Supplementary Information

Supplementary Table 1 Statistics of the curated ChEMBL36 activity datasets used in this study.

Dataset version	Task	# Activities	# Targets	# Molecules	# Mixed-label molecules
Unaugmented	Classification	662,770	6,608	514,612	12,140
Unaugmented	Regression	353,594	4,975	196,623	/
Augmented	Classification	1,170,590	6,608	514,612	403,623
Augmented	Regression	521,566	4,975	196,623	/

Note: Mixed-label molecules refer to the number of unique molecules that are associated with both positive and negative labels across different targets or assays.

Supplementary Table 2 Performance comparison on the benchmark DUD-E. Values are reported as mean/median across targets. (All models are retrained from scratch)

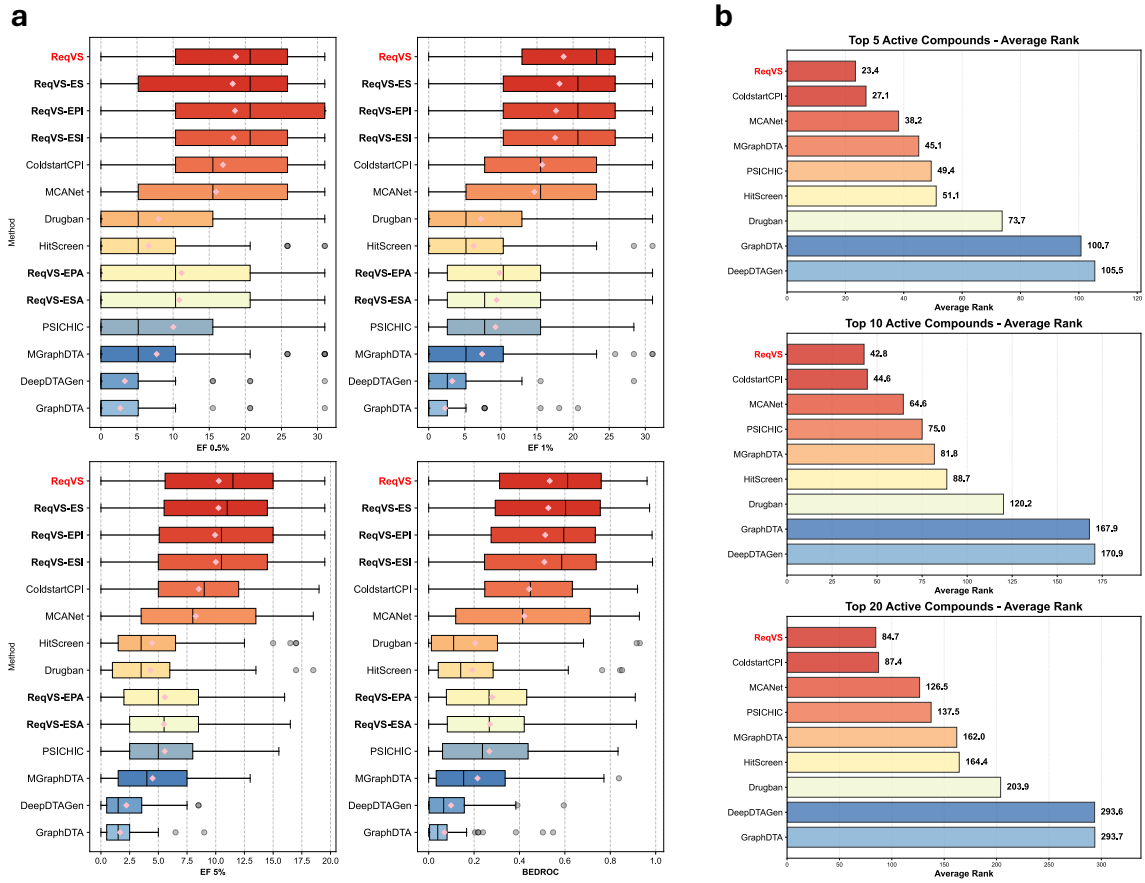
Method	EF0.1%		EF0.5%		EF1%		EF5%		AUROC		BEDROC	
	Mean	Med.	Mean	Med.	Mean	Med.	Mean	Med.	Mean	Med.	Mean	Med.
GraphDTA [1]	4.42	0.00	3.15	0.84	2.64	1.06	1.86	1.23	0.537	0.627	0.051	0.026
MGraphDTA [2]	13.09	4.32	10.96	5.03	9.63	5.39	5.12	3.68	0.672	0.671	0.170	0.117
PSICHIC [3]	16.29	8.70	13.45	7.83	11.54	8.90	6.44	5.89	0.758	0.795	0.209	0.176
DeepDTAGen [4]	5.98	0.00	4.91	1.66	4.37	1.98	3.05	2.00	0.590	0.605	0.085	0.044
DrugBAN [5]	12.37	0.00	10.45	4.25	8.68	3.93	4.61	3.49	0.682	0.723	0.157	0.081
MCANet [6]	32.52	34.74	26.08	20.64	21.60	15.33	8.33	7.35	0.742	0.782	0.350	0.258
HitScreen [7]	14.77	6.51	11.02	4.30	9.55	4.45	4.95	3.62	0.708	0.722	0.170	0.095
ColdstartCPI [8]	33.02	38.61	27.98	25.42	22.79	18.61	9.20	9.27	0.813	0.865	0.377	0.350
ReqVS-ESA	22.33	14.23	15.50	10.23	12.52	7.54	5.92	5.18	0.724	0.750	0.216	0.147
ReqVS-EPA	22.48	15.75	16.28	10.65	13.02	7.78	6.16	5.47	0.727	0.754	0.225	0.161
ReqVS-ESI	38.43	41.00	31.74	28.28	26.99	22.79	10.19	**11.58**	0.806	0.866	0.437	0.418
ReqVS-EPI	39.68	43.41	32.51	30.81	27.25	23.07	10.04	11.49	0.798	0.857	0.440	0.423
ReqVS-ES	41.64	51.62	33.51	32.59	28.16	25.33	**10.43**	11.40	**0.821**	0.866	0.454	0.465
ReqVS	**42.65**	**52.38**	**34.29**	**35.12**	**28.58**	**27.95**	10.40	11.44	0.820	**0.871**	**0.459**	**0.494**

Note: Bold values indicate the best performance in each metric.

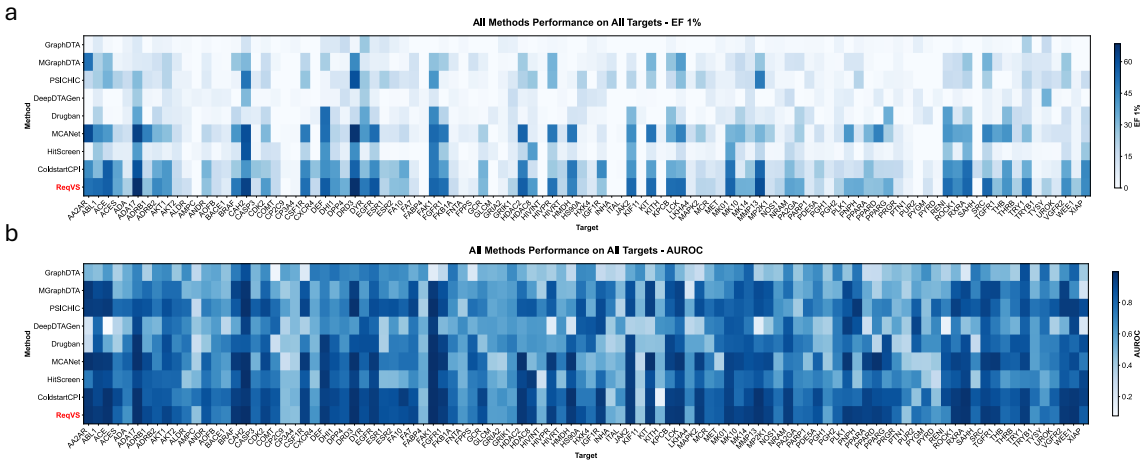
Supplementary Table 3 Performance comparison on the benchmark DEKOIS2.0. Values are reported as mean/median across targets. EF0.1% is not reported because the limited number of compounds per target in DEKOIS2.0 makes this metric statistically meaningless. (All models are retrained from scratch)

Method	EF0.5%		EF1%		EF5%		AUROC		BEDROC	
	Mean	Med.	Mean	Med.	Mean	Med.	Mean	Med.	Mean	Med.
GraphDTA [1]	2.68	0.00	2.26	0.00	1.67	1.50	0.558	0.558	0.069	0.040
MGraphDTA [2]	7.72	5.17	7.40	5.17	4.50	4.00	0.699	0.735	0.215	0.154
PSICHIC [3]	10.01	5.17	9.25	7.75	5.56	5.00	0.750	0.805	0.267	0.237
DeepDTAGen [4]	3.32	0.00	3.25	2.58	2.22	1.50	0.560	0.559	0.099	0.066
DrugBAN [5]	7.97	5.17	7.24	5.16	4.29	3.50	0.678	0.710	0.206	0.110
MCANet [6]	15.95	15.50	14.64	15.50	8.27	8.00	0.763	0.851	0.421	0.414
HitScreen [7]	6.57	5.17	6.25	5.17	4.46	3.50	0.710	0.744	0.192	0.141
ColdstartCPI [8]	16.90	15.50	15.72	15.50	8.53	9.00	0.814	0.880	0.442	0.448
ReqVS-ESA	10.84	10.33	9.41	7.75	5.49	5.50	0.735	0.784	0.271	0.267
ReqVS-EPA	11.16	10.33	9.82	10.33	5.57	5.00	0.736	0.788	0.280	0.266
ReqVS-ESI	18.37	**20.67**	17.51	20.67	10.01	10.50	0.817	0.891	0.509	0.586
ReqVS-EPI	18.56	**20.67**	17.64	20.67	9.93	10.50	0.814	0.886	0.513	0.586
ReqVS-ES	18.24	**20.67**	18.12	20.67	10.24	11.00	**0.828**	**0.890**	0.527	0.603
ReqVS	**18.69**	**20.67**	**18.69**	**23.25**	**10.27**	**11.50**	0.825	0.888	**0.533**	**0.612**

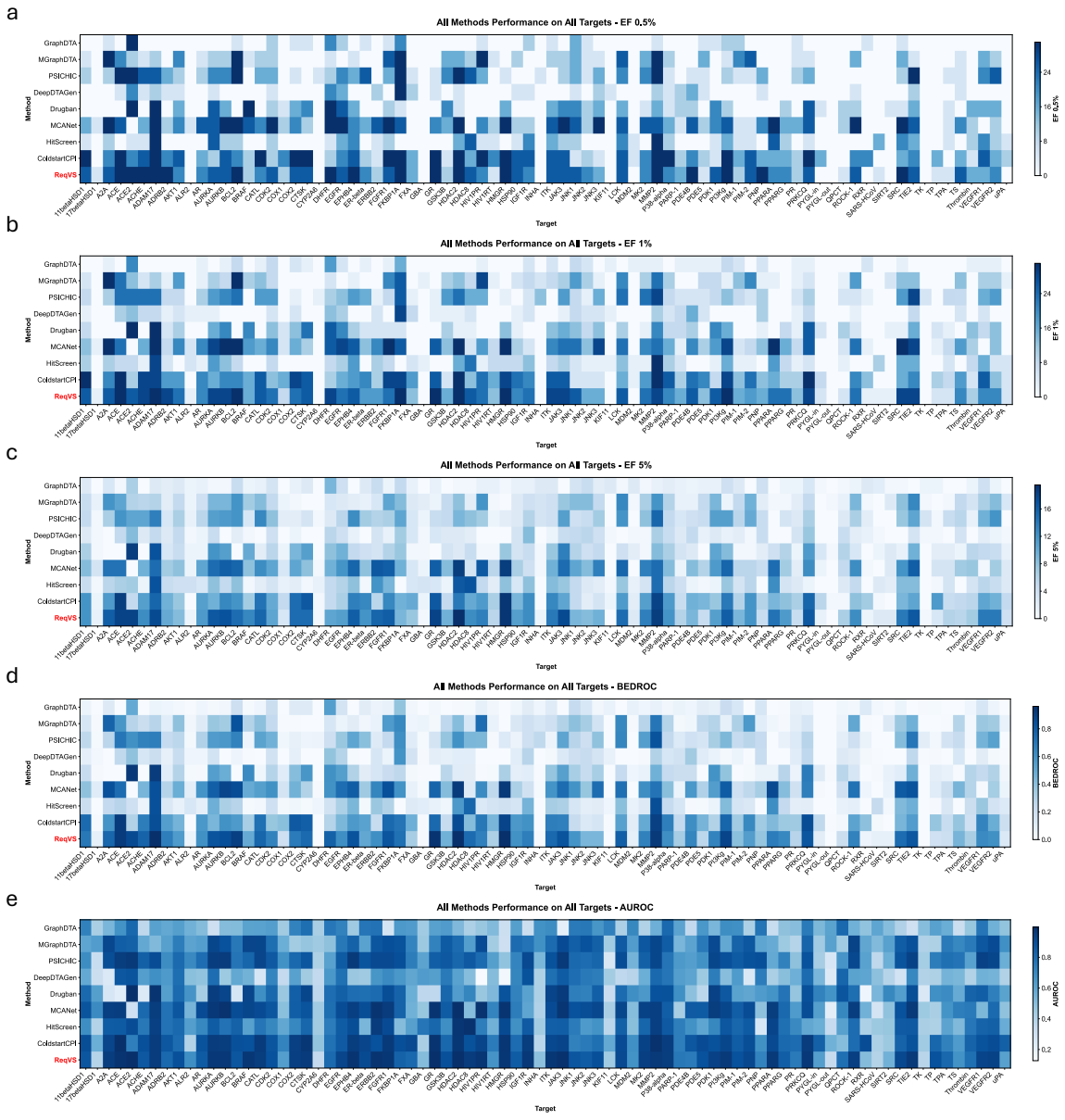
Note: Bold values indicate the best performance in each metric.



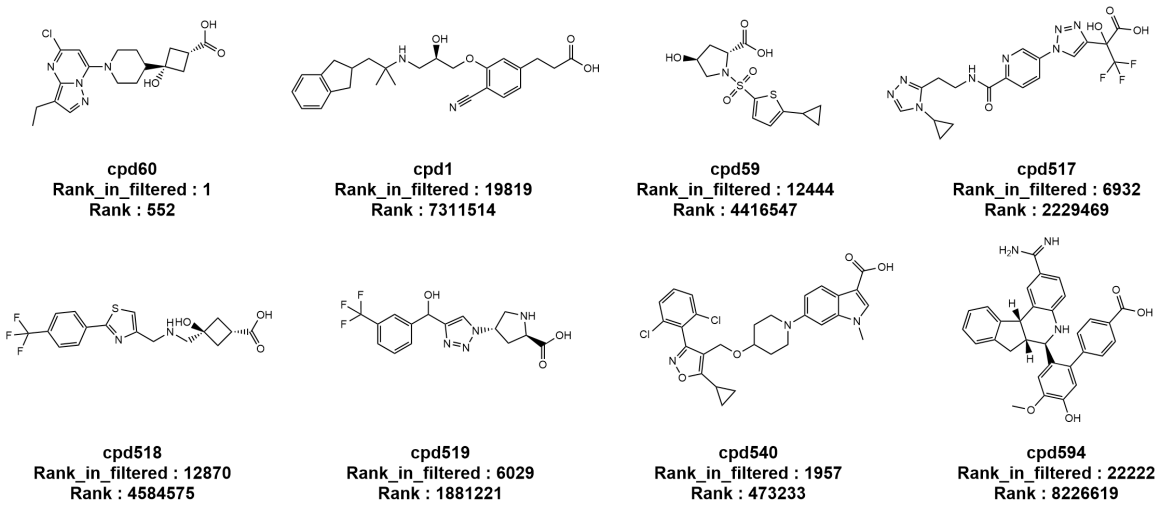
Supplementary Figure 1 Performance evaluation and ablation study of ReqVS on the DEKOIS2.0 benchmark.a Comparison of ReqVS against its framework variants and single constituent models across four key metrics: EF0.5%, EF1%, EF5%, and BEDROC. EF0.1% is not reported because the limited number of compounds per target in DEKOIS2.0 makes this metric statistically meaningless. Pink diamonds denote the mean value for each metric; a median line is included in each box-whisker plot. **b** Comparison of the Average Rank for the top 5, 10, and 20 active compounds across all test targets. Lower values indicate a superior capability to prioritize active molecules at the absolute top of the screening list.



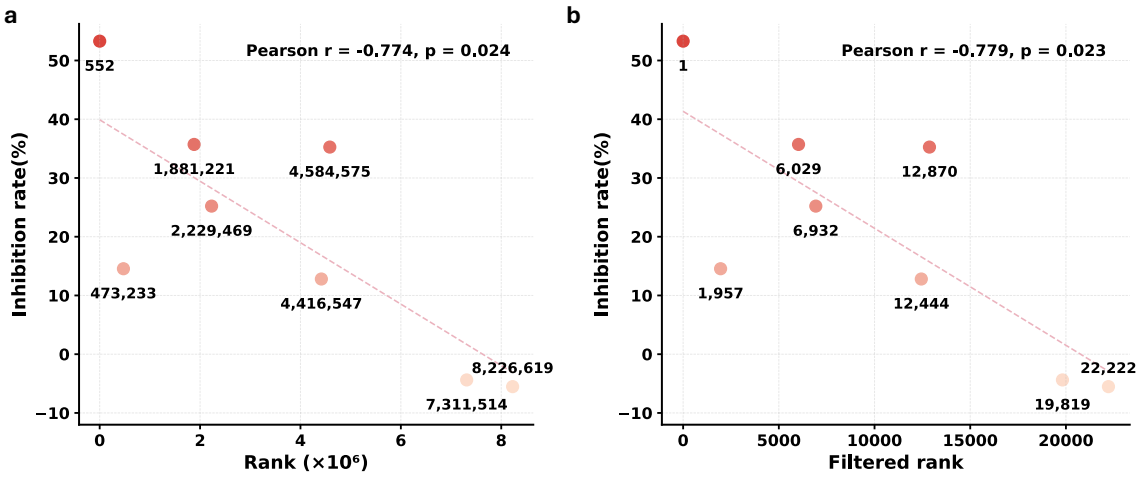
Supplementary Figure 2 Target-wise performance heatmaps of ReqVS and 8 contained methods across all targets in the DUD-E dataset. The heatmaps across targets in terms of **a** EF1%, and **b** AUROC, respectively.



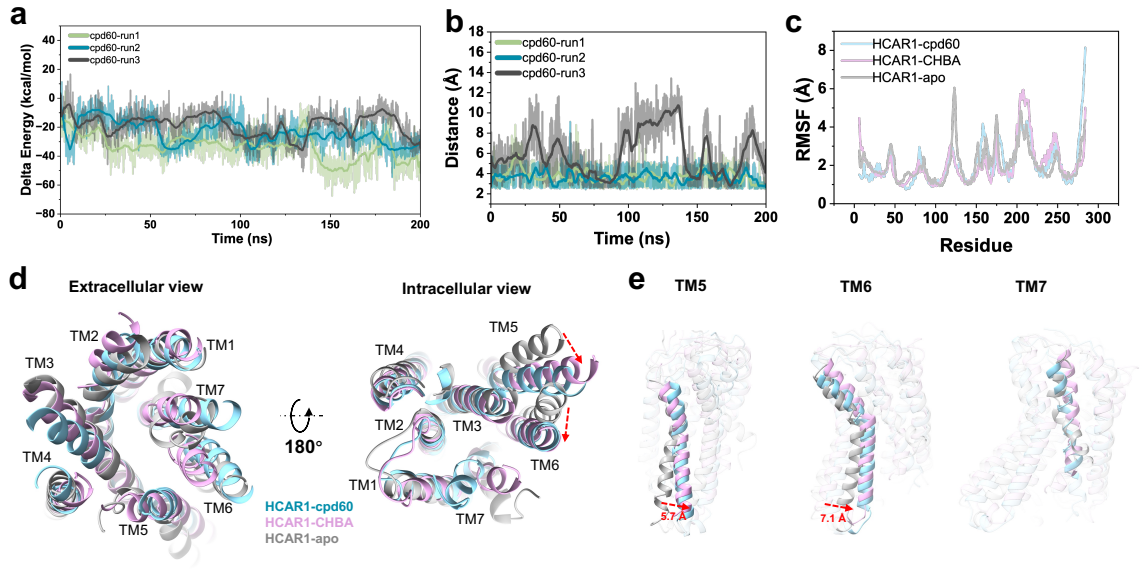
Supplementary Figure 3 Target-wise performance heatmaps of ReqVS and 8 contained methods across all targets in the DEKOIS2.0 dataset. The heatmaps across targets in terms of **a** EF0.5% , **b** EF1% , **c** EF5% , **d** BEDROC, and **e** AUROC, respectively.



Supplementary Figure 4 Chemical structures and virtual screening rankings of the eight candidate compounds.



Supplementary Figure 5 Correlation between ReqVS predicted ranks and inhibition rate of HCAR1. **a** Scatter plot illustrating the relationship between the raw virtual screening rank ($\times 10^6$) of the eight candidate compounds and their experimental inhibition (%). A significant negative correlation is observed, with a Pearson $r = -0.774$, $p = 0.024$. **b** Scatter plot displaying the relationship between the candidate compounds' inhibition rate and their filtered rank after applying expert-guided structural constraints. The correlation is maintained with Pearson $r = -0.779$, $p = 0.023$.



Supplementary Figure 6 Additional MD analyses of the cpd60-HCAR1, related to Figure 8. a Binding free energy of the cpd60–HCAR1 complex during three independent MD simulations. **b** Distance between the chlorine atom of cpd60 and the OE2 atom of Glu171. **c** RMSF of HCAR1 in the cpd60-bound, CHBA-bound, and apo states. **d** Structural superposition of apo-, CHBA-bound, and cpd60-bound HCAR1 conformations at 100 ns after alignment to the TM1–4 core, shown from extracellular and intracellular views. **e** Comparison of the conformational changes of TM5–TM7 among apo-, CHBA-bound, and cpd60-bound HCAR1 at 100 ns.

We propose the following proposition to derive the approximation error bound when using sequence-based entropy as a practical **surrogate** for interaction uncertainty, i.e.,

$$I(Y; S \mid s_t, s_c) \approx H(Y \mid s_t, s_c).$$

Proposition 1 (Surrogate Approximation Error) *The approximation error of using sequence-based entropy as a surrogate for interaction uncertainty satisfies*

$$H(Y \mid s_t, s_c) - I(Y; S \mid s_t, s_c) = H(Y \mid S). \tag{1}$$

Therefore, the approximation error is bounded by

$$0 \leq H(Y \mid s_t, s_c) - I(Y; S \mid s_t, s_c) = H(Y \mid S) \leq H(Y). \tag{2}$$

Under the assumption that $Y \in \{0, 1\}$, the entropy of a binary random variable is bounded by

$$H(Y) \leq \log_2 2 = 1. \tag{3}$$

Therefore, the surrogate approximation error satisfies

$$0 \leq H(Y \mid s_t, s_c) - I(Y; S \mid s_t, s_c) \leq 1. \tag{4}$$

This result shows that the worst-case approximation error is bounded by one bit, which corresponds to the maximum uncertainty of a binary interaction outcome.

References

[1] Nguyen, T. *et al.* Graphdta: predicting drug–target binding affinity with graph neural networks. *Bioinformatics* **37**, 1140–1147 (2021).

[2] Yang, Z., Zhong, W., Zhao, L. & Chen, C. Y.-C. Mgraphdta: deep multiscale graph neural network for explainable drug–target binding affinity prediction. *Chemical science* **13**, 816–833 (2022).

[3] Koh, H. Y., Nguyen, A. T., Pan, S., May, L. T. & Webb, G. I. Physicochemical graph neural network for learning protein–ligand interaction fingerprints from sequence data. *Nature Machine Intelligence* **6**, 673–687 (2024).

[4] Shah, P. M. *et al.* Deepdtagen: a multitask deep learning framework for drug-target affinity prediction and target-aware drugs generation. *Nature Communications* **16**, 5021 (2025).

[5] Bai, P., Miljković, F., John, B. & Lu, H. Interpretable bilinear attention network with domain adaptation improves drug–target prediction. *Nature Machine Intelligence* **5**, 126–136 (2023).

[6] Bian, J., Zhang, X., Zhang, X., Xu, D. & Wang, G. Mcanet: shared-weight-based multihead-crossattention network for drug–target interaction prediction. *Briefings in Bioinformatics* **24**, bbad082 (2023).

[7] Chen, G. *et al.* Hitscreen: A sequence-based drug virtual screening approach using data augmentation and protein language models. *Journal of Chemical Information and Modeling* **65**, 10152–10166 (2025).

[8] Zhao, Q. *et al.* Coldstartcpi: Induced-fit theory-guided dti predictive model with improved generalization performance. *Nature Communications* **16**, 6436 (2025).