



FINAL PROJECT



JANUARY 14, 2024

STUDENT NAME: ERIFEOLUWA JAMGBADI
Computational Biology and Drug Design

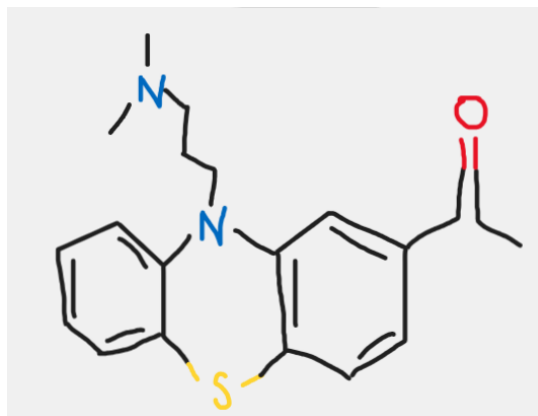
Table of Contents

Question 1.....	2
Question 2.....	4
Question 2a.....	4
Question 2b.....	5
Question 2c.....	9
Question 2d.....	12
Question 2e.....	15
Question 2f.....	18
References.....	19

Question 1

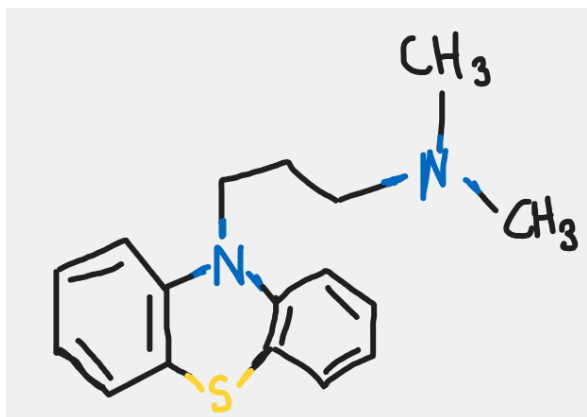
The **structural difference** between Acepromazine and Promazine are as follows:

Acepromazine C₁₉H₂₂N₂O₂S



- Polar atoms: Has polar atoms nitrogen and oxygen.
- bonds: It holds 47 bonds, 25 non-H bonds, 13 multiple bonds, 5 rotatable bonds, 1 double bond, 2 aromatic bond, 3 six-membered ring(s), 2 ten-membered ring(s), 1 ketone(s) (aromatic), 1 tertiary amine (aliphatic), 1 tertiary amine(s) (aromatic) and 1 sulfide(s)
[<https://www.molinstincts.com/structure/ACEPROMAZINE-cstr-CT1001409751.html>]
- rotatable bonds: it has 5 rotatable bonds.
- aromatic rings: it is a polycyclic aromatic compound, and it is a linear tricyclic system that consists of a two benzene rings joined by a para-thiazine ring.
- net charged atoms: More likely to become charged in acidic circumstances.
- Molecular formula is C₁₉H₂₂N₂O₂S.

Promazine



- Polar atoms: has polar atoms.
- Bonds: total of 42 bonds, 22 non-H bonds, 12 multiple bonds, 4 rotatable bonds, 12 aromatic bonds, 3 six-membered rings, 2 ten-membered rings, 1 tertiary amine(aliphatic), 1 tertiary amine (aromatic), and 1 sulfide [1]
- rotatable bonds: **4** rotatable bonds
- aromatic rings: same as acepromazine

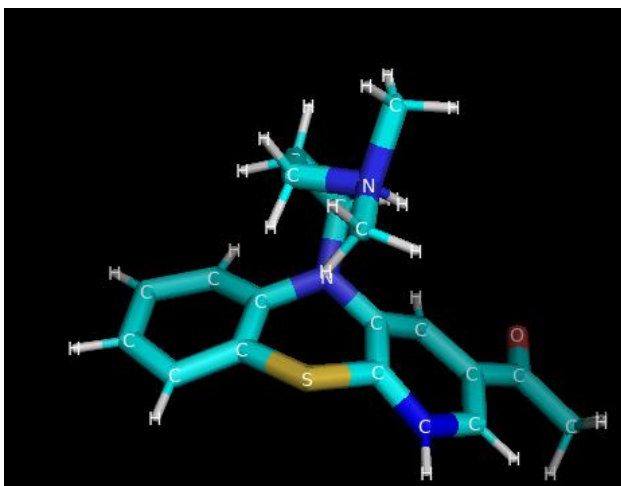
To summarise they both have different number of bonds, with acepromazine holding 47 bonds, while promazine has 42 bonds in total. The atoms attached to their nitrogen differs from each other as can be seen in the images above. Promazine is a 10-(3-dimethylaminopropyl)-2-(methylthio)phenothiazine structure, while acepromazine is a 1-methyl-2-phenothiazin-10-ylpropan-2-ylamine structure.

Question 2

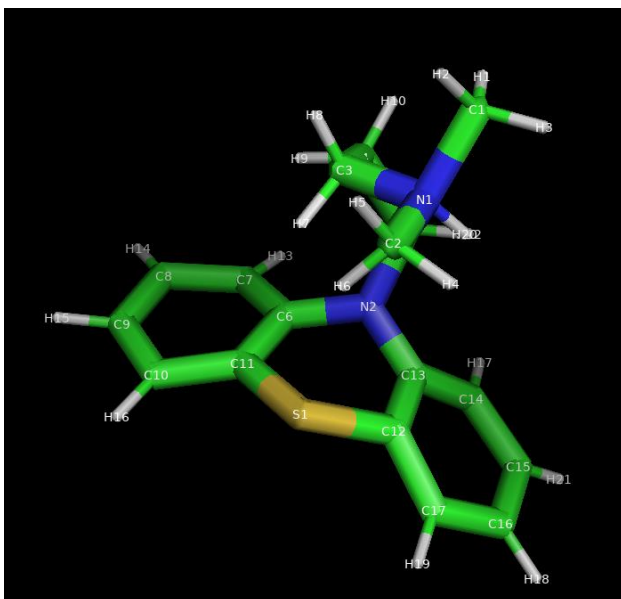
Question 2a

Yes the number of **torsdof** is 5 in the .pdbqt file for Acepromazine and for promazine, it is 4, which is the same the description mentioned for the rotatable bond description above in question 1.

For the acepromazine, the **atom types** are the same between my description and the one produced with amdock. The amdock as seen in the image contains the nitrogen, oxygen, sulfur(s), the carbon and hydrogen atoms.



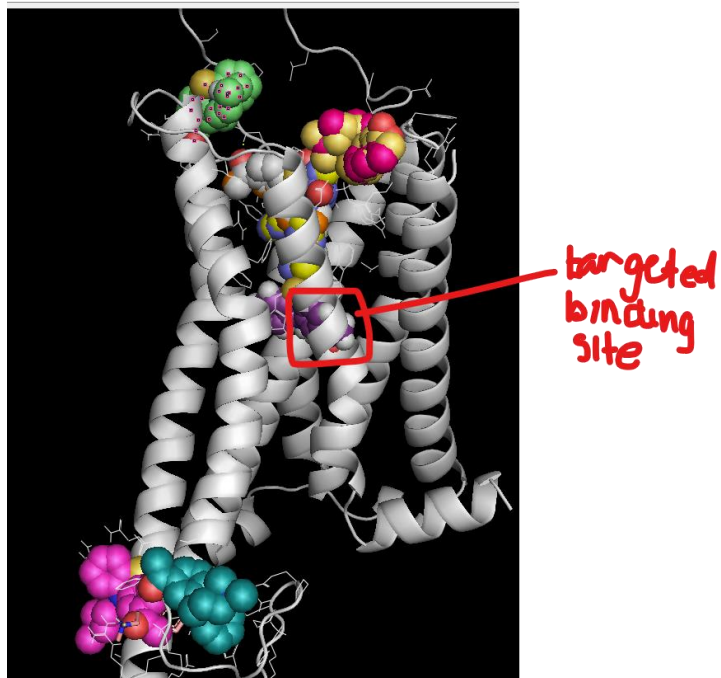
For the promazine, it has the same atoms, the 2 benzene rings, the 2 nitrogen in blue, the 2 ch₃ at the end of one of the nitrogen and the structure is the same as well.



Question 2b

I selected the grid center coordinates for both protein targets by using automatic and got 10 poses for the binding sites. I then visualised all the poses in pymol. I took into consideration extra information found about the binding sites of GPCR.

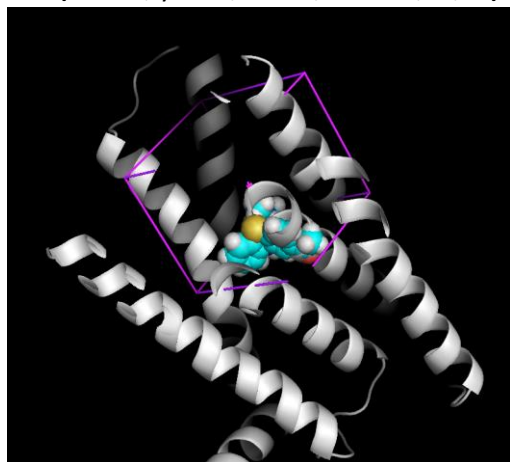
The binding sites of the the GPCR is within plasma membrane called a cavity[2]. However it is to be noted that ligands would also be able to bind elsewhere, but the cavity (7 transmembrane alpha helic is where the binding often occur) is an important place. It is important to take note or use the 2nd extracellular loop when taking into consideration binding sites for the GPCR. Using these facts as well as the automatic 10 poses, I was able to pick the best binding site for the ligands. I got the coordinates of the best box using the information I previously mentioned, then I rounded some of the number and increased the size of the number to include some ligands that might be in similar coordinates position.



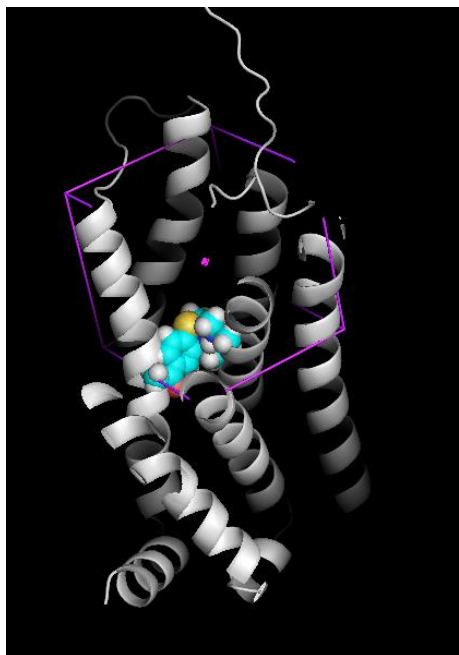
I turned the poses in a sphere shape to determine if they are located within an alpha helic cavity. When none of the 10 poses position contained the box, I used the nearest position of my targeted binding site to the one closest to the pose and increased the size of the box, but the Acepromazine ligand above for d1 was used for demonstrative purposes.

My final grid coordinates are as follows:

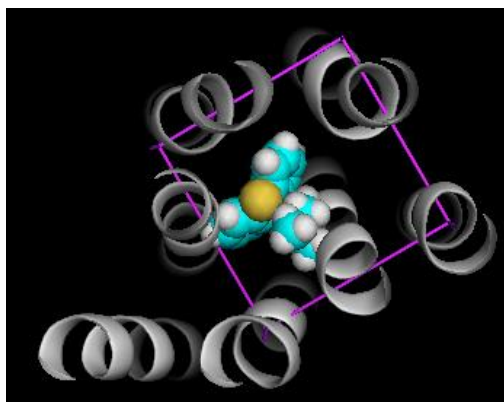
AD2(x = 1.8, y =0.1, z =6.3, size=17,17,17)



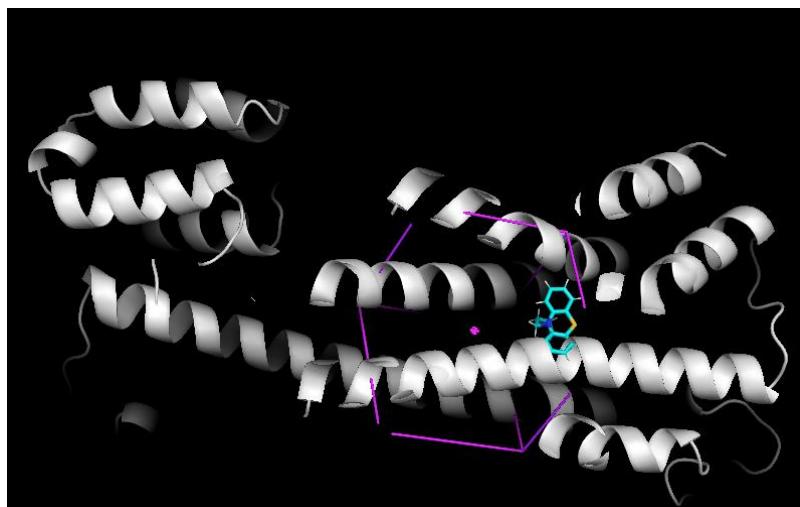
AD1(x = -0.3, y =3.6, z =8.9, size=20,20,20)



PD1(x = -0.3, y =3.6, z =8.5, size=17,17,17)



PD2 (x = 2, y =-0.4, z =-8.8, size=17,17,17)



Question 2c

Docking Calculations were preformed on both **50** and **100GA** for **2,500,000** evaluations. I have attached the results of the docking calculations done for each of the protein-ligand for both 50 and 100GA. Based on the screenshot below and the green ligend efficiency, I believe that I made good choices for binding sites, so it led to good results.

AD150

If I was narrowing down the pose just based on these results, I would narrow it down to pose 1 and 8, only because they all have green ligand efficiency, so they're good poses, but pose 1 has the best affinity and best ligand efficiency, while pose 8 has the lowest ki by a huge difference.

Import

Load Data

C:\Users\SAMSUNG\OneDrive - USI\USI\Modules\2YEAR\Computational Biology and Drug Design\FINAL PROJECT\AD 150\AD 150.amdock

Results File

Data Result

Target: D1_receptor

Pose	Affinity(kcal/mol)	Estimated Ki	Ki Units	Ligand Efficiency
1	-9.52	105.15	nM	-0.41
2	-9.02	244.51	nM	-0.39
3	-8.92	289.47	nM	-0.39
4	-8.75	385.67	nM	-0.38
5	-8.59	505.24	nM	-0.37
6	-8.34	770.45	nM	-0.36
7	-8.26	881.83	nM	-0.36
8	-8.1	1.16	uM	-0.35

PD150

Based on the results below the best pose would be 2 because it has the lowest estimate ki.

Import

Load Data

C:\Users\SAMSUNG\OneDrive - USI\USI\Modules\2YEAR\Computational Biology and Drug Design\FINAL PROJECT\PD 150\PD 150.amdock

Results File

Data Result

Target: D1_receptor

Pose	Affinity(kcal/mol)	Estimated Ki	Ki Units	Ligand Efficiency
1	-9.11	210.05	nM	-0.46
2	-7.98	1.41	uM	-0.40
3	-7.37	3.96	uM	-0.37

AD250

It was easy to determine the best pose here, it would be pose 1. Pose one has the lowest estimate ki, best affinity and ligand efficiency.

Import

Load Data

C:\Users\SAMSUNG\OneDrive - USI\USI\Modules\2YEAR\Computational Biology and Drug Design\FINAL PROJECT\AD250\AD250.amdock

Results File

Data Result

Target: D2_receptor

Pose	Affinity(kcal/mol)	Estimated Ki	Ki Units	Ligand Efficiency
1	-9.62	88.82	nM	-0.42
2	-8.85	325.77	nM	-0.38
3	-8.79	360.49	nM	-0.38
4	-8.54	549.72	nM	-0.37
5	-8.53	559.08	nM	-0.37
6	-8.28	852.56	nM	-0.36
7	-8.27	867.07	nM	-0.36

PD250

One reason why I think the results for this is bad is because the binding site may have been the wrong choice. The energy used is way too much and the affinity is very high, indicating that the poses here are not good.

Target: D2_receptor				
Pose	Affinity(kcal/mol)	Estimated Ki	Ki Units	Ligand Efficiency
1	2.76	105.46	-	0.14
2	3.35	285.47	-	0.17
3	6.97	128540.03	-	0.35
4	7.54	336395.93	-	0.38
5	11.14	146441915.42	-	0.56
6	31.74	184310193594123467882496.00	-	1.59
7	39.21	55092159862731446474216833024.00	-	1.96
8	60.21	136199459406393920554969628091068739718479...	-	3.01

AD1100

The difference in estimate ki is big, so it would be between pose 1 and pose 12. These are all goodposes as indicated by the green ligand efficiency.

Import

Load Data

C:\Users\SAMSUNG\OneDrive - USI\USI\Modules\2YEAR\Computational Biology and Drug Design\FINAL PROJECT\AD1100\AD1100.amdock

Results File

Data Result

Target: D1_receptor

Pose	Affinity(kcal/mol)	Estimated Ki	Ki Units	Ligand Efficiency
1	-9.34	142.48	nM	-0.41
2	-9.28	157.66	nM	-0.40
3	-9.16	193.06	nM	-0.40
4	-8.77	372.87	nM	-0.38
5	-8.6	496.78	nM	-0.37
6	-8.43	661.87	nM	-0.37
7	-8.43	661.87	nM	-0.37
8	-8.41	684.60	nM	-0.37
9	-8.39	708.10	nM	-0.36
10	-8.25	896.84	nM	-0.36
11	-8.23	927.63	nM	-0.36
12	-7.95	1.49	uM	-0.35

PD1100

Import

Load Data

C:\Users\SAMSUNG\OneDrive - USI\USI\Modules\2YEAR\Computational Biology and Drug Design\FINAL PROJECT\PD1100\PD1100.amdock

Results File

Data Result

Target: D1_receptor

Pose	Affinity(kcal/mol)	Estimated Ki	Ki Units	Ligand Efficiency
1	-9.1	213.63	nM	-0.45
2	-8.44	650.80	nM	-0.42
3	-8.11	1.14	uM	-0.41
4	-7.98	1.41	uM	-0.40
5	-7.52	3.07	uM	-0.38

AD2100

It has a really good affinity of -10.07cal/mol and it has a low estimated ki for pose 1 and 6, indicating good binding interaction between D2 and acepromazine.

Import

Load Data

C:\Users\SAMSUNG\OneDrive - USI\USI\Modules\2YEAR\Computational Biology and Drug Design\FINAL PROJECT\AD2100\AD2100.amdock

Results File

Data Result

Target: D2_receptor

Pose	Affinity(kcal/mol)	Estimated Ki	Ki Units	Ligand Efficiency
1	-10.07	41.56	nM	-0.44
2	-9.66	83.02	nM	-0.42
3	-8.88	309.69	nM	-0.39
4	-8.61	488.47	nM	-0.37
5	-8.57	522.58	nM	-0.37
6	-8.04	1.28	uM	-0.35

PD2100

Even with an increased time from 50 → 100 runs, the results for the promazine for d2 doesn't look good or anywhere close to it being optimal. Its like way out of range. Based on this, I either think, its because I picked a bad binding site for the promazine to D2 receptor or it promazine and D2 don't match. This result below won't be taken into consideration for the best pose as the reason can be seen clearly below in the image, no optimal pose is reached.

Import

Load Data

C:\Users\SAMSUNG\OneDrive - USI\USI\Modules\2YEAR\Computational Biology and Drug Design\FINAL PROJECT\PD2100\PD2100.amdock

Results File

Data Result

Target: D2_receptor

Pose	Affinity(kcal/mol)	Estimated Ki	Ki Units	Ligand Efficiency
1	2.85	122.76	-	0.14
2	3.02	163.56	-	0.15
3	6.54	62208.06	-	0.33
4	6.6	68837.79	-	0.33
5	12.72	2107714775.01	-	0.64
6	13.74	11788973647.60	-	0.69
7	31.06	58493362016160904642560.00	-	1.55
8	38.31	12061078139206746325515239424.00	-	1.92
9	60.4	187692171173874993066381823784011441784225...	-	3.02

Question 2d

The ligand that **binds best** to D1 is acepromazine based on the results I achieved above and the one that binds best to D2 is acepromazine based on the results I achieved above (but I wanted to say promazine because in reality promazine has affinity to D2 receptors).

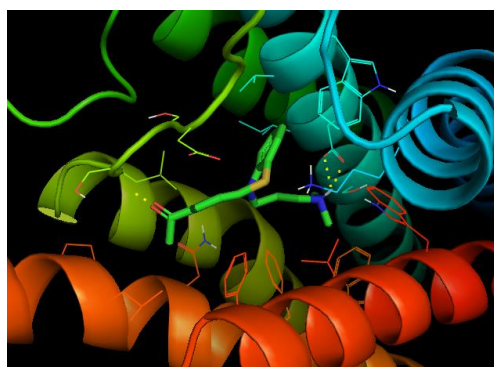
The **dlg** for each of the docking procedure is attached with the homework file and can be found in the results of each folder.

Since the goal of this assignment is to find if the ligands are good binders to the proteins, I placed more emphasis on a lower estimated k_i because they are a good indicator of binding affinity. Although the emphasis was placed on a lower estimate k_i , I also looked at other things such as the affinity, ligand efficiency, but I obtained good results for these, so I had to weigh them on a scale and evaluate, which is another reason for the emphasis on the lower estimate k_i .

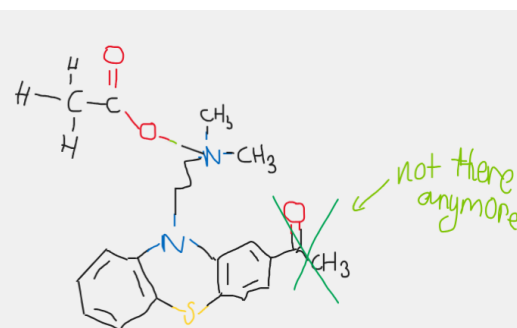
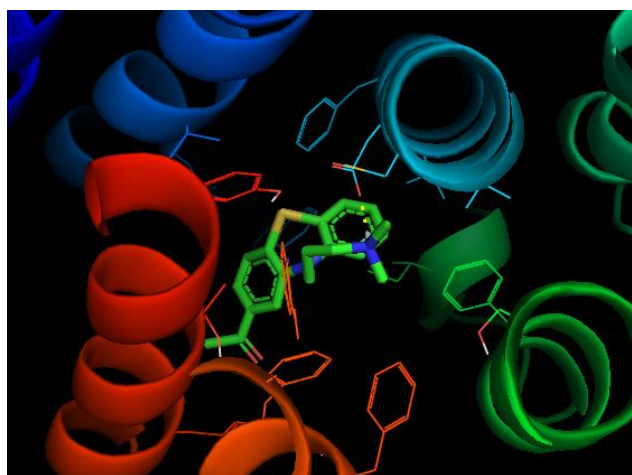
The **best binding pose** are as follows:

D1 50 runs pose 1 for acepromazine.

Partial bonds are formed during the interaction. Between the od1 and od2 on the side of the nitrogen.



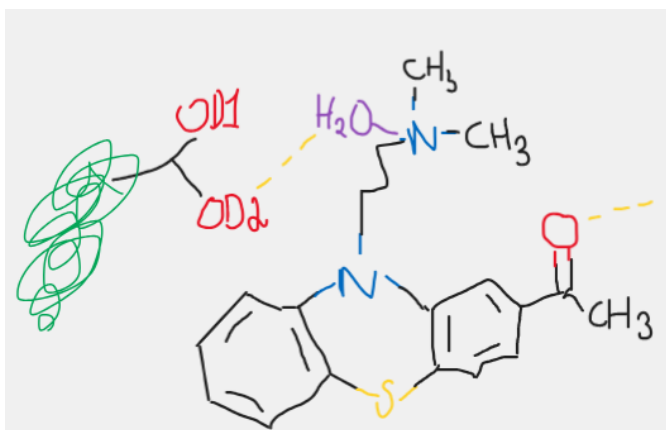
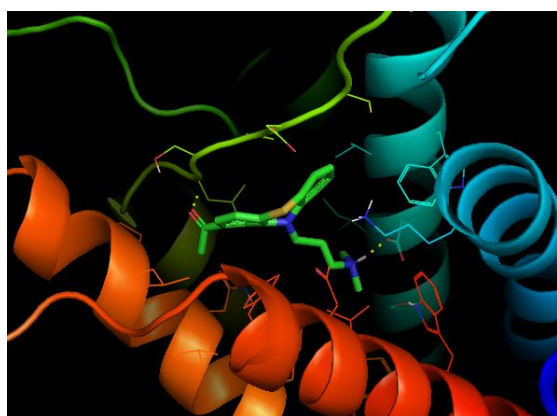
D2 50 runs is pose 1 for acepromazine.



As can be seen above shows the interaction that takes place between acepromazine and D2 for pose 1. The image on the right shows the structure of the interaction. The new interactions produce a new connection at nitrogen and removes it from the oxygen and ch3 from the benzene and gives it to the nitrogen instead. I assume because the bond is broken on the benzene side because nitrogen is less electromagnetic than oxygen, so that means that oxygen will attract more electrons to itself, which is what happened in this case, where we see the electrons go to the oxygen. The combination of nitrogen and oxygen results in partial charges for the atoms.

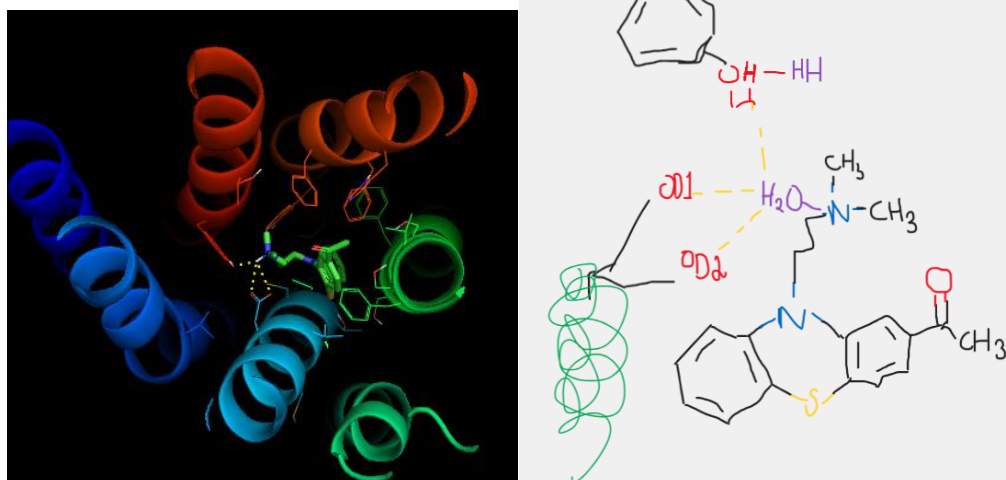
D1 100 runs is pose 1 for acepromazine.

In the image below



In the image above a weak partial charge is formed with the oxygen when an interaction takes place with pose 1 between acepromazine and D1. On the other side of the acepromazine a bond between OD2 and the nitrogen takes place. The water allows interaction with charged or polar molecules, allowing it to dissolve or aid easier chemical reaction. These elements are produced during the interaction that takes place with this ligand and D1.

D2 100 runs is pose 1 for acepromazine.



As can be seen from the image above on the right, it shows the new elements that are produced when the interaction between the acepromazine and D2 take place. The green just shows one of the curly things in the protein. New hydrogen bonds, water (H₂O), hydroxyl group (OH), oxygen (D1 and D2) and a benzene ring are produced. There are partial bonds which are weaker between the H₂O and the new elements produced. This means that the interaction between those elements connected to the H₂O is weaker, only the H₂O has a strong connection to the rest of acepromazine. This means that the interaction is weak between the protein and the ligand.

Then to break down the above runs and compare with each other to get the best pose for D1 and D2, regardless of the run is shown below as follows:

D1 is pose 1 for acepromazine for 50 runs and D2 is pose 1 for acepromazine for 100 runs. This is because it has more interactions with the protein, an affinity of -9.52, an estimated K_i of 105.15 and a ligand efficiency of -0.41. All this indicates it has the better affinity for binding with D1.

Question 2e

The histogram for the found poses for each ligand-protein/receptor poses are as follows:

AD150

Although pose 8 looked favourable in the docking calculation with its low estimate k_i , the histogram seems to indicate otherwise about it being a good choice to pick for the best pose. Pose 2 looks good as well, but because pose 1 has better values for the affinity, estimated k_i and ligand efficiency, I will pick pose 1 for this binding.

```
Number of distinct conformational clusters found = 8, out of 50 runs,
Using an rmsd-tolerance of 2.0 Å

CLUSTERING HISTOGRAM
```

Clus	Lowest	Run	Mean	Num	Histogram
-ter	Binding		Binding	in	
Rank	Energy		Energy	Clus	5 10 15 20 25 30 35
1	-9.52	18	-9.20	10	#####
2	-9.02	42	-8.75	14	#####
3	-8.92	11	-8.86	7	#####
4	-8.75	26	-8.59	2	##
5	-8.59	12	-8.50	3	###
6	-8.34	21	-8.03	11	#####
7	-8.26	7	-8.26	1	#
8	-8.10	34	-7.73	2	##

```
Number of multi-member conformational clusters found = 7, out of 50 runs.
```

PD150

Looking at the clustering histogram, there's clearly an over representation of pose 1 with 48/50 runs. Using the previous result of affinity and estimated k_i , I wanted to pick pose 1, but the results below show an indicator or something wrong.

```
Number of distinct conformational clusters found = 3, out of 50 runs,
Using an rmsd-tolerance of 2.0 Å

CLUSTERING HISTOGRAM
```

Clus	Lowest	Run	Mean	Num	Histogram
-ter	Binding		Binding	in	
Rank	Energy		Energy	Clus	5 10 15 20 25 30 35
1	-9.11	16	-8.76	48	#####
2	-7.98	40	-7.98	1	#
3	-7.37	44	-7.37	1	#

```
Number of multi-member conformational clusters found = 1, out of 50 runs.
```

AD250

Number of distinct conformational clusters found = 7, out of 50 runs,
Using an rmsd-tolerance of 2.0 Å

CLUSTERING HISTOGRAM

Clus	Lowest	Run	Mean	Num	Histogram
-ter	Binding		Binding	in	
Rank	Energy		Energy	Clus	5 10 15 20 25 30 35
1	-9.62	15	-9.24	8	#####
2	-8.85	30	-8.66	4	####
3	-8.79	20	-8.38	23	#####
4	-8.54	11	-7.93	8	#####
5	-8.53	8	-8.04	3	###
6	-8.28	48	-8.08	3	###
7	-8.27	34	-8.27	1	#

Number of multi-member conformational clusters found = 6, out of 50 runs.

PD250

Number of distinct conformational clusters found = 8, out of 50 runs,
Using an rmsd-tolerance of 2.0 Å

CLUSTERING HISTOGRAM

Clus	Lowest	Run	Mean	Num	Histogram
-ter	Binding		Binding	in	
Rank	Energy		Energy	Clus	5 10 15 20 25 30 35
1	+2.76	43	+5.45	25	#####
2	+3.35	27	+9.17	6	#####
3	+6.97	6	+8.96	7	#####
4	+7.54	36	+7.54	1	#
5	+11.14	49	+11.14	1	#
6	+31.74	13	+36.95	8	#####
7	+39.21	15	+39.21	1	#
8	+60.21	47	+60.21	1	#

Number of multi-member conformational clusters found = 4, out of 50 runs.

AD1100

Before when I saw the results from the docking calculations, for me the best pose was between 1 and 6, but after seeing the clustering histogram, I can conclude that the best pose for adpromazine-D1 is pose 1. Although the value for the histogram for pose 2 is good, using the affinity and estimated k_i , pose 1 is the better option.

Number of distinct conformational clusters found = 12, out of 100 runs,
Using an rmsd-tolerance of 2.0 Å

CLUSTERING HISTOGRAM

Clus	Lowest	Run	Mean	Num	Histogram
-ter	Binding		Binding	in	
Rank	Energy		Energy	Clus	5 10 15 20 25 30 35
1	-9.34	86	-8.99	29	#####
2	-9.28	41	-8.81	32	#####
3	-9.16	46	-9.16	1	#
4	-8.77	2	-8.59	10	#####
5	-8.60	43	-8.51	3	###
6	-8.43	23	-8.43	1	#
7	-8.43	72	-7.97	15	#####
8	-8.41	39	-8.41	1	#
9	-8.39	9	-8.39	1	#
10	-8.25	63	-8.25	1	#
11	-8.23	16	-7.92	2	###
12	-7.95	45	-7.79	4	####

Number of multi-member conformational clusters found = 7, out of 100 runs.

PD1100

Based on the docking calculations above for this, I wanted to use pose 1 for comparison to get the best pose for D1 for 100 runs, but it seems based on the histogram below, there might have been an over representation of pose 1.

Number of distinct conformational clusters found = 5, out of 100 runs, Using an rmsd-tolerance of 2.0 Å									
CLUSTERING HISTOGRAM									
Clus	Lowest	Run	Mean	Num	Histogram				
-ter	Binding		Binding	in					
Rank	Energy		Energy	Clus	5	10	15	20	25 30 35
1	-9.10	35	-8.81	85	#####				
2	-8.44	25	-8.44	1	#				
3	-8.11	62	-7.92	8	#####				
4	-7.98	77	-7.97	2	##				
5	-7.62	81	-7.48	4	###				

Number of multi-member conformational clusters found = 4, out of 100 runs.

AD2100

Pose 2 and 5 look stable here, with pose 3 overshooting the histogram.

Number of distinct conformational clusters found = 6, out of 100 runs, Using an rmsd-tolerance of 2.0 Å									
CLUSTERING HISTOGRAM									
Clus	Lowest	Run	Mean	Num	Histogram				
-ter	Binding		Binding	in					
Rank	Energy		Energy	Clus	5	10	15	20	25 30 35
1	-10.07	79	-8.87	9	#####				
2	-9.66	12	-9.42	14	#####				
3	-8.88	79	-8.42	53	#####				
4	-8.61	84	-8.56	9	#####				
5	-8.57	8	-8.11	14	#####				
6	-8.04	53	-8.04	1	#				

Number of multi-member conformational clusters found = 5, out of 100 runs.

PD2100

If it wasn't for the not optimal results obtained in the docking calculations, the histogram looks stable, but the lowest binding energy is an indicator that is not good to consider this for best bind affinity for D2.

Number of distinct conformational clusters found = 9, out of 100 runs, Using an rmsd-tolerance of 2.0 Å									
CLUSTERING HISTOGRAM									
Clus	Lowest	Run	Mean	Num	Histogram				
-ter	Binding		Binding	in					
Rank	Energy		Energy	Clus	5	10	15	20	25 30 35
1	+2.88	69	+7.15	22	#####				
2	+3.02	13	+5.86	40	#####				
3	+6.54	74	+7.90	7	#####				
4	+6.60	97	+9.85	15	#####				
5	+12.72	98	+12.72	1	#				
6	+13.74	68	+13.88	3	###				
7	+31.06	32	+40.83	9	#####				
8	+38.31	12	+39.28	2	##				
9	+60.40	51	+60.40	1	#				

Number of multi-member conformational clusters found = 7, out of 100 runs.

My docking calculations does converge one of the reasons is because I used ran all the experiments with 2,500,000 evaluations thus allowing for convergency. I also checked the log out of the docking calculation and I looked at the RMSD TABLE and in this table I looked at the reference rmsd. Through this I was able to use it as a measure for the convergence of the algorithm. You can see that which each run, the value is gradually decreasing, meaning the docking calculation is becoming more stable with the runs.

Question 2f

Yes, promazine exhibited selectivity towards D1 than D2 **[based on the results found in my docking calculation]**. As can be seen from the docking results and clustering histogram when comparing promazine between D1 and D2, the results for D2 for promazine was not good and could not be used, but while for the D1, it produced good poses which indicated potential binding sites in D1. Also, for the clustering histogram, for D1, overshooting happens which could indicate a favourable pose in D1. In theory and reality, promazine is meant to exhibit selectivity to D2 [3] rather than D1 leading me to believe I went wrong somewhere with picking the boxing size for promazine for D2.

Acepromazine produced good results for both receptors obtaining good results. So, the ligand acepromazine, it did not exhibit selectivity.

References

1. Promazine structure - c17h20n2s - over 100 million chemical compounds: CCDDS Mol. Available at: <https://www.molinstincts.com/structure/promazine-cstr-CT1000478155> (Accessed: 15 January 2024).
2. G protein-coupled receptor (2023a) Wikipedia. Available at: https://en.wikipedia.org/wiki/G_protein-coupled_receptor (Accessed: 15 January 2024).
3. NCATS INXIGHT drugs — PROMAZINE. Available at: <https://drugs.ncats.io/drug/O9M39HTM5W> (Accessed: 15 January 2024).