Supplementary Information

Ligand based pharmacophoric discovery of new CLK1 inhibitors

Vidya Jyothi Alli[a,b#], Swapnil Anil Sule[a#], Darna Mounika[a,b], Sai Satya Sri Pulla[a], Pawan Yadav[a] & Surender Singh Jadav*[a,b]

[a] Department of Natural Products and Medicinal Chemistry, CSIR-Indian Institute of Chemical Technology Tarnaka, Uppal Road, Hydrabad 500 037, India

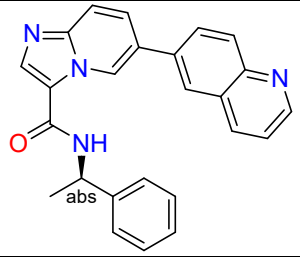
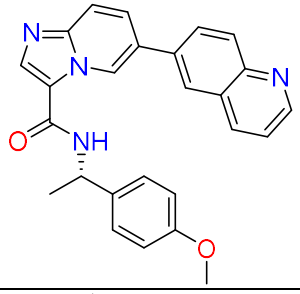
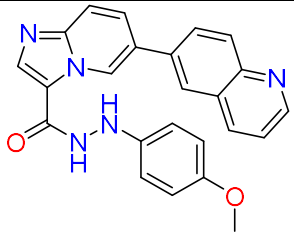
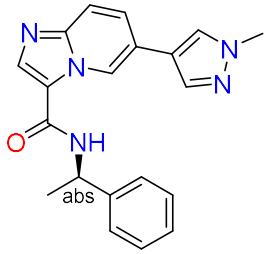
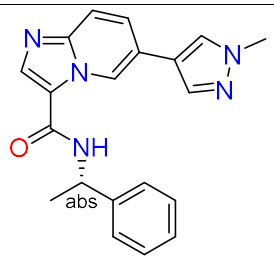
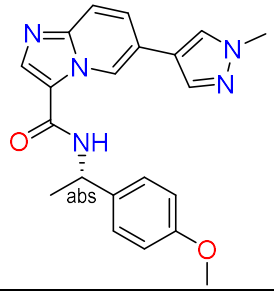
[b] Academy of Scientific and Innovative Research (AcSIR), CSIR-HRDC Campus, Postal Staff College Area

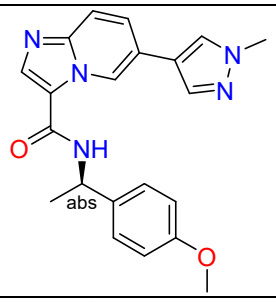
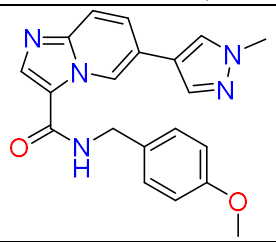
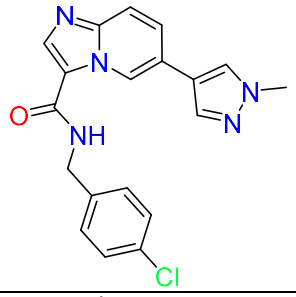
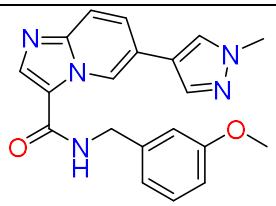
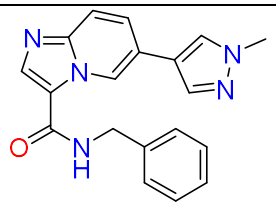
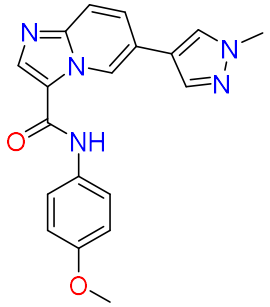
Sector 19, Kamla Nehru Nagar, Ghaziabad 201 002, Uttar Pradesh, India
E-mail: surenderjs@iict.res.in

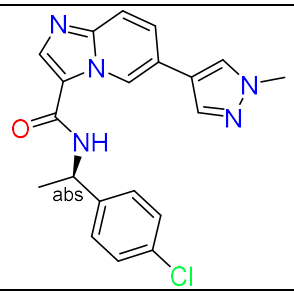
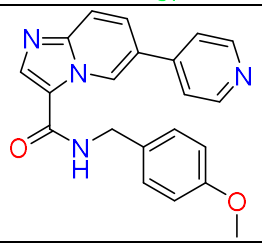
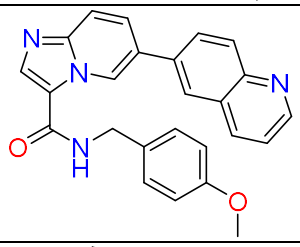
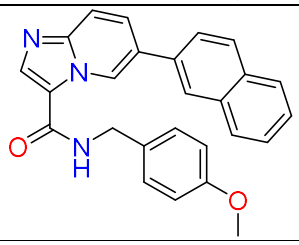
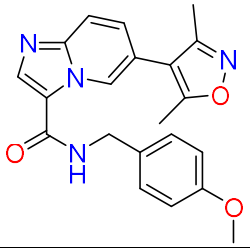
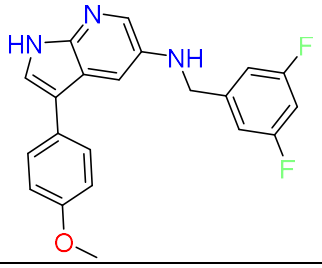
Received 25 September 2024; accepted(revised) 27 December 2024

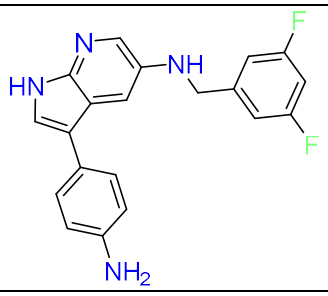
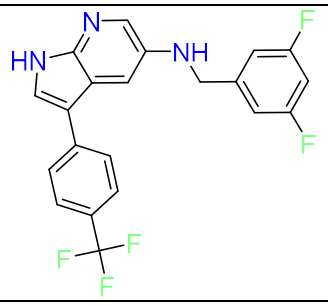
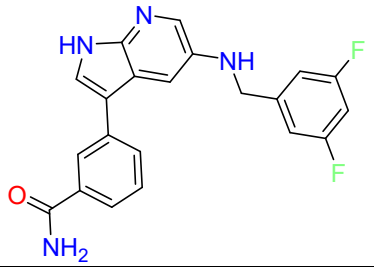
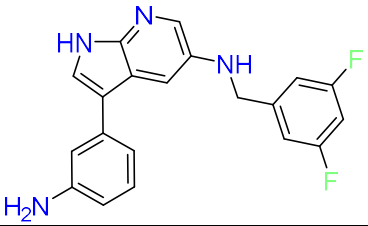
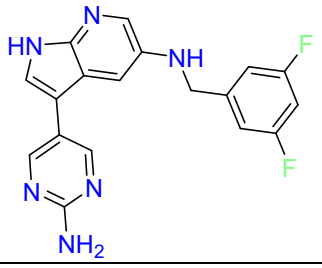
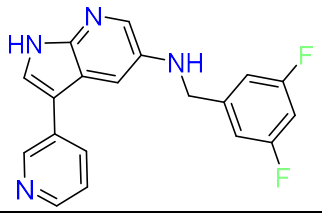
Contents

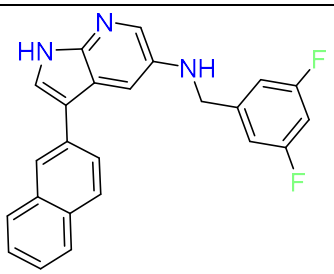
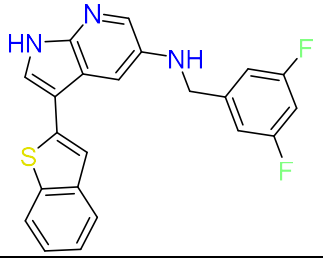
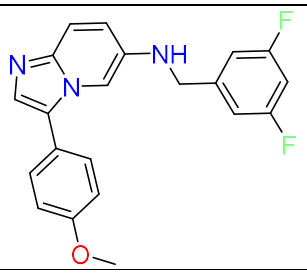
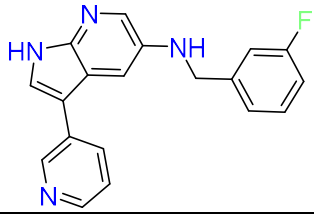
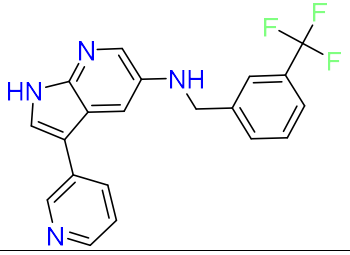
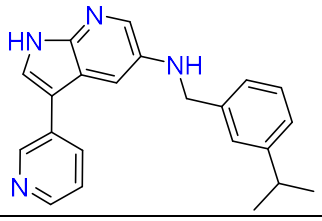
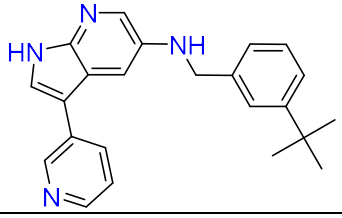
Table S1. Ligands of test set.

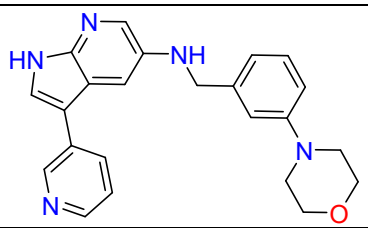
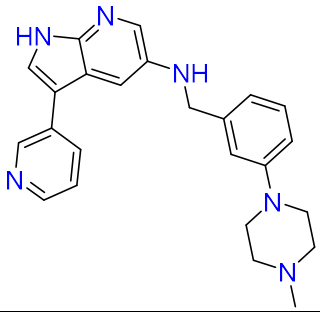
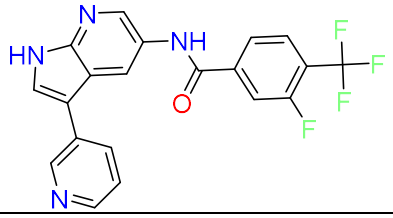
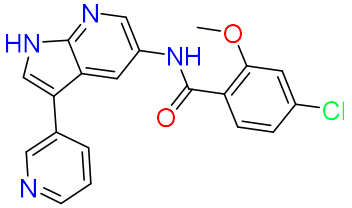
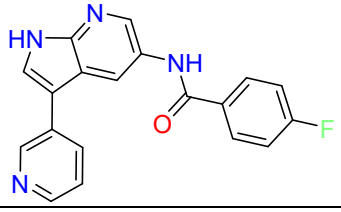
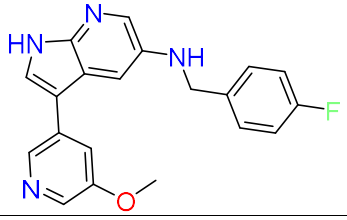
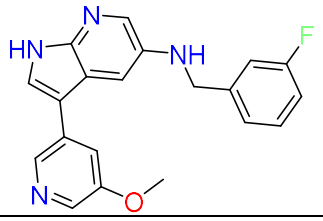
S.No	Structure	IC_{50} (nM)	pIC_{50}
T1		6	8.221
T2		44	7.356
T3		5	8.301
T4		14	7.853
T5		296	6.528
T6		809	6.092

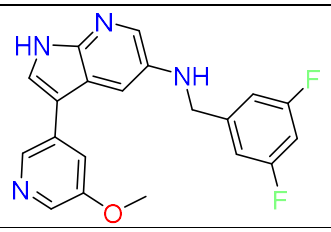
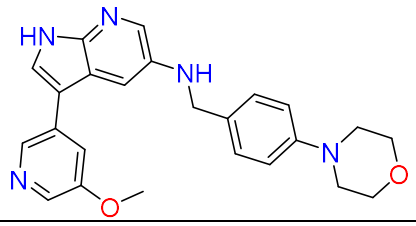
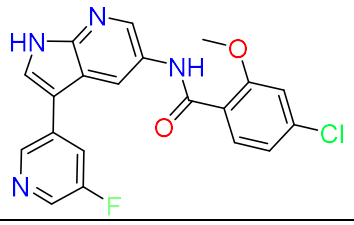
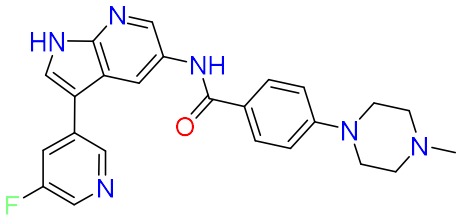
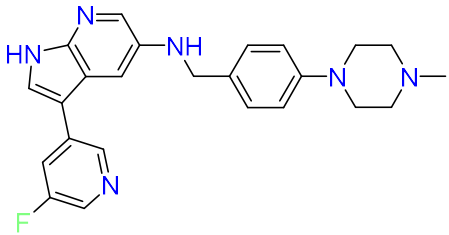
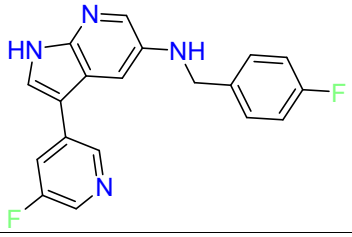
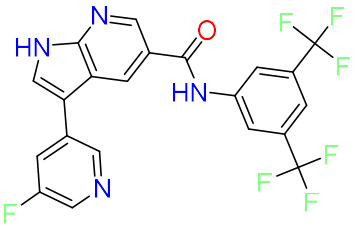
T7		8	8.096
T8		6	8.221
T9		14	7.853
T10		11	7.958
T11		12	7.920
T12		7	8.154

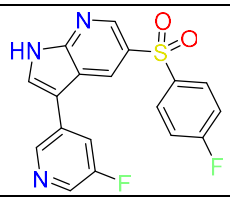
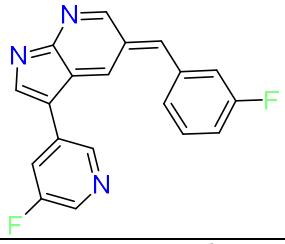
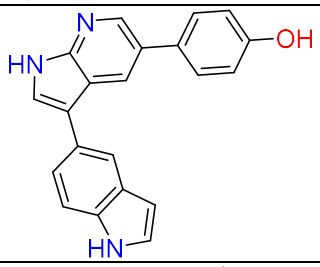
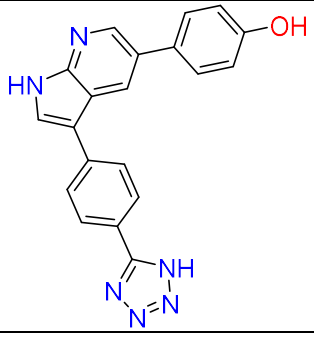
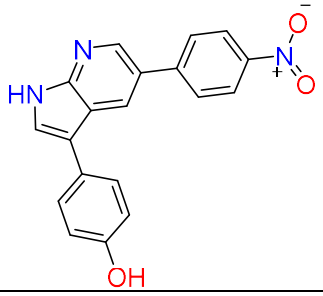
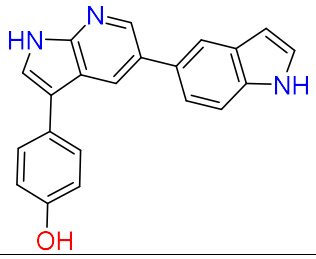
T13		399	6.399
T14		13	7.886
T15		4	8.397
T16		10000	5
T17		10000	5
T18		15	7.823

T19		89	7.050
T20		458	6.339
T21		762	6.118
T22		66	7.180
T23		75	7.124
T24		18	7.744

T25		476	6.322
T26		252	6.598
T27		569	6.244
T28		15	7.823
T29		45	7.346
T30		23	7.638
T31		45	7.346

T32		70	7.154
T33		13	7.886
T34		89	7.050
T35		25	7.602
T36		31	7.508
T37		15	7.823
T38		17	7.769

T39		16	7.795
T40		30	7.522
T41		66	7.180
T42		5	8.301
T43		42	7.376
T44		64	7.193
T45		1867	5.728

T46		667	6.175
T47		456	6.341
T48		117	6.931
T49		85	7.070
T50		187	6.728
T51		26	7.585

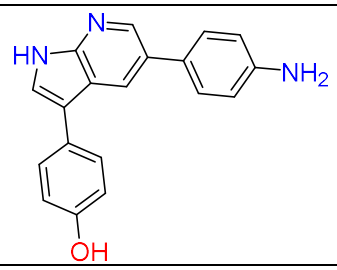
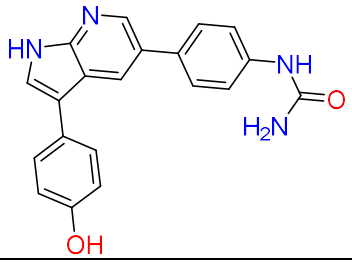
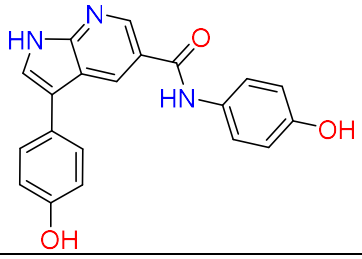
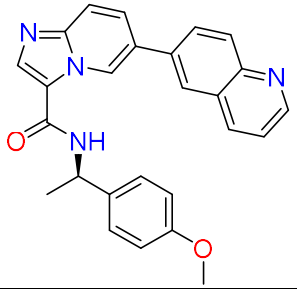
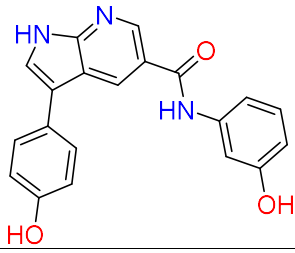
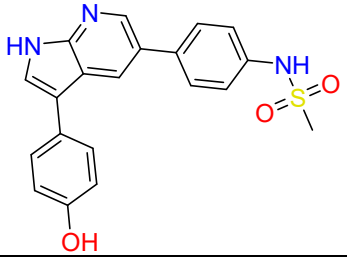
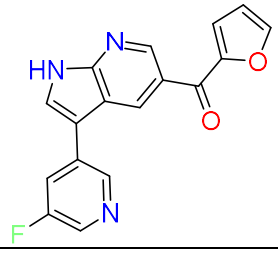
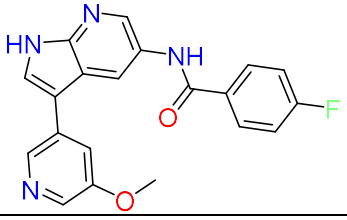
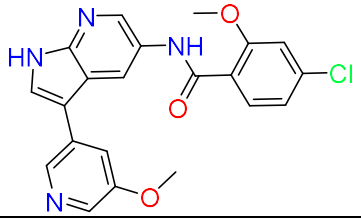
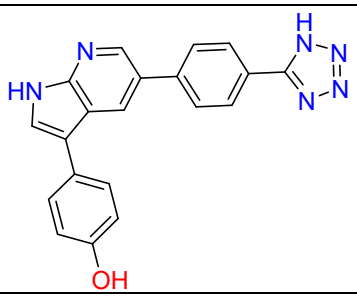
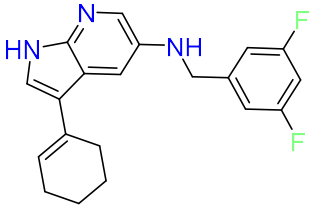
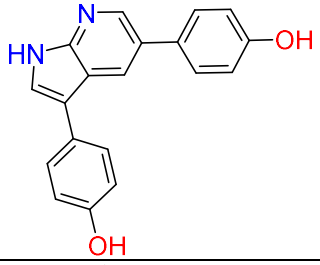
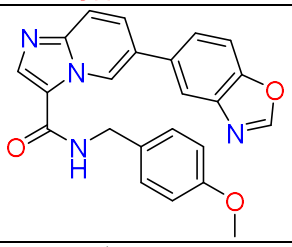
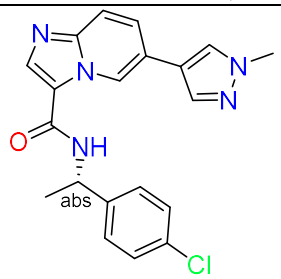
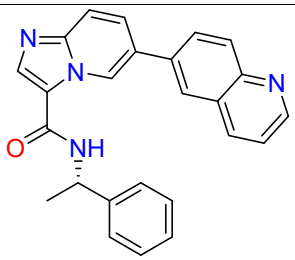
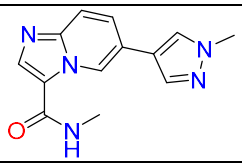
T52		34	7.468
T53		9.6	8.017
T54		12	7.920

Table S2. Ligands of training set.

S.no	Structure	IC_{50} (nM)	pIC_{50}
V1		6	8.221
V2		34	7.468
V3		44	7.356
V4		389	6.410
V5		24	7.619
V6		5	8.301

V7		70	7.154
V8		324	6.489
V9		84	7.318
V10		122	6.913
V11		11	7.958
V12		52	7.283
V13		23	7.638

V14		8	8.096
V15		35	7.455

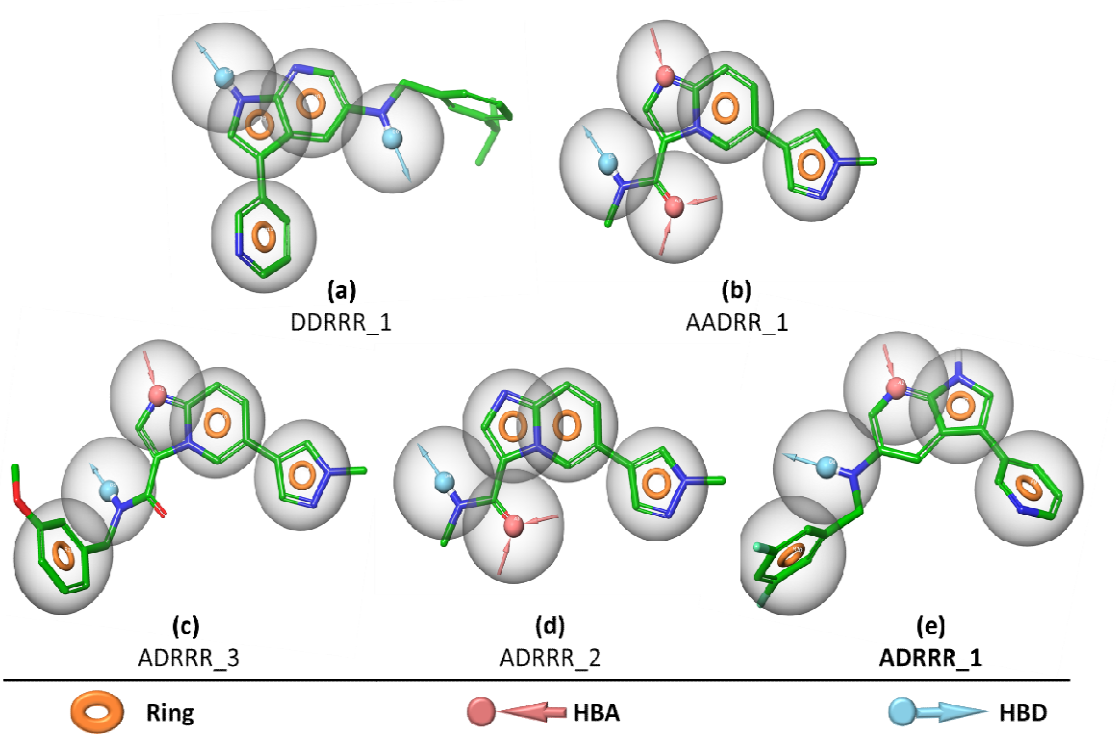


Figure S1. Generated pharmacophore models with their respective reference ligands.

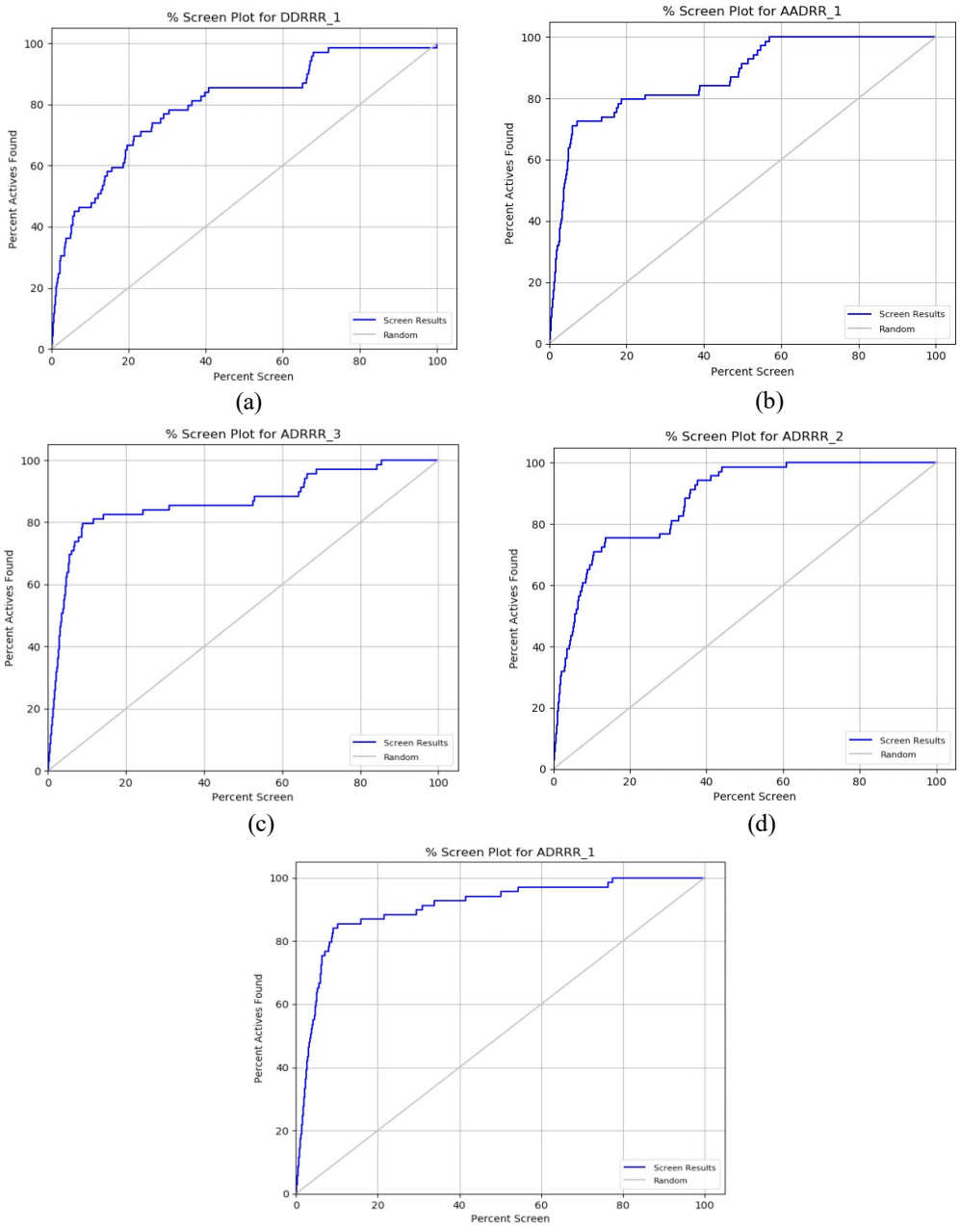


Figure S2. % Screen plots of all Pharmacophore models.

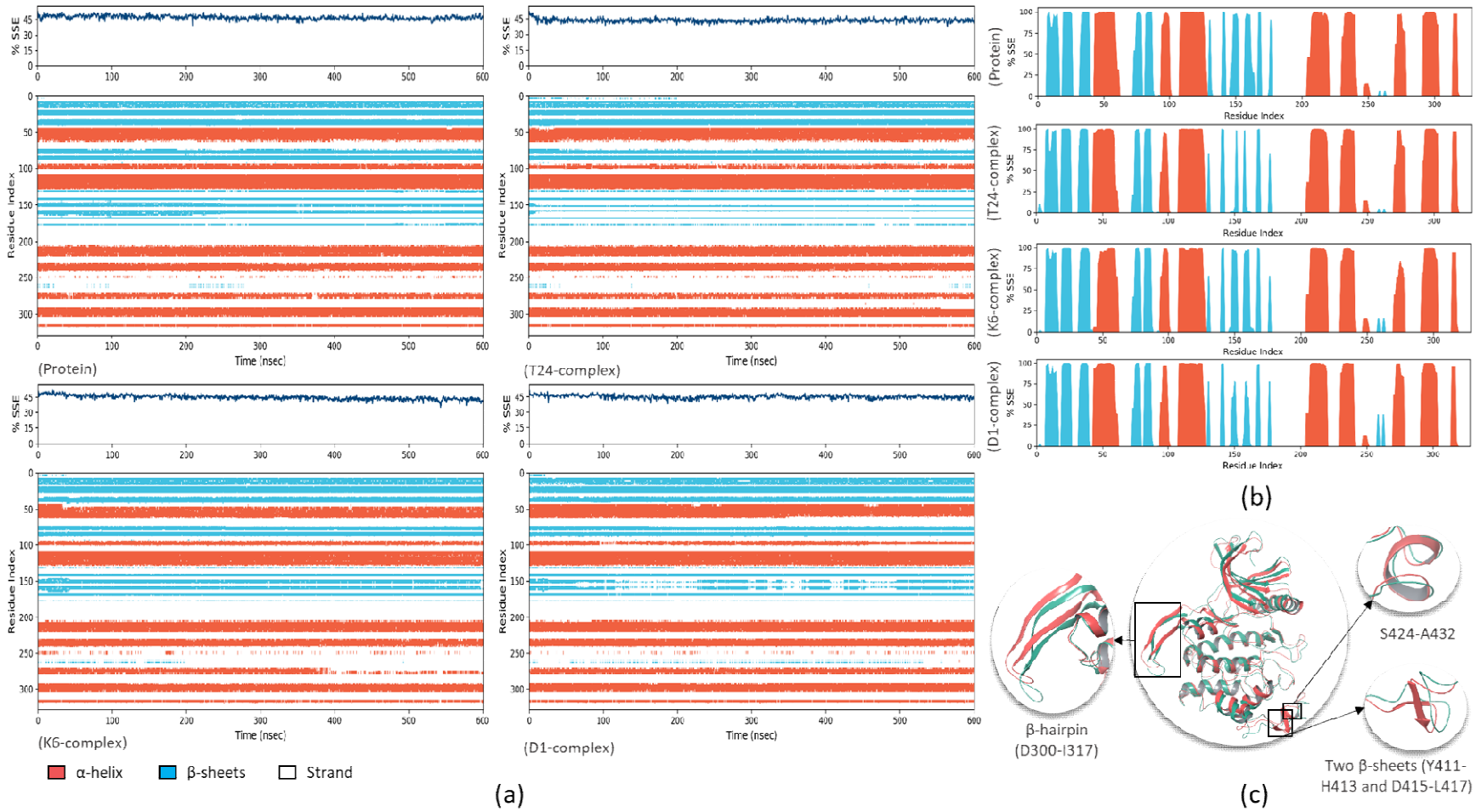


Figure S3. a) Timeline graph of protein secondary structure elements (SSE); b) Histogram of protein %SSE; and c) Major dynamic secondary structures of CLK1.

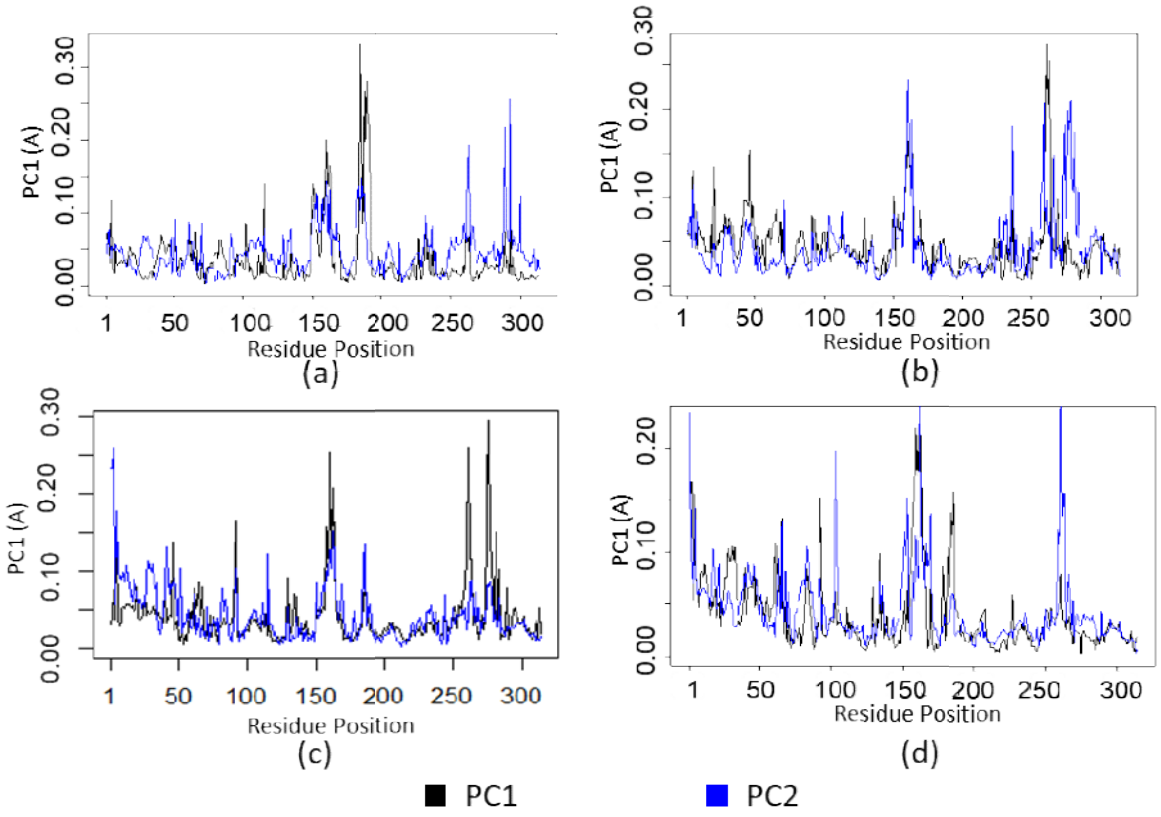


Figure S4. Principle components with respect to residues of a) Apo; b) T24-complex; c) K6-complex; and d) D1-complex, respectively.

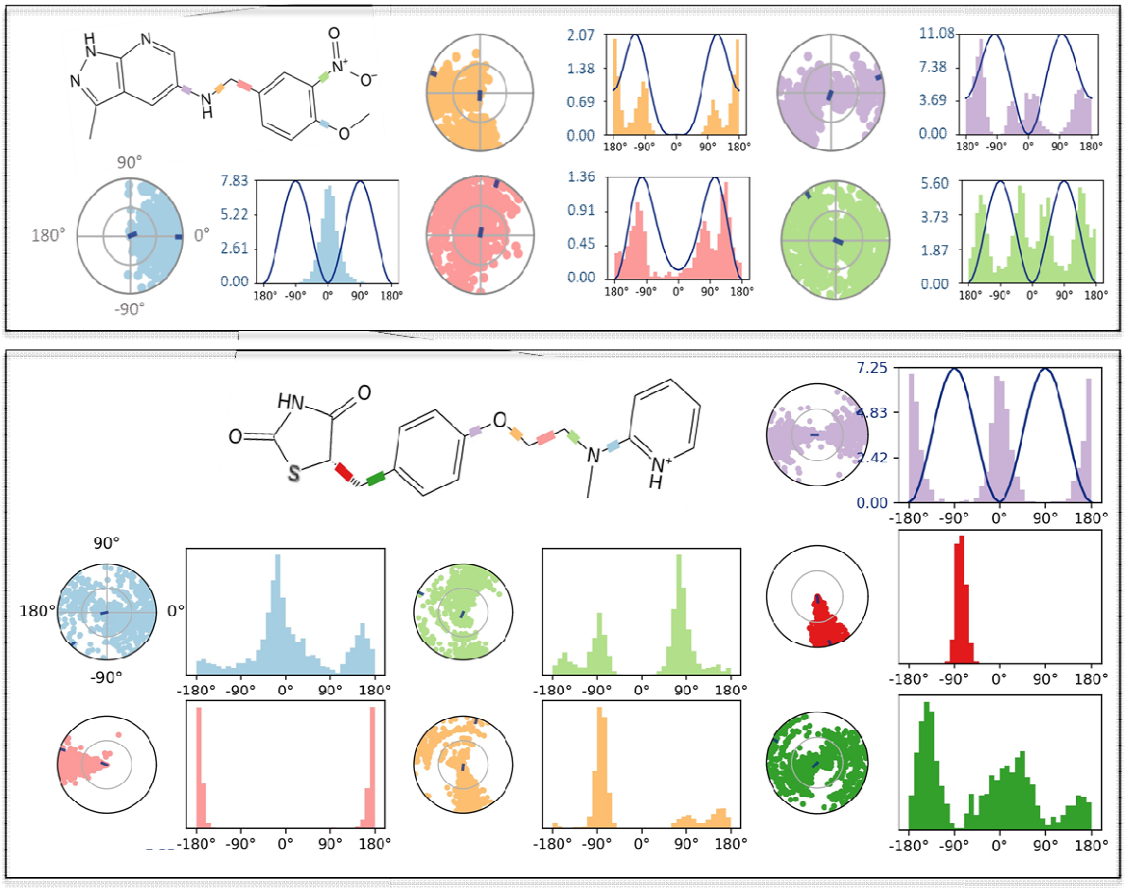


Figure S5. Ligand torsions of **K6** and **D1**.

Table S3. Pharmacokinetic properties of selected ligands obtained from PreADMET and ProTox-II free online servers.

	Property	Value	
		K6	**D1**
ADME properties	Mol Wt	313.315	357.43
	HBA	20	26
	HBD	2	1
	No of rotatable bonds	5	7
	TPSA	108.65	96.83
	logP	3.39	2.82
	BBB	0.0548	0.0123
	Caco2	2.335	28.616
	CYP_2C19 inhibition	-	-
	CYP_2C9 inhibition	Inhibitor	Inhibitor
	CYP_2D6 inhibition	-	-
	CYP_2D6 substrate	-	-
	CYP_3A4 inhibition	-	-
	CYP_3A4 substrate	Weakly	Weakly
	HIA	86.629	97.449
	MDCK	0.0641	2.076
	Pgp_inhibition	Inhibitor	Inhibitor
	Plasma protein binding	79.969	91.236
	Skin permeability	-4.4805	-3.443
Toxicity	Toxicity class	4	4
	LD_{50}	1600 mg/kg	1000 mg/kg
	Toxicity target	Immunotoxicity, Mutagenicity	Immunotoxicity, MMP
	Carcino_Mouse	negative	Positive
	Carcino_Rat	negative	negative
	hERG_inhibition	medium_risk	High risk
	TA100_10RLI	positive	negative
	TA100_NA	positive	negative
	TA1535_10RLI	negative	positive
	TA1535_NA	positive	negative
Druglikeness	CMC like rule	No violation	No violation
	Lead like rule	No violation	2 violations (Mol.Wt, lopP value)
	MDDR like rule	1 violation (5 rotatable bonds only)	No violation
	Rule of five	No violation	No violation
	WDI like rule	No violation	1 violation (Kier-α3)

*mutagenicity in the four selected strains of *Salmonella typhimurium:* TA100_10RLI, TA100_NA, TA1535_10RLI, and TA1335_NA,

Table S4. Pharmacokinetic properties of selected ligands obtained from QikProp

	Property	Value	
		K6	**D1**
ADME properties	Mol Wt	313.315	357.426
	HBA	4.75	5.25
	HBD	2	1
	No of rotatable bonds	5	7
	PSA	108.903	94.075
	logPo/w	1.935	3.44
	logS	-3.721	-4.639
	logBB	-1.604	-1.121
	Caco2	107.503	476.623
	% Human oral absorption	74.636	95.021
	MDCK	44.405	335.723
	Plasma protein binding	0.09	0.277
	Skin permeability	-4.358	-2.444
	hERG_inhibition	-4.469	-5.581
	CNS	-2	-2

Table S5. Top targets for predicted by Swiss Target prediction server

Identified targets for K6		
Target	**Common name**	**Uniprot ID**
Carbonic anhydrase II	CA2	P00918
Carbonic anhydrase VII	CA7	P43166
Carbonic anhydrase I	CA1	P00915
Carbonic anhydrase VI	CA6	P23280
Carbonic anhydrase XII	CA12	O43570
Carbonic anhydrase XIV	CA14	Q9ULX7
Carbonic anhydrase IX	CA9	Q16790
Carbonic anhydrase IV	CA4	P22748
Carbonic anhydrase XIII	CA13	Q8N1Q1
Carbonic anhydrase VA	CA5A	P35218
Inosine-5'-monophosphate dehydrogenase 2	IMPDH2	P12268
Sodium channel protein type IX alpha subunit	SCN9A	Q15858
Rho-associated protein kinase 1	ROCK1	Q13464
Receptor-type tyrosine-protein phosphatase beta	PTPRB	P23467
MAP kinase ERK2	MAPK1	P28482
DNA-dependent protein kinase	PRKDC	P78527
Tyrosine-protein kinase HCK	HCK	P08631
PI3-kinase p110-gamma subunit	PIK3CG	P48736
PI4-kinase beta subunit	PI4KB	Q9UBF8
Anandamide amidohydrolase	FAAH	O00519
Mineralocorticoid receptor	NR3C2	P08235
Type-1 angiotensin II receptor	AGTR1	P30556
Prostanoid EP3 receptor	PTGER3	P43115
Fructose-1,6-bisphosphatase	FBP1	P09467
G protein-coupled receptor 44	PTGDR2	Q9Y5Y4
Inhibitor of nuclear factor kappa B kinase beta subunit	IKBKB	O14920
Ephrin receptor	EPHB4	P54760
Tyrosyl-DNA phosphodiesterase 2	TDP2	O95551
Butyrylcholinesterase	BCHE	P06276
Acetylcholinesterase	ACHE	P22303
Adenosine A2b receptor	ADORA2B	P29275
Multidrug resistance-associated protein 1	ABCC1	P33527
Matrix metalloproteinase 2	MMP2	P08253
Aurora kinase B/Inner centromere protein	INCENP AURKB	Q9NQS7 Q96GD4
ALK tyrosine kinase receptor	ALK	Q9UM73
Perforin-1	PRF1	P14222

Insulin-like growth factor I receptor	IGF1R	P08069
Dipeptidyl peptidase IV	DPP4	P27487
Eukaryotic translation initiation factor 2-alpha kinase 3	EIF2AK3	Q9NZJ5
Rho-associated protein kinase 2	ROCK2	O75116
Caspase-7	CASP7	P55210
Hexokinase type IV	GCK	P35557
Kinesin-like protein 1	KIF11	P52732
Cyclin-dependent kinase 5	CDK5R1 CDK5	Q15078 Q00535
CDC7/DBF4 (Cell division cycle 7-related protein kinase/Activator of S phase kinase)	DBF4 CDC7	Q9UBU7 O00311
CDC7	CDC7	O00311
Beta-adrenergic receptor kinase 2	GRK3	P35626
Cyclophilin A	PPIA	P62937
Glucagon receptor	GCGR	P47871
Progesterone receptor	PGR	P06401
Complement factor D	CFD	P00746
G-protein coupled receptor kinase 2	GRK2	P25098
G protein-coupled receptor kinase 5	GRK5	P34947
Insulin receptor	INSR	P06213
Histone deacetylase 8	HDAC8	Q9BY41
5'-nucleotidase	NT5E	P21589
Tyrosine-protein kinase SYK	SYK	P43405
Dual specificity protein phosphatase 3	DUSP3	P51452
Metabotropic glutamate receptor 4	GRM4	Q14833
Matrix metalloproteinase 13	MMP13	P45452
Matrix metalloproteinase 9	MMP9	P14780
Matrix metalloproteinase 1	MMP1	P03956
Protein-tyrosine phosphatase 1B	PTPN1	P18031
T-cell protein-tyrosine phosphatase	PTPN2	P17706
Tyrosine-protein kinase Lyn	LYN	P07948
Tyrosine-protein kinase TXK	TXK	P42681
Low molecular weight phosphotyrosine protein phosphatase	ACP1	P24666
Serine/threonine-protein kinase RAF	RAF1	P04049
Interleukin-8 receptor B	CXCR2	P25025
Tyrosine-protein kinase JAK1	JAK1	P23458
Tyrosine-protein kinase TYK2	TYK2	P29597
Monoamine oxidase A	MAOA	P21397
NAD-dependent deacetylase sirtuin 3	SIRT3	Q9NTG7
NAD-dependent deacetylase sirtuin 2	SIRT2	Q8IXJ6
NAD-dependent deacetylase sirtuin 1	SIRT1	Q96EB6
Heat shock factor protein 1	HSF1	Q00613

Hematopoietic prostaglandin D synthase	HPGDS	O60760
Vanilloid receptor	TRPV1	Q8NER1
Mitogen-activated protein kinase kinase kinase 8	MAP3K8	P41279
Thyrotropin-releasing hormone receptor (by homology)	TRHR	P34981
Cyclooxygenase-1	PTGS1	P23219
Cholecystokinin B receptor	CCKBR	P32239
Carnitine O-palmitoyltransferase 1, liver isoform	CPT1A	P50416
Macrophage colony stimulating factor receptor	CSF1R	P07333
Carnitine O-palmitoyltransferase 1, muscle isoform	CPT1B	Q92523
Matrix metalloproteinase 3	MMP3	P08254
Dual specificity protein kinase TTK	TTK	P33981
Ribosomal protein S6 kinase alpha 5	RPS6KA5	O75582
Caspase-1	CASP1	P29466
15-hydroxyprostaglandin dehydrogenase [NAD+]	HPGD	P15428
Epoxide hydrolase 1	EPHX1	P07099
Cyclin-dependent kinase 2/cyclin A	CDK2 CCNA1 CCNA2	P24941 P78396 P20248
Dual-specificity tyrosine-phosphorylation regulated kinase 1A	DYRK1A	Q13627
Thrombin and coagulation factor X	F10	P00742
Protein-glutamine gamma-glutamyltransferase	TGM2	P21980
Coagulation factor XIII	F13A1	P00488
GABA receptor alpha-5 subunit	GABRA5	P31644
Histone deacetylase 3	HDAC3	O15379
Histone deacetylase 6	HDAC6	Q9UBN7
Vascular endothelial growth factor receptor 1	FLT1	P17948
Identified targets for D1		
Peroxisome proliferator-activated receptor gamma	PPARG	P37231
Thromboxane-A synthase	TBXAS1	P24557
Monoamine oxidase B	MAOB	P27338
Carbonic anhydrase II	CA2	P00918
Type-1 angiotensin II receptor (by homology)	AGTR1	P30556
Peroxisome proliferator-activated receptor alpha	PPARA	Q07869
Peroxisome proliferator-activated receptor delta	PPARD	Q03181
Bile salt export pump	ABCB11	O95342
Free fatty acid receptor 1	FFAR1	O14842
15-hydroxyprostaglandin dehydrogenase [NAD+]	HPGD	P15428
Matrix metalloproteinase 8	MMP8	P22894
Phospholipase A2 group IIA	PLA2G2A	P14555
Macrophage colony stimulating factor receptor	CSF1R	P07333
Matrix metalloproteinase 3	MMP3	P08254
Glucocorticoid receptor	NR3C1	P04150

Tyrosine-protein kinase LCK	LCK	P06239
Ileal bile acid transporter	SLC10A2	Q12908
Prostanoid DP receptor	PTGDR	Q13258
Cyclin-dependent kinase 9	CDK9	P50750
Calpain 1	CAPN1 CAPNS1	P07384 P04632
Calpain 2	CAPN2	P17655
MAP kinase p38 alpha	MAPK14	Q16539
Methionyl-tRNA synthetase	MARS	P56192
Dihydroorotate dehydrogenase	DHODH	Q02127
Endothelin receptor ET-A	EDNRA	P25101
Epoxide hydratase	EPHX2	P34913
Hydroxycarboxylic acid receptor 2	HCAR2	Q8TDS4
Lysine-specific demethylase 4C	KDM4C	Q9H3R0
Protein-tyrosine phosphatase 1B	PTPN1	P18031
Cannabinoid receptor 1	CNR1	P21554
Cannabinoid receptor 2	CNR2	P34972
Caspase-3	CASP3	P42574
Receptor-type tyrosine-protein phosphatase gamma	PTPRG	P23470
Intercellular adhesion molecule (ICAM-1), Integrin alpha-L/beta-2	ITGAL ICAM1 ITGB2	P20701 P05362 P05107
Adenosine kinase (by homology)	ADK	P55263
CDGSH iron-sulfur domain-containing protein 1	CISD1	Q9NZ45
Prostanoid EP4 receptor (by homology)	PTGER4	P35408
Intercellular adhesion molecule-1	ICAM1	P05362
Metabotropic glutamate receptor 5	GRM5	P41594
Selectin E	SELE	P16581
Kappa Opioid receptor	OPRK1	P41145
Receptor-type tyrosine-protein phosphatase beta	PTPRB	P23467
Protein-tyrosine phosphatase 1C	PTPN6	P29350
Protein-tyrosine phosphatase G1	PTPN12	Q05209
Leukocyte common antigen	PTPRC	P08575
Receptor-type tyrosine-protein phosphatase F (LAR)	PTPRF	P10586
Prostanoid EP1 receptor	PTGER1	P34995
Chymotrypsin C	CTRC	Q99895
Induced myeloid leukemia cell differentiation protein Mcl-1	MCL1	Q07820
Lymphocyte differentiation antigen CD38	CD38	P28907
Arachidonate 5-lipoxygenase	ALOX5	P09917
Carbonic anhydrase VII	CA7	P43166
Glycogen synthase kinase-3 beta	GSK3B	P49841
Carbonic anhydrase XII	CA12	O43570

6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 4	PFKFB4	Q16877
5-lipoxygenase activating protein	ALOX5AP	P20292
Dual-specificity tyrosine-phosphorylation regulated kinase 1A	DYRK1A	Q13627
Dual specificty protein kinase CLK1 (by homology)	CLK1	P49759
Dual specificity protein kinase CLK2 (by homology)	CLK2	P49760
Dual-specificity tyrosine-phosphorylation regulated kinase 2	DYRK2	Q92630
Dual-specificity tyrosine-phosphorylation regulated kinase 3	DYRK3	O43781
Dual specificity tyrosine-phosphorylation-regulated kinase 1B	DYRK1B	Q9Y463
Epidermal growth factor receptor erbB1	EGFR	P00533
Sigma opioid receptor (by homology)	SIGMAR1	Q99720
Tyrosine-protein kinase SYK	SYK	P43405
Tyrosine-protein kinase ABL	ABL1	P00519
Serum albumin	ALB	P02768
Cathepsin (V and K)	CTSV	O60911
Cathepsin L	CTSL	P07711
Apoptosis regulator Bcl-2	BCL2	P10415
Dual specificity mitogen-activated protein kinase kinase 1	MAP2K1	Q02750
Muscarinic acetylcholine receptor M4	CHRM4	P08173
Alpha-2a adrenergic receptor	ADRA2A	P08913
Alpha-2b adrenergic receptor	ADRA2B	P18089
Dopamine D1 receptor	DRD1	P21728
Muscarinic acetylcholine receptor M2	CHRM2	P08172
Norepinephrine transporter	SLC6A2	P23975
Serotonin transporter	SLC6A4	P31645
Neurokinin 2 receptor	TACR2	P21452
Dopamine D3 receptor	DRD3	P35462
Delta opioid receptor	OPRD1	P41143
Muscarinic acetylcholine receptor M3	CHRM3	P20309
Adenosine A3 receptor	ADORA3	P0DMS8
Matrix metalloproteinase 16	MMP16	P51512
Liver glycogen phosphorylase	PYGL	P06737
Matrix metalloproteinase 9	MMP9	P14780
Matrix metalloproteinase 1	MMP1	P03956
Matrix metalloproteinase 2	MMP2	P08253
Matrix metalloproteinase 14	MMP14	P50281
Sodium channel protein type IX alpha subunit	SCN9A	Q15858
Matrix metalloproteinase 12	MMP12	P39900
Angiotensin II receptor	AGTR2	P50052
Apoptosis regulator Bcl-X	BCL2L1	Q07817
c-Jun N-terminal kinase 1	MAPK8	P45983
Carbonic anhydrase I	CA1	P00915
Serine/threonine-protein kinase mTOR	MTOR	P42345

Fibroblast growth factor receptor 1	FGFR1	P11362
Prostanoid EP2 receptor (by homology)	PTGER2	P43116
Thromboxane A2 receptor	TBXA2R	P21731
Nerve growth factor receptor Trk-A	NTRK1	P04629