Team name	PharmAI GmbH | pharm.ai
Team member(s) (firstname lastname; …)	Joachim Haupt; Florian Kaiser; Michael Schroeder
Affiliation	PharmAI GmbH, Tatzberg 47, 01307 Dresden, Germany
Contact email	haupt@pharm.ai
Contact phone number (optional)	
Protein targets (for example: 3CLPro/Nsp5, BoAT1, Fc Receptor, Furin, IL6R, M protein, Nspx, OrfXx, N, E, etc…) | 3 required	SARS-nCoV2 Main Protease (UniProt: P0DTD1) Human Furin (UniProt: P09958) Human ACE2 (UniProt: Q9BYF1)

Section 1: methods & metrics

Describe what methods you have used, how they are independent from one another, what your workflow was, how you performed the cross-correlation between your methods. If applicable, please report estimated performance metrics of your methods, such as accuracy, sensitivity, false-discovery rate, etc., and how those metrics were obtained (e.g. cross-validation). Please provide key references if available.

Methods:

PharmAI uses own proprietary, knowledge-based algorithms that exploit hidden information in protein structures. It is a combination of several methods, including: protein binding site similarity (https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0065894), protein-ligand interaction similarity (https://www.nature.com/articles/s41598-017-11924-4), and sophisticated chemical similarity. The PharmAI *DiscoveryEngine* allows to quickly select from millions of compounds to obtain a ranked list specific for the given target. The results from the independent methods are ranked and combined using an empirical P-value. In a recent benchmark study, this so-called *Focused Library* approach, achieved a hit rate of 6% and identified 7 new lead compounds cGMP-dependent 3',5'-cyclic phosphodiesterase inhibition. For more details see https://www.pharm.ai/2020/04/27/rapid-identification-of-novel-pde2-inhibitor/ and https://www.biorxiv.org/content/10.1101/2020.04.22.021360v. The unique approach of our *DiscoveryEngine* guarantees a maximal scaffold diversification to discover new chemical entities. Please understand that we cannot disclose any more details about the algorithm as this is our main IP.

Section 2: targets

Describe for each protein target: why you chose it, from which source you obtained it (e.g., insidecorona.net / covid.molssi.org / rcsb.org) and why this is the best quality structure, if any pre-processing (e.g., energy minimization, residue correction, alternative folding, …) was performed.

We decided for three independent targets which – if appropriately hit by a small molecule – combat SARS-nCoV2 infection at different molecular mechanism. The *DiscoveryEngine* requires PDB structures as input, which were carefully selected for each target. No pre-processing of the input structures was performed.

Target 1: SARS-nCoV2 Main Protease (https://www.uniprot.org/uniprot/P0DTD1)

The main protease of the virus poses one of the most promising targets for virus inactivation. It is responsible for cleavage of viral protein products post cell entry and replication. If this enzyme can be inhibited, viral reproduction is stopped. We used multiple crystal structures with high resolution available on the Protein Data Bank, e.g. 6lu7 and 6y2f.

Target 2: Human Furin (https://www.uniprot.org/uniprot/P09958)

Another promising target is human furin. Being a ubiquitous endoprotease, it should be druggable with appropriate small molecules. Furin plays an important role in SARS-nCoV2 infection as it cleaves viral spike protein, which allows then for host cell entry. Again, the best high-quality crystal structures for Furin were carefully selected from the Protein Data Bank, e.g. 6hzb or 4ryd.

Target 3: Human Angiotensin-converting enzyme 2 (https://www.uniprot.org/uniprot/Q9BYF1)

Finally, we decided to target the human Angiotensin-converting enzyme 2 (ACE2) as an alternative mode of action for new small molecules. The viral spike protein attaches to ACE2 and thus this target poses the first level of defense which is exploited by SARS-nCoV2. A potential strategy to stop virus infection is to block ACE2 attachment. Because ACE2 is a cell surface protein, small molecules are not required to enter the cell for successful treatment. We used the structure 1r4l as starting point for the *DiscoveryEngine* screening.

Section 3: libraries

Describe which libraries you have used, how they were combined, if any compounds were removed / added, why additions are relevant, any unique features of your library, etc. Please provide the sources you obtained the libraries from (if publicly available). Describe the procedure of data preparation (removal of duplicates, standardization, etc). Indicate if different libraries were used for different targets, and why. If possible, provide a download link to your version of the library.

Library 1:

We composed a library of exactly 5,053,436 compounds, by combining the SWEETLEAD library with your internal Merck library that we requested from Prof. Hermans (https://seafile.unistra.fr/f/eacbff40a914496b98d8/?dl=1). Eventual duplicates were filtered by computing standard InChI Keys, no compounds were removed or added. The composed library was used for all targets and can be found for download here: https://drive.google.com/file/d/1qzXscknfd-k8z2i8VPaFs9w1Cc0yrfGC/view?usp=sharing

Section 4: results

Briefly describe your key findings, any interesting trends in your data, a description of your top 5 compounds for each target. If possible, provide a link to a code and/or data repository. Please do not submit randomly selected compounds!

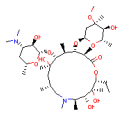
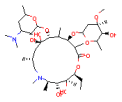
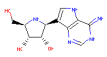
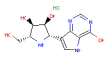
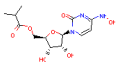
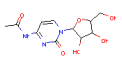
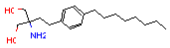
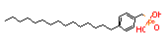
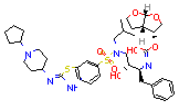
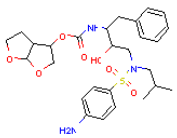
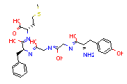
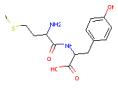
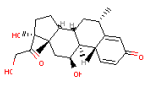
Results:

The top hits for the main protease and Furin show an enrichment of anti-viral compounds. The top hit for Furin is in clinical trial against COVID-19. For ACE2, our top-ranked hit Suramin was recently found to inhibit SARS-CoV-2 infection in cell culture, most likely by binding to ACE2 (https://www.biorxiv.org/content/10.1101/2020.05.06.081968v1).

Target	Rank	InChI Key	Description
SARS-nCoV2 Main Protease	1	AHLPHDHHMVZTML-UHFFFAOYSA-N	DL-Ornithine (Ornithine-containing ribosomal peptides are anti-viral)
	2	AIONOLUJZLIMTK-UHFFFAOYSA-N	(±)-hesperetin (anti-viral)
	3	ALYNCZNDIQEVRV-IDEBNGHGSA-N	4-Aminobenzoic Acid-13C6 (anti-viral)
	4	AYCPARAPKDAOEN-UHFFFAOYSA-N	Anti-Cancer Activity
	5	BDUHCSBCVGXTJM-UHFFFAOYSA-N	Nutlin-3 (Anti-cancer, anti-viral)
Human Furin	1-3	ACTOXUHEUCPTEW-DWDMVGJGSA-N	Spiramycin I and stereoisomers thereof (The similar macrolide Azithromycin are in clinical trials for COVID-19)
	4	CDOJPCSDOXYJJF-CBTAGEKQSA-N	diacetylchitobiose
	5	CMWTZPSULFXXJA-UHFFFAOYSA-N	Naproxen (anti-influenza)
	6	DVZARZBAWHITHR-SFHVURJKSA-O	
	7	FSBIGDSBMBYOPN-UHFFFAOYSA-N	DL-Canavanine (anti-viral)
Human ACE2	1	FIAFUQMPZJWCLV-UHFFFAOYSA-N	Suramin (inhibits SARS-CoV-2 infection)
	2	FQYBTYFKOHPWQT-UHFFFAOYSA-N	
	3	KCSKCIQYNAOBNQ-YBSFLMRUSA-N	Biotin sulfoxide
	4	MXZPGYFBZHBAQM-UHFFFAOYSA-N	
	5	QAQREVBBADEHPA-SSERFMASSA-N	

Other comments:

For the compounds listed as 'in clinical trial' for COVID-19 on DrugBank (https://www.drugbank.ca/covid-19), we identified the closest hits in any of the three targets according to our predictions. The 10 clinical trial drugs with highest similarity to any of the hits are presented in the following table:

URL	Name	Description	DrugBank Structure	Jedi InChI Key	Jedi Structure
https://www.drugbank.ca/drugs/DB00207	Azithromycin	A macrolide antibiotic used to treat a variety of bacterial infections.		MQTOSJVFKKJCRP-CZAPWWKRSA-N	
https://www.drugbank.ca/drugs/DB15660	N4-Hydroxycytidine	A cytidine analog being investigated to treat COVID-19.		NIDVTARKFBZMOT-UHFFFAOYSA-N	
https://www.drugbank.ca/drugs/DB01264	Darunavir	A HIV protease inhibitor used in the treatment of human immunodeficiency virus (HIV) infection in patients with history of prior antiretroviral therapies.		CJBJHOAVZSMMDJ-UHFFFAOYSA-N	
https://www.drugbank.ca/drugs/DB00608	Chloroquine	An antimalarial drug used to treat susceptible infections with P. vivax, P. malariae, P. ovale, and P. falciparum. It is also used for second line treatment for rheumatoid arthritis.		GYFRJEFMZQFVLO-UHFFFAOYSA-N	
https://www.drugbank.ca/drugs/DB11676	Galidesivir	Galidesivir is an adenosine analogue that has been investigated for use against Zaire Ebolavirus. In animal studies, galidesivir was effective in increasing the survival rates from infections caused by various...		WEIAMZKHBCLFOG-QPAIBFMUSA-N	
https://www.drugbank.ca/drugs/DB15661	EIDD-2801	An orally bioavailable isopropylester cytidine analog being investigated to treat COVID-19.		NIDVTARKFBZMOT-UHFFFAOYSA-N	
https://www.drugbank.ca/drugs/DB08868	Fingolimod	A sphingosine 1-phosphate receptor modulator used to treat patients with the relapsing-remitting form of multiple sclerosis (MS) and studied to manage lung complications of COVID-19.		XWVBRTAQHMKHCE-UHFFFAOYSA-N	
https://www.drugbank.ca/drugs/DB15623	TMC-310911	TMC-310911 (also known as ASC-09) is a novel investigational protease inhibitor (PI) that is structurally similar to the currently available darunavir. It is being investigated for use in HIV-1 infections....		CJBJHOAVZSMMDJ-UHFFFAOYSA-N	
https://www.drugbank.ca/drugs/DB12668	Metenkefalin	An investigational endogenous opioid being studied for the treatment of COVID-19.		PESQCPHRXOFIPX-UHFFFAOYSA-N	
https://www.drugbank.ca/drugs/DB00959	Methylprednisolone	A corticosteroid used to treat inflammation or immune reactions across a variety of organ systems, endocrine conditions, and neoplastic diseases.		WHBHBVVOGNECLV-RCSFYLKVSA-N