Computational molecular modeling of Paxlovid binding

Gordon Chalmers

Complex Carbohydrate Research Center

University of Georgia, Athens, GA, 30602

gordoncs@uga.edu, ORCID: 0000-0003-0254-352X

Submitted to European Journal of Medicinal Chemistry

Abstract:

Computational binding modeling of the drug Paxlovid to the 3CL-protease of SARS-Cov-2 is presented. Paxlovid contains 2 molecule ingredients: Nirmatrelvir and Ritonavir. This is a detailed distributional analysis using docking of both small molecules, to the target protease and to the CYP 3A4 liver enzyme. The atomic score interactions are quantified both in the non-covalent binding and in the reversible covalently bound Nirmatrelvir. Conformational analysis is performed for both of the small molecules in binding to the proteins and without using molecular dynamics. The computational binding is used in good comparisons to x-ray determined structures. The total free energy calculations are presented in terms of docking scores and translated to energy at room temperature.

Keywords: protein-ligand interactions, docking software, small molecule inhibitors, high performance computing, drug design, CADD

Introduction

In this work a complete computational binding analysis is given of the small molecule treatment Paxlovid for early SARS-Cov-2 infection which is useful for a thorough understanding of the drug interactions. SARS-Cov-2 remains a continual threat to the population even after 2 years during which time its molecular structure and amino acid variations have been extensively documented [1]. Although vaccines have been successfully developed and used, there still remains the development of a viable therapeutic [2] that has no potential risk to the general individual. This is motivating in finding new small molecules, possibly unrelated to known molecules and without repurposing.

The search for orally ingested therapeutics has generated 2 drug US EUA approvals, with the successful Pfizer's Paxlovid [3] and Merck's Molnupiravir [4]. Paxlovid refers to Nirmatrelvir in the presence of Ritonavir. Nirmatrelvir, i.e. PF-07321332, is a small molecule that binds to a particular region of the SARS-Cov-2 main protease, blocking proteolytic cleavage of some of the SARS-Cov-2 polyproteins [1], hence viral replication at an early stage of infection. X-ray structures of Nirmatrelvir bound to the SARS-Cov-2 3CL protease, aka Mpro, are available at the Protein Data Bank [5], and the one PDB ID 7SI9 is used [5,5 to 6]. The molecule is a descendant of an earlier created molecule, PF-00835231 [6], created during the SARS-Cov-1 outbreak in 2003. (The molecule PF-00835231 was created in 2003, and is from an early year in modern computational chemistry in molecular simulation and modeling with computers. PF-07321332, Nirmatrelvir, is a modification.) Nirmatrelvir's precursor is an injectable and was

modified, turned into an orally ingested pill, which gained FDA EUA approval as PF-07321332 in Paxlovid in Nov 2021. However, the approval is for patients with higher risk of hospitalization or death after infection. The inhibition of viral replication to avoid serious illness has a window of ~5 to ~10 days post-infection, after that the virus has multiplied enough in the body that inhibition of viral replication has much less effect.

A serious drawback of Nirmatrelvir is that it metabolizes too quickly in the liver [7]. The cytochrome P450 class (aka CYP ###: gene family, subfamily, isoform/individual type) of enzymes is common, occurring in different species and contain a heme iron network that makes them function as oxygenases [8,9,10]. There are 57 encodings for CYPs in the human body, of which 6 are liver enzymes: CYP 1A2, 2C9, 2C19, 2D6, 2E1, 3A4, some of which are also found in the abdominal gut, and estimated to account for 80% of drug metabolism. These enzymes are the main source of metabolism and excretion of drugs in the human body. These, primarily CYP 3A4, eliminate Nirmatrelvir from the body and reduce the plasma concentration in the bloodstream.

Ritonavir is a well-known inhibitor of most of these enzymes and is commonly used to increase the concentration of other drugs in the body by non-specifically blocking the enzyme's normal metabolic function. It was initially developed as a protease inhibitor, in particular to HIV, but is now commonly used as an activator of other drugs by inhibiting the CYP enzymes, amongst other proteins. The molecule Ritonavir is large and flexible and sticks to almost any protein; it is not a precision designed small molecule, but the body can metabolize it. The problem with Ritonavir, and as a result with the Paxlovid therapeutic, is that by inhibiting these enzymes it also blocks the metabolism of a number of other molecules and drugs [11], including commonly used anti-depressants and cholesterol medicines. The use of Ritonavir could send a normal medication or dose of prescribed medication into a toxic level. There are a variety of side effects of Ritonavir [12]. As a result, Paxlovid is primarily intended for high risk groups in its FDA approval [13], which is about 2% of the population in the US. It was announced in June 2022 to be effective primarily only in this group.

Method

Throughout this work Corina Classic [14] is used to generate mol2 files from SMILES expressions [15], and the Cambridge Crystallographic Data Center GOLD software [16] is used for protein-ligand docking calculations. Matlab [17] is used for scripting, implementation of software, and analysis. The computational molecular analysis uses these software in a large randomly sampled docking distributional method of analysis [18].

Paxlovid is a combination of 2 molecules: Nirmatrelvir and Ritonavir. Paxlovid in the literature is often referred to Nirmatrelvir in reference to the combination with the Ritonavir molecule.

Results: Nirmatrelvir

The non-isomeric Nirmatrelvir molecule in SMILES notation is,

CC1(C2C1C(N(C2)C(=O)C(C(C)(C)C)NC(=O)C(F)(F)F)C(=O)NC(CC3CCNC3=O)C#N)C eqn 1

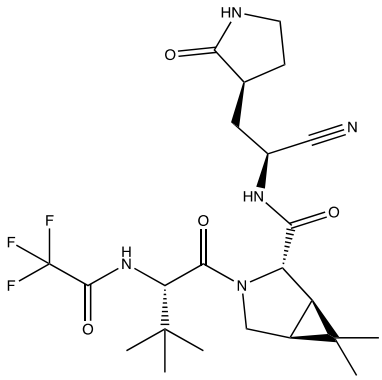
This non-isomeric molecule has 32 stereoisomers, one of which is the Nirmatrelvir molecule,

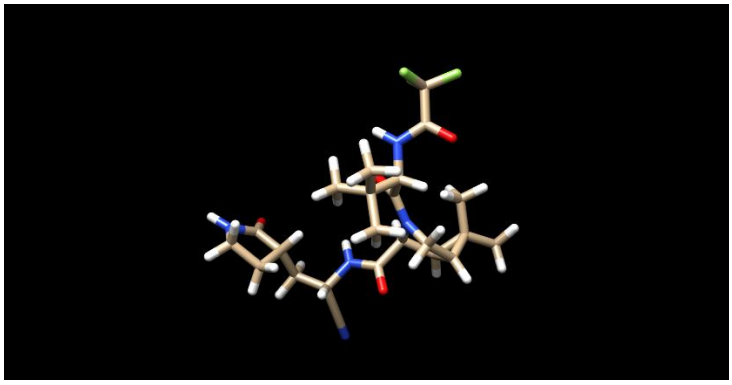
CC1([C@@H]2[C@H]1[C@H](N(C2)C(=O)[C@H](C(C)(C)C)NC(=O)C(F)(F)F)C(=O)N[C@@H](C[C@@H]3CC
NC3=O)C#N)C eqn 2

and is shown in Figure 1.

Figure 1: (a) Nirmatrelvir. (b) Nirmatrelvir in Chimera.

(a) (b)





Nirmatrelvir has the molecular and ADME properties:

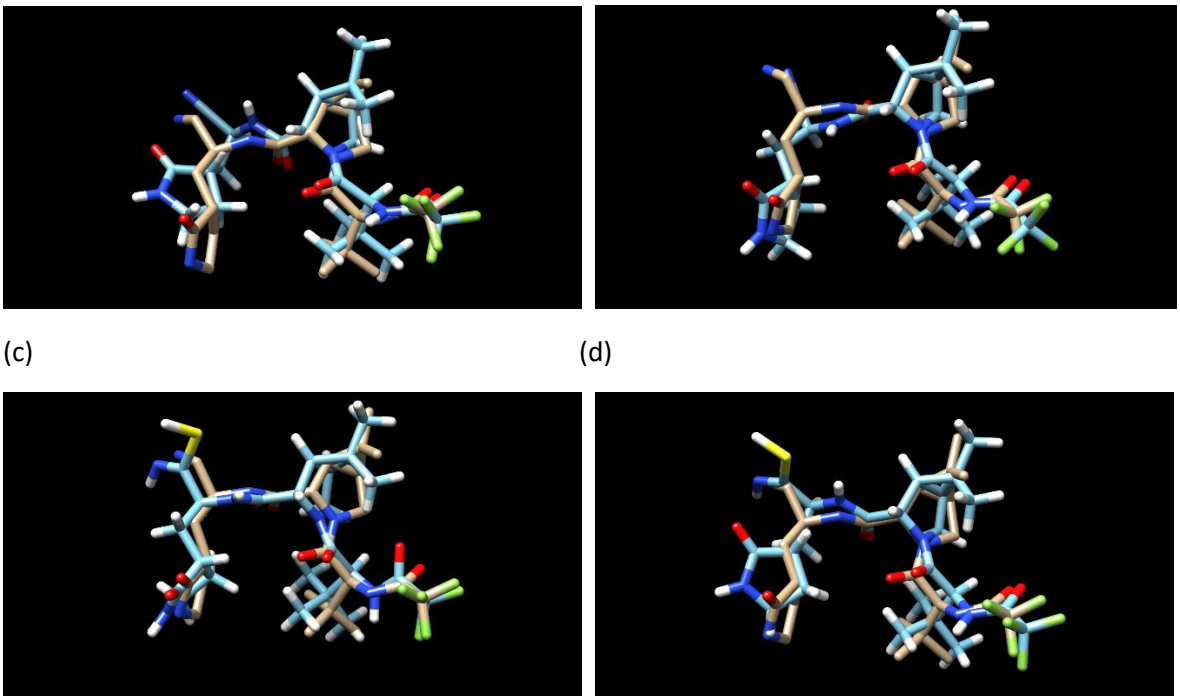
Mass (Da)	Heavy atoms	Rotat. bds	Hbond-don	Hbond-acc
499	35	9	3	9

It is borderline in satisfying the well-known ADME requirements [19-23] given these characteristics.

The x-ray crystal structure PDB ID 7SI9 [5,6] of Nirmatrelvir bound to the SARS-Cov-2 was used for docking and also in validating the docked poses; the x-ray protein structures were fully protonated in GOLD/HERMES before docking calculations. The computational docked pose is close to that of the complete x-ray structure, including the eclipsed trifluoromethyl group. The non-covalent and covalently docked pose overlayed with the 7SI9 ligand is illustrated in Figure 2, and these will be discussed. The trifluoromethyl is in an eclipsed position which is unusual because typically this is higher in energy than non-eclipsed; Studies of these groups have been done with this eclipsed position and found to be more stable in classes of compounds [24]. In both (a,b) and (c,d) 2 conformations are presented, the highest scoring docked Nirmatrelvir/Mpro pose and a pose at the start of the continuous set of scored poses, generated by CCDC GOLD out of multiple runs. The RMSDs are given in the caption and are comparable to the PDB 7SI9 2.0 x-ray resolution [6].

Figure 2: Nirmatrelvir from the x-ray 7SI9 superimposed (tan) with the computationally non-covalently docked molecule (blue). The highest scoring pose (a) at 72.64 and a nearby one at (b) 70.15. 72.68 is one of .05% isolated results (11 total out of 20040) higher in score and away from the continuous distribution. (c) The covalently docked highest scoring pose at 116.32 (blue) with the 7SI9 x-ray ligand (tan). There are .1% (21 total) with scores $\geq$112.5. (d) Score at 112.02 modeled ligand. Protons were not added to the (tan) ligand from the xray structure and the white off of sulfur in (c,d) represents CYS 145 covalently bonded, and not a hydrogen.

(a) (b)

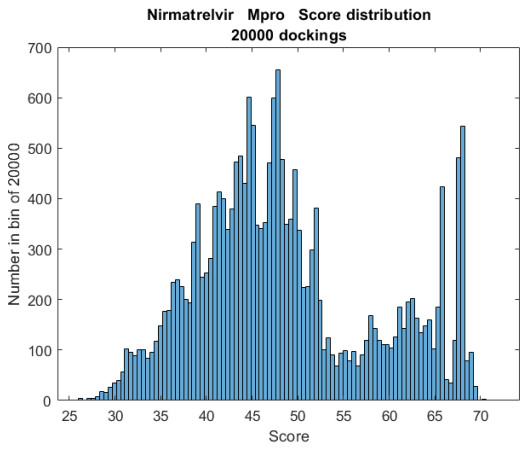


(c) (d)

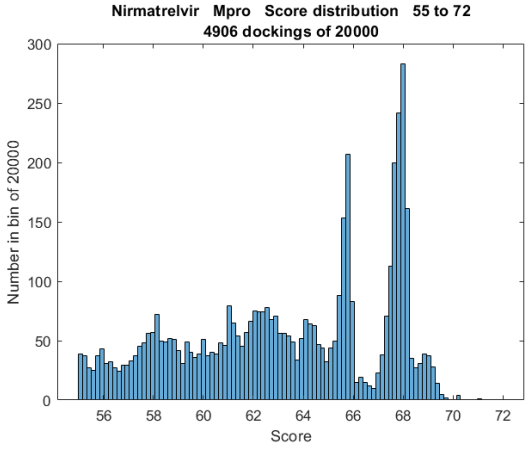
GOLD is a non-deterministic genetic algorithm and running it multiple times samples the possible binding states, local and global minima of the binding interaction model fitness function. Large sets of docking jobs of Nirmatrelvir were done to sample the distribution of states quantified by docking score, i.e. binding interaction energy. The histogram of the docking scores of Nirmatrelvir/Mpro is shown in Figure 3. There are seemingly 2 dominant modes of ligand-protein interaction, a broad one at 46 and a narrow one at 68 GOLD PLP score. A finer bin resolution points to additional sub-structure about these scores. Figure 3(b) has the distribution of 4906 (of 20000) in the range of 55 to 71; Note that 72.68 is an outlier, due from too much specificity in the protein-ligand interaction for the ligand to wiggle into. Figure 3(c,d) show a 4-exp|| fit to this part of the distribution and characterizes score-wise in energy the distinct ligand conformations.

Figure 3: Histogram of GOLD docking job scores of Nirmatrelvir to Mpro. (a) distribution from 20000 docking jobs, and (b) the same in the score range ≥55.

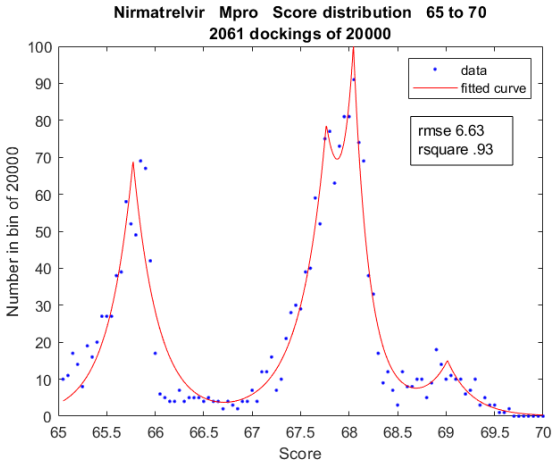
(a)



(b)



(c)



(d)

Multi-exp|| fit, 75 bins in scores [65.0,70.0]

4 binding modes,

$$\sum_{i=1}^{4} a_i \, e^{-|x-b_i|/c_i}$$

a = 69.0, 69.0, 73.3, 14.0

b = 65.8, 67.8, 68.0, 69.0 max 72.64

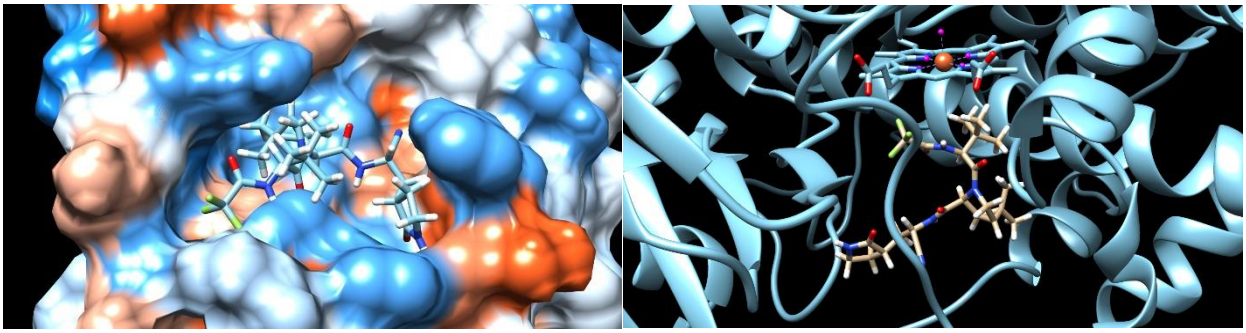
c = .3, .3, .1, .2

A Chimera view of Nirmatrelvir in its near highest Mpro pose and it bound to CYP 3A4 are in Figure 4. The distributions of docking runs to both proteins, Mpro and CYP 3A4, are in Figures 3 and 5. Several structural characteristics of Nirmatrelvir are: it has a fused fused bicyclo-[3.1.0]-azahexane ring, a trifluoromethyl group which is eclipsed when bound to Mpro, and a nitrile extension.

The SARS-Cov-2 3CL protease (Mpro) binding site is an X-shaped cavity which can be described as: 3 ends are caves and the fourth is an exit valley. This is the region of the Mpro relevant to viral replication which is targeted and partly blocked by Nirmatrelvir or other inhibitors. This molecule, when bound, primarily fills in 2 caves. In Figure 4(a) the blue-to-white-to-orange coloring are hydrophilic to hydrophobic amino acids.

Figure 4: (a) Nirmatrelvir non-covalently bound to Mpro. (b) Nirmatrelvir non-covalently bound to CYP 3A4. Cysteine 145 is near the nitrile N40, 3.64 Angstroms. Note the hydrophilic (blue) and -phobic (orange) regions.

(a) (b)

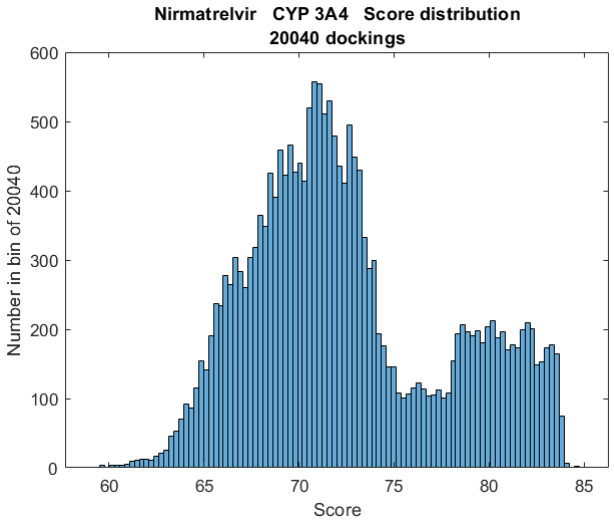


Note: In this image of Nirmatrelvir bound to Mpro, the solvent accessible protein surface (SAS) is displayed. This has the advantage of clearly showing the binding cavity. However, in this image and in general there could be a hydrogen appearing to penetrate the surface. This is the case with Nirmatrelvir and Mpro in the atoms N19, N36 farthest, nearest nitrogens from the trifluoromethyl group. These both are hydrogen bond donors (info on next page) with Mpro atoms 1315, 1112 in hydrophilic GLU 166, hydrophobic PHE 140. There is no structural clash or overlap, and this is common in ligand-protein Chimera images.

Nirmatrelvir is extracted from the body through its interaction with enzymes in the liver, primarily the CYP 3A4. This enzyme is the primary means of elimination of pharmaceuticals in the body. If the binding to this enzyme is too high then the lifetime of Nirmatrelvir will be short and not effective at small dose to inhibit SARS-Cov-2 viral replication. This in fact is an issue with Nirmatrelvir, and Ritonavir was added to inhibit normal liver function and activate the active drug. The score distribution of Nirmatrelvir and CYP 3A4 is given in Figure 5.

The CYP 3A4 cavity is a large featureless region that penetrates well into the structure. It doesn't have the structural features that typical target proteins have such as pockets, holes, valleys, and others. As a result it binds to generally anything and the question is how much. This largeness and featurelessness is the reason that the distribution has very wide peaks without much structure. While score can increase in lowering atom number in a cavity with lots of structure, generally with CYP 3A4 the more atoms the higher score. Nirmatrelvir has its highest score of 84.4, and it has 35 atoms (2.4 per atom).

Figure 5: 20000 dockings of Nirmatrelvir with the enzyme CYP 3A4. Score distributions.



The docking information of highest scoring Nirmatrelvir to Mpro and CYP 3A4 is

	Score	S(PLP)	S(hbond)	S(cho)	S(metal)	DE(clash)	DE(tors)	Intcor
Mpro	70.15	-64.46	2.82	0	0	0	1.46	.16
3A4	84.44	-82.76	1.00	0	0	0	1.00	0.68

The particular state of Nirmatrelvir at highest score of 72.64 is an outlier in the histogram of docking runs, and that part of the distribution doesn't get any smoother with additional runs. First, it is the highest scoring conformer. Second, there are 11 out of these between 70 and 72.6. The ~3 score spread and discreteness gap in states is due to a strong physical restriction of binding. Binding in these poses is like flying a small plane into a mountain tunnel in high wind. It seems unlikely in thermal conditions for a ligand to bind there although possible. Physically there is a discrete ring rotation in the conformation that would make this conformer stick out from the continuous bumps at 68, 69 in Figure 3.

The geometric pose agrees with the x-ray crystal structures of covalently bound Mpro to Nirmatrelvir [5,6]. X-ray samples of protein-ligand complexes are unique in energy minimization however due to crystal creation at low entropy conditions. The very much larger density of states at 67-69 with a spread of ±3, Figure 1(b) is thermally realistic, with a typical score of 68.

Non-covalent Nirmatrelvir hydrogen bonds

There are 4 ligand-protein hydrogen bonds to Mpro in the pose at 70.15,

Donor molecule	Donor type	Donor heavy atom	Donor atom	Acceptor molecule	Acceptor atom	Score	Amino acid
L1	H	19	56	P1	1315	1.00	GLU 166
L1	H	36	64	P1	1112	.61	PHE 140
P1	H	1136	2470	L1	40	.60	GLY 143
P1	H	1312	3711	L1	12	.91	GLU 166

and 6 in the highest scoring, but isolated, docked pose at 72.68,

Donor molecule	Donor type	Donor heavy atom	Donor atom	Acceptor molecule	Acceptor atom	Score	Amino acid
L1	H	19	56	P1	1315	1.00	GLU 166
L1	H	36	64	P1	1123	.40	LEU 141
P1	H	1136	2470	L1	38	.32	GLY 143
P1	H	1140	2475	L1	38	.22	SER 144
P1	H	1146	2467	L1	38	1.00	CYS 145
P1	H	1312	3711	L1	12	1.00	GLU 166

The ligand atoms that form hydrogen bonds in the 1st binding mode is found by adding the hydrogen bond contribution to the atom interaction, column 5 in the GOLD output file HBond(atom).S, over all of the output docking files. The cutoff for this binding mode is 66 and there are 1434 with higher total scores. The frequency of hydrogen bonds in these states is: 1434, 1434, 1075, 315, 1260, and 1527 for atoms 12, 19, 28, 36, 38, and 40. Note that if an atom could be in 2 hydrogen bonds where the pose score would be greater than 1. The hydrogen bonds in the 1st mode involve,

O12, N19, N28, N36, N40 and somewhat O38

The same calculation goes for the amino acid participation in hydrogen bonds. Adding all of these 1434 terms per heavy atom (in the protein atom score section of the GOLD output file) in the protein gives,

Atom number	Count	Amino acid	Atom
183	254	THR 26	O
312	1075	HIS 41	NE2
1135	1304	ASN 142	ND2
1136	1420	GLY 143	N
1312	1434	GLU 166	N
1315	1434	GLU 166	O
Non-zero and ≤ 30			
THR 25 OG1	PHE 140 O	LEU 141 O	SER 144 N
SER 144 OG	CYS 145 N	GLU 166 OE1	

The 1st binding mode is characterized primarily by

ASN 142, GLY 143, GLU 166 and partly THR 26, HIS 41.

This non-covalent hydrogen bond count in the highest binding mode is substantially different than from the covalently bound molecule to CYS 145. The THR 26 is present with non-covalent interaction but no contribution in the covalently bound Nirmatrelvir highest binding mode. The fact that LEU 141, SER 144, and CYS 145 showed up in the highest scoring pose and not in the common set is due to being an isolated state with score 72.68.

The difference in score from Nirmatrelvir-Mpro to Nirmatrelvir-CYP 3A4 is 84.4-70.7 = 13.8. This is large in comparing binding choices of 2 targets, approximately of 2.1 kCal in binding energy. GOLD PLP scores

are dimensionless numbers representing the interaction of docked ligand to the protein; the rough correspondence between PLP score differences and energy differences is found heuristically by known calculations and is roughly 6.5 PLP for 1 kCal for various non-covalently bound molecules. The binding energy is important in biochemical terms of ligand in the on-/off- states. In terms of k_D or on-off ligand binding fraction, 1 kCal is commonly referred to be a factor of 10, 10^x, although this is not a linear or rigorous relation.

Atomic level interactions

The total interaction for a ligand bound to a protein can be broken down at the atomic level, then further in types of atomic interactions. Distributions of any can be calculated from large sets of docking runs. There are many reasons why this information is relevant. A chemical interpretation of the individual atomic binding characterizes overall protein-ligand binding. Identifying badly fitting atoms can be used in improved molecular design. Identifying atom and regions of good interaction can explain binding in terms of positional attachment and longevity of the ligand semi-bound state (in an on-/off- dynamic setting).

An average PLP score per atom of many protein-ligand complexes leads to the following heuristic,

Total atom modeled score	
1 or <1.6	Naturally occurring ligands
2 or < 2.1	Common in pharmaceuticals
2.5	Specific to cavity
3 or < 3.5	Very high
>3.6	Exceptionally high, hydrogen bonds, buried, metals, …

Total GOLD docking score of a ligand pose is characterized in the gold_soln.mol2 output files by,

Score S(PLP) S(hbond) S(cho) S(metal) DE(clash) DE(tors) intcor

with higher score correlated linearly with protein-ligand binding energy. The total score is further broken into the contributing interaction types in GOLD output files. The GOLD usage keeps the early_termination flag to zero, and 8 islands, 200000 maxops, and a population size of 75 to 100 was used.

Furthermore, the total score is a summation of the individual ligand atom contributions and their different modeled interaction types. These individual scores are listed in the output GOLD files as,

AtomID ChemScore_PLP.Hbond ChemScore_PLP.CHO ChemScore_PLP.Metal PLP.S(hbond)

PLP.S(metal) PLP.S(buried) PLP.S(nonpolar) PLP.S(repulsive) PLP.total

And two examples are,

1 0.00 0.00 0.00 0.00 0.00 -0.26 -3.92 0.00 4.18

20 2.00 0.00 0.00 -2.37 0.00 -0.53 0.00 0.00 8.90

Nirmatrelvir has 35 heavy atoms, and in the 1st binding mode seen in Figure 3(a,b), scores ≥ 66. The atomic score distributions are generated from the output gold.mol2 files. Color-coding is used to illustrate in the molecule structure the amount of interaction from individual atoms, Figure 6 coming from the more detailed atomic score distributions in Figure 7. The -bi,-tri, high refer to bi-modal, tri-modal, high interaction or very high interaction (anomalously so); the atom color is the larger peak in states, and the second peak descriptor, e.g., bi-, is smaller. A plus or minus means the peak contains the color change point, and the +(-) means the smaller side is in that color. These distributions and color coding of the molecule give a tremendous amount of binding information for a set of states such as the highest binding mode.

The atomic level distributions of scores are given in Figure 7. The total atomic score in a sampled ligand docked pose could go negative, and the total comes from adding the individual contributions modeling the different protein-ligand atom interactions. One example is a repulsed carbon atom (methyl) in a strongly hydrophilic amino acid neighborhood, the non-polar component or the repulsive component of the interactions could be negative. If a total atom score is centered about a negative valued peak score, less than some value such as -1.5, then this means that the atom isn't welcome in that location from the binding mode. In Figure 7, the range of scores in the distributions is [-5,5]. There aren't any negative valued peaks in this case, but there are some negative atomic scores.

Nirmatrelvir is not a particularly strong non-covalent binder with average PLP score per atom of 2.0. This is seen in detail by the number of light colored blue atoms in Figure 6. However, upon inspection of Figure 6, notice that there are 4 separated anchor points in purple with adjacent red and the trifluoro group is red. The branching is not highly interacting but there are separated stably bound regions.

Figure 6: Interactions of individual Nirmatrelvir atoms within the Mpro cavity from the 1st binding mode, total scores ≥ 66.

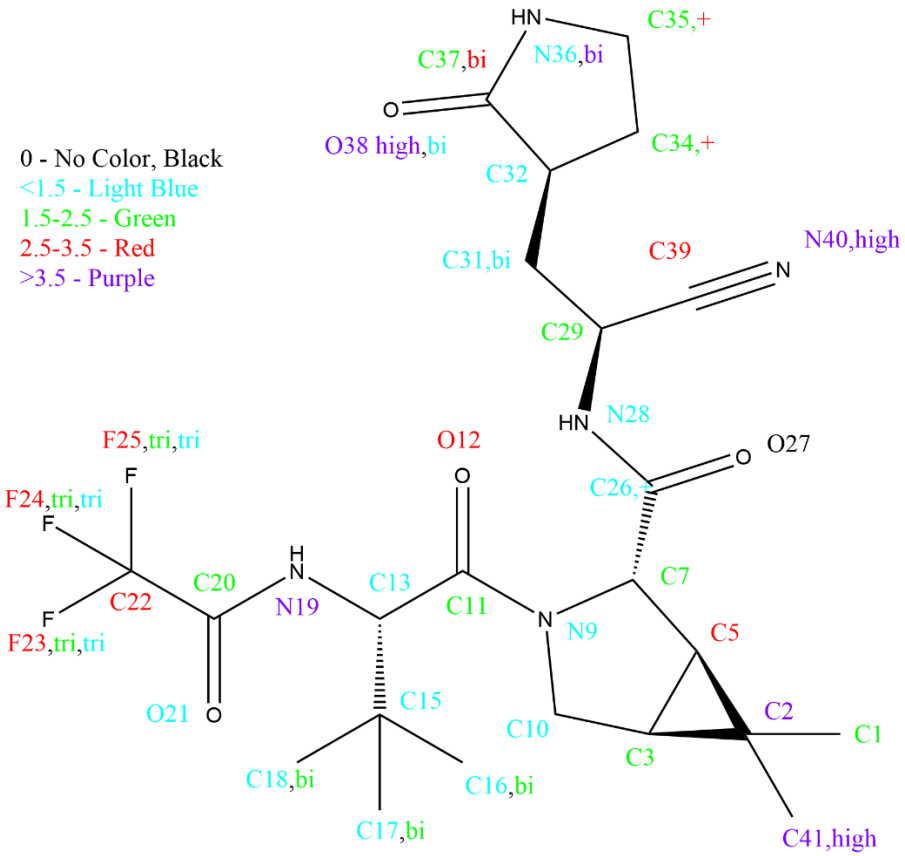
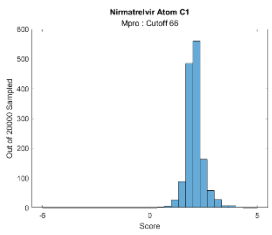
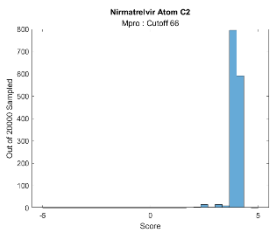


Figure 7: Atomic score distributions for 35 heavy atoms of Nirmatrelvir non-covalently bound to Mpro in the highest binding mode. There are 1434 docked poses with total score ≥ 66 out of 20000. Domain is -5 to 5. The size of the Word multi-figs is so that they can be zoomed in on with readability.

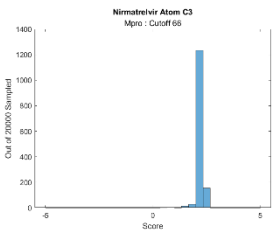
C1



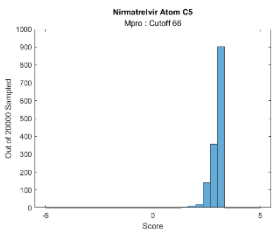
C2



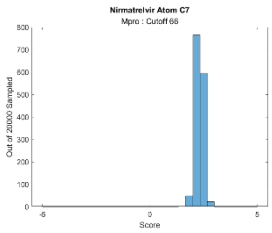
C3



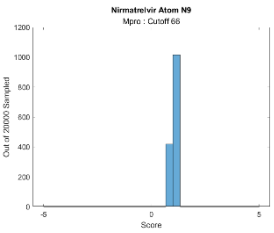
C5



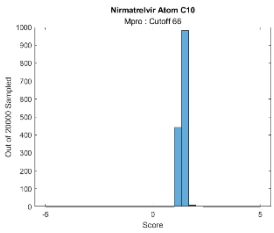
C7



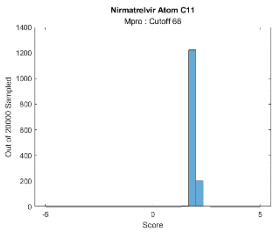
N9



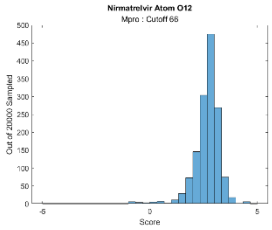
C10



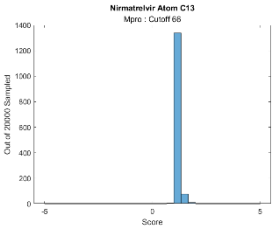
C11



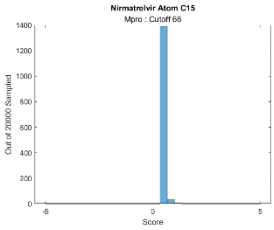
O12 C13 C15 C16



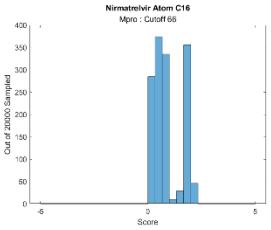
C17



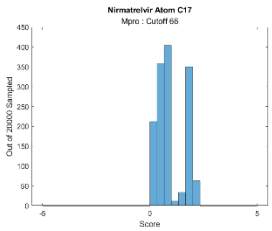
C18



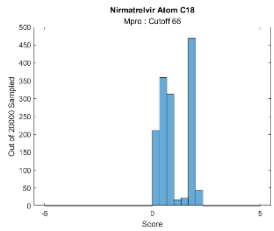
C19



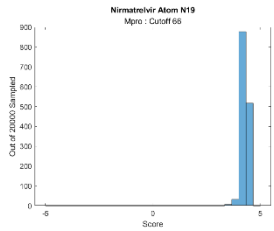
C20



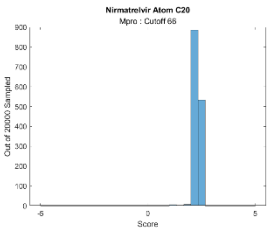
O21



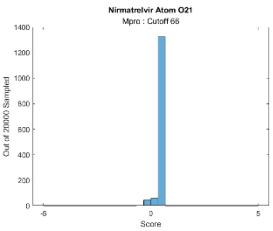
C22



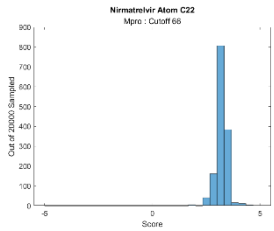
F23



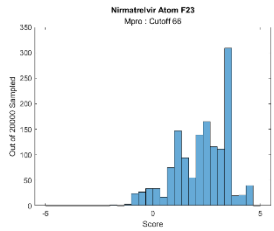
F24



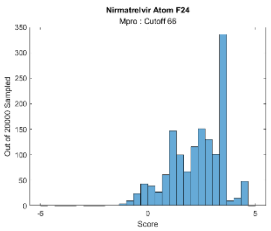
F25



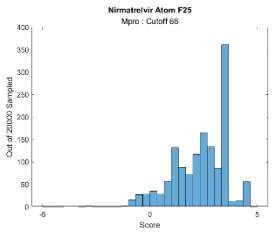
C26



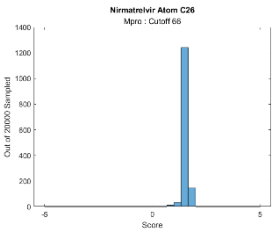
O27



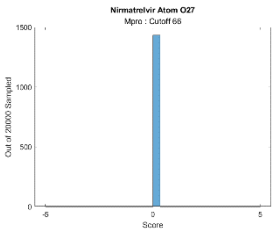
N28



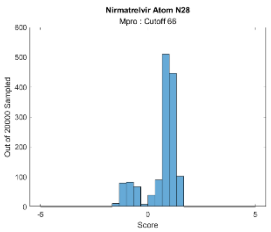
C29



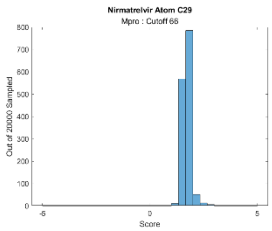
C31



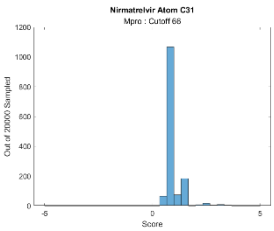
C32



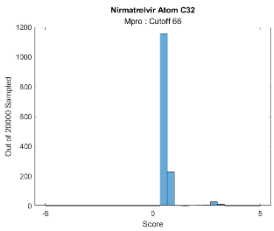
C34



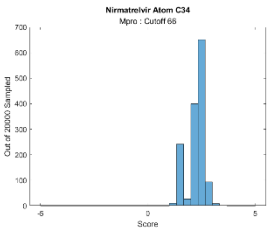
C35



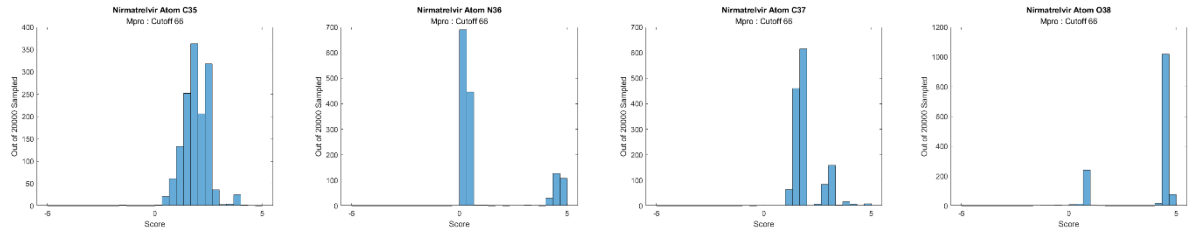
N36



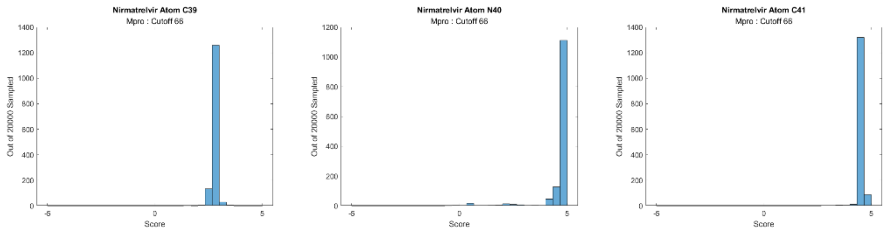
C37



O38



C39 N40 C41



Atoms C17, C18, N19, C20 are all near each other and score very high, also O38 and C41. N36 doesn't interact at all except for a small set of poses where the nitrogen has strong hydrogen bonding, and this is the case of the isolated binding pose score of 70, away from the bulk in the 1st binding mode at 66-69. N19, N28, O38, N40 all are high scores, but not on the extreme side of ≥5 which would mean also buried and more interactive. In general the hydrogen bonding atoms are not the primary sources of Nirmatrelvir's score as can be seen by the high 4 to 5 contribution of various carbons and non-hydrogen bonded oxygens and nitrogen. Some potentially interacting nitrogens and oxygens are very weak, such as N9, O21, O27, N28, and N36. In particular, =O27 has no interaction which then allows it to be involved as ketone instigator in the covalent binding of Nirmatrelvir. Several of the ligand atoms given the location in either of hydro -phobic (nonpolar) or -philic (polar) neighborhoods and buried scores show good or bad placement with high or low scores.

The protein region near the branch N9 to F25 is hydrophilic – blue in figure 1(a). In the color coded Figure 6 and distributions in Figure 7 the O12,N19 and F23,F24,F25 atoms are all highly interacting (red or purple) except for O21 which average (blue).

In the 2 docked poses in Figure 1(a) and 1(b) with scores 72.68 and 70.13, the distance between N40 and CYS 145 sulfur atoms are 3.50 and 3.68 A. These distances are pointed out in relevance to the thionate reaction that covalently binds the two. The distance from these two poses are approximately 2 times the N-S bond distance of 1.8 A.

Atomic level interaction information is very useful to understand the origins and locality of ligand binding to the protein cavity. It is encouraged to examine the superimposed Nirmatrelvir/Mpro complex at the atomic level in a molecular viewer with the Figures 6 and 7 distributions in mind.

Covalent

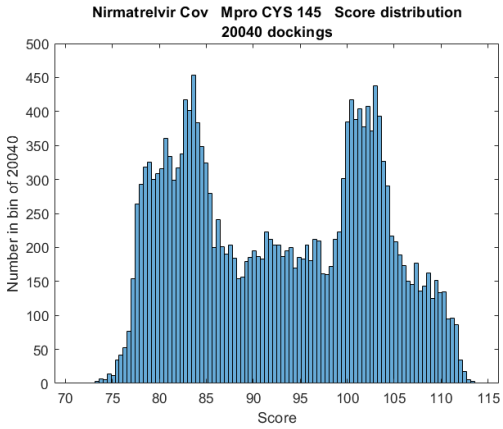
When binding to the Mpro cavity, a thioimidate adduct sub-structure is formed when the nitrile N40 interacts with the sulfur of Cysteine 145. Although these and neighboring atoms are involved in the reaction, other atoms have an effect in anchoring the ligand to the protein in order to initiate the reaction. GOLD is used in covalently attaching the (H deleted) CYS 145 to the (H ignored) N40 and the total score distribution is shown in Figure 8.

The thionate reaction of CYS 145 with the nitrile is,

-C(C#N)- + S- - -> -C(C(=N)S-145)- eqn 3

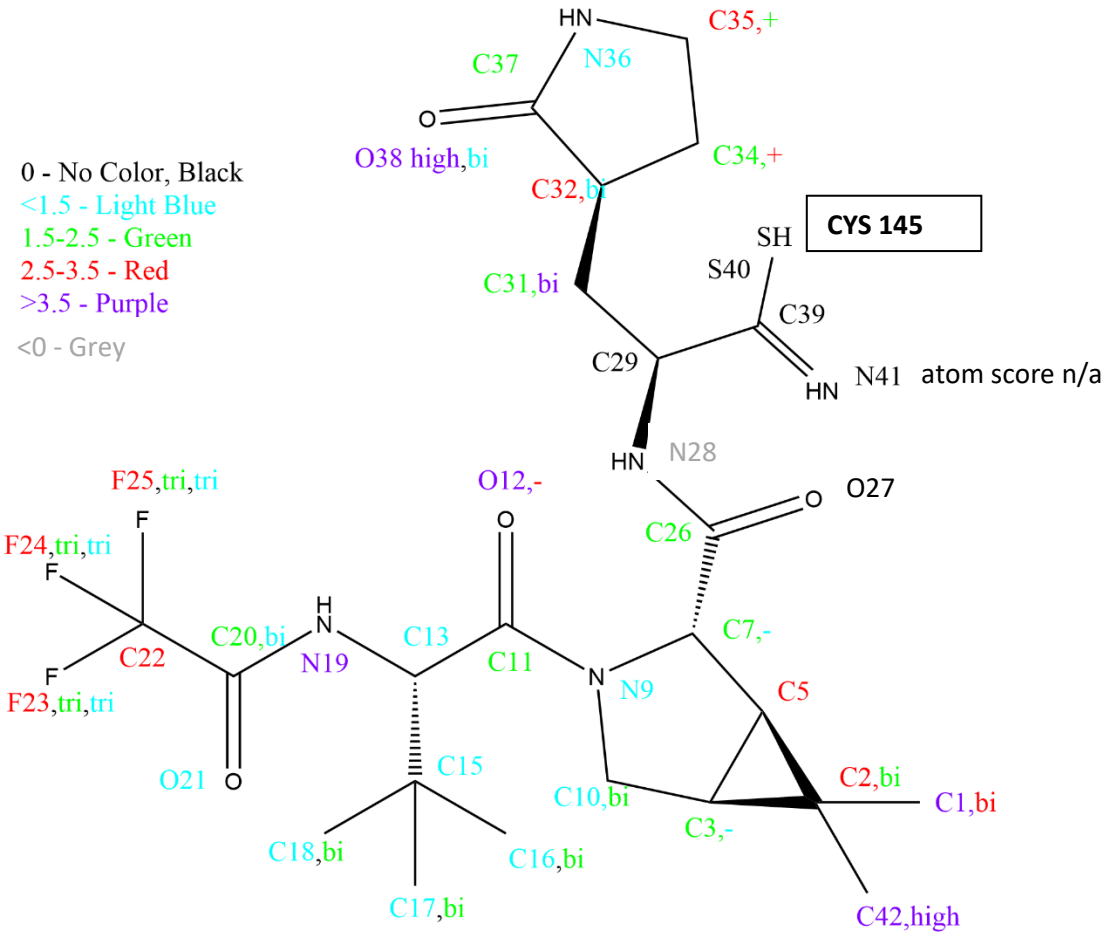
with hydrogen number preserved. Keeping track of the bond energies, C#N 891 kJ/mol, SH 347 in the initial and C=N 615, CS 259, NH 391 in the final, gives a bond energy change of 1268-1238=30 kJ/mol. This translates to a difference of 30/4.184=7.17 kcal. Given the heuristic of 6.5 score to 1 kcal, 47 is covalent score difference from 7.17 kcal. The highest non-covalent interaction score of ~70 is now in the covalently bonded distribution ~116, for a difference of 46. Surprisingly, this matches with the bond energy change of Δscore 47 and the approximate heuristic of 6.5 per 1 kcal. (Details of covalent interaction is included in the GOLD output when the per_atom_score flag is 1.)

Figure 8: Score distribution of Nirmatrelvir N40 covalently bound to CYS 145 S.



The strength of individual atoms change pre- and post- covalent binding. This analysis is possible by using a distributional analysis of the docking runs, and in the highest binding modes of each. The non-covalently bound structure in Figures 1(a) or 6 changes into that of covalently bound in Figure 9, where all poses with scores ≥ 97 are used. The atomic total score distributions are shown in Figure 10.

Figure 9: Covalently bound Nirmatrelvir of the nitrile N40 to the sulfur in Cysteine 145 of Mpro, a thioidimate. The 8146 docking runs with scores ≥ 97 of the covalently bound molecule give atomic score distributions labeled by color.

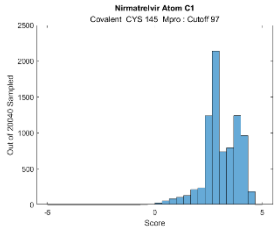


Upon comparing the color coded bound Nirmatrelvir in Figure 9 with that of the non-covalently bound, it is seen that most of the interactions haven't changed much. The region near the functional group containing C39 and the sulfur have atom scores generally increased. This contains C32, C31, N28, and C26. Also the anchor point of O12 has shifted to extremely high. Overall, the anchor points in purple and red haven't changed and a couple of atoms have increased, O12 and C31.

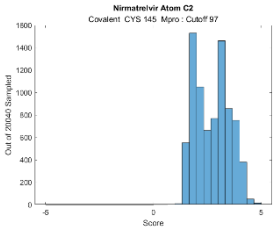
Furthermore, there are generally more bi-modal distributions in the highest scoring covalently bound conformation than the non-covalent, as seen in the color coded Figures 6 and 9. There is less flexibility in the covalently bound highest scoring conformation and as a result there is more specificity just by less room for the molecule to flex into. Any structural modifications of Nirmatrelvir could benefit from detailed knowledge of the atomic score distributions in both binding poses.

Figure 10: Individual atom score distributions of the covalently bound Nirmatrelvir to Mpro, from poses with total scores in the first mode, ≥ 97, having 8146 total in each histogram. Atom scores are shown from -5 to 5 in the distributions. The atoms C29, C39, S40, and N41 are not listed due to their proximity to the covalent bonding and interpretation from ligand atom and protein atom scores.

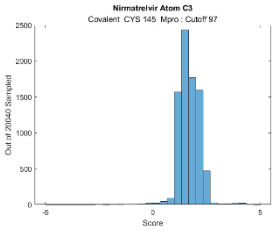
C1



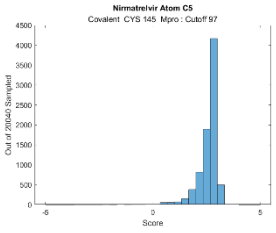
C2



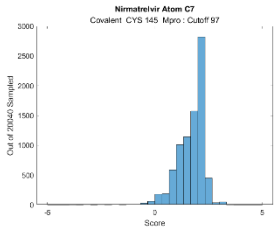
C3



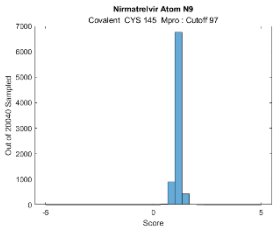
C5



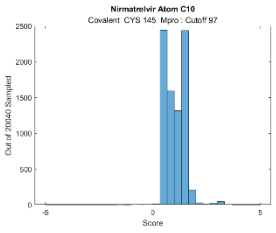
C7



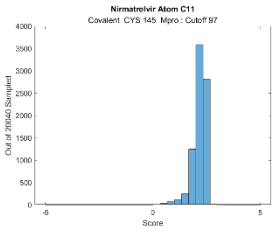
N9



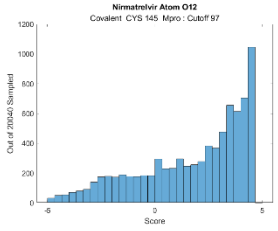
C10



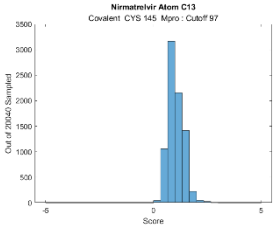
C11



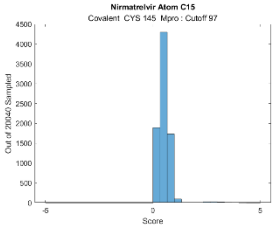
O12



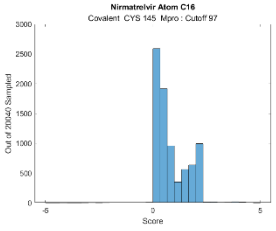
C13



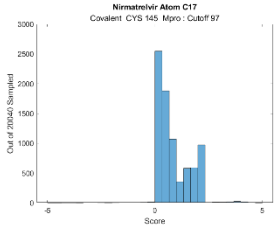
C15



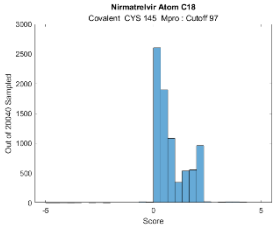
C16



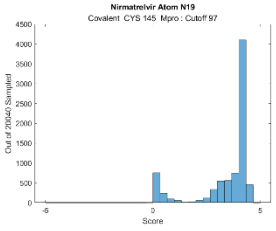
C17



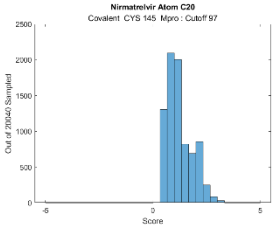
C18



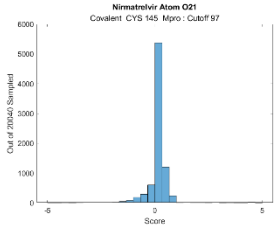
N19



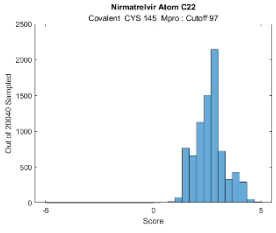
C20



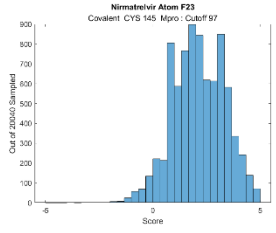
O21



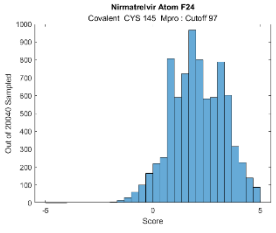
C22



F23



F24

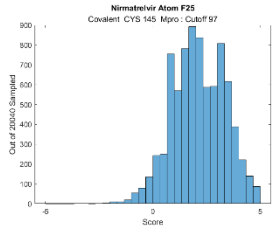


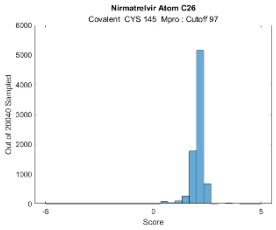
F25

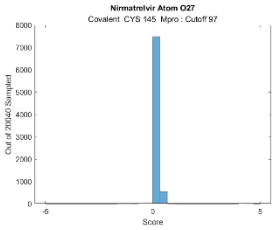
C26

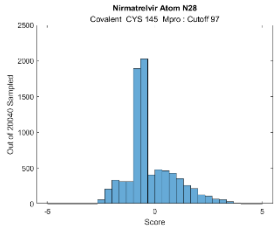
O27

N28

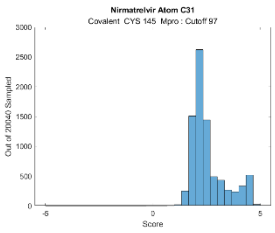


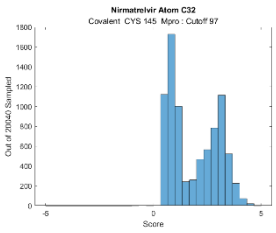


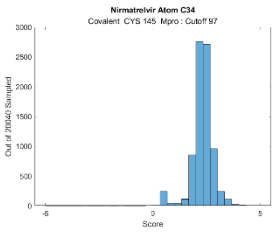


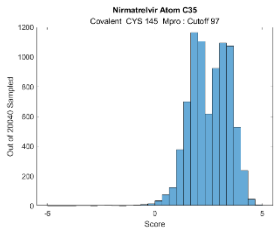


C31 C32 C34 C35

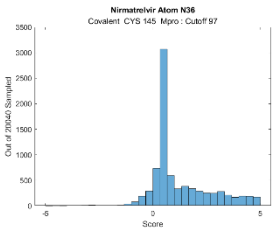


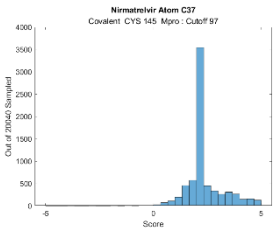


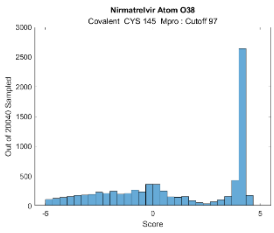


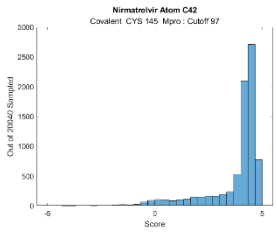


N36 N37 O38 C42









Note that O27 is almost zero in interaction as it was in the non-covalent analysis. Atom O27 is not interactive in Nirmatrelvir. It has no role in interaction with Mpro as far as binding is concerned.

N28, with its hydrogen, is mostly negative in the covalently bound molecule where it was slightly interacting (<1.5) in the non-covalent, but not much negative with score still ≥ -1.5. This score does include a hydrogen bond contribution in which this atom contributes in the covalent binding (next in paper). Next, O38 and O12 also have a negative component in their distributions with a wider spread. These 2 also have hydrogen bond contributions, as acceptors, and all 3 are polar adjacent to nonpolar amino acids. The negativity means that 1 to 3 of these atoms are repulsive in the overall binding of some of the ligands in the highest binding mode of score ≥ 97.

Covalent Nirmatrelvir Hydrogen bonds

Overall there is good agreement with the hydrogen bond results of the analysis with the x-ray structure results in [3][6]. There are 6 ligand-protein hydrogen bonds in the pose at 112.02,

Donor molecule	Donor type	Donor heavy atom	Donor atom	Acceptor molecule	Acceptor atom	Score	Amino acid
L1	H	19	57	P1	1315	.99	GLU 166
L1	H	28	58	P1	1297	.83	HIS 164
L1	H	36	65	P1	1123	.14	LEU 141
P1	H	1145	2466	L1	38	.05	SER 144
P1	H	1146	2467	L1	41	.94	CYS 145
P1	H	1312	3710	L1	12	.66	GLU 166

of which 2 are weak, LUE 141 and SER 144, and a total of 3.66. In the isolated binding pose at 116.32 there are 4 hydrogen bonds with one different,

Donor molecule	Donor type	Donor heavy atom	Donor atom	Acceptor molecule	Acceptor atom	Score	Amino acid
L1	H	28	58	P1	1297	.33	HIS 164
L1	H	36	65	P1	1112	.94	PHE 140
P1	H	1146	2467	L1	41	.92	CYS 145
P1	H	1312	3710	L1	12	.68	GLU 166

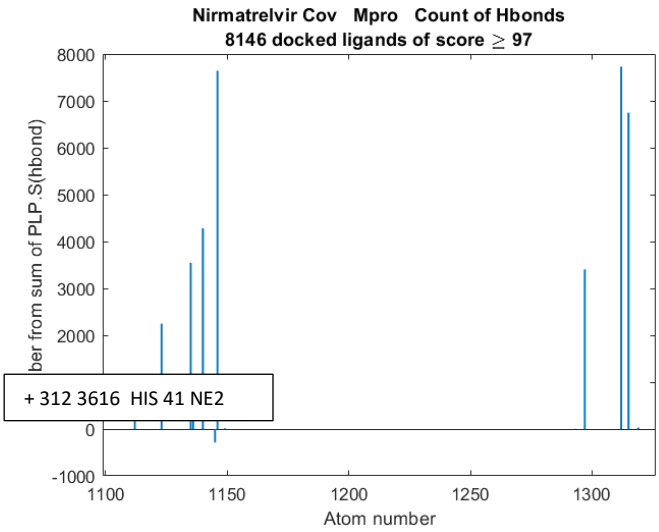
for a total of 2.87. Both of these modeled covalently bound poses have the CYS 145 (1 bond in [3][6]) and GLU 166 (1 bond in [3] and both bonds in [6]), which is also present in the analysis of [6] (figure 4f) and [3] (figure 2e,f). Both also have HIS 163 which is not present computationally, and HIS 163 in the analysis of [6] and not pointed out in [3] but is in the binding here. There is a weak SER 144 bond in [3] and there is also here. PHE 140 is in the computational binding model but not either of the these papers. GLN 189 is not but THR 190 is there and weak [6].

Instead of examining single docking poses, the hydrogen bonds are examined distributionally in the range [97,115] (and checked in [110,115]). The histograms are similar in both and the only atoms which have these are O12, N19, N28, N36, and N41. N36 and N41 each have 2 bonds in half of the poses. These 5 atoms are the primary ligand hydrogen bonded atoms in [3][6].

The Mpro amino acid residues involved from the 8146 docked ligands with total score ≥ 97 are listed next. These are counted using column 5 of the Gold.PLP.Protein.Score.Contributions in the GOLD output file. (Although uncommon, modeled hydrogen bond interactions can be negative due to their relative angular bond and atom type depedence.) The number per amino acid is,

Atom number	Count	Amino acid	Heavy atom
312	2616	HIS 41	NE2
1112	936	PHE 140	O
1123	2248	LEU 141	O
1135	3547	ASN 142	ND2
1136	871	GLY 143	N
1140	4282	SER 144	N
1145	-287	SER 144	OG
1146	7643	CYS 145	N
1297	3404	HIS 164	O
1312	7732	GLU 166	N
1315	6748	GLU 166	O
1496	17	GLN 189	NE2

Figure 11: Hbonds shown as a count over the Mpro atoms from 1100 to 1325. The atom numbers 312 and 1496 are outside this range. The distinct groupings are neighboring aminos. Not shown is HIS 41 with total 2616.



The covalent highest binding mode is characterized by,

HIS 41, LEU 141, ASN 142, SER 144, CYS 145, HIS 164, GLU 166 and by ½ PHE 140, GLY 143

Only ASN 142 and GLU 166 are major components in both non-covalent and covalent binding, and GLY 143 present in ~10% in the latter. This difference could be due to the spatial closeness of the covalently bound ligand to Mpro.

Their presence in the computational binding generally agree with [3] and [6] and there are a few non-matches. [6] has HIS 163, MET 165, and THR 190 (to a fluorine) listed in figure 4(f). Several amino acids appear in the computation that aren't listed in [3][6]: LEU 141, ASN 142, SER 144 in [3].

The modeling presented here is over a range of scores, within the primary docking mode, and the x-ray structure is from a low entropic energetic state of the protein-ligand complex. The use of distributions over a small range of protein-ligand energies matches if one considers the wiggle of an x-ray determined epitope within the cavity. Any small discrepancies could be due to 1) fixed x-ray structure versus dynamic modeling, 2) the use of different x-ray structures, 3) computational modeling of the bound ligand including additional explicit waters. Lastly, in [6] (whose x-ray structure 7SI9 is used here) the relevance of Mpro's HIS 164 in the structure activity is pointed out in their analysis, and the modeled binding here with Mpro contains this amino acid.

The distance contacts $\leq 3.2\ A$ stated as hydrogen bonds in [3] and [6] aren't necessarily hydrogen bonds in the GOLD computational modeling; hydrogen bonds require not only distance but also angular bond orientation (being a dipole-dipole interaction) depending on the donor and acceptor atom type. There are distances reported in [3][6] (figure 2e,f and 4f) which are not reported as hydrogen bonds here, but the hydrogen bond agreement generally matches between x-ray and computational.

Entropic contribution to free energy

From the internal binding energy and entropy, the Hemholtz free energy F=E-TS can be calculated and also broken into conformational modes using the distribution of states. In this overview of Paxlovid, the entropy calculation and discussion is mostly taken from a concurrent paper [18]. The GOLD PLP docking distributions are calculated in terms of score. An approximate linear relation, heuristically valid across

scores 40 to 90 in different protein-ligand complexes, between energy and score is used to translate these into energy (in kcals) $E = \gamma Z$, with Z the docking score.

Two example calculations are given, non-covalently bound Nirmatrelvir and covalently bound. The zero energy state is referred to as free ligand. The entropic energy null contribution of ground state is made by assuming a unique ground bound state. The numerically calculated distribution is used, and no curve is used or required. An example calculation of E and S is given from the distribution in Figure 3(a), Nirmatrelvir to Mpro. The constants are

$$\text{k\_B} = 1.38\ 10^{-23}\ \text{J/K} \qquad \text{eqn 4}$$

$$\text{mole} = 6.02\ 10^{23} \qquad \text{eqn 5}$$

$$1\ \text{J} = 2.3901\ 10^{-4}\ \text{kcal} \qquad \text{eqn 6}$$

and at T=298 K,

$$1/k_BT = 1.86\ \text{mol/kcal}, \quad \exp(1\ kcal/k_BT) = 6.45 \qquad \text{eqn 7}$$

6.5 is the heuristic coefficient between score and energy given x10 for 1 kcal in various IC50 measurement comparisons. This is at room temperature. Equations 6 and 8 are, using the numerical distribution with 200 bins,

$$\langle E(298\ K)\rangle = -10.19\ \text{kcal/mol} \qquad \text{eqn 8}$$

$$-T\langle S\rangle = 1.56\ \text{kcal/mol} \qquad \text{eqn 9}$$

for a total of -8.63 kcal/mol. Binding estimates of Nirmatrelvir to the main protease from MD simulations range in the -7.7 to -10 range.

The probabilities of the states less than about 58 are negligibly populated, being about 12/6.5=1.84 kcal, or one width of 1 kcal/k_BT and this is seen in the calculated distribution and thermally weighted Boltzmann $p(Z_i)$ values [18]. The mean of the distribution is score 49 and to get that from an expected energy requires T well beyond the boiling point of water, in the tens of thousands of Kelvins. Nirmatrelvir does have 9 rotatable bonds, is somewhat flexible even when bound, and these aspects show in the distributions of Figures 3(a) and 8; this shows up in the size of the entropic term, which is still much smaller but non-negligible in the total free energy.

In the covalent bonding of Nirmatrelvir to Mpro, the Mpro CYS 145 forms a thioimidate adduct from the nitrile bonding, -C(=N)S-, through the reaction

$$\text{-C(C\#N)- + S- - -> -C(C(=N)S-145)-} \qquad \text{eqn 10}$$

In the case of covalently bound Nirmatrelvir and using its numerical distribution Figure 8,

$$\langle E(298\ K)\rangle = -16.53\ \text{kcal/mol} \qquad \text{eqn 11}$$

$$-T\langle S\rangle = 2.17\ \text{kcal/mol} \qquad \text{eqn 12}$$

for a total of -14.36 kcal/mol. The covalent in-bond energy change from eqn 2 is 7.17 kcal/mol. There are 2 additional rotatable bonds in the covalent coming from the thiodinate; the entropy is slightly more

and the distribution in Figure 6 less defined. The entropic contribution of the covalently bound Nirmatrelvir is 13% of the internal binding energy, and in the non-covalently bound molecule it is 15.3%.

The distributions are reproducible and valid at any temperature. The accuracy of the energy calculation depends on the coefficient from score to energy, γ, and the linearity. Both of these appears correct to 1st order in the relation of docking score to experiment.

Results: Ritonavir

Studies of the metabolism of Nirmatrelvir revealed that it doesn't last for long in the body. The liver enzymes are actively working fast with it. As a result, plasma concentrations of it decrease too rapidly. 2 possible solutions to maintaining sufficient presence to have an inhibitory effect on viral replication via the target Mpro are: 1) use higher and more frequent doses, 2) add an inhibitor to the garbage collection process by blocking the action of the CYP enzymes, in particular 3A4. Both have drawbacks: 1) unwanted interactions with other proteins, 2) blocking the liver enzymes from normal function could result in increased levels of other medications to the point of toxicity. Dose changes for either purpose obviously means more interactions with a variety of proteins, not only Mpro.

Ritonavir (Norvir) is a well-known inhibitor of CYP 3A4, and most of the CYP liver enzymes. Initially developed for treating HIV, it became well known as a CYP inhibitor and is currently used in conjunction with other drugs to maintain or raise their concentrations in the body. It is a very large small molecule that breaks most of Lipinski's Rule of 5 ADME conditions. It has mass of 720 Da and 18 rotatable bonds. It is so flexible and large that it potentially sticks to almost anything; it is not a precision small molecule inhibitor of a protease. Ritonavir is also a protease class specific inhibitor and was considered in conjunction with Liponavir as an inhibitor of Mpro due to latter's cavity structure to that of the Hepatitus C protease.

The canonical representation of Ritonavir is,

CC(C)C1=NC(=CS1)CN(C)C(=O)N[C@@H](C(C)C)C(=O)N[C@@H](CC2=CC=CC=C2)C[C@@H]([C@H](CC3=CC=CC=C3)NC(=O)OCC4=CN=CS4)O eqn 13

Molecular properties of Ritonavir are:

Mass (Da)	Heavy atoms	Rotat. Bds	Hbond-don	Hbond-acc	P_{ow}
720	50	18	4	9	6

Ritonavir was docked both to Mpro and to CYP 3A4. Docked to Mpro and CYP 3A4

	Score	S(PLP)	S(hbond)	S(cho)	S(metal)	DE(clash)	DE(tors)	Intcor
Mpro	100.37	-99.24	2.90	0	0	3.85	2.61	1.49
3A4	142.18	-133.97	2.00	0	.99	0	3.24	2.77

The molecule and its image, its highest docking pose to Mpro, and its highest scoring pose to CYP 3A4 are shown in Figure 12. The score distributions of both protein docking runs is given in Figure 13. The distributions are wide and very normal; this is due to the flexibility of 18 rotatable bonds giving it the

ability to stick in almost any conformation. The peaks/widths pertaining to Mpro and CYP 3A4 are centered at 57 and 102.

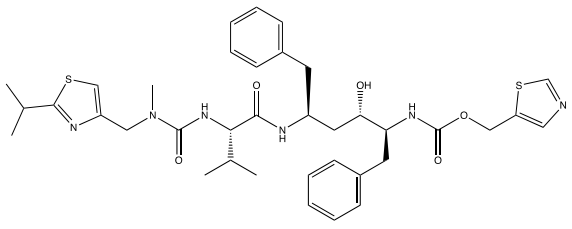
The docking of Ritonavir to CYP 3A4 results in an extremely high PLP docking score, 142. This confirms its use as an effective inhibitor. It essentially fills the cavity adjacent to the heme iron, see Figure 12(d). Note that it has 50 heavy atoms and its size is primarily responsible for the high PLP docking (interaction) score. Per atom the score is 2.56 but the filling of the cavity and the number of atoms is the primary reason for its large interaction with CYP 3A4. By contrast, Nirmatrelvir has 35 atoms, a score of 84.11, and a similar score per atom of 2.41 to CYP 3A4.

The SARS-Cov-2 Mpro is structurally similar to a class of proteases that Ritonavir binds to, and Liponavir-Ritonavir, has been examined in inhibiting Mpro. The highest docking score is 100.37 and appears quite much, especially when compared to Nirmatrelvir's score of 68 or 70.7. There aren't any isolated conformers. However, as mentioned, Ritonavir is large and flexible and can generally interact and fit somewhat to general proteases and proteins. Upon inspection of the docked Ritonavir to Mpro, it extends into one X-cave and the surface tunnel. It partly enters another X-cave and covers over, not in, the 3rd cave. It possibly enters the target region somewhat ecessary to inhibit biochemical viral replication, but it is so flexible and non-specific that it will in biological conditions flex into almost any orientation nearby and almost freely considering the distribution in Figure 13 and its width. This is a good example of why localization of small molecule to the protein is more important than a high docking score in realistic molecular design, however both being important. Specificity, however, does usually go with high score. Ritonavir was found in conjunction with Liponavir, and also in other combinations, not to be an effective drug in blocking SARS-Cov-2 [25].

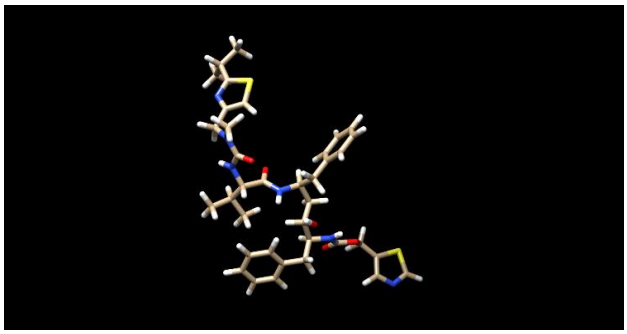
Ritonavir's use in blocking the metabolism of other drugs is well known. However, its large effect of 'stuffing the waste basket until full' of CYP 3A4 (and other CYP liver enzymes) so that that a molecule can not be cleared by the liver has the disadvantage of blocking a variety of normal liver functions and of other drugs/small molecules. This a reason that the drug Paxlovid is FDA approved only for high risk groups, i.e., those individuals with predisposition to severe Covid-19, who can accept the risk of Ritonavir. This disadvantage could be eliminated or alleviated by finding other potential small molecule candidates that don't require Ritonavir or other drugs to inhibit ordinary bodily function. This can be accomplished by finding molecules with higher non-covalent interaction with the Mpro cavity and less to the CYP 3A4 enzyme, whether covalent or non-covalent.

Figure 12: (a) Ritonavir. (b) Ritonavir in Chimera. (c) Ritonavir in its highest docked pose to Mpro at 100. (d) Ritonavir bound to CYP 3A4 at score 142.

(a)



(b)



(c) (d)

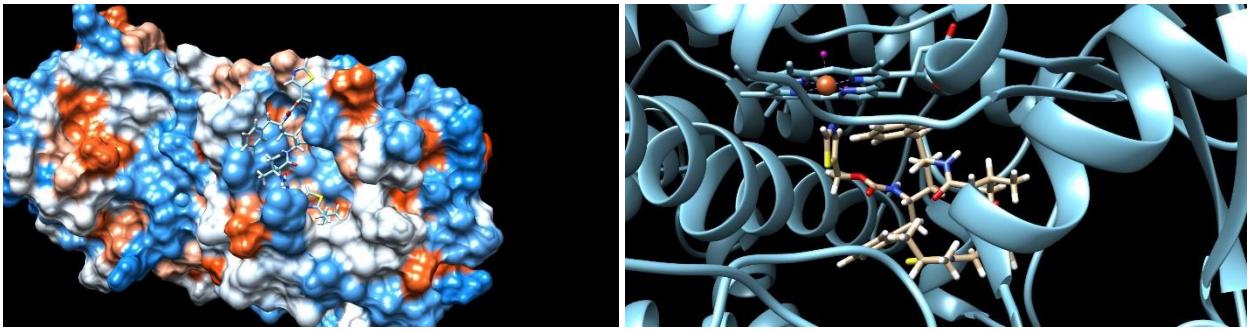
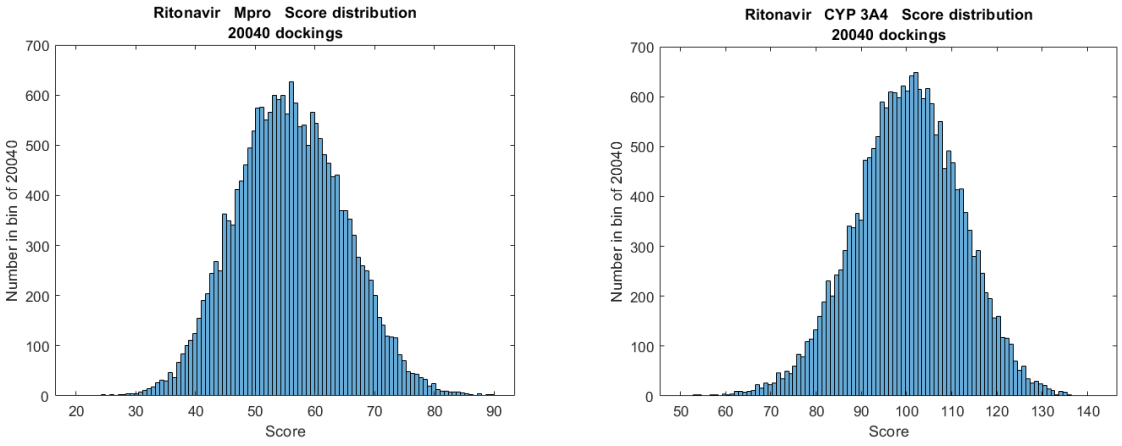


Figure 13: Score distributions of (a) Ritonavir/Mpro and (b) Ritonavir/CYP 3A4. Due to its flexibility, 18 rotatable bonds, and size of 50 heavy atoms, Ritonavir is not specific as seen in the single peak width. It is an effective blocker of the CYP 3A4 cavity however, and as a result, an activator of other co-used molecules.

(a) (b)



Discussion

Paxlovid and its 2 ingredients, Nirmatrelvir and Ritonavir, are examined computationally in interaction to the SARS-Cov-2 3CL (Mpro) protease and the liver enzyme CYP 3A4. The pre-covalent and covalent binding of Nirmatrelvir to Mpro is considered, the non-covalent ligand to CYP 3A4 and the binding of Ritonavir to Mpro and CYP 3A4 are examined in detail using large sets of docking runs. Ligand conformations, binding modes, atomic-level interactions, hydrogen bonding, and overall interaction and types are used to characterize in detail the protein-ligand interactions. Docking distributions are used to quantify the different interactions and conformations as opposed to a single highest scoring docked ligand. The use of distributions places the analysis in an interacting protein-ligand dynamics setting including rotatable amino acid side-chains. This is without the use of molecular dynamics and with orders of magnitude less computation and time spent.

Comparison of modeled Nirmaltrevir with the 2.0 Angstrom x-ray structure 7SI9 is made. There is good agreement spatially with the highest binding mode, both when non-covalent and covalently bound. In

general, the physical pre-covalent binding orientation of a ligand is important to facilitate the chemical covalent reaction. The ligand-cavity hydrogen bonding is examined and it agrees with 2 x-ray covalently bound structures (7SI9, and including those of 7RFR, 7RFS, 7RFU, 7RFW) due to the distributions and specificity of highly scoring epitopes. There are differences between the non-covalent and covalent binding, especially in the hydrogen bonding, and these are pointed out.

The molecule Nirmatrelvir is not very strongly non-covalently interacting with the relevant Mpro cavity to inhibit viral replication of SARS-Cov-2 virus, due to a score per atom of 2.0. It does have 2 important features: 1) it has separated anchor points of high interaction with the Mpro protein which attract it in an orientation to 2) covalently bind via a thionate reaction with CYS 145, generating a thioidinate adduct. Then these high interaction anchor points are maintained after binding. The energy gained in covalent binding is 7.1 kcal. Using atomic interaction distribution analysis, the highest binding conformational mode pre-covalently is shown atomically to be similar in interaction to the post-covalent binding. The hydrogen bonding across the amino acids, although maintaining 2 aminos out of the sets in contributing hydrogen bonds, is different. These features are relevant in determining stability and lifetime of the covalently bound molecule, Nirmatrelvir. In addition, it is noticed that there are interesting points in the molecule in possible alteration of the overall non-covalent binding and possible pharmacological properties.

This work gives a detailed analysis of a protein-ligand system. In this case, it is about an EUA drug, Paxlovid (Nirmatrelvir and Ritonavir) for treating early infection from SARS-Cov-2. The method can be used with any complex and is informative. Results from this type of analysis can be used to improve, via molecular modification, the binding of molecules, pharmacological properties, and also the identification of relevant conformations and their populations in different environments. Additionally, if a molecule is projected to be a covalent binder, then this type of analysis can be used together with the non-covalent analysis to modify it ahead of and including covalent binding to a better pharmaceutical.

Acknowledgements

G.C. thanks Christian Heiss, James H. Prestegard, and David Crich for useful discussions. G.C. is grateful for the use of the Sapelo2 cluster at the Georgia Advanced Computing Resource Center and for the hospitality at the Complex Carbohydrate Research Center.

Statements and declarations

There are no conflicts of interest. Several paragraphs (half a page) and 3 figures were taken from [18], the latter written for a different purpose, that of a docking methodology. This paper is a much greater and expanded analysis than [26], which has been split into 3 new papers and no longer submitted for publication.

Supplementary information

Contact the author for any files or scripts used in this work and paper. Each set of protein-ligand docking runs makes many files, which are well-organized, but totaling GBs for each set.

References

1. PDBe-KB COVID-19 Data Portal at Protein Data Bank in Europe. (2021). Retrieved from: https://www.ebi.ac.uk/pdbe/covid-19

2. Hall MD, Anderson JM, Anderson A, Baker D, Bradner J, Brimacombe KR, Campbell EA, Corbett KS, Carter K, Cherry S, Chiang L, Cihlar T, de Wit E, Denison M, Disney M, Fletcher CV, Ford-Scheimer SL, Götte M, Grossman AC, Hayden FG, Hazuda DJ, Lanteri CA, Marston H, Mesecar AD, Moore S, Nwankwo JO, O'Rear J, Painter G, Saikatendu KS, Schiffer CA, Sheahan TP, Shi P, Smyth HD, Sofia MJ, Weetall M, Weller SK, Whitley R, Fauci AS, Austin CP, Collins FS, Conley AJ, Davis MI. *The Journal of Infectious Diseases* **2021**, 224(S1)1:S1–S21. https://doi.org/10.1093/infdis/jiab305

3. Owen DA, et. al. An oral SARS-Cov-2 Mpro inhibitor clinical candidate for the treatment of Covid-19. *Science*, 374(575): 1586-1593, 2 Nov 2021. DOI: science.org/doi/10.1126/science.abl4784.

Making Paxlovid. Retrieved from: https://www.science.org/content/blog-post/making-paxlovid

4. Fischer W, et. al. Molnupiravir, an Oral Antiviral Treatment for COVID-19. 17 June 2021. DOI: https://doi.org/10.1101/2021.06.17.21258639/ PMCID: PMC8219109 PMID: 34159342

Molnupiravir. Retrieved from PubChem: https://pubchem.ncbi.nlm.nih.gov/compound/145996610.

5. Nirmatrelvir/Mpro complexes:

PDB ID 7SI9. Retrieved from the Protein Data Bank: https://www.rcsb.org/structure/7SI9

PDB ID 7VH8. Retrieved from the PDB: https://www.rcsb.org/structure/7VH8

PDB ID 7TEO. Retrieved from the PDB: https://www.rcsb.org/structure/7TE0

6. Kneller DW, Li H, Phillips G, Weiss KL, Zhang Q, Arnould MA, Jonsson CB, Surendranathan S, Parvathareddy J, Blakeley MP, Coates L, Louis JM, Bonnesen PV, Kovalevsky A. Covalent narlaprevir- and boceprevir-derived hybrid inhibitors of SARS-CoV-2 main protease. *Nat. Commun*. **2022**, 13: 2268. DOI: 10.1038/s41467-022-29915-z, PMID: 35477935, PMCID: PMC9046211

7. Nirmatrelvir predecessor PF-00835231. Retrieved from PubChem: https://pubchem.ncbi.nlm.nih.gov/compound/11561899.

8. P450 (CYP) enzymes. Retrieved from: https://en.wikipedia.org/wiki/Cytochrome_P450#:~:text=Cytochromes%20P450%20(CYPs)%20are%20a,for%20hormone%20synthesis%20and%20breakdown.

9. Meunier B, de Visser SP, Shaik S. Mechanism of oxidation reactions catalyzed by cytochrome P450 enzymes. *Chem. Rev.* 2004, 104, 9, 3947–3980. DOI: 10.1021.cr020443g

10. McDonnell AM, Dang CH. Basic review of the cytochrome P450 system. *J. Adv. Pract. Oncol.* 2013 Jul-Aug; 4(4): 263–268. DOI: 10.6004/jadpro.2013.4.4.7, PMID: 25032007, PMCID: PMC4093435

11. Drug interactions with Ritonavir. Retrieved from: https://www.covid19treatmentguidelines.nih.gov/therapies/statement-on-paxlovid-drug-drug-interactions/

12. Ritonavir side effects. Retrieved from: https://clinicalinfo.hiv.gov/en/drugs/ritonavir/patient

Retrieved from: https://medlineplus.gov/druginfo/meds/a696029

13. Paxlovid FDA authorization drug sheet. Retrieved from: https://www.fda.gov/media/155050/download/

14. Sadowski J, Gasteiger J., Klebe G. Comparison of Automatic Three-Dimensional Model Builders Using 639 X-Ray Structures. *J. Chem. Inf. Comput. Sci.* **1994**, 34,4, 1000-1008 DOI: 10.1021/ci00020a039

Schwab CH. Conformations and 3D pharmacophore searching. *Drug Discovery Today: Technologies* 2010, 7(4), Winter **2010**, e245-e253 DOI: 10.1016/j.ddtec.2010.10.003

Molecular Networks GmbH, Altamira, LLC. (2021). Corina. Retrieved from MN-AM Corina: https://www.mn-am.com/products/corina

Schwab CH. Molecular Structure Representation in Chemoinformatics Applications, Schwab CH, BigChem Autumn 2017 School, Modena, Italy. Received from http://bigchem.eu/sites/default/files/School3_Schwab.pdf

15. Weininger D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci*. **1988**, 28(1), 31-36. DOI:10.1021/ci00057a005

Weininger D, Weininger A, Weininger J. SMILES. 2. Algorithm for generation of unique SMILES notation. *J. Chem. Inf. Comput. Sci.* **1998**, 29(2), 97-101. DOI:10.1021/ci00062/a008

Weininger D. Smiles. 3. Depict. Graphical depiction of chemical structures. *J. Chem. Inf. Comput. Sci.* **1990**, 30(3), 237-243. DOI:10.1021/ci00067a005

16. Jones G, Willett P, Glen RC, Leach AR, Taylor R. (1997). Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, 267(3), 727-748. DOI: 10.1006/jmbi.1996.0897 PMID: 9126849

Cambridge Crystallographic Data Centre. (2021). CCDC Discovery GOLD. Retrieved from GOLD Protein Ligand Docking Software: https://www.ccdc.cam.ac.uk/solutions/csd-discovery/Components/Gold/

17. MathWorks, Inc. MATLAB. (2020b). Retrieved from MathWorks: https://www.mathworks.com/

18. Chalmers G. Dynamic docking in protein-ligand modeling. Submitted to Journal of Computational Chemistry, June 2022. DOI: 10.26434/chemrxiv-2022-6m0q5

19. Lipinski CA, Lombardo F, Dominy WB, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **2001**, 46(1-3), 3-26. DOI: 10.1016/s0169-409x(00)00129-0

Lipinski C. Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discovery Today Technologies* **2004**, 1(4), 337-341. DOI: 10.1016/j.ddtec.2004.11.007, PMID: 24981612

20. Ghose AK, Viswanadhan VN, Wendoloski JJ. A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. *J. Comb. Chem*. **1999**, 1(1), 55-68. DOI: 10.1021/cc9800071, PMID: 10746014

21. Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kobble KD. Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* **2002**, 45(12), 2615-23. PMID: 12036371 DOI: 10.1021/jm020017n, PMID: 12036371

22. Congreve M, Carr R, Murray C, Jhoti H. A 'rule of three' for fragment-based lead discovery? *Drug Discov. Today* **2003**, 8(19), 876-877. DOI: 10.1016/S1359-6446(03)02831-9, PMID: 14554012

23**.** Benet LZ, Hosey CM, Ursu O, Oprea TI. BDDCS, the Rule of 5 and Drugability. *Adv Drug Deliv Rev*. 2016, 101, 89-98. DOI: 10.1016/j.addr.2016.05.007, PMID: 27182629, PMCID: PMC4910824

24. Møllendal H. A Microwave and Quantum Chemical Study of (Trifluoromethyl)thiolacetic Acid, CF3COSH, a Compound with an Unusual Double-Minimum Potential. *J. Phys. Chem. A* **2007**, 111, 1891-1898. DOIi: 10.1021/jp0677290, PMCID: PMC7486594, PMID: 32918656

25. Cao B, Wang Y, Wen D, Liu W, Wang J, Fan G, Ruan L, Song B, Cai Y, Wei M, Li X, Xia J, Chen N, Xiang J, Yu T, Bai T, Xie X, Zhang L, Li C, Yuan Y, Chen H, Li H, Huang H, Tu S, Gong F, Liu Y, Wei Y, Dong C, Zhou F, Gu X, Xu J, Liu Z, Zhang Y, Li H, Shang L, Wang K, Li K, Zhou X, Dong X, Qu Z, Lu S, Hu X, Ruan S, Luo S, Wu J, Peng L, Cheng F, Pan L, Zou J, Jia C, Wang J, Liu X, Wang S, Wu X, Ge Q, He J, Zhan H, Qiu F, Guo L, Huang C, Jaki T, Hayden FG, Horby PW, Zhang D, Wang. A trial of lopinavir-ritonavir in adults hospitalized with severe Covid-19. *N. Engl. J. Med.* **2020.** DOI: 10.1056/NEJMoa2001282, PMID: 3218746, PMCID: PMC7121492

26. Chalmers, G. Computational study of Paxlovid in Ligand GA. ChemRxiv preprint. DOI: 10.26434/chemrxiv-2022-p2phq. Not in review.