Supporting Information

Table S1. Schrödinger Maestro Docking score (kcal/mol) of compounds against SARS-Cov-2 RNA-dependent-polymerase (PDB ID: 7BV2)

Molecular number	Docking score	Glide ligand efficiency	MMGBSA dG Bind(kcal/mol)
Reference	-7.150	-0.170	-65.16
a1	-9.354	-0.312	-45.09
a2	-10.403	-0.4	-43.68
a3	-10.163	-0.462	-41.31
a4	-9.419	-0.41	-34.35
a5	-8.828	-0.384	-26.69
a6	-8.793	-0.419	-25
a7	-9.509	-0.432	-22.12
a8	-10.533	-0.479	-21.03

Where Reference means our reference molecular, a1-a8 is our candidates generated by our model. Docking score and Glide ligand efficiency and MMGBSA dG Bind are conducted by the Schrödinger software. The receptor is 7Bv2 in the PDB database. For the generation of antiviral drugs of nucleoside analogs, we generated 5000 molecules against SARS-Cov-2 RNA-dependent-polymerase and filtered them for molecular docking scoring and MMGBSA binding energy calculation by using Schrödinger software, we finally selected the above a1-a8 molecules based on docking pose, scoring score, and MMGBSA energy, ligand effectiveness, as shown in Table S1.

Table S2. Schrödinger Maestro Docking score (kcal/mol) of compounds against SARS-CoV-2 3CLpro (PDB ID: 6W63)

Molecular number	Docking score	Glide ligand efficiency	MMGBSA dG Bind(kcal/mol)
Reference	-7.624	-0.206	-36.09
b1	-8.593	-0.2	-58.45
b2	-7.923	-0.273	-56.12
b3	-7.147	-0.204	-53.32
b4	-8.013	-0.276	-53.31
b5	-8.065	-0.288	-52.72
b6	-8.105	-0.225	-52.47
b7	-8.553	-0.231	-52.38
b8	-7.416	-0.212	-52.28
b9	-8.142	-0.214	-50.74
b10	-7.325	-0.203	-50.31
b11	-7.061	-0.208	-50.28
b12	-8.263	-0.236	-49.94
b13	-7.671	-0.202	-49.63
b14	-8.06	-0.212	-49.01
b15	-7.729	-0.215	-47.84
b16	-8.018	-0.206	-47.29
b17	-7.188	-0.2	-46.78
b18	-7.848	-0.201	-45.99

Where Reference means our reference molecular, b1-b18 is our candidates generated by our model. Docking score and Glide ligand efficiency and MMGBSA dG Bind are conducted by the Schrödinger software. The receptor is 6w63 in the PDB database. For the generation of antiviral drugs of non-nucleoside analogs, we generated 5000 molecules against SARS-CoV-2 3CLpro and filtered them for molecular docking scoring and MMGBSA binding energy calculation by using Schrödinger software, we finally selected the above b1-b18 molecules based on docking pose, scoring score, and MMGBSA energy, ligand effectiveness, as shown in Table S2.

Transformer-based molecular generative model for antiviral drug design

Jiashun Mao

The Interdisciplinary Graduate Program in Integrative Biotechnology and Translational Medicine, Yonsei University, Incheon 21983, Republic of Korea

Jianmin Wang

The Interdisciplinary Graduate Program in Integrative Biotechnology and Translational Medicine, Yonsei University, Incheon 21983, Republic of Korea

Amir Zeb

Faculty of Natural and Basic Sciences, University of Turbat, Balochistan, Pakistan

Kwang-Hwi Cho

School of Systems Biomedical Science, Soongsil University, Seoul, Republic of Korea

Haiyan Jin

The Interdisciplinary Graduate Program in Integrative Biotechnology and Translational Medicine, Yonsei University, Incheon 21983, Republic of Korea

Jongwan Kim

Bioinformatics and Molecular Design Research Center (BMDRC), Incheon 21983

Department of Biotechnology, Yonsei University, Seoul 03722, Korea

Onju Lee

The Interdisciplinary Graduate Program in Integrative Biotechnology and Translational Medicine, Yonsei University, Incheon 21983, Republic of Korea

the corresponding authors：

Yunyun Wang: wangyunyun91@ntu.edu.cn

School of Pharmacy and Jiangsu Province Key Laboratory for Inflammation and Molecular Drug Target, Nantong University, Nantong 226001, Jiangsu, PR China

Kyoung Tai No: ktno@yonsei.ac.kr

The Interdisciplinary Graduate Program in Integrative Biotechnology and Translational Medicine, Yonsei University, Incheon 21983, Republic of Korea

Jiashun Mao: jiashun_mao@yonsei.ac.kr

The Interdisciplinary Graduate Program in Integrative Biotechnology and Translational Medicine, Yonsei University, Incheon 21983, Republic of Korea