

## Supporting Information

# Potential Inhibitors for Novel Coronavirus Protease Identified by Virtual Screening of 606 Million Compounds

André Fischer, Manuel Sellner, Santhosh Neranjan, Martin Smieško\*, and Markus A. Lill\*

*Computational Pharmacy Group, Department of Pharmaceutical Sciences, Klingelbergstrasse 50, 4056 Basel*

*Corresponding author contact: martin.smiesko@unibas.ch, markus.lill@unibas.ch*

**Abstract:** The rapid outbreak of the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in China followed by its spread around the world poses a serious global concern for public health with almost two million people tested positive and more than hundred thousand of fatalities up to date. To this date, no specific drugs or vaccines are available to treat SARS-CoV-2 despite its close relation to the SARS-CoV-1 virus that caused a similar epidemic in 2003. Thus, there remains an urgent need for the identification and development of specific antiviral therapeutics to conquer SARS-CoV-2. To conquer viral infections, the inhibition of proteases essential for proteolytic processing of viral polyproteins is a conventional therapeutic strategy. In order to find novel inhibitors, we computationally screened a compound library of over 606 million compounds for binding at the recently solved crystal structure of the main protease ( $M^{pro}$ ) of SARS-CoV-2. A screening of such a vast chemical space for SARS-CoV-2  $M^{pro}$  inhibitors has not been reported before. After shape screening, two docking protocols were applied followed by the determination of molecular descriptors relevant for pharmacokinetics to narrow down the number of initial hits. Next, molecular dynamics simulations were conducted to validate the stability of docked binding modes and comprehensively quantify ligand binding energies. After evaluation of off-target binding, we report a list of 12 purchasable compounds, with binding affinity to the target protease that is predicted to be more favorable than that of the cocrystallized peptidomimetic compound. In order to quickly advise ongoing therapeutic intervention for patients, we evaluated approved antiviral drugs and other protease inhibitors to provide a list of 9 compounds for drug repurposing. Furthermore, we identified the natural compounds (-)-taxifolin and rhamnetin as potential inhibitors of  $M^{pro}$ . Rhamnetin is already commercially available in pharmacies.

## Table of Contents

### Supporting Results and Discussion

Virtual Screening Procedures: Figures S1-S5, Tables S1-S3, Pharmacokinetic parameters 2

Final Selection of Compounds: Figures S6-S8, Tables S4-S6 11

### Supporting Materials and Methods

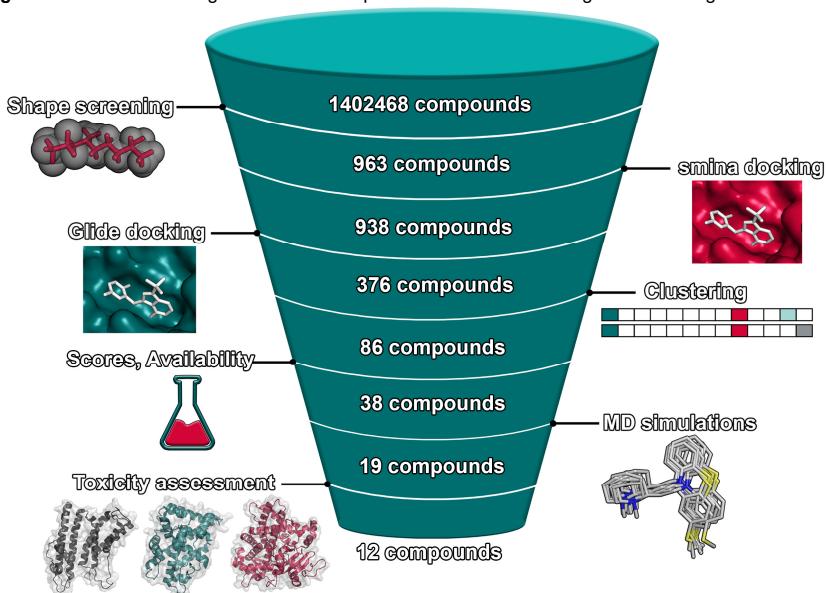
Docking and shape screening: Table S7 13

Supporting References 13

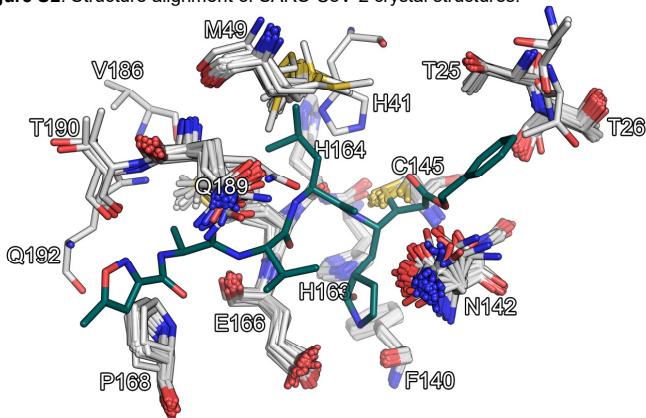
## Supporting Results and Discussion

### Virtual screening procedures

**Figure S1.** Virtual screening workflow for compounds with molecular weight above 500 g/mol.

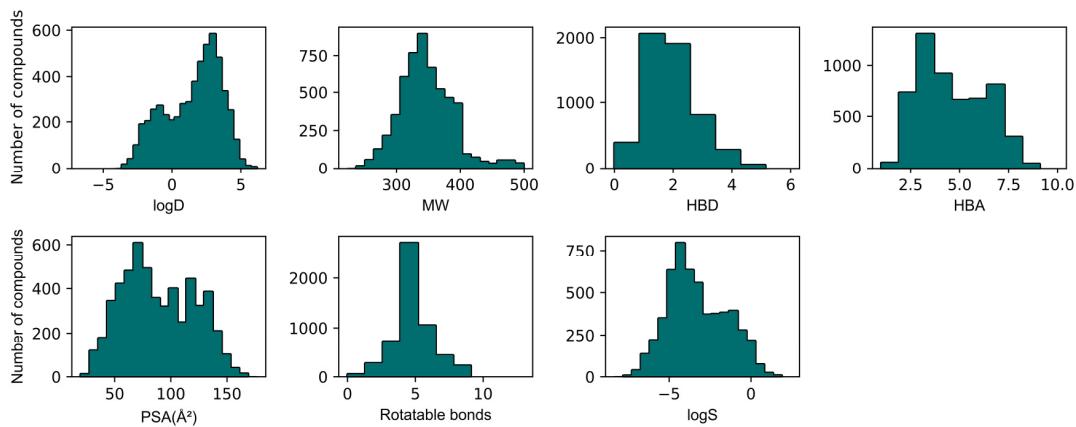


**Figure S2.** Structure alignment of SARS-CoV-2 crystal structures.



The following crystal structures were aligned: 6LU7, 6Y2G, 5M03, 5R84, 5R83, 5R82, 5R80, 5R7Z, Mpro-x0072, Mpro-x0104, Mpro-x0107, Mpro-x0161, Mpro-x0195, Mpro-x0305, Mpro-x0354, Mpro-x0387, Mpro-x0434, Mpro-x0540, Mpro-x0678, Mpro-x0689, Mpro-x0691, Mpro-x0692, Mpro-x0734, Mpro-x0749, Mpro-x0752, Mpro-x0755, Mpro-x0759, Mpro-x0769, Mpro-x0770, Mpro-x0774, Mpro-x0786, Mpro-x0820, Mpro-x0830, Mpro-x0831, Mpro-x0874, Mpro-x0946, Mpro-x0967, Mpro-x0978, Mpro-x0981, Mpro-x0991, Mpro-x0995, Mpro-x1077, Mpro-x1093, Mpro-x1249, Mpro-x1308, Mpro-x1311, Mpro-x1334, Mpro-x1336, Mpro-x1348, Mpro-x1351, Mpro-x1358, Mpro-x1374, Mpro-x1375, Mpro-x1380, Mpro-x1382, Mpro-x1384, Mpro-x1385, Mpro-x1386, Mpro-x1392, Mpro-x1402, Mpro-x1412, Mpro-x1418, Mpro-x1425, Mpro-x1458, Mpro-x1478, and Mpro-x1493. The cocrystallized ligand N3 of the protein structure 6LU7 is shown and was used to determine residues involved in ligand binding.

**Figure S3.** Distribution of pharmacokinetically relevant descriptors for all hits supplied to the Glide SP docking protocol.



Abbreviations: HBA, hydrogen bond donors, HBD, hydrogen bond donors

## Pharmacokinetic parameters

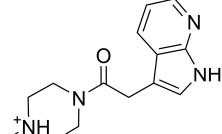
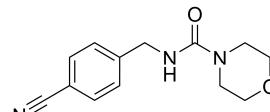
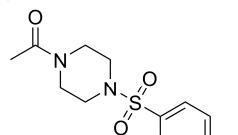
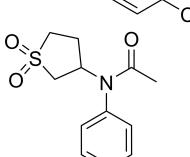
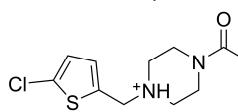
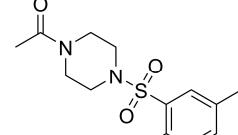
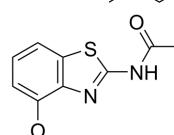
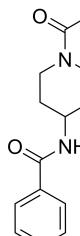
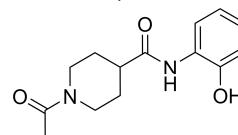
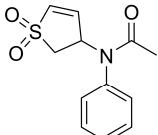
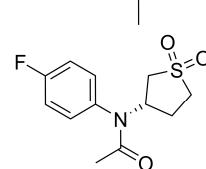
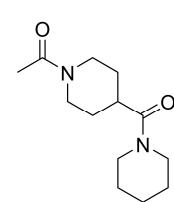
The pharmacokinetic descriptors of these compounds are shown in Figure S3. Due to our selection criterion for extracting compounds from the ZINC database, drug-likeness with respect to molecular weight (MW) was guaranteed. Several compounds, however, violated commonly accepted criteria such as the distribution coefficient logD and polar surface area (PSA). Those compounds were eliminated from further consideration. Due to the catalytic function of the target enzyme, peptides and peptidomimetics are widely applied in targeting proteases.<sup>1</sup> Nonetheless, disadvantages of peptides or peptidomimetics include limited oral bioavailability due to their large MW, PSA, and high number of rotatable bonds as well as poor metabolic stability and higher production cost.<sup>2</sup> Therefore, the development of small-molecules with balanced and favorable pharmacokinetic properties facilitating oral absorption offers a promising alternative.

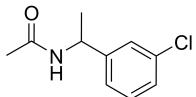
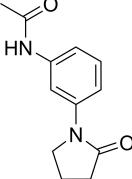
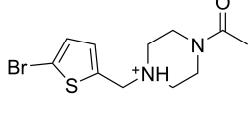
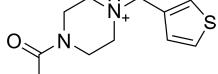
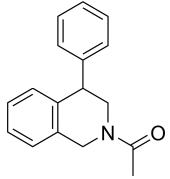
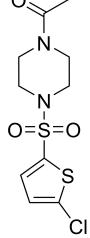
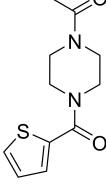
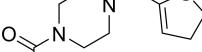
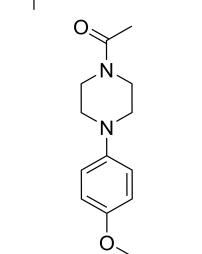
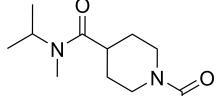
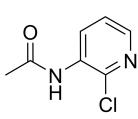
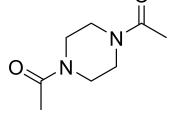
**Table S1.** Redocking of crystallographic ligands.

Crystal structure	Ligand 2D structure	RMSD Glide (Å)	RMSD smina (Å)
Mpro-x0072		4.78	4.94
Mpro-x0104		5.97	2.70
Mpro-x0107		4.25	2.52
Mpro-x0161		3.37	2.21
Mpro-x0195		2.95	2.91
Mpro-x0305		3.49	3.34
Mpro-x0354		5.79	7.40
Mpro-x0387		4.32	5.49
Mpro-x0434		1.25	2.05
Mpro-x0678		8.85	7.62
		2.57	2.04

Mpro-x0689		4.58	5.10
Mpro-x0691		4.62	5.87
Mpro-x0692		2.71	2.68
Mpro-x0734		6.34	5.70
Mpro-x0749		4.20	4.23
Mpro-x0752		8.06	5.44
Mpro-x0755		7.42	6.11
Mpro-x0759		4.51	3.57
Mpro-x0769		4.69	5.98
Mpro-x0770		2.33	2.45

Mpro-x0774		4.38	4.6
Mpro-x0786		9.61	5.63
Mpro-x0820		4.96	6.18
Mpro-x0830		2.71	2.90
Mpro-x0831		2.90	3.12
Mpro-x0874		2.45	3.00
Mpro-x0946		3.13	3.65
Mpro-x0967		2.77	4.15
Mpro-x0978		6.36	2.70
Mpro-x0981		7.54	6.96
Mpro-x0991		4.41	5.71
Mpro-x0995		7.76	2.22
Mpro-x1077		5.06	5.37

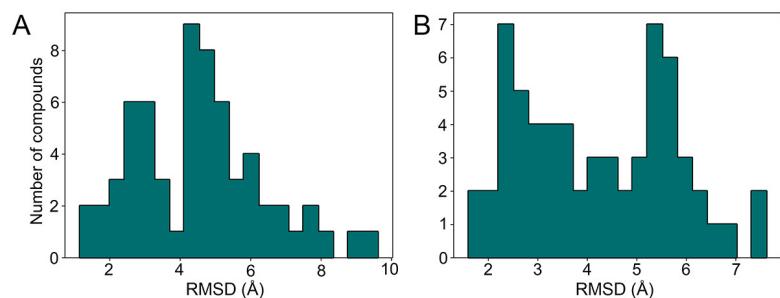
Mpro-x1093		1.76	4.27
Mpro-x1249		4.14	3.43
Mpro-x1308		4.83	6.33
Mpro-x1311		5.02	3.31
Mpro-x1334		4.54	4.91
Mpro-x1336		4.97	5.33
Mpro-x1348		7.02	4.58
Mpro-x1351		5.73	5.81
Mpro-x1358		5.53	5.30
Mpro-x1374		3.09	3.28
Mpro-x1375		4.18	4.61
Mpro-x1380		5.98	5.74

Mpro-x1382		1.69	1.60
Mpro-x1384		6.06	5.23
Mpro-x1385		4.57	5.07
Mpro-x1386		2.25	2.45
Mpro-x1392		4.07	3.73
Mpro-x1402		4.88	5.70
Mpro-x1412		4.28	3.64
Mpro-x1418		1.15	2.44
Mpro-x1425		6.90	6.52
Mpro-x1458		4.96	2.37
Mpro-x1478		3.47	3.97
Mpro-x1493		2.16	5.44

6LU7		3.09	1.62
5R7Z		5.94	2.72
5R81		3.12	3.03
6Y2F		4.97	4.40
6Y7M		2.75	2.33

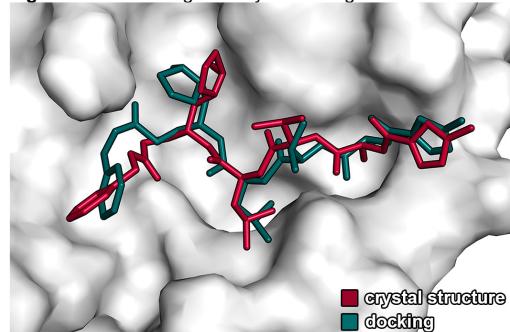
The lowest RMSD of the cocrystallized ligands compared to all five ensemble structures is reported. Crystal structures lacking the standard PDB notation were derived from the diamond webpage.<sup>3</sup>

**Figure S4.** RMSD distribution of redocking crystallographic ligands



(A) RMSD distribution of the Glide SP protocol. (B) RMSD distribution of the smina protocol.

**Figure S5.** Redocking of cocrystallized ligand N3.



The cocrystallized ligand N3 (PDB ID 6LU7) was docked with the smina docking protocol to one of the ensemble structures.





ZINC000097971922	-66.7 ± 6.5	-8.2	-7.9	0.48	584.6	3.0	-4.1	163.0	10	n/a	In-Stock
ZINC000150383182	-64.7 ± 5.7	-7.1	-8.4	0.46	667.8	2.9	-4.5	133.8	16	n/a	In-Stock
ZINC000223764140	-64.7 ± 6.2	-7.5	-7.9	0.47	519.1	7.4	-9.3	58.2	8	n/a	In-Stock
ZINC000036158329	-64.0 ± 6.1	-7.5	-8.3	0.50	504.4	3.4	-6.5	78.5	6	n/a	For-Sale
ZINC000103418421	-60.4 ± 5.1	-8.3	-9.1	0.46	550.6	6.1	-7.4	55.8	5	n/a	For-Sale
ZINC000012412803	-60.0 ± 5.5	-7.3	-8.1	0.47	504.4	3.0	-2.2	74.9	5	n/a	In-Stock
ZINC000003408051	-58.2 ± 4.6	-7.4	-7.2	0.47	533.2	5.5	-7.8	104.0	6	n/a	In-Stock
ZINC000103120251	-58.1 ± 5.6	-7.8	-8.1	0.48	534.4	5.3	-7.0	75.5	6	n/a	For-Sale
ZINC000100004777	-57.4 ± 4.6	-7.7	-8.5	0.49	525.4	7.3	-9.0	61.4	4	n/a	In-Stock
ZINC000102757914	-55.7 ± 4.6	-7.8	-8.3	0.50	539.8	6.7	-9.0	63.0	4	n/a	In-Stock
ZINC000103140857	-55.7 ± 5.3	-7.8	-7.9	0.48	538.4	5.2	-7.3	75.5	6	n/a	For-Sale
ZINC000408583162	-53.2 ± 6.1	-7.3	-7.8	0.45	532.2	7.1	-9.0	42.9	6	n/a	In-Stock
ZINC000100408112	-51.0 ± 7.0	-7.1	-8.5	0.48	569.4	8.6	-9.5	72.5	6	n/a	In-Stock
ZINC000034735622	-47.5 ± 3.1	-7.5	-7.7	0.45	501.9	5.9	-8.1	52.7	6	n/a	For-Sale
ZINC000101601966	-44.0 ± 3.6	-9.2	-9.3	0.45	569.6	-1.0	-0.6	142.2	7	n/a	In-Stock
ZINC000514383436	-42.6 ± 3.9	-7.5	-8.0	0.46	572.6	4.5	-7.1	110.5	10	n/a	On-Demand

Compounds in italics are hits reported in the main text. For some compounds no toxicity assessment was performed (n/a).

<sup>a</sup>Ligand free binding energy predicted by MM/GBSA approach (excluding entropic contributions) with standard deviation;

<sup>b</sup>Lowest docking score from docking against ensemble of five structures;

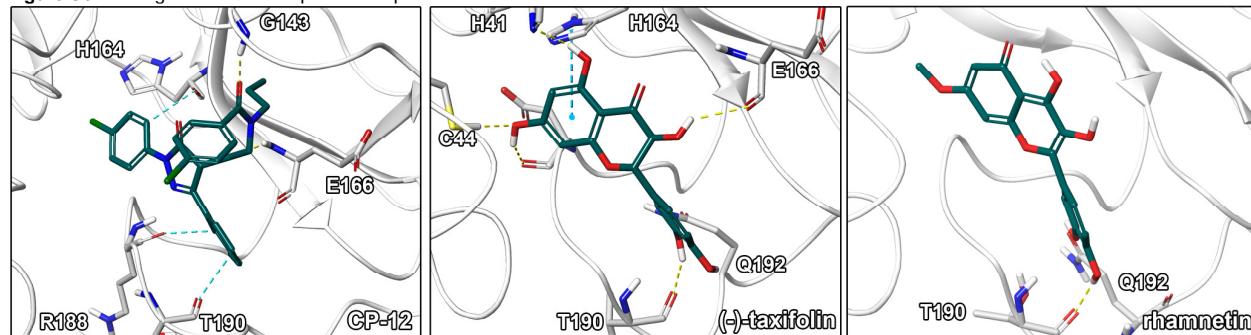
<sup>c</sup>Number of rotatable bonds;

<sup>d</sup>Toxic potential determined in the VirtualToxLab;

<sup>e</sup>Availability of the compounds according to the ZINC database.<sup>4</sup>

## Final Selection of Compounds

**Figure S6.** Binding modes of the reported compounds.



The binding modes were assessed for the MD frame with lowest binding free energy according to MM/GBSA post-processing.

**Table S4.** Number of hydrogen bonds of the proposed compounds.

Compound	Hydrogen bonds
Apixaban	3
Nelfinavir	7
Glecaprevir	4
Lorecivint	5
Rivaroxaban	5
Betrixaban	1
Saquinavir	4
Voxilaprevir	3
Amprenavir	4
Average repurposing	4.0
CP-1	5
CP-2	5
CP-3	3
CP-4	1
CP-5	3
CP-6	5
CP-7	2
CP-8	4
CP-9	3
CP-10	3
CP-11	2
CP-12	2
(-)taxifolin	6
Average shape screen	3.4

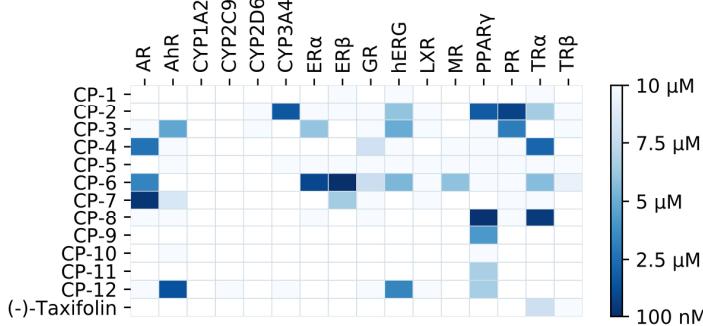
The number of hydrogen bonds was determined for the MD frame with lowest binding free energy.

**Table S5.** Natural compounds determined in main screening.

ZINC ID	Trivial name	$\Delta G$ (kcal/mol) <sup>a</sup>
ZINC00000105082	(-)taxifolin	-53.3 ± 5.1
ZINC000003875620	Rhamnetin	-52.4 ± 3.5
n/a	(+)-taxifolin	-39.5 ± 5.7

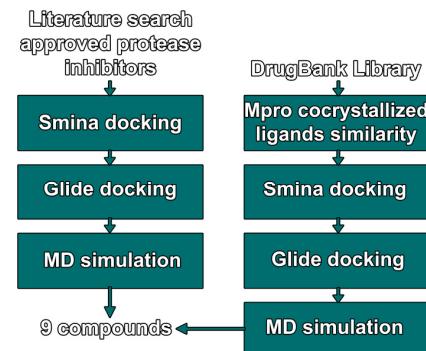
<sup>a</sup>Ligand free binding energy predicted by MM/GBSA approach (excluding entropic contributions) with standard deviation.

**Figure S7.** Off-target binding profile of the reported compounds.



The off-target binding profile was obtained from the VirtualToxLab.<sup>5,6</sup>

**Figure S8.** Schematic depiction of the screening for drug repurposing.



**Table S6.** Results from the screen for drug repurposing.

Trivial Name	DrugBank ID	Indication <sup>a</sup>	$\Delta G$ (kcal/mol) <sup>b</sup>	Glide score (kcal/mol) <sup>c</sup>	Smina Score (kcal/mol) <sup>c</sup>	Origin <sup>d</sup>	Approval status <sup>e</sup>
Apixaban	DB06605	Anticoagulant	-84.0 ± 5.5	-7.7	-8.7	Literature	yes
Nelfinavir	DB00220	Antiviral	-80.6 ± 8.2	-8.4	-9.2	Literature	yes
Glecaprevir	DB13879	Antiviral	-80.3 ± 5.2	-5.9	-9.1	Literature	yes
Lorecavint	DB14883	Inflammation	-79.7 ± 5.0	-7.0	-9.2	Similarity search	no
Rivaroxaban	DB06228	Anticoagulant	-77.2 ± 4.1	-6.8	-7.9	Literature	yes
Betrixaban	DB12364	Anticoagulant	-73.3 ± 3.8	-7.0	-8.1	Literature	yes
Saquinavir	DB01232	Antiviral	-71.5 ± 8.8	-8.2	-8.7	Literature	yes
Voxilaprevir	DB12026	Antiviral	-66.5 ± 4.6	-7.2	-8.3	Literature	yes
Amprenavir	DB00701	Antiviral	-66.5 ± 6.8	-6.7	-7.4	Literature	yes
Telaprevir	DB05521	Antiviral	-63.3 ± 7.7	-7.6	-7.5	Literature	yes
Simeprevir	DB06290	Antiviral	-62.4 ± 3.7	-8.1	-9.0	Literature	yes
Darunavir	DB01264	Antiviral	-61.4 ± 6.9	-7.5	-8.2	Similarity search	yes
n/a	DB07991	Antiestrogen	-61.4 ± 5.1	-7.5	-7.8	Similarity search	no
(R)-boceprevir	DB08873	Antiviral	-60.7 ± 5.8	-7.2	-7.6	Literature	yes
N3	n/a	Antiviral	-59.3 ± 7.6	-9.9	-7.7	Protein Data Bank	no
n/a	DB01810	Cancer	-58.8 ± 4.4	-8.6	-7.3	Similarity search	no
n/a	DB06890	Antifungal	-55.7 ± 3.2	-6.2	-5.9	Similarity search	no
Asunaprevir	DB11586	Antiviral	-55.1 ± 5.8	-6.3	-8.0	Literature	yes
n/a	DB07222	Unknown	-53.6 ± 4.0	-6.6	-6.8	Similarity search	no
Tipranavir	DB00932	Antiviral	-52.8 ± 4.4	-6.0	-8.3	Literature	yes
Indinavir	DB00224	Antiviral	-52.4 ± 6.0	-6.5	-8.5	Literature	yes
Difenpiramide	DB13371	Unknown	-52.0 ± 4.5	-6.2	-7.1	Similarity search	no
Grazoprevir	DB11575	Antiviral	-51.1 ± 7.1	-6.1	-8.9	Literature	yes
Topiroxostat	DB01685	Unknown	-50.5 ± 3.6	-6.4	-7.1	Similarity search	no
(S)-boceprevir	DB08873	Antiviral	-49.7 ± 3.9	-7.4	-7.8	Literature	yes
n/a	DB07781	Unknown	-48.0 ± 3.6	-5.4	-6.4	Similarity search	no
Atazanavir	DB01072	Antiviral	-47.8 ± 9.2	-6.4	-7.3	Literature	yes
Pyroxamide	DB12847	Unknown	-47.3 ± 3.9	-3.0	-6.3	Similarity search	no
MSX-122	DB12715	Unknown	-46.7 ± 4.4	-4.8	-6.5	Similarity search	no
Paritaprevir	DB09297	Antiviral	-46.4 ± 6.0	-8.1	-9.5	Literature	yes
Lopinavir	DB01601	Antiviral	-46.1 ± 8.3	-8.9	-7.8	Literature	yes
Ritonavir	DB00503	Antiviral	-45.8 ± 7.5	-7.5	-8.0	Literature	yes
n/a	DB07152	Unknown	-44.4 ± 4.1	-6.4	-8.8	Similarity search	no
N-acetylserotonin	DB04275	Unknown	-41.3 ± 5.2	-6.6	-6.3	Similarity search	no
n/a	DB04601	Unknown	-38.2 ± 4.1	-6.8	-7.4	Similarity search	no
Dapivirine	DB08639	Antiviral	-36.6 ± 3.0	-6.5	-7.4	Similarity search	no
Dansylamide	DB02866	Unknown	-35.5 ± 4.1	-5.5	-5.9	Similarity search	no
Melatonin	DB01065	Insomnia	-33.7 ± 3.3	-5.9	-5.9	Similarity search	yes
Indane-5-sulfonamide	DB08165	Unknown	-28.2 ± 2.8	-6.3	-5.7	Similarity search	no
Fosamprenavir	DB01319	Antiviral	-25.3 ± 4.3	-6.9	-7.6	Literature	yes
n/a	DB03468	Unknown	-21.9 ± 3.5	-6.2	-5.9	Similarity search	no
n/a	DB07114	Unknown	-14.6 ± 4.8	-5.8	-5.6	Similarity search	no

<sup>a</sup>Indication retrieved from the DrugBank database.<sup>7</sup>

<sup>b</sup>Ligand free binding energy predicted by MM/GBSA approach (excluding entropic contributions) with standard deviation;

<sup>c</sup>Lowest docking score from docking against ensemble of five structures;

<sup>d</sup>Workflow from which the compound was selected for MD;

<sup>e</sup>Approval status of the compound according to the DrugBank database.<sup>7</sup>

## Supporting Materials and Methods

### Docking and shape screening

**Table S7.** SARS-CoV-1 inhibitors derived from the PubChem database for template selection in shape screening.

PubChem Compound ID	Activity type	Activity ( $\mu\text{M}$ )	Method	PubChem bioassay ID	Selection <sup>a</sup>
127045229	$\text{IC}_{50}$	6.7	Fluorimetric assay	1304461	no
127045221	$\text{IC}_{50}$	8.6	Fluorimetric assay	1304461	no
127043644	$\text{IC}_{50}$	6.4	Fluorimetric assay	1304461	no
127043642	$\text{IC}_{50}$	5.8	Fluorimetric assay	1304461	yes
127043641	$\text{IC}_{50}$	6.0	Fluorimetric assay	1304461	no
25256829	$\text{IC}_{50}$	5.4	FRET	417985	yes
45271832	$\text{IC}_{50}$	8.1	FRET	430140	no
45271831	$\text{IC}_{50}$	9.3	FRET	430140	no
45271826	$\text{IC}_{50}$	5.2	FRET	430140	no
45270979	$\text{IC}_{50}$	4.1	FRET	430140	yes
44517748	$\text{IC}_{50}$	2.7	FRET	430140	no
5280343	$\text{IC}_{50}$	8.1	FRET	430140	no
2131982	$K_i$	9.9	FRET	596993	no
2131972	$K_i$	9.1	FRET	596993	yes

<sup>a</sup>Depiction if compound was used in shape screening.

## Supporting References

- (1) Pillaiyar, T.; Manickam, M.; Namasivayam, V.; Hayashi, Y.; Jung, S. H. An Overview of Severe Acute Respiratory Syndrome-Coronavirus (SARS-CoV) 3CL Protease Inhibitors: Peptidomimetics and Small Molecule Chemotherapy. *J. Med. Chem.* **2016**, *59*, 6595–6628.
- (2) Recio, C.; Maione, F.; Iqbal, A. J.; Mascolo, N.; De Feo, V. The Potential Therapeutic Application of Peptides and Peptidomimetics in Cardiovascular Disease. *Front. Pharmacol.* **2017**, *7*, 1–11.
- (3) Diamond. Main protease structure and XChem fragment screen <https://www.diamond.ac.uk/covid-19/for-scientists/Main-protease-structure-and-XChem/Downloads.html> (accessed Mar 20, 2020).
- (4) Sterling, T.; Irwin, J. J. ZINC 15 - Ligand Discovery for Everyone. *J. Chem. Inf. Model.* **2015**, *55*, 2324–2337.
- (5) Vedani, A.; Dobler, M.; Smieško, M. VirtualToxLab - A Platform for Estimating the Toxic Potential of Drugs, Chemicals and Natural Products. *Toxicol. Appl. Pharmacol.* **2012**, *261*, 142–153.
- (6) Vedani, A.; Dobler, M.; Hu, Z.; Smieško, M. OpenVirtualToxLab-A Platform for Generating and Exchanging in Silico Toxicity Data. *Toxicol. Lett.* **2015**, *232*, 519–532.
- (7) Wishart, D. S.; Feunang, Y. D.; Guo, A. C.; Lo, E. J.; Marcu, A.; Grant, J. R.; Sajed, T.; Johnson, D.; Li, C.; Sayeeda, Z.; et al. DrugBank 5.0: A Major Update to the DrugBank Database for 2018. *Nucleic Acids Res.* **2018**, *46*, D1074–D1082.