Supplementary Material

Structure-based virtual screening discovers novel PKMYT1 inhibitors

Haoyu Zhang[a,b,c], Jinyu Yu[a,b,c], Ziheng Yang[a,b,c], Zhiqiang Guo[a,b,c], Rui Liu [a,b,c], Qiaohua Qin[a,b,c], Yixiang Sun[a,b,c], Nian Liu[a,b,c], Zixuan Gao[a,b,c], Dongmei, Zhao [a,b,c]*, Maosheng Cheng [a,b,c]

[a] *Key Laboratory of Structure-Based Drug Design & Discovery of Ministry of Education, Shenyang Pharmaceutical University, Shenyang 110016, China*

[b] *School of Pharmaceutical Engineering, Shenyang Pharmaceutical University, Shenyang 110016, China*

[c] *Key Laboratory of Intelligent Drug Design and New Drug Discovery of Liaoning Province, Shenyang Pharmaceutical University, Shenyang 110016, China*

*Corresponding author: Prof. Dongmei Zhao, medchemzhao@163.com.

The final virtual screening results, including docking scores and MMGBSA calculations.

Table S1. The results of hit compounds via Virtual screening.

Name	CAS	docking score	MMGBSA dG Bind (kcal /mol)
(-)-Gallocatechin gallate	4233-96-9	-10.933	-81.54
Secoisolariciresinol diglucoside	148244-82-0	-14.646	-77.5
Poliumoside	94079-81-9	-14.033	-76.84
Cistanoside A	93236-42-1	-11.208	-76.15
Pyrithioxin	1098-97-1	-10.339	-75.4
Orientin	28608-75-5	-13.675	-74.65
Oroxin B	114482-86-9	-15.768	-73.61
Tubuloside A	112516-05-9	-15.34	-72.33
Epimedin B	110623-73-9	-12.18	-69.44
Vitexin -4"-O-glucoside	178468-00-3	-14.501	-68.86
Mulberroside A	102841-42-9	-12.418	-68.26
(-)-Epigallocatechin Gallate	989-51-5	-13.446	-68.16
Emodin-8-glucoside	23313-21-5	-10.501	-67.82
Mirificin	103654-50-8	-11.313	-67.68
Baohuoside I	113558-15-9	-11.367	-66.82
2"-O-Galloylhyperin	53209-27-1	-14.565	-66.5
1,4-Dicaffeoylquinic acid	1182-34-9	-10.823	-66.17
Isochlorogenic acid B	14534-61-3	-13.883	-66
Rosavin	84954-92-7	-12.279	-65.96
Vitexin	3681-93-4	-12.348	-65.36
Riboflavin	83-88-5	-10.714	-65.26
Silymarin	65666-07-1	-12.088	-64.66
Proanthocyanidins	4852-22-6	-12.248	-64.23
Pectolinarin	28978-02-1	-13.794	-63.13
Plantamajoside	104777-68-6	-13.115	-62.34
D-Stachyose tetrahydrate	10094-58-3	-16.923	-61.98
polygalaxanthone III	162857-78-5	-10.018	-61.8
Icaritin	118525-40-9	-10.034	-61.54
Aurantio-obtusin	67979-25-3	-9.447	-61.52
glucosylvitexin	76135-82-5	-14.307	-61.4
Pueraria glycoside	117060-54-5	-10.015	-61.2
(-)-Epicatechin gallate	1257-08-5	-12.267	-61.05
Dihydromyricetin	27200-12-0	-11.293	-60.79
Diosmin	520-27-4	-12.913	-60.75
Diosmetin-7-O-beta-D-glucopyranoside	20126-59-4	-10.631	-60.36
Flavin Adenine Dinucleotide Disodium	84366-81-4	-13.307	-59.82
Gefitinib	184475-35-2	-9.305	-59.55
Dronedarone	141626-36-0	-9.978	-59.53
Rhaponiticin	155-58-8	-10.046	-59.04
Eriocitrin	13463-28-0	-11.624	-58.85
Rutin	153-18-4	-12.265	-58.79
Doxorubicin hydrochloride	25316-40-9	-10.314	-58.3
Amygdalin	29883-15-6	-10.414	-57.98
Isorhamnetin 3-O-neohesperidin	55033-90-4	-11.074	-57.5

Name	CAS	docking score	MMGBSA dG Bind (kcal /mol)
Wogonoside	51059-44-0	-10.656	-57.16
Icariin	489-32-7	-10.245	-57.14
Harpagoside	19210-12-9	-9.374	-57.03
Hesperidin	520-26-3	-13.547	-56.9
Trans-Zeatin Riboside	6025-53-2	-10.856	-56.79
Paromomycin Sulfate	1263-89-4	-9.417	-56.68
Piceid	65914-17-2	-9.323	-56.4
Quercitrin	522-12-3	-11.341	-56.39
Naringin dihydrochalcone	18916-17-1	-10.897	-56.29
9'''-MethyllithosperMate B	1167424-32-9	-13.276	-56.16
Curculigoside	85643-19-2	-9.96	-56.13
Neohesperidin	13241-33-3	-12.013	-56.09
Rhoifolin	17306-46-6	-11.431	-55.97
Naringin	10236-47-2	-10.38	-55.43
Rhodiosin	86831-54-1	-11.953	-55.2
Tetrahydro Curcumin	36062-04-1	-9.545	-55.09
Asperuloside	14259-45-1	-9.374	-55.06
Narirutin	14259-46-2	-14.71	-54.9
Oroxin A	57396-78-8	-13.848	-54.62
Plantagoside	78708-33-5	-14.235	-54.44
tetrahydroxyl diphenylethylene-2-o-gluco	55327-45-2	-10.86	-54.4
Methyl hesperidin	11013-97-1	-13.021	-54.3
Genistin	529-59-9	-10.454	-53.79
Spinosin	72063-39-9	-11.946	-53.78
Isoorientin	4261-42-1	-10.622	-53.41
Isochlorogenic acid C	32451-88-0	-12.226	-52.99
Astragalus polyphenols	82373-94-2	-9.901	-52.99
Adenosine	58-61-7	-9.974	-52.55
Neodiosmin	38665-01-9	-11.179	-52.53
Hyperoside	482-36-0	-12.923	-52.4
4,5-Dicaffeoylquinic acid	57378-72-0	-12.312	-51.82
kaempferide	491-54-3	-9.848	-51.59
Verbascoside	61276-17-3	-13.556	-51.43
Luteolin	491-70-3	-9.943	-51.13
Salvianolic Acid C	115841-09-3	-12.12	-50.92
Orientin	28608-75-5	-10.713	-50.79
ISOQUERCITRIN	21637-25-2	-11.626	-50.77
Manninotriose	13382-86-0	-11.863	-50.6
Adenosine 5'-monophosphate	61-19-8	-10.777	-50.32
Thiocolchicoside	602-41-5	-9.769	-50.09
Albiflorin	39011-90-0	-9.306	-49.86
Vitexin-2"-O-rhamnoside	64820-99-1	-11.332	-49.53
Picroside I	27409-30-9	-9.428	-49.49
Neomangiferin	64809-67-2	-12.991	-49.46
Anhydroicaritin	38226-86-7	-9.861	-49.35
Forsythoside A	79916-77-1	-13.512	-49.34
Tenuifoliside A	139726-35-5	-11.137	-49.06
2"-O-Galloylhyperin	53209-27-1	-10.549	-49.06
Isochlorogenic acid A	2450-53-5	-9.872	-48.81

Name	CAS	docking score	MMGBSA dG Bind (kcal /mol)
Esculin	531-75-9	-9.483	-48.47
Neochlorogenic acid	906-33-2	-9.289	-48.34
Didymin	14259-47-3	-12.634	-48.25
N6-methyladenosine	1867-73-8	-10.49	-48.18
Thymidine	50-89-5	-9.384	-47.46
8-O-Acetylharpagide	6926-14-3	-10.824	-47.24
Quercetin Dihydrate	6151-25-3	-10.487	-47.22
Diosmetin	520-34-3	-9.302	-47.05
Phlorizin	60-81-1	-11.26	-46.91
Troxerutin	7085-55-4	-12.304	-46.53
Luteolin-7-glucuronide	29741-10-4	-12.331	-46.44
Agarotetrol	69809-22-9	-9.401	-46.37
Casanthranol	8024-48-4	-9.634	-46.09
Neohesperidin Dihydrochalcone	20702-77-6	-9.899	-45.81
Isorhamnetin	480-19-3	-9.598	-45.59
Chlorogenic Acid	327-97-9	-10.189	-45.47
Raffinose	17629-30-0	-12.924	-45.18
Vitexin	3681-93-4	-10	-45.08
3,6'-Disinapoyl sucrose	139891-98-8	-12.344	-44.77
Kaempferitrin	482-38-2	-10.661	-44.39
Echinacoside	82854-37-3	-13.331	-43.96
Phlorizin dihydrate	7061-54-3	-11.106	-43.72
(-)-3,5-Dicaffeoyl quinic acid	89919-62-0	-9.739	-43.64
Kaempferol	520-18-3	-9.772	-43.36
Maltopentaose	34620-76-3	-14.443	-43.3
Picroside II	39012-20-9	-9.593	-42.91
Thymidine 5'-monophosphate disodium salt	33430-62-5	-9.502	-42.08
Vaccarin	53452-16-7	-10.192	-42.06
2'-Deoxyadenosine 5'-monophosphate	653-63-4	-9.619	-41.87
Neomangiferin	64809-67-2	-10.683	-41.82
D-Cellobiose	528-50-7	-9.454	-41.67
Morin	480-16-0	-9.563	-41.63
Picroside III	64461-95-6	-11.174	-41.34
Baicalein	491-67-8	-10.325	-40.84
Arctiin	20362-31-6	-9.973	-40.83
Breviscapin	116122-36-2	-11.682	-40.79
(-)-Epigallocatechin Gallate	989-51-5	-9.661	-40.76
Scutellarein	529-53-3	-9.646	-39.86
Guanosine	118-00-3	-9.648	-39.74
Tubercidin	69-33-0	-9.258	-39.73
polygalaxanthone III	162857-78-5	-9.409	-39.66
Forsythin	487-41-2	-11	-39.47
Nystose	13133-07-8	-14.003	-39.24
Sophoricoside	152-95-4	-9.267	-38.95
Chicoric acid	70831-56-0	-9.775	-38.09
D-Pantethine	16816-67-4	-9.687	-36.86
2"-O-Galloylhyperin	53209-27-1	-10.896	-36.25
(-)-Gallocatechin gallate	4233-96-9	-9.791	-35.28
Uridine 5'-monophosphate	58-97-9	-9.738	-35.13

Name	CAS	docking score	MMGBSA dG Bind (kcal /mol)
(-)-Epicatechin gallate	1257-08-5	-9.343	-35.01
Scutellarin	27740-01-8	-11.413	-34.24
Aloin	1415-73-2	-11.083	-33.82
5'-Cytidylic acid	63-37-6	-9.318	-33.75
9'-Methyl lithospermate B	1167424-31-8	-9.456	-33.49
Stevioside	57817-89-7	-11.891	-33
Plantagoside	78708-33-5	-11.263	-31.23
2''-O-Galloylhyperin	53209-27-1	-10.183	-31.07
Salvianolic acid B	115939-25-8	-12.545	-29.55
Parishin B	174972-79-3	-11.624	-29.5
Calceolarioside B	105471-98-5	-11.6	-28.83
Loganic acid	22255-40-9	-9.559	-27.23
Plantagoside	78708-33-5	-9.324	-26.76
Hyperoside	482-36-0	-9.477	-26.45
5'-UMP disodium salt	3387-36-8	-9.607	-26.43
Cyclic AMP	60-92-4	-9.656	-25.84
Secoxyloganin	58822-47-2	-9.898	-25.82
Naringin dihydrochalcone	18916-17-1	-10.009	-23.86
polygalaxanthone III	162857-78-5	-10.56	-23.14
Phlorizin	60-81-1	-9.318	-20.75
Adenosine 5'-monophosphate	61-19-8	-10.048	-19.8
UDP disodium salt	27821-45-0	-9.701	-19.58
Bilirubin	635-65-4	-10.103	-17.18
Thymidine 5'-monophosphate disodium salt	33430-62-5	-9.868	-16.35
Phlorizin dihydrate	7061-54-3	-10.703	-15.88
Monotropein	5945-50-6	-10.952	-13.6
Neohesperidin Dihydrochalcone	20702-77-6	-10.611	-12.36
Uridine 5'-monophosphate	58-97-9	-9.844	-11.48
2'-Deoxyuridine 5'-monophosphate disodium salt	42155-08-8	-9.706	-11.14
UDP disodium salt	27821-45-0	-9.522	7.78

Decoy data set for virtual screening model validation

Table S2. The smiles structure of Positive compounds.

Compd[a]	Smiles
1	O=C(N)C1=C(N)[N@]([C@]2=C(C)C=CC(O)=C2C)C3=C1C=C(C)C(C)=N3
2	O=C(C1=CN=C(NC2=CC(N3CCN(CCO)CC3)=NC(C)=N2)S1)NC4=C(C)C=CC=C4Cl
3	O=C(N1CC2[N](CC1)=NC=C2)C3=C(N)[N@]([C@]4=C(C)C=CC(O)=C4C)C5=C3C=C(C6CC6)C(C)=N5
4	CN(C1=C(C=C2C3=C(Cl)C=CC=C3Cl)C=NC(NC4=NC=C(OC5=NN(CC)C=C5)C=C4)=N1)C2=O
5	NC1=C(C(NC(C2=CN(C)N=C2)=C3)=O)C(C3=NC=N4)=C4[N@@]1[C@@]5=C(C)C=CC(O)=C5C
6	CC(C=CC(O)=C1C)=[C@]1[C@@]2=C(C#N)C(C(Cl)=C(C3=CN=C(N4[C@@H](C5)CO[C@@H]5C4)N=C3)N6)=C6N=C2
7	O=C(N1CC2[N](CC1)=NC=C2)C3=C(N)[N@]([C@]4=C(C)C=CC(O)=C4C)C5=C3C=C(C)C(C)=N5
8	O=C(N1CC2[N](CC1)=NC=C2Cl)C3=C(N)[N@]([C@]4=C(C)C=CC(O)=C4C)C5=C3N=C(C)C(C)=N5
9	O=C(N1CC2[N](CC1)=NC=C2C#N)C3=C(N)[N@]([C@]4=C(C)C=CC(O)=C4C)C5=C3N=C(C)C(C)=N5
10	NC1=C(C(N)=O)C(C2=NC=CS2)=NN1C3=C(C)C=CC(O)=C3C
11	NC1=C(C(N)=O)C(C2=NC(C)=CS2)=NN1C3=C(C)C=CC(O)=C3C
12	NC1=C(C(N)=O)C(C2=NC(C(F)(F)F)=CS2)=NN1C3=C(C)C=CC(O)=C3C
13	O=C(N)C1=C(N)N(C2=C(C)C=CC(O)=C2C)C3=C1N=C(C=CC=C4Br)C4=N3
14	O=C(N)C1=C(N)N(C2=C(C)C=CC(O)=C2C)C3=C1N=C(C=CC=C4C#N)C4=N3
15	O=C(N)C1=C(N)N(C2=C(C)C=CC(O)=C2C)C3=C1N=C(C=C(Br)C=C4)C4=N3
16	O=C(N)C1=C(N)N(C2=C(C)C=CC(O)=C2C)C3=C1N=C(C=C(C#N)C=C4)C4=N3
17	O=C(N)C1=C(N)N(C2=C(C)C=CC(O)=C2C)C3=C1N=C(C=C(C4=CC=NN4C)C=C5)C5=N3
18	COC(C=C1)=CN=C1NC2=NC3=C(C=C(C4=C(C)C(O)=CC=C4C)C(N3C)=O)C=N2
19	O=C(NC(C1=CC=CC=C1)=N2)C3=C2C(C4=NC=CS4)=CC(C5=C(C=NN6)C6=CC=C5)=C3N
20	O=C(NC=N1)C2=C1C(C3=NC=CS3)=CC(C4=C(C=NN5)C5=CC=C4C)=C2N
21	BrC1=CC(C2=C(C=NN3)C3=CC=C2)=C(N)C4=C1N=C(C5=CN=C(N6CCNCC6)C=C5)NC4=O

[a]all of negative compounds were generate via these compounds by DUDE web.

The dynamics simulations data of hit compounds and positive control.

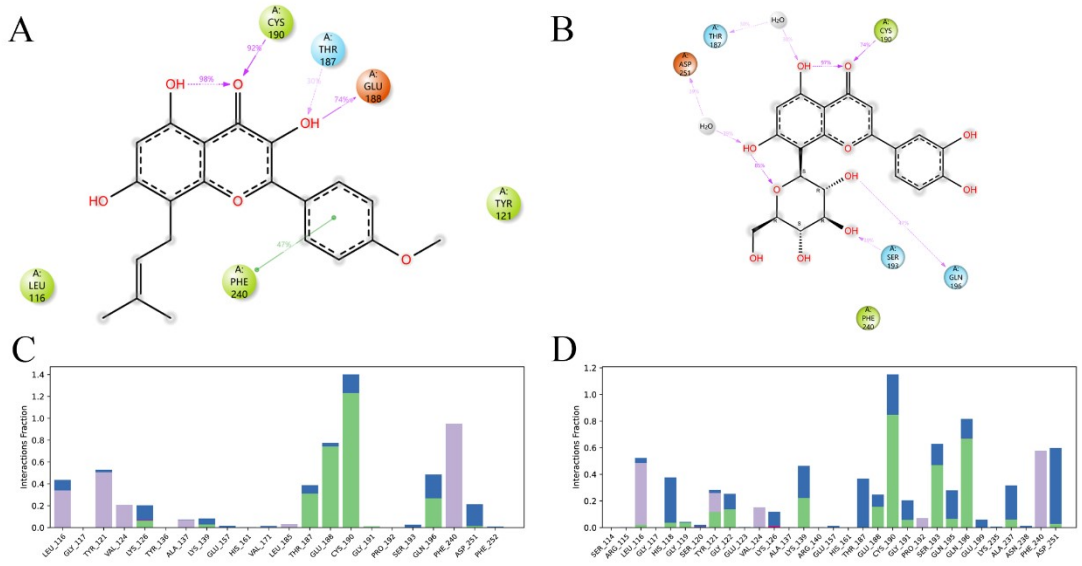

Fig. S1 2D interaction contacts diagram of (A) Icaritin and (B) Genistin, Histogram displaying of interaction fractions between protein and ligand through MD trajectory for (C) Icaritin and (D) Genistin.

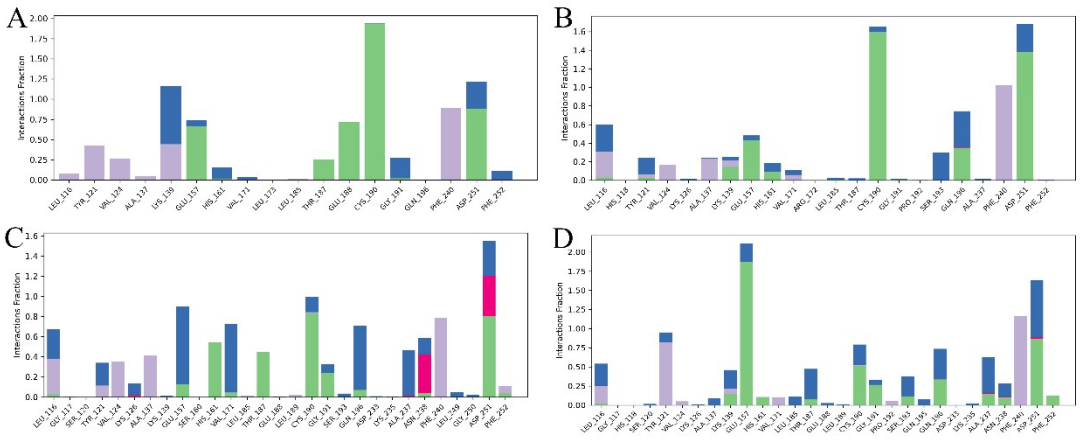

Fig. S2 Histogram displaying of interaction fractions between protein and ligand through MD trajectory for (A) RP-6306, (B) Luteolin, (C) EGCG and (D) GCG.

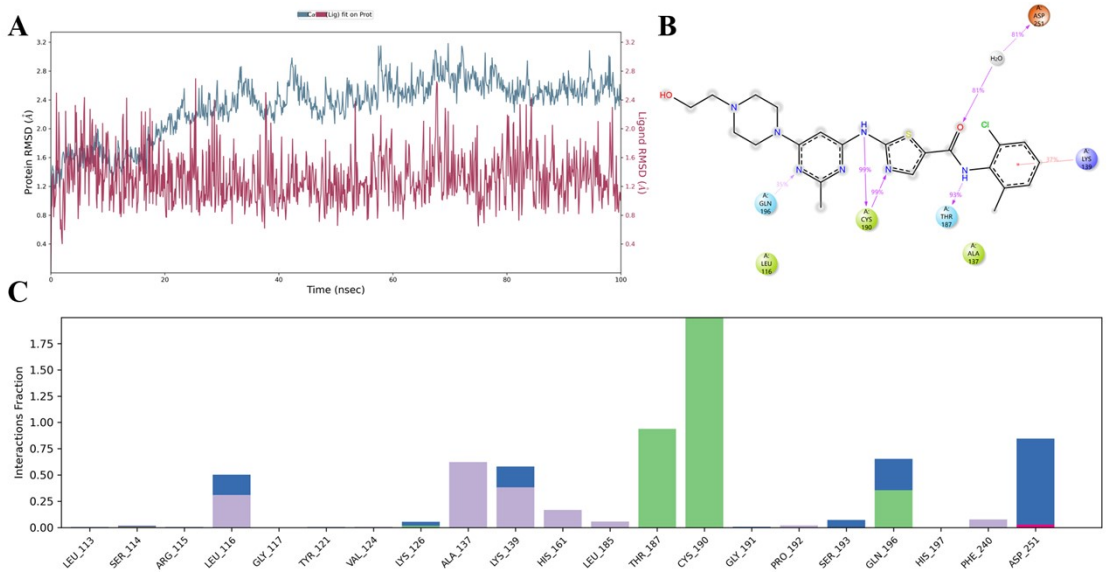

Fig. S3 MD simulations analyses of **Dasatinib**. (A) RMSD plots of 5VCV. The above plot shows the RMSD evolution of a protein. Ligand RMSD (right Y-axis) indicates how stable the ligand is with respect to the protein and its binding pocket. (B) Protein-ligand contacts of 5VCV. (C) Histogram displaying of 5VCV.

Raw data from compound kinase activity tests

Table S3 Test concentration and inhibition rate of EGCG

Concentration (nM)	1 (Inhibition rate %)	2 (Inhibition rate %)	3 (Inhibition rate %)
10000	90.6081	88.4624	102.336
1000	97.0403	91.9284	101.882
200	83.6153	89.3351	73.9298
100	29.7895	20.6302	10.9172
20	5.8468	5.6435	7.4633
8	13.0232	8.7564	-0.3341
1.6	8.5323	0.1557	-0.0959

Table S4 Test concentration and inhibition rate of GCG

Concentration (nM)	1 (Inhibition rate %)	2 (Inhibition rate %)	3 (Inhibition rate %)
10000	91.1944	93.3276	101.614
1000	95.1757	98.8724	101.196
200	71.5008	69.2304	75.4422
100	23.1747	19.2255	17.3848
20	11.2653	9.4473	14.2025
8	9.3872	10.599	7.0183
1.6	10.607	0.1088	7.2377

Table S5 Test concentration and inhibition rate of Luteolin

Concentration (nM)	1 (Inhibition rate %)	2 (Inhibition rate %)	3 (Inhibition rate %)
10000	94.3423	98.0374	100.401
2000	70.9295	45.9267	85.1788
400	11.6908	2.4964	13.3276
80	2.3864	10.1222	9.5818
16	2.6408	-10.2165	8.8092
1.6	6.4631	0.9268	6.1909

Table S6 Test concentration and inhibition rate of RP-6306

Concentration (nM)	1 (Inhibition rate %)	2 (Inhibition rate %)	3 (Inhibition rate %)
1000	96.2896	97.6238	98.2138
200	97.9096	95.7139	97.775
40	98.7672	99.1856	98.3973
8	94.6417	94.597	93.2614
1.6	78.406	67.5346	60.8334
0.32	22.6529	15.8324	11.129
0.064	4.2735	1.4415	0.8177
0.0128	1.8042	1.2886	0.7713

Table S7 Test concentration and inhibition rate of Genistin

Concentration (nM)	1 (Inhibition rate %)	2 (Inhibition rate %)	3 (Inhibition rate %)
100000	57.8404	61.0138	58.2416
50000	41.1270	44.5854	41.4200
10000	15.6763	21.7683	24.2202
2000	13.8721	22.7485	15.1127
400	10.4627	14.8606	10.8415
80	-1.1880	1.5044	13.2167

Table S8 Test concentration and inhibition rate of Icaritin

Concentration (nM)	1 (Inhibition rate %)	2 (Inhibition rate %)	3 (Inhibition rate %)
100000	59.9511	65.0404	48.9980
50000	34.6897	48.6212	41.6845
10000	3.6226	19.5036	22.0069
2000	1.5641	0.1499	17.3470
400	17.5021	11.5058	13.5330
80	14.6591	18.8085	18.9148