The 30 cocaine-addiction related datasets are collected from ChEMDL database (https://www.ebi.ac.uk/chembl/), which involve 30 cocaine-addiction protein targets. The labels are binding affinities to these targets.

Dataset	Target ChEMBLID	Dataset size	Binding affinity range (kcal/mol)
AKR1B1	CHEMBL1900	725	[-13.99,-5.80]
AKT1	CHEMBL4282	2859	[-14.04,-5.69]
APP	CHEMBL2487	1141	[-13.84,-5.86]
CACNA1B	CHEMBL4478	352	[-10.91,-5.86]
CSNK2A1	CHEMBL3629	737	[-13.89,-5.85]
DHFR	CHEMBL202	960	[-14.27,-5.49]
DPP4	CHEMBL284	3883	[-14.31,-5.45]
FGFR1	CHEMBL3650	2060	[-13.63,-5.45]
FYN	CHEMBL1841	459	[-12.95,-5.45]
GBA	CHEMBL2179	398	[-12.81,-5.47]
HDAC1	CHEMBL325	4608	[-14.18,-5.85]
HTR1A	CHEMBL214	4342	[-14.79,-5.73]
HTR2A	CHEMBL224	4307	[-15.00,-5.45]
LCK	CHEMBL258	1855	[-15.41,-5.45]
LYN	CHEMBL3905	468	[-13.22,-6.33]
MAPKAPK2	CHEMBL2208	789	[-12.68,-5.92]
MDM2	CHEMBL5023	1745	[-14.59,-5.94]
MINK1	CHEMBL5518	364	[-13.13,-6.41]
NET	CHEMBL222	2981	[-14.63,-5.47]
NTRK1	CHEMBL2815	2783	[-14.04,-5.45]
NTRK2	CHEMBL4898	566	[-14.04,-5.45]
NTRK3	CHEMBL5608	355	[-13.58,-5.45]
PLG	CHEMBL1801	937	[-15.00,-5.45]
PRKCD	CHEMBL2996	792	[-14.04,-6.13]
SERT	CHEMBL228	4327	[-15.00,-5.64]
SLC5A2	CHEMBL3884	1231	[-13.39,-6.01]
STAT3	CHEMBL4026	670	[-11.12,-5.89]
SYK	CHEMBL2599	3175	[-15.41,-5.94]
TDO2	CHEMBL2140	291	[-10.91,-5.97]
VCP	CHEMBL1075145	323	[-11.53,-6.27]

References

[1] Machine learning analysis of cocaine addiction informed by DAT, SERT, and NET-based interactome networks. *arXiv preprint arXiv:2201.00114*, 2022.