

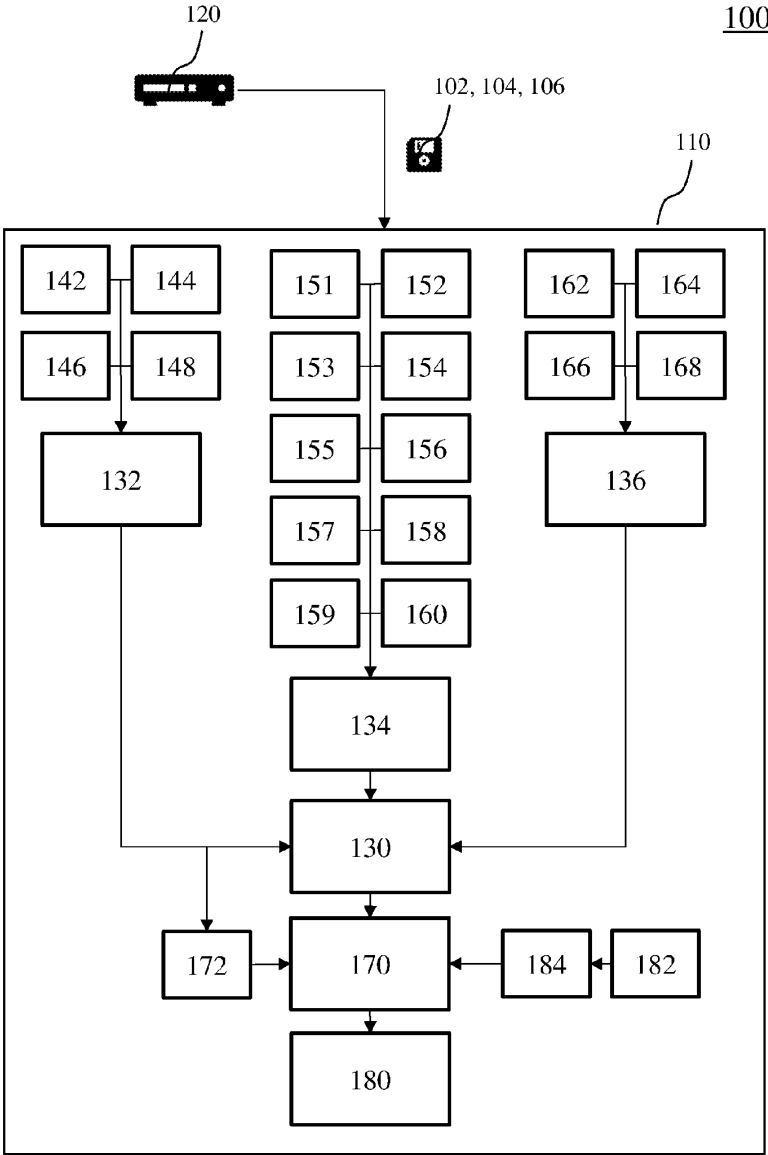


FIG. 1

200

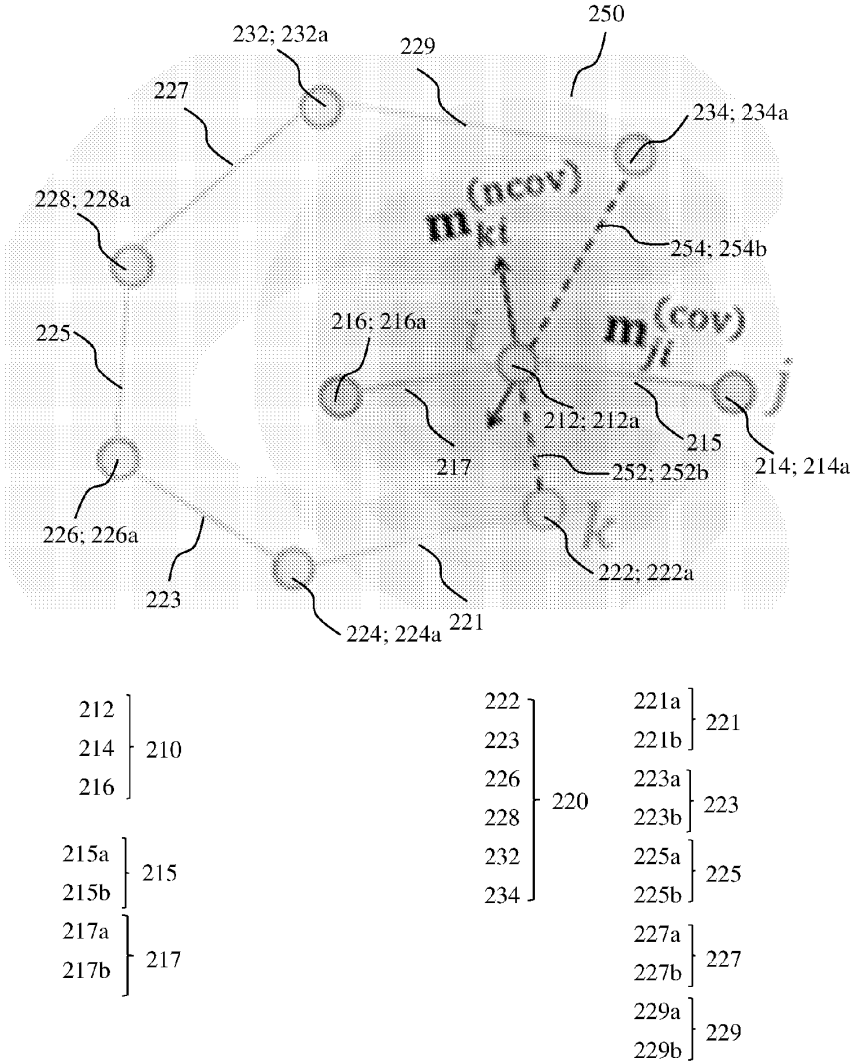


FIG. 2

Atom Features		Types	Sizes	Descriptions
212a, 214a, 216a 222a, 224a, 226a, 228a, 232a, 234a	Atom type	One-hot	10	Heavy atom type [C, N, O, S, F, P, Cl, Br, I, others]
	Degree	One-hot	7	Number of covalent bonds [0, 1, 2, 3, 4, 5, 6]
	Implicit valence	One-hot	7	Implicit valence of the atom [0, 1, 2, 3, 4, 5, 6]
	Hybridization	One-hot	5	[sp, sp ² , sp ³ , sp ^{3d} , sp ^{3d2}]
	Aromatic	Binary	1	Whether the atom is part of an aromatic system
Hydrogens		Integer	5	Number of connected hydrogens [0, 1, 2, 3, 4]
Intra-Molecular Bond Features		Types	Sizes	Descriptions
215a, 217a 221a, 223a, 225a, 227a, 229a	bond type	one-hot	4	[single, double, triple, aromatic]
	conjugation	binary	1	whether the bond is conjugated [0/1]
	ring	binary	1	whether the bond is in ring [0/1]
	stereo	one-hot	4	[StereoNone, StereoAny, StereoZ, StereoE]
Intermolecular Interaction Features between atoms		Types	Sizes	Descriptions
λ ⁽⁰⁾ (d)	Distance (d) in Coulomb force	real-valued vector	64	$\left[\exp\left(-\frac{1}{(1+\phi_{12})^2}\right)\ d^{1,2}-\phi_0\ ^2\right], \dots, \exp\left(-\frac{1}{(\phi_{12}-\phi_{13})^2}\right)\ d^{1,2}-\phi_{12}\ ^2\right], \phi_0=(6-5d/63)^{-2}$
	Distance (d) in London dispersion forces	real-valued vector	64	$\left[\exp\left(-\frac{1}{(1+\phi_{12})^2}\right)\ d^{1,2}-\phi_0\ ^2\right], \dots, \exp\left(-\frac{1}{(\phi_{12}-\phi_{13})^2}\right)\ d^{1,2}-\phi_{12}\ ^2\right], \phi_0=(6-5d/63)^{-6}$

FIG. 3

400

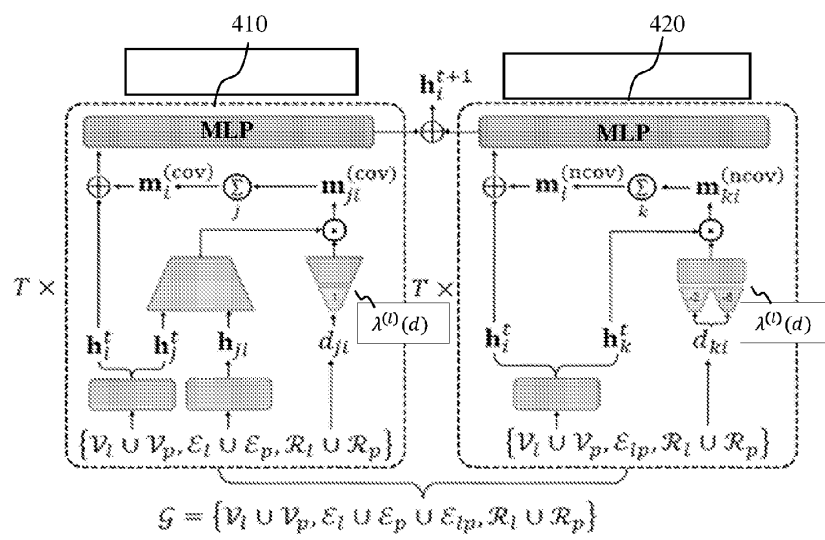


FIG. 4

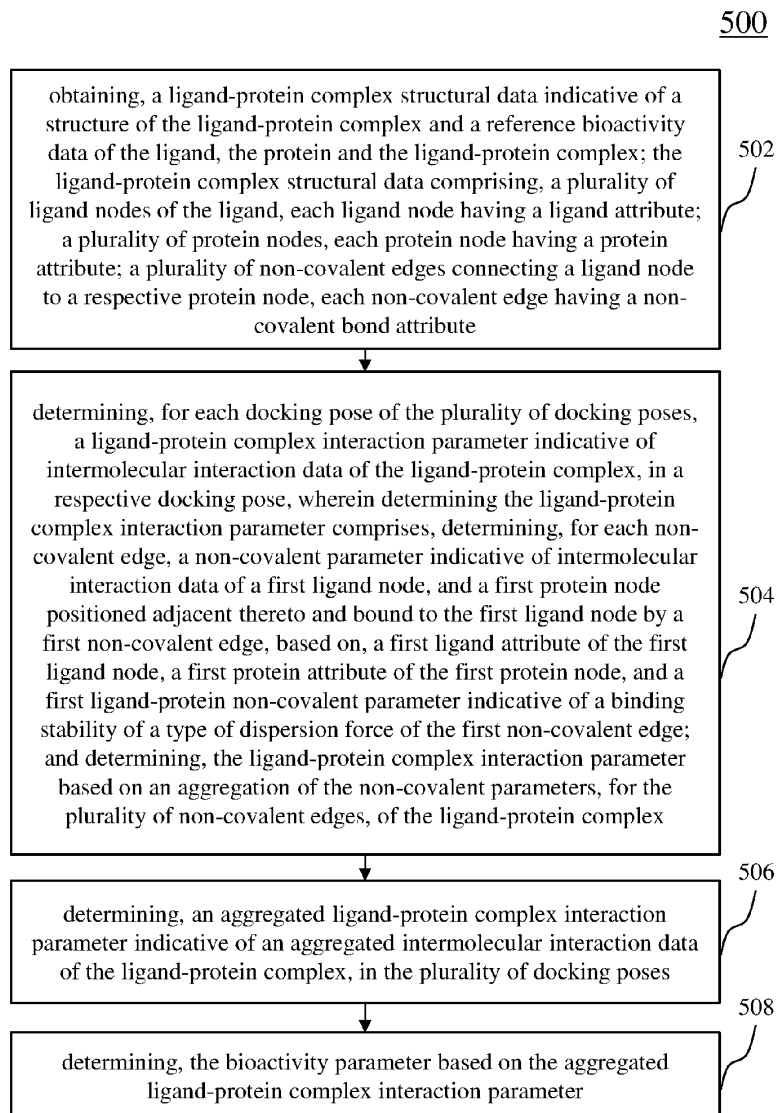
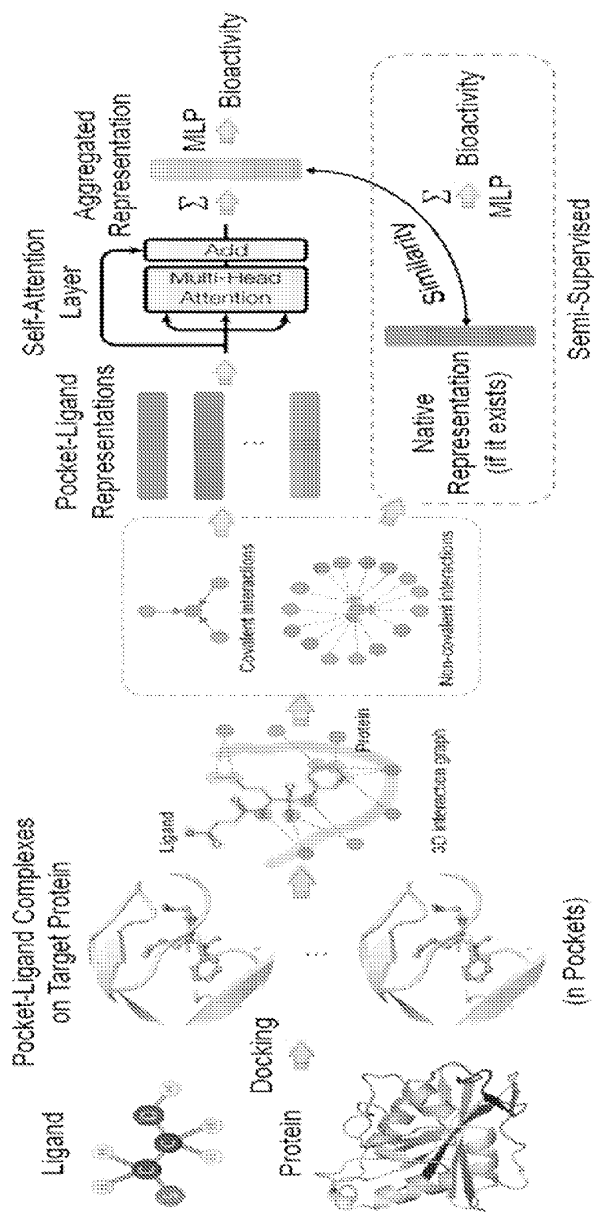


FIG. 5



600

FIG. 6

700

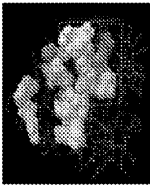
Pockets	Pocket 1	Pocket 2	Pocket 3	Pocket 4
Binding Sites	10, 16-22	31-36	55-57, 65	71
Grid Boxes				
Pocket	Pocket 5	Pocket 6	Pocket 7	
Binding Sites	77-79	117-124	ALL	
Grid Boxes				

FIG. 7

800

Data Sources	Data Types	Methods	Evaluation Metrics		
			r (%)	RMSE (Å)	τ_B (%)
Ligand Structure	2D Graph	GCN	0.2596	1.6651	0.1977
	3D Graph	Uni-Mol	0.3193	1.8373	0.2480
Protein-Ligand QVina-W-Docked Complexes	3D Graph	GIGN	0.4552	1.4937	0.3201
Protein-Ligand Vina-Docked Complexes			0.5092	1.5803	0.3786

FIG. 8

IDs	Targets	Protein Names	Assay Types	PDB IDs ¹	# Binding Sites	# Molecules	# Pockets ²	# Poses /Pocket ³
I1	CHEMBL202	Dihydrofolate reductase	IC50	1hoz	8	957	7	1
I2	CHEMBL3976	Dipeptidyl peptidase 2	IC50	4c06	3	1694	2	4
I3	CHEMBL333	72 kDa type IV collagenase	IC50	1c47	24	3656	6	3
I4	CHEMBL2971	JAK2 HUMAN	IC50	3ags	3	6207	3	4
I5	CHEMBL279	Vascular endothelial growth factor receptor 2	IC50	lyw0	3	9573	3	4
E1	CHEMBL3820	Hexokinase-4	EC50	309n	11	997	6	3
E2	CHEMBL4422	Free fatty acid receptor 1	EC50	5ar1	2	1693	3	3
E3	CHEMBL235	Peroxisome proliferator-activated receptor gamma	EC50	1zgy	4	3611	4	2

¹ The PDB IDs were selected with high resolution and with their ligands that bound most binding sites. ² The number of handpicked pockets based on binding site clusters. ³ The number of involved top poses in each pocket, selected based on docking scores.

FIG. 9

Data Sources		Data Types	Methods	r (↑)	RMSE (↓)	τ_h (↑)
Score	AutoDock-Vina Score					
		MLP				
Ligand	2D Graph	GCN		0.2596	1.6651	0.1977
		GAT		0.3409	1.5935	0.2316
		GIN (Pre-trained)		0.4223	1.3344	0.2820
		MPNN		0.3753	1.4800	0.2393
		Weave		0.0430	1.6337	0.0411
		Neural FP		0.2007	1.8375	0.1367
		Attentive FP		0.2261	1.5786	0.1600
		SGCN		0.0937	1.6267	0.0841
3D Graph	UniMol (Pre-trained)					
			0.3193	1.8373	0.2480	
Pocket + Ligand	3D Graph	DTIGN w/o edge feature & self-attention				
				0.5092	1.5803	0.3786
		DTIGN w/o self-attention				
				0.5065	1.2645	0.3544
		DTIGN Unsupervised				
		0.5916	1.2823	0.4165		
		DTIGN Semi-Supervised (~3%) [†]				
				0.6116	1.1874	0.4507

[†] In the training set, about 3% of the active ligands possess native structures, which contribute as supervised information in semi-supervised learning.

FIG. 10

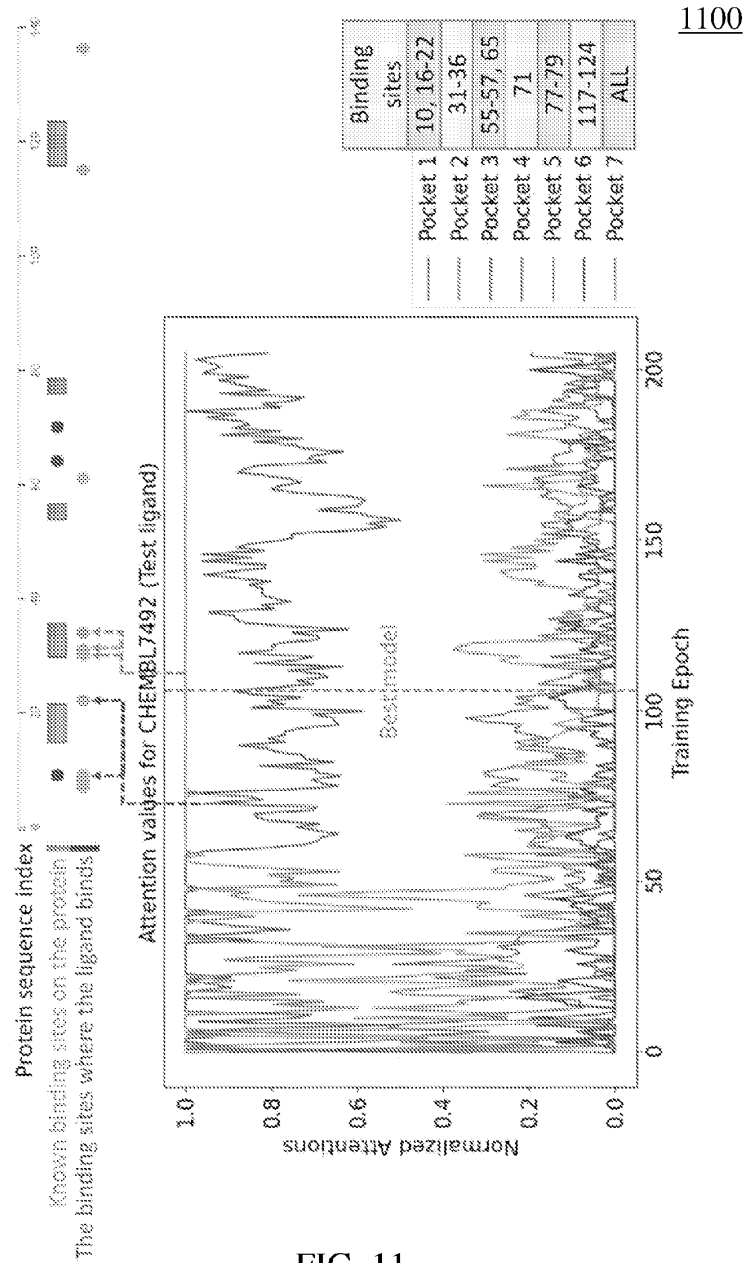


FIG. 11

1200

Datasets	# Native Structures	# Native Pockets in Average	# IoU ≥ 0.5	Mean IoU ¹
Training	17	4	13	0.55
Test	9	4	6	0.56

¹ While computing the mean Intersection over Union (IoU), “Attention Pockets” are determined by the highest average attention scores during training, with the number of “Attention Pockets” aligned with “Native Pockets”. For drug-target pairs with multiple native conformations, IoU is evaluated individually between each “Native Pockets” of native conformation and “Attention Pockets”.

FIG. 12

1300

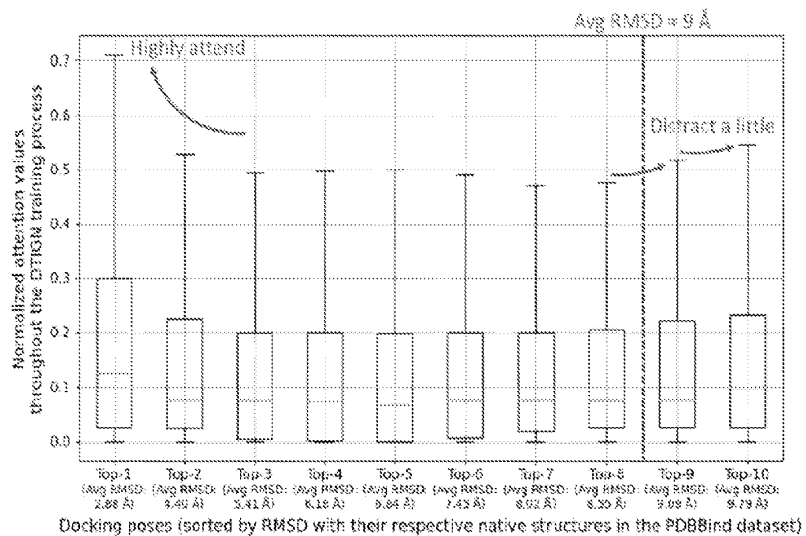


FIG. 13

1400

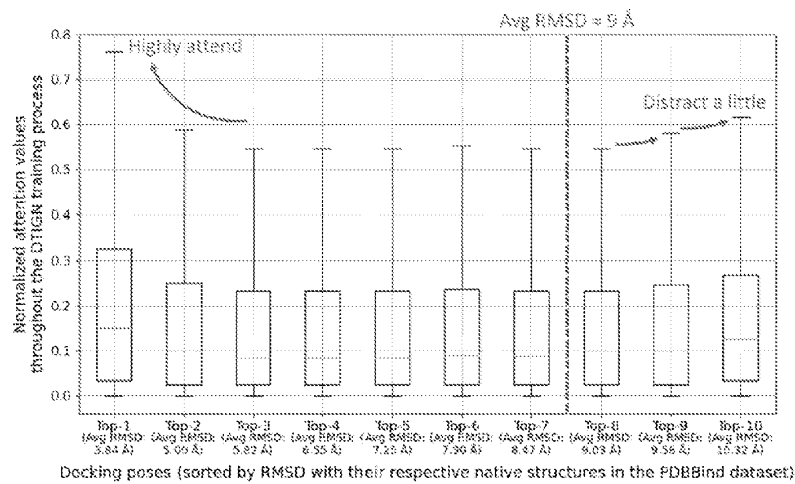


FIG. 14

1500

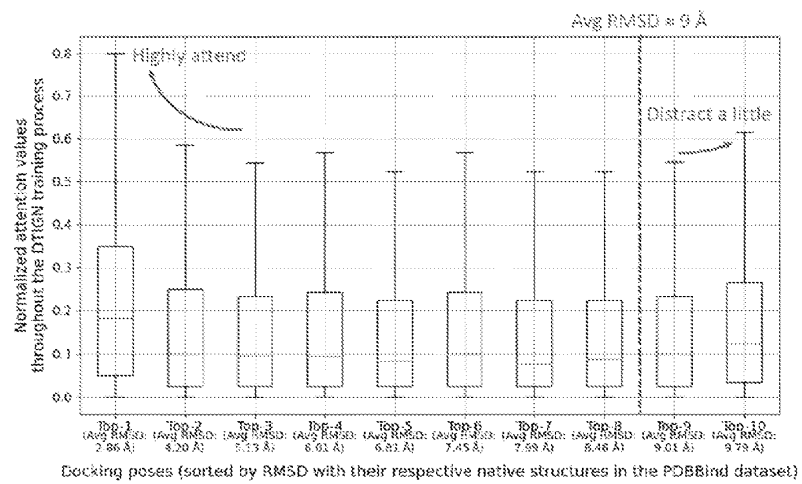


FIG. 15

1600

Data	Data Type	Method	F1	F2	F3	F4	F5	E1	E2	E3	Average
Score-based	AutoDock-Vina Score	MLP	0.3858	0.0246	0.0665	0.1205	0.1745	-0.0100	0.3325	0.1443	0.1548
Ligand-based	2D Graph	GCN	0.2596	0.1344	0.4340	0.3973	0.3433	0.1601	0.3540	0.1697	0.2814
		GAT	0.3409	0.4932	0.9039	0.2406	0.2788	0.1649	0.3944	0.2016	0.3133
		GIN (Pre-trained)	0.4223	0.1707	0.3466	0.3066	0.2417	0.1401	0.2396	0.1654	0.2466
		MPNN	0.3733	0.2395	0.3648	0.1942	0.0928	0.0691	0.3157	0.0883	0.2175
		Weave	0.0430	0.4000	0.3337	0.2393	0.2149	0.1153	0.3967	0.2345	0.2394
		Neural FP	0.2907	0.0011	0.4095	0.3197	0.1588	0.0843	0.3334	0.0862	0.1905
		Attentive FP	0.2261	0.2030	0.3777	0.3902	0.3145	0.0220	0.4000	0.2320	0.2748
		SGCN	0.0037	0.3845	0.1601	0.1757	0.1820	0.1528	0.3465	0.0825	0.1847
	3D Graph	UniMot (Pre-trained)	0.3193	0.0451	0.2298	0.1516	0.1959	0.0966	0.4158	0.1928	0.2060
Interaction-based	3D Graph	DTIGN (Ours)	0.5916	0.5332	0.4737	0.3887	0.3487	0.2802	0.4618	0.3160	0.4142

FIG. 16

1700

Data	Data Type	Method	F1	F2	F3	F4	F5	E1	E2	E3	Average
Score-based	Vina Scores	MLP	1.5344	2.0807	2.0394	1.4773	1.1827	0.8641	0.9282	1.3202	1.4246
Ligand-based	2D Graph	GCN	1.6651	1.9815	1.5389	1.2213	1.1334	0.9010	0.8182	1.2738	1.3167
		GAT	1.5935	1.7103	1.5217	1.1802	1.1715	0.9700	0.8278	1.2603	1.2794
		GIN *	1.3344	1.9587	1.5386	1.2143	1.1844	0.9570	0.9457	1.3187	1.3065
		MPNN	1.4800	1.9113	1.6893	1.1743	1.1846	0.8752	0.8295	1.4227	1.3209
		Weave	1.6337	1.9423	1.6326	1.1762	1.1764	0.9990	0.8728	1.2861	1.3399
		Neural FP	1.8375	1.8250	1.5837	1.3622	1.2075	0.8863	1.1840	1.2962	1.3978
		Attentive FP	1.5786	1.8754	1.6390	1.1998	1.1485	0.9346	0.8544	1.2635	1.3117
		SGCN	1.6267	1.8401	1.8817	1.2668	1.1856	1.0230	0.9060	1.3370	1.3834
	3D Graph	UniMot *	1.8373	1.6848	1.7296	1.1948	1.3928	1.8867	0.9067	1.2975	1.4913
Interaction-based	3D Graph	DTIGN (Ours)	1.2823	1.6719	1.4735	1.1666	1.1251	0.8551	0.8641	1.2313	1.2087

* Pre-trained.

FIG. 17

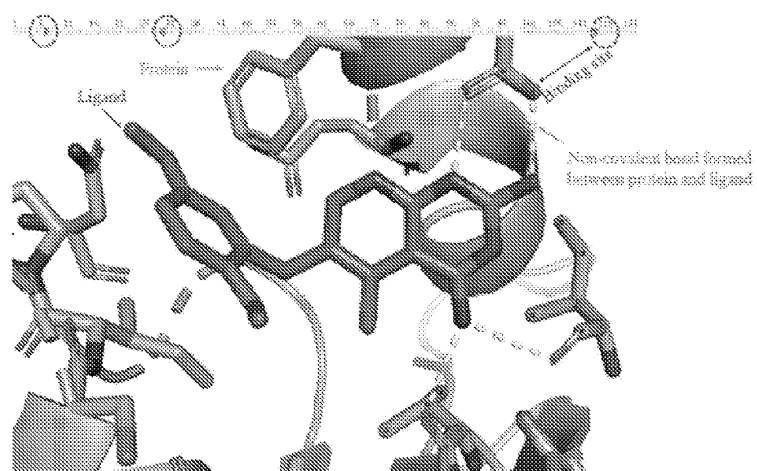
1800

Data	Data Type	Method	I1	I2	I3	I4	I5	E1	E2	E3	Average
Score-based	Vinn Scores	MLP	0.2921	-0.0191	0.0201	0.0824	0.0905	-0.0575	0.2792	0.1494	0.1059
		GCN	0.1977	0.0488	0.3118	0.2908	0.2039	-0.1224	0.2431	0.1663	0.1675
Ligand-based	2D Graph	GAT	0.2316	0.3690	0.2818	0.1505	0.1683	-0.0596	0.3198	0.1849	0.2058
		GIN [*]	0.2820	-0.1099	0.2415	0.2048	0.1375	-0.1099	0.1548	0.1621	0.1204
		MPNN	0.2393	0.2283	0.2182	0.1180	0.0476	0.0688	0.2152	0.0683	0.1505
		Weave	0.0411	-0.2845	0.2407	0.1565	0.1241	-0.1022	0.2327	0.2004	0.0761
		Neural FP	-0.1367	-0.0346	0.2853	0.2255	0.1178	-0.0638	0.2519	0.0282	0.0842
		Attentive FP	0.1600	0.0892	0.2570	0.2976	0.1927	-0.0340	0.3068	0.1713	0.1801
		SGCN	0.0841	-0.3600	0.0689	-0.1456	0.1769	-0.1456	0.1769	0.0566	-0.0110
		UniMot [*]	0.2480	0.0628	0.1361	0.0152	0.1362	0.0152	0.2661	0.0983	0.1223
Interaction-based	3D Graph	DTIGN (Ours)	0.4165	0.3926	0.3140	0.2725	0.2261	0.1387	0.3296	0.2714	0.2952

^{*} Pre-trained.

FIG. 18

(A)

1900

(B)

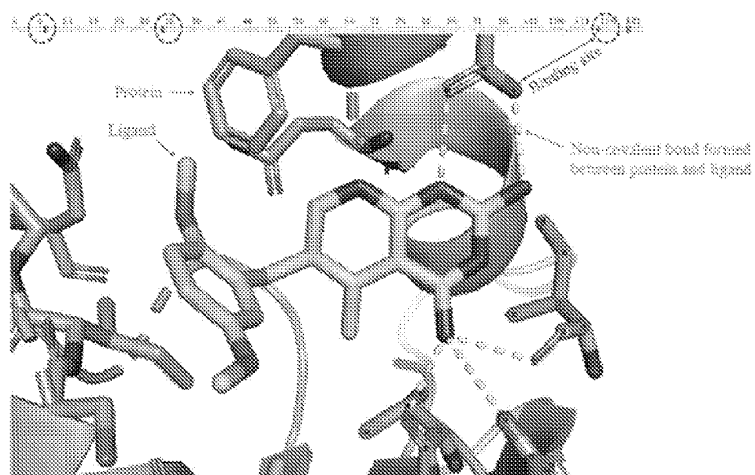


FIG. 19

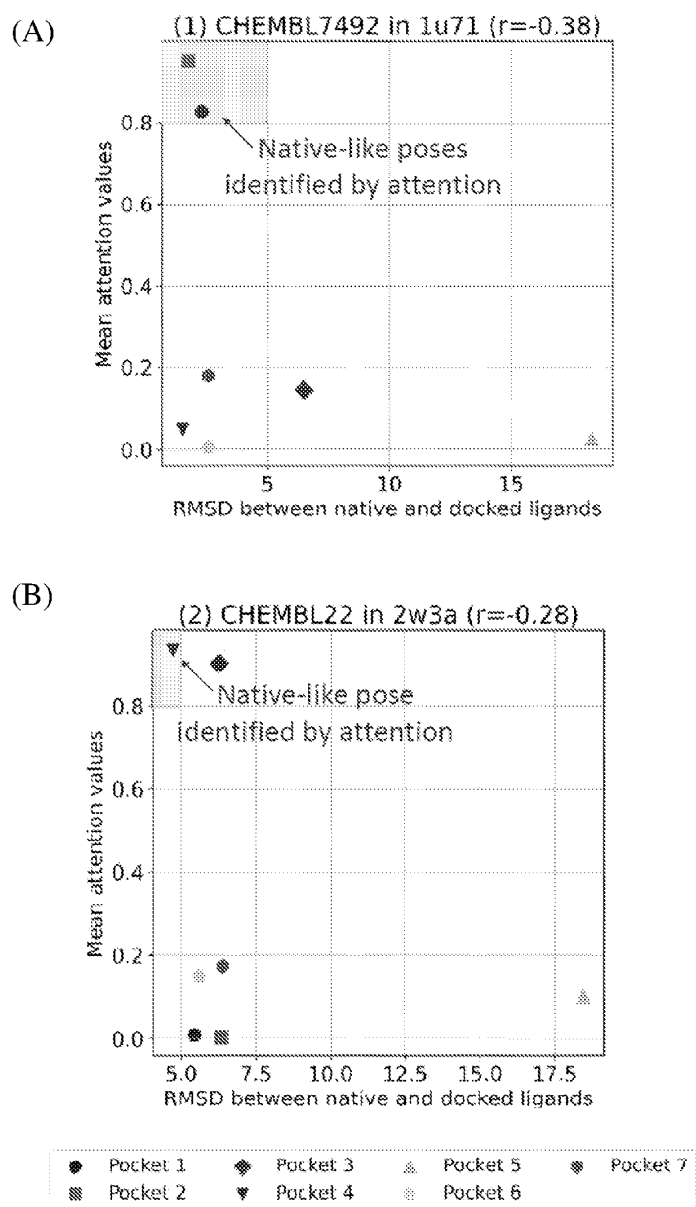


FIG. 20