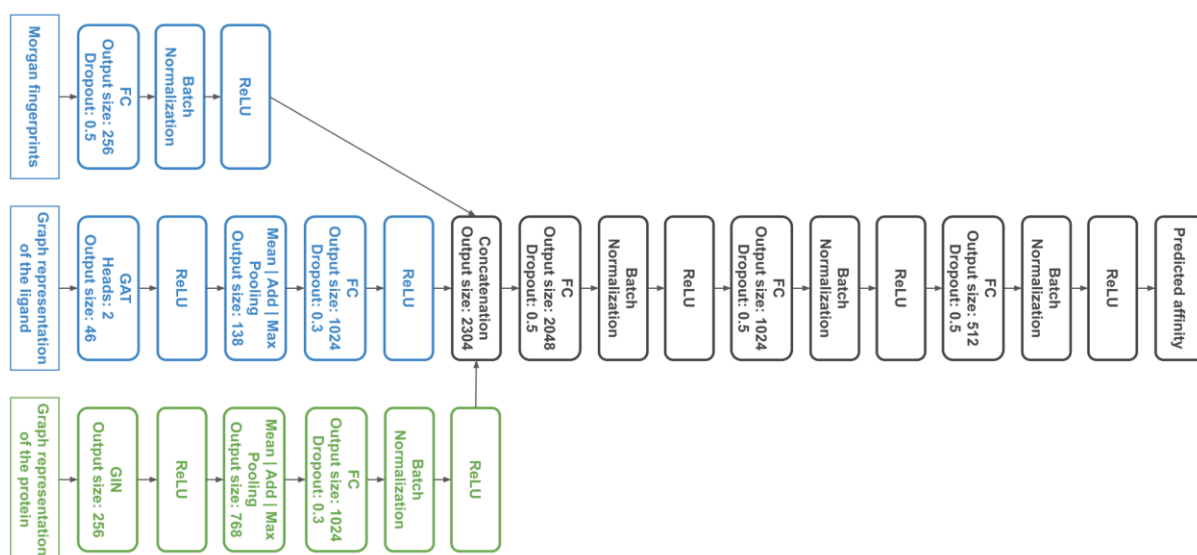


# 3DProtDTA: the deep learning model for drug-target affinity prediction based on the residue-level protein graphs

## Supplementary Information

Taras Voitsitskiy, Roman Stratiichuk, Ihor Koleiev, Leonid Popryho, Zakhar Ostrovsky, Pavlo Henitsoi, Ivan Khropachev, Volodymyr Vozniak, Roman Zhytar, Diana Nechepurenko, Semen Yesylevskyy, Alan Nafiev, Serhii Starosyla



**Figure S1.** 3DProtDTA model architecture tuned by Tree-structured Parzen Estimator algorithm.

dataset	experiment	case	fold	mse	ci	rm2
Davis	ligand features	Graph+Morgan	1	0.178	0.910	0.759
Davis	ligand features	Graph+Morgan	2	0.166	0.920	0.769
Davis	ligand features	Graph+Morgan	3	0.172	0.918	0.725
Davis	ligand features	Graph+Morgan	4	0.165	0.923	0.730
Davis	ligand features	Graph+Morgan	5	0.182	0.913	0.733

Davis	ligand features	Graph	1	0.177	0.912	0.765
Davis	ligand features	Graph	2	0.173	0.918	0.759
Davis	ligand features	Graph	3	0.177	0.920	0.713
Davis	ligand features	Graph	4	0.161	0.921	0.758
Davis	ligand features	Graph	5	0.182	0.912	0.744
Davis	ligand features	Morgan	1	0.189	0.908	0.744
Davis	ligand features	Morgan	2	0.176	0.907	0.758
Davis	ligand features	Morgan	3	0.181	0.919	0.684
Davis	ligand features	Morgan	4	0.173	0.918	0.724
Davis	ligand features	Morgan	5	0.186	0.911	0.719
Davis	graph pooling	mean+add+max	1	0.178	0.910	0.759
Davis	graph pooling	mean+add+max	2	0.166	0.920	0.769
Davis	graph pooling	mean+add+max	3	0.172	0.918	0.725
Davis	graph pooling	mean+add+max	4	0.165	0.923	0.730
Davis	graph pooling	mean+add+max	5	0.182	0.913	0.733
Davis	graph pooling	mean	1	0.173	0.912	0.752
Davis	graph pooling	mean	2	0.168	0.919	0.754
Davis	graph pooling	mean	3	0.177	0.921	0.695
Davis	graph pooling	mean	4	0.163	0.920	0.744
Davis	graph pooling	mean	5	0.181	0.914	0.725
Davis	graph pooling	add	1	0.173	0.913	0.753
Davis	graph pooling	add	2	0.166	0.921	0.769
Davis	graph pooling	add	3	0.175	0.918	0.738
Davis	graph pooling	add	4	0.162	0.924	0.752

Davis	graph pooling	add	5	0.181	0.915	0.763
Davis	graph pooling	max	1	0.185	0.912	0.757
Davis	graph pooling	max	2	0.174	0.925	0.757
Davis	graph pooling	max	3	0.177	0.917	0.732
Davis	graph pooling	max	4	0.175	0.920	0.745
Davis	graph pooling	max	5	0.195	0.914	0.738
Davis	ligand GNN	GAT	1	0.178	0.910	0.759
Davis	ligand GNN	GAT	2	0.166	0.920	0.769
Davis	ligand GNN	GAT	3	0.172	0.918	0.725
Davis	ligand GNN	GAT	4	0.165	0.923	0.730
Davis	ligand GNN	GAT	5	0.182	0.913	0.733
Davis	ligand GNN	GCN	1	0.172	0.913	0.775
Davis	ligand GNN	GCN	2	0.168	0.917	0.766
Davis	ligand GNN	GCN	3	0.182	0.922	0.704
Davis	ligand GNN	GCN	4	0.166	0.924	0.738
Davis	ligand GNN	GCN	5	0.179	0.913	0.743
Davis	ligand GNN	GIN	1	0.170	0.915	0.775
Davis	ligand GNN	GIN	2	0.163	0.916	0.778
Davis	ligand GNN	GIN	3	0.173	0.918	0.724
Davis	ligand GNN	GIN	4	0.167	0.923	0.771
Davis	ligand GNN	GIN	5	0.175	0.914	0.742
Davis	ligand GNN	GINE	1	0.171	0.917	0.768
Davis	ligand GNN	GINE	2	0.165	0.917	0.760
Davis	ligand GNN	GINE	3	0.176	0.918	0.714

Davis	ligand GNN	GINE	4	0.168	0.919	0.764
Davis	ligand GNN	GINE	5	0.179	0.915	0.765
Davis	ligand GNN	GMF	1	0.174	0.917	0.765
Davis	ligand GNN	GMF	2	0.173	0.922	0.736
Davis	ligand GNN	GMF	3	0.178	0.922	0.722
Davis	ligand GNN	GMF	4	0.163	0.921	0.750
Davis	ligand GNN	GMF	5	0.181	0.916	0.762
Davis	protein GNN	GAT	1	0.170	0.914	0.748
Davis	protein GNN	GAT	2	0.166	0.920	0.742
Davis	protein GNN	GAT	3	0.177	0.923	0.705
Davis	protein GNN	GAT	4	0.158	0.921	0.747
Davis	protein GNN	GAT	5	0.188	0.913	0.747
Davis	protein GNN	GCN	1	0.173	0.912	0.752
Davis	protein GNN	GCN	2	0.166	0.915	0.774
Davis	protein GNN	GCN	3	0.172	0.921	0.710
Davis	protein GNN	GCN	4	0.162	0.924	0.739
Davis	protein GNN	GCN	5	0.185	0.914	0.737
Davis	protein GNN	GIN	1	0.178	0.910	0.759
Davis	protein GNN	GIN	2	0.166	0.920	0.769
Davis	protein GNN	GIN	3	0.172	0.918	0.725
Davis	protein GNN	GIN	4	0.165	0.923	0.730
Davis	protein GNN	GIN	5	0.182	0.913	0.733
Davis	protein GNN	GINE	1	0.178	0.913	0.790
Davis	protein GNN	GINE	2	0.169	0.917	0.770

Davis	protein GNN	GINE	3	0.178	0.917	0.727
Davis	protein GNN	GINE	4	0.168	0.925	0.762
Davis	protein GNN	GINE	5	0.185	0.911	0.752
Davis	protein GNN	GMF	1	0.186	0.915	0.736
Davis	protein GNN	GMF	2	0.178	0.918	0.760
Davis	protein GNN	GMF	3	0.182	0.916	0.718
Davis	protein GNN	GMF	4	0.174	0.918	0.760
Davis	protein GNN	GMF	5	0.193	0.911	0.734

**Table S2. Detailed 5-fold cross-validation results for the KIBA dataset.**

dataset	experiment	case	fold	mse	ci	rm2
KIBA	ligand features	Graph+Morgan	1	0.137	0.897	0.805
KIBA	ligand features	Graph+Morgan	2	0.147	0.892	0.789
KIBA	ligand features	Graph+Morgan	3	0.145	0.895	0.796
KIBA	ligand features	Graph+Morgan	4	0.154	0.890	0.772
KIBA	ligand features	Graph+Morgan	5	0.146	0.890	0.777
KIBA	ligand features	Graph	1	0.150	0.886	0.788
KIBA	ligand features	Graph	2	0.162	0.883	0.774
KIBA	ligand features	Graph	3	0.165	0.883	0.764
KIBA	ligand features	Graph	4	0.172	0.876	0.762
KIBA	ligand features	Graph	5	0.159	0.879	0.776
KIBA	ligand features	Morgan	1	0.151	0.891	0.783
KIBA	ligand features	Morgan	2	0.160	0.888	0.771
KIBA	ligand features	Morgan	3	0.158	0.888	0.776
KIBA	ligand features	Morgan	4	0.167	0.881	0.756

KIBA	ligand features	Morgan	5	0.156	0.885	0.764
KIBA	graph pooling	mean+add+max	1	0.137	0.897	0.805
KIBA	graph pooling	mean+add+max	2	0.147	0.892	0.789
KIBA	graph pooling	mean+add+max	3	0.145	0.895	0.796
KIBA	graph pooling	mean+add+max	4	0.154	0.890	0.772
KIBA	graph pooling	mean+add+max	5	0.146	0.890	0.777
KIBA	graph pooling	mean	1	0.143	0.893	0.792
KIBA	graph pooling	mean	2	0.151	0.890	0.776
KIBA	graph pooling	mean	3	0.151	0.892	0.794
KIBA	graph pooling	mean	4	0.160	0.886	0.751
KIBA	graph pooling	mean	5	0.149	0.889	0.795
KIBA	graph pooling	add	1	0.138	0.895	0.795
KIBA	graph pooling	add	2	0.147	0.892	0.793
KIBA	graph pooling	add	3	0.143	0.894	0.778
KIBA	graph pooling	add	4	0.154	0.888	0.748
KIBA	graph pooling	add	5	0.142	0.891	0.797
KIBA	graph pooling	max	1	0.139	0.897	0.804
KIBA	graph pooling	max	2	0.154	0.892	0.767
KIBA	graph pooling	max	3	0.150	0.894	0.796
KIBA	graph pooling	max	4	0.157	0.886	0.777
KIBA	graph pooling	max	5	0.146	0.891	0.787
KIBA	ligand GNN	GAT	1	0.137	0.897	0.805
KIBA	ligand GNN	GAT	2	0.147	0.892	0.789
KIBA	ligand GNN	GAT	3	0.145	0.895	0.796

KIBA	ligand GNN	GAT	4	0.154	0.890	0.772
KIBA	ligand GNN	GAT	5	0.146	0.890	0.777
KIBA	ligand GNN	GCN	1	0.138	0.897	0.796
KIBA	ligand GNN	GCN	2	0.151	0.893	0.790
KIBA	ligand GNN	GCN	3	0.147	0.894	0.800
KIBA	ligand GNN	GCN	4	0.159	0.887	0.784
KIBA	ligand GNN	GCN	5	0.143	0.890	0.794
KIBA	ligand GNN	GIN	1	0.140	0.897	0.796
KIBA	ligand GNN	GIN	2	0.148	0.891	0.793
KIBA	ligand GNN	GIN	3	0.149	0.895	0.791
KIBA	ligand GNN	GIN	4	0.158	0.887	0.774
KIBA	ligand GNN	GIN	5	0.145	0.889	0.788
KIBA	ligand GNN	GINE	1	0.138	0.897	0.801
KIBA	ligand GNN	GINE	2	0.146	0.895	0.786
KIBA	ligand GNN	GINE	3	0.147	0.895	0.792
KIBA	ligand GNN	GINE	4	0.156	0.889	0.783
KIBA	ligand GNN	GINE	5	0.145	0.891	0.797
KIBA	ligand GNN	GMF	1	0.138	0.897	0.784
KIBA	ligand GNN	GMF	2	0.149	0.892	0.764
KIBA	ligand GNN	GMF	3	0.145	0.896	0.796
KIBA	ligand GNN	GMF	4	0.155	0.887	0.774
KIBA	ligand GNN	GMF	5	0.143	0.892	0.790
KIBA	protein GNN	GAT	1	0.144	0.896	0.794
KIBA	protein GNN	GAT	2	0.152	0.892	0.772

KIBA	protein GNN	GAT	3	0.147	0.892	0.792
KIBA	protein GNN	GAT	4	0.154	0.888	0.769
KIBA	protein GNN	GAT	5	0.144	0.891	0.798
KIBA	protein GNN	GCN	1	0.140	0.897	0.790
KIBA	protein GNN	GCN	2	0.150	0.893	0.783
KIBA	protein GNN	GCN	3	0.150	0.893	0.789
KIBA	protein GNN	GCN	4	0.156	0.888	0.783
KIBA	protein GNN	GCN	5	0.148	0.891	0.790
KIBA	protein GNN	GIN	1	0.137	0.897	0.805
KIBA	protein GNN	GIN	2	0.147	0.892	0.789
KIBA	protein GNN	GIN	3	0.145	0.895	0.796
KIBA	protein GNN	GIN	4	0.154	0.890	0.772
KIBA	protein GNN	GIN	5	0.146	0.890	0.777
KIBA	protein GNN	GINE	1	0.138	0.898	0.799
KIBA	protein GNN	GINE	2	0.150	0.893	0.789
KIBA	protein GNN	GINE	3	0.145	0.894	0.797
KIBA	protein GNN	GINE	4	0.155	0.888	0.776
KIBA	protein GNN	GINE	5	0.143	0.891	0.798
KIBA	protein GNN	GMF	1	0.140	0.896	0.794
KIBA	protein GNN	GMF	2	0.150	0.892	0.780
KIBA	protein GNN	GMF	3	0.149	0.891	0.782
KIBA	protein GNN	GMF	4	0.157	0.887	0.761
KIBA	protein GNN	GMF	5	0.146	0.890	0.790



**Table S3. 5-fold mean and standard deviation for each experiment, Davis dataset.**

dataset	experiment	case	mse_mean	ci_mean	rm2_mean	mse_std	ci_std	rm2_std
Davis	ligand features	Graph+Morgan	0.173	0.917	0.743	0.007	0.005	0.020
Davis	ligand features	Graph	0.174	0.917	0.748	0.008	0.004	0.021
Davis	ligand features	Morgan	0.181	0.913	0.726	0.007	0.006	0.028
Davis	graph pooling	mean+add+max	0.173	0.917	0.743	0.007	0.005	0.020
Davis	graph pooling	mean	0.172	0.917	0.734	0.007	0.004	0.024
Davis	graph pooling	add	0.171	0.918	0.755	0.007	0.004	0.012
Davis	graph pooling	max	0.181	0.918	0.746	0.009	0.005	0.011
Davis	ligand GNN	GAT	0.173	0.917	0.743	0.007	0.005	0.020
Davis	ligand GNN	GCN	0.174	0.918	0.745	0.007	0.005	0.028
Davis	ligand GNN	GIN	0.170	0.917	0.758	0.005	0.004	0.024
Davis	ligand GNN	GINE	0.172	0.917	0.754	0.006	0.001	0.023
Davis	ligand GNN	GMF	0.174	0.920	0.747	0.007	0.003	0.018
Davis	protein GNN	GAT	0.172	0.918	0.738	0.012	0.004	0.018
Davis	protein GNN	GCN	0.172	0.917	0.742	0.009	0.005	0.023
Davis	protein GNN	GIN	0.173	0.917	0.743	0.007	0.005	0.020
Davis	protein GNN	GINE	0.176	0.917	0.760	0.007	0.005	0.023
Davis	protein GNN	GMF	0.183	0.916	0.742	0.007	0.003	0.018

**Table S4. 5-fold mean and standard deviation for each experiment, KIBA dataset.**

dataset	experiment	case	mse_mean	ci_mean	rm2_mean	mse_std	ci_std	rm2_std
KIBA	ligand features	Graph+Morgan	0.146	0.893	0.788	0.006	0.003	0.014
KIBA	ligand features	Graph	0.161	0.881	0.773	0.008	0.004	0.011
KIBA	ligand features	Morgan	0.158	0.887	0.770	0.006	0.004	0.010
KIBA	graph pooling	mean+add+max	0.146	0.893	0.788	0.006	0.003	0.014
KIBA	graph pooling	mean	0.151	0.890	0.781	0.006	0.003	0.019
KIBA	graph pooling	add	0.145	0.892	0.782	0.006	0.003	0.020
KIBA	graph pooling	max	0.149	0.892	0.786	0.007	0.004	0.015
KIBA	ligand GNN	GAT	0.146	0.893	0.788	0.006	0.003	0.014
KIBA	ligand GNN	GCN	0.148	0.892	0.793	0.008	0.004	0.006
KIBA	ligand GNN	GIN	0.148	0.892	0.789	0.007	0.004	0.009
KIBA	ligand GNN	GINE	0.146	0.894	0.792	0.006	0.003	0.007

KIBA	ligand GNN	GMF	0.146	0.893	0.782	0.006	0.004	0.013
KIBA	protein GNN	GAT	0.148	0.892	0.785	0.005	0.003	0.014
KIBA	protein GNN	GCN	0.149	0.893	0.787	0.006	0.003	0.004
KIBA	protein GNN	GIN	0.146	0.893	0.788	0.006	0.003	0.014
KIBA	protein GNN	GINE	0.146	0.893	0.792	0.006	0.004	0.010
KIBA	protein GNN	GMF	0.149	0.891	0.781	0.006	0.003	0.013

**Table S5. 5-fold mean for each experiment, the average for the KIBA and Davis datasets.**

experiment	case	mse	ci	rm2
ligand features	Graph+Morgan	0.159	0.905	0.765
ligand features	Graph	0.168	0.899	0.760
ligand features	Morgan	0.170	0.900	0.748
graph pooling	mean+add+max	0.159	0.905	0.765
graph pooling	mean	0.162	0.903	0.758
graph pooling	add	0.158	0.905	0.769
graph pooling	max	0.165	0.905	0.766
ligand GNN	GAT	0.159	0.905	0.765
ligand GNN	GCN	0.161	0.905	0.769
ligand GNN	GIN	0.159	0.905	0.773
ligand GNN	GINE	0.159	0.906	0.773
ligand GNN	GMF	0.160	0.906	0.764
protein GNN	GAT	0.160	0.905	0.762
protein GNN	GCN	0.160	0.905	0.765
protein GNN	GIN	0.159	0.905	0.765
protein GNN	GINE	0.161	0.905	0.776
protein GNN	GMF	0.166	0.903	0.761

**Table S6. The number of proteins annotated by the top 20 InterPro entries for the Davis dataset.**

InterPro ID	InterPro Name	Number of Proteins
IPR011009	Kinase-like_dom_sf	375
IPR000719	Prot_kinase_dom	361
IPR017441	Protein_kinase_ATP_BS	323
IPR008271	Ser/Thr_kinase_AS	247
IPR001245	Ser-Thr/Tyr_kinase_cat_dom	106

IPR008266	Tyr_kinase_AS	80
IPR020635	Tyr_kinase_cat_dom	79
IPR013783	Ig-like_fold	41
IPR000961	AGC-kinase_C	38
IPR036860	SH2_dom_sf	28
IPR000980	SH2	28
IPR001452	SH3_domain	25
IPR036028	SH3-like_dom_sf	24
IPR003961	FN3_dom	23
IPR036116	FN3_sf	23
IPR017892	Pkinase_C	22
IPR036179	Ig-like_dom_sf	22
IPR011993	PH-like_dom_sf	21
IPR007110	Ig-like_dom	21
IPR003599	Ig_sub	19

**Table S7. The number of proteins annotated by the top 20 InterPro entries for the KIBA dataset.**

<b>InterPro ID</b>	<b>InterPro Name</b>	<b>Number of Proteins</b>
IPR011009	Kinase-like_dom_sf	221
IPR000719	Prot_kinase_dom	213
IPR017441	Protein_kinase_ATP_BS	202
IPR008271	Ser/Thr_kinase_AS	136
IPR001245	Ser-Thr/Tyr_kinase_cat_dom	72
IPR020635	Tyr_kinase_cat_dom	62
IPR008266	Tyr_kinase_AS	62
IPR000961	AGC-kinase_C	32
IPR013783	Ig-like_fold	29
IPR036860	SH2_dom_sf	26
IPR000980	SH2	26
IPR017892	Pkinase_C	23
IPR036028	SH3-like_dom_sf	21
IPR001452	SH3_domain	21
IPR007110	Ig-like_dom	19
IPR036179	Ig-like_dom_sf	19
IPR011993	PH-like_dom_sf	17
IPR046349	C1-like_sf	17
IPR002219	PE/DAG-bd	17
IPR003599	Ig_sub	16

**Table S8. Protein residue-level graphs similarity between AlphaFold and experimental structures.**

<b>UniProt ID</b>	<b>PDB ID</b>	<b>n nodes, AlphaFold</b>	<b>n nodes, experimental</b>	<b>n intersecting nodes</b>	<b>n edges, AlphaFold</b>	<b>n edges, experimental</b>	<b>n intersecting edges</b>
Q9P1W9	4x7q	255	240	237	1416	1360	1282
O95819	4zp5	265	253	253	1452	1544	1360
P06241	2dq7	254	262	252	1466	1452	1385
P49760	6fyl	317	340	317	1866	2002	1801
Q16539	4f9y	285	275	275	1622	1596	1514
Q9UHD2	4euu	302	298	298	1618	1658	1572
P05771	2i0e	330	326	323	1878	1794	1650
Q15118	2q8g	231	219	219	1214	1210	1166