DETERMINING A PHARMACEUTICAL HIT FOR CORONAVIRUS-19 PROTEASE USING DOCKING STUDIES AND MOLECULAR SCANS

Kerin Ingegneri

ABSTRACT

The spread of Coronavirus-19 caused a global pandemic resulting in roughly 3.14 million deaths this past year. Our efforts here are to design an antiviral drug using molecular docking studies to target the main protease enzyme used by the virus. This is a known pharmaceutical target meaning the virus can not replicate without this enzyme. There are no known human homologues of this protease thus reducing potential side effects. By employing computer software systems, we have generated a model to produce docking studies using six different criteria evaluating the virtual compounds. The virtual compounds that we employ are drug like and similar in chemical moieties to known inhibitors. The goal is to dock structures readily available to purchase and test in vitro. Then using a pivot table from excel, the duplicates of the virtual compounds with the binding criteria are revealed. These docking studies reveal how tight the virtual compounds are binding at the active site along with structural kinetic data and the end goal is to find that pharmacological hit.

PHARMACEUTICAL HIT INTRODUCTION

- Molecular docking studies with crystal structure and virtual compounds

- Generates a binding energy ΔG value looking at molecular interactions with virtual compound and active site

- Virtual compound given ten different poses/conformations evaluated via six different criteria

- Top 5% pasted into Excel using Pivot table look for duplicates

PHARMACEUTICAL HIT INTRODUCTIONS

- Generate pose at active site further evaluate, molecular interactions

- Buy compounds

- Test on enzyme

- Optimize hit using virtual compounds

- Rational Drug design Structure Activity Relationships (SAR)

CORONAVIRUS-19
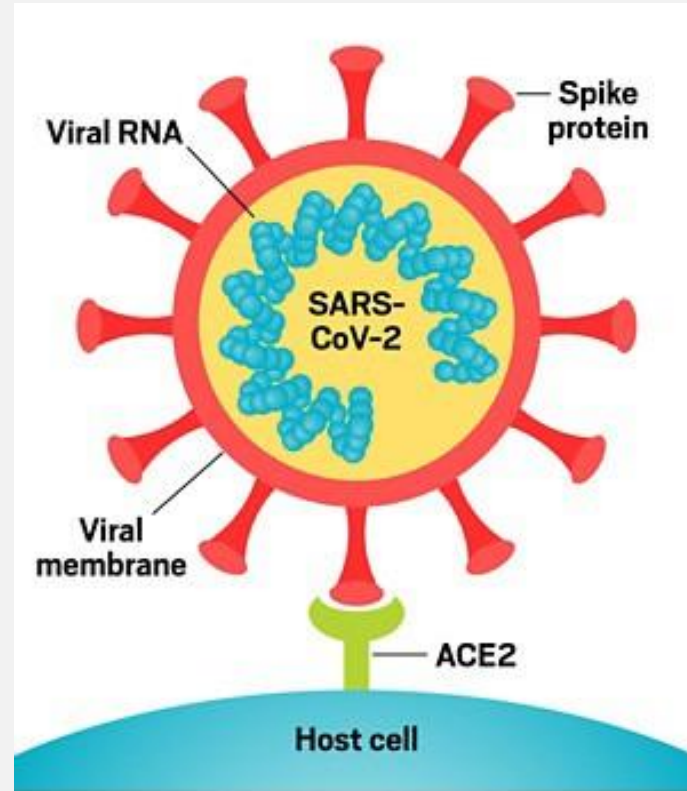
- Objective is to inhibit protease enzyme that generates peptides that are vital to the replication of the virus

CORONAVORUS-19

✓ Protease -A type of enzyme that breaks down proteins into smaller proteins or smaller protein units, such as peptides or amino acids.

✓ Known pharmaceutical target

✓ The protease is essential, but has _no human homologues_

✓ So inhibitors of the protease have less of a chance of hitting a human protease

CORONAVIRUS-19

✓ This enzyme processes a polyprotein chain coded by the virus's RNA, chopping up the chain into functional proteins that the virus then uses to assemble itself and multiply.

✓ Disrupting this key piece of the virus's self-replication machinery could bring an infection screeching to a halt
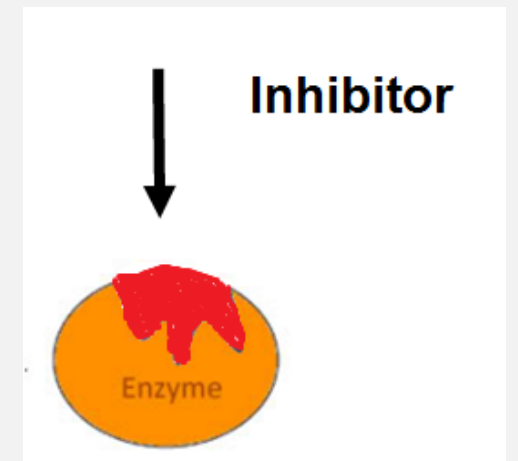
ENZYMES

The Lock and Key Theory of Enzymes and Substrates

Substrate= "Key"   Enzyme= "Lock"   Active Site= "Key hole"
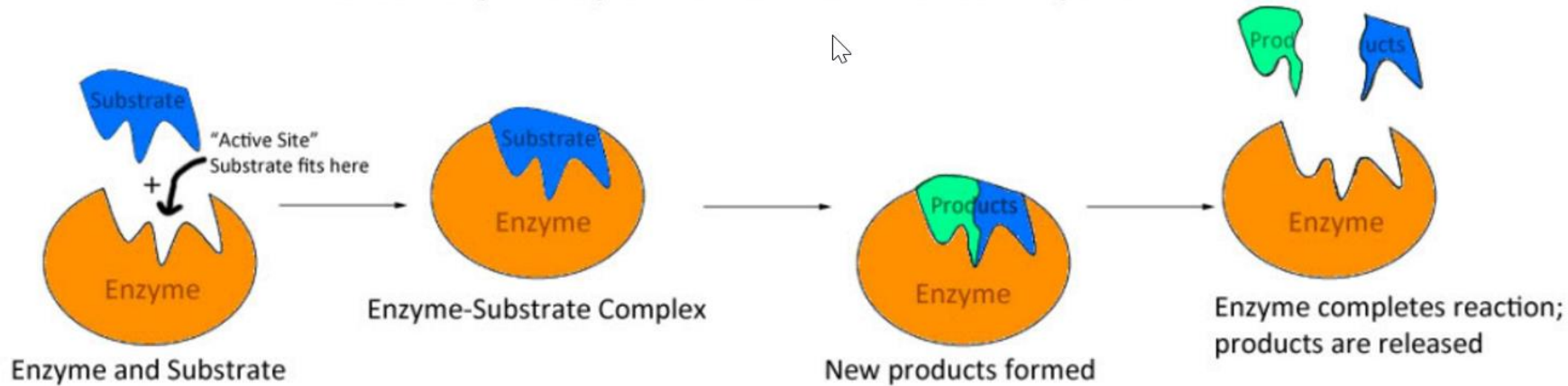
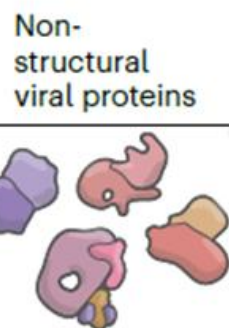✓Bind molecule/inhibitor to stop enzymes function treat patient
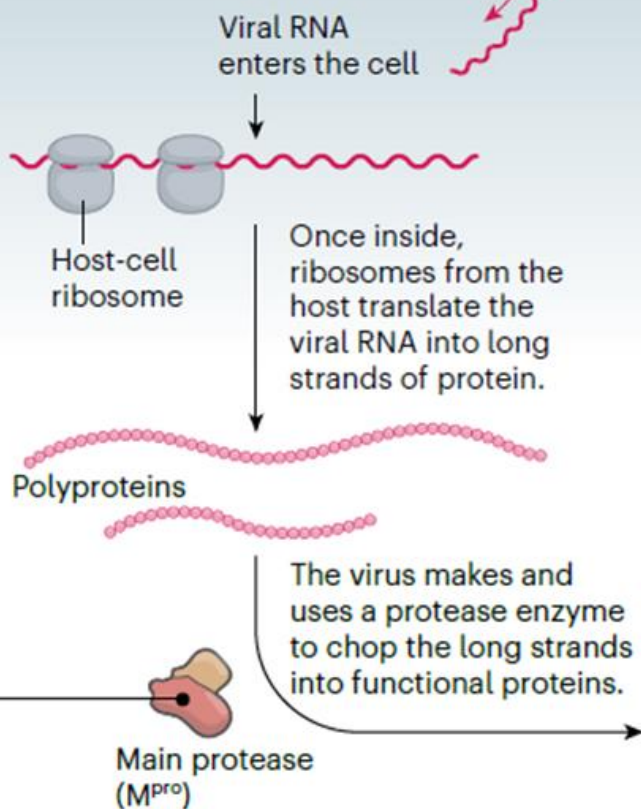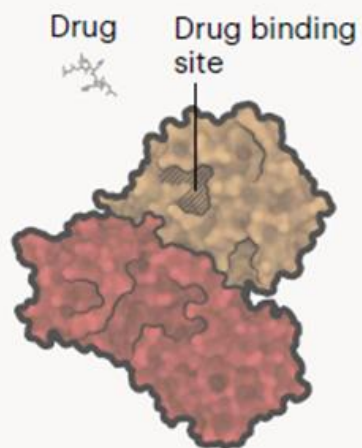
✓Valid pharmaceutical target
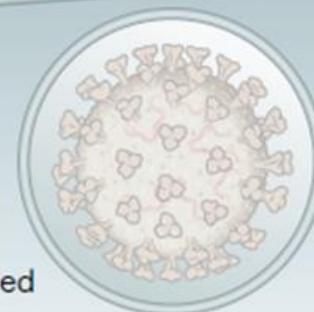
BREAKING THE CYCLE

Once the virus has fused with a host cell, the virus injects genetic material and uses the host machinery to make copies of itself. Many teams are studying viral proteins involved in replication.

Human cell

Targeting M^{pro}
If the virus can't build its main components, it can't replicate. Researchers are trying to find drugs that block the cutting action of M^{pro} by fitting into its two active sites – little dimples on the surface.

Drug   Drug binding site

Viral RNA enters the cell

Host-cell ribosome

Once inside, ribosomes from the host translate the viral RNA into long strands of protein.

Polyproteins

The virus makes and uses a protease enzyme to chop the long strands into functional proteins.

Main protease (M^{pro})

Non-structural viral proteins

RdRp produces viral RNA and RNA that is translated into new virus proteins by the host.

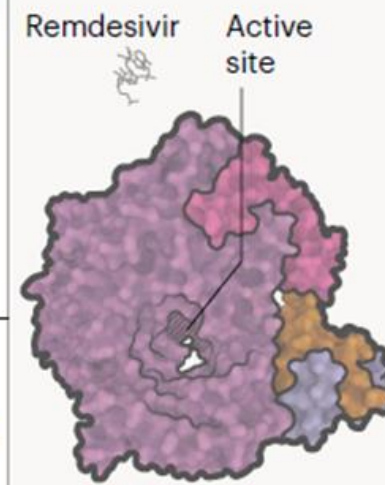Proteins involved in replication and transcription form a complex within a host vesicle.

New virus packaged and released

Targeting RdRp
This enzyme makes copies of the full viral genome. Several novel compounds and approved drugs, such as remdesivir, bind to its active site.

Remdesivir   Active site

RNA-dependent RNA polymerase (RdRp)

MOLECULAR DOCKING

✓Sort out duplicates between the five different binding criteria

✓Compare pose to crystal ligand pose

✓Look for hydrogen bonding interactions

✓Analyze fit in binding pockets look for cavities in enzyme

✓ Look for acceptable pose buy compound test on in vitro assay

✓Generate pharmacological hit

FIRST STEPS

Known inhibitor
S-score produced = -10.8

✔ With MOE software system, the crystal structure of the enzyme's active site and the known inhibitor produce an S-score of -10.8

✔ This represents the binding energy

✔ Known covid inhibitors are scanned through the software to make sure they are suitable for binding
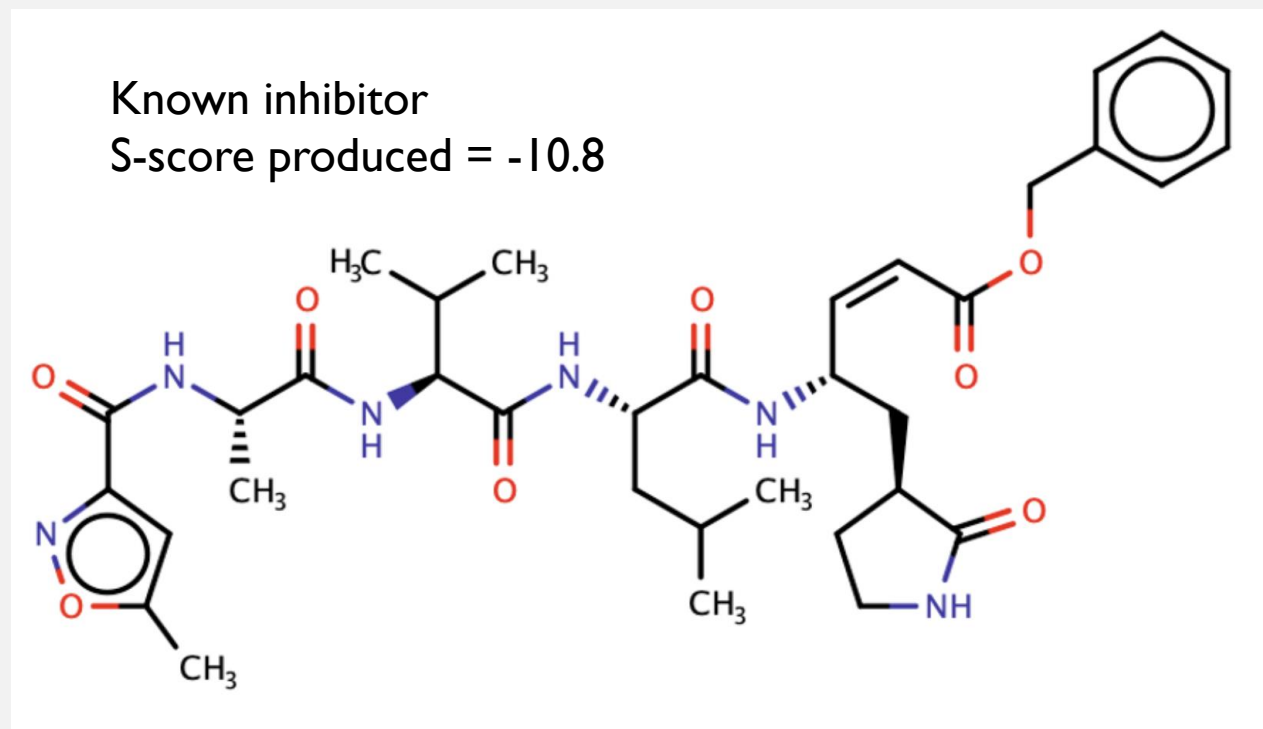
Active site with known inhibitor
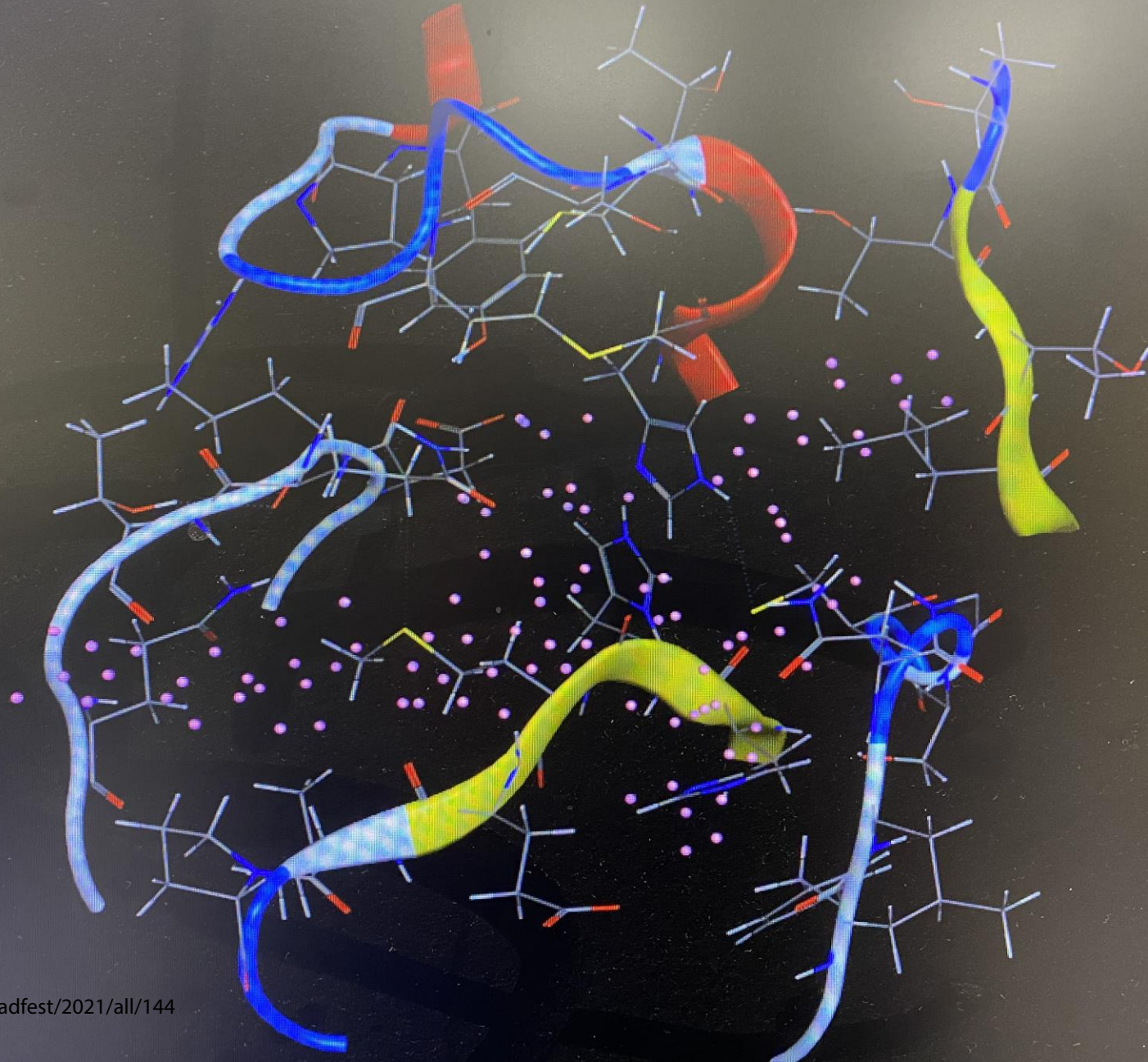
CONTINUED FOR KNOWN INHIBITORS

-10 or lower shows good binding energy

Compound 091
S-score = -10.5
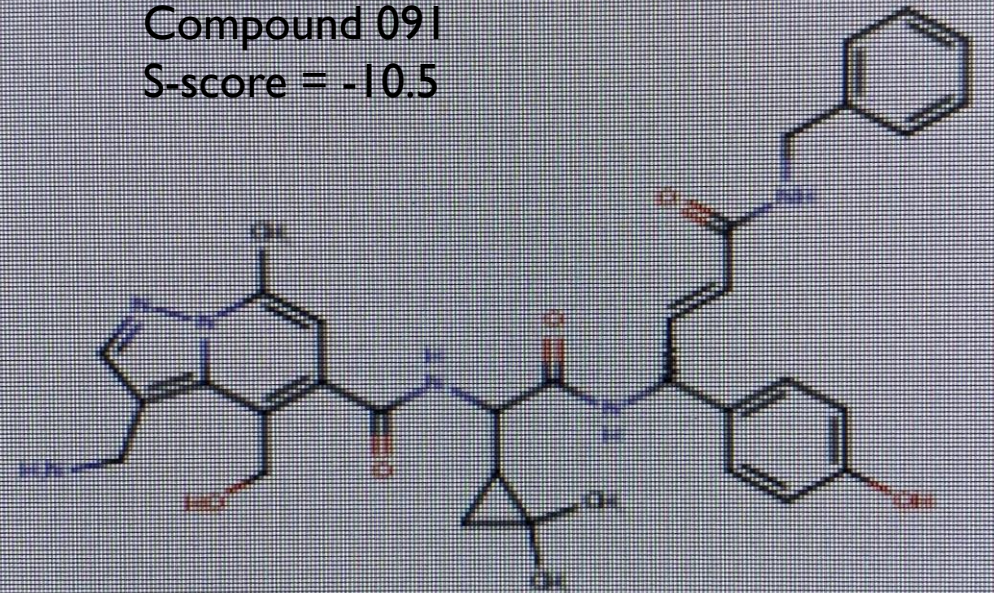
Compound 060
S-score = -10.2
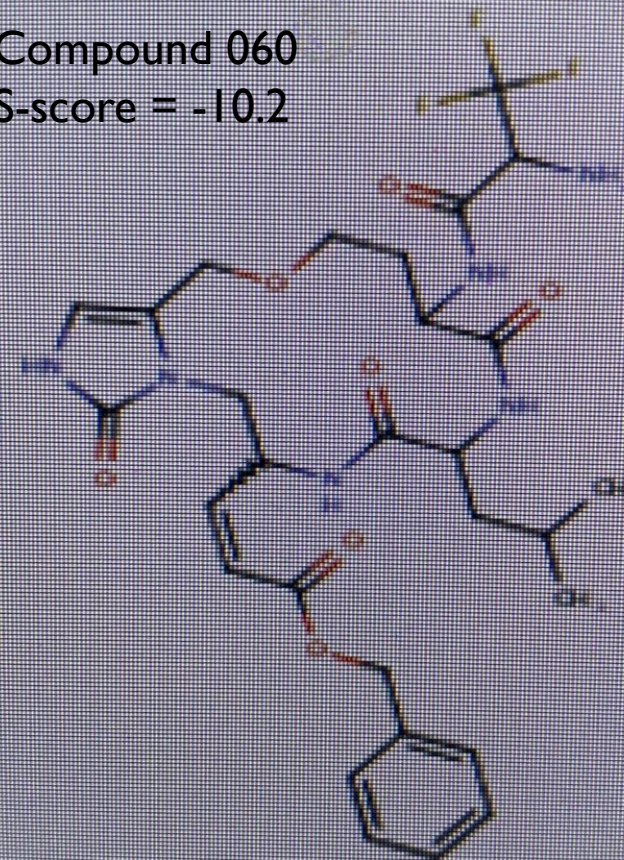
ARE THEY DRUG-LIKE IN NATURE?

- SMILES codes of these compounds were put through Swiss ADME

  - Do compounds violate Lipinski's rules?

  - Do the compounds have a MW > 500?

DO THE COMPOUNDS FOLLOW LIPINSKI RULES?

Criteria for drug like molecules

1. Molecular mass less than 500 Dalton
2. High lipophilicity (expressed as LogP less than 5)
3. Less than 5 hydrogen bond donors
4. Less than 10 hydrogen bond acceptors
5. Molar refractivity should be between 40-130

Examples:

- 091 = breaks rules #1 and 5
- 060 = breaks rule #1
- 033 = breaks rule #1

RESCORE

Pivot table for known inhibitors

✓ Compounds were rescored using 5 different criteria important for binding

✓ A pivot table was produced

✓ Compounds 72 and 91 had the most hits

Sum of Appearance	Column Labels					
Row Labels	Affinity dG	ASE	GBVI/WSA d	London dG	S	Grand Total
Grand Total	49	49	49	49	49	245
INSCoV-072	6	4	5	4	5	24
INSCoV-091	2	5	6	2	5	20
INSCoV-033		5	4	6	2	17
INSCoV-090	2	3	4	2	5	16
INSCoV-031	3	2	4	3	2	14
INSCoV-029		2	5	2	3	12
INSCoV-060		1	4	2	3	10
INSCoV-092	1	2	1	3	3	10
INSCoV-030	1	2	3	1	2	9
INSCoV-095	6	3				9
INSCoV-053	2		3	1	2	8
INSCoV-058	2	1		2	1	6
INSCoV-035			2	1	2	5
INSCoV-083	2	2		1		5
INSCoV-085	1			3	1	5
INSCoV-002		2	1		1	4
INSCoV-003		2	1	1		4
INSCoV-032	1		2		1	4
INSCoV-051	2	2				4
INSCoV-081		1	1		2	4
INSCoV-094	3				1	4

PROTEASE INHIBITORS

name of protease inhibitors	smiles
ciluprevir	O=C2N[C@]7(C(=O)O)C[C@H]7/C=C\CCCCC[C@H](NC(=O)OC
ritonavir	CC(C)c4nc(CN(C)C(=O)N[C@@H](C(C)C)C(=O)N[C@@H](Cc1c
tipranavir	CCC[C@]1(CC(/O)=C(\C(=O)O1)[C@H](CC)c3cccc(NS(=O)(=O)c
saquinavir	O=C(N)C[C@H](NC(=O)c1nc2c(cc1)cccc2)C(=O)N[C@@H](Cc3c
lopinavir	O=C(N[C@@H](Cc1ccccc1)[C@@H](O)C[C@@H](NC(=O)[C@@
calmidazolium	C1=CC(=CC=C1C(C2=CC=C(C=C2)Cl)N3C=C[N+](=C3)CC(C4=C(C
glycyrrhizin	O=C(O)[C@H]7O[C@@H](O[C@@H]6[C@@H](O)[C@H](O)[C
nelfinavir	O=C(c1cccc(O)c1C)N[C@@H](CSc2ccccc2)[C@H](O)CN4[C@H]
boceprevir	O=C(N3[C@H](C(=O)NC(C(=O)C(=O)N)CC1CCC1)[C@H]2C(C)([
GC376	CC(C)C[C@@H](C(=O)N[C@@H](CC1CCNC1=O)C(O)S(=O)(=O)
GC373	CC(C)CC(C(=O)NC(CC1CCNC1=O)CO)NC(=O)OCC2=CC=CC=C2
remdesivir	CCC(COC(=O)[C@@H](NP(=O)(Oc1ccccc1)OC[C@H]1O[C@@](
PF-07304814	[H][C@@]3(C[C@H](NC(=O)[C@H](CC(C)C)NC(=O)c2cc1c(C(=O
PF-00835231	[H][C@@]3(CC(NC(=O)[C@H](CC(C)C)NC(=O)c2cc1c(OC)cccc1[
from asinex pro inibitor list2	NC[C@@H]3CC[C@@H](C(NC(=O)COc1ccccc1F)c2ccccn2)CC3
from asinex pro inibitor list1	NC(=O)c3cnn4c(O)cc(C2CC(=O)N(c1ccc(F)cc1)C2)nc34
from asinex pro inibitor list3	Cc4cccc(CN1CCCCC1c3cc(O)nc(c2ccccn2)n3)n4
from asinex pro inibitor list5	COc4ccc(c3ccc(CC2CN(C(=O)CN1CCCC1)CCNC2=O)cc3)cc4
from asinex pro inibitor list6	NC[C@@H]3CC[C@@H](C(NC(=O)COc1ccccc1F)c2ccccn2)CC3
D-Phe-Pro-p-Amidinobenzylamine	N=C(N)c3ccc(CNC(=O)[C@@H]1CCCN1C(=O)[C@H](N)Cc2ccccc2
camostat mesylate	CN(C)C(=O)COC(=O)CC1=CC=C(C=C1)OC(=O)C2=CC=C(C=C2)N=C(
nafamostat mesilate	C1=CC(=CC=C1C(=O)OC2=CC3=C(C=C2)C=C(C=C3)C(=N)N)N=C(N
hydroxychloroquine	Clc1cc2nccc(c2cc1)NC(C)CCCN(CC)CCO
chloroquine	Clc1cc2nccc(c2cc1)NC(C)CCCN(CC)CC
3CL pro inhibitor	[C@@H](C(N([HH0])[C@@H](C[C@@H]1CCN([HH0])C1=O)C=O)
Hirsutenone- dual 3CLpro PLpro inhibitor.	c1cc(c(cc1CC/C=C/C(=O)CCc2ccc(c(c2)O)O)O)O
Tideglusib	c1ccc(cc1)Cn2c(=O)n(sc2=O)c3cccc4c3cccc4
Ebselen	c1ccc(cc1)n2c(=O)c3ccccc3[se]2

- Next I compiled a list of protease inhibitors

- Converted this into an SDF file

- Protease inhibitors are scanned through the software to make sure they are suitable for binding

17

WASHED PROTEASE INHIBITORS

- Receive the S-score

- Compound 35 and 1 have an s-score of -10

- Look to see how they apply to Lipinski's rules

https://digitalcommons.sacredheart.edu/acadfest/2021/all/144

	mol	rseq	mseq	s	rms
1	Compound 35	1	34	-10.6664	
2	Compound 1	1	1	-10.0720	
3	Compound 4	1	4	-9.6826	
4	Compound 2	1	2	-9.6185	
5	Compound 35	1	34	-9.6009	
6	Compound 1	1	1	-9.5637	
7	Compound 35	1	34	-9.5487	
8	Compound 2	1	2	-9.4725	
9	Compound 13	1	13	-9.4231	
10	Compound 4	1	4	-9.3807	
11	Compound 1	1	1	-9.2399	
12	Compound 14	1	14	-9.2213	
13	Compound 2	1	2	-9.2135	
14	Compound 13	1	13	-9.1735	
15	Compound 9	1	9	-9.1554	
16	Compound 1	1	1	-9.1380	
17	Compound 9	1	9	-9.0764	
18	Compound 4	1	4	-9.0362	
19	Compound 4	1	4	-8.9732	
20	Compound 35	1	34	-8.9179	
21	Compound 13	1	13	-8.9044	
22	Compound 35	1	34	-8.8933	
23	Compound 1	1	1	-8.8730	
24	Compound 9	1	9	-8.8633	
25	Compound 1	1			

❖ Both break 2 of Lipinski's rules

Compound 35

Compound 1

FUTURE WORK

- Create similarity searches of the compounds with the highest binding energy
  - In doing so, stick to finding compounds that do not violate Lipinski's rules
- Pharmacophore elucidation
  - Create a drug that follows Lipinski's rules and has high binding energy to the main protease for drug therapy