Methods Article

Progress in Drug Discovery & Biomedical Science

HH PUBLISHER

Illustrated step by step protocol to perform molecular docking: Human estrogen receptor complex with 4-hydroxytamoxifen as a case study

Shafi Ullah Khan[1*], Nafees Ahemad[1, 2], Lay-Hong Chuah[1,3], Rakesh Naidu[4] and Thet Thet Htar[1*]

[1]School of Pharmacy, Monash University Malaysia, Jalan Lagoon Selatan, Bandar Sunway, 47500 Subang Jaya, Malaysia
[2]Tropical Medicine and Biology Multidisciplinary Platform, Monash University Malaysia, Jalan Lagoon Selatan, Bandar Sunway, 47500 Subang Jaya, Malaysia
[3]Advanced Engineering Platform, Monash University Malaysia, Jalan Lagoon Selatan, Bandar Sunway, 47500 Subang Jaya, Malaysia
[4]Jeffrey Cheah School of Medicine and Health Sciences, Monash University Malaysia, Jalan Lagoon Selatan, Bandar Sunway, 47500 Subang Jaya, Malaysia

Abstract: Molecular docking is one of the most frequently used technique in structure-based drug design. Molecular docking can predict the binding-conformation and interactions of small molecule to the appropriate binding site within the target protein. This tutorial aimed to design a step by step protocol to get the basic insight into the molecular docking calculations employing very simple and easy to follow procedure.

Keywords: Molecular docking; case study; AutoDock; human estrogen receptor; protein-ligand interactions

***Correspondence:** Shafi Ullah Khan, School of Pharmacy, Monash University Malaysia, Jalan Lagoon Selatan, Bandar Sunway, 47500 Subang Jaya, Malaysia; shafiullapharmd@gmail.com;shafi.khan1@moansh.edu. Thet Thet Htar, School of Pharmacy, Monash University Malaysia, Jalan Lagoon Selatan, Bandar Sunway, 47500 Subang Jaya, Malaysia; thet.thet.htar@monash.edu.

Received: 2nd Feb 2020
Accepted: 2nd March 2020
Published Online: 12th March 2020

Citation: Khan SU, Ahemad N, Chuah L-H, *et al.* Illustrated step by step protocol to perform molecular docking: Human estrogen receptor complex with 4-hydroxytamoxifen as a case study. Prog Drug Discov Biomed Sci 2020; 3(1): a0000054. https://doi.org/10.3687/pddbs.a0000054.

Background and Introduction

Molecular docking is one of the most frequently used methods in structure-based drug design, due to its ability to predict the binding-conformation of small-molecule ligands to the appropriate target binding site[1]. Molecular docking assists in the process of computer-aided drug designing by considering every possible conformation of the protein and ligand molecule[2]. Docking has a considerable advantage when it comes to the study of protein interactions. To date, many molecular docking protocols have been reported, but to the best knowledge, no illustrated protocol has been reported[3]. Herein, we reported the most straightforward approach for performing a molecular docking study. We designed this protocol for users who does not have any prior experience of molecular docking study. To make this tutorial applicable and generalize, we started from downloading and installation of all freely available pre-requisite software, which was followed by subsequent illustrated step by step methodology.

If the protocol described here is not applicable for your target of interest, a thorough literature review on the architecture and assembly of the target protein and the amino acid residues at the binding site is required. Sometimes active sites also comprised of metals, co-factor, or conserved water molecule, which need to be treated very carefully[4].

Before setting to implement this tutorial for any target or protein or receptor of interest, keep in mind that this tutorial is elementary, to make new user used to different task execution in Molecular docking by using Autodock Tool[5], autogrid4, autogrid4[6] and Discovery Studio[7].

We believe that this illustrated step by step tutorial will be helpful for all novice users who have an interest in drug designing, *in silico* screening, virtual screening, and binding of the ligand with protein structure. Advice to reader and user is that follow every step to execute this tutorial your-own and then change accordingly for your protein or target of interest[6,8].

In this tutorial, we use the human estrogen receptor alpha ligand-binding domain co-crystallized with 4-hy-

droxytamoxifen. The reason for selecting this receptor is that it plays a vital role in the physiological development and function of a variety of organ systems to varying degrees, including the reproductive, central nervous, skeletal, and cardiovascular systems[9].

Methodology

This methodology described step by step protocol to perform molecular docking study of human oestrogen receptor with 4-hydroxytamoxifen from basic required tools up-to the visualization steps of protein-drug interactions.

Downloading and installation of basic tools

Basic required tool required for completing this protocol are Window 8 or 10, freely available software such as MGLtools, binary files of Autodock4 and Autogrid4 and Discovery Studio Visualizer. MGL tools can be downloaded from below link (accessed checked on 30-08-2019) http://mgltools.scripps.edu/downloads/downloads/tars/releases/REL1.5.6/mgltools_win32_1.5.6_Setup.exe. For downloading the binary files of Autodock and Autogrid go to below link (accessed checked on 30-08-2019) http://autodock.scripps.edu/downloads/autodock-registration/tars/dist426/autodocksuite-4.2.6.i86Windows.exe .

To download Discovery Studio Visualizer, go to

https://www.3dsbiovia.com/products/collaborative-science/biovia-discovery-studio/visualization-download.php (accessed checked on 30-08-2019) and after filling the required form, a download link will be sent to your provided email. After downloading, install all three software in default settings.

After successful installation of required tools, the next steps with required files are mentioned in schematic flow chart as shown in Figure 1.

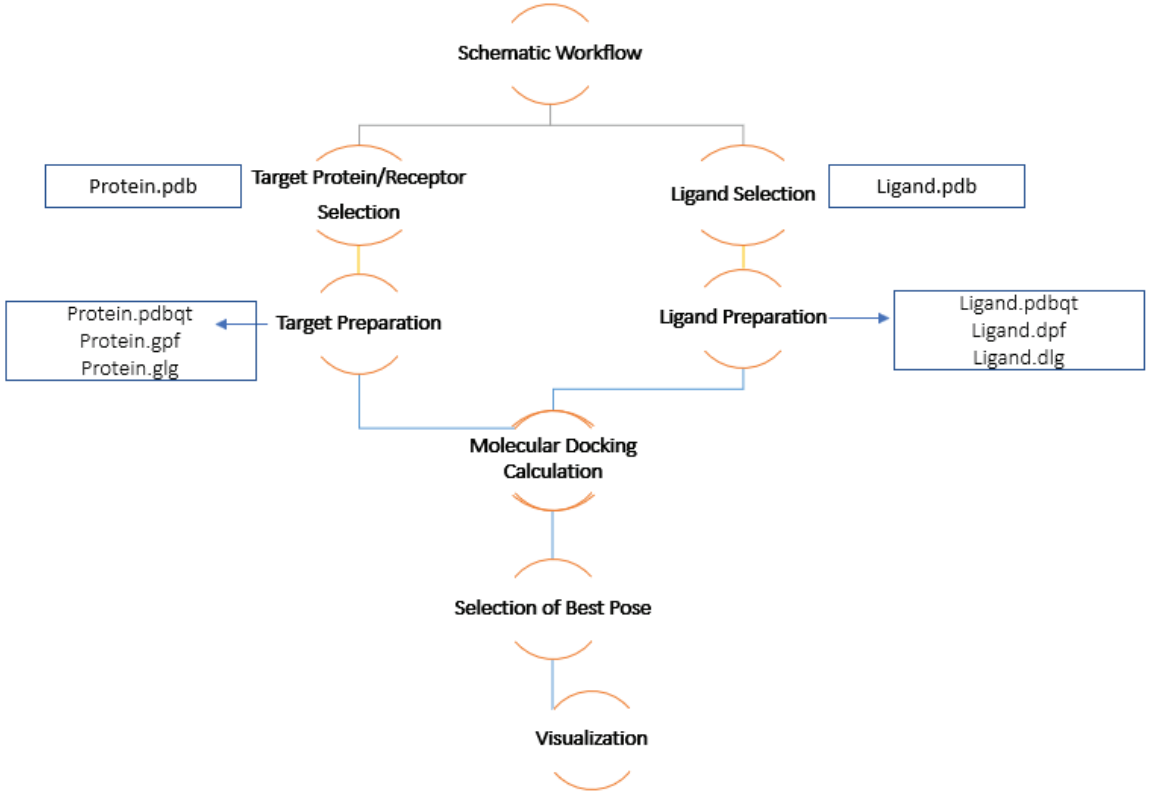

Figure 1. Schematic flowchart for performing Molecular docking studies.

Selection of target protein

Selection of target protein depends on the aim of the study, ***RCSB Protein Data Bank (***https://www.rcsb.org/) is one of the most widely used database from which three dimensional structure of target protein can be downloaded. In the absence of protein structure in database, other tools of structure prediction (homology modelling, ab initio modelling etc.) can be used (which is beyond this protocol). In this protocol we use Human Estrogen Receptor Alpha Ligand-Binding Domain in Complex With 4-Hydroxytamoxifen having PDB ID of **3ERT** as shown in Figure 2 and 3.

Clicking on 3ERT, will open a new pop-up window, first click on ***Download Files*** on right hand side and then in download menu click on ***PDB Format.*** It will download **3ert.pdb** file in download directory. Open **3ert.pdb** (can be any other protein in case of different target of one interest) in Discovery Studio Visualiser

Figure 2. RCSB Protein Data Bank website display.

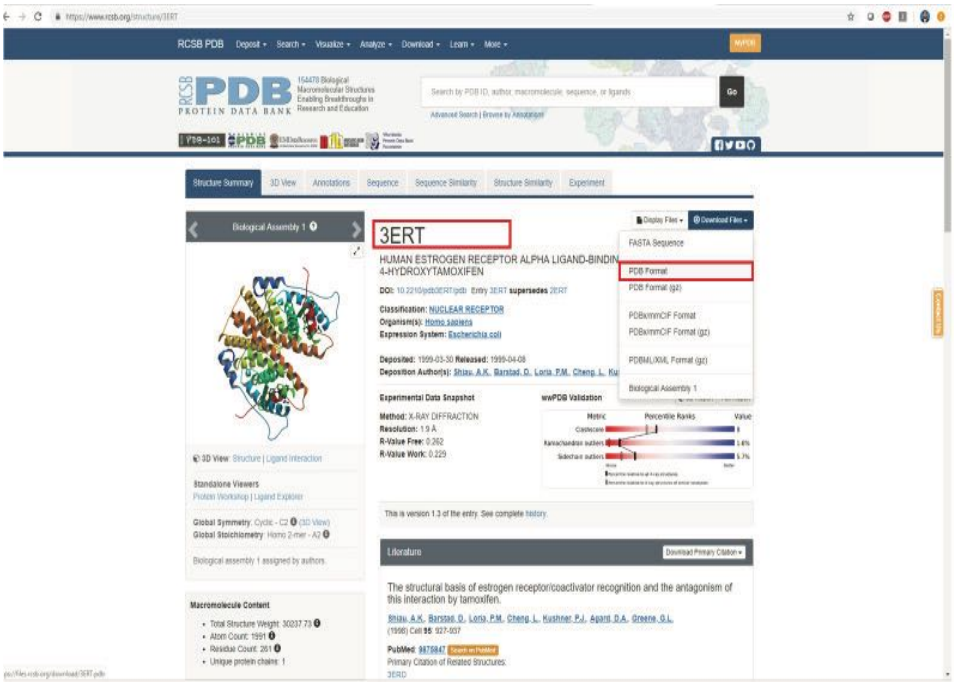

Figure 3. Downloading the pdb format file of protein.

Open **3ert.pdb** in Discovery Studio Visualiser

Go to **View > Hierarchy** or simply Pressing ***Ctrl+H*** on keyboard will open ***Hierarchy window*** on left side of graphical window as shown and highlighted in Red rectangle in Figure 4.

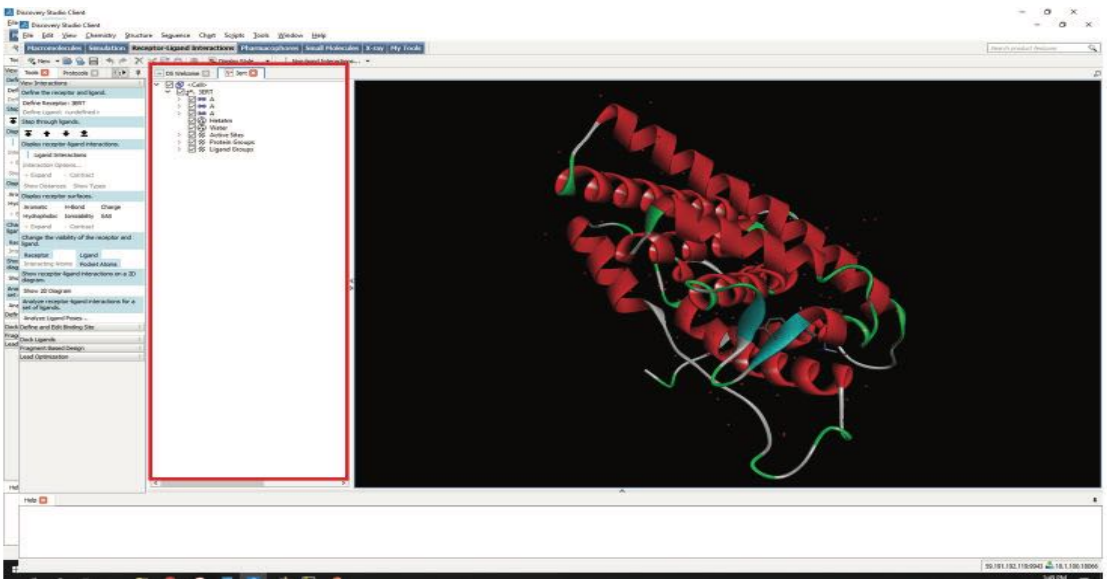

Figure 4. Graphical representation of protein.

You will find that there are three A, correspond to different entities like 1st one is comprised of amino acid (Leu306-Pro552), second one is Co-crystal ligand (4-OHT600) and third one is comprised of H_2O (HOH1-HOH79). Using the inbound ligand, one can easily locate the active site dimension and coordination for Grid box generation.

Finding the *xyz* dimension and coordination of Co-crystal ligand (4-OHT600)

Expand 2nd A in the hierarchy window and **right click on 4OHT600 will pop-up** small window **as shown in Figure 5a and 5b**. At the end of the pop-up window Click on Attributes of OHT600.

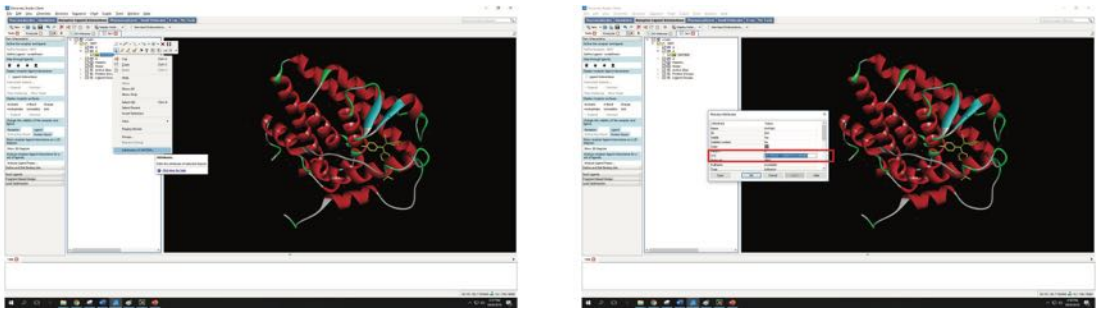

Figure 5a and 5b. Steps for finding the xyz coordinate of active site or bound ligands.

Note down the XYZ coordinates from this small pop-up window. In this case of OHT600 it is (X=31.574552, Y=1.590379 and Z=25.599483).

Selection of Ligand

After noting down the XYZ dimension and coordinate, we need to remove water and extract co-crystal ligand (in this case OHT600) for redocking and optimization of docking protocol. Ligands and small molecules can be sketch using different software like MarvinSketch, ChemDraw as well various online servers frequently using one is **Pubchem** online server database. In this we are using OHT600, so select 2nd A chain in the hierarchy window and do cut paste in new window of Discovery studio as shown in Figure 6.

Click on **File** on left side and Save as **lligand.pdb** file in working directory (One can save with different name to differentiate different ligands from one another).

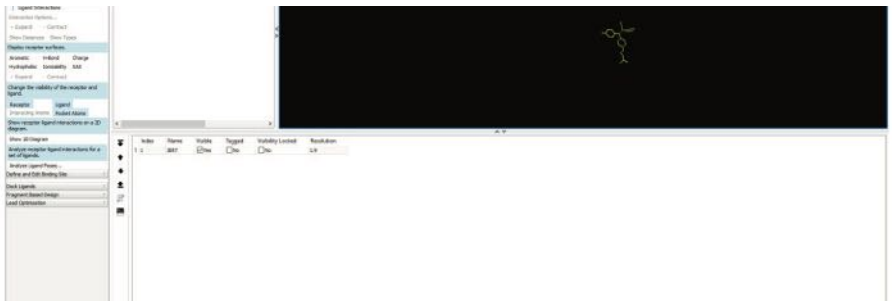

Figure 6. Extracting ligand structure obtained from crystal structure of 3ert.pdb.

Delete the 3rd A chain in hierarchy window will delete all water molecule

Note: This is not the case for every protein, some time in protein structure, few conserve water molecules need to be retained which involves in the stabilization of ligand-protein complex. Now this protein is free from heteroatoms i.e water and co-crystal ligands.

Click on **File** on left side and save prepared structure as **protein.pdb** file in the working directory. (One can also save with different name to distinguish between different protein from each another)

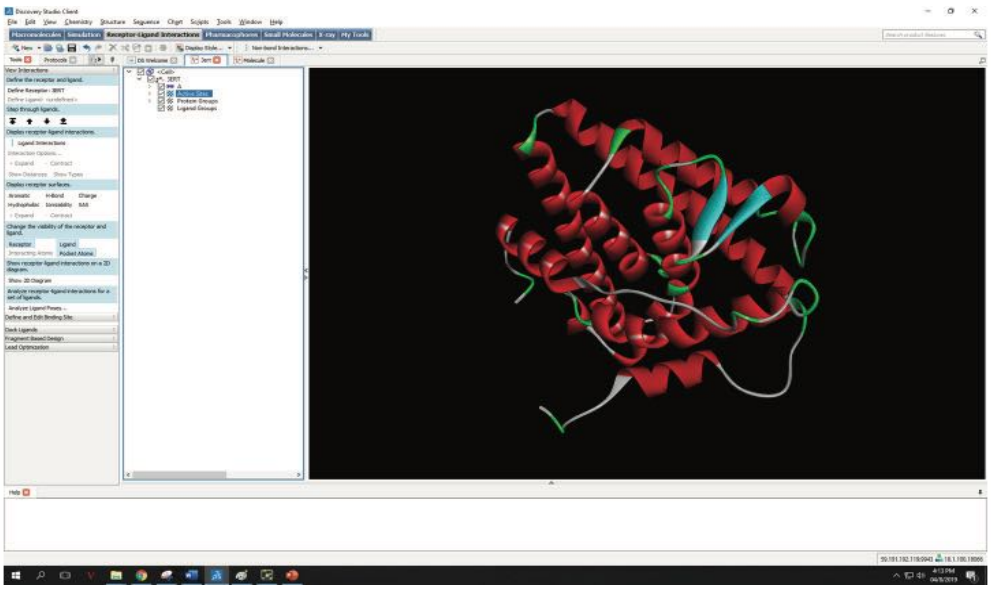

Figure 7. Prepared protein structure without water and bound ligands.

Crosscheck **protein.pdb** and **ligand.pdb** files are present within the specified working directory or folder (in this case our working directory is named as **Tutorial**).

Preparation of Target Protein or Receptor

After successful downloading of ligand and protein files, next step is the preparation of pdbqt format files of both protein as well as ligand. Pdbqt file for ligands can also be generated using OPEBBABEL but for here MGLTools will be used for preparation of PDBQT files of both Protein and ligand (protein.pdbqt, ligand.pdbqt). Along with pdbqt file, required Grid parameter file (protein.gpf) and Docking Parameter file (ligand.dpf) will be generated using MGLtools (AutoDock Tools).

First Open AutoDockTools-1.5.6 as shown below as AutoDock Tools Display

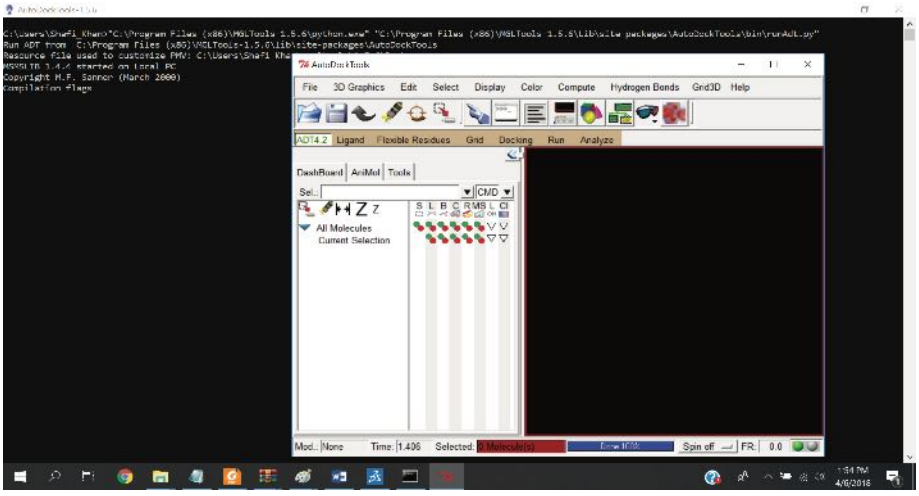

Figure 8. Graphical display of AutodockTools.

Preparation of Protein files (protein.pdbqt)

Open **File** and then click on **Read Molecule**

Go to specified working folder where preotin.pdb was saved

Select and Open protein.pdb

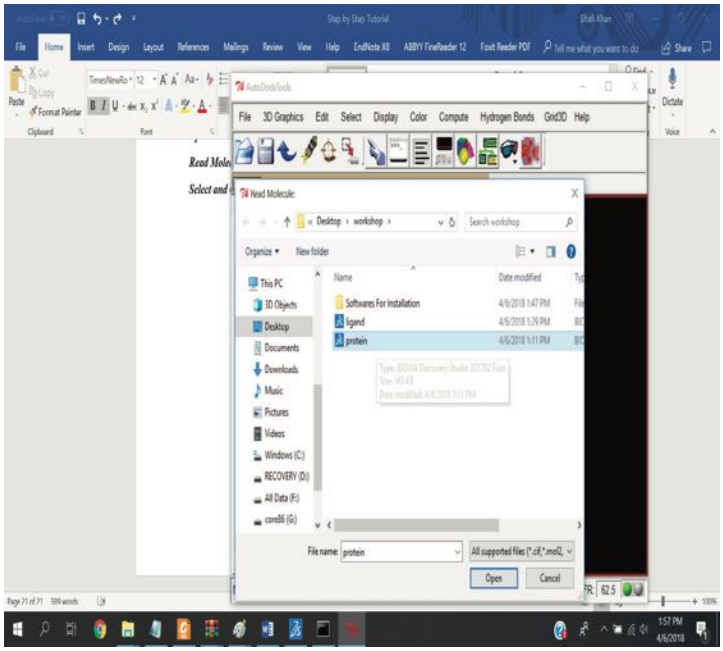

Figure 9. Selecting protein.pdb from working directory.

Click on **Edit,** followed by click on **Hydrogens** and then **Add**

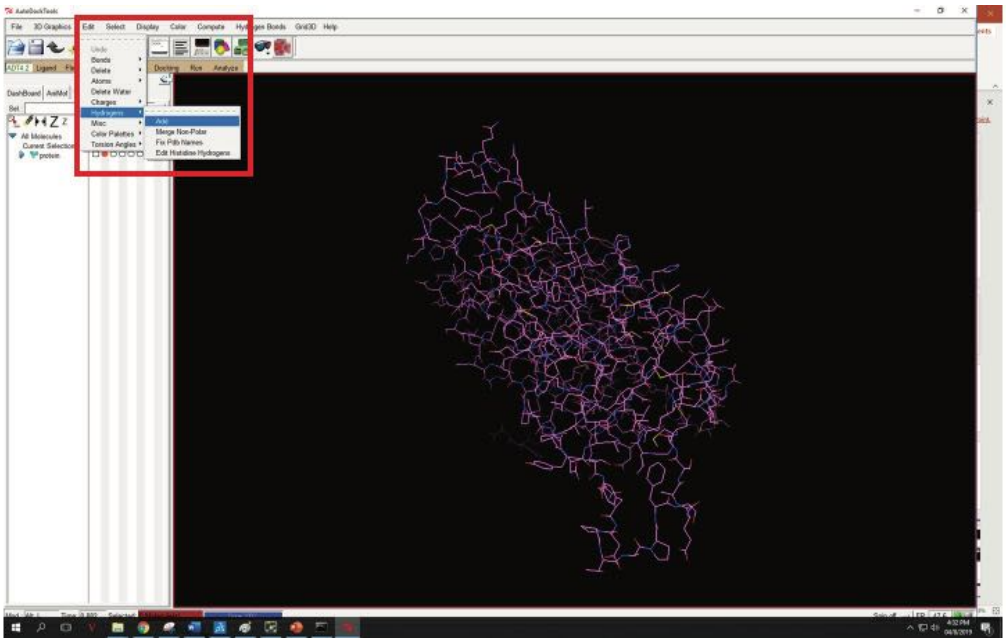

Figure 10. Addition of Hydrogen to protein structure.

Click ***Polar Only*** in next pop-up window and then Click ***OK***

In next step, go to ***Edit*** followed by click on ***Charges*** and then add on ***Kollman Charges*** and compute Gasteiger charges and click **OK** in pop-up window

In next step, first click on ***Grid*** followed by ***Macromolecules*** and then click on ***Choose*** and click on ***protein*** in pop-up window and finally click on ***Select Molecule.*** *Click* ***OK*** when warning window pop-up and then Save ***protein.pdbqt*** in working directory. After saving macromolecule in ***protein.pdbq*** format (you will note that even the colour get changed from pink to grey).

After saving **protein.pdbqt file** in specified folder, next step is the preparation ligand.pdbqt file. Note: (Be careful with case sensitive letter and naming of protein naming as it may affects the follow-up commands execution). Close and re-open AutodockTool to avoid any confusion

Preparation of Ligand (Ligand.pdbqt)

Open ***Ligand*** and Click on ***Input*** and then ***Open***

Change file format from**. pdbqt** to **.pdb** in next pop-up window as shown in Figure.

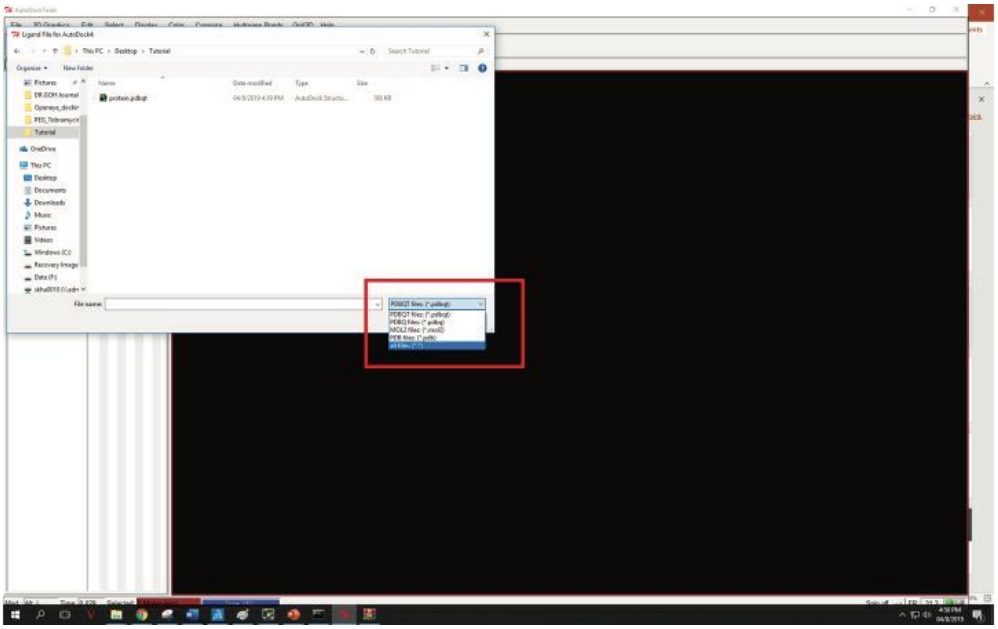

Figure 11. Changing the extension format before opening of ligand.pdb file from working directory.

Select ***ligand.pdb*** and then click ***OPEN***

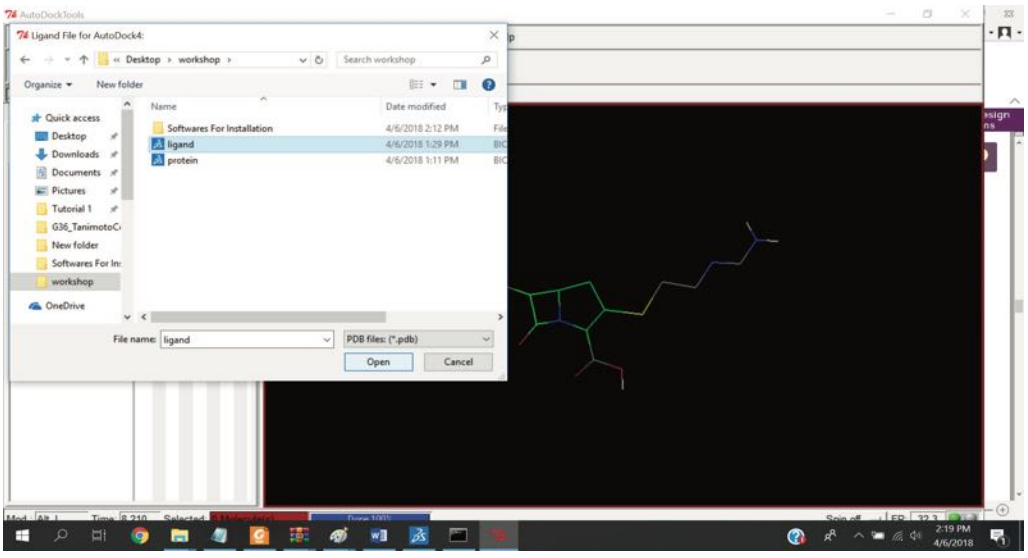

Figure 12. Selecting and opening of ligand.pdb file from working directory.

Click ***OK*** in pop-up window of ***summary of ligand***

Again, click on ***Ligand*** menu then Click on ***Torsion Tree followed by*** *Clicking on* ***Detect Root*** as shown in Figure.

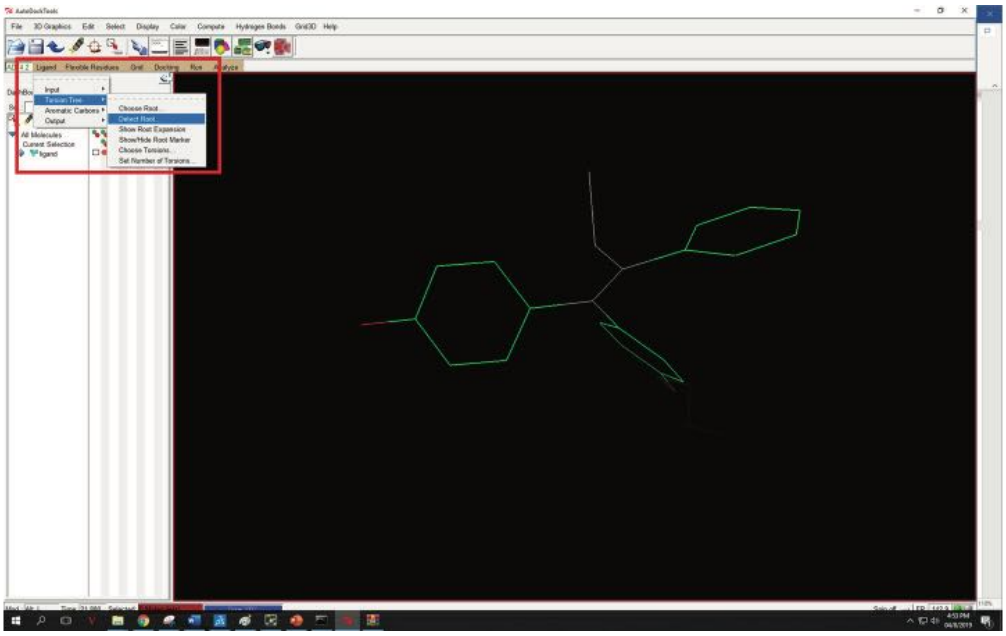

Figure 13. Detection of root in ligand structure.

Again, ***click on Ligand*** then Click on ***Torsion Tree*** followed by clicking ***Set Number of Torsions***

Set number of active torsions between ***1 to 8*** in pop-up window (only in this case we select 8 i.e maximum no. of torsion), then Click on ***Dismiss***

Again, Open **Ligand** menu then Click on **Output,** followed by clicking **Save as PDBQT**

Save Ligand file as **ligand.pdbqt** in specified working folder where protein.pdbqt was saved.

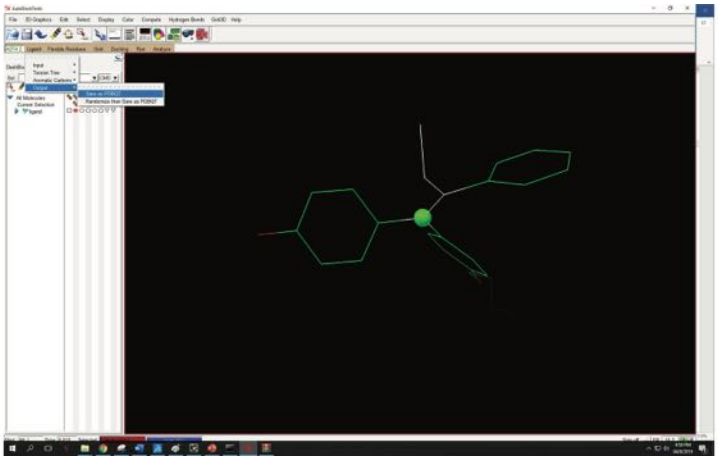

Figure 14. Ligand structure ready for saving in ligand.pdbqt format.

After preparation and saving of protein.pdbqt and ligand.pdbqt files, next step is the Preparation of Grid Parameter File (protein.gpf)

Preparation of Grid Parameter File (protein.gpf)

First, open **protein.pdbqt** and **ligand.pdbqt** one by one in AutodockTool

Then go to **Grid** and click on **Set Map Types** followed by clicking **Choose Ligand** as shown in Figure below.

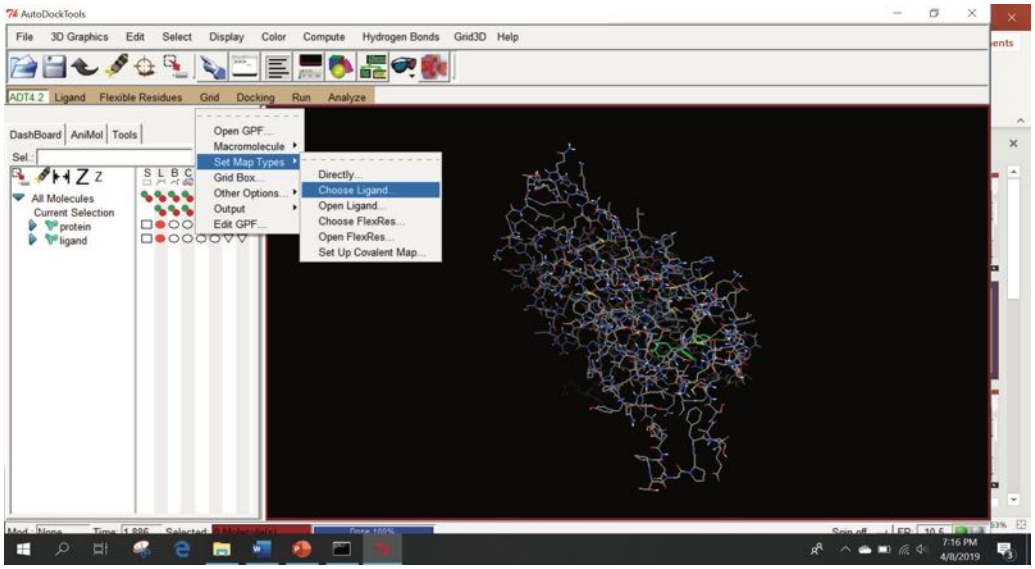

Figure 15. Setting the map type of ligand.

In subsequent pop-up window, first Click on **Ligand and** then **c**lick on **Select Ligand**

Again, Open **Grid** menu and click on Grid **Box.** Choose **x,y,z** dimension as **40 x 40 x 40** with **spacing** of **0.375. Additionally,** use Center Grid Box coordinate x, y, z center 31.570, 1.590, 25.590 respectively.

Note: these XYZ dimension and coordination, we gotten previously described from the attributes of co-crystal ligand. These coordinates would be different in case of different protein as per to the active site position.

After selection of predefined XYZ dimensions and coordination.

Go to **File** in Grid option window and select **Close saving current.** It will the save the adjusted coordinates and dimension as shown in Figure 16.

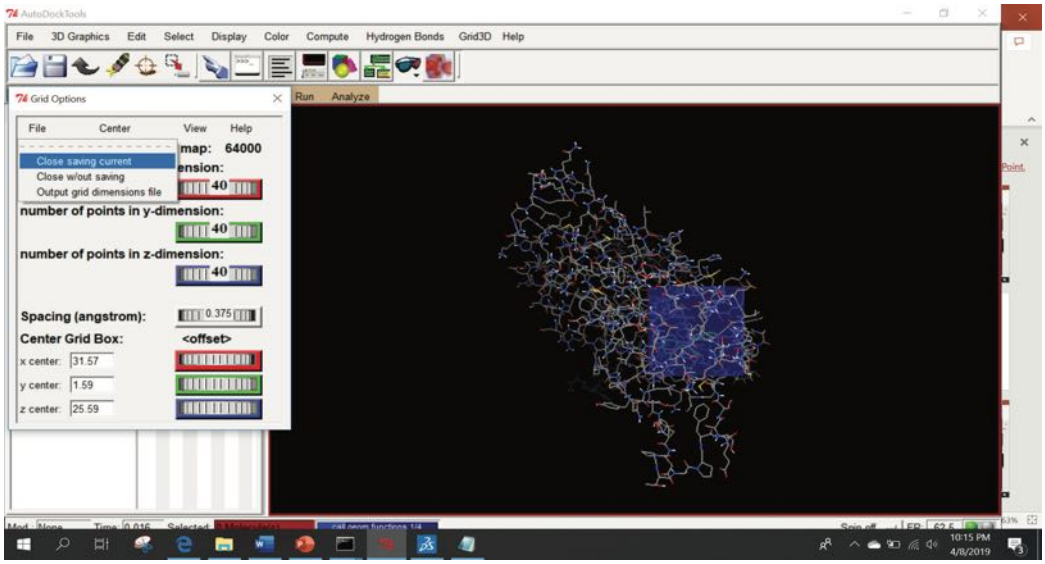

Figure 16. Adjustment of xyz coordinate and grid box volume.

Again, Open **Grid** menu and click on **Output** followed by clicking on **Save GPF**

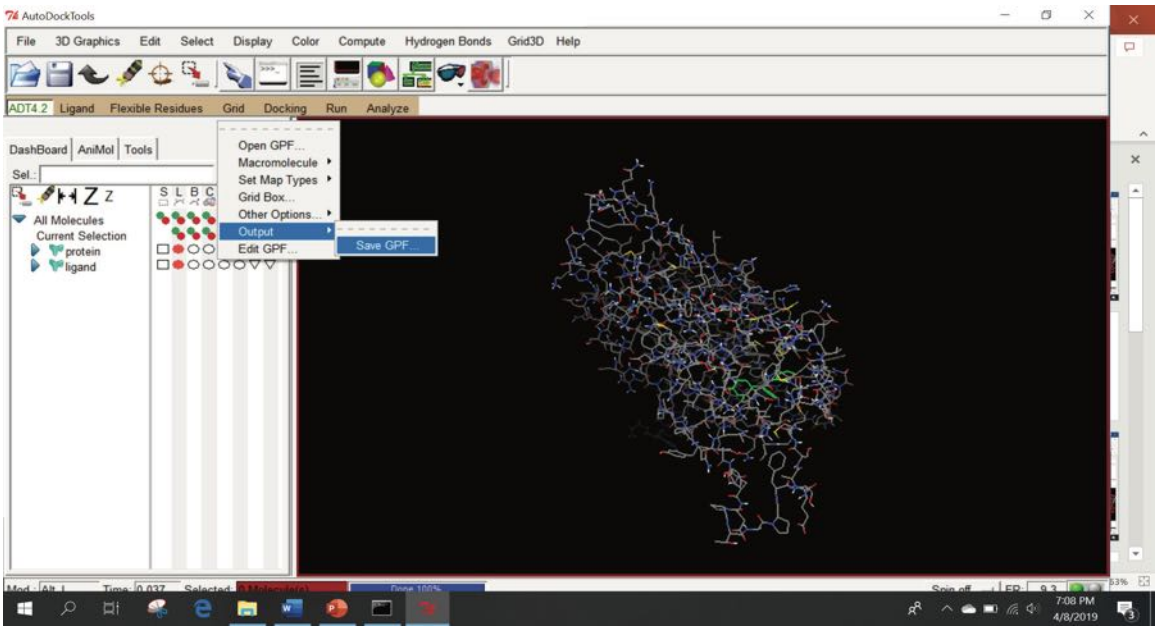

Figure 17. Saving the protein in. gpf format.

Name the file as **protein.gpf** in working folder where other files of **protein.pdbqt** and **ligand.pdbqt** were saved.

Preparation of Docking Parameter File (ligand.dpf)

Open the **Docking** menu and click on **Macromolecules** followed by Clicking **Set Rigid Filename** as shown in Figure.

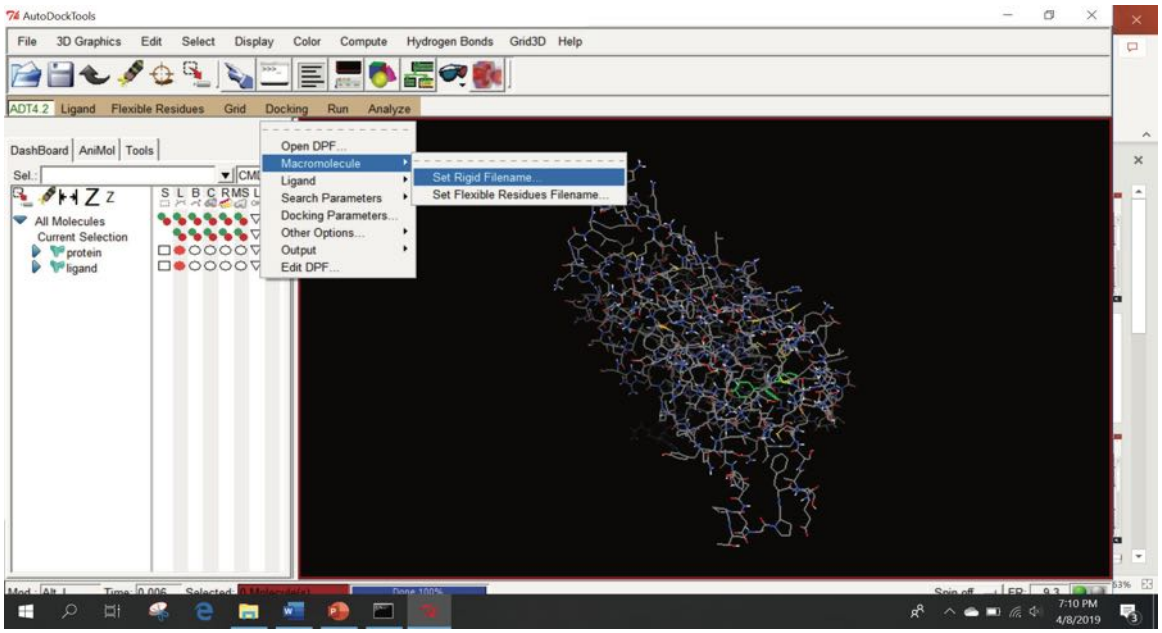

Figure 18. Selection of protein.pdbqt file for from docking manu.

From pop-up window, go to working directory and select **protein.pdbqt** and then **c**lick **Open**

Again, go to **Docking** menu and click on **Ligand** followed by clinking **Choose.**

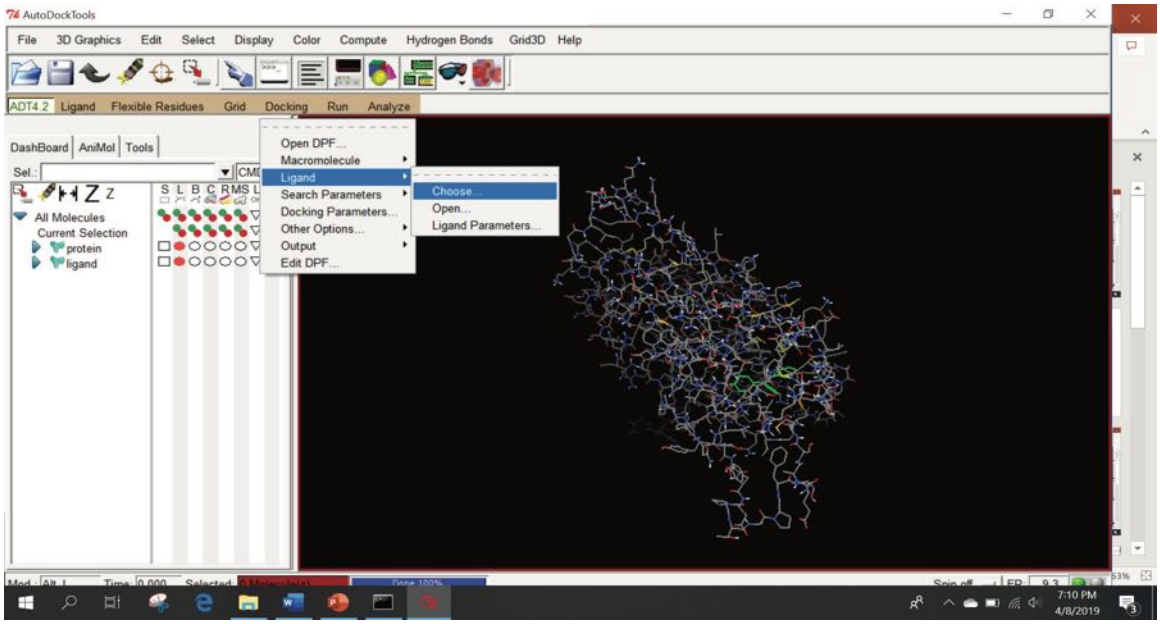

Figure 19. Selection of ligand from docking menu for docking calculation.

In the pop-up window, first select **Ligand** and then click **Select Ligand.**

Again, go to **Docking** menu and then **Search Parameters** followed by clicking on **Genetic Algorithm.** In the pop-up window one can change the number of GA, population size**, in** this case we will use number of GA runs of 10, and maximum number of evals as medium as shown in Figure.

Click on **Accept**

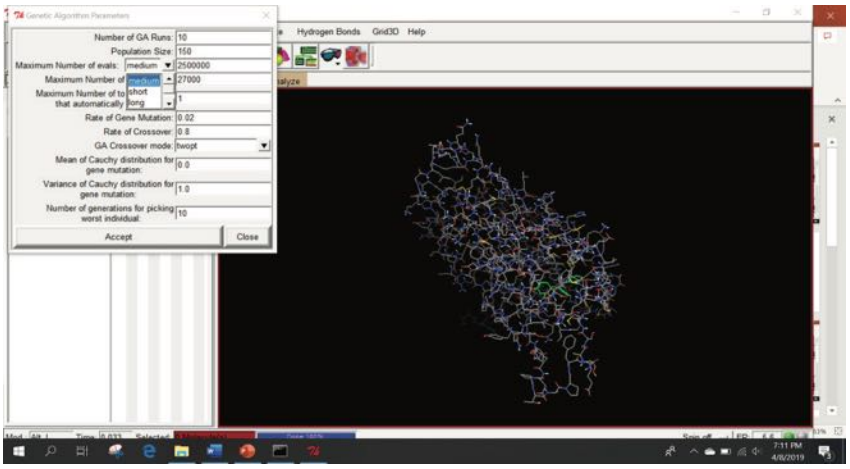

Figure 20. Setting number of runs and poses of ligand to be generated in docking calculation.

Again, go to **Docking** and then click on **Docking parameters** followed by clicking on **Accept** in pop-up window

Again, go to **Docking** and then click on **Output** followed by clicking **LamarkianGA (4.2)**

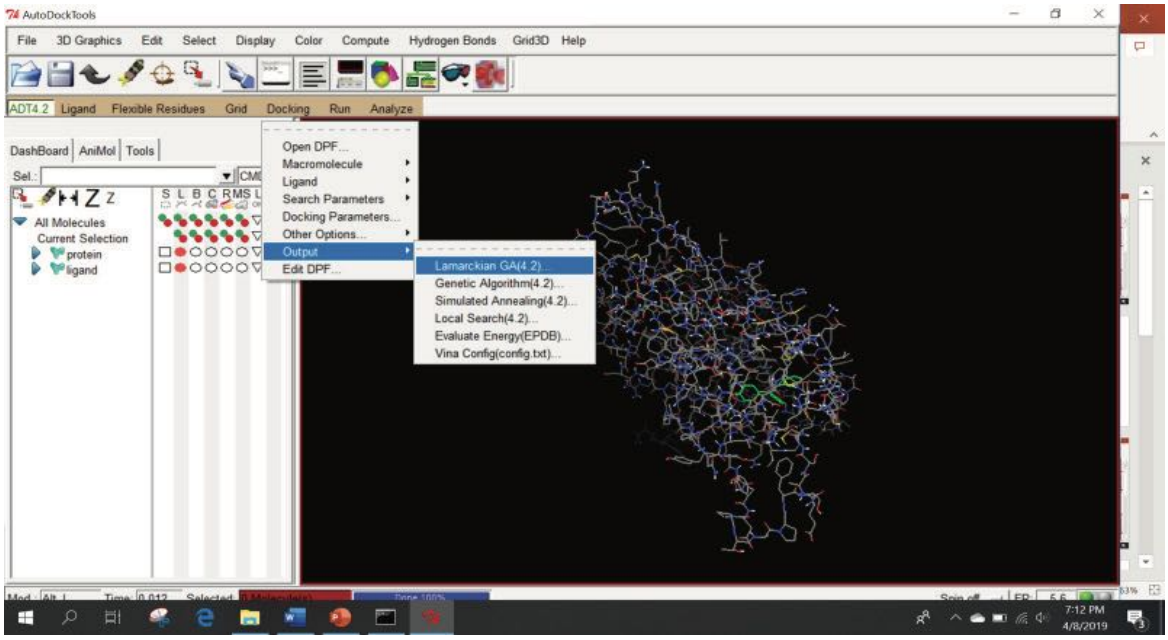

Figure 21. Saving with ligand file with docking parameter in. dpf format.

Name the File as ligand or antthing you wish, but in this case, we save as **ligand.dpf** file

Note: Make sure to save protein.gpf and ligand.dpf in the same folder or working directory where protein.pdbqt and ligand .pdbqt and protein.gpf files were already saved.

Now along with protein.pdb and ligand.pdb, four additional files ***protein.pdbqt***, ***ligand.pdbqt***, ***protein.gpf*** and ***ligand.dpf*** will be present in the working directory or folder as shown in Figure.

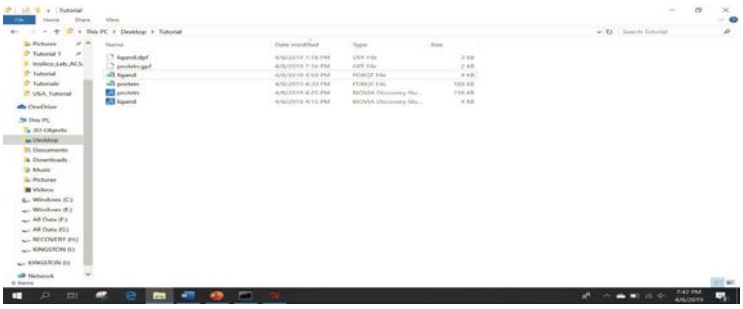

Figure 22. Pre-requisite file of protein and ligand required for docking calculation.

Preparing Autogrid and AutoDock execution setup (Made easy for beginner)

First, locate the installed Autodock and Autogrid in your default installation directory.

Default installation directory could be **"C:\Program Files (x86)\The Scripps Research Institute"**

Copy that ***Autodock4.exe*** and ***autogrid4*.exe** files from Installation folder into the working directory (In current case, it is working directory on Desktop\Tutorial"

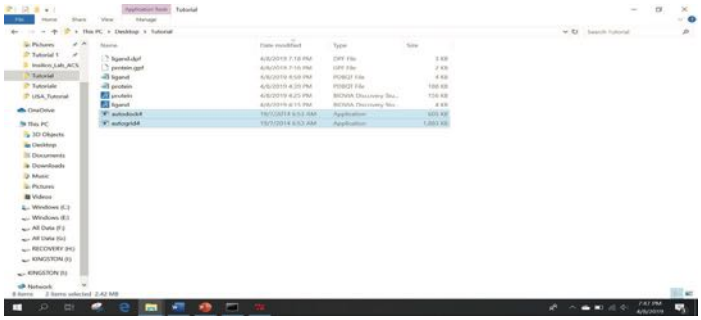

Figure 23. Copy and pasting the Autogrid.exe and Autodock.exe from installation folder to working directory.

After Preparation of Required files such as protein.pdbqt, ligand.pdbqt, protein.gpf and ligand.dpf file as well copying binary file of Autodock4.exe and autogrid4.exe into the working directory, next step is running Autogrid4 and autodock4 comond

Using Command Prompt (command line in window) for Molecular Docking

Open **CMD,** from **Window START** search for **CMD** and click on Command prompt

Go to working directory where required files are present using command i.e

"cd working directory address" which is ***cd C:\Users\Shafi_Khan\Desktop\Tutorial*** in this case study

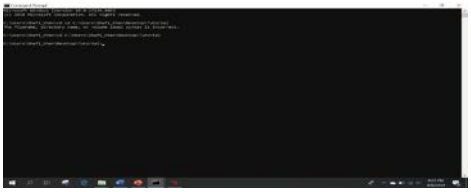

Figure 24. Setting the working directory location in command line.

Now make sure, you are in the workshop directory where other pre-requisite files are presents.

Use below command exactly for running the **Autogrid** to convert ***protein.gpf*** into ***protein.glg***

autogrid4.exe -p protein.gpf -l protein.glg

Note: "Depend on protein name and working directory this can be change accordingly"

***Blinking** at the start of line indicates that the autogrid process is running

Normally, it will take few minutes to complete depend on the performance of the system.

When processing is finished the curser will shift back to the specified working directory location

Use below command exactly for running the **Autodock** to convert ***ligand.dpf*** into ***ligand.dlg***

autodock4.exe -p ligand.dpf -l ligand.dlg

***Blinking** at the start of the line indicates that Autodock process is running

It will take quite some time to complete (from minutes to hours) depends on the performance of system

When processing is finished, the cursor will go shift back to the working directory location.

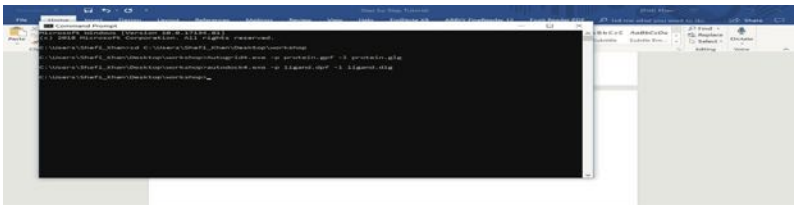

Figure 25. Command line window after successful execution of autogrid and autodock commands.

After completion of **autodock** and **autogrid** processing, both ***protein.glg*** as well as ***ligand.dlg*** along with map files for each atom could be found in working directory folder as shown below

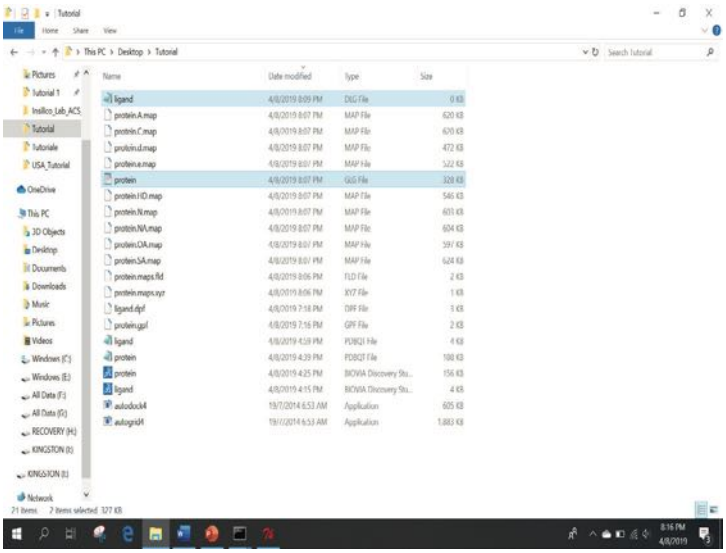

Figure 26. All map files, protien.glg and ligand.dlg files in working directory.

Results Analysis

For analysing the docking results and retrieving the binding interaction of Ligand-protein complex .pdb, open **AutoDock** then click on **Analyze** menu, followed by **Docking** and **Open**

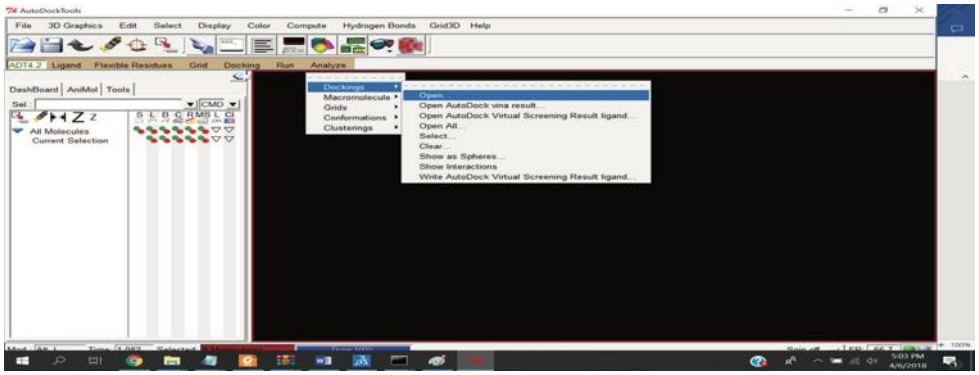

Figure 27.

In pop-up window, select ligand.dlg and click on Open

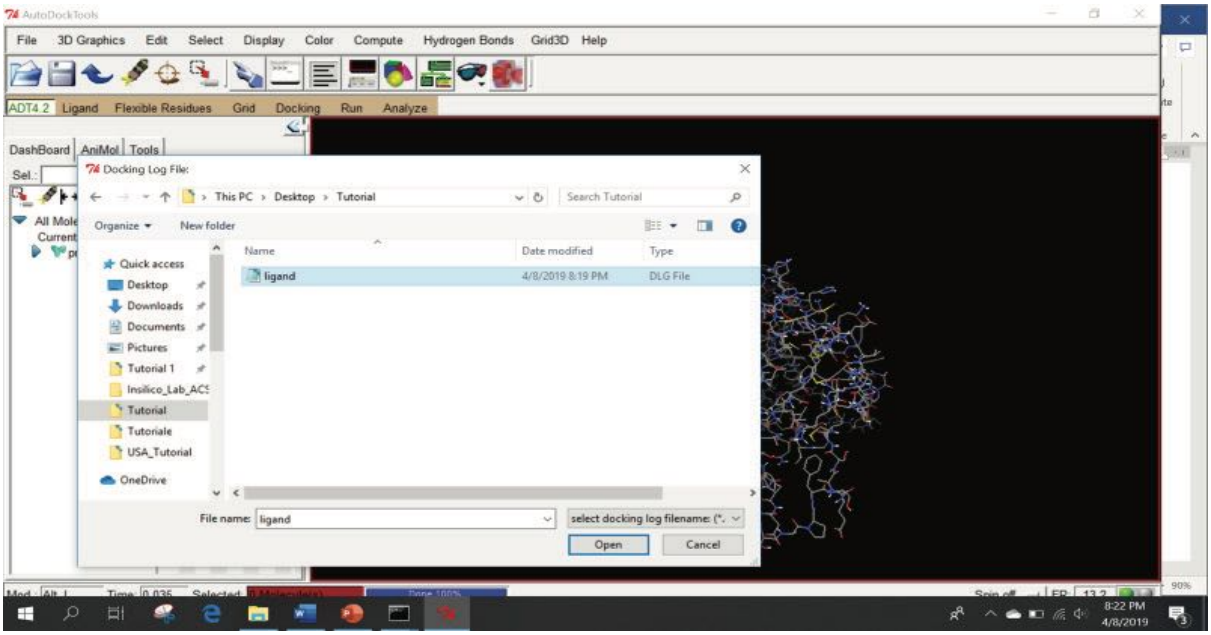

Figure 28. Selection and opening window for ligand.dlg file.

Click **OK**

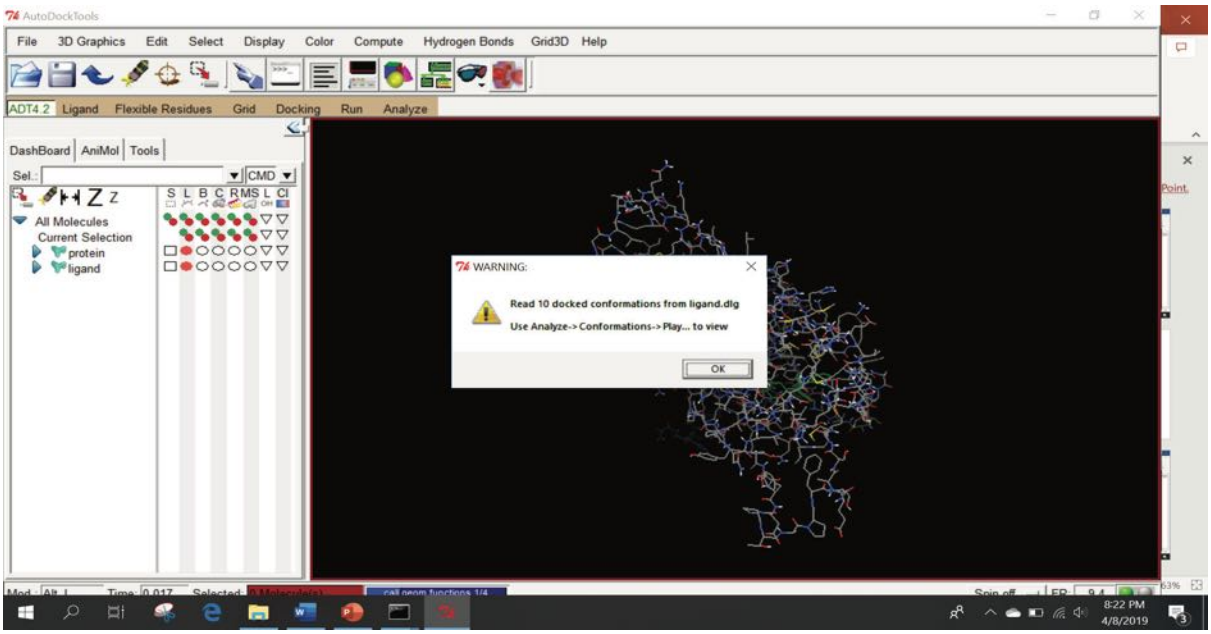

Figure 29. Pop-up window after ligand.dlg selction.

In next step, click on **Analyze** followed by click on ***Macromoleule*** and then Click on **Open**

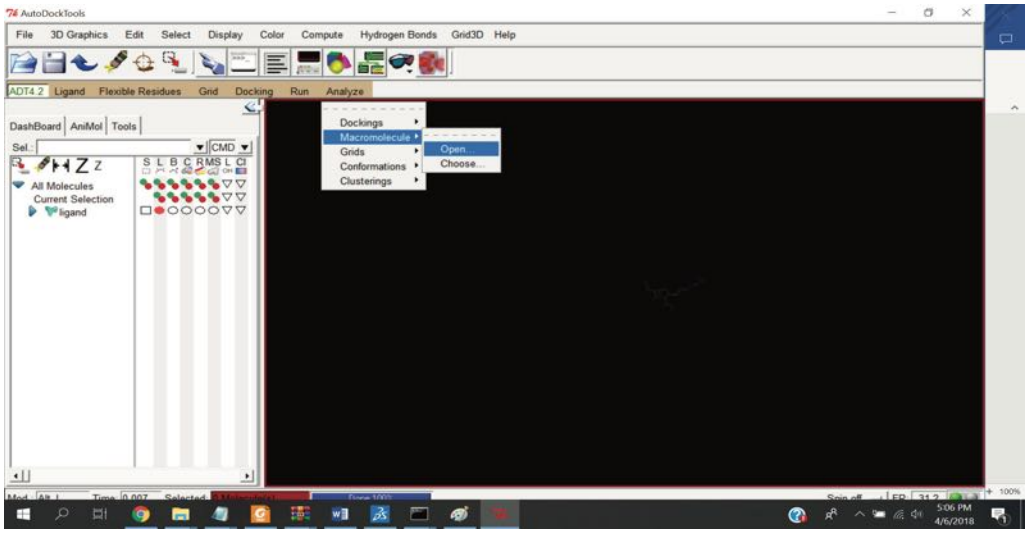

Figure 30. Selection of macromolecule from analyze menu.

Select **protein.pdbqt** and then Click **Open**

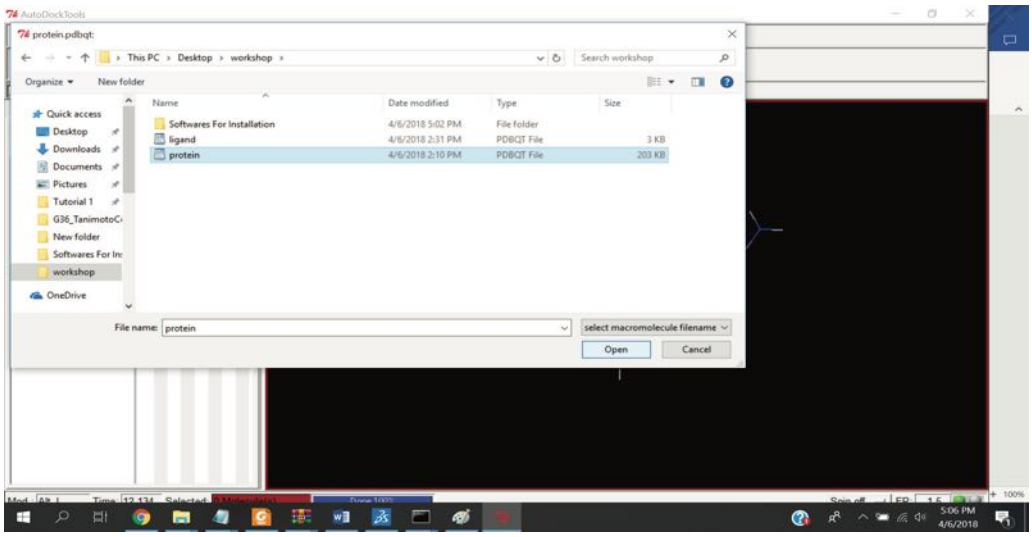

Figure 31. Selection of protein.pdbqt file for opening in Autodocktool.

Again, go to **Analyze,** followed by **c**lick on **Conformations** and then click on **Play**

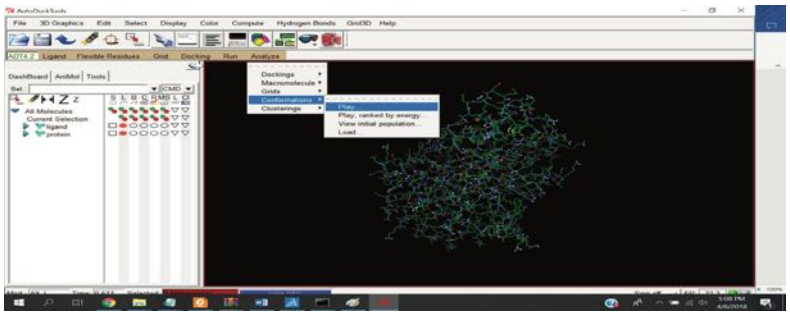

Figure 32. For playing the conformation of ligand.

Click on "**&"** setplay option sign, followed by click on **show info** and then **c**lick on **Write current.**

Above steps will generate a pdbqt file of corresponding docked pose (in this case pose 1 is the best because it has lowest binding free energy of -11.33) in default start-up directory of Autodock tool as shown in below.

Save as ***ligand_1. pdbqt*** in working directory (can be named differently depend on the conformation number with best binding energy).

Note: Lower the binding energy better would be the ligand-protein complex

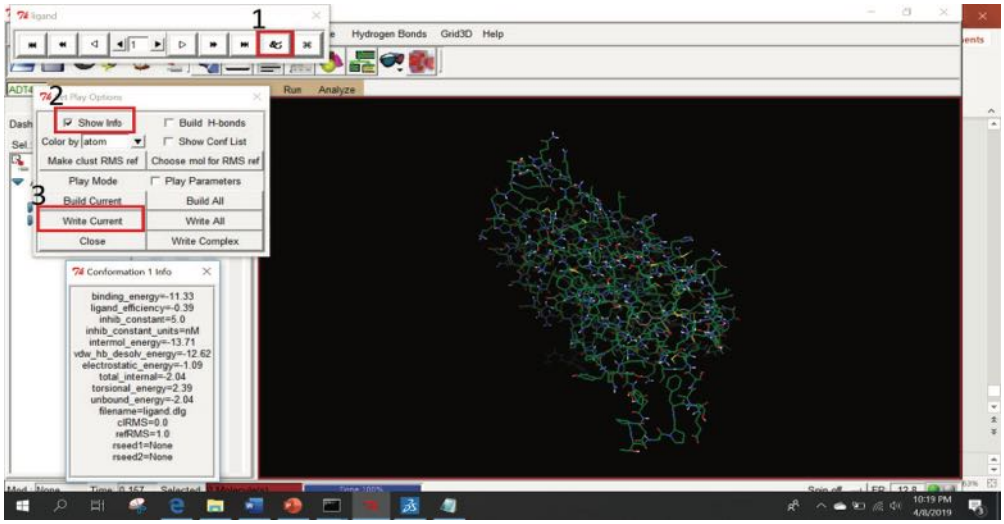

Figure 33. Sequence of steps (1–3) to follow for displaying the set play option, show info and generating the best docked pose.

Click " ▬► " play sign to observe each conformation or pose from 1 to 10 and note down the difference in free binding energy for each conformation.

Click on ***Write Current*** for generating the current individual docked pose of ligand (in this case saved as ligand_1.pdbqt).
Analysis of Nonbonding Interactions

Open **protein.pdbqt** in Discovery Studio then drag and drop **ligand_1. pdbqt** from working directory into graphical window of discovery studio as shown in Figure 34.

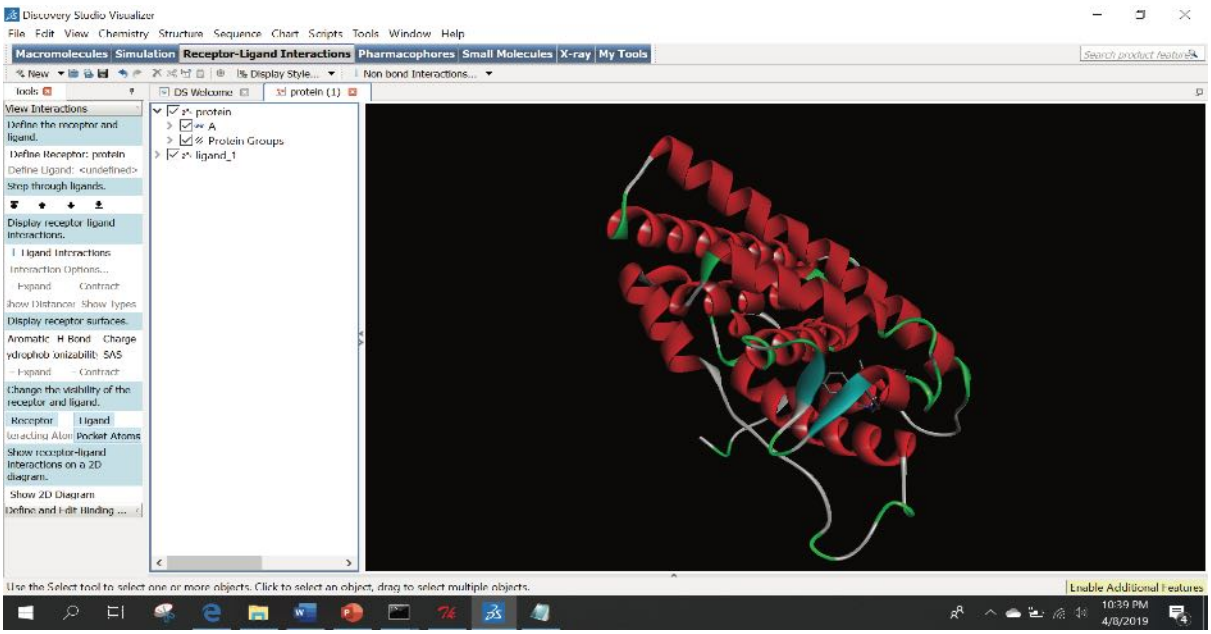

Figure 34. Opening and displaying the protein and best docked ligand in DS.

Click **Scripts** in Discovery studio followed by clicking **Ligand Interactions** and then click on **Show Ligand Binding Site Atoms** as illustrated in Figure 35.

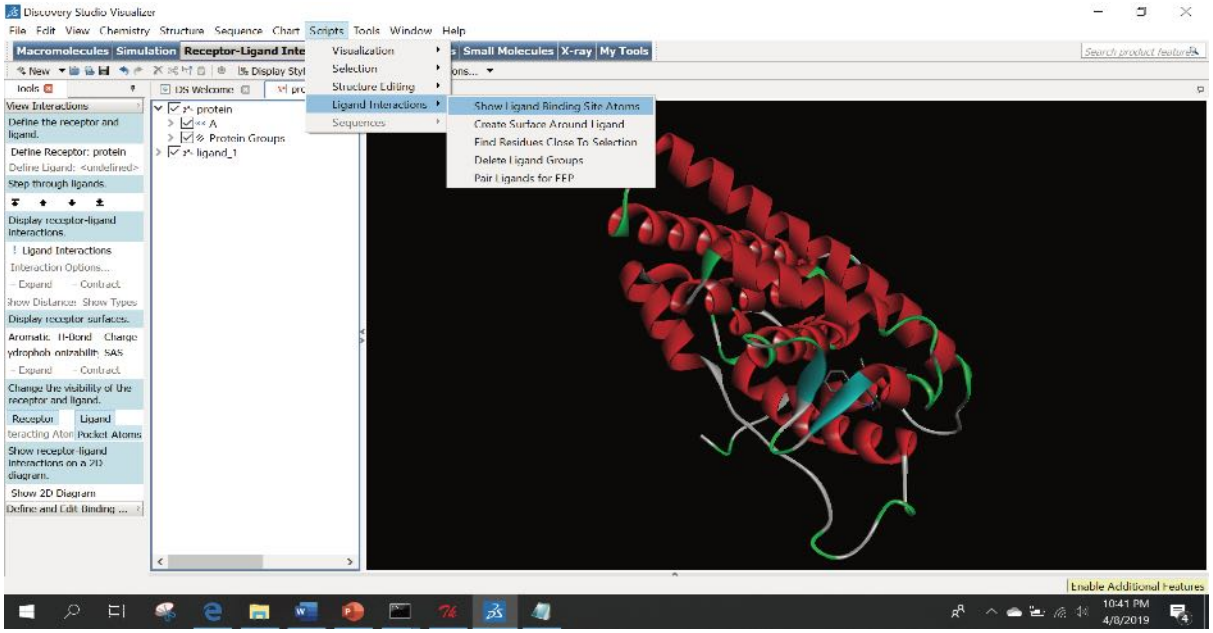

Figure 35. Step to follow for showing the ligand binding site atom.

Right click in **graphical window** and select **Labels** and Click **Add**

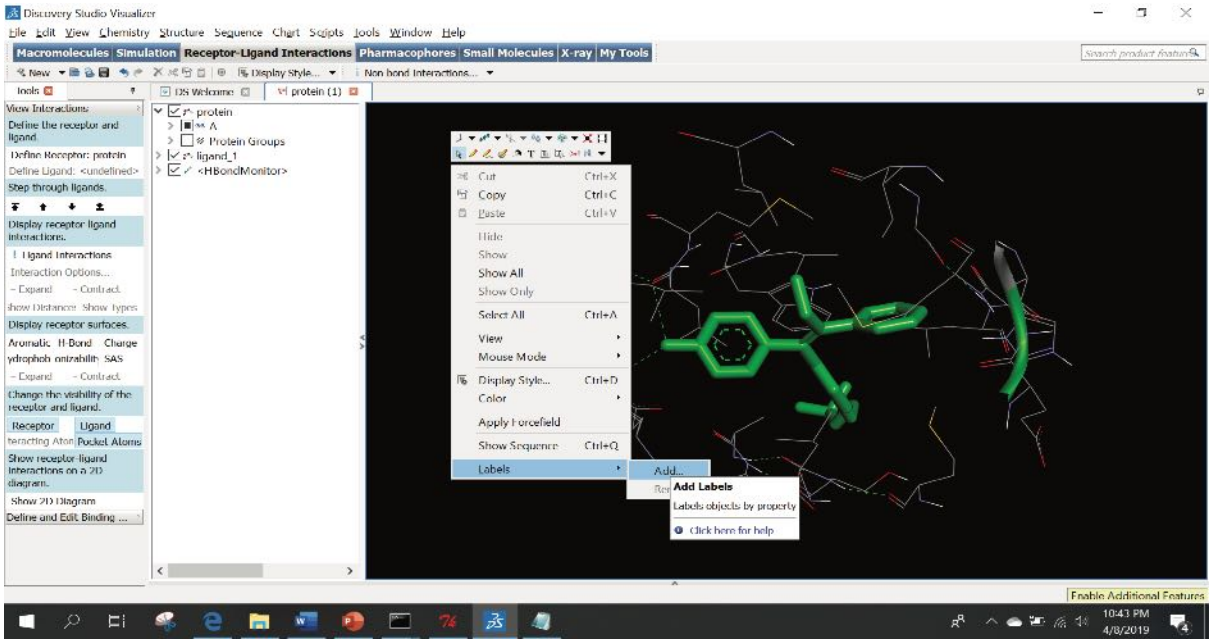

Figure 36. To label the name of interacting amino acid residues.

Select **Object: Amino Acid,** followed by selecting **Attributes: 3 Letter & ID#** and click **OK**

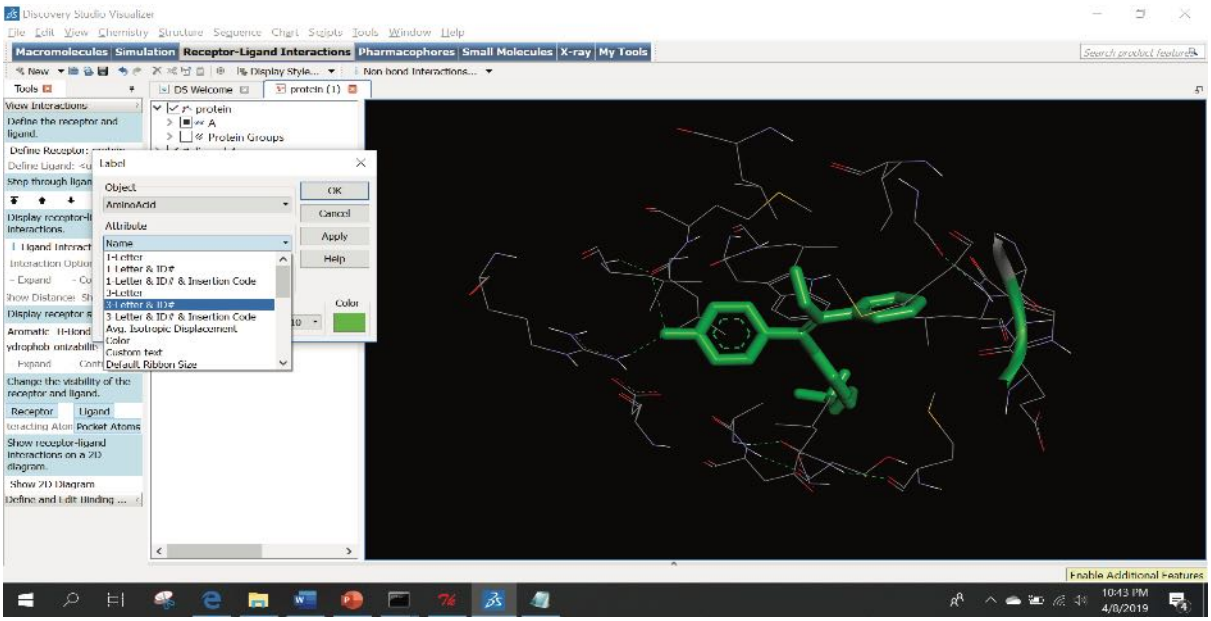

Figure 37. Step to follow for labelling different format of naming attributes.

For Saving the interactions diagrams as Image file, go to **File** and then **Save As**

Change file format form ***Discovery Studio files*** *to* ***Image Files***

Name the new file as ***ligand_1_3D***

Click ***Save*** and then ***OK*** to Save in the working directory

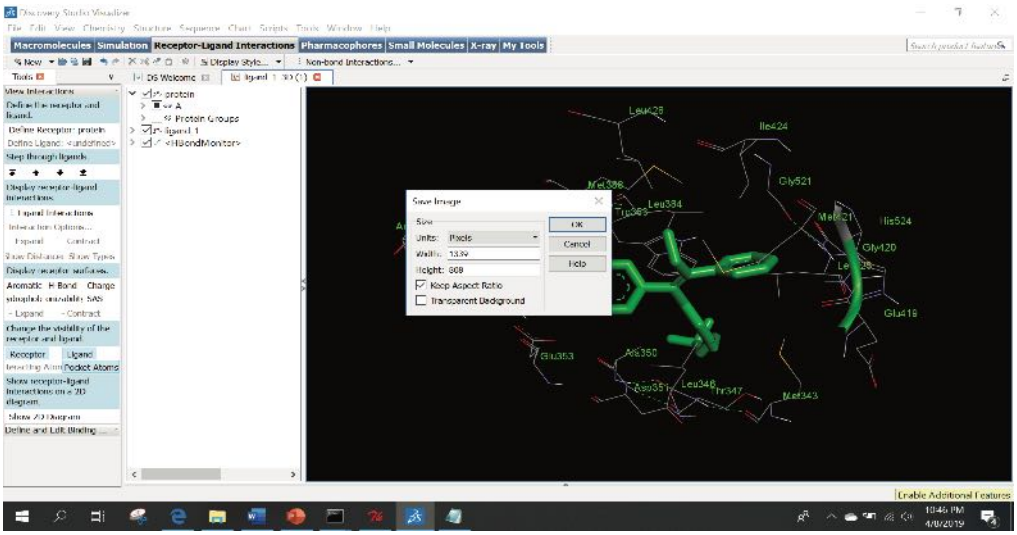

Figure 38. To save the 3D interaction diagram in image format.

Click OK

On left hand side in the ligand interaction window click on show 2D interaction will also generated the 2D image of binding interaction between the ligand and protein

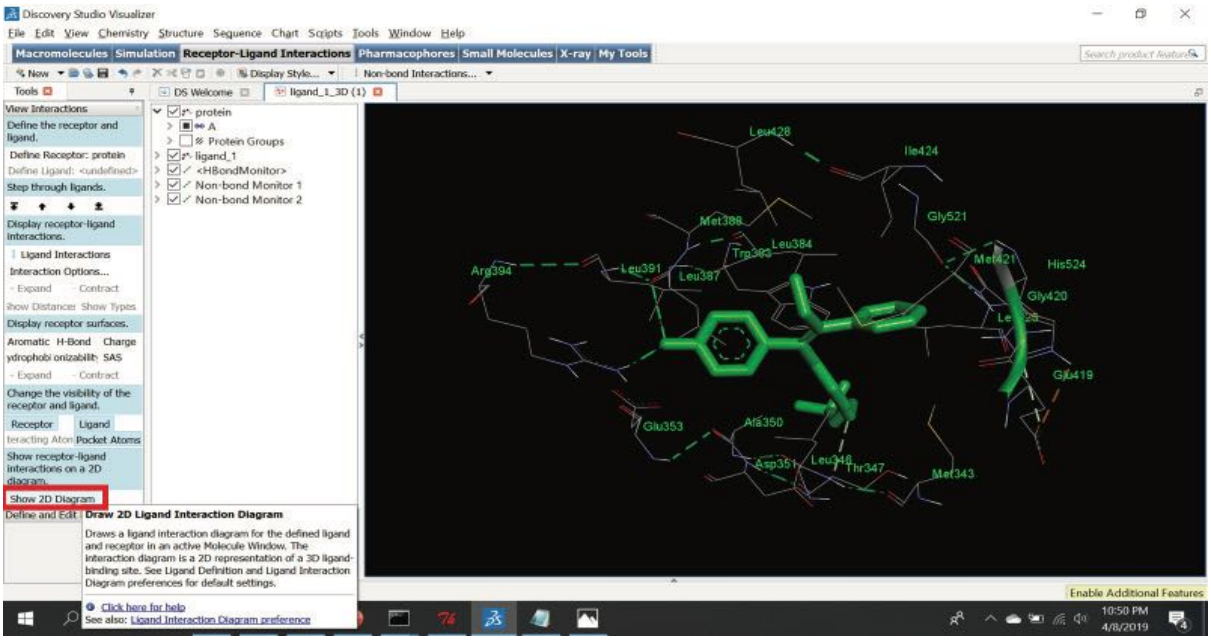

Figure 39. Step to follow for displaying the 2D interaction between protein and bound ligand.

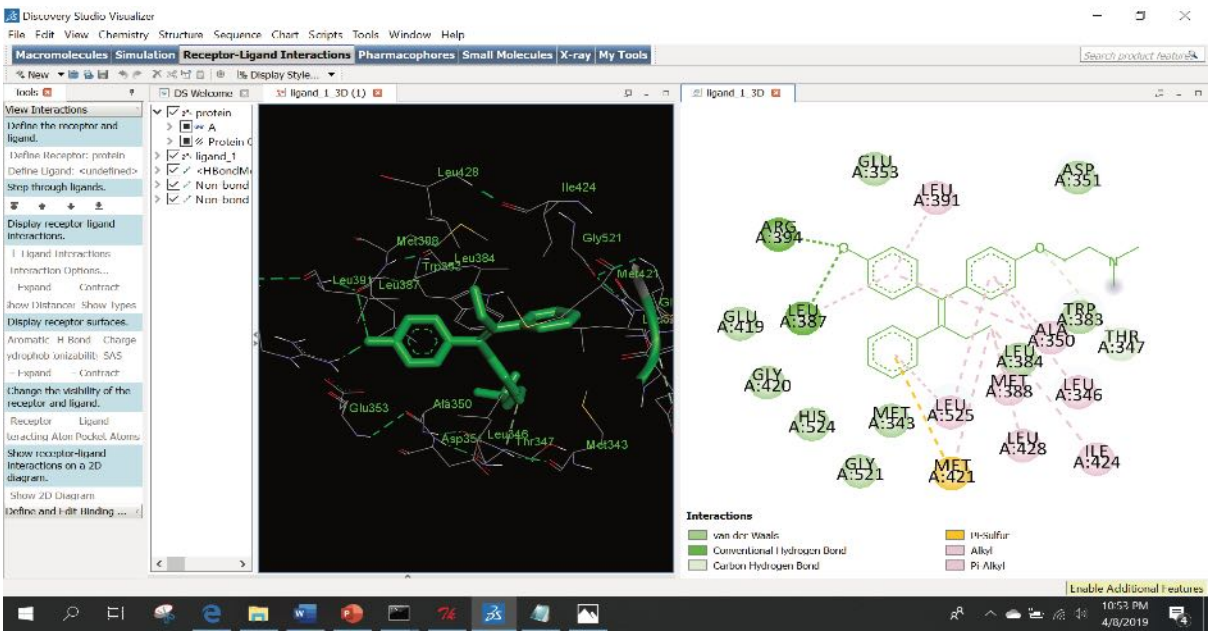

Figure 40. 3D and 2D binding Interactions of dicked ligand with amino acid residue of protein.

Analyse critically the binding interactions within 4OHT re-docked one and co-crystal one and check whether there is any significant difference within these two or not.

If there is no significant difference between these two, decipher that the followed docking protocol seems deem enough for further molecular docking of other compounds.

Conclusion

Molecular docking gives valuable insight into the binding affinity and interactions of ligand within the target protein. In this protocol, starting from no experience in molecular docking studies, an illustrated tutorial was designed to assist any researcher interested in molecular docking studies. Starting from the installation of pre-requisite software, accessing the protein and ligand structure, identifying the active site grid box coordination, generation of required files, execution of docking, selection of best docked pose and binding interaction, all step are illustrated in simple possible way for demonstration and execution process. We hope that this tutorial will make possible molecular docking of small compound within target protein. We are confident that this tutorial will be a convenient start of molecular docking for all researchers who could not have access to commercial and licensed software for performing molecular docking studies.

Conflict of Interest

The authors declare that there is no conflict of interest in this work.

Acknowledgement

The authors would like to thank the Ministry of Higher Education (FRGS, grants no. FRGS/1/2015/ST01/MUSM/03/1) for financial support.

Reference

1. Yuriev E, Agostino M and Ramsland PA. Challenges and advances in computational docking: 2009 in review. J Mol Recognit 2011; 24(2): 149–164.
2. Yuriev E, Holien J and Ramsland PA. Improvements, trends, and new ideas in molecular docking: 2012–2013 in review. J Mol Recognit 2015; 28(10): 581–604.
3. Rizvi SM, Shakil S and Haneef M. A simple click by click protocol to perform docking: AutoDock 4.2 made easy for non-bioinformaticians. J EXCLI 2013; 12–831.
4. Erickson JA, Jalaie M, Robertson DH, et al. Lessons in molecular recognition: The effects of ligand and protein flexibility on molecular docking accuracy. J Med Chem 2004; 47(1): 4555.
5. Morris GM, Huey R and Olson AJ. Using autodock for ligand□receptor dock zing. Curr Protoc Bioinf 2008; 24(1): 814.
6. Goodsell DS, Morris GM and Olson AJ. Automated docking of flexible ligands: Applications of AutoDock. J Mol Recognit 1996; 9(1): 15.
7. Biovia DS. Discovery studio visualizer. San Diego, CA, USA. 2018.
8. Ferreira L, Santos R, Oliva G, et al. Molecular docking and structure-based drug design strategies. Molecules 2015; 20(7): 13384 13421.
9. Coward P, Lee D, Hull MV, et al. 4-Hydroxytamoxifen binds to and deactivates the estrogen-related receptor γ. Proceedings of the National Academy of Sciences 2001; 98(15): 88808884.