(Please note that a copy of the information contained in this document is also available in the Microsoft TEAMS NCI CBIIT page.)

1. Common Datasets and Software Suitable for COVID-19 Modeling

Top-5 commonly used databases for small-molecule, natural compounds, biologics and nucleosides.

- **ZINC:** https://zinc15.docking.org/
    - ZINC, a free database of commercially-available compounds for virtual screening. ZINC contains over 230 million purchasable compounds in ready-to-dock, 3D formats. ZINC also contains over 750 million purchasable compounds you can search for analogs in under a minute.
    - Good Database for small-molecule, biologics, natural compounds, marketed drugs and nucleosides
- **PubChem**: https://pubchem.ncbi.nlm.nih.gov/
    - PubChem contains 103 M compounds.
    - Good Database for small-molecule, biologics, natural compounds, marketed drugs and nucleosides
- **SWEETLEAD:** https://simtk.org/projects/sweetlead
    - The SWEETLEAD database has been created to provide an exhaustive and highly curated resource for chemical structures of the world's approved medicines, illegal drugs, and isolates from traditional medicinal herbs
- **ChEMBL**: https://www.ebi.ac.uk/chembl/
    - Contains 2M compounds
    - Good Database for small-molecule, biologics, natural compounds, marketed drugs and nucleosides
- **DrugBank:** https://www.drugbank.ca/
    - contains 13,536 drug entries including 2,630 approved small molecule drugs, 1,372 approved biologics (proteins, peptides, vaccines, and allergenics), 131 nutraceuticals and over 6,358 experimental (discovery-phase) drugs.
    - Good Database for small-molecule, natural compounds, marketed drugs and nucleosides

Databases are recently used for COVID-19 Modeling

DB	DB used/mentioned in the following COVID-19 articles	About the DB
ZINC	Repurposing Therapeutics for COVID-19: Supercomputer-Based Docking to the SARS-CoV-2 Viral Spike Protein and Viral Spike Protein-Human ACE2 Interface (link) Anti-HCV, Nucleotide Inhibitors, Repurposing Against COVID-19 (link) The potential chemical structure of anti-SARS-CoV-2 RNA-dependent RNA polymerase (link) Rapid Identification of Potential Inhibitors of SARS-CoV-2 Main Protease by Deep Docking of 1.3 Billion Compounds (link) Unrevealing Sequence and Structural Features of Novel Coronavirus Using in silico Approaches: The Main Protease as Molecular Target (link)	ZINC is a free Opensource database of commercially available compounds for virtual screening. ZINC contains over 230 million purchasable compounds in ready-to-dock, 3D formats. ZINC also contains over 750 million purchasable compounds you can search for analogs in under a minute.
PubChem	In Silico Screening of Chinese Herbal Medicines With the Potential to Directly Inhibit 2019 Novel Coronavirus (link)	PubChem is an open chemistry database (103 M compounds) at the National Institutes of Health (NIH). "Open" means that you can put your scientific data in PubChem and that others may use it. Since the launch in 2004, PubChem has become a key chemical information resource for scientists, students, and the general public. Each month our website and programmatic services provide data to several million users worldwide. PubChem mostly contains small molecules, but also larger molecules such as nucleotides, carbohydrates, lipids, peptides,

		and chemically-modified macromolecules. We collect information on chemical structures, identifiers, chemical and physical properties, biological activities, patents, health, safety, toxicity data, and many others.
SWEETLEAD	Repurposing Therapeutics for COVID-19: Supercomputer-Based Docking to the SARS-CoV-2 Viral Spike Protein and Viral Spike Protein-Human ACE2 Interface (link)	The SWEETLEAD database has been created to provide an exhaustive and highly curated resource for chemical structures of the world's approved medicines, illegal drugs, and isolates from traditional medicinal herbs. This database has been built using a consensus generating scheme pulling data from several public chemical databases (such as PubChem, ChemSpider, PharmGKB, etc.), as detailed in the publication.

Here is a detailed summary of additional potential molecular databases:

- **SuperDRUG2:** http://cheminfo.charite.de/superdrug2/
    - SuperDRUG2 database is a unique, one-stop resource for approved/marketed drugs, containing more than 4,600 active pharmaceutical ingredients.
- **DRUGCENTRAL:** http://drugcentral.org/
    - DrugCentral is online drug information resource created and maintained by Division of Translational Informatics at University of New Mexico in collaboration with the IDG. DrugCentral provides information on active ingredients chemical entities, pharmaceutical products, drug mode of action, indications, pharmacologic action. It contains ~250,000 compounds
- **Molport:** https://www.molport.com/shop/screeening-compound-database
    - The MolPort database contains data and prices for over 7 million compounds purchasable from stock and over 20 million made-to-order compounds.

Software commonly used (based on the number of publications for COVID-19 host protein interactions):

Docking:

- **AutoDock Vina:** http://vina.scripps.edu/
    - flexible ligand-receptor docking
- **Dock** http://dock.compbio.ucsf.edu

    - predict binding modes of small molecule-protein complexes
    - search databases of ligands for compounds that mimic the inhibitory binding interactions of an experimentally validated inhibitor
    - search databases of ligands for compounds that bind a particular site of a specific protein
    - search databases of ligands for compounds that bind nucleic acid targets
    - examine possible binding orientations of protein-protein and protein-DNA complexes
    - help guide synthetic efforts by examining small molecules that are computationally derivatized
- **idock:** https://github.com/HongjianLi/idock
    - idock is a standalone tool for structure-based virtual screening powered by fast and flexible ligand docking.

MD simulation and protein modeling software:

- **GROMACS:** http://www.gromacs.org/:
    - GROMACS is a versatile package to perform molecular dynamics, i.e. simulate the Newtonian equations of motion for systems with hundreds to millions of particles. It is designed for biochemical molecules like proteins, lipids and nucleic acids that have a lot of complicated bonded interactions.
- **AMBER:** https://ambermd.org/:
    - Amber is a suite of biomolecular MD simulation programs.
- **NAMD:** https://www.ks.uiuc.edu/Research/namd/:
    - Parallel molecular dynamics code designed for high-performance simulation of large biomolecular systems.

Modeling Software:

- **SwissModel:** https://swissmodel.expasy.org:
    - SWISS-MODEL is a fully automated protein structure homology-modelling server, accessible via the ExPASy web server, or from the program DeepView (Swiss Pdb-Viewer). The purpose of this server is to make protein modelling accessible to all life science researchers worldwide

- **Phyre2:** http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index
    - Phyre2 is a suite of tools available on the web to predict and analyze protein structure, function and mutations
- **ROSETTA:** https://www.rosettacommons.org/software
    - Rosetta software suite includes algorithms for computational modeling and analysis of protein structures. It has enabled notable scientific advances in comp. biology, including de novo protein design, enzyme design, ligand docking and structure prediction of biological macromolecules and macromolecular complexes.
- **I-TASSER:** https://zhanglab.ccmb.med.umich.edu/I-TASSER/
    - I-TASSER is a hierarchical approach to protein structure and function prediction. I-TASSER is the top-ranking automatic prediction software in the CASP competition.

2. Potential Lead Compounds (small-molecules/peptide/nucleosides) for Training/modeling

Repurposing Therapeutics for COVID-19:

Repurposing Therapeutics for COVID-19: Supercomputer-Based Docking to the SARS-CoV-2 Viral Spike Protein and Viral Spike Protein-Human ACE2 Interface

https://chemrxiv.org/articles/Repurposing_Therapeutics_for_the_Wuhan_Coronavirus_nCov-2019_Supercomputer-Based_Docking_to_the_Viral_S_Protein_and_Human_ACE2_Interface/11871402

Micholas Dean Smith and Jeremy Smith used Summit (IBM AC922 Summit – Oak Ridge Leadership Computing Facility) to screen a small molecule library (~ 8000 compounds) and carried out Docking and Molecular Dynamics simulations to identify hits that could bind to the main "spike" protein (aka S-protein) of the caronavirus. You can read the article, Repurposing Therapeutics for COVID-19: Supercomputer-Based Docking to the SARS-CoV-2 Viral Spike Protein and Viral Spike…, from The Preprint Server for Chemistry, ChemRxiv. You can watch the following video on how the drugs could bind to virus-proteins and disrupt the virus host binding interactions, https://www.olcf.ornl.gov/wp-content/uploads/2020/03/corona_split_video.mp4?_=1. The video link is part of the following ORNL article, ORNL Team Enlists World's Fastest Supercomputer to Combat the Coronavirus – Oak Ridge Leadership Computing Facility.

What software/Database was used?

- **AutoDock Vina**: http://vina.scripps.edu/
- MD software **GROMACS**: http://www.gromacs.org/

- DB: **SWEETLEAD**: https://simtk.org/projects/sweetlead

What do we have?

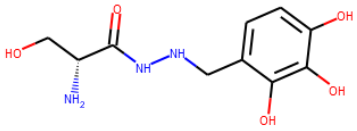
- Ligands (~ 9,120) in AutoDock vina format (pdbqt)
- Receptor structure

ID	smiles	zinc_id	Name
1	Cc1cccn2c(=O)c(-c3nn[nH]n3)cnc12	ZINC000005783214	Permirolast
2	N[C@H](CO)C(=O)NNCc1ccc(O)c(O)c1O	ZINC000003830273	Benserazide
3	O=c1cc(-c2ccc(O)c(O)c2)oc2cc(O)cc(O)c12	ZINC000018185774	Luteolin-monoarbinoside (NP)
4	C/C(=N\NC(=O)c1ccncc1)C(=O)O	ZINC000004974291	Pyruvic-acid-Calcium-isoniazid
5	O=c1c(O)c(-c2ccc(O)c(O)c2)oc2cc(O)cc(O)c12	ZINC000003869685	Quercetol;quercetin (NP)
6	NC(=O)[C@@H]1CCCN1C(=O)[C@H](Cc1c[nH]cn1)NC(=O)[C@@H]1CCC(=O)N1	ZINC000004096261	Protirelin
7	CN1C[C@H](O)C2=C/C(=N/NC(N)=O)C(=O)C=C21	ZINC000100029428	Carbazochrome
8	O=C1CN(/N=C/c2ccc([N+](=O)[O-])o2)C(=O)N1	ZINC000003875368	Nitrofurantoin
9	N[C@H](CO)C(=O)NNCc1ccc(O)c(O)c1O	ZINC000003830273	Benserazide
10	CN1C[C@@H](O)C2=C/C(=N/NC(N)=O)C(=O)C=C21	ZINC000100045148	Carbazochrome
11	C[C@H](O)[C@H](O)[C@H]1CNc2nc(N)[nH]c(=O)c2N1	ZINC000013585233	Sapropterin
12	Nc1ncnc2c1ncn2[C@@H]1O[C@H](CO)[C@@H](O)[C@@H]1O	ZINC000000970363	Vidarbine
13	O=C1C[C@@H](c2ccc(O)c(O)c2)Oc2cc(O)cc(O)c21	ZINC000000058117	NP:Eriodictyol
14	C[C@]1(Cn2ccnn2)[C@H](C(=O)O)N2C(=O)C[C@H]2S1(=O)=O	ZINC000003787060	Tazobactum
15	NC(N)=N/C(N)=N\CCc1ccccc1	ZINC000005851063	Phenformin-hcl
16	CN1C[C@@H](O)C2=C/C(=N/NC(N)=O)C(=O)C=C21	ZINC000100045148	Carbazochrome
17	CN1C[C@@H](O)C2=C/C(=N/NC(N)=O)C(=O)C=C21	ZINC000100045148	Carbazochrome
18	N#C[C@@H]1CCCN1C(=O)CNC12C[C@@H]3C[C@H](CC(O)(C3)C1)C2	ZINC000100003507	Vildagliptin
19	Oc1ccc(C[C@H]2NCCc3cc(O)c(O)cc32)cc1	ZINC000000896041	Demethyl-coclaurine (NP)

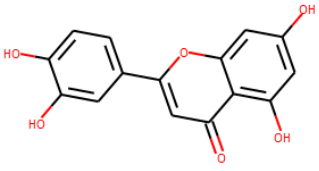
Please note, NP, means Natural Product

Here are the chemical structures of the above 19 compounds

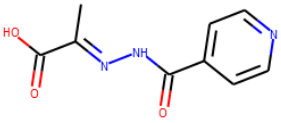
Permirolast



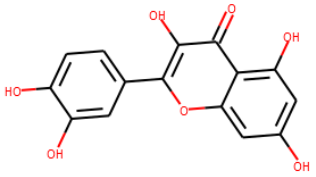
Benserazide



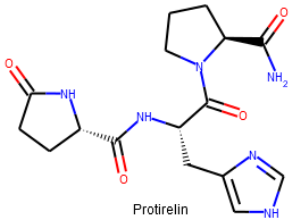
NP-Luteolin-monoarbinoside



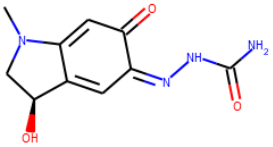
Pyruvic-acid-Calcium-isoniazid



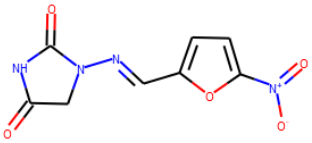
NP-Quercetol;quercetin



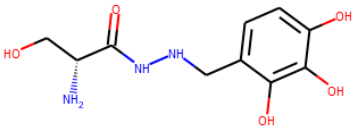
Protirelin



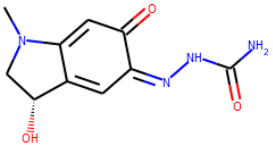
Carbazochrome



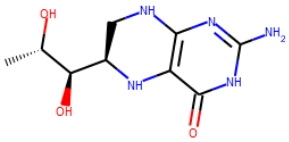
Nitrofurantoin



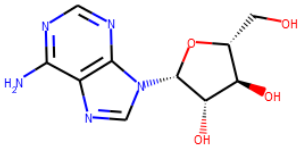
Benserazide



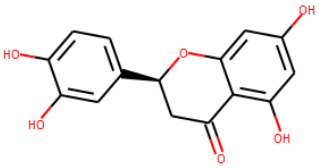
Carbazochrome



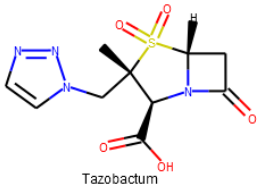
Sapropterin



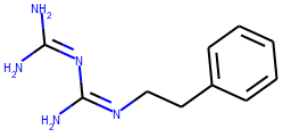
Vidarbine



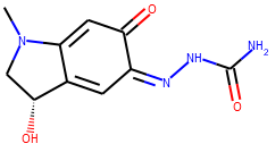
NP:Eriodictyol



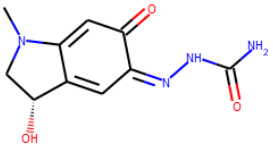
Tazobactum



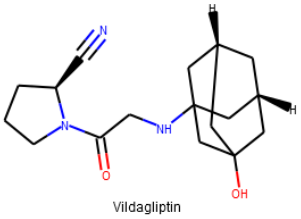
Phenformin-hcl



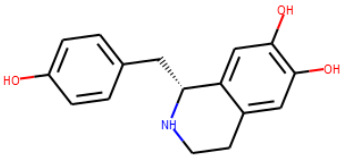
Carbazochrome



Carbazochrome



Vildagliptin



NP:Demethyl-coclaurine

Anti-HCV, Nucleotide Inhibitors, Repurposing Against COVID-19

In this study, sequence analysis, modeling, and docking are used to build a model for Wuhan COVID-19 RdRp. Additionally, the newly emerged Wuhan HCoV RdRp model is targeted by anti-polymerase drugs, including the approved drugs Sofosbuvir and Ribavirin.

Key findings: The results suggest the effectiveness of Sofosbuvir, IDX-184, Ribavirin, and Remidisvir as potent drugs against the newly emerged HCoV disease.

https://pubmed.ncbi.nlm.nih.gov/32119961/

What software was used?

Swiss Model web server is used to build a model for RdRp

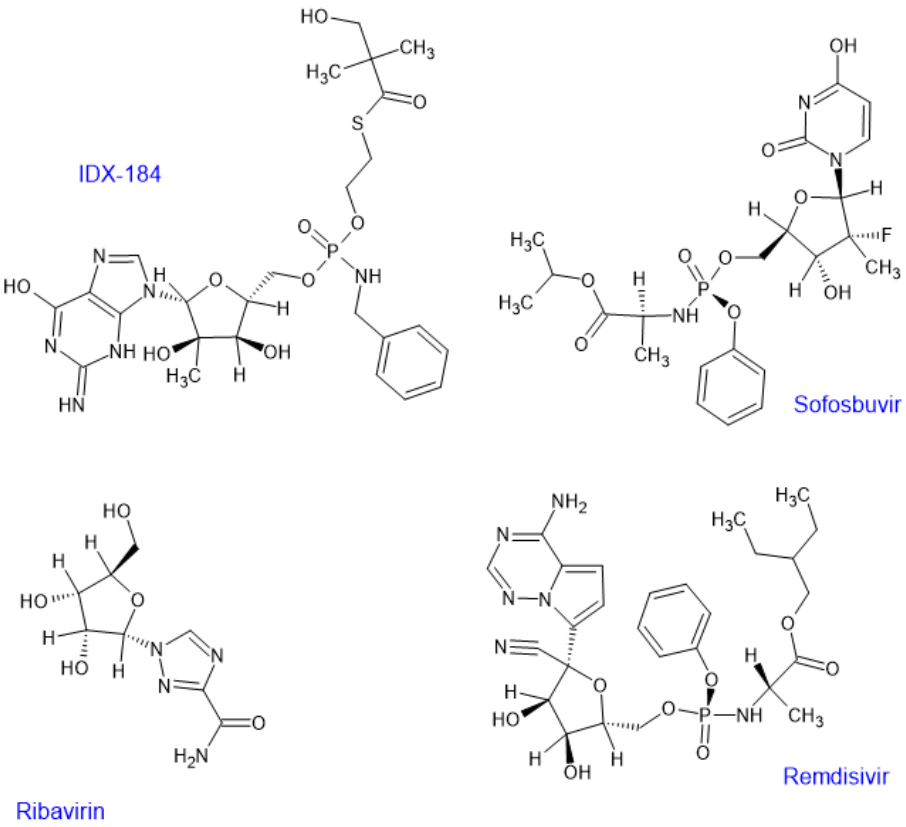
Molprobity web for structure analysis

SCIGRESS, https://www.fqs.pl/en/chemistry/products/scigress, is used to minimize the model and to perform molecular docking experiments.

The minimization of the model is performed using the MM3 force field

AutoDock Vina for docking

These small molecules was selected but this can be downloaded from any compound libraries like PubChem, ZINC, DrugBank etc.



The potential chemical structure of anti-SARS-CoV-2 RNA-dependent RNA polymerase.

RNA-dependent RNA polymerase (RdRp) is an important protease that catalyzes the replication of RNA from RNA template and is an attractive therapeutic target. In this study, we screened these chemical structures from traditional Chinese medicinal compounds proven to show antiviral activity in severe acute respiratory syndrome coronavirus (SARS-CoV) and the similar chemical structures through a molecular docking study to target RdRp of SARS-CoV-2, SARS-CoV, and Middle East respiratory syndrome coronavirus (MERS-CoV).

https://pubmed.ncbi.nlm.nih.gov/32167173

What software was used?

idock for docking, (https://github.com/HongjianLi/idock)

"Achilles" Blind Docking Server, available at: http://bio-hpc.eu/software/blind-docking-server/.

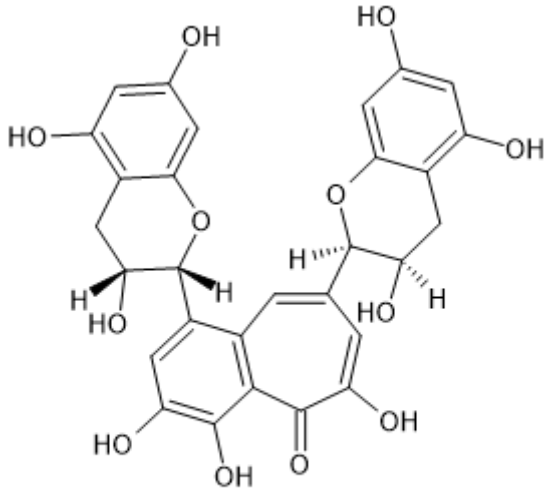
What database was used for ligands?

ZINC

What compounds were used for docking?

Eighty-three chemical structures from traditional Chinese medicinal compounds and their similar structures were retrieved from ZINC15 database.

(Ravi: It is not clear what these 80 Chinese Medicinal Compounds are)



Theaflavin

Theaflavin was found to be the potent compound

In Silico Screening of Chinese Herbal Medicines With the Potential to Directly Inhibit 2019 Novel Coronavirus

Objective: In this study we execute a rational screen to identify Chinese medical herbs that are commonly used in treating viral respiratory infections and also contain compounds that might directly inhibit 2019 novel coronavirus (2019-nCoV), an ongoing novel coronavirus that causes pneumonia.

Results: Of the natural compounds screened, 13 that exist in traditional Chinese medicines were also found to have potential anti-2019-nCoV activity. Further, 125 Chinese herbs were found to contain 2 or more of these 13 compounds. Of these 125 herbs, 26 are classically catalogued as treating viral respiratory infections. Network pharmacology analysis predicted that the general in vivo roles of these 26 herbal plants were related to regulating viral infection, immune/inflammation reactions and hypoxia response.

https://pubmed.ncbi.nlm.nih.gov/32113846/

What software was used?
AutoDock (v4.0) and **SwissModel** for Docking and Modeling respectively.

What DB was used for small-molecules?

PubChem: https://pubchem.ncbi.nlm.nih.gov/

These molecules were reported to inhibit viral entry, and were docked with spike proteins

Table 1. The molecules and their docking proteins, binding energy (kcal/mol).

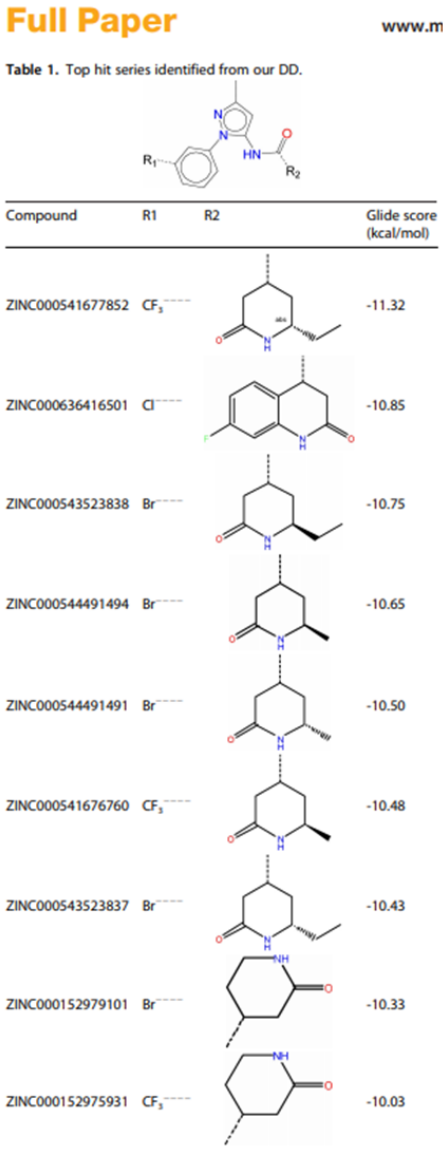
No.	Molecular name	Targets or inhibition	Reference	Docking (binding energy)		
				PLpro	3CLpro	Spike
M1	Betulinic acid	Replication, 3CLpro	[16]	Undo	−4.23	Undo
M2	Coumaroyltyramine	PLpro and 3CLpro	[11], [20]	−3.22	−4.18	Undo
M3	Cryptotanshinone	PLpro and 3CLpro	[18]	−5.25	−6.23	Undo
M4	Desmethoxyreserpine	Replication, 3CLpro, and entry	[6]	Undo	−3.52	Undo
M5	Dihomo-γ-linolenic acid	3CLpro	[7]	Undo	−3.88	Undo
M6	Dihydrotanshinone I	Entry, and spike protein	[28]	Undo	Undo	−5.16
M7	Kaempferol	PLpro and 3CLpro	[11]	−2.15	−6.01	Undo
M8	Lignan	Replication, 3CLpro	[16]	Undo	−4.27	Undo
M9	Moupinamide	PLpro	[20]	−3.05	Undo	Undo
M10	N-cis-feruloyltyramine	PLpro and 3CLpro	[11], [20]	−3.11	−4.31	Undo
M11	Quercetin	PLpro and 3CLpro	[20]	−4.62	−6.25	Undo
M12	Sugiol	Replication, 3CLpro	[16]	Undo	−6.04	Undo
M13	Tanshinone IIa	PLpro and 3CLpro	[18]	−5.02	−5.17	Undo

3CLpro: 3C-like protease; PLpro: papain-like protease.

Rapid Identification of Potential Inhibitors of SARS-CoV-2 Main Protease by Deep Docking of 1.3 Billion Compounds

https://pubmed.ncbi.nlm.nih.gov/32162456/

In the current study we applied DD to all 1.3 billion compounds from ZINC15 library to identify top 1,000 potential ligands for SARS-CoV-2 Mpro protein. The compounds are made publicly available for further characterization and development by scientific community.



Full Paper www.m

Table 1. Top hit series identified from our DD.

Compound	R1	R2	Glide score (kcal/mol)
ZINC000541677852	CF_3		-11.32
ZINC000636416501	Cl		-10.85
ZINC000543523838	Br		-10.75
ZINC000544491494	Br		-10.65
ZINC000544491491	Br		-10.50
ZINC000541676760	CF_3		-10.48
ZINC000543523837	Br		-10.43
ZINC000152979101	Br		-10.33
ZINC000152975931	CF_3		-10.03

Table 1. continued

Compound	R1	R2	Glide score (kcal/mol)
ZINC001627499877	Br		-9.32
ZINC001362111980	Cl		-9.13

Computers and viral diseases. Preliminary bioinformatics studies on the design of a synthetic vaccine and a preventative peptidomimetic antagonist against the SARS-CoV-2 (2019-nCoV, COVID-19) coronavirus

The goal of this work was to find a short section or sections of viral protein sequence suitable for preliminary design proposal for a peptide synthetic vaccine and a peptidomimetic therapeutic, and to explore some design possibilities

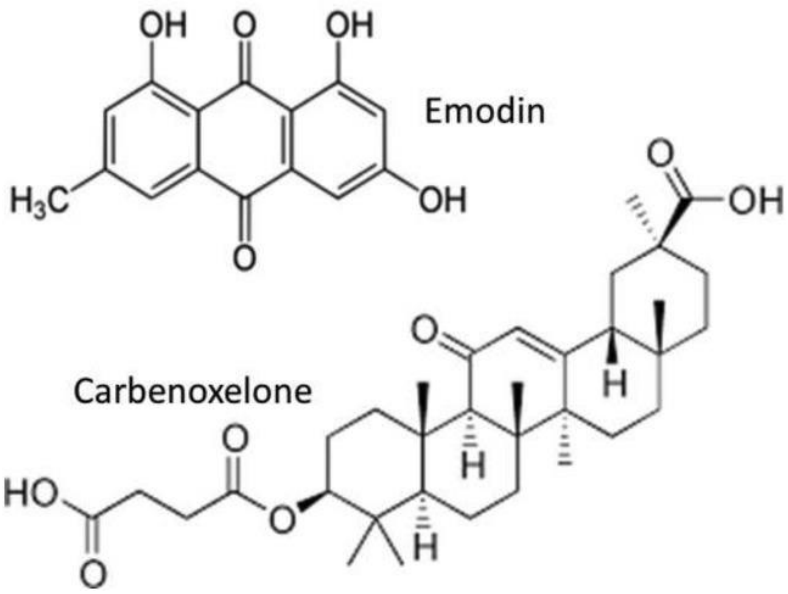
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7094376/

From the point of synthesis of a peptide as a plausible analogue of an immunogenic part (“epitope”) of a protein for development of diagnostics and the peptide of interest is:

(NH_3^+)-GPSKRSFIEDLLFNKVTLAC-(COO^-)

The rationale is that the section KRSFIEDLLFNKV is exposed as associated with S2′ at the surface but highly conserved as shown in the second (i.e. “FIEDLL”) alignment in Section 4.3.

Potential small molecule inhibitors



The first-in-class peptide binder to the SARS-CoV-2 spike protein

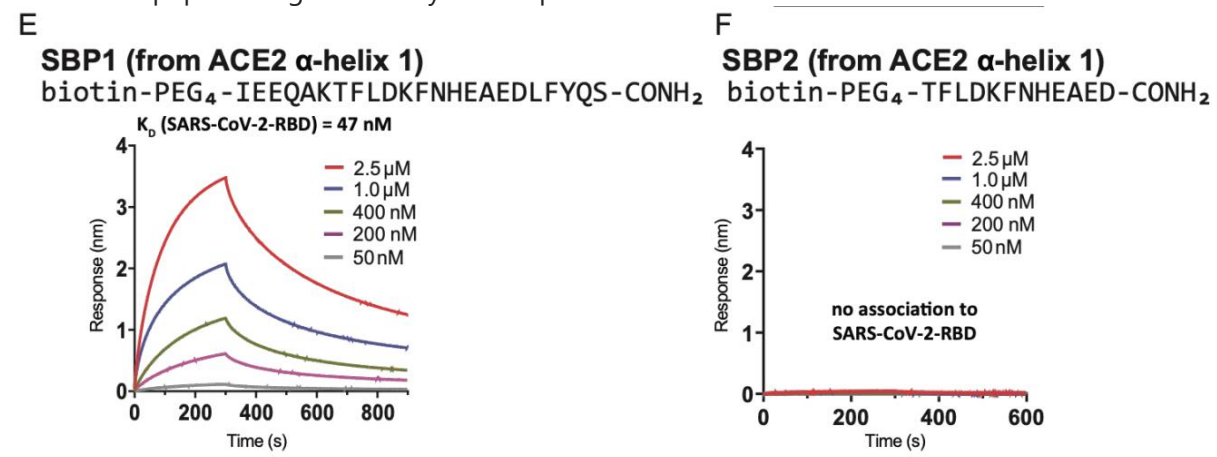
https://www.biorxiv.org/content/10.1101/2020.03.19.999318v1.article-info

Using molecular dynamics simulations based on the recently solved ACE2 and SARS38 CoV-2-RBD co-crystal structure, we observed that the ACE2 peptidase domain (PD) α1 helix is important for binding SARS-CoV-2-RBD. Using automated fast-flow peptide synthesis, we chemically synthesized a 23-mer peptide fragment of the ACE2 PD α1 helix composed entirely of proteinogenic amino acids. Chemical synthesis of this human derived sequence was complete in 1.5 hours and after work up and isolation >20 milligrams of pure material was obtained. Bio-layer 43 interferometry revealed that this peptide specifically associates with the SARS-CoV-2-RBD with low nanomolar affinity. This peptide binder to SARS-CoV-2-RBD provides new avenues for COVID-19 treatment and diagnostic modalities

by blocking the SARS-CoV-2 spike protein interaction with ACE2 and thus precluding virus entry into human cells.

Software used: MD NAMD and LC-MS experiments

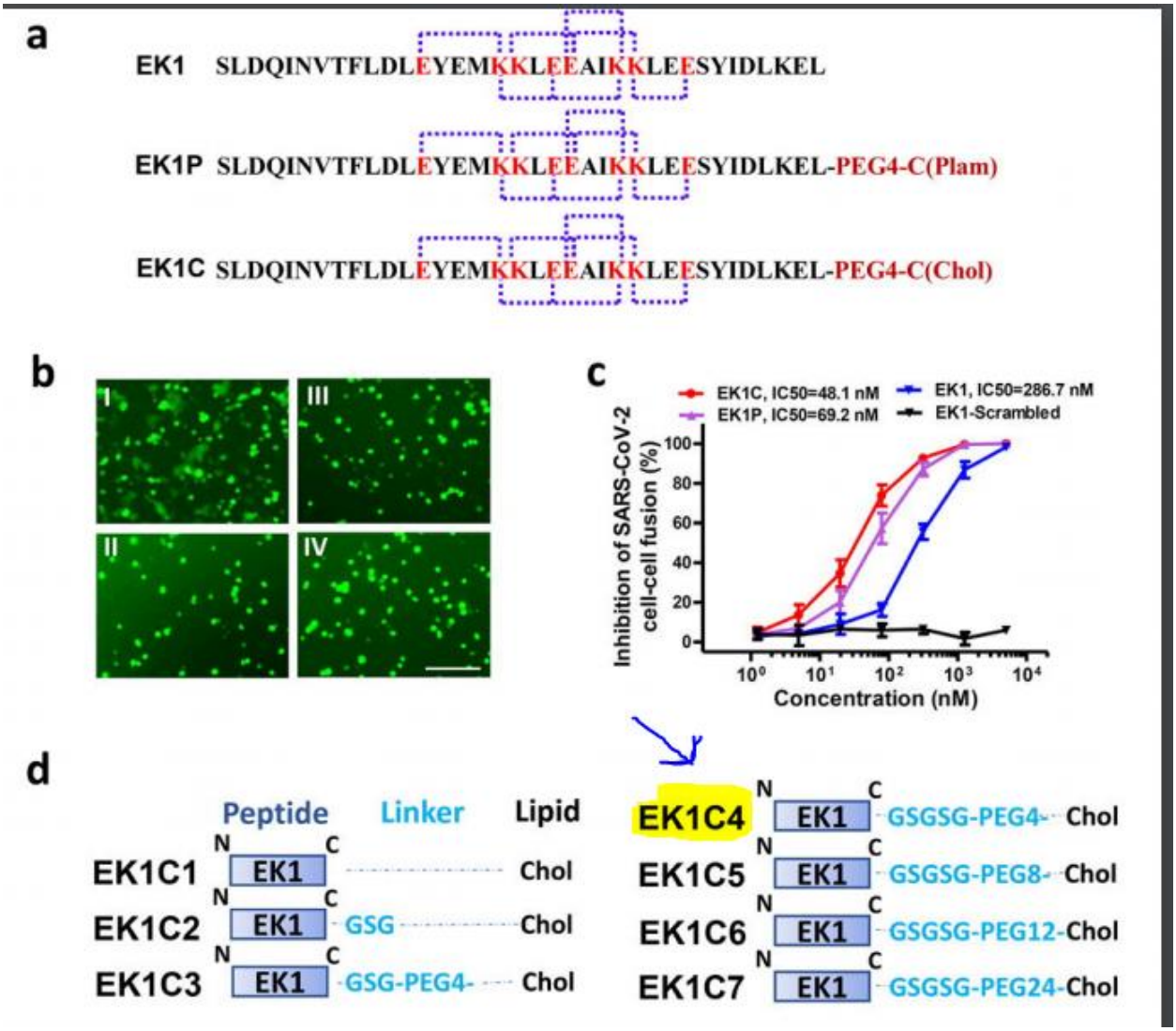
Here are the peptide fragments they have reported.



Inhibition of SARS-CoV-2 infection (previously 2019-nCoV) by a highly potent pan-coronavirus fusion inhibitor targeting its spike protein that harbors a high capacity to mediate membrane fusion

https://www.biorxiv.org/content/10.1101/2020.03.09.983247v1.article-info

we found that one of the lipopeptides, EK1C, exhibited highly 115 potent inhibitory activity against SARS-CoV-2 S-mediated membrane fusion and PsV 116 infection, about 240- and 150-fold more potent than EK1 peptide, respectively



Unrevealing Sequence and Structural Features of Novel Coronavirus Using in silico Approach

https://pubmed.ncbi.nlm.nih.gov/32210741

Our results showed that several HIV inhibitors such as lopinavir, ritonavir, and saquinavir produce strong interaction with the active site of SARS-CoV-2 main protease. Furthermore, broad library protease inhibitors obtained from PubChem and ZINC (www.zinc.docking.org) were evaluated. Our analysis revealed 20 compounds that could be clustered into three groups based on their chemical features. Then, these structures could serve as leading compounds to develop a series of derivatives optimizing their activity against SARS-CoV-2 and other coronaviruses.

What DB was used?

In order to contribute with further studies related to developing more effective drugs, in this work was evaluated a broad library of protease inhibitors available in the ZINC database (over 100 compounds) and PubChem (over 200 compounds).

What Software was used?
Achilles Blind Docking server: https://bio-hpc.ucam.edu/achilles/
VINA: http://vina.scripps.edu/
VegaZZ 3.1.0.21: https://www.ddl.unimi.it/cms/index.php?Software_projects:VEGA_ZZ
NAMD (for MD simulations): https://www.ks.uiuc.edu/Research/namd/

Binding energy (BE) values of the best 20 compounds selected as potential inhibitors of SARS-CoV-2 protease. Compound structures were obtained from ZINC database (Bold) and PubChem (italics), lowest BE compound of each group are shown in red.

Compound ID	BE (kcal/mol)	Molecular formula	MW	F	S	Br	Cl	RNH2	R2NH	R3N	ROH	RCOR	RCOOR	ROR	RINGS	AROMATIC
213039	-8,1	C27H37N3O7S	547,6636	0	1	0	0	1	1	0	1	0	1	2	4	2
4369144	-8,2	C30H34N6O3	526,6294	0	0	0	0	0	1	1	0	0	0	0	5	3
444214	-8,1	C28H31F5N4O4	582,5622	5	0	0	0	0	3	1	1	0	0	0	4	3
444603	-8,7	C25H34N6O5S	530,6397	0	1	0	0	0	1	2	1	0	0	0	4	2
444743	-8,3	C23H29BrN6O5S	581,4826	0	1	1	0	0	1	1	0	0	0	0	4	2
444745	-9,3	C31H35N7O3	553,6547	0	0	0	0	0	2	1	0	0	0	0	6	4
444756	-8	C37H55ClN8O5	727,3362	0	0	0	1	0	2	4	2	0	0	0	5	2
457409	-8	C28H36N2O9S	576,6584	0	1	0	0	0	1	0	1	0	1	4	5	2
71627394	-7,9	C28H37N3O8S	575,6737	0	1	0	0	0	1	0	1	0	1	2	4	2
9543419	-8,1	C30H39N7O6	593,674	0	0	0	0	0	2	1	1	0	0	0	2	1
ZINC001014061065	-7,6	C15H17N5O4	331,33	0	0	0	0	0	1	1	1	0	0	0	3	2
ZINC001014061063	-7,6	C18H28N4O3	348,44	0	0	0	0	0	1	1	0	0	0	0	3	1
ZINC001014061081	-8,7	C16H17N7O2	339,35	0	0	0	0	0	1	1	0	0	0	0	4	3
ZINC000923097446	-7,6	C17H21N5O3	343,38	0	0	0	0	0	1	2	0	1	0	1	3	2
ZINC001014061043	-7,6	C18H22N4O3	342,39	0	0	0	0	0	1	1	0	0	0	0	3	2
ZINC001014061084	-8,1	C16H17N7O2	339,35	0	0	0	0	0	1	1	0	0	0	0	4	3
ZINC001014061083	-8,3	C16H17N7O2	339,35	0	0	0	0	0	1	1	0	0	0	0	4	3
ZINC000923097334	-7,7	C17H18N4O2	310,35	0	0	0	0	0	1	1	0	1	0	0	3	3
ZINC001014061082	-7,7	C16H17N7O2	339,35	0	0	0	0	0	1	1	0	0	0	0	4	3
ZINC001435413766	-7,7	C20H24FN3O2	357,42	1	0	0	0	0	0	1	0	0	1	0	4	2
ZINC001014061061	-7,8	C18H28N4O3	348,44	0	0	0	0	0	1	1	0	0	0	0	3	1

F= fluoride, S= sulphurous, Br= bromide, Cl=chloride, RNH2= primary amine, R2NH= secondary amine, R3N= tertiary amine, ROH= alcohol, RCOR= ketone, RCOOR= ester, ROR= ether

Protease inhibitors also identified in the paper

Protease inhibitor	Binding energy (Kcal/mol) for SARS CoV-2 protease	Binding energy (Kcal/mol) for SARS CoV protease	Binding energy (Kcal/mol) for HIV-1 protease
Saquinavir	-9.6	-8.1	-9.7
Lopinavir	-9.1	-8.4	-11.4
Tipranavir	-8.7	-7.7	-10.3
Darunavir	-8.2	-6.7	-10.9
Amprenavir	-7.6	-7.3	-9.2
Atazanavir	-7.2	-7	-9.8
Ritonavir	-6.9	-7.1	-9.4

3. Structure-based/peptide Modeling COVID-19

Emerging WuHan (COVID-19) coronavirus: glycan shield and structure prediction of spike glycoprotein and its interaction with human CD26section

The study modelled homo-trimer structure of COVID-19 spike glycoprotein in both closed (ligand-free) and open (ligand-bound) conformation, which is involved in host cell adhesion
https://pubmed.ncbi.nlm.nih.gov/32178593/

Modeling

- **SwissModel:** https://swissmodel.expasy.org/

Protein-Protein docking:

- **Cluspro:** https://cluspro.bu.edu/login.php

COVID-19 spike-host cell receptor GRP78 binding site prediction

https://pubmed.ncbi.nlm.nih.gov/32169481/

- DOI: 10.1016/j.jinf.2020.02.026

The study has modeled the COVID-19 spike binding site to the cell-surface receptor (Glucose Regulated Protein 78 (GRP78)) is predicted using combined molecular modeling docking and structural bioinformatics. The COVID-19 spike protein is modeled using its counterpart, the SARS spike.

2019-nCoV (Wuhan virus), a novel Coronavirus: human-to-human transmission, travel-related cases, and vaccine readiness

https://pubmed.ncbi.nlm.nih.gov/32088679/

Protein-protein docking of Receptor Binding Domain, Phylogenetic analysis and docking

Software used:

- **Phyre2** was used for protein modeling
- **Haddock2** was used for docking

The potential chemical structure of anti-SARS-CoV-2 RNA-dependent RNA polymerase.

https://pubmed.ncbi.nlm.nih.gov/32167173

RNA-dependent RNA polymerase (RdRp) is an important protease that catalyzes the replication of RNA from RNA template and is an attractive therapeutic target. In this study, we screened these chemical structures from traditional Chinese medicinal compounds proven to show antiviral activity in severe acute respiratory syndrome coronavirus (SARS-CoV) and the similar chemical structures through a molecular docking study to target RdRp of SARS-CoV-2, SARS-CoV, and Middle East respiratory syndrome coronavirus (MERS-CoV).

Simulation of the Clinical and Pathological Manifestations of Coronavirus Disease 2019 (COVID-19) in Golden Syrian Hamster Model: Implications for Disease Pathogenesis and Transmissibility

https://pubmed.ncbi.nlm.nih.gov/32215622/

Molecular docking on the binding between angiotensin-converting enzyme 2 (ACE2) of common laboratory mammals and the receptor-binding domain of the surface spike protein of SARS-CoV-2 suggested that the golden Syrian hamster is an option.

Software Used:

I-TASSER: https://zhanglab.ccmb.med.umich.edu/I-TASSER/

Rosetta: https://www.rosettacommons.org/software

Coronavirus treatment: Vaccines/drugs in the pipeline for COVID-19

https://www.clinicaltrialsarena.com/analysis/coronavirus-mers-cov-drugs/