Experiment-5: Structure Based Drug Designing : Molecular Docking using Autodock

- **AIM:** To retrieve the Protein and Ligand structure from PDB Database and perform molecular docking studies and find Interactions between them.
- **TOOL & FILES:**

  PDB: http://www.rcsb.org
  Autodock v1.5.6
  Autodock.exe
  Autogrid.exe
- **Molecular Docking:**

Molecular docking is a method which predicts the orientation of bonding one molecule with another molecule to form a stable complex. It is a frequently used method in structure based drug designing to predict the binding conformation of ligand with the target binding site. Docking Is a term used for computational schemes that attempt to find the "best" matching between two molecules: a receptor and a ligand. In the docking, the program obtains the image of the binding site from the molecular surface of the macromolecule and ligand molecules are mapped on to the binding and then docking energies and scores have been evaluated.

- **METHODOLOGY:**

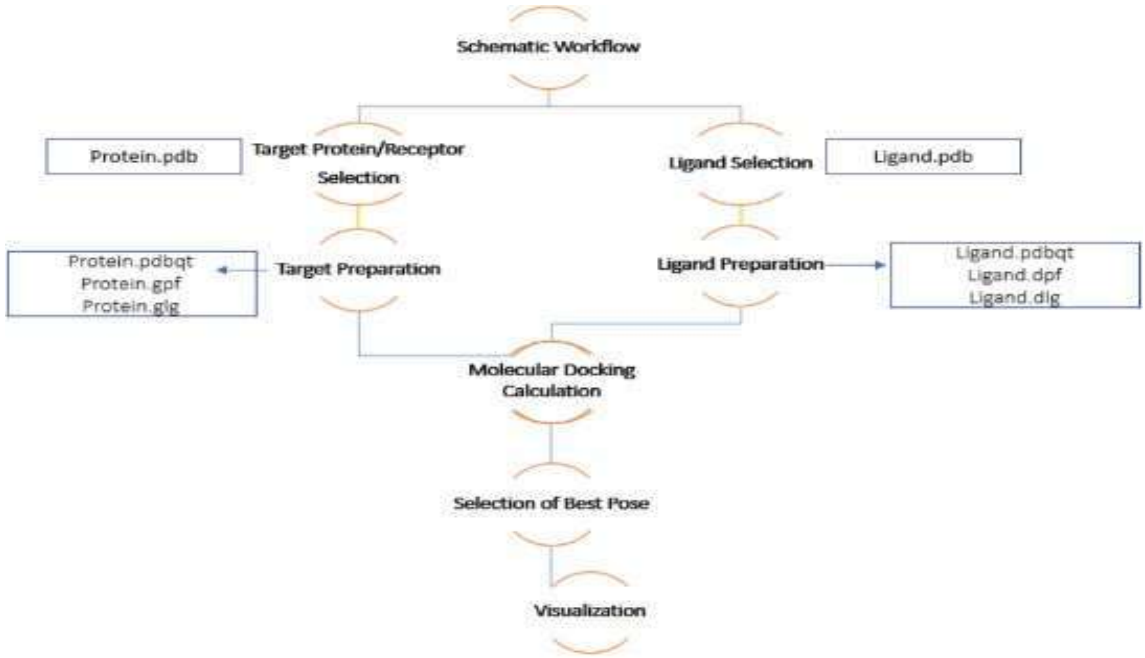


Figure 1. Schematic flowchart for performing Molecular docking studies

Figure 24: Schematic Workflow of Molecular Docking

1. Retrieve of Protein:

i. Go to PDB database and search for **1CRK** protein structure
ii. Download the structure in .pdb file format
iii. Open PyMOL and upload the protein file
iv. Extract the Chain A and save this as .pdb file

2. Retrieve or Draw the Ligand:

i. Download the ligand which is mention in literature
ii. Or, Draw the ligand using Marvin Sketch and download in .sdf format
iii. Save it as .pdb format

3. Setting Path in Autodock:

i. Open Autodock1.5.6
ii. Go to file and then Preferences
iii. Set the directory path from the folder address where the all .pdb files along with the Autodock.exe and Autogrid.exe files are present

4. Protein Preparation:

a. Protein preparation is a process keeping protein structure ready for docking.
b. AutoDock is based on the United Atom force-field of AMBER, which uses only polar hydrogens.
c. This helps to reduce the number of atoms that must be modeled explicitly during the docking, thus speeding up the calculation
d. Polar hydrogens are hydrogen atoms that are bonded to electronegative atoms like oxygen and nitrogen. That makes protein more interactive to ligand.
e. Read the 1crkA.pdb file in AutoDockTools-1.5.6

STEPS:

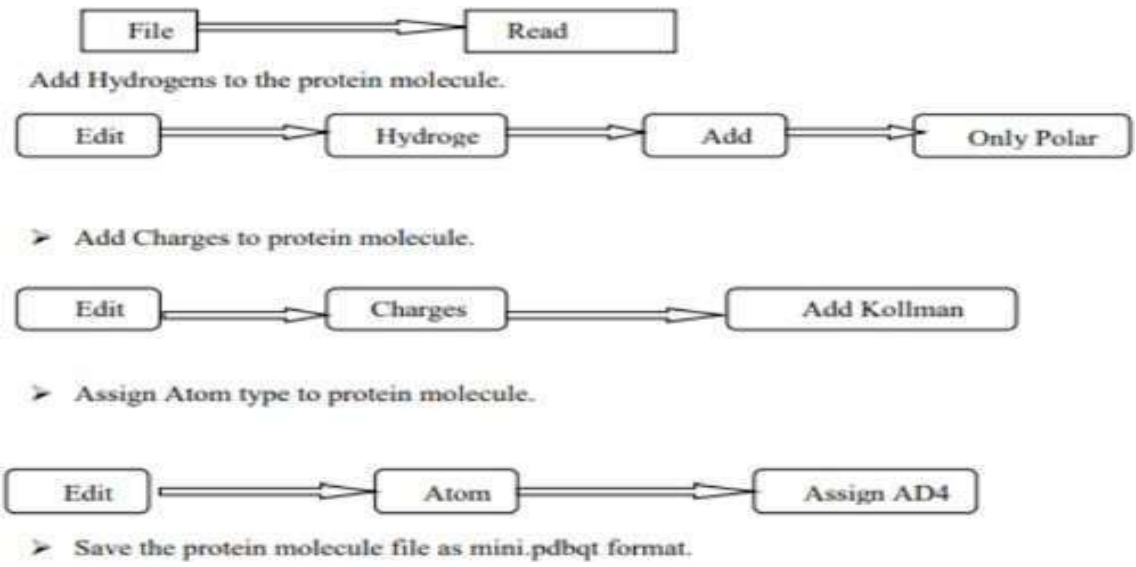


Figure 25: Steps in Protein Preparation

5. Ligand Prearation:

Ligand preparation is a process of making ligand structure ready for docking. This is achieved by detecting the root and adding charge. It is done to make ligand to interact at specific pockets of protein structure.

STEPS: ➢ Read the ligand structure in AutoDockTools-1.5.6

➢ Detect the root for ligand structure.

➢ Find out the charges in ligand.

➢ Save the ligand file as ligand.pdbqt format

Figure 26: Steps in Ligand Preparation

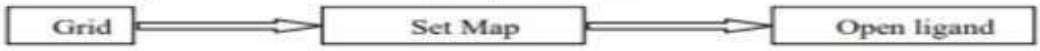
6. Grid BOX Generation:

- Grid box preparation is setting the predicted interaction site of ligand to the target pocket.
- Box is adjusted at the interactive site on X, Y, Z co-ordinations.
- Here use both 1crkA.pdbqt and ligand.pdbqt
- Read the ligand structure from ligand.pdbqt to Grid Box Set the Grid box at the predicted binding pocket, by adjusting X, Y, Z dimension parameters.

STEPS:

❖ Read protein molecule from mini.pdbqt to grid.

❖ Read ligand molecule from ligand.pdbqt to grid.



❖ Adjust the Grid Box at possible interaction site.

❖ Set the Grid Box in right position or around the binding site of the protein.

File ➔ close saving current.

Grid ➔ Output ➔ save GPF

Figure 27: Steps in Grid Box Generation

7. Docking:

❖ **Docking Parameters:-**

➢ Read protein molecule from mini.pdbqt file to docking

➢ Read ligand structure from ligand.pdbqt file to docking.

➢ Set the Search parameters by using Genetic algorithm.

➢ Set the Docking parameters by using docking options.

Docking ➔ Output ➔ Lamarckian GA (4.2)

Figure 28: Steps in Docking

❖ **RUN:-**

➢ Run auto grid using autogrid.exe programming file.

➢ Run auto dock using autodock.exe programming file.

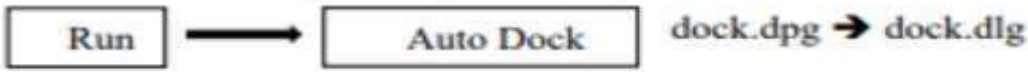


Figure 29: Steps in Run the Autogrid and Autodock

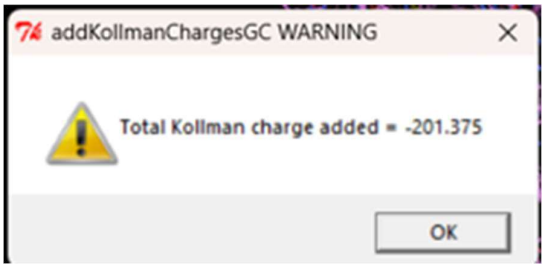
Molecular Docking Result:-

- Protein preparation
- Non polar Hydrogen were added to the protein structure
- Add Charge (Kollman charges)
- Add Atom type ()
- The prepared protein was saved in 1crkA.pdbqt format.
- Ligand preparation
- Root of the ligand molecule was detected.
- Add charges ()
- Save the prepared ligand in ligand1.pdbqt

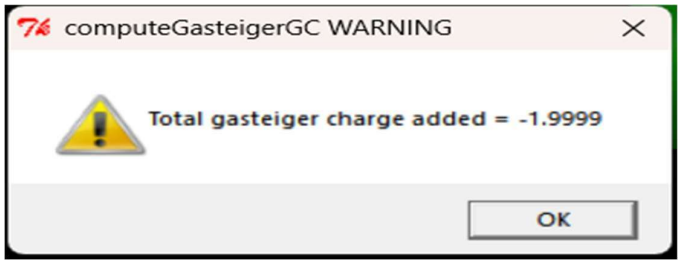
8. Analysis:

i. Go to Analyze and then macromolecule then click conformation and play
ii. I Show info check box mark show H bond , GA run which has lowest binding energy and then Write complex to save it as doc_complex.pdbqt
iii. Go to file, read molecule to open the dock_complex.pdbqt file
iv. Go to file again, output and save the file as .pdb format
v. Visualize in PyMOL

- **RESULT:**



(a)Kollman charges



(b)Gasteiger charge

Grid box generation:-

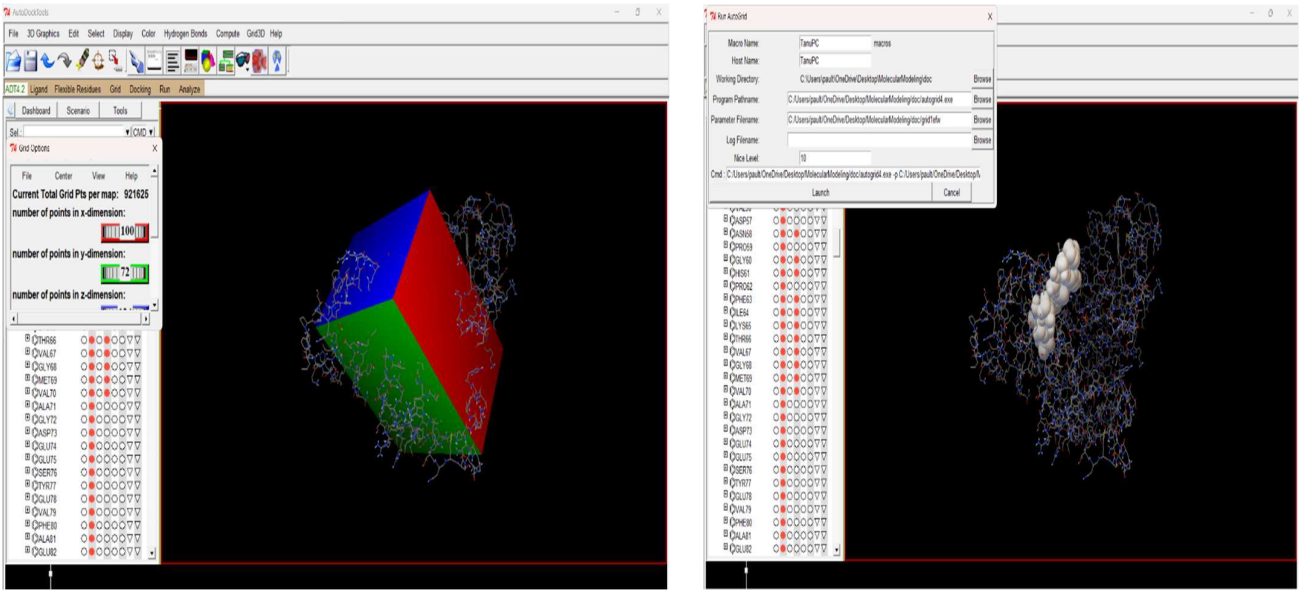


Figure 30: Grid Box

Autodock output file:-

```
        FINAL GENETIC ALGORITHM DOCKED STATE
        ____________________________________


Detailed state:  trans 38.678 35.662 11.220 quatxyzw -0.239051 0.000732 -0.527026 0.815535 center 37.591 34.440 8.368 ntor 0
State:    38.678  35.662  11.220  -0.413  0.001 -0.911 70.719

DOCKED: MODEL        1
DOCKED: USER    Run = 1
DOCKED: USER    DPF = C:/Users/pault/OneDrive/Desktop/MolecularModeling/doc/dock1.dpf
DOCKED: USER
DOCKED: USER    Estimated Free Energy of Binding    =    -8.66 kcal/mol  [=(1)+(2)+(3)-(4)]
DOCKED: USER    Estimated Inhibition Constant, Ki   =  445.37 nM (nanomolar)  [Temperature = 298.15 K]
DOCKED: USER
DOCKED: USER    (1) Final Intermolecular Energy     =    -8.66 kcal/mol
DOCKED: USER        vdW + Hbond + desolv Energy     =    -1.72 kcal/mol
DOCKED: USER        Electrostatic Energy            =    -6.94 kcal/mol
DOCKED: USER    (2) Final Total Internal Energy     =    +0.00 kcal/mol
DOCKED: USER    (3) Torsional Free Energy           =    +0.00 kcal/mol
DOCKED: USER    (4) Unbound System's Energy  [=(2)] =    +0.00 kcal/mol
DOCKED: USER
DOCKED: USER
DOCKED: USER    NEWDPF move Aligand.pdbqt
DOCKED: USER    NEWDPF about 37.591000 34.440000 8.368000
DOCKED: USER    NEWDPF tran0 38.677557 35.661690 11.219689
DOCKED: USER    NEWDPF quaternion0 -0.239051 0.000732 -0.527026 0.815535
DOCKED: USER    NEWDPF axisangle0 -0.413078 0.001264 -0.910695 70.719369
DOCKED: USER    NEWDPF quat0 -0.413078 0.001264 -0.910695 70.719369
DOCKED: USER  keepresnum = 1
DOCKED: USER
DOCKED: REMARK  0 active torsions:
DOCKED: REMARK  status: ('A' for Active; 'I' for Inactive)
DOCKED: USER                                 x       y       z     vdW  Elec       q    Type
DOCKED: USER                              _______ _______ _______ _____ _____    ______ ____
DOCKED: ROOT
DOCKED: ATOM      1  P    PO4 A 400      38.676  35.662  11.220 +0.24 +1.33    +0.436 P
DOCKED: ATOM      2  O1   PO4 A 400      40.083  35.547  10.684 -0.44 -2.21    -0.609 OA
DOCKED: ATOM      3  O2   PO4 A 400      38.256  37.109  11.229 -0.31 -1.69    -0.609 OA
DOCKED: ATOM      4  O3   PO4 A 400      38.631  35.121  12.618 -0.93 -2.13    -0.609 OA
DOCKED: ATOM      5  O4   PO4 A 400      37.740  34.868  10.347 -0.29 -2.25    -0.609 OA
DOCKED: ENDROOT
DOCKED: TORSDOF 0
DOCKED: TER
DOCKED: ENDMDL
_______________________________________________________________________________
```

Figure 31: .dlg file - Autodock output fie

Cluster Analysis:-

```
                CLUSTER ANALYSIS OF CONFORMATIONS
                _________________________________

Number of conformations = 10

RMSD cluster analysis will be performed using the ligand atoms only (5 / 5 total atoms).

Outputting structurally similar clusters, ranked in order of increasing energy.


_______________________________________________________________________________

Number of distinct conformational clusters found = 1,  out of 10 runs,
Using an rmsd-tolerance of 2.0 A
```

Figure 32: .dlg file - Cluster analysis

Clustering Histogram:-

```
        CLUSTERING HISTOGRAM
        ____________________

_____________________________________________________________________________
     |             |     |             |     |
Clus | Lowest      | Run | Mean        | Num | Histogram
-ter | Binding     |     | Binding     | in  |
Rank | Energy      |     | Energy      | Clus|    5    10   15   20   25   30   35
_____|_____________|_____|_____________|_____|____:____|____:____|____:____|____:___
   1 |      -8.68  |   8 |      -8.67  |  10 |##########
_____|_____________|_____|_____________|_____|_____________________________________


Number of multi-member conformational clusters found = 1, out of 10 runs.
```

Figure 33: .dlg file - Cluster Histogram

RMSD table:-

RMSD TABLE

Rank	Sub-Rank	Run	Binding Energy	Cluster RMSD	Reference RMSD	Grep Pattern
1	1	8	-8.68	0.00	2.68	RANKING
1	2	10	-8.68	0.03	2.69	RANKING
1	3	9	-8.68	0.01	2.68	RANKING
1	4	4	-8.67	0.03	2.68	RANKING
1	5	7	-8.67	0.07	2.70	RANKING
1	6	3	-8.67	0.06	2.69	RANKING
1	7	6	-8.67	0.09	2.70	RANKING
1	8	5	-8.67	0.09	2.71	RANKING
1	9	1	-8.66	0.07	2.70	RANKING
1	10	2	-8.66	0.12	2.72	RANKING

Figure 34: .dlg file - Root Mean Square Deviation Table

Entropy and Statistical Mechanical Analysis:

```
        INFORMATION ENTROPY ANALYSIS FOR THIS CLUSTERING
        ________________________________________________


Information entropy for this clustering = 0.00  (rmstol = 2.00 Angstrom)

_____________________________________________________________________


        STATISTICAL MECHANICAL ANALYSIS
        _______________________________


Partition function, Q =    10.15           at Temperature, T = 298.15 K
Free energy,        A ~ -1372.91 kcal/mol   at Temperature, T = 298.15 K
Internal energy,    U =    -8.67 kcal/mol   at Temperature, T = 298.15 K
Entropy,            S =     4.58 kcal/mol/K at Temperature, T = 298.15 K

_____________________________________________________________________
```

Figure 35: .dlg file - Entropy and statistical mechanical analysis

Lowest Energy Docked Conformation:-

```
        LOWEST ENERGY DOCKED CONFORMATION from EACH CLUSTER
        __________________________________________________


Keeping original residue number (specified in the input PDBQ file) for outputting.

MODEL        8
USER    Run = 8
USER    Cluster Rank = 1
USER    Number of conformations in this cluster = 10
USER
USER    RMSD from reference structure        = 2.684 A
USER
USER    Estimated Free Energy of Binding     =    -8.68 kcal/mol  [=(1)+(2)+(3)-(4)]
USER    Estimated Inhibition Constant, Ki    =  432.51 nM (nanomolar)  [Temperature = 298.15 K]
USER
USER    (1) Final Intermolecular Energy      =    -8.68 kcal/mol
USER        vdW + Hbond + desolv Energy      =    -1.76 kcal/mol
USER        Electrostatic Energy             =    -6.92 kcal/mol
USER    (2) Final Total Internal Energy      =    +0.00 kcal/mol
USER    (3) Torsional Free Energy            =    +0.00 kcal/mol
USER    (4) Unbound System's Energy  [=(2)] =    +0.00 kcal/mol
USER
USER
USER
USER    DPF = C:/Users/pault/OneDrive/Desktop/MolecularModeling/doc/dock1.dpf
USER    NEWDPF move      Aligand.pdbqt
USER    NEWDPF about     37.591000 34.440000 8.368000
USER    NEWDPF tran0     38.702633 35.650398 11.181502
USER    NEWDPF axisangle0        0.374142 0.002073 0.927369 -70.918092
USER    NEWDPF quaternion0       0.217047 0.001203 0.537986 -0.814530
USER
USER                               x       y       z     vdW    Elec       q      RMS
ATOM      1  P    PO4 A 400       38.701  35.651  11.182 +0.25 +1.33    +0.436     2.684
ATOM      2  O1   PO4 A 400       40.124  35.545  10.686 -0.44 -2.22    -0.609     2.684
ATOM      3  O2   PO4 A 400       38.253  37.088  11.125 -0.31 -1.68    -0.609     2.684
ATOM      4  O3   PO4 A 400       38.630  35.162  12.598 -0.94 -2.12    -0.609     2.684
ATOM      5  O4   PO4 A 400       37.804  34.805  10.316 -0.32 -2.23    -0.609     2.684
TER
ENDMDL
```

Figure 36: .dlg file - Lowest energy docked conformation

Best stable conformation with minimum energy :-

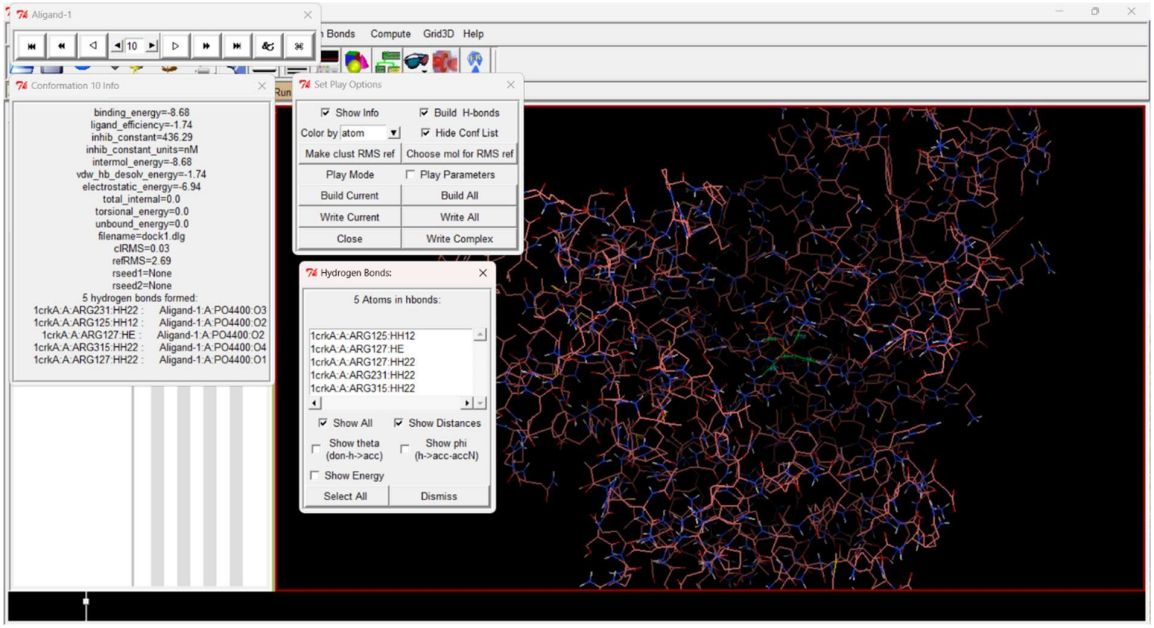


Figure 37: Hydrogen bonds and distances

Visualization of conformation in Pymol tool :-

1. target protein and ligand interaction:

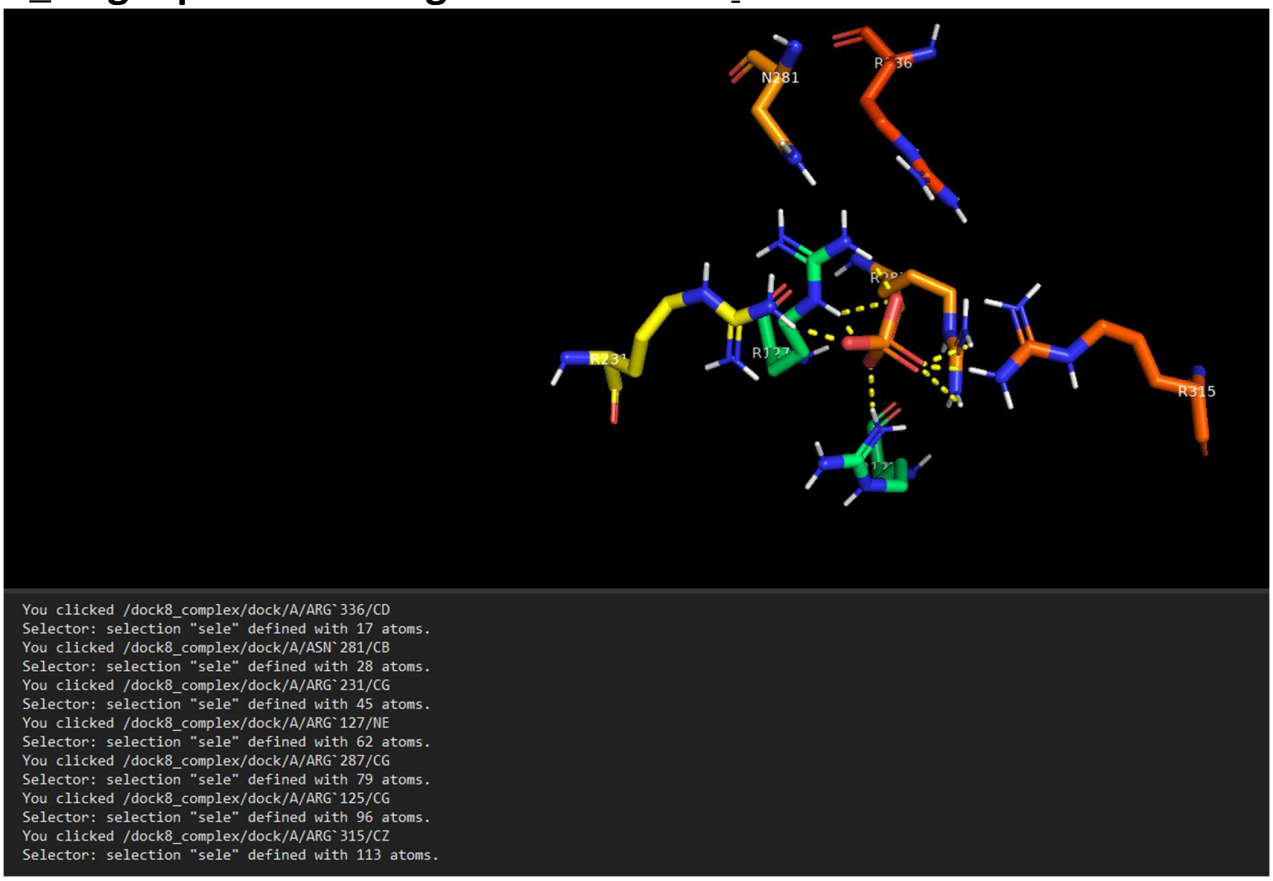


Figure 38: Target protein and ligand interactions

2. Possible polar contacts:

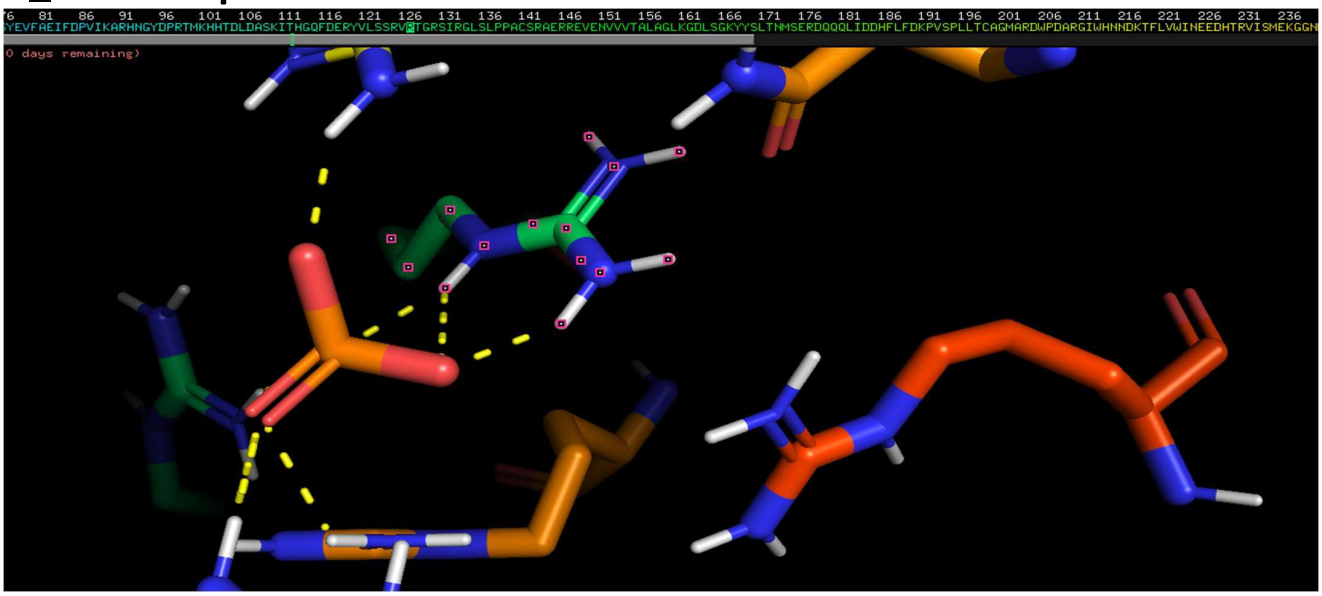


Figure 39: Polar contact between target protein residue and ligand

- **INTERPRETATION:**

- Total Kollman Charge: -201.375
- Total Gasteiger Charge: -1.9999
- Algorithm : Lamark Genetic Algorithm
- Total number of conformation: 10 conformations
- Best interaction and minimum energy conformation: conformation 8
- Lowest binding energy: -8.68 (found in 8^{th} run)
- Cluster rank: 1
- Interactions between target site and ligand molecule:- Number of hydrogen bonds:5 bonds
- Name of the residues interacting: - ARG231, ARG125, ARG121, ARG315, ARG127.

The lowest binding energy in the RMSD table corresponds to the most stable and favourable bind conformation, such as stronger interactions , optimal orientation, stable complex and higher affinity.

Target protein and ligand interaction and polar contacts(hydrogen bonds) can be visualized in PyMOL , which represented in yellow doted lines.