Supplementary materials

Table S1: Summary of common molecular docking programs

Program	Year Published	Organisation	Description	Webservice	License
AutoDock[1]	1990	The Scripps Research Institute	Automated docking of ligand to macromolecule by Lamarckian Genetic Algorithm and Empirical Free Energy Scoring Function	No	Open source (GNU GPL)
AutoDock *vina*[2]	2010	The Scripps Research Institute	New generation of AutoDock	No	Open source (Apache License)
UCSF DOCK[3]	1988	University of California-San Francisco	Based on Geometric Matching Algorithm	No	Freeware for academic use
FlexX[4]	2001	BioSolveIT	Incremental build based docking program	No	Commercial
FRED[5]	2003	OpenEye Scientific	Systematic,exhaustive,nonstochastic examination of all possible poses within the protein active site combined with scoring Function	No	Freeware for academic use
Glide[6]	2004	Schrödinger	Exhaustive search based docking program	No	Commercial
GOLD[7]	1995	Collaboration between the University of Sheffield, GlaxoSmithKlinepl c and CCDC	Genetic algorithm based, flexible ligand, partial flexibility for protein	No	Commercial
LeDock[8]	2016	Lephar	Program for fast and accurate flexible docking of small molecules into a protein	No	Freeware for academic use
LigandFit[9]	2003	BioVia	CHARMm based docking program	No	Commercial
MOE	2008	Chemical Computing Group	Docking application within MOE; choice of placement methods (including alpha sphere methods) and scoring functions (including London dG)	No	Commercial
PSI-DOCK[10]	2006	Peking University	Pose-Sensitive Inclined (PSI)-DOCK	No	Academic
rDock[11]	1998 (commercial)2006 (academic)[14]2012 (open source)[15]	Vernalis R&D(commercial)University of York (academic)University of Barcelona (open source)	HTVS process of small molecules against proteins and nucleic acids, binding mode prediction	No	Open source (GNU LGPL) (formerly commercial, academic)
SEED[12]	1999	University of Zurich	Automated docking of fragments with evaluation of free energy of binding including electrostatic solvation effects in the continuum dielectric approximation (generalized Born)	No	Open source (GNU GPL)
smina[13]	2012	University of Pittsburgh	A customized fork of AutoDock *vina*with a better support scoring function and a high-performance energy minimization	No	Open source (Apache License)
Surflex-Dock[14]	2003	Tripos	Based on an idealized active site ligand (a protomol)	No	Commercial
SwissDock[15]	2011	Swiss Institute of Bioinformatics	Webservice to predict interaction between a protein and a small molecule ligand	Available	Free to use webservice for academic usage

Figure S1

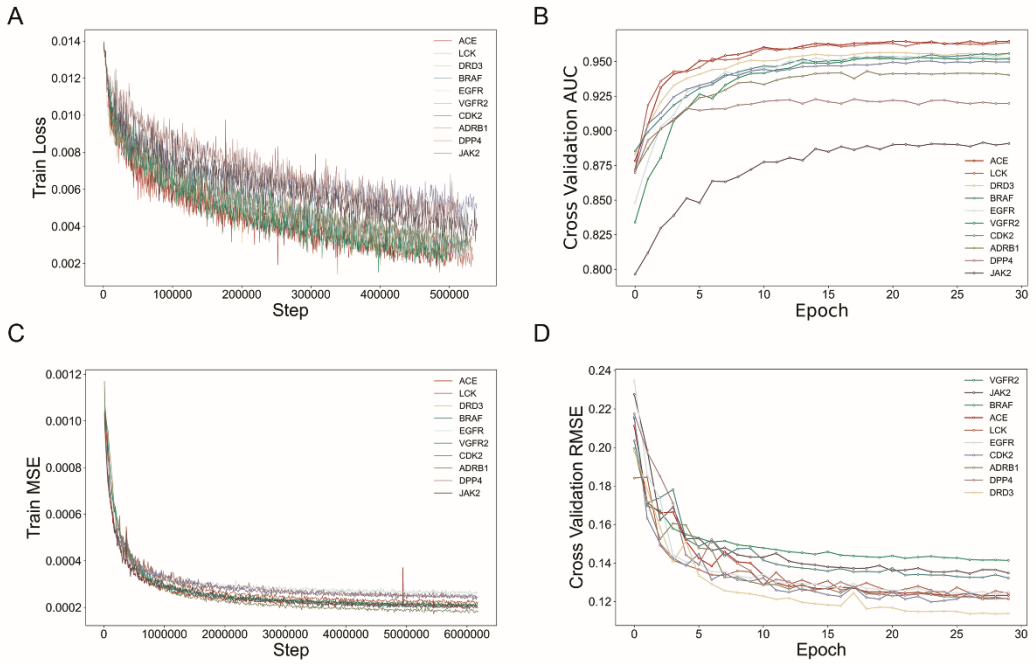

Figure S1: The train Loss(classification model), validation AUC(classification model), train MSE(regression model) and validation RMSE(regression model) of 10 selected targets

Figure S2

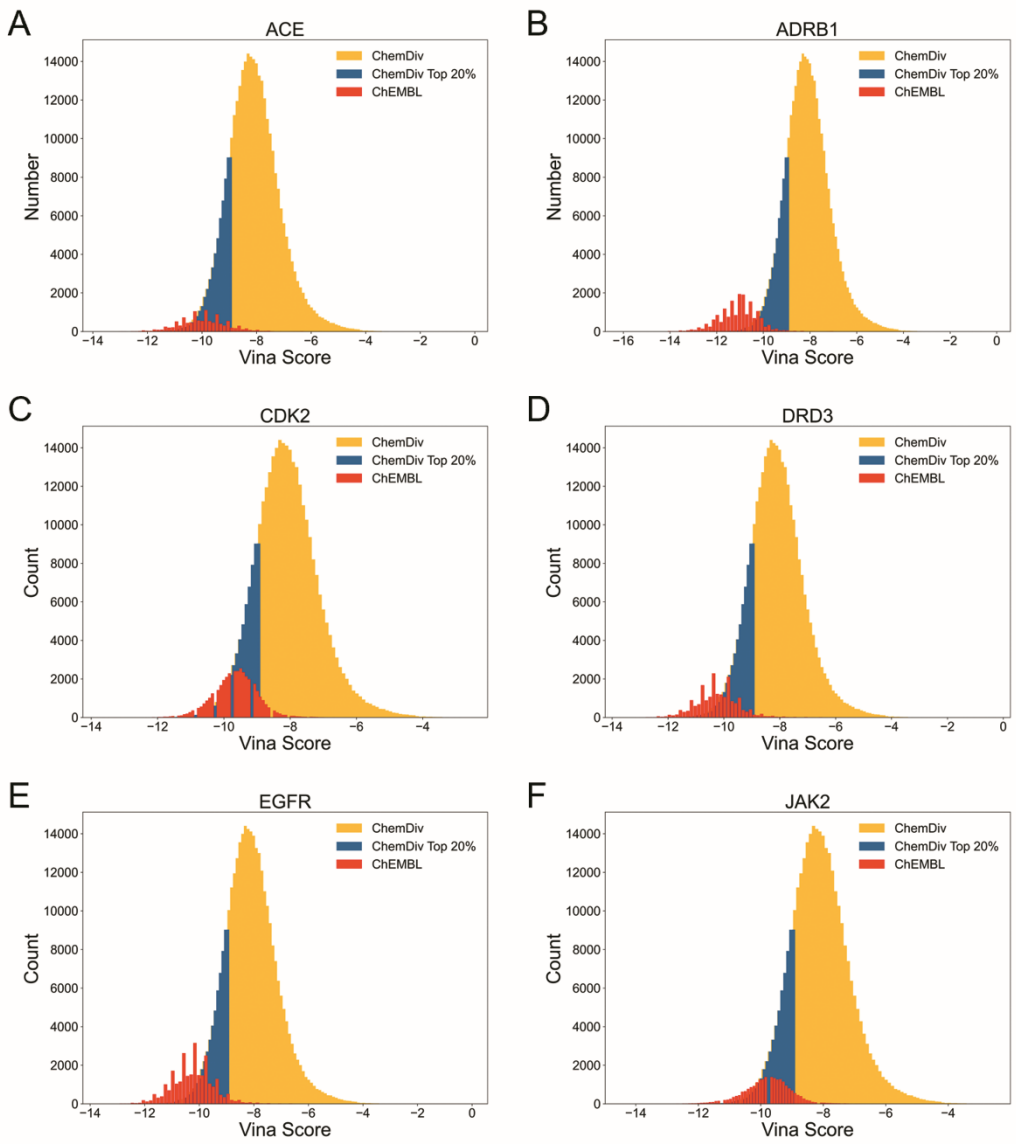

Figure S2: Docking score distribution of ChEMBL recalled and ChemDiv compounds.

Figure S3

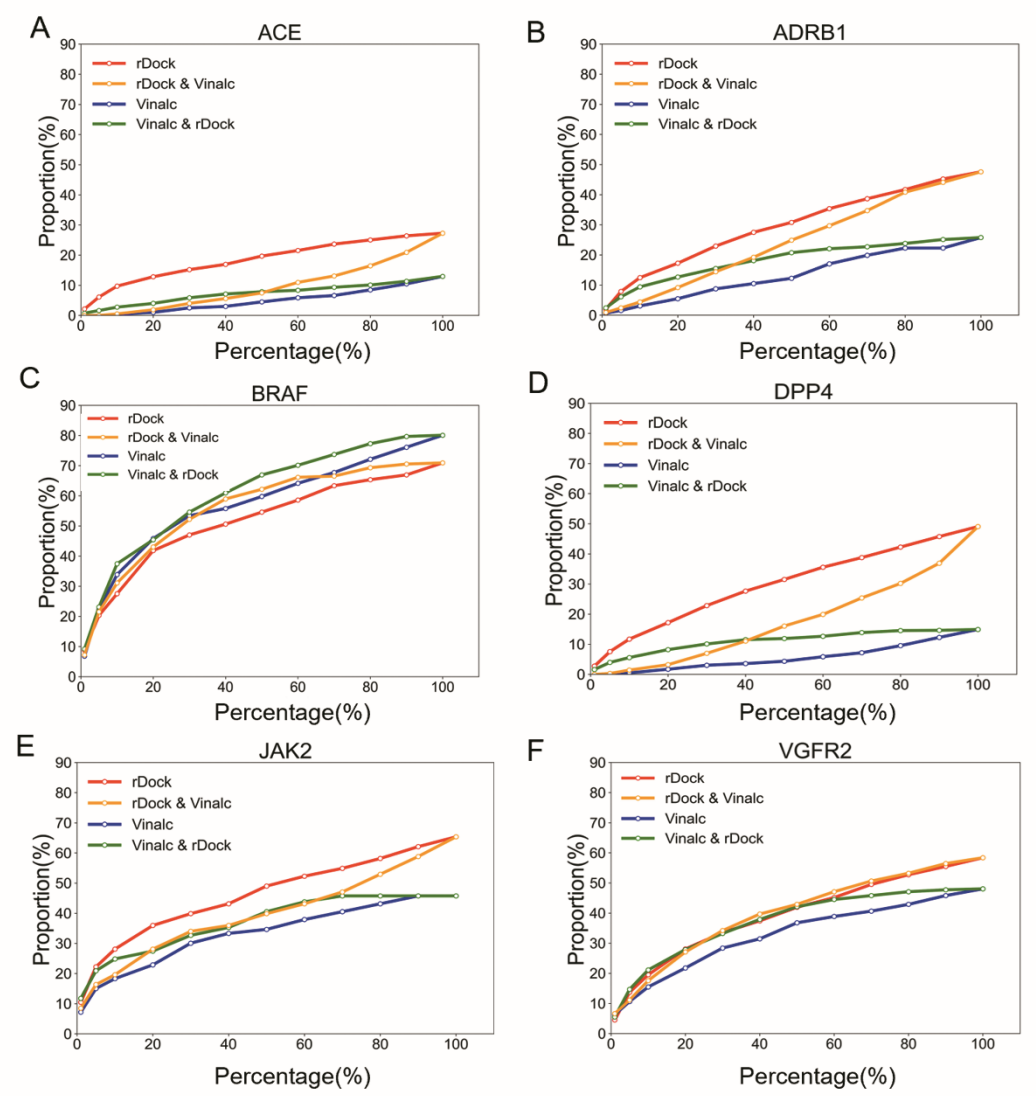

Figure S3: Enrichment ratio of DUD-E active compound for two docking programs.

References:

1. Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S. and Olson, A. J. (2009) Autodock4 and autodocktools4: Automated docking with selective receptor flexiblity. Journal of Computational Chemistry. 16, 2785-2791

2. Trott, O., and Olson, A. J. (2010) Autodock vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of Computational Chemistry. 31, 455-461

3. (2015) Dock 6: Impact of new features and current docking performance. Journal of Computational Chemistry. 36, 1132-1156

4. Rarey, M., Kramer, B., Lengauer, T., and Klebe, G. (1996) A fast flexible docking method using an incremental construction algorithm. Journal of Molecular Biology. 261, 470-489

5. McGann, and Mark. (2011) Fred pose prediction and virtual screening accuracy. Journal of Chemical Information & Modeling. 51, 578-596

6. Halgren, T. A., Murphy, R. B., Friesner, R. A., B Ea Rd, H. S., Frye, L. L., Pollard, W. T., and Banks, J. L. (2004) Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. Journal of Medicinal Chemistry. 47, 1750-1759

7. Curt D . Haffner, James M. Lenhard, Aaron B. Miller, and Darryl L. McDougald. (2004) Structure-based design of potent retinoid x receptor alpha agonists. Journal of Medicinal Chemistry. 47, 2010-2029

8. Liu, N., and Xu, Z. (2019) Using ledock as a docking tool for computational drug design. IOP Conference Series Earth and Environmental Science. 218,

9. Venkatachalam, C. M., Jiang, X., Oldfield, T., and Waldman, M. (2003) Ligandfit: A novel method for the shape-directed rapid docking of ligands to protein active sites. Journal of Molecular Graphics & Modelling. 21, 289-307

10. Pei, J., Wang, Q., Liu, Z., Li, Q., Yang, K., and Lai, L. (2006) Psi-dock: Towards highly efficient and accurate flexible ligand docking. Proteins. 62, 934-946

11. Ruiz-Carmona, S., Alvarez-Garcia, D., Foloppe, N., Garmendia-Doval, A. B., Juhos, S., Schmidtke, P., Barril, X., Hubbard, R. E., and Morley, S. D. (2014) Rdock: A fast, versatile and open source program for docking ligands to proteins and nucleic acids. PLOS Computational Biology. 10, e1003571-e1003578

12. N Majeux, M. S., J Apostolakis, C Ehrhardt, A Caflisch. (1999) Exhaustive docking of molecular fragments on protein binding sites with electrostatic solvation. Proteins: Structure. Function and Genetics. 37, 88-105

13. DR Koes, M. B., CJ Camacho. (2011) Lessons learned in empirical scoring with smina from the csar 2011 benchmarking exercise. Journal of Chemical Information and Modeling. 53, 1893-1904

14. Spitzer, R., and Jain, A. N. (2012) Surflex-dock: Docking benchmarks and real-world application. Journal of Computer-Aided Molecular Design. 26, 687-699

15. Hetényi, C., and Spoel, D. (2011) Toward prediction of functional protein pockets using blind docking and pocket search algorithms. Protein Science. 20,