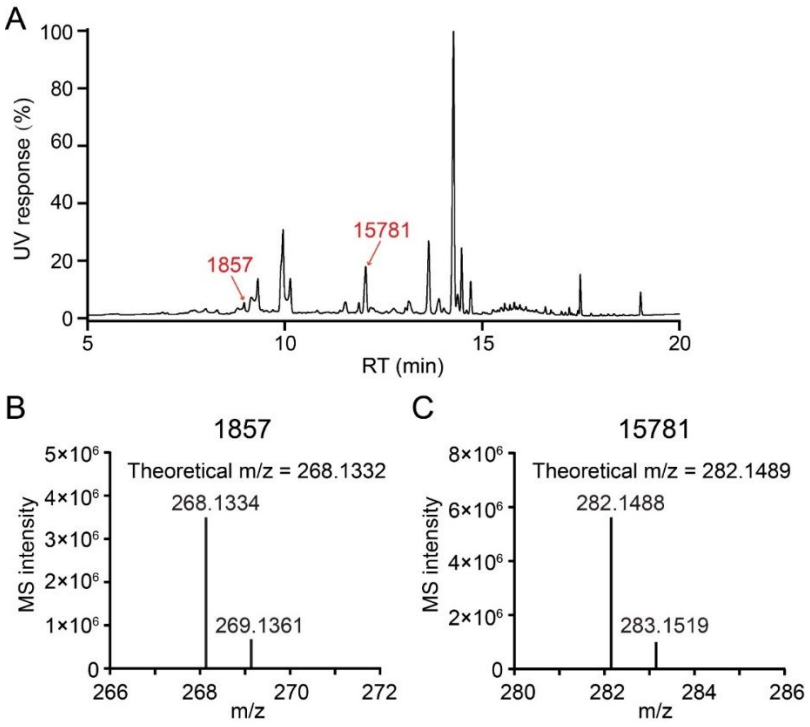
G

	AD standard concentration μM	TR standard concentration μM	TR standard concentration μM
	0	0	0
	0.16	0.63	0.31
	0.63	1.25	0.63
	1.25	2.5	1.25
			2.5
5-HT	0.63	0.03	-
5-MeO-DMT	-	-	0.24

Supplementary Figure 4. Concentrations of 5-HT and 5-MeO-DMT in the crude extracts of AD and TR determined by a standard spike-in quantification approach. MS response curves of 5-HT or 5-MeO-DMT titrated into the extracts of AD (A) or TR (B, C). Each compound was identified by HRMS measurement of the crude extract (D, E, F, left) in agreement with the standard (D, E, F, right) and retention time matching (data not shown). (G) Concentration of each compound in the original AD or TR extract was determined by extrapolating the standard curve to the origin and deriving it from the linear regression (values shown in the two rows at the bottom).



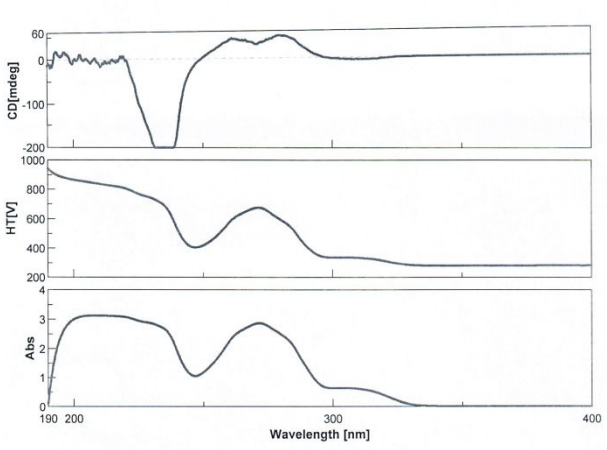
Supplementary Figure 5. Fractionation of ST crude extract and affinity MS screening of each fraction for 5-HT_{2C} ligands. (A) LC-UV chromatograms of the total extract and its three fractions from the extract (F1, F2, F3) at 254 nm. (B) Representative LC-MS chromatograms of each ST fraction. (C) Representative LC-MS chromatograms of 5-HT_{2C} target samples from the affinity MS screening of each ST fraction.



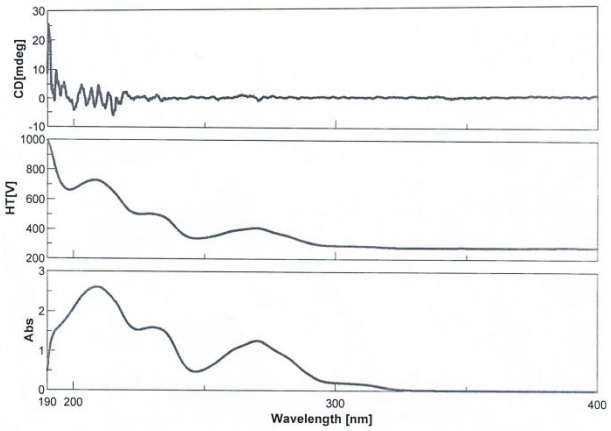
Supplementary Figure 6. LC-UV/MS guided isolation of 1857 and 15781. (A) LC-UV chromatogram of 1857 and 15781 in the ST fraction F3 at 254 nm. (B, C) HRMS spectra of 1857 and 15781 isolated from the ST extract with measured monoisotopic mass very close to the theoretical value.

Supplementary Figure 7. CD spectroscopy analysis of four aporphines.

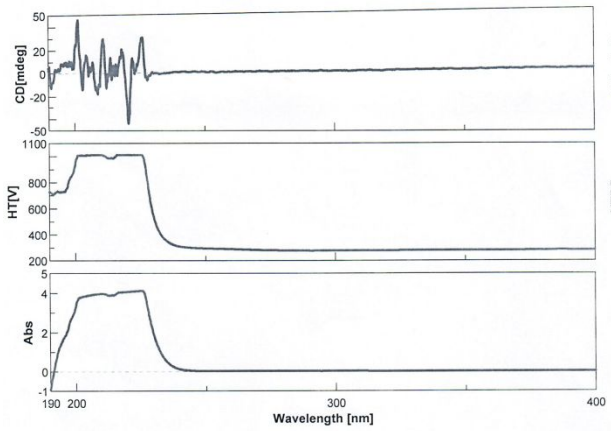
CD spectrum of 1857



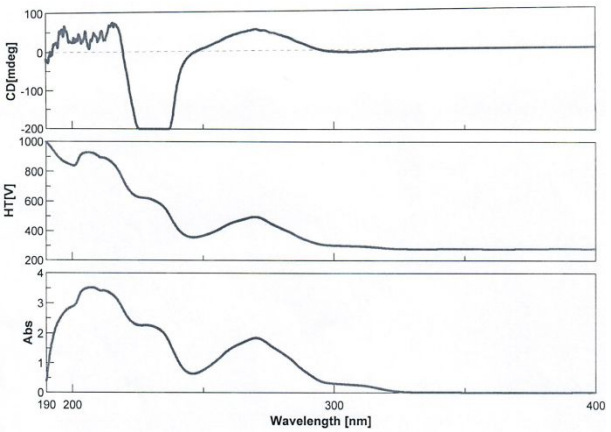
CD spectrum of 15781

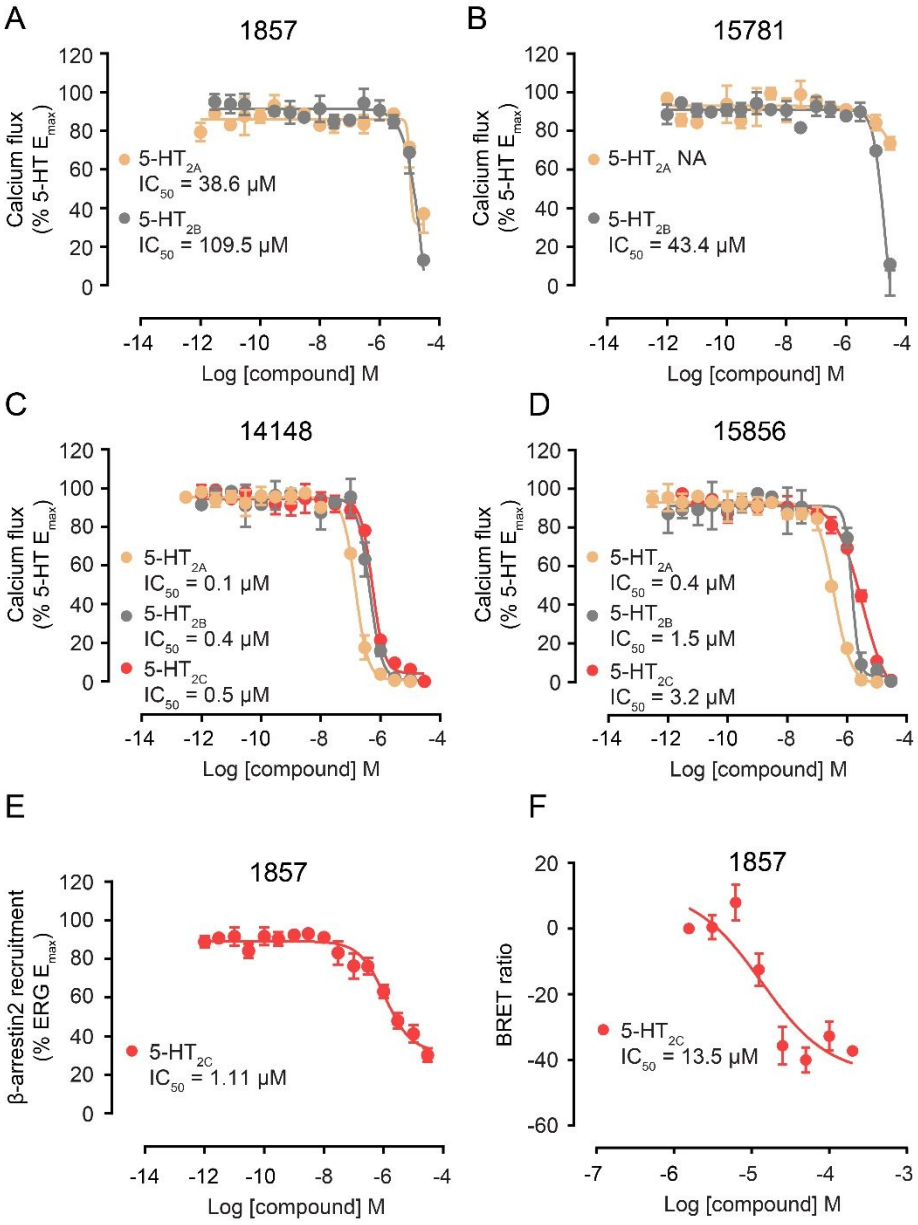


CD spectrum of 14148

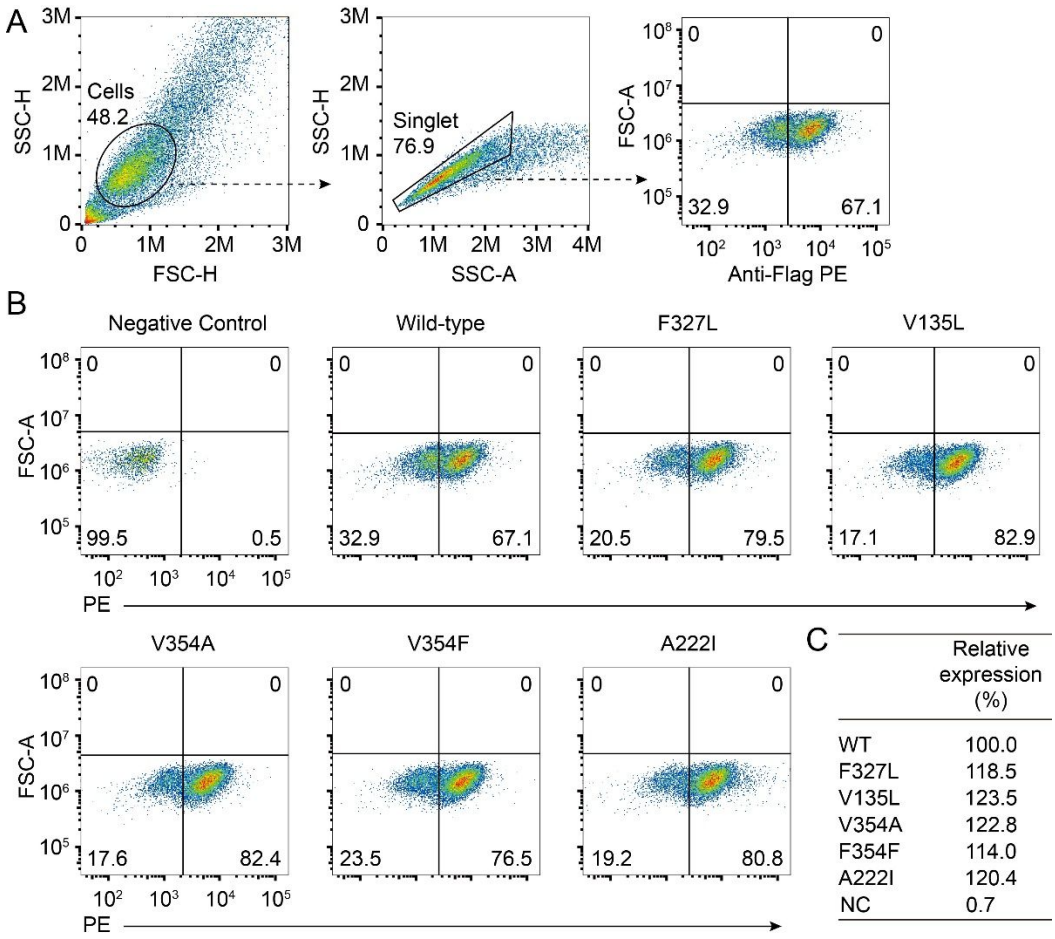


CD spectrum of 15856





Supplementary Figure 8. Functional profiles of four aporphines active at 5-HT$_2$ subfamily members.. Weak inhibition of Gq-mediated calcium mobilization in 5-HT$_{2A}$ and 5-HT$_{2B}$ expressing cells in response to 5-HT by 1857 (A) and 15781 (B). Note that 15781 did not antagonize G_q activity in 5-HT$_{2A}$ expressing cells under the same condition. Non-selective inhibition of G_q-mediated calcium mobilization in 5-HT$_{2A}$, 5-HT$_{2B}$ and 5-HT$_{2C}$ expressing cells in response to 5-HT by 14148 (C) and 15856 (D). Inhibition of β-arrestin2 recruitment in 5-HT$_{2C}$ expressing cells by 1857 as measured by Tango (E) or BRET (F) assays. Data represent means ± SEM of three independent experiments performed in triplicate. NA, no measurable activity.



Supplementary Figure 9. Cell surface expression of wild-type and mutant 5-HT_{2C} receptors. (A) Gating strategy of cell sorting. (B) FACS analysis of wild-type and mutant 5-HT_{2C} receptors. (C) Relative surface expression of wild-type and mutant 5-HT_{2C} receptors.

Supplementary Note 1.

Identification of known 5-HT$_{2C}$ agonists from herbal extracts

The established affinity MS workflow was first applied to screening 5-HT$_{2C}$ ligands from crude extracts of AD and TR that displayed the highest potency among all TCM herbs studied. The purified receptor was incubated with either extract and underwent the same affinity MS procedure as described above. Representative total ion chromatograms for the crude extract, 5-HT$_{2C}$ target and control samples are shown in Supplementary Fig. 3A. A targeted metabolomics data mining strategy previously developed by us[1] was implemented to process the affinity MS screening data for individual extracts. All LC-MS features that matched the peak characteristics of compounds registered in the TCM herb database (TCMHD)[2] were assigned to be herbal constituents. To identify these putative ligands, we determined the binding index (BI) for each of 704 and 485 assigned constituents in AD and TR extracts, respectively (Supplementary Table 1). Screening hits were selected if their mean BI values were above 2.0 ($P < 0.05$, $n = 4$)[3, 4]. In the end, three and ten initial hits were identified from AD and TR extracts, respectively (Supplementary Fig. 3B).

Interestingly, serotonin, the natural ligand for all 5-HT family members, turned out to be a top-ranking hit with high BI values in the screening (Supplementary Fig. 3B, Supplementary Table 1). A serotonin analogue 5-methoxy-*N*, *N*-dimethyltryptamine (5-MeO-DMT) known as 5-HT$_{2C}$ agonist[5] was also identified from TR screening (Supplementary Fig. 3B). MSMS spectral matching and retention time consistency between the putative ligand in 5-HT$_{2C}$ and pure standard further confirmed the structural identity of two compounds (Supplementary Fig. 3C). Of note is that neither serotonin nor

5-MeO-DMT has been reported to be present in these two herbs, suggesting that the affinity MS screen in combination with metabolomics data mining is a powerful tool to identify unknown bioactive constituents in natural products.

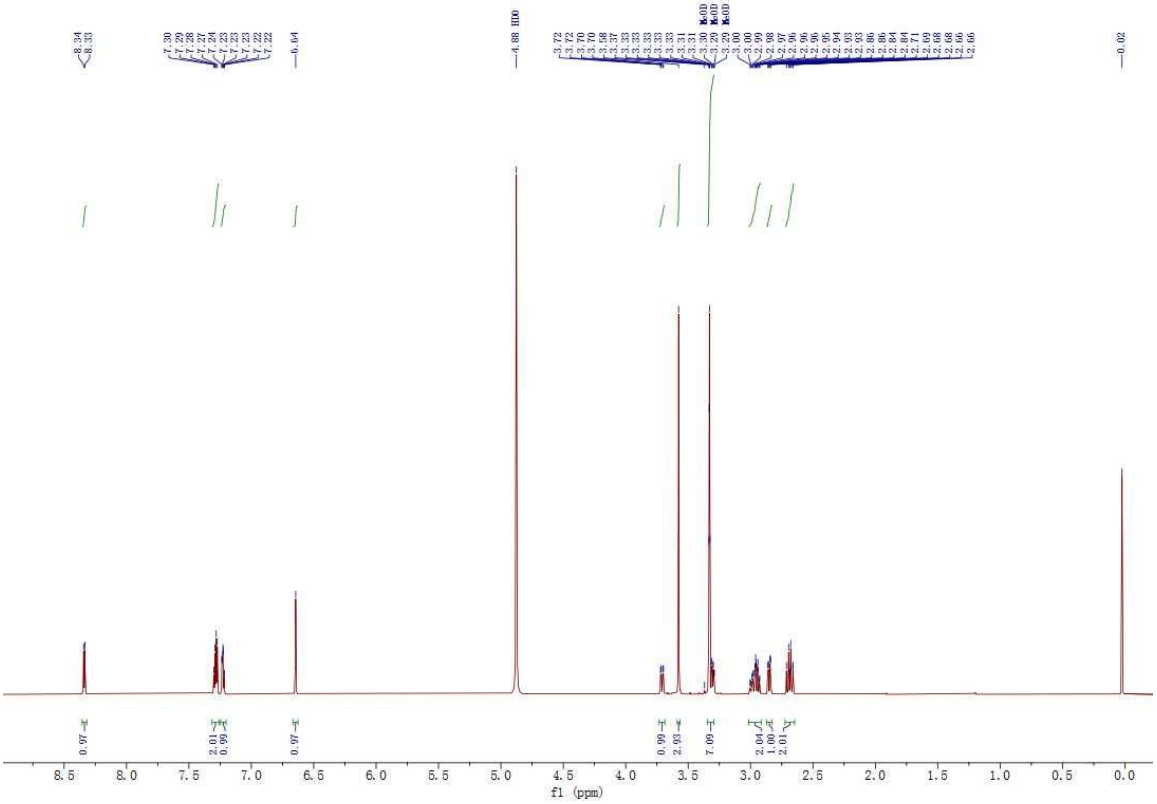
To examine how much of the total bioactivity of AD and TR is attributed to serotonin and 5-MeO-DMT, we determined the percentage of each compound in the herbal extract (Supplementary Fig. 4). According to their measured concentration in the extract and potency of the pure agonist (Supplementary Fig. 3D), we estimated to what extent serotonin or 5-MeO-DMT accounted for the total agonism of an extract (Supplementary Fig. 3E). It turned out that serotonin alone or serotonin combined with 5-MeO-DMT contributed to 100% of the overall 5-HT_{2C} activity of AD and TR (Supplementary Fig. 3E). As serotonin is present in certain plants for growth regulation and defense[6], we speculate that serotonin and its analogue are also biosynthesized in AD and TR. When analyzing additional extracts from other herbs with strong 5-HT_{2C} activities (RV and SR), serotonin was also detected at 0.02%-0.05%, close to that of TR. Thus, these two herbs were abandoned for further experimentation. Taken together, affinity MS screen conveniently identified known agonists accountable for the bioactivity detected in specific herbs, thereby prompting us to change the focus to other medicinal plants.

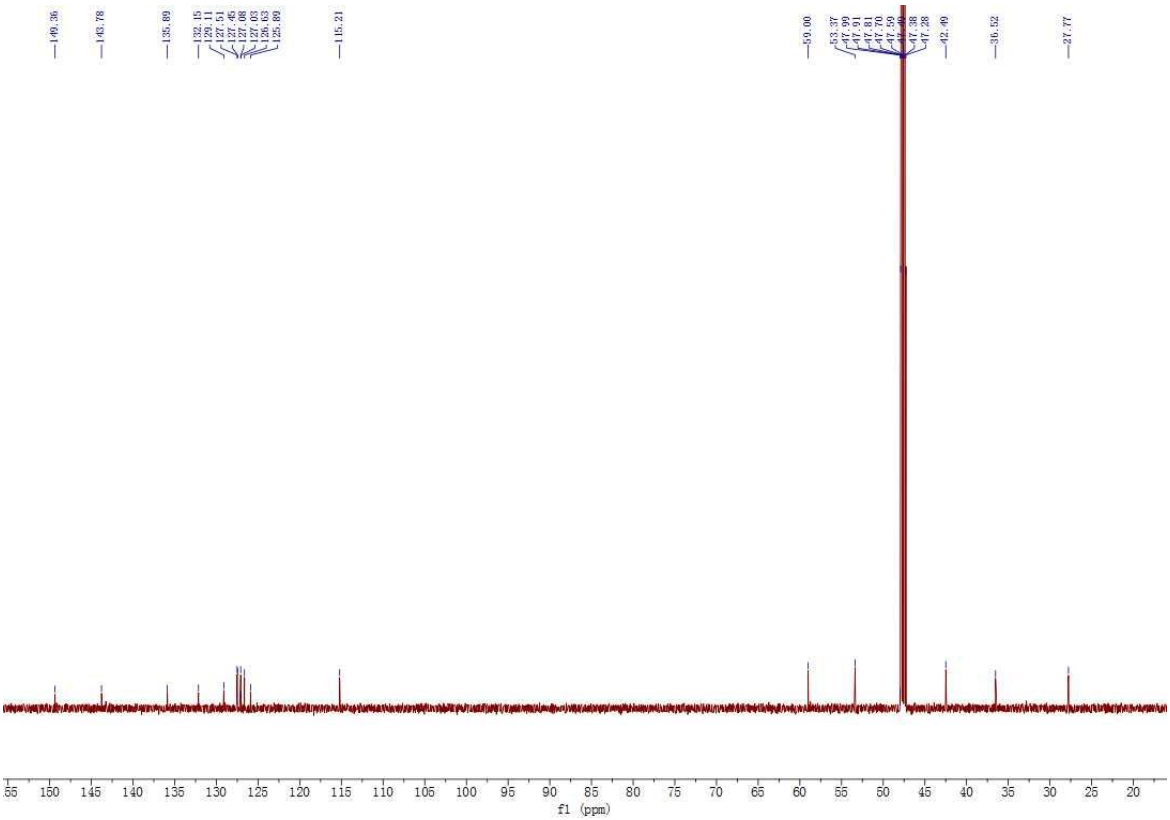
Supplementary Note 2.

(*R*)-Asimilobine (1857).^{1}H NMR (CD_3OD) δ 8.34 (d, *J* = 7.9 Hz, 1H), 7.31-7.26 (m, 2H), 7.23 (td, *J* = 7.3, 1.3 Hz, 1H), 6.64 (s, 1H), 3.71 (dd, *J* = 14.1, 4.5 Hz, 1H), 3.58 (s, 3H), 3.31-3.29 (m, 1H), 2.92-3.00 (m, 2H), 2.85 (dd, *J* = 13.8, 4.5 Hz, 1H), 2.69-2.66 (m, 2H); ^{13}C NMR (CD_3OD) δ 149.3, 143.7, 135.9, 132.2, 129.1, 127.5, 127.5, 127.1, 127.0, 126.6, 125.9, 115.2, 59.0, 53.4, 42.5, 36.5, 27.8. HRMS calcd for $C_{17}H_{17}NO_2^+$ ([M + H]$^+$): 268.1332; found, 268.1336. $[\alpha]_D^{20}$ −92.0 (c 0.1, MeOH)[7-9].

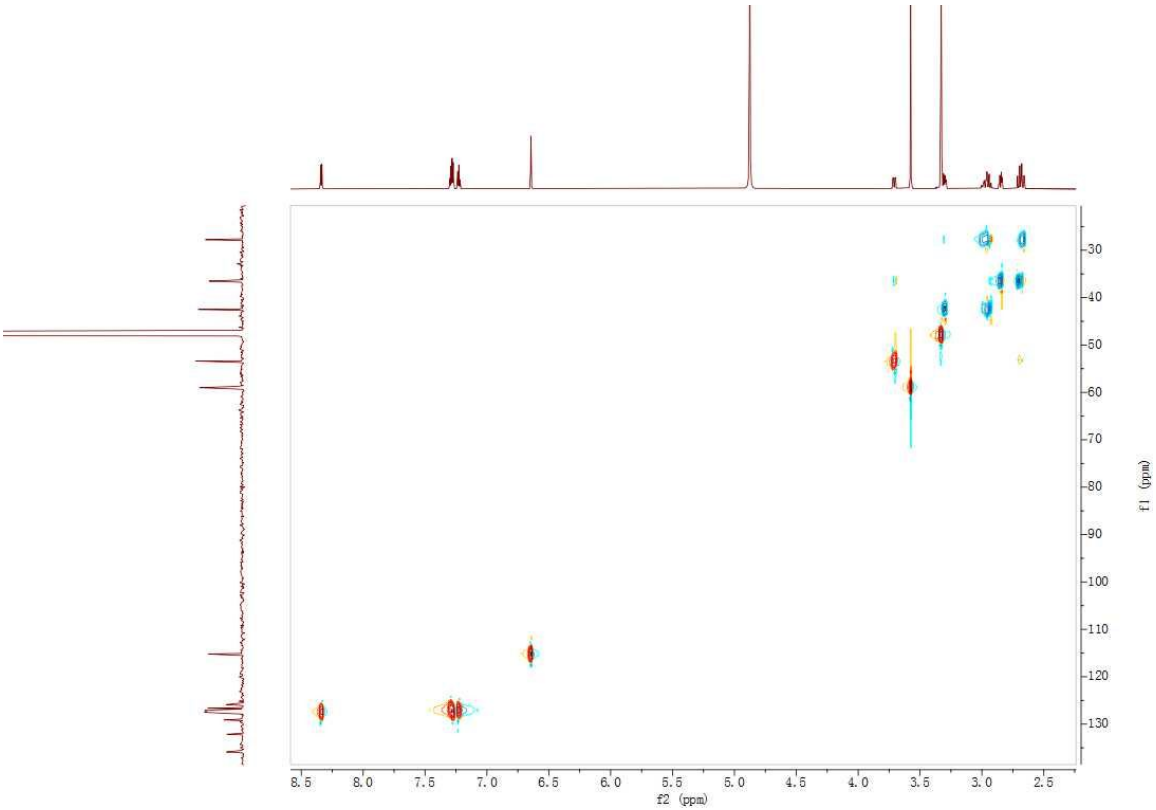
Nornuciferine (15781).^{1}H NMR (CD_3OD) δ 8.34 (d, *J* = 7.9 Hz, 1H), 7.34 (dd, *J* = 14.9, 7.3 Hz, 2H), 7.29 (t, *J* = 7.3 Hz, 1H), 6.88 (s, 1H), 4.15 (dd, *J* = 14.1, 4.1 Hz, 1H), 3.90 (s, 3H), 3.66 (s, 3H), 3.65-3.61 (m, 1H), 3.30-3.20 (m, 2H), 3.04 (dd, *J* = 13.6, 4.4 Hz, 1H), 3.00 (dd, *J* = 17.0, 4.0 Hz, 1H), 2.89 (t, *J* = 13.6 Hz, 1H); ^{13}C NMR (CD_3OD) δ 153.4, 145.7, 133.7, 131.4, 128.2, 127.8, 127.7, 127.2, 126.8, 126.3, 123.5, 111.6, 59.2, 55.0, 53.0, 41.5, 34.3, 25.7. HRMS calcd for $C_{18}H_{20}NO_2^+$ ([M + H]$^+$): 282.1489 m/z; found, 282.1491 m/z. $[\alpha]_D^{20}$ −1.5 (c 0.1, MeOH)[7].

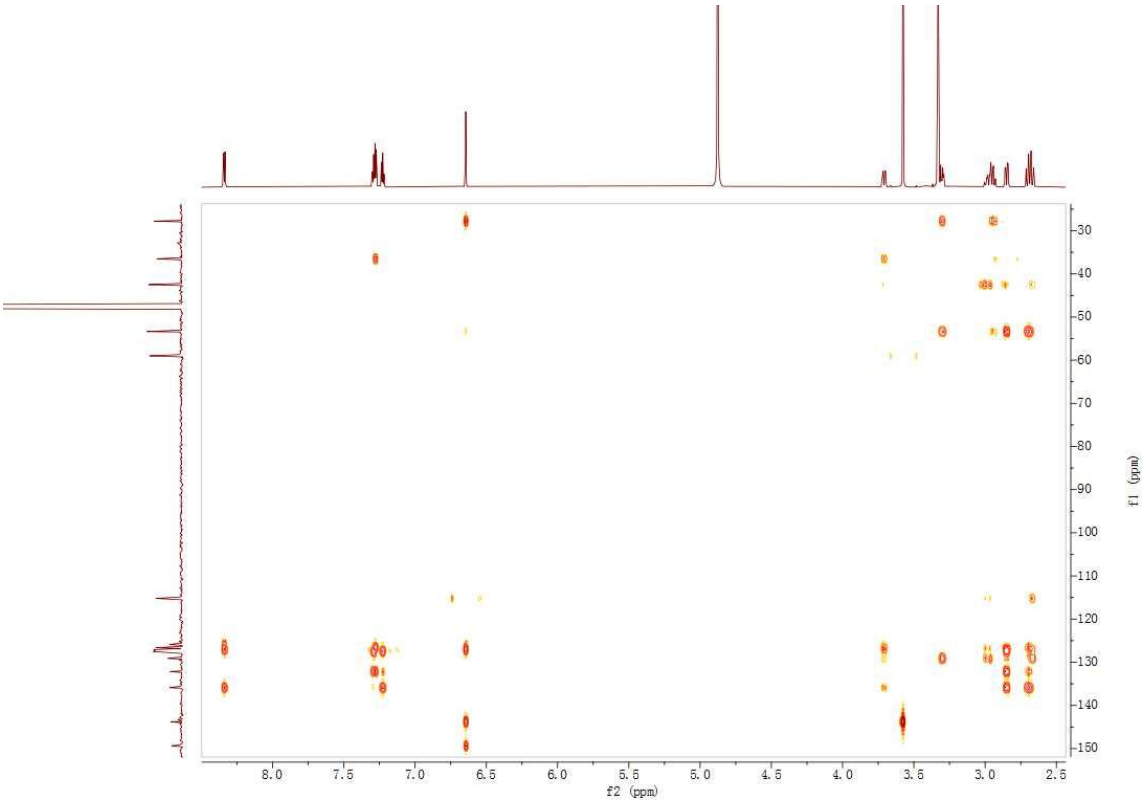
1D-NMR spectra for 1857



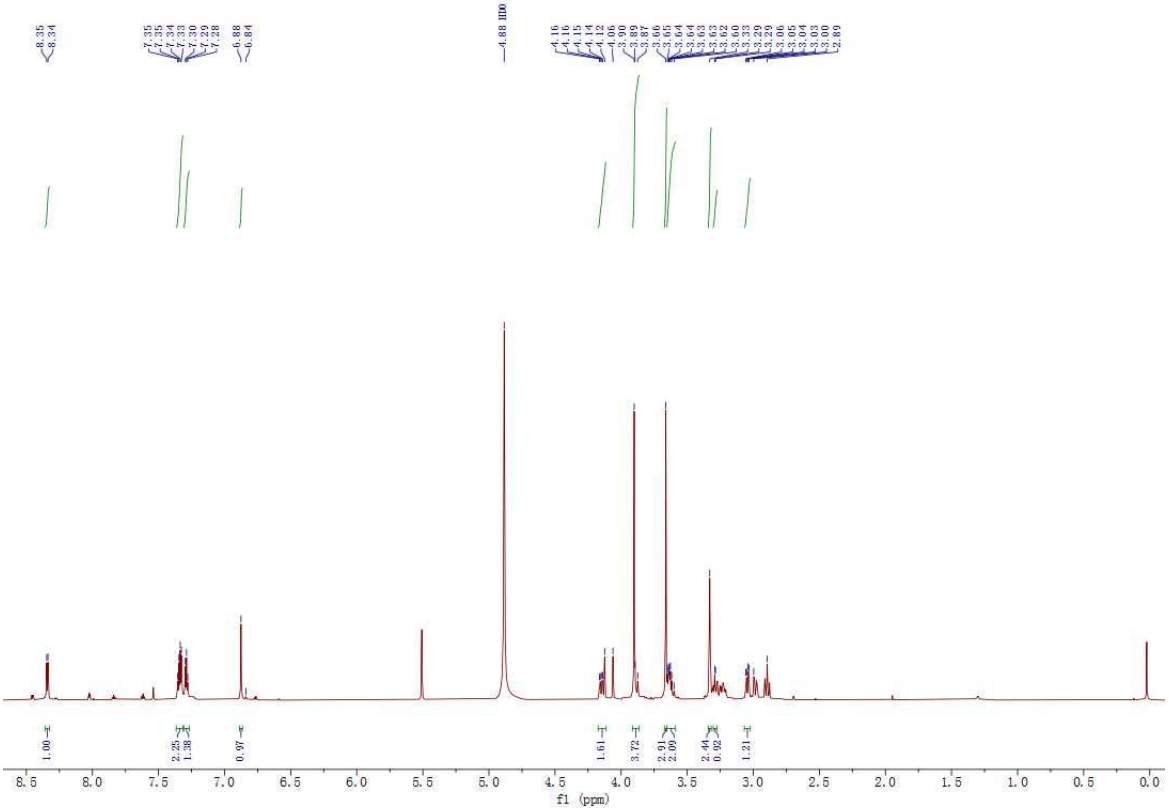


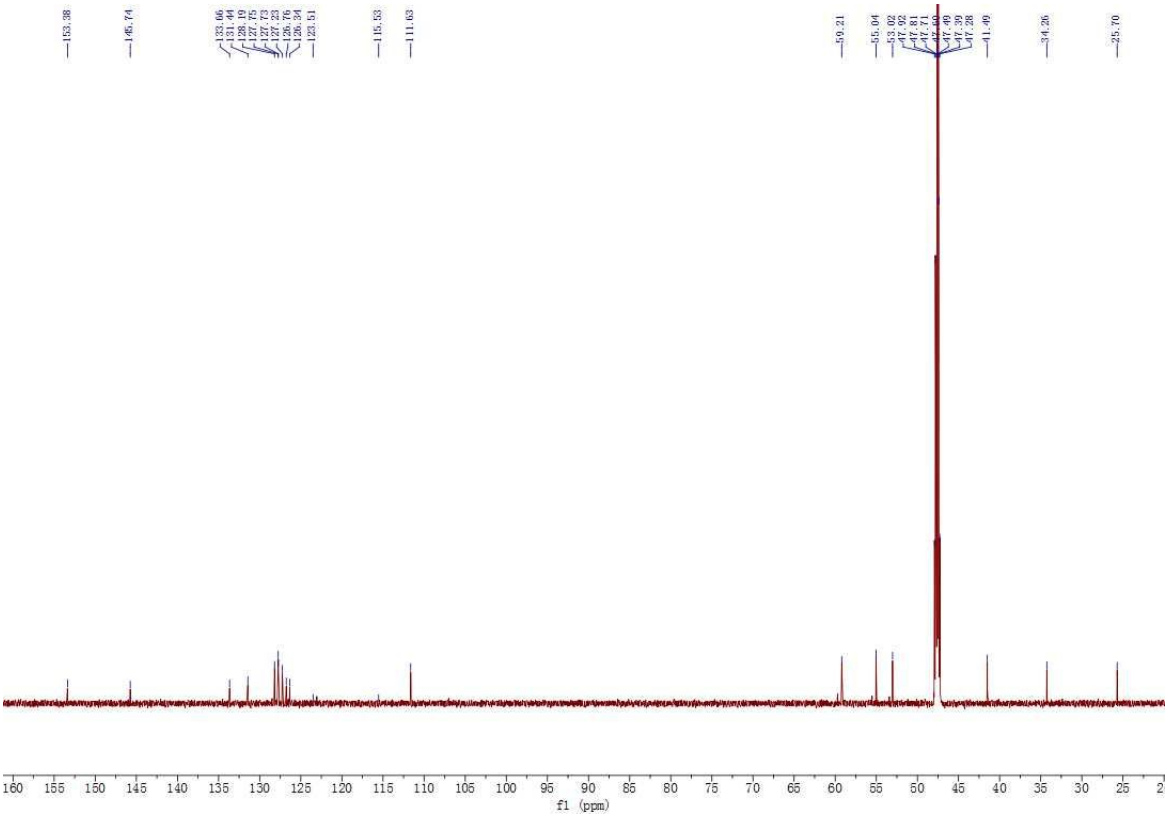
2D-NMR spectra for 1857



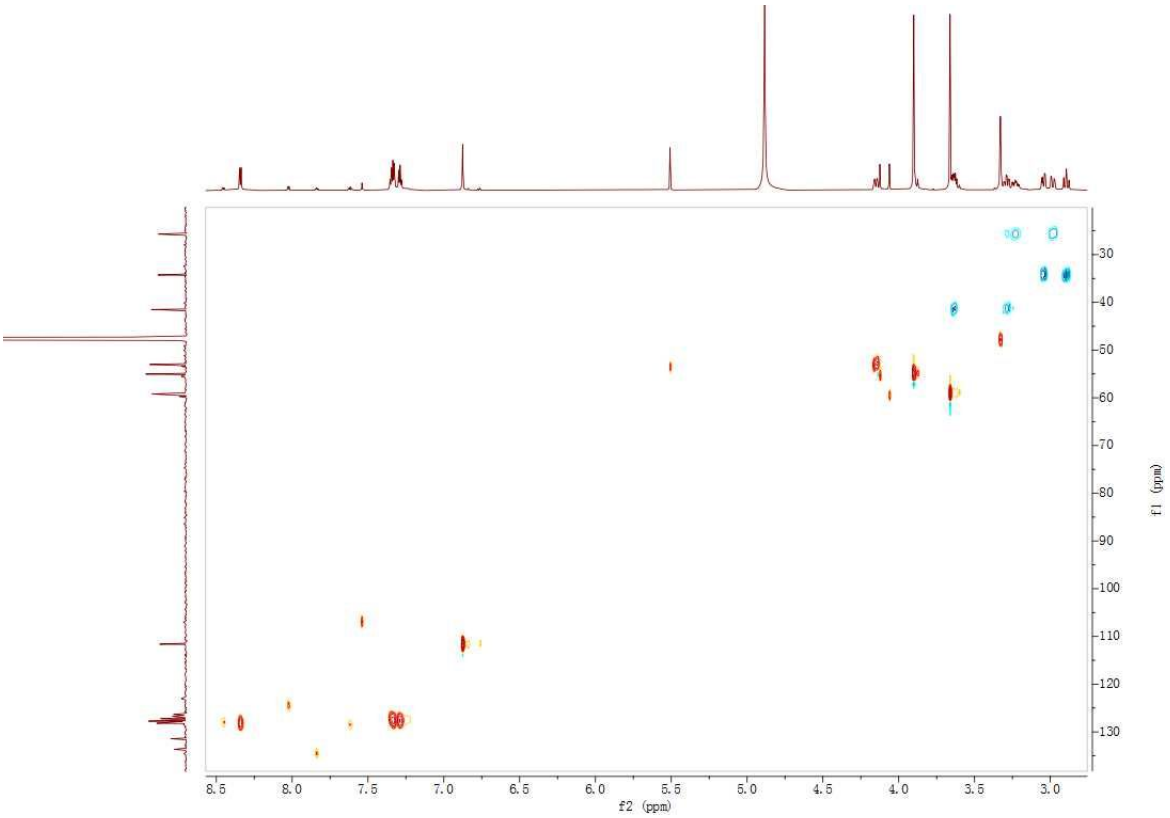


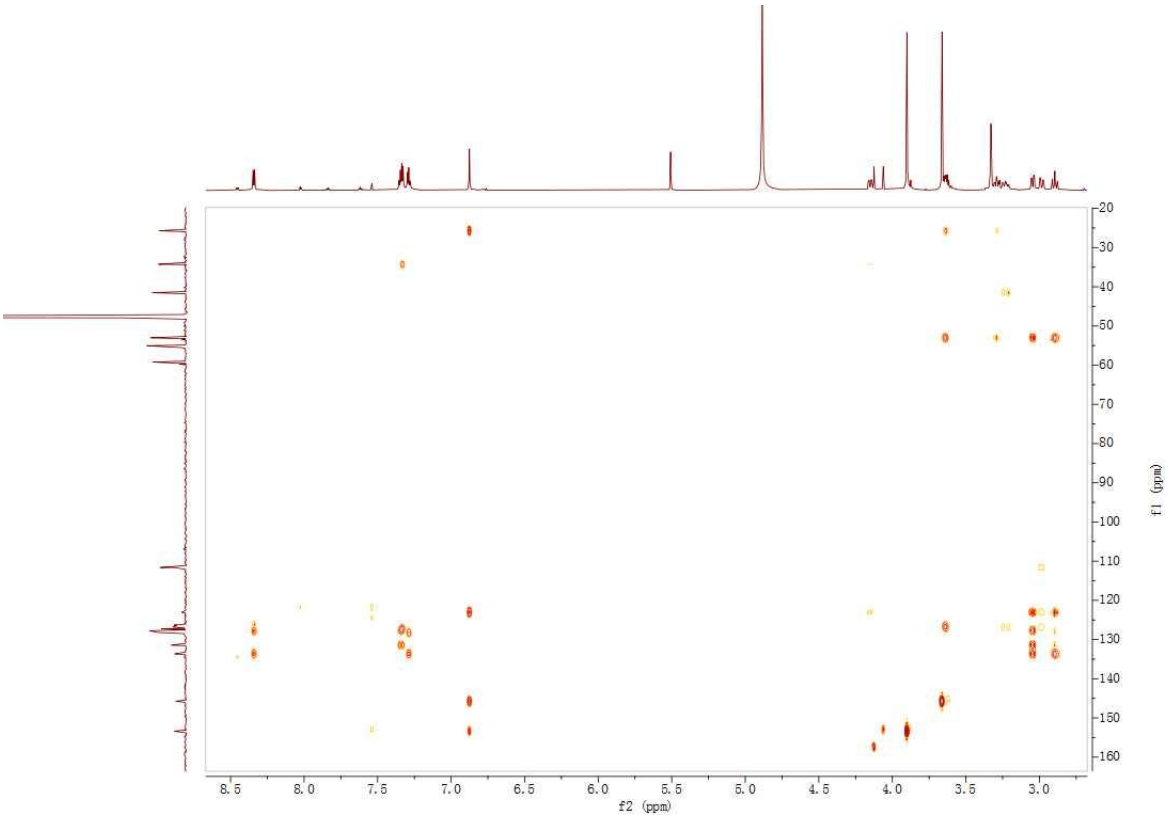
1D-NMR spectra for 15781





2D-NMR spectra for 15781





Supplementary Note 3.

The potential role of 1857 in elucidating the structural basis of preferential G protein signaling

By virtue of the rapid advances in GPCR structural biology, the molecular basis of functional selectivity of certain biased ligands has started to unfold. In particular, Stevens and Roth groups have revealed distinct structural features of 5-HT_{2B} bound to ergotamine or LSD that are both strong β–arrestin biased agonists[10, 11]. Another elegant work on the design of β-arrestin biased ligands based on aminergic GPCR structures further pointed out that the key contacts at transmembrane helix 5 (TM5) and extracellular loop 2 (ECL2) may be responsible for G protein and β-arrestin signaling, respectively[12]. Furthermore, structural characterization of 5-HT_{2B} in complex with diverse ligands has illuminated important structural determinants essential for receptor activation and biased agonism[13]. In particular, Leu362$^{7.35}$ in TM7 of 5-HT_{2B} appears to be a crucial determinant of preference for G protein or β-arrestin2 recruitment[13]. As for the novel biased agonist 1857 which acts on the same serotonin receptor system, we identified residues in TM5 (A222$^{5.46}$) and TM7 (V354$^{7.39}$) of 5-HT_{2C} that interact with the ligand and have profound effects on its biased agonism. Our finding is consistent with the previous notion that ligand engagement of TM5 and TM7 is vital to G protein biased signaling for 5-HT_{2B} and other aminergic GPCRs[12, 13]. Moreover, inability of 1857 to elicit β-arrestin2 recruitment may be attributed to its weak interaction with ECL2, a structural region posited to be critical for arrestin bias[12]. Therefore, 1857 may serve as a desirable probe for elucidating the structural basis of preferential G protein signaling and its contribution to therapeutic effects mediated by 5-HT_{2C}. Of note, our docking

analysis did not fully explain the regulatory roles of certain residues such as V354, which may suggest the presence of alternative conformations or dynamics of the receptor not seen in crystal structures.

References

1. Wang, Z. et al. Efficient ligand discovery from natural herbs by integrating virtual screening, affinity mass spectrometry and targeted metabolomics. *Analyst* **144**, 2881-2890 (2019).
2. He, M., Yan, X., Zhou, J. & Xie, G. Traditional Chinese medicine database and application on the Web. *J Chem Inf Comput Sci* **41**, 273-277 (2001).
3. Qin, S. et al. High-throughput identification of G protein-coupled receptor modulators through affinity mass spectrometry screening. *Chem Sci* **9**, 3192-3199 (2018).
4. Lu, Y. et al. Accelerating the Throughput of Affinity Mass Spectrometry-Based Ligand Screening toward a G Protein-Coupled Receptor. *Anal Chem* **91**, 8162-8169 (2019).
5. Blair, J.B. et al. Effect of ring fluorination on the pharmacology of hallucinogenic tryptamines. *J Med Chem* **43**, 4701-4710 (2000).
6. Ramakrishna, A., Giridhar, P. & Ravishankar, G.A. Phytoserotonin: a review. *Plant Signal Behav* **6**, 800-809 (2011).
7. Yang, Z.D. et al. Synthesis and structure-activity relationship of nuciferine derivatives as potential acetylcholinesterase inhibitors. *Med Chem Res* **23**, 3178-3186 (2014).
8. Rollinger, J.M. et al. Taspine: Bioactivity-guided isolation and molecular ligand-target insight of a potent acetylcholinesterase inhibitor from Magnolia x soulangiana. *J Nat Prod* **69**, 1341-1346 (2006).
9. Leboeuf, M. et al. Alkaloids of Annonaceae. XXXV. Alkaloids of Desmos Tiebaghiensis. *J Nat Prod* **45**, 617-623 (1982).
10. Wacker, D. et al. Structural features for functional selectivity at serotonin receptors. *Science* **340**, 615-619 (2013).
11. Wacker, D. et al. Crystal Structure of an LSD-Bound Human Serotonin Receptor. *Cell* **168**, 377-389 e312 (2017).
12. McCorvy, J.D. et al. Structure-inspired design of beta-arrestin-biased ligands for aminergic GPCRs. *Nat Chem Biol* **14**, 126-134 (2018).
13. McCorvy, J.D. et al. Structural determinants of 5-HT2B receptor activation and biased agonism. *Nat Struct Mol Biol* **25**, 787-796 (2018).