Support Information. A Graph Convolutional Network-based screening strategy for rapid identification of SARS-CoV-2 cell-entry inhibitors

Peng Gao,[†]Miao Xu,[†]Qi Zhang,[‡,†]Catherine Z Chen,[†]Hui Guo,[†]Yihong Ye,[*,‡]Wei Zheng,[*,†]and Min Shen[*,†]

[†]The National Center for Advancing Translational Sciences (NCATS), National Institutes of Health (NIH), MD 20850, USA

[‡]National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institutes of Health (NIH), MD 20850, USA

E-mail: yihongy@niddk.nih.gov; zhengwei@mail.nih.gov; shenmin@mail.nih.gov

Structural elucidation of the newly identified compounds

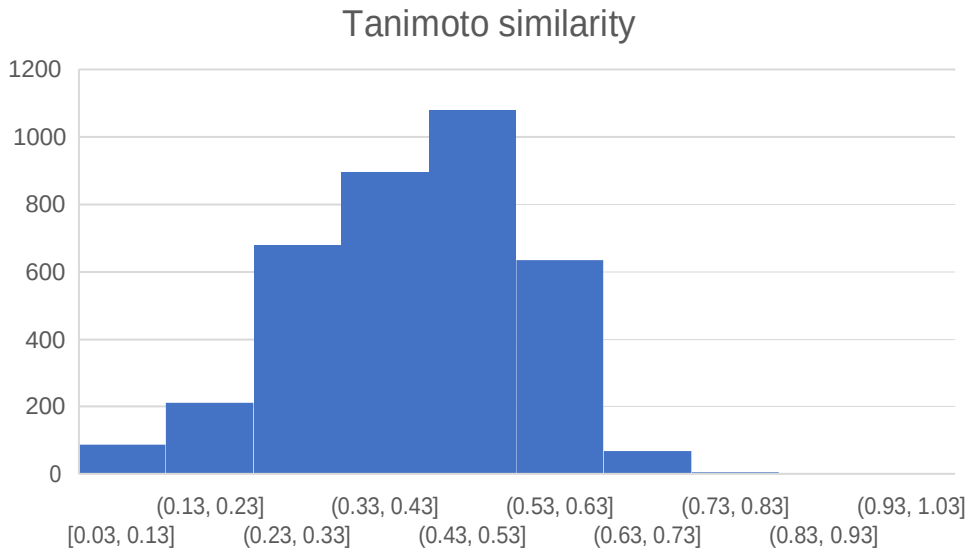

Figure S. 1. Tanimoto similarity analysis of the 367 experimentally identified compounds showing activity within endocytosis assay; the type of the applied fingerprint is Extended Connectivity Fingerprint (ECFP).

Table S. 1. The confusion matrix of the prediction results of validation data set by GCN (SchNet[1]) architecture.

	Active[a)]	Inactive[a)]
Active[b)]	28	17
Inactive[b)]	11	313

[a)] Actual active & inactive labels; [b)] Predicted active & inactive labels.

Table S. 2. Tanimoto similarity of the 10 newly identified compounds with the 367 previously confirmed compounds; the type of the applied fingerprint is Extended Connectivity Fingerprint (ECFP).

	Compound	Tanimoto similarity
1	NCGC00115755-02	0.37
2	NCGC00119962-01	0.34
3	NCGC00159478-04	0.33
4	NCGC00411138-01	0.38
5	NCGC00411588-01	0.38
6	NCGC00411611-01	0.38
7	NCGC00411705-01	0.37
8	NCGC00411718-01	0.38
9	NCGC00411727-01	0.38
10	NCGC00411733-01	0.38

References

(1) Schütt, K. T.; Sauceda, H. E.; Kindermans, P.-J.; Tkatchenko, A.; Müller, K.-R. SchNet – A deep learning architecture for molecules and materials. *The Journal of Chemical Physics* **2018**, *148*, 241722.