Supporting Information for

Discovery of Selective, Substrate-Competitive and Passive Membrane Permeable Glycogen Synthase Kinase-3β Inhibitors: Synthesis, Biological Evaluation, and Molecular Modeling of New C-Glycosylflavones

Zhibin Liang and Qing X. Li*

Department of Molecular Biosciences and Bioengineering, University of Hawaii at Manoa, Honolulu, Hawaii 96822, United States

*Corresponding author. Tel: +1 (808) 965-2011; Fax: +1 (808) 965-3542. E-mail address: qingl@hawaii.edu

Table of Contents

S1. Docking Method Validation by Re-docking and Cross-docking Experiments with Known GSK-3β Inhibitors 3

S2. Inhibitions and Docking Scores of C-Glycosylflavones and Reference Inhibitors to GSK3β 5

S3. Predicted Docking Poses of Compounds 1, 9, 17, 21, 3, and 4 with GSK-3β (PDB Code 1PYX) 7

S4. Summary of Molecular Interaction Distances of Compound 30 with GSK3β (PDB Code 1PYX) in Molecular Docking 8

S5. GSK-3α Homology Modeling Method 9

S6. Protein Sequence Alignment Between Human GSK-3α (UniProt Code P49840) and Human GSK-3β (PDB Code 1PYX) 11

S7. HPLC Analysis for Purity Determination of Compounds 5-31 12

S8. In Vitro Kinase Selectivity Screening for Compound 30 17

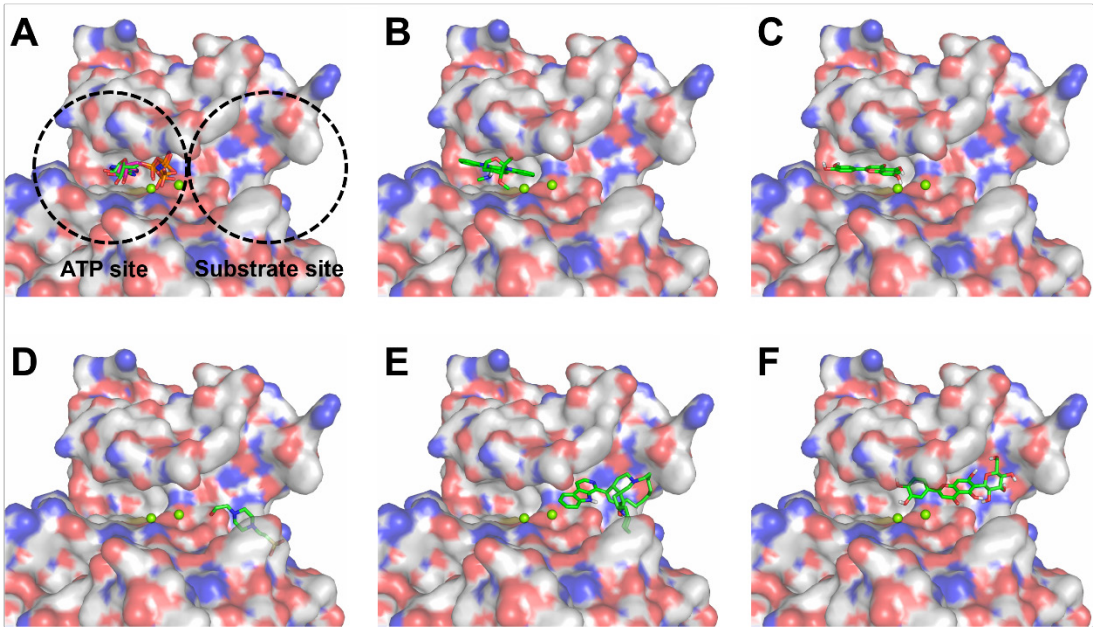
S9. LC-MS Method for Determination of the PAMPA Permeability 19

S10. Structures of the Reference Compounds 20

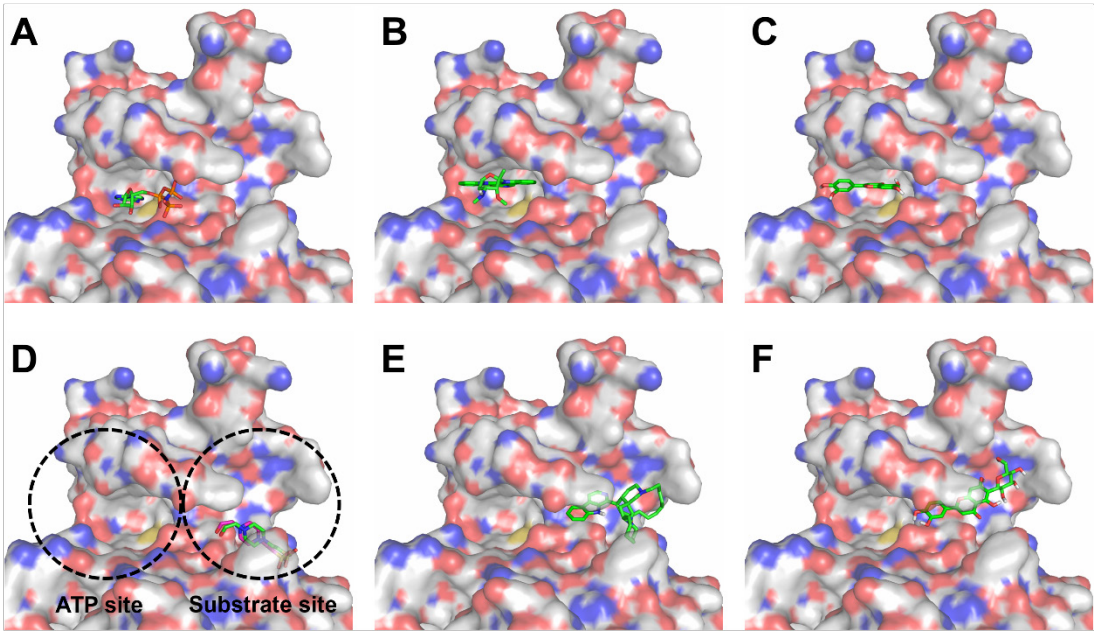
S1. Docking Method Validation by Re-docking and Cross-docking Experiments with Known GSK-3β Inhibitors

We chose two X-ray crystallographic structures of GSK-3β (PDB codes 1PYX[1] and 1H8F[2]). The docking method was validated by re-docking and cross-docking experiments using both ATP-competitive and substrate-competitive inhibitors.

Re-docking the native ligand ANP to 1PYX resulted in a binding pose at the ATP site of GSK-3β with a RMSD of 0.33 Å as compared to its original crystal structure.[1] Under the same parameters, docking with the known ATP-competitive inhibitors, staurosporine and luteolin, led to the binding to the ATP site of GSK-3β, of which the binding poses and positions are similar to the reported data.[1, 3] The ligands HEPES, manzamine A and isoorientin are known GSK-3β inhibitors bound to the substrate site. Docking these ligands to 1PYX confirmed their binding to the substrate site, and the predicted poses are comparable to the reported data.[2, 4, 5] Cross-docking of the same ligands to the 1H8F demonstrated that ANP, staurosporine, and luteolin reside the ATP site of GSK-3β, whereas HEPES, manzamine A, and isoorientin are in the substrate site. Re-docking experiment using the native ligand HEPES of 1H8F resulted in a binding pose at the substrate site of GSK-3β with a RMSD of 0.62 Å as compared to its original crystal structure.[2]



Predicted docking poses of ANP (A), staurosporine (B), luteolin (C), HEPES (D), manzamine A (E), and isoorientin (F) in GSK-3β (PDB code 1PYX). Ligands in magenta color are original structures in X-ray crystallographic complex, ligands in green color are docking structures.

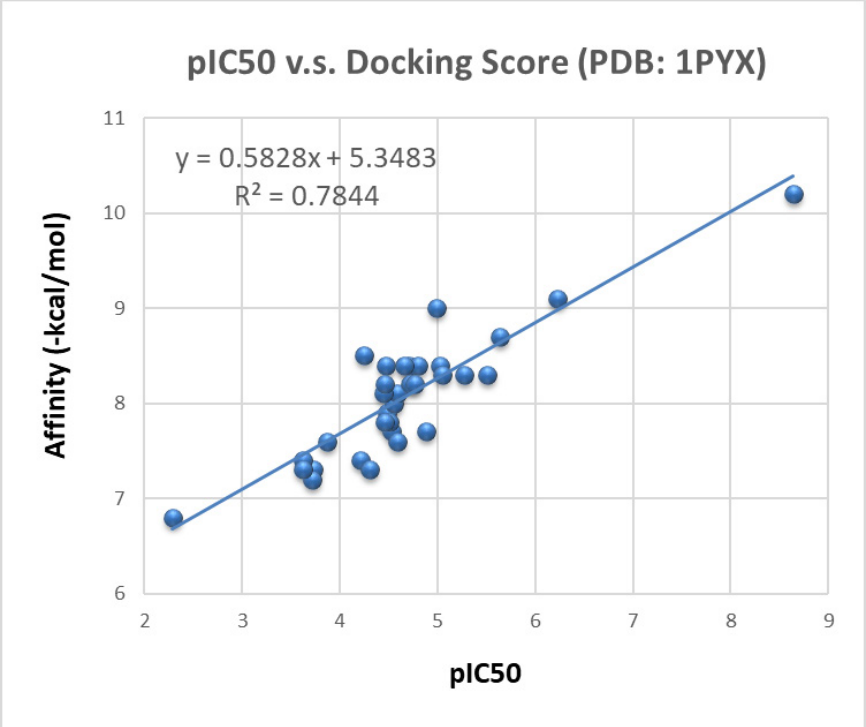


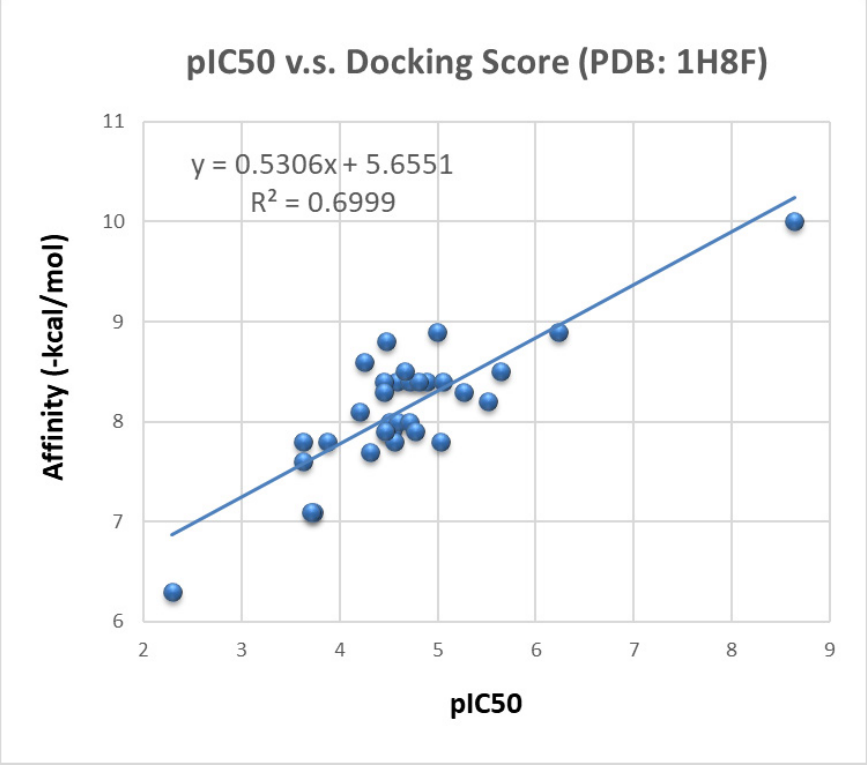
Predicted docking poses of ANP (A), staurosporine (B), luteolin (C), HEPES (D), manzamine A (E), and isoorientin (F) into GSK-3β (PDB code 1H8F). Ligands in magenta color are original structures in X-ray crystallographic complex, ligands in green color are docking structures.

S2. Inhibitions and Docking Scores of *C*-Glycosylflavones and Reference Inhibitors to GSK3β

cmpd	GSK3β inhibition	Docking Score [b]: Affinity (-kcal/mol)		cmpd	GSK3β inhibition	Docking Score [b]: Affinity (-kcal/mol)	
no.	**IC_{50} (μM) [a]**	**1PYX**	**1H8F**	**no.**	**IC_{50} (μM) [a]**	**1PYX**	**1H8F**
1	184.9 ± 1.4	7.3	7.1	**21**	15.8 ± 1.2	8.4	8.4
2	194.1 ± 1.0	7.2	7.1	**22**	35.6 ± 1.3	8.1	8.4
3	5153 ± 31	6.8	6.3	**23**	35.1 ± 1.0	8.2	8.3
4	3.1 ± 1.3	8.3	8.2	**24**	56.2 ± 1.2	8.5	8.6
5	239.2 ± 1.2	7.4	7.8	**25**	21.7 ± 1.2	8.4	8.5
6	237.3 ± 1.4	7.3	7.6	**26**	34.8 ± 1.0	7.8	7.9
7	135.0 ± 1.3	7.6	7.8	**27**	19.3 ± 0.8	8.2	8.0
8	29.2 ± 1.1	7.7	7.9	**28**	49.5 ± 1.2	7.3	7.7
9	5.4 ± 0.1	8.3	8.3	**29**	17.2 ± 0.9	8.2	7.9
10	33.3 ± 1.1	7.9	7.9	**30**	0.59 ± 0.04	9.1	8.9
11	31.1 ± 1.2	7.8	8.0	**31**	2.3 ± 0.5	8.7	8.5
12	28.0 ± 1.1	8.0	7.8	**Staurosporine**	0.0023 ± 0.0006	10.2	10.0
13	9.4 ± 0.9	8.4	7.8	**Manzamine A**	10.2 [c]	9.0	8.9
14	25.8 ± 1.1	7.6	8.0	**ANP**	n/a	9.7	9.3
15	61.9 ± 1.2	7.4	8.1	**HEPES**	n/a	4.4	4.5
16	13.1 ± 1.1	7.7	8.4				
17	9.0 ± 1.3	8.3	8.4				
18	26.0 ± 1.2	8.1	8.4				
19	33.5 ± 0.8	8.4	8.8				
20	19.7 ± 1.1	8.4	8.4				

[a] IC_{50} values were the mean of quadruplicate of each of two independent experiments. [b] Docking scores were binding affinities calculated by AutoDock Vina. [c] IC_{50} value from ref 6.

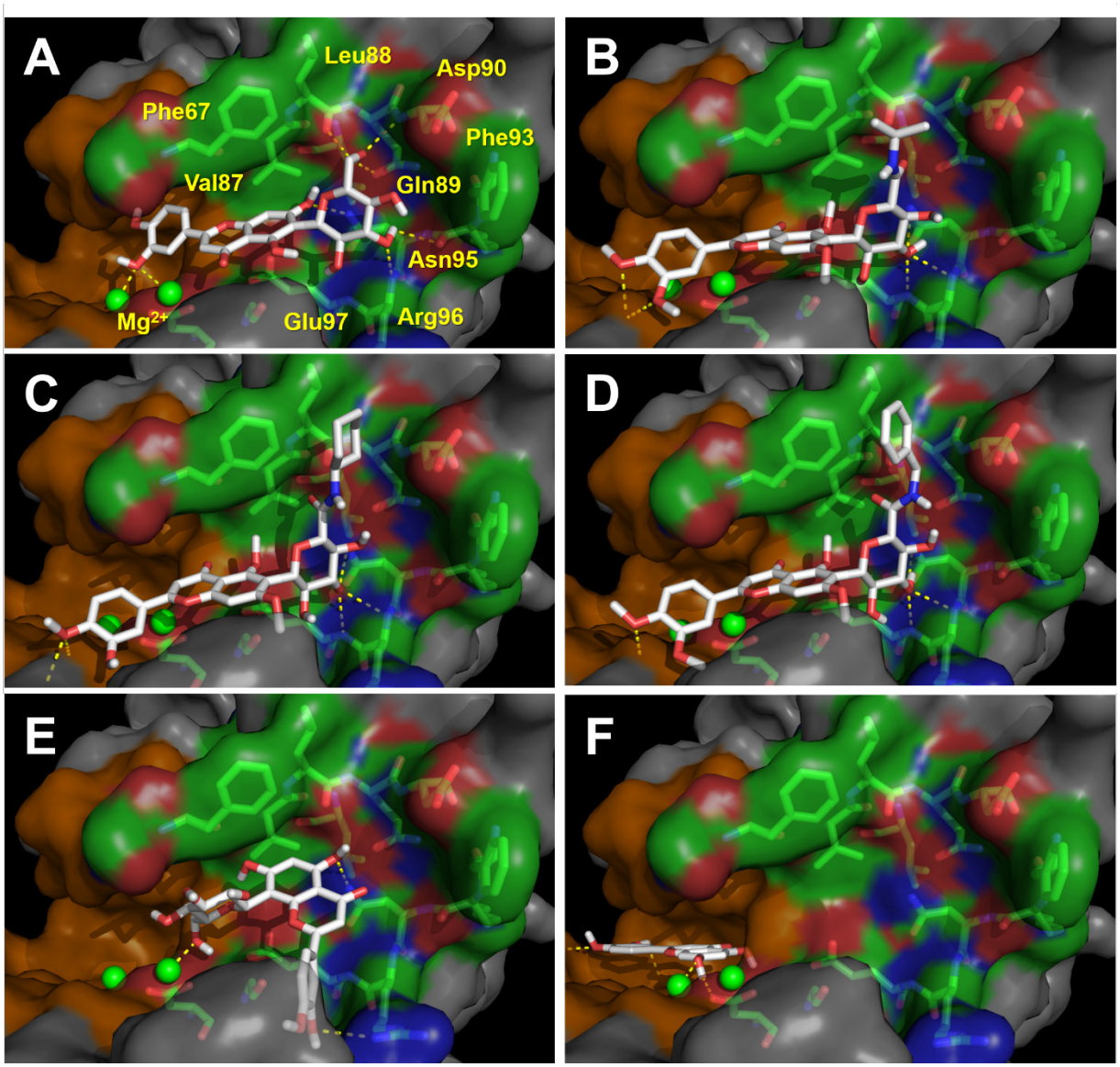




Linear plots of pIC50 (-LogIC50) versus calculated binding affinity (PDB codes 1PYX and 1H8F) for GSK-3β inhibitors

S3. Predicted Docking Poses of Flavones with GSK-3β (PDB Code 1PYX)

Docking poses of flavones **1** (A), **9** (B), **17** (C), **21** (D), **3** (E), and **4** (F) with GSK-3β (PDB code 1PYX). Surface representation of the ATP site of GSK-3β is shown in orange, the substrate site of GSK-3β is shown in green. Within the substrate site, surface features are shown with acidic residues in red, basic residues in blue, and hydrophobic residues in green. Mg^{2+} ions are shown as bright green spheres. The dotted lines represent interactions with Mg^{2+} ions and hydrogen bonds with side chain residues of GSK-3β.



S4. Summary of Molecular Interaction Distances of Compound 30 with GSK3β (PDB Code 1PYX) in Molecular Docking

Interaction type	Interaction residue	Distance (Å)
Hydrogen bonds	Carbonyl O···HN (Phe67)	2.3
	Methoxy O···HN (Lys183)	3.1, 3.2
	Glycone HO···HN (Leu88)	3.4
	Glycone HO···HN (Asn95)	1.8, 2.6
	Glycone OH···O=C (Leu88)	2.8
	Glycone OH···O=C (Gln89)	3.2
	Glycone OH···O=C (Glu97)	2.8
	Glycone OH···O=C (Gly202)	2.5
Hydrophobic interactions	Isopropyl Me (Phe67)	
	Methoxy Me (Val87)	
	Isopropyl CF_3 (Leu88)	
π-cation interaction	Catechol B-ring···NH_3^+ (Lys183)	4.1 (angle 71.5°)
Orthogonal multipolar interactions	Isopropyl F···C=O (Leu88)	3.3
	Isopropyl F···C=O (Gln89)	3.6
	Isopropyl F···C=O (Asp90)	3.3, 3.8
Fluorine polar interaction	Isopropyl F···HN (Asp90)	2.6

S5. GSK-3α Homology Modeling Method

The SWISS-MODEL[7] template library (SMTL version 2017-09-21, PDB release 2017-09-15) was searched with BLAST and HHBlits for evolutionary related structures matching the target sequence GSK-3α (UniProt Code P49840). For each identified template, the template's quality has been predicted from features of the target-template alignment. The templates with the highest quality have then been selected for model building. Overall 4470 templates were found to match the target sequence. The top templates are:

	Name	Title	Coverage	Identity	Method	Oligo State	Ligands
☐	5air.1.A	LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED PROTEIN 6, GLYCOGEN SYNTHASE KINASE-3 BETA		81.87	X-ray, 2.5Å	homo-dimer	None
☐	5air.1.B	LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED PROTEIN 6, GLYCOGEN SYNTHASE KINASE-3 BETA		81.87	X-ray, 2.5Å	homo-dimer	None
☑ ✓	1pyx.1.A	Glycogen synthase kinase-3 beta		82.97	X-ray, 2.4Å	homo-dimer	4 x MG, 2 x ANP
☐	1q5k.1.A	Glycogen synthase kinase-3 beta		82.97	X-ray, 1.9Å	homo-dimer	2 x TMU
☐	4j1r.1.A	Glycogen synthase kinase-3 beta		82.97	X-ray, 2.7Å	monomer	1 x K, 1 x I5R
☐	1q3d.1.B	GLYCOGEN SYNTHASE KINASE-3 BETA		82.97	X-ray, 2.2Å	homo-dimer	2 x STU
☐	1j1b.1.A	Glycogen synthase kinase-3 beta		82.97	X-ray, 1.8Å	homo-dimer	2 x ANP
☐	4acg.1.A	GLYCOGEN SYNTHASE KINASE-3 BETA		82.97	X-ray, 2.6Å	homo-dimer	2 x 6LQ
☐	4e7w.1.A	Glycogen Synthase Kinase 3		69.16	X-ray, 3.3Å	monomer	None
☐	4e7w.1.A	Glycogen Synthase Kinase 3		70.39	X-ray, 3.3Å	monomer	None
☐	1h27.1.A	CELL DIVISION PROTEIN KINASE 2		40.14	X-ray, 2.2Å	hetero-oligomer	None
☐	1h1p.2.A	CELL DIVISION PROTEIN KINASE 2		40.14	X-ray, 2.1Å	hetero-oligomer	1 x CMG
☐	1h28.1.A	CELL DIVISION PROTEIN KINASE 2		40.14	X-ray, 2.8Å	hetero-oligomer	None
☐	1h24.2.A	CELL DIVISION PROTEIN KINASE 2		40.14	X-ray, 2.5Å	hetero-oligomer	None
☐	1h1q.2.A	CELL DIVISION PROTEIN KINASE 2		40.14	X-ray, 2.5Å	hetero-oligomer	1 x 2A6
☐	1vyw.2.A	CELL DIVISION PROTEIN KINASE 2		39.43	X-ray, 2.3Å	hetero-oligomer	1 x 292

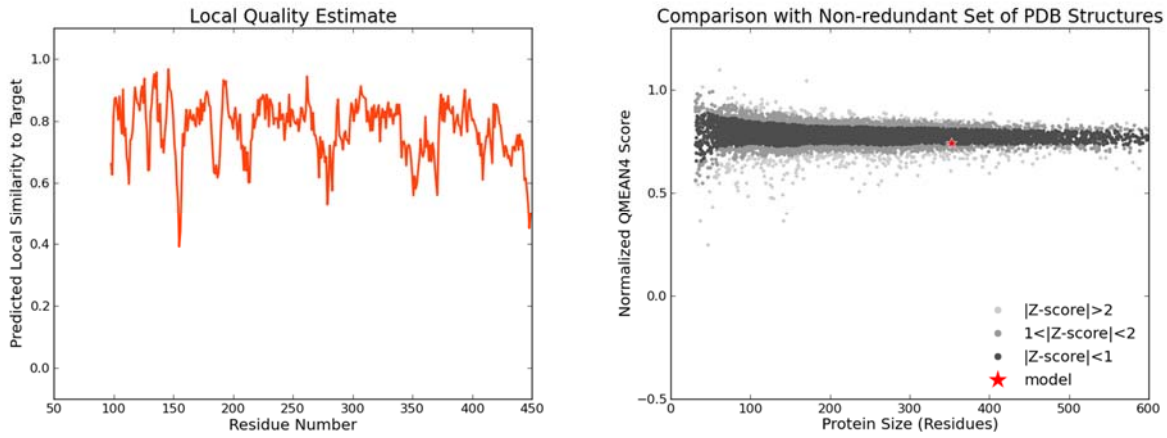
The GSK-3α homology model was built based on the GSK-3β template PDB code 1PYX:

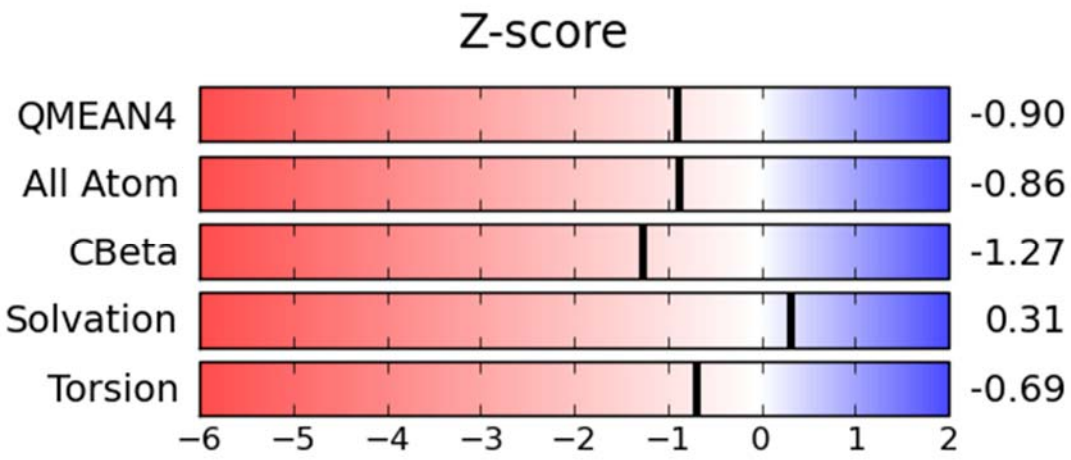
Template	Sequence identity	Oligo-state	Found by	Method	Resolution	Sequence similarity	Range	Coverage	Description
1pyx.1.A	82.97	homodimer	BLAST	X-ray	2.40 Å	0.55	98–449	0.77	GSK-3β

Built with	Oligo-State	Ligands	GMQE	QMEAN
ProMod3 Version 1.0.2.	monomer	Mg^{2+} ion	0.72	-0.90

Model Quality Estimation:

The global and per-residue model quality has been assessed using the QMEAN scoring function. For improved performance, weights of the individual QMEAN terms have been trained specifically for SWISS-MODEL.





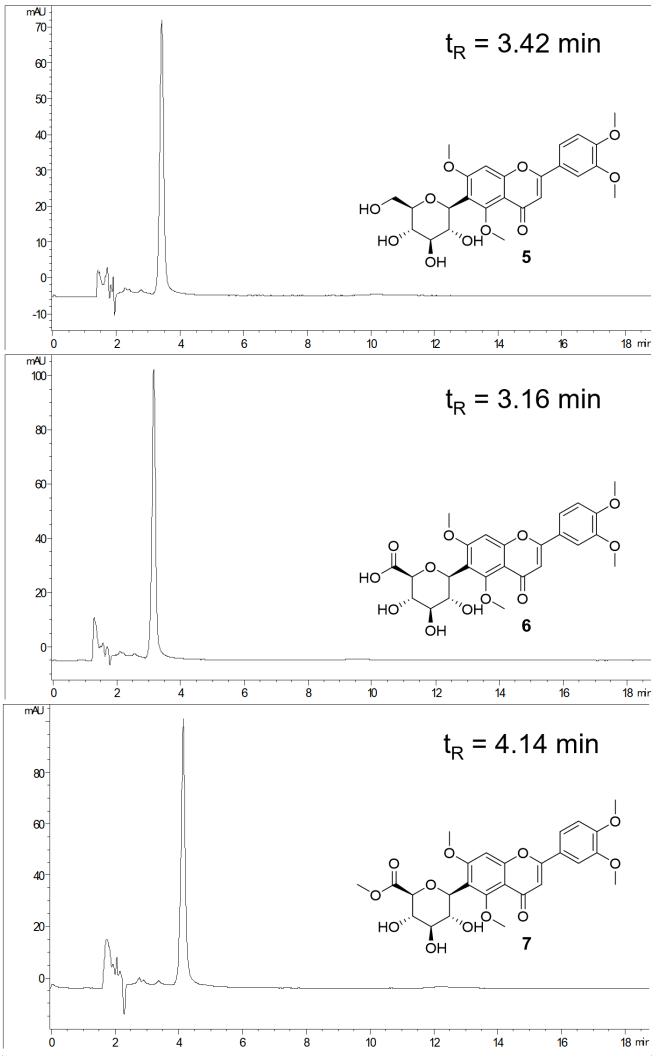
S6. Protein Sequence Alignment Between Human GSK-3α (UniProt Code P49840) and Human GSK-3β (PDB Code 1PYX)

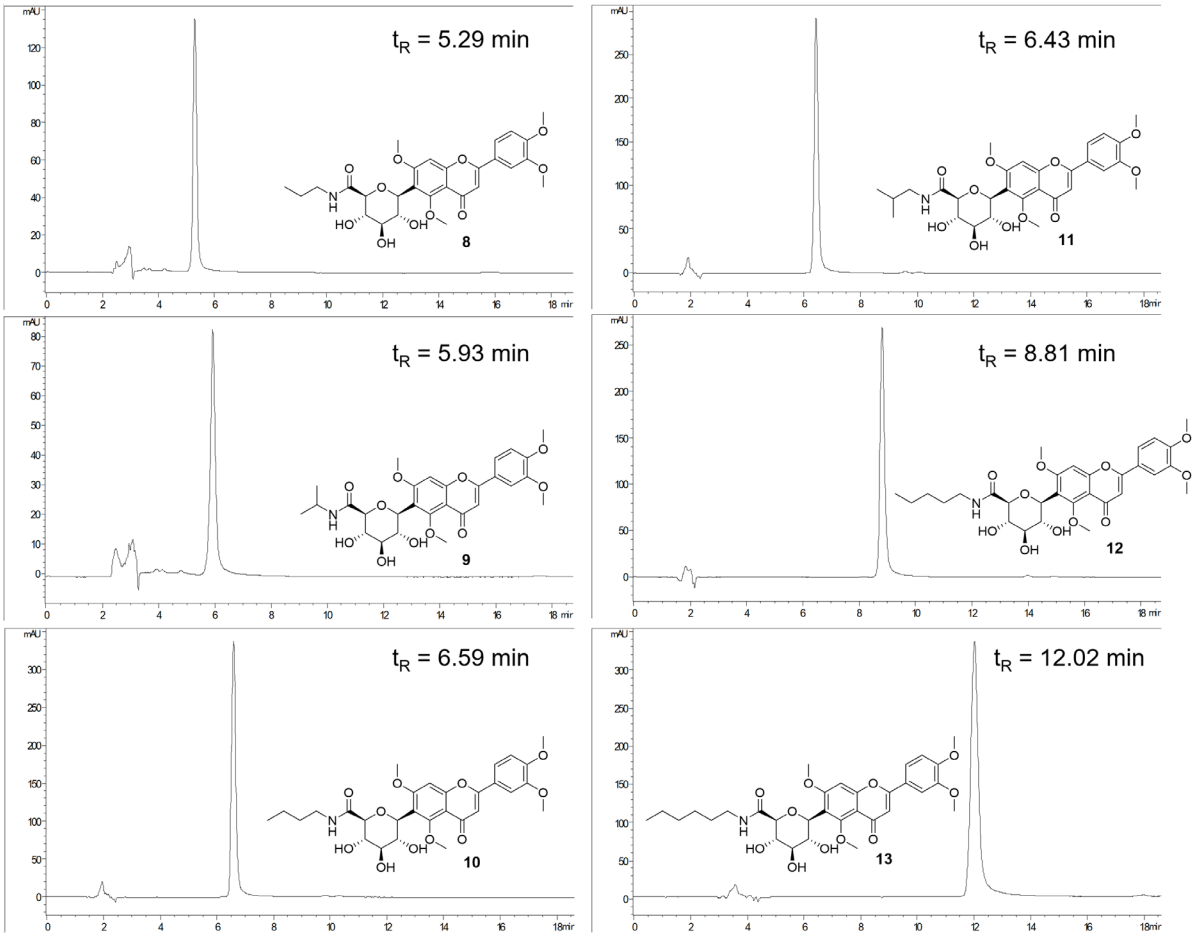
Identical residues are unpainted, nonidentical residues are yellow highlighted, unmatched residues are green highlighted; key residues comprised of ATP site are colored in red, key residues comprised of substrate site are colored in blue. It is obvious that the kinase catalytic domain comprised of both ATP and substrate pocket sequences (D121-H208 and C241-L270 in GSK-3α; D58-H145 and C178-L207 in GSK-3β) are conserved as high as 93% identical. Only nine amino acid residues are nonidentical in these regions between two isoforms.

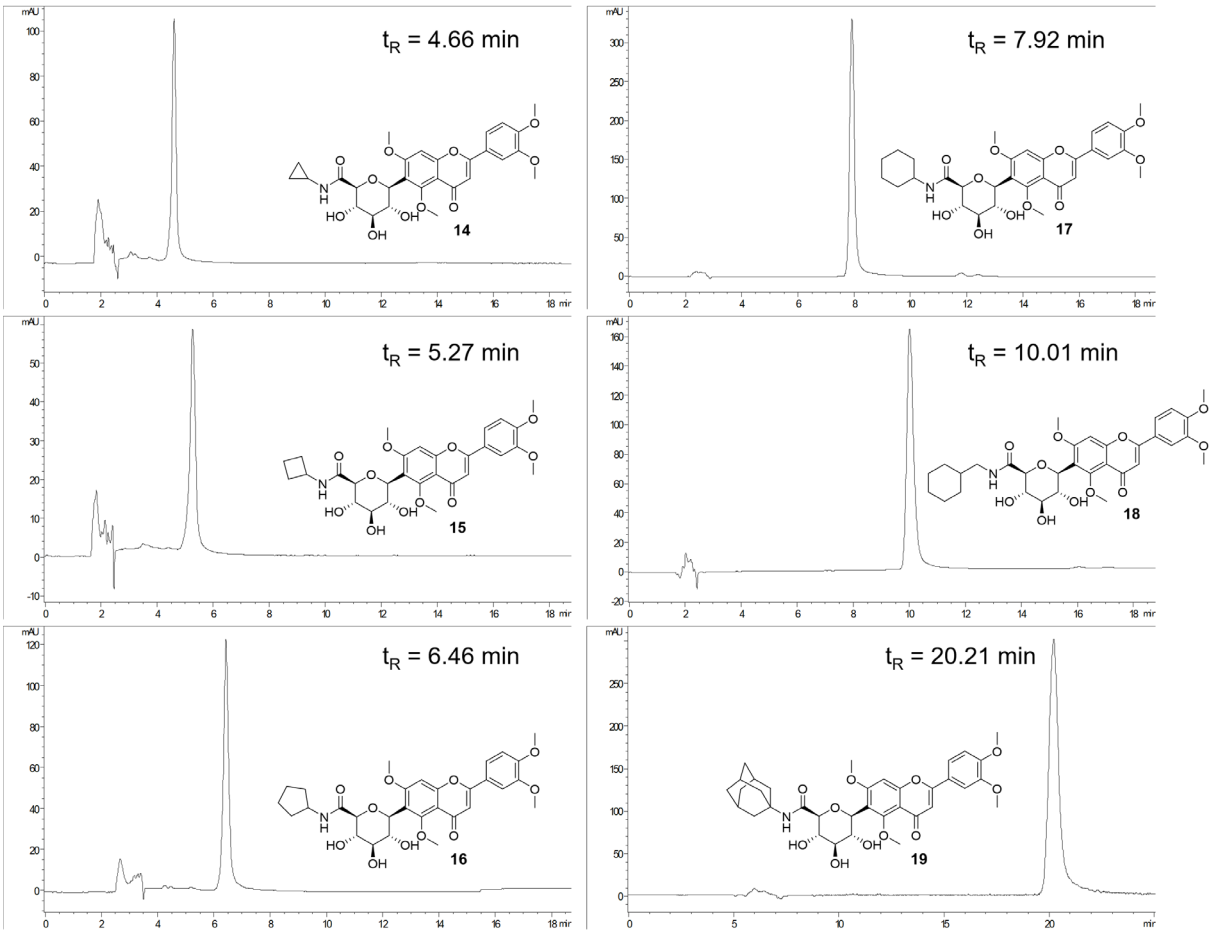
P49840_hGSK3A	MSGGGPSGGGPGGSGRARTSSFAEPGGGGGGGGGGPGGSASGPGGTGGGKASVGAMGGGV	60
1PYX_hGSK3B	--	
P49840_hGSK3A	GASSSGGGPGGSGGGGSGGPGAGTSFPPPGVKLGRDSGKVTTVVATLGQGPERSQEVAYT	120
1PYX_hGSK3B	--------------------------------VSRDKDGSKVTTVVATPGQGPDRPQEVSYT	57
P49840_hGSK3A	DIKVIGNGSFGVVYQARLAETRELVAIKKVLQDKRFKNRELQIMRKLDHCNIVRLRYFFY	180
1PYX_hGSK3B	DTKVIGNGSFGVVYQAKLCDSGELVAIKKVLQDKRFKNRELQIMRKLDHCNIVRLRYFFY	117
P49840_hGSK3A	SSGEKKDELYLNLVLEYVPETVYRVARHFTKAKLTIPILYVKVYMYQLFRSLAYIHSQGV	240
1PYX_hGSK3B	SSGEKKDEVYLNLVLDYVPETVYRVARHYSRAKQTLPVIYVKLYMYQLFRSLAYIHSFGI	177
P49840_hGSK3A	CHRDIKPQNLLVDPDTAVLKLCDFGSAKQLVRGEPNVSYICSRYYRAPELIFGATDYTSS	300
1PYX_hGSK3B	CHRDIKPQNLLLDPDTAVLKLCDFGSAKQLVRGEPNVSYICSRYYRAPELIFGATDYTSS	237
P49840_hGSK3A	IDVWSAGCVLAELLLGQPIFPGDSGVDQLVEIIKVLGTPTREQIREMNPNYTEFKFPQIK	360
1PYX_hGSK3B	IDVWSAGCVLAELLLGQPIFPGDSGVDQLVEIIKVLGTPTREQIREMNPNYTEFKFPQIK	297
P49840_hGSK3A	AHPWTKVFKSRTPPEAIALCSSLLEYTPSSRLSPLEACAHSFFDELRCLGTQLPNNRPLP	420
1PYX_hGSK3B	AHPWTKVFRPRTPPEAIALCSRLLEYTPTARLTPLEACAHSFFDELRDPNVKLPNGRDTP	357
P49840_hGSK3A	PLFNFSAGELSIQPSLNAILIPPHLRSPAGTTTLTPSSQALTETPTSSDWQSTDATPTLT	480
1PYX_hGSK3B	ALFNFTTQELSSNPPLATILIPPHARIQAAASTPTNATAA----------------------	397
P49840_hGSK3A	NSS	483
1PYX_hGSK3B	---	

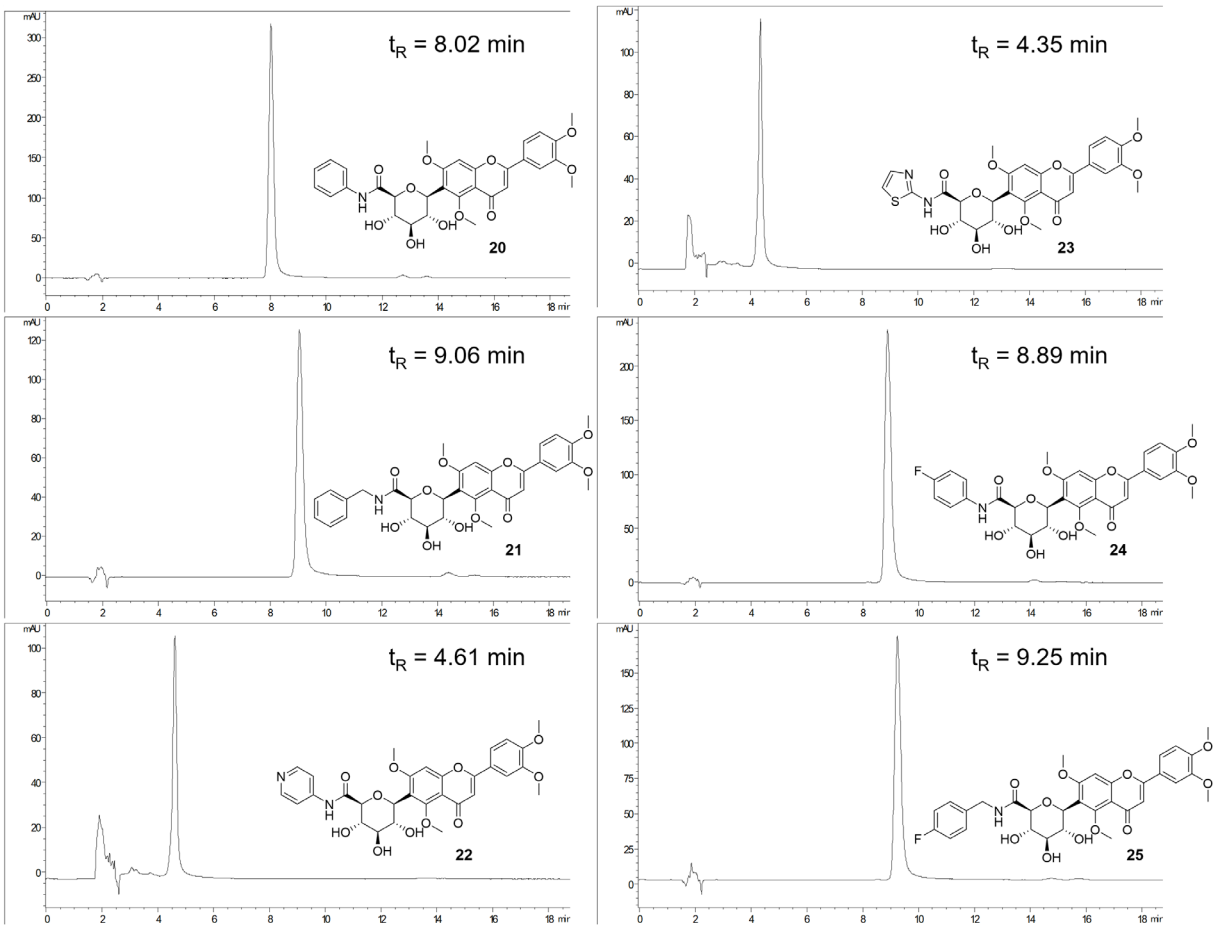
S7. HPLC Analysis for Purity Determination of Compounds 5-31

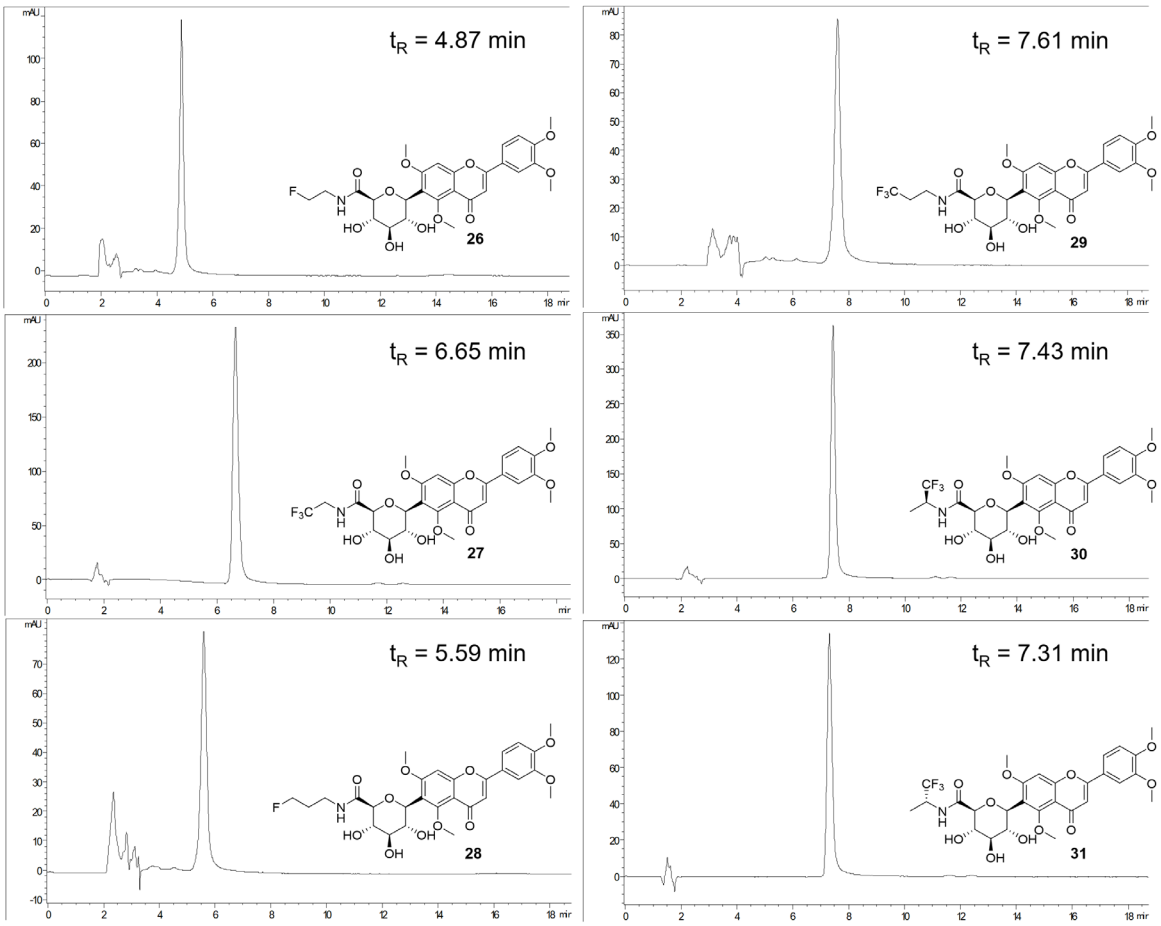
The purity of compounds **5-31** was determined by analytical reverse phase Agilent HPLC with a diode array detector (Waters XSELECT CSH Fluoro-Phenyl column, 150 × 4.6 mm, 3.5 μm, isocratic elution with 35% aqueous acetonitrile containing 0.1% formic acid for over 20 min, detection at 210 nm, flow rate of 0.8 mL/min). All tested compounds were over 95% purity.




t_R = 5.29 min
8
t_R = 6.43 min
11
t_R = 5.93 min
9
t_R = 8.81 min
12
t_R = 6.59 min
10
t_R = 12.02 min
13


t_R = 4.66 min
14
t_R = 7.92 min
17
t_R = 5.27 min
15
t_R = 10.01 min
18
t_R = 6.46 min
16
t_R = 20.21 min
19


t_R = 8.02 min
20
t_R = 4.35 min
23
t_R = 9.06 min
21
t_R = 8.89 min
24
t_R = 4.61 min
22
t_R = 9.25 min
25


t_R = 4.87 min
26
t_R = 7.61 min
29
t_R = 6.65 min
27
t_R = 7.43 min
30
t_R = 5.59 min
28
t_R = 7.31 min
31

S8. In Vitro Kinase Selectivity Screening for Compound 30

The Kinase Selectivity Profiling System includes kinase and substrate pairs. Percentage kinase activities upon treatment of compound **30** at 5 μM in a panel of human protein kinases were measured by the ADP-Glo Kinase Assay Kit according to the manufacturer's protocol. The kinase inhibitor staurosporine was used at 1 μM as a reference control. Each data point was collected in quadruplicate of two independent experiments.

Assay Kinase	Substrate/Co-Factor	Kinase Activity (%)
ERK2	MBP	101.3
GSK3α	GSK3 Substrate	92.5
GSK3β	GSK3 Substrate	7.7
JNK1	p38 Substrate	101.4
JNK3	p38 Substrate	74.5
p38α	p38 Substrate	71.1
p38β	p38 Substrate	75.2
p38δ	p38 Substrate	71.2
p38γ	p38 Substrate	77.5
CDK1/CyclinA2	Histone H1 Protein	55.1
CDK2/CyclinE1	Histone H1 Protein	77.5
CDK3/CyclinE1	Histone H1 Protein	70.4
CDK5/p25	Histone H1 Protein	101.3
CDK5/p35	Histone H1 Protein	95.5
CDK6/CyclinD3	Histone H1 Protein	70.1
CDK9/CyclinK	PDKtide	98.7
CLK3	MBP	88.5
AKT1	AKT (PKB) Substrate	85.5
p70S6Kβ	RSK Substrate	74.3
PDK1	PDKtide	62.3
PKA	Kemptide	94.5
PKC	Neurogranin Peptide Substrate	91.8
PRKG1	RSK Substrate	71.5
ROCK1	S6K Substrate	90.8
RSK2	RSK Substrate	91.3

AMPK A1/B1/G1	SAMStide	97.6
AMPK A1/B1/G2	SAMStide	108.1
AMPK A2/B1/G1	SAMStide	84.7
CAMK2α	Autocamtide-2	75.9
CAMK2γ	Autocamtide-2	94.9
CAMK4	Autocamtide-2	92.8
DAPK1	MBP	77.2
STK33	MBP	72.0
Aurora A	MBP	89.2
Aurora B	MBP	65.4
CK2α1	Casein	105.0
DNA-PK	DNA-PK Peptide Substrate	103.4
CK1α1	De-Phospho Casein	82.3
CK1ε	De-Phospho Casein	77.7
CK1γ1	Casein	70.1
VRK2	Casein	92.5

S9. LC-MS Method for Determination of the PAMPA Permeability

At the end of the incubation period, 10 μL sample from each donor and acceptor well was quantified by LC-ESI-QTOF-MS analysis under a positive mode and used a 7-point calibration with appropriate dilution of the samples. A sum of pseudo-molecular ions $[M+H]^+$, $[M+Na]^+$ and $[M+K]^+$ for each target analyte was extracted from total ion chromatogram (TOC) for quantification. Samples were run in quadruplicate.

Permeability of the compounds (P_e) were calculated using the following formula:

Permeability (cm/s): $P_e = \{-\ln[1-C_A(t)/C_{eq}]\}/[A*(1/V_D+1/V_A)*t]$

where A = filter area (0.3 cm^2), V_D = donor well volume (0.3 mL), V_A = acceptor well volume (0.2 mL), t = incubation time (5 hours = 18000 seconds), $C_A(t)$ = compound concentration in acceptor well at time t, $C_D(t)$ = compound concentration in donor well at time t,

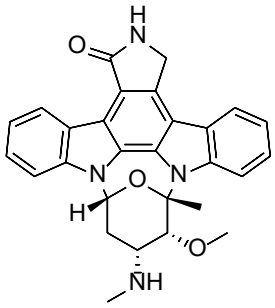
$C_{eq} = [C_D(t)*V_D+C_A(t)*V_A]/(V_D+V_A)$.

Mass retention (*R*%) is calculated using the following formula:

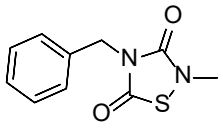
$R\% = \{1 - [C_D(t)*V_D+C_A(t)*V_A]/(C0*V_D)\}*100$

where C0 is the initial compound concentration in donor well.

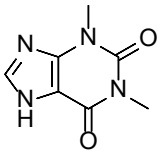
S10. Structures of the Reference Compounds



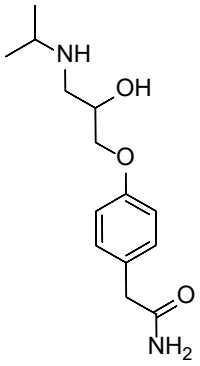
Staurosporine



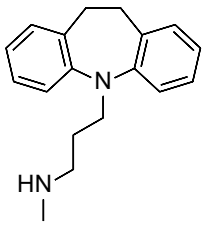
TDZD-8



Theophylline



Atenolol



Desipramine

References

(1) Bertrand, J. A.; Thieffine, S.; Vulpetti, A.; Cristiani, C.; Valsasina, B.; Knapp, S.; Kalisz, H. M.; Flocco, M. Structural characterization of the GSK-3β active site using selective and non-selective ATP-mimetic inhibitors. *J. Mol. Biol.* **2003,** 333, 393-407.

(2) Dajani, R.; Fraser, E.; Roe, S. M.; Young, N.; Good, V.; Dale, T. C.; Pearl, L. H. Crystal structure of glycogen synthase kinase 3β: Structural basis for phosphate-primed substrate specificity and autoinhibition. *Cell* **2001,** 105, 721-732.

(3) Johnson, J. L.; Rupasinghe, S. G.; Stefani, F.; Schuler, M. A.; Gonzalez de Mejia, E. Citrus flavonoids luteolin, apigenin, and quercetin inhibit glycogen synthase kinase-3β enzymatic activity by lowering the interaction energy within the binding cavity. *J. Med. Food* **2011,** 14, 325-333.

(4) Peng, J.; Kudrimoti, S.; Prasanna, S.; Odde, S.; Doerksen, R. J.; Pennaka, H. K.; Choo, Y.-M.; Rao, K. V.; Tekwani, B. L.; Madgula, V.; Khan, S. I.; Wang, B.; Mayer, A. M. S.; Jacob, M. R.; Tu, L. C.; Gertsch, J.; Hamann, M. T. Structure-activity relationship and mechanism of action studies of manzamine analogues for the control of neuroinflammation and cerebral infections. *J. Med. Chem.* **2010,** 53, 61-76.

(5) Liang, Z.; Zhang, B.; Su, W. W.; Williams, P. G.; Li, Q. X. C-Glycosylflavones alleviate tau phosphorylation and amyloid neurotoxicity through GSK3β inhibition. *ACS Chem. Neurosci.* **2016,** 7, 912-923.

(6) Hamann, M.; Alonso, D.; Martín-Aparicio, E.; Fuertes, A.; Pérez-Puerto, M. J.; Castro, A.; Morales, S.; Navarro, M. L.; del Monte-Millán, M.; Medina, M.; Pennaka, H.; Balaiah, A.; Peng, J.; Cook, J.; Wahyuono, S.; Martínez, A. Glycogen synthase kinase-3 (GSK-3) inhibitory activity and structure–activity relationship (SAR) studies of the manzamine alkaloids. Potential for Alzheimer's disease. *J. Nat. Prod.* **2007,** 70, 1397-1405.

(7) Biasini, M.; Bienert, S.; Waterhouse, A.; Arnold, K.; Studer, G.; Schmidt, T.; Kiefer, F.; Cassarino, T. G.; Bertoni, M.; Bordoli, L.; Schwede, T. SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **2014,** 42, W252-W258.