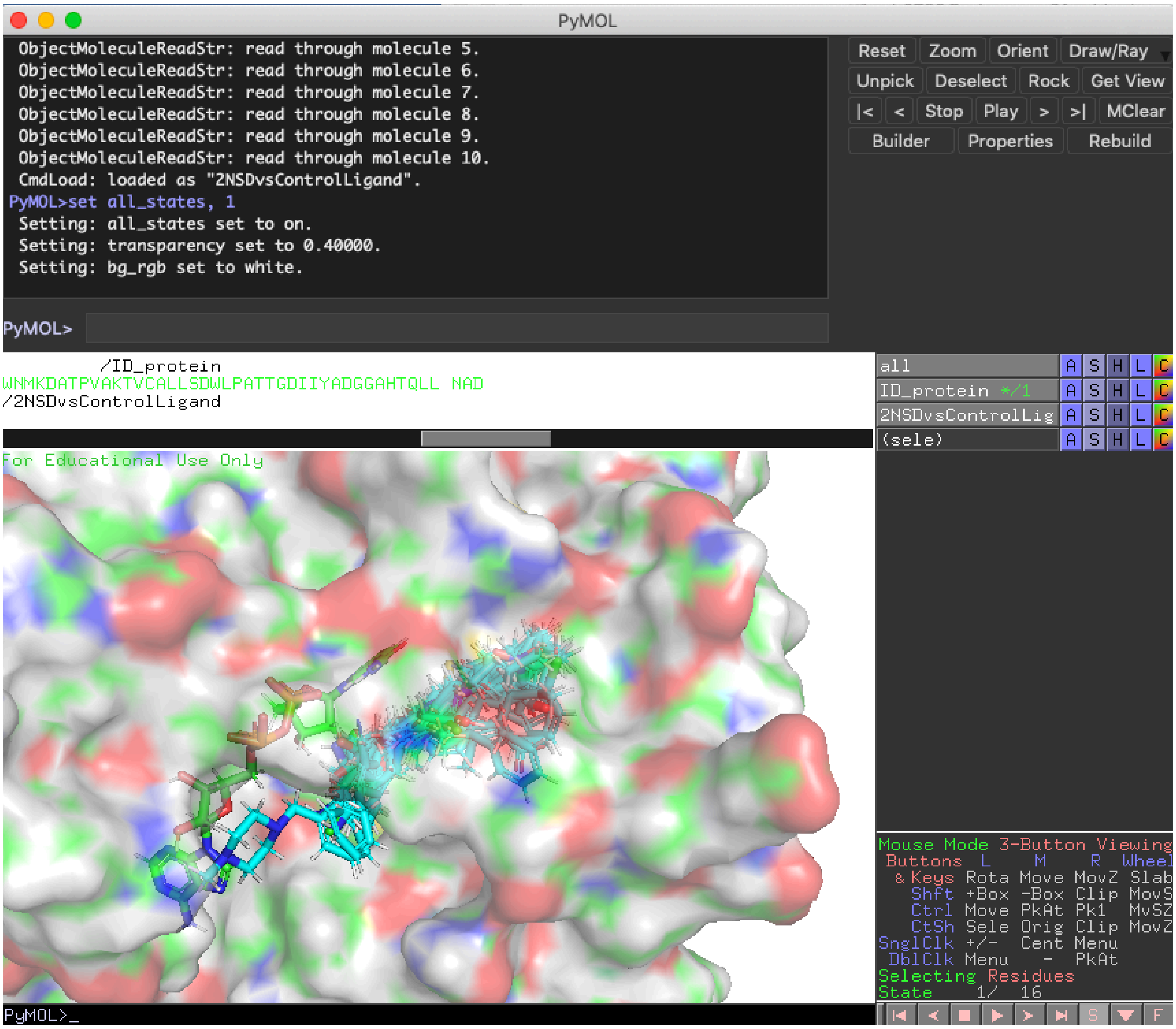
Sandy Bakheet Research Page 2020

Week 5:

July 1st, 2020



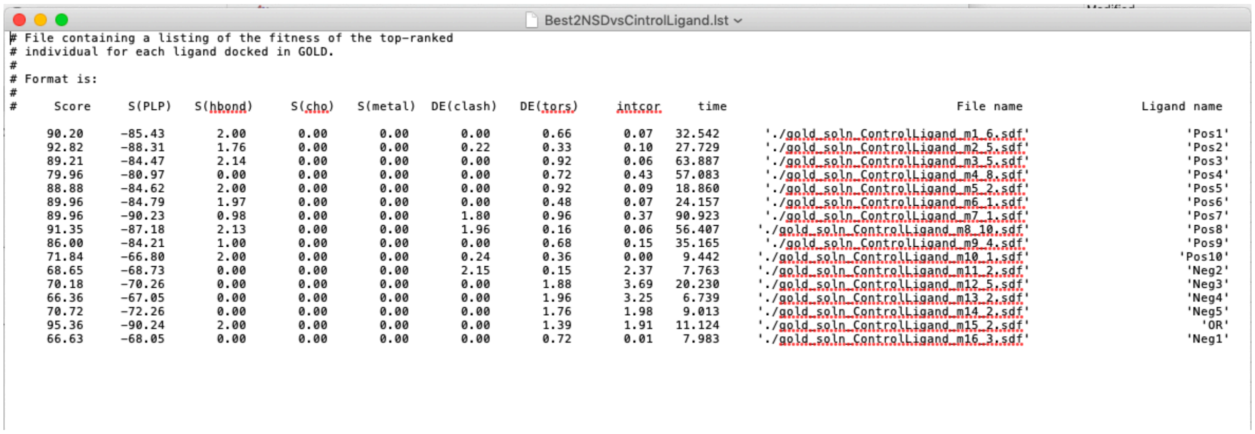
```
ObjectMoleculeReadStr: read through molecule 5.
ObjectMoleculeReadStr: read through molecule 6.
ObjectMoleculeReadStr: read through molecule 7.
ObjectMoleculeReadStr: read through molecule 8.
ObjectMoleculeReadStr: read through molecule 9.
ObjectMoleculeReadStr: read through molecule 10.
CmdLoad: loaded as "2NSDvsControlLigand".
PyMOL>set all_states, 1
Setting: all_states set to on.
Setting: transparency set to 0.40000.
Setting: bg_rgb set to white.
```

```
        /ID_protein
WNMKDATPVAKTVCALLSDWLPATTGDIIYADGGAHTQLL NAD
/2NSDvsControlLigand
```

```
Mouse Mode 3-Button Viewing
 Buttons  L    M    R  Wheel
  & Keys Rota Move MovZ Slab
    Shft +Box -Box Clip MovS
    Ctrl Move PkAt Pk1  MvSZ
    CtSh Sele Orig Clip MovZ
SnglClk +/-  Cent Menu
 DblClk Menu   -  PkAt
Selecting Residues
State      1/  16
```

Figure 1. PyMol image of all the control ligands docked in the active site plus the NADH ligand for the InhA drug target.

Analysis: Overall, all the ligands bound in the ACP site showing good scores. Since the ligands had good docking scores, they show stability through aromatic rings and more Van der Waals interactions.



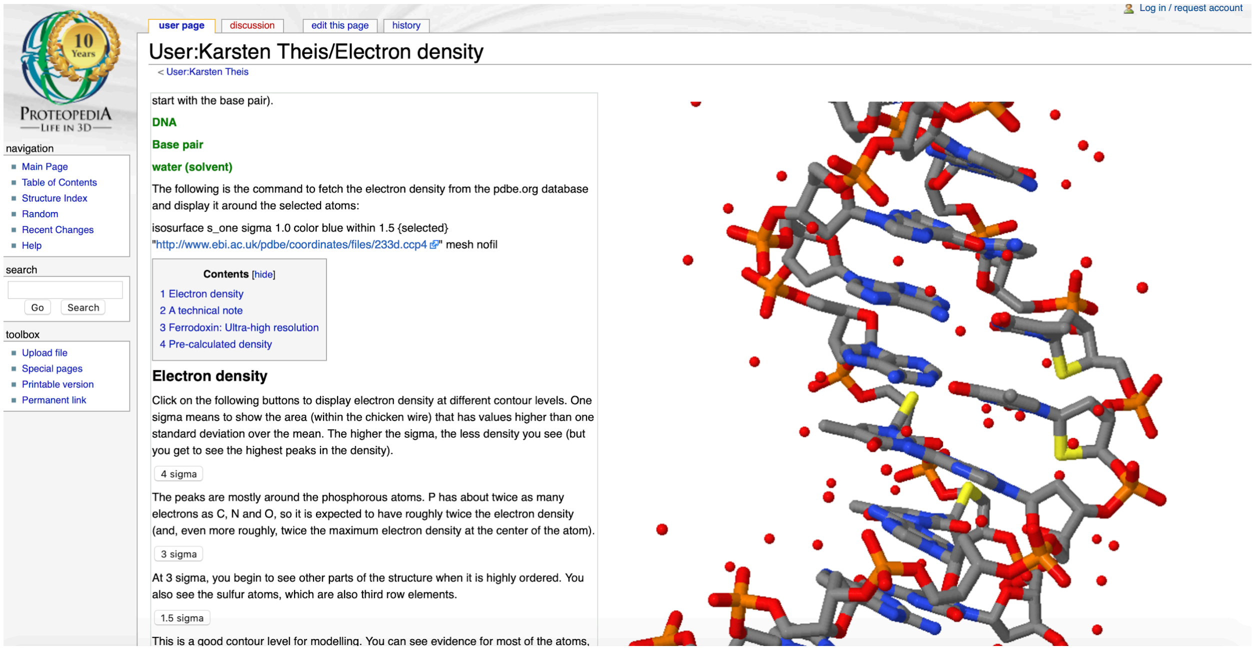
```
# File containing a listing of the fitness of the top-ranked
# individual for each ligand docked in GOLD.
#
# Format is:
#
#      Score      S(PLP)   S(hbond)      S(cho)   S(metal)  DE(clash)   DE(tors)      intcor     time                                   File name                Ligand name

     90.20      -85.43       2.00        0.00       0.00       0.00       0.66        0.07   32.542     './gold_soln_ControlLigand_m1_6.sdf'                        'Pos1'
     92.82      -88.31       1.76        0.00       0.00       0.22       0.33        0.10   27.729     './gold_soln_ControlLigand_m2_5.sdf'                        'Pos2'
     89.21      -84.47       2.14        0.00       0.00       0.00       0.92        0.06   63.887     './gold_soln_ControlLigand_m3_5.sdf'                        'Pos3'
     79.96      -80.97       0.00        0.00       0.00       0.00       0.72        0.43   57.083     './gold_soln_ControlLigand_m4_8.sdf'                        'Pos4'
     88.88      -84.62       2.00        0.00       0.00       0.00       0.92        0.09   18.860     './gold_soln_ControlLigand_m5_2.sdf'                        'Pos5'
     89.96      -84.79       1.97        0.00       0.00       0.00       0.48        0.07   24.157     './gold_soln_ControlLigand_m6_1.sdf'                        'Pos6'
     89.96      -90.23       0.98        0.00       0.00       1.80       0.96        0.37   90.923     './gold_soln_ControlLigand_m7_1.sdf'                        'Pos7'
     91.35      -87.18       2.13        0.00       0.00       1.96       0.16        0.06   56.407    './gold_soln_ControlLigand_m8_10.sdf'                        'Pos8'
     86.00      -84.21       1.00        0.00       0.00       0.00       0.68        0.15   35.165     './gold_soln_ControlLigand_m9_4.sdf'                        'Pos9'
     71.84      -66.80       2.00        0.00       0.00       0.24       0.36        0.00    9.442    './gold_soln_ControlLigand_m10_1.sdf'                       'Pos10'
     68.65      -68.73       0.00        0.00       0.00       2.15       0.15        2.37    7.763    './gold_soln_ControlLigand_m11_2.sdf'                        'Neg2'
     70.18      -70.26       0.00        0.00       0.00       0.00       1.88        3.69   20.230    './gold_soln_ControlLigand_m12_5.sdf'                        'Neg3'
     66.36      -67.05       0.00        0.00       0.00       0.00       1.96        3.25    6.739    './gold_soln_ControlLigand_m13_2.sdf'                        'Neg4'
     70.72      -72.26       0.00        0.00       0.00       0.00       1.76        1.98    9.013    './gold_soln_ControlLigand_m14_2.sdf'                        'Neg5'
     95.36      -90.24       2.00        0.00       0.00       0.00       1.39        1.91   11.124    './gold_soln_ControlLigand_m15_2.sdf'                          'OR'
     66.63      -68.05       0.00        0.00       0.00       0.00       0.72        0.01    7.983    './gold_soln_ControlLigand_m16_3.sdf'                        'Neg1'
```

Figure 2. Bestranking.lst of the docking of Enoyl ACP Reductase is shown with the scores of the positive and the negative control ligands.

Analysis: Overall, the scores were a lot better than I had expected. The highest score is the original ligand followed by the positive controls. This means I chose good positive control ligands in place of the ACP site.

June 28, 2020



User:Karsten Theis/Electron density

< User:Karsten Theis

start with the base pair).

DNA

Base pair

water (solvent)

The following is the command to fetch the electron density from the pdbe.org database and display it around the selected atoms:

isosurface s_one sigma 1.0 color blue within 1.5 {selected} "http://www.ebi.ac.uk/pdbe/coordinates/files/233d.ccp4" mesh nofil

Contents [hide]

1 Electron density
2 A technical note
3 Ferrodoxin: Ultra-high resolution
4 Pre-calculated density

Electron density

Click on the following buttons to display electron density at different contour levels. One sigma means to show the area (within the chicken wire) that has values higher than one standard deviation over the mean. The higher the sigma, the less density you see (but you get to see the highest peaks in the density).

4 sigma

The peaks are mostly around the phosphorous atoms. P has about twice as many electrons as C, N and O, so it is expected to have roughly twice the electron density (and, even more roughly, twice the maximum electron density at the center of the atom).

3 sigma

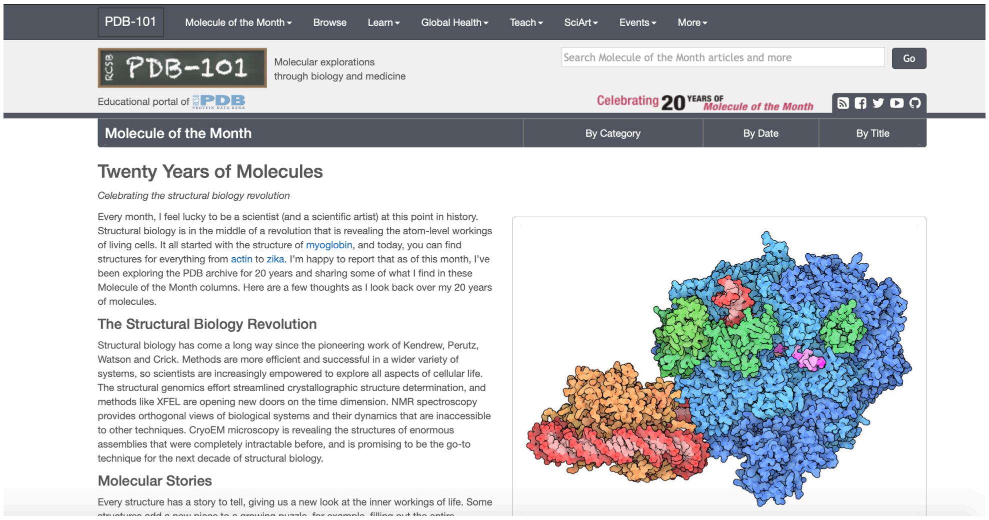
At 3 sigma, you begin to see other parts of the structure when it is highly ordered. You also see the sulfur atoms, which are also third row elements.

1.5 sigma

This is a good contour level for modelling. You can see evidence for most of the atoms,

Figure 1. Image of the proteopedia website showing information about electron density

Analysis: Electron density is higher with lower sigma because the sigma bond can be concentrated more between the two nuclei.



Twenty Years of Molecules

Celebrating the structural biology revolution

Every month, I feel lucky to be a scientist (and a scientific artist) at this point in history. Structural biology is in the middle of a revolution that is revealing the atom-level workings of living cells. It all started with the structure of myoglobin, and today, you can find structures for everything from actin to zika. I'm happy to report that as of this month, I've been exploring the PDB archive for 20 years and sharing some of what I find in these Molecule of the Month columns. Here are a few thoughts as I look back over my 20 years of molecules.

The Structural Biology Revolution

Structural biology has come a long way since the pioneering work of Kendrew, Perutz, Watson and Crick. Methods are more efficient and successful in a wider variety of systems, so scientists are increasingly empowered to explore all aspects of cellular life. The structural genomics effort streamlined crystallographic structure determination, and methods like XFEL are opening new doors on the time dimension. NMR spectroscopy provides orthogonal views of biological systems and their dynamics that are inaccessible to other techniques. CryoEM microscopy is revealing the structures of enormous assemblies that were completely intractable before, and is promising to be the go-to technique for the next decade of structural biology.

Molecular Stories

Every structure has a story to tell, giving us a new look at the inner workings of life. Some

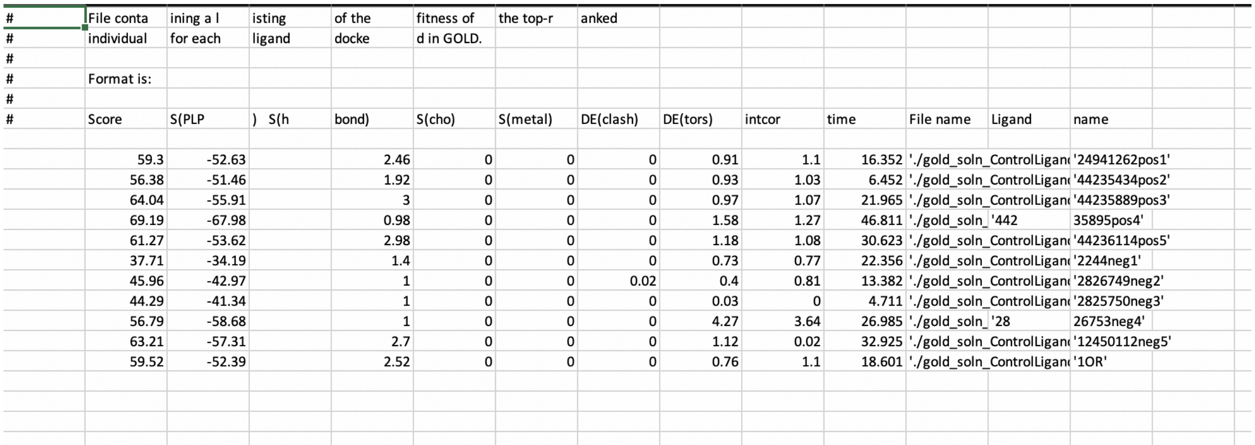
Figure 2. Image of the PDB101 website showing evolution of molecular exploration

Analysis: It is amazing how many molecules that have been pictured and entered into this database has impacted global health.

Week 4:

Good job Sandy!   Good display of all the controls docked into the structure at once. . Thx, - Dr. B

June 26, 2020



#	File conta	ining a l	isting	of the	fitness of	the top-r	anked						
#	individual	for each	ligand	docke	d in GOLD.								
#													
#	Format is:												
#													
#	Score	S(PLP	) S(h	bond)	S(cho)	S(metal)	DE(clash)	DE(tors)	intcor	time	File name	Ligand	name
	59.3	-52.63		2.46	0	0	0	0.91	1.1	16.352	'./gold_soln_ControlLigan		'24941262pos1'
	56.38	-51.46		1.92	0	0	0	0.93	1.03	6.452	'./gold_soln_ControlLigan		'44235434pos2'
	64.04	-55.91		3	0	0	0	0.97	1.07	21.965	'./gold_soln_ControlLigan		'44235889pos3'
	69.19	-67.98		0.98	0	0	0	1.58	1.27	46.811	'./gold_soln_	'442	35895pos4'
	61.27	-53.62		2.98	0	0	0	1.18	1.08	30.623	'./gold_soln_ControlLigan		'44236114pos5'
	37.71	-34.19		1.4	0	0	0	0.73	0.77	22.356	'./gold_soln_ControlLigan		'2244neg1'
	45.96	-42.97		1	0	0	0.02	0.4	0.81	13.382	'./gold_soln_ControlLigan		'2826749neg2'
	44.29	-41.34		1	0	0	0	0.03	0	4.711	'./gold_soln_ControlLigan		'2825750neg3'
	56.79	-58.68		1	0	0	0	4.27	3.64	26.985	'./gold_soln_	'28	26753neg4'
	63.21	-57.31		2.7	0	0	0	1.12	0.02	32.925	'./gold_soln_ControlLigan		'12450112neg5'
	59.52	-52.39		2.52	0	0	0	0.76	1.1	18.601	'./gold_soln_ControlLigan		'1OR'

Figure 1. Bestranking.lst of the docking of papain like structure of SARS CoV2 is shown with the scores of the positive and the negative control ligands.

Analysis: The scores were not expected because a good score would have been in the 100 value, but even the ligand found for alignment had a lower score than the negative controls. The alignment between 6W9C and 2IEB gave an RMSD value of 0.732, so it is a good alignment between the structure, therefore good fitting of the ligand. One reason could be because of its rotation in the active site in a different structure.

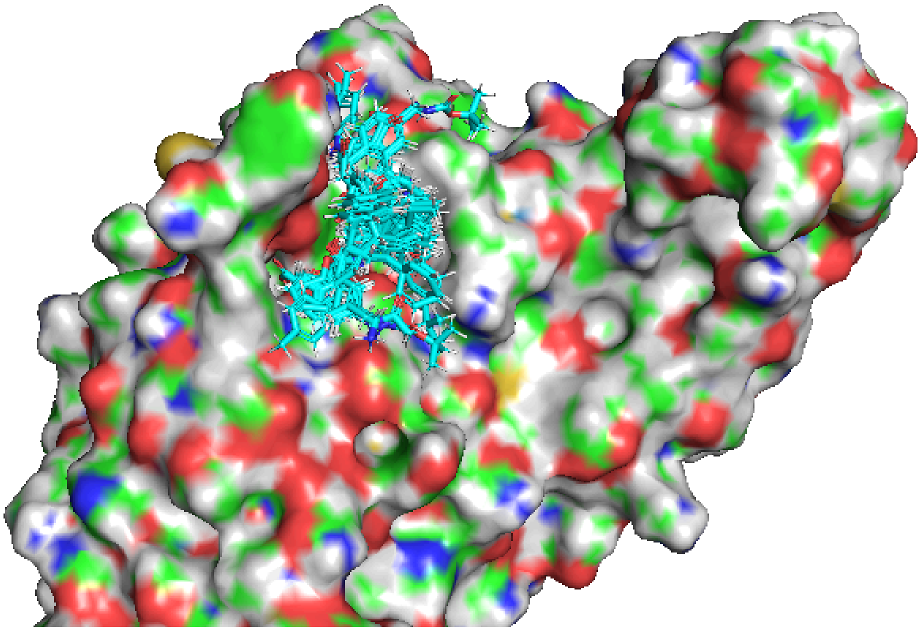
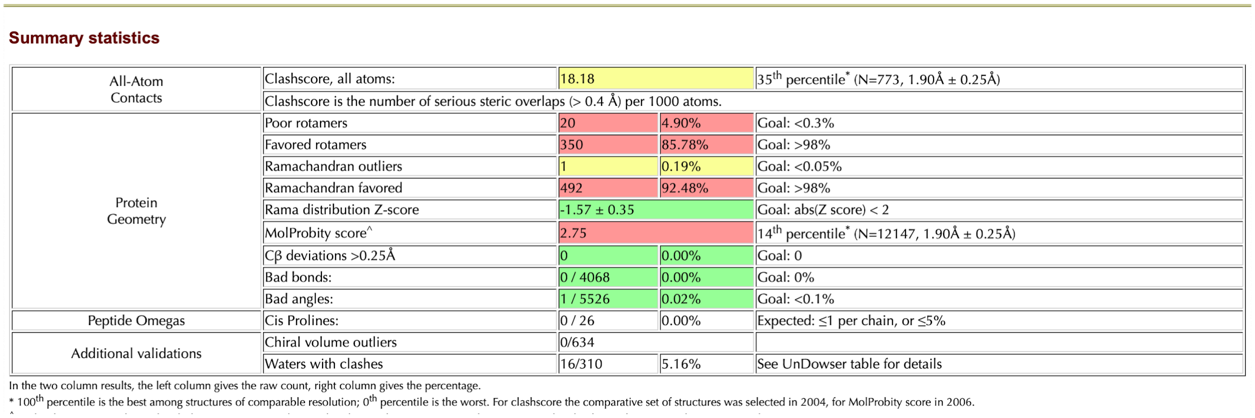


Figure 2. PyMol image of all the control ligands docked in the active site for the papain-like structure of SARS-CoV2.

Analysis: All of the ligands are bound to the active site with visible rotation. Some ligands have shown precision. More aromatic rings are shown on some structures which shows more stability. The larger ligands had more Van Der Waals interaction which could account for the better docking score.



Summary statistics

All-Atom Contacts	Clashscore, all atoms:	18.18		35th percentile* (N=773, 1.90Å ± 0.25Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	20	4.90%	Goal: <0.3%
	Favored rotamers	350	85.78%	Goal: >98%
	Ramachandran outliers	1	0.19%	Goal: <0.05%
	Ramachandran favored	492	92.48%	Goal: >98%
	Rama distribution Z-score	-1.57 ± 0.35		Goal: abs(Z score) < 2
	MolProbity score^	2.75		14th percentile* (N=12147, 1.90Å ± 0.25Å)
	Cβ deviations >0.25Å	0	0.00%	Goal: 0
	Bad bonds:	0 / 4068	0.00%	Goal: 0%
	Bad angles:	1 / 5526	0.02%	Goal: <0.1%
Peptide Omegas	Cis Prolines:	0 / 26	0.00%	Expected: ≤1 per chain, or ≤5%
Additional validations	Chiral volume outliers	0/634		
	Waters with clashes	16/310	5.16%	See UnDowser table for details

In the two column results, the left column gives the raw count, right column gives the percentage.

* 100th percentile is the best among structures of comparable resolution; 0th percentile is the worst. For clashscore the comparative set of structures was selected in 2004, for MolProbity score in 2006.

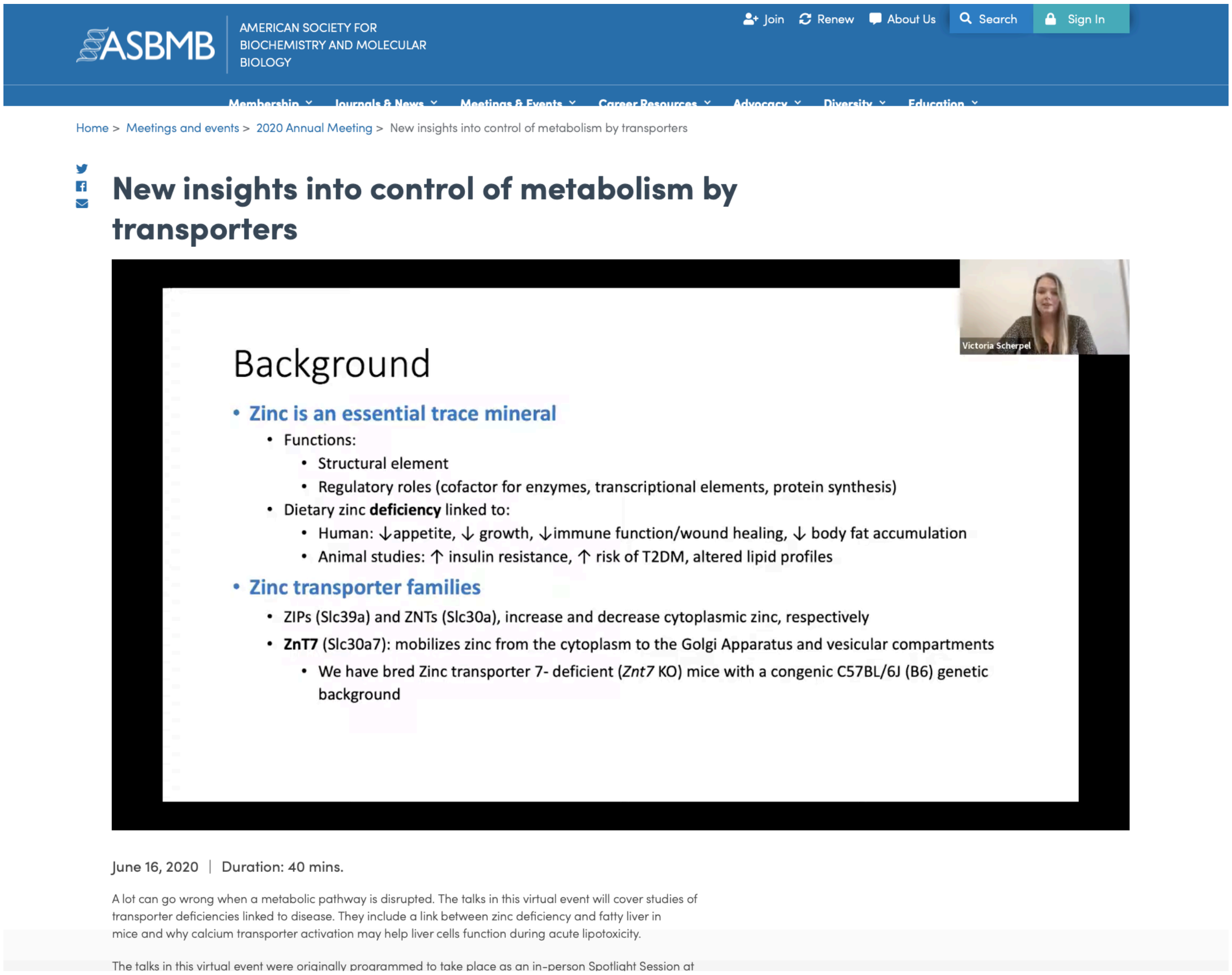
Figure 3. Screenshot of the Multi-Criterion Kinemage of Enoyl-acyl carrier protein reductase, InhA (*Mycobacterium tuberculosis)* from Molprobity is shown

Analysis: Overall, the structure is not so bad since there is half green/half red boxes. This could be due to the NADH being so old, so it could reflect later for virtual screening.

Week 3:

Great work Sandy. Nice pymol work. - Dr. B

June 19, 2020



New insights into control of metabolism by transporters

Background

- Zinc is an essential trace mineral
  - Functions:
    - Structural element
    - Regulatory roles (cofactor for enzymes, transcriptional elements, protein synthesis)
  - Dietary zinc **deficiency** linked to:
    - Human: ↓appetite, ↓ growth, ↓immune function/wound healing, ↓ body fat accumulation
    - Animal studies: ↑ insulin resistance, ↑ risk of T2DM, altered lipid profiles
- Zinc transporter families
  - ZIPs (Slc39a) and ZNTs (Slc30a), increase and decrease cytoplasmic zinc, respectively
  - **ZnT7** (Slc30a7): mobilizes zinc from the cytoplasm to the Golgi Apparatus and vesicular compartments
    - We have bred Zinc transporter 7- deficient (*Znt7* KO) mice with a congenic C57BL/6J (B6) genetic background

June 16, 2020 | Duration: 40 mins.

A lot can go wrong when a metabolic pathway is disrupted. The talks in this virtual event will cover studies of transporter deficiencies linked to disease. They include a link between zinc deficiency and fatty liver in mice and why calcium transporter activation may help liver cells function during acute lipotoxicity.

The talks in this virtual event were originally programmed to take place as an in-person Spotlight Session at

Figure 1. Screenshot of ASBMB website showing a recorded virtual meeting about control of metabolism by transporters

Analysis: This experiment of zinc deficiency that has been tested on mice has made me realize the significance of zinc in our body. The researcher has used DNA methylation as a way to express genes without changing the sequence. She has used the bisulfite conversion procedure to determine the methylated part of a sequence which is then followed by PCR.

June 18, 2020

VDS of the FRI v. 8
Analyzing DNA Sequence

```
GATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGT
AGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGA
TAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTC
GGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACT
GAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGA
CAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGG
AAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATT
TTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTT
ACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGA
TTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAAC
GACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCC
TCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAA
AGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGC
TTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCA
CACAGGAAACAGCTATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACT
CAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATGGTCGACCTGCAGGCGGCCGCACTAGT
GATTTTGATTGATTGAAGGAGAAATATCATGAGTGTGATCGCTAAACAAATGACCTACAAGGTTT
ATATGTCAGGCACGGTCAATGGACACTACTTTGAGGTCGAAGGCGATGGAAAAGGAAAGCCTTAC
GAGGGGGAGCAGACGGTAAAGCTCACTGTCACCAAGGGTGGACCTCTGCCATTTGCTTGGGATAT
TTTATCACCACTGTCTCAATACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGATT
ATGTAAAGCAGTCATTCCCTGAGGGATATACATGGGAGAGGATCATGAACTTTGAAGATGGTGCA
GTGTGTACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACAATGTCAAAATCTC
TGGTGTGAACTTTCCTCCCAATGGACCTGTTATGCAGAAGAAGACACAGGGCTGGGAACCCAACA
CTGAGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTTGGAA
GGAGGTGGTTACTATTTGTGTGAATTCAAATCTACTTACAAGGCAAAGAAGCCTGTGAGGATGCC
AGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAGTCACAACAAGGATTACACATTTGTTG
AGCAGTGTGAAATATCCATTGCACGCCACTCTTTGCTCGGTCATCACCATCACCATCACTAAAAA
TCCCGCGGCCATGGCGGCCGGGAGCATGCGACGTCGGGCCCAATTCGCC
```

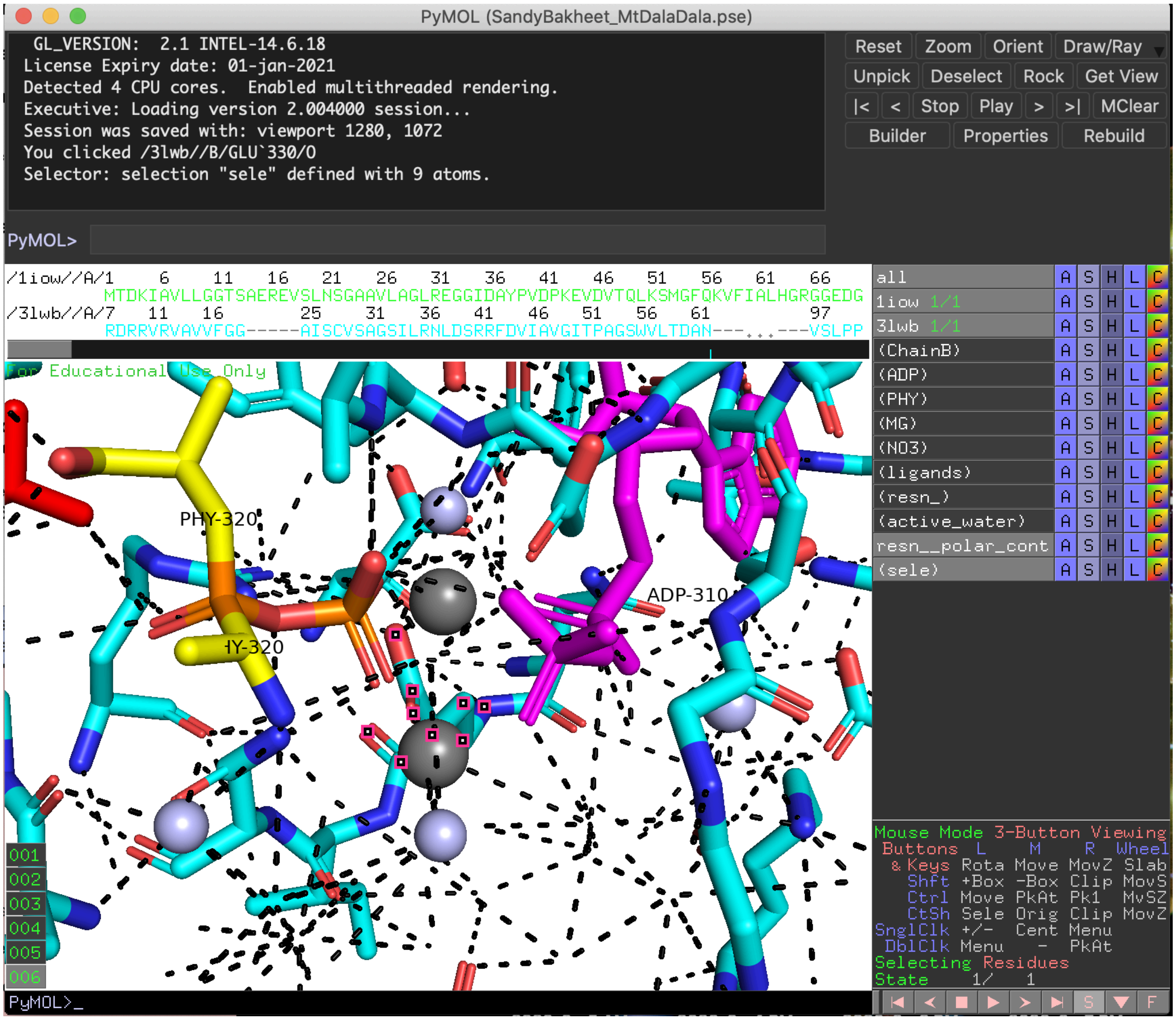
Be sure not to duplicate the TA when you put the sequences together – because the M13Rev read will have it in there as well. The best way to do this is to delete the TA as you insert your sequence in place.

Underline the M13Rev sequence once it is in the pGEM T sequence.

Once you have inserted the appropriate section into the vector, highlight:
The Start codon: ATG
The Shine-Dalgarno sequence: AGGAGA
The Stop codon: TAA
The M13 Reverse (-27) primer site **CAGGAAACAGCTATGAC**
Underlined and **bold**
Also find you're his tag and color it this color
Congratulations! - Now, you have the complete sequence for pGBR22!

Figure 2. Image of DNA Sequence protocol shows inserting M31R into pGEM T sequence.

Analysis: The protocol fully explains how the process of inserting a piece of DNA in a plasmid is done. My sequence, M31R, was inserted into the known sequence, pGEMT, after being cut. One of the things I have learned is that when choosing a reading frame, it is best to choose a long sequence that goes on a few lines not a couple of codons. His tag should appear at the end.



```
  GL_VERSION:  2.1 INTEL-14.6.18
 License Expiry date: 01-jan-2021
 Detected 4 CPU cores.  Enabled multithreaded rendering.
 Executive: Loading version 2.004000 session...
 Session was saved with: viewport 1280, 1072
 You clicked /3lwb//B/GLU`330/0
 Selector: selection "sele" defined with 9 atoms.
PyMOL>
```

```
/1iow//A/1    6    11   16   21   26   31   36   41   46   51   56   61   66
         MTDKIAVLLGGTSAEREVSLNSGAAVLAGLREGGIDAYPVDPKEVDVTQLKSMGFQKVFIALHGRGGEDG
/3lwb//A/7   11   16         25   31   36   41   46   51   56   61          97
         RDRRVRVAVVFGG-----AISCVSAGSILRNLDSRRFDVIAVGITPAGSWVLTDAN---...---VSLPP
```

```
PyMOL>_
```

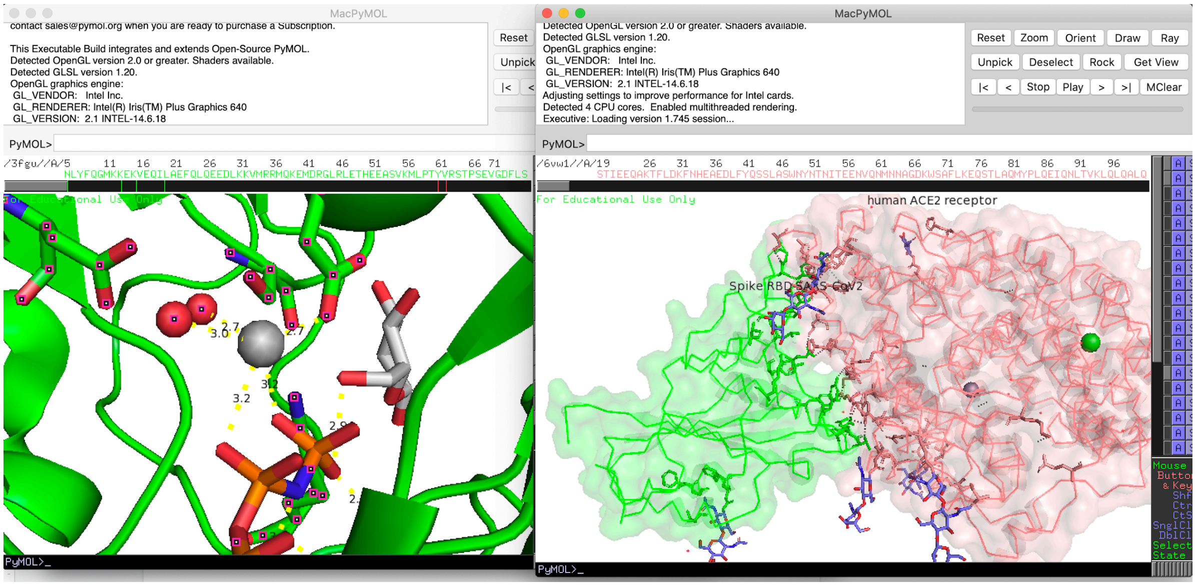
Figure 3. Image of PyMol shows PyMol scripting work of MtDalaDala.

Analysis: PyMol scripting was made using the medium step of the guide. Overall, it was not too bad since most of the steps were practiced from previous protocols such as PyMol 1, 2, and 3.

Week 2:

Week 2 - great work Sandy- Dr. B

June 10, 2020



```
contact sales@pymol.org when you are ready to purchase a Subscription.

This Executable Build integrates and extends Open-Source PyMOL.
Detected OpenGL version 2.0 or greater. Shaders available.
Detected GLSL version 1.20.
OpenGL graphics engine:
 GL_VENDOR:   Intel Inc.
 GL_RENDERER: Intel(R) Iris(TM) Plus Graphics 640
 GL_VERSION:  2.1 INTEL-14.6.18
```

```
Detected OpenGL version 2.0 or greater. Shaders available.
Detected GLSL version 1.20.
OpenGL graphics engine:
 GL_VENDOR:   Intel Inc.
 GL_RENDERER: Intel(R) Iris(TM) Plus Graphics 640
 GL_VERSION:  2.1 INTEL-14.6.18
Adjusting settings to improve performance for Intel cards.
Detected 4 CPU cores.  Enabled multithreaded rendering.
Executive: Loading version 1.745 session...
```

Figure 1. Two PyMol windows show PyMol Scripting work of 3FGU on the left and 6VW1 on the right.

Analysis: PyMol Scripting of both structures were made using the Granular step of the guide. Overall, the granular step was easy because almost all instructions had its own line command that needed to be typed.

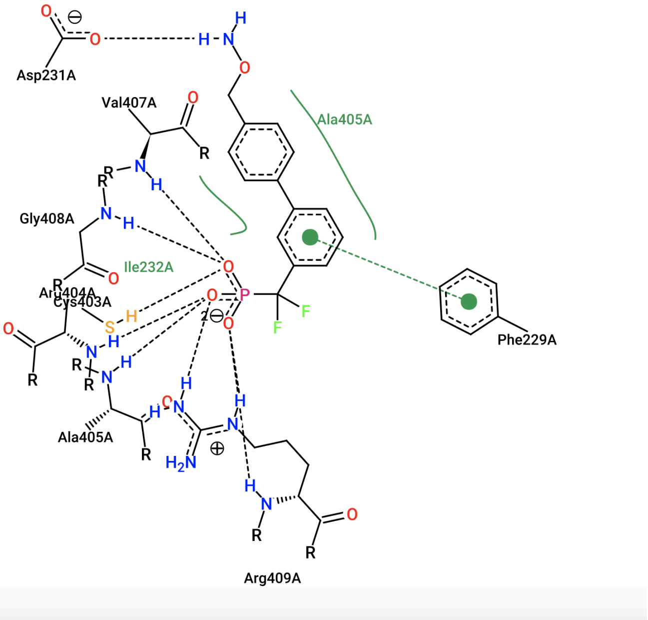
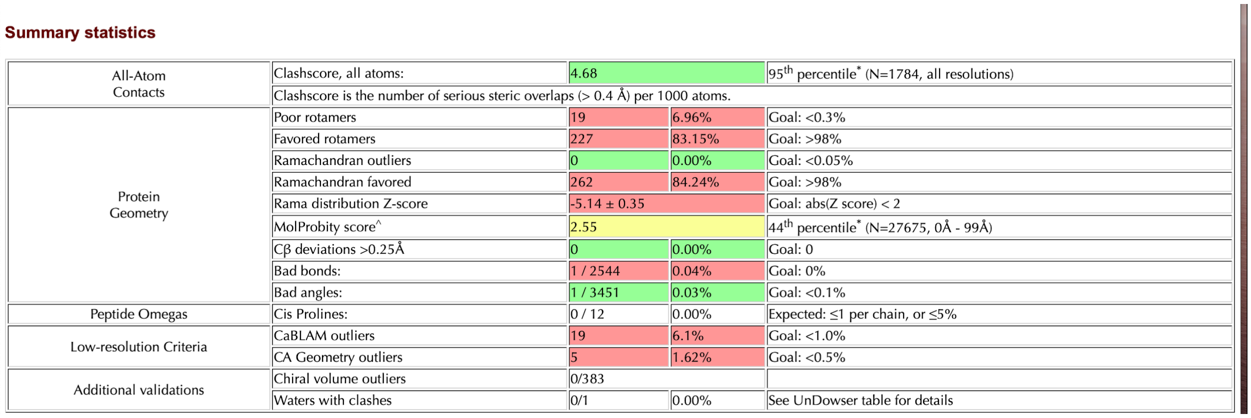


Figure 2. The Poseview Image of YopH was made using the University of Hamburg website and PDB ID.

Analysis: The image shown above is a 2D image of YopH which is helpful in showing different intermolecular interactions not present in 3D images. The poseview tool is a quick and easy way to see interactions for drug targets.

June 11, 2020



Summary statistics

All-Atom Contacts	Clashscore, all atoms:	4.68		95th percentile* (N=1784, all resolutions)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	19	6.96%	Goal: <0.3%
	Favored rotamers	227	83.15%	Goal: >98%
	Ramachandran outliers	0	0.00%	Goal: <0.05%
	Ramachandran favored	262	84.24%	Goal: >98%
	Rama distribution Z-score	-5.14 ± 0.35		Goal: abs(Z score) < 2
	MolProbity score^	2.55		44th percentile* (N=27675, 0Å - 99Å)
	Cβ deviations >0.25Å	0	0.00%	Goal: 0
	Bad bonds:	1 / 2544	0.04%	Goal: 0%
	Bad angles:	1 / 3451	0.03%	Goal: <0.1%
Peptide Omegas	Cis Prolines:	0 / 12	0.00%	Expected: ≤1 per chain, or ≤5%
Low-resolution Criteria	CaBLAM outliers	19	6.1%	Goal: <1.0%
	CA Geometry outliers	5	1.62%	Goal: <0.5%
Additional validations	Chiral volume outliers	0/383		
	Waters with clashes	0/1	0.00%	See UnDowser table for details

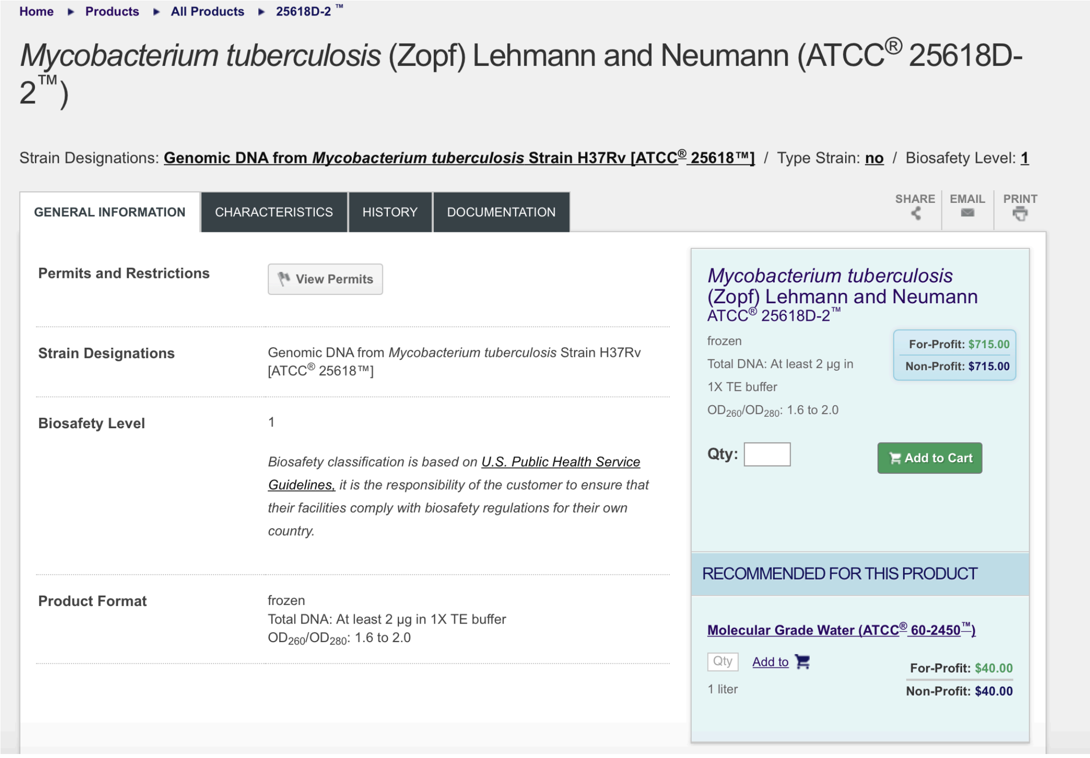
Figure 3. Screenshot of the Multi-Criterion Kinemage of papain-like protease of SARS CoV-2 from Molprobity is shown

Analysis: Overall, the structure is not so good since there is more red than green colored boxes meaning it was far from the ideal structure. This low quality will definitely show later for virtual screening.

Sandy - good work - put in Reverse Chronological order (newest stuff at top) - Thx, Dr. B

Week 1:

June 5, 2020



Mycobacterium tuberculosis (Zopf) Lehmann and Neumann (ATCC® 25618D-2™)

Strain Designations: **Genomic DNA from *Mycobacterium tuberculosis* Strain H37Rv [ATCC® 25618™]** / Type Strain: **no** / Biosafety Level: **1**

Permits and Restrictions	View Permits
Strain Designations	Genomic DNA from *Mycobacterium tuberculosis* Strain H37Rv [ATCC® 25618™]
Biosafety Level	1 *Biosafety classification is based on U.S. Public Health Service Guidelines, it is the responsibility of the customer to ensure that their facilities comply with biosafety regulations for their own country.*
Product Format	frozen Total DNA: At least 2 µg in 1X TE buffer OD_{260}/OD_{280}: 1.6 to 2.0

Mycobacterium tuberculosis (Zopf) Lehmann and Neumann
ATCC® 25618D-2™

frozen
Total DNA: At least 2 µg in
1X TE buffer
OD_{260}/OD_{280}: 1.6 to 2.0

For-Profit: $715.00
Non-Profit: $715.00

Qty:

RECOMMENDED FOR THIS PRODUCT

Molecular Grade Water (ATCC® 60-2450™)

Qty
1 liter

For-Profit: $40.00
Non-Profit: $40.00

Figure 1. ATCC Website shows the price for the target Mycobacterium Tuberculosis Strain H37Rv

Analysis: After choosing an organism, a target was chosen by reading the abstract of a few articles. One target that seemed promising was the enoyl-acyl carrier protein reductase for the organism Mycobacterium Tuberculosis. Then to prove that this drug is a good fit, all the criteria must be met. First, the ATCC website shown in figure 1 was used to find the affordability of the drug, and it seems to be at a reasonable price and with a readily available substrate, NADH. Another criteria is to see how similar the target is to humans. By going to the NCBI Blast, the percent identity of the target to humans was less than 30% which is also promising given it is not that similar to humans. So far, all the criteria seems to have been met.