Electronic Supplementary Material (ESI) for MedChemComm.

Supporting Information

Discovery of indoleamine 2,3-dioxygenase inhibitors using machine learning based virtual screening

Hongao Zhang, Wei Liu, Zhihong Liu, Yingchen Ju, Mengyang Xu, Yue Zhang, Xinyu Wu, Qiong Gu*, Zhong Wang*, and Jun Xu*

Research Center for Drug Discovery, School of Pharmaceutical Sciences, Sun Yat-sen University, Guangzhou 510006, China

Table S1 Linear correlation coefficients and *p*-values of IDO inhibitory activity (pIC_{50}) and properties.

Descriptor	*R*	*p*-value
O counts	0.399	9.887×10^{-21}
N counts	-0.082	0.067
Mass	0.340	4.582×10^{-15}
HBA	0. 281	$1. 289\times10^{-10}$
HBD	0. 195	$1. 083\times10^{-5}$
pKa	0. 038	0. 410
AlogP	0.016	0.720
logD	-0.008	0.853
Solubility	-0.177	6.538×10^{-5}

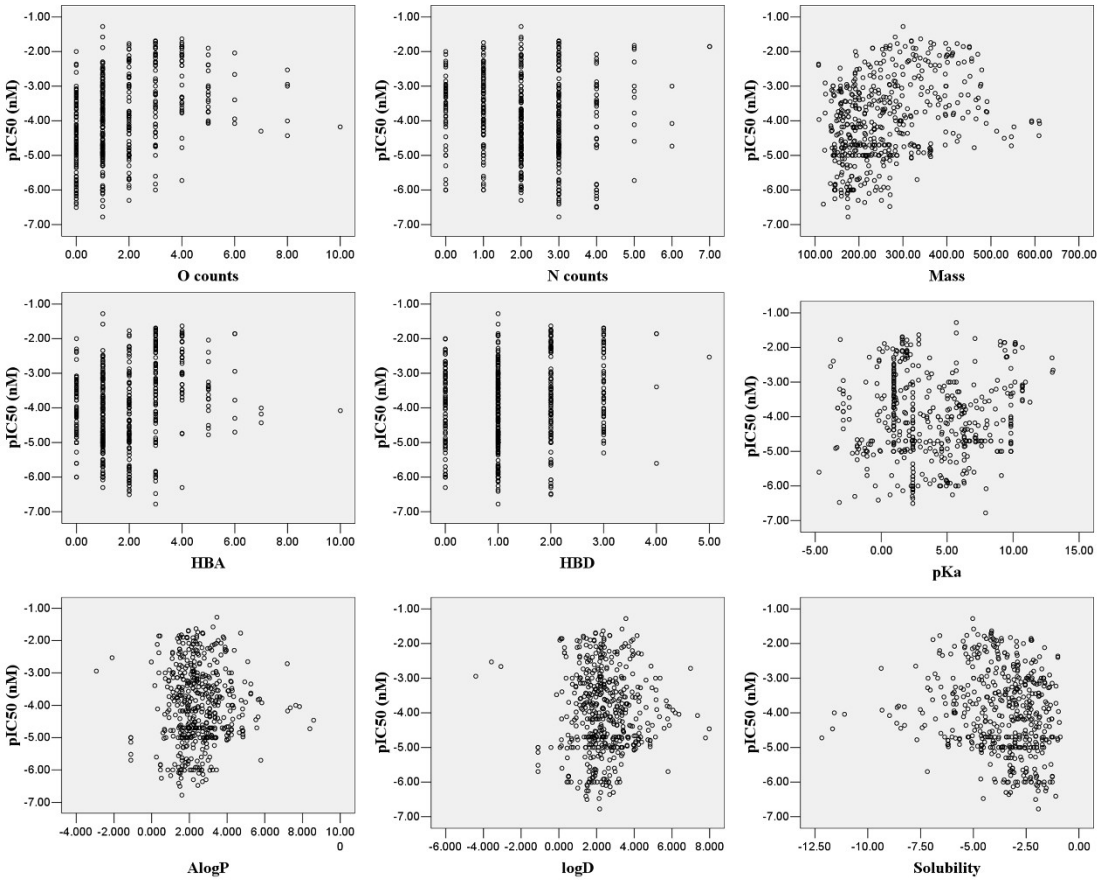
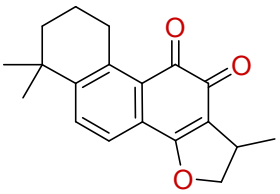
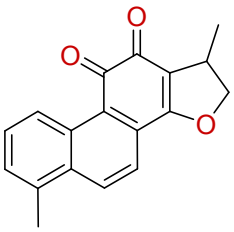
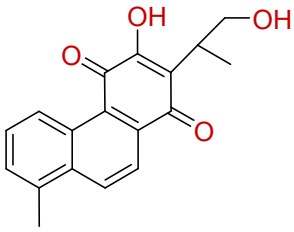
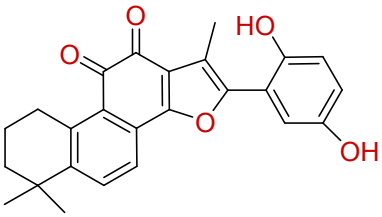
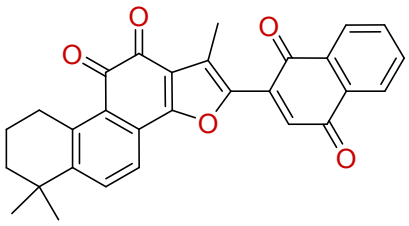
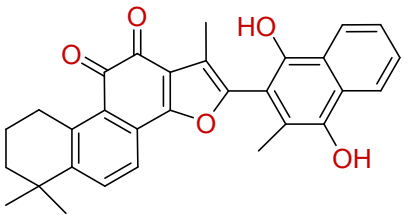
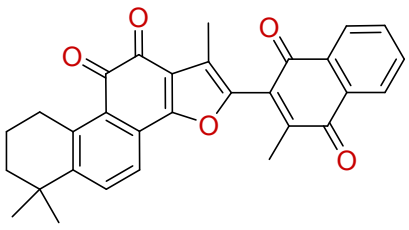
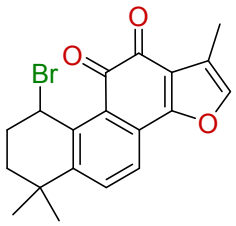
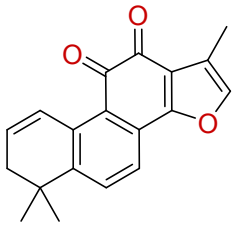
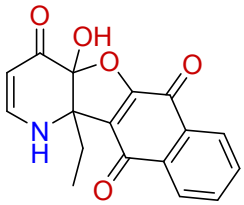
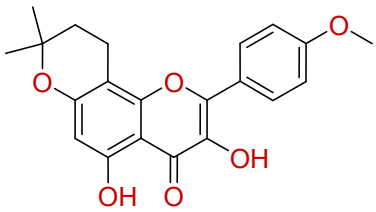
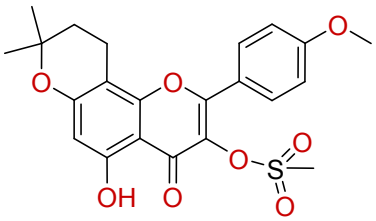
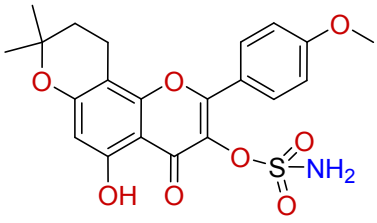
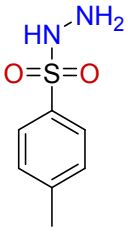
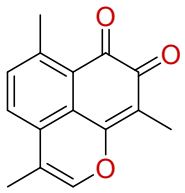
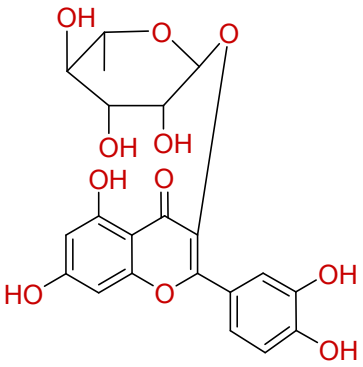
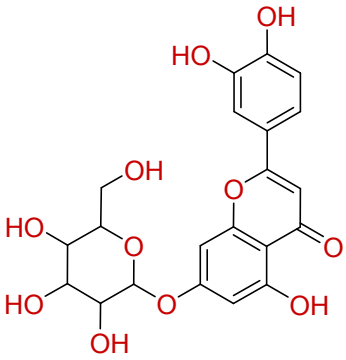
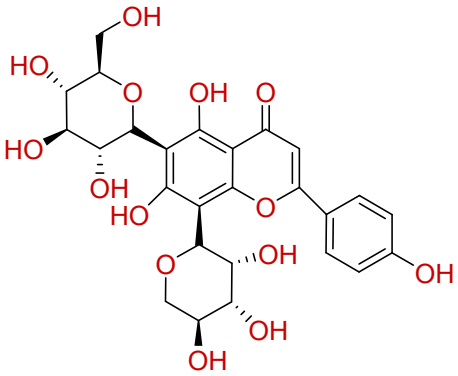
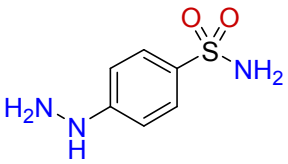


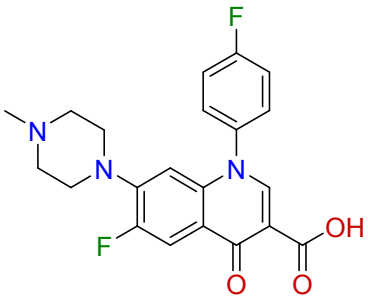
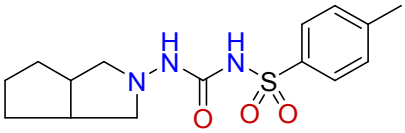
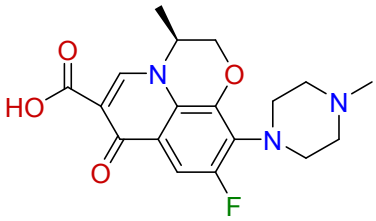
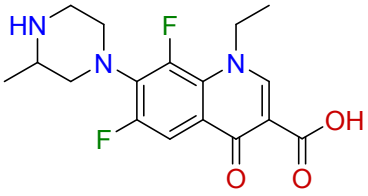
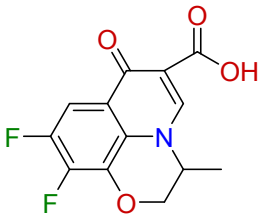
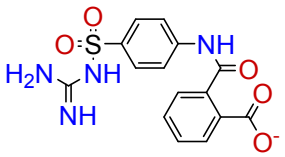
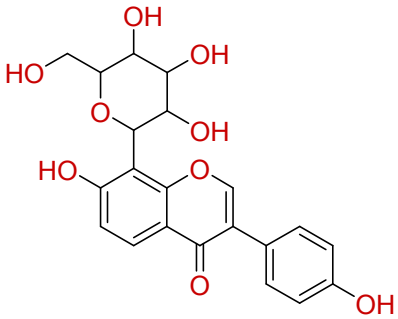
Fig. S1 Scatter plots of IDO inhibitory activity (pIC50) and properties.

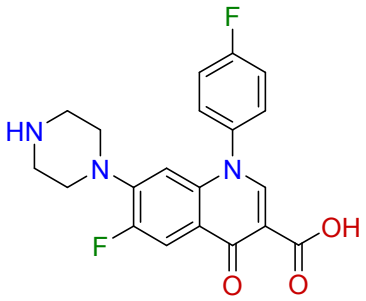
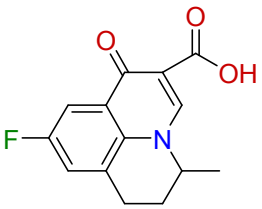
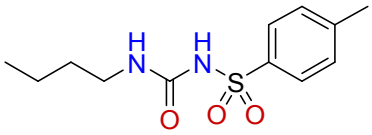
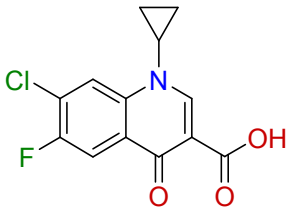
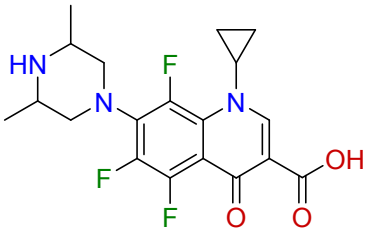
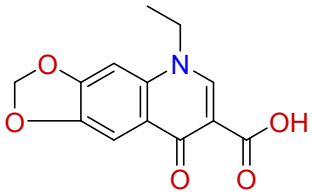
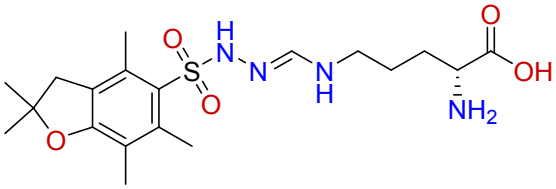
Table S2 Structure and IDO inhibitory activity of all the selected virtual hits.

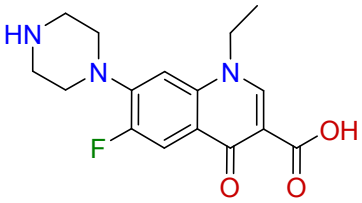
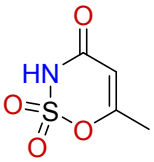
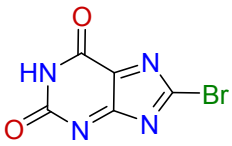
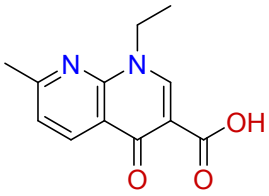
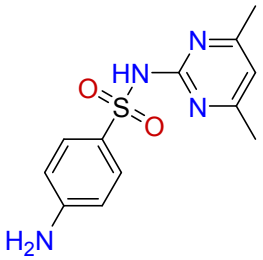
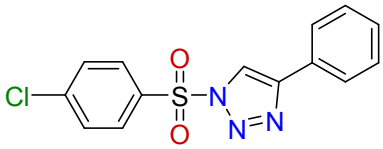
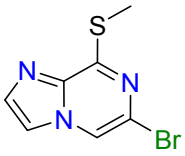
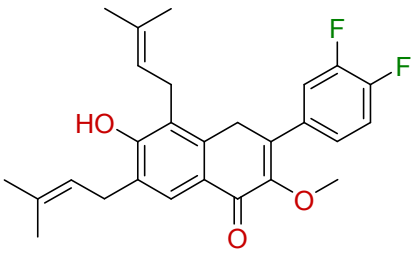
ID	Compound	Structure	IDO inhibition (%, 10 μM)
1	SYSU-00437		-15.87
2	SYSU-00440		65.56
3	SYSU-00441		16.88
4	SYSU-00460		1.34
5	SYSU-00463		1.08
6	SYSU-00464		43.86
7	SYSU-00465		35.09

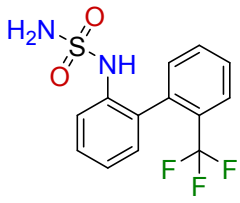
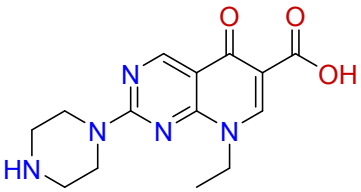
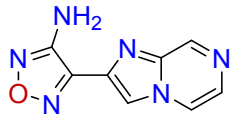
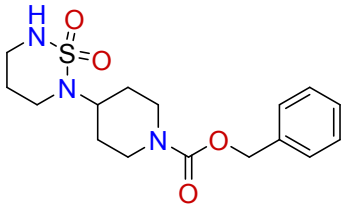
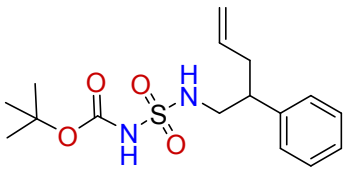
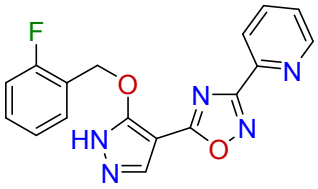
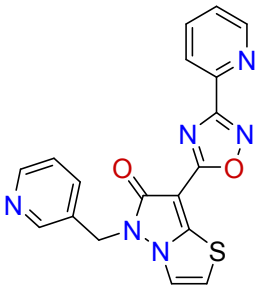
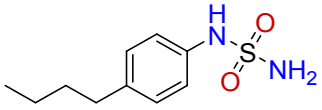
8	SYSU-00468		153.90
9	SYSU-00469		55.37
10	SYSU-00504		9.66
11	SYSU-00655		0.32
12	SYSU-00658		-4.72
13	SYSU-00659		4.62
14	SYSU-00660		18.51

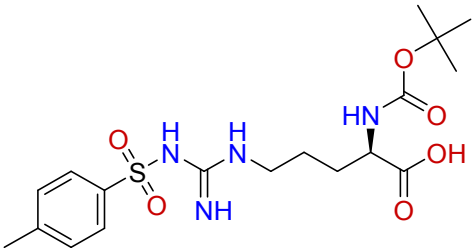
15	SYSU-00877		40.72
16	SYSU-10147		-4.80
17	SYSU-10241		-6.16
18	SYSU-10348		3.33
19	SYSU-20119		29.17

20	SYSU-20247		51.75
21	SYSU-20300		31.69
22	SYSU-20319		319.42
23	SYSU-20325		57.77
24	SYSU-20329		0.81
25	SYSU-20412		27.58
26	SYSU-20419		184.89

27	SYSU-20434		9.07
28	SYSU-20476		-22.14
29	SYSU-20538		34.73
30	SYSU-20562		6.25
31	SYSU-20591		0.67
32	SYSU-20594		23.08
33	SYSU-21653		14.79

34	SYSU-21711		65.93
35	SYSU-22061		-22.49
36	SYSU-22210		26.38
37	SYSU-22645		-15.54
38	SYSU-22878		20.02
39	SYSU-23369		22.75
40	SYSU-23661		9.20
41	SYSU-23696		11.99

42	SYSU-24222		9.02
43	SYSU-24421		6.02
44	SYSU-24766		26.76
45	SYSU-24892		-2.31
46	SYSU-25109		-18.80
47	SYSU-25113		9.27
48	SYSU-25123		3.64
49	SYSU-25145		-20.61

50	SYSU-25585		-15.30

Molecular Dynamics (MD) Simulations methods. All the proteins and ligands in complexes were then modeled using the ff14SB force field and general AMBER force field (gaff2) respectively. Missing protein hydrogen atoms were added using the program LEaP embedded in AMBER16. The partial atomic charges and missing force-field parameters of these ligands were obtained from the restrained electrostatic potential (RESP) charge based on HF/6-31G* calculation with the Gaussian09 package and antechamber suite. The system was then solvated in a TIP3P water box with a minimum distance of 10 Å between any protein atom and the edge of the box. The system was neutralized by adding of Cl^- counterions. The amino acid side chain protonation states were assigned using default AMBER protonation states. The molecular dynamics simulations were carried out using the AMBER16 molecular simulation package. First, 4,000 cycles of minimization (2,000 cycles of steepest descent and 2,000 cycles of conjugate gradient) were carried out to relax the solvent, while all protein and ligands atoms were constrained by a potential of 3,000 kcal/mol/Å^2. Second, another minimization stage was conducted with the protein backbone and heavy atoms of the substrates constrained (500 kcal/mol/Å^2). Each complex was then submitted to 4,000 cycles of energy minimization without any constraint (2,000 cycles of steepest descent and 2,000 cycles of conjugate gradient). Subsequently, each system was gradually heated from 0 to 300 K over a period of 50 ps, followed by another 100 ps of NPT MD simulations at 300 K. Afterward, 20 ns of NVT MD simulation with a target temperature of 300 K were performed for each system to produce trajectories. The SHAKE algorithm was employed to constrain the bonds involving hydrogen atoms, and a time step of 2.0 fs was used for all simulations.

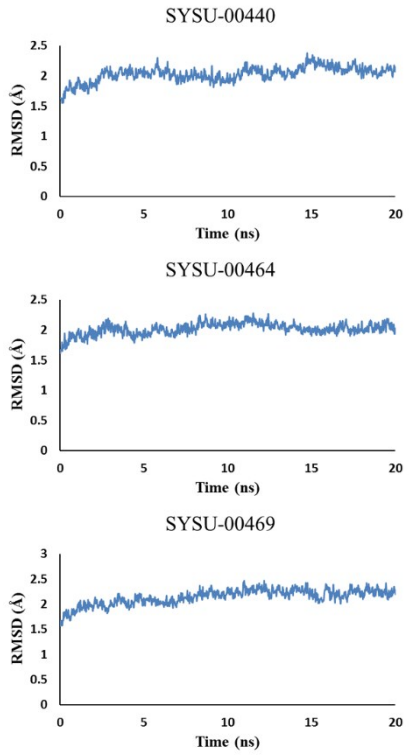


Fig. S2 RMSD curve of all system in 20 ns.

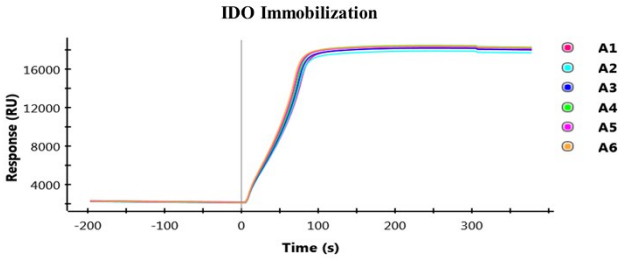


Fig. S3 Sensorgrams for immobilization of IDO.

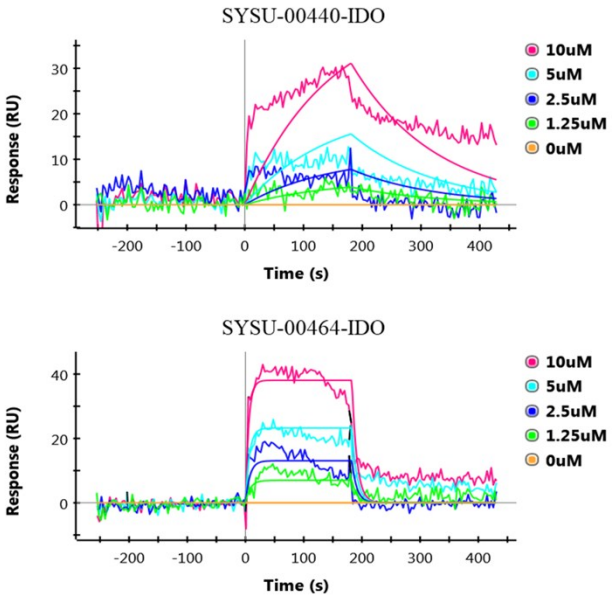


Fig. S4 Sensorgrams for the interactions of SYSU-00440 and SYSU-00464 with the IDO.