3 DAY TRAINING CUM WORKSHOP IN DRUG DISCOVERY TECHNOLOGY

Govt. Madhav Science P.G.College Ujjain (M.P.)

8-10 SEPTEMBER, 2016

HANDS ON TRAINING SESSIONS

COORDINATOR

Dr. Kalpana Virendra Singh
Asstt. Professor Chemistry,
P.G. Department of Chemistry and Pharmaceutical Chemistry,
Govt. Madhav Science P.G.College Ujjain (M.P.)
Phone: 9713389538
eMail: singhkalpana297@gmail.com

The workshop is organized as joint venture of three departments Department of Pharmaceutical Chemistry, Department of Biotechnology & Department of Bioinformatics, Govt. Madhav Science P.G. College, Ujjain, MP, India in association with BioDiscovery Group, India.

BioDiscovery GROUP

BioDiscovery GROUP

WORKSHOP & SEMINAR TOPICS- 2/3/5 Days

DRUG DISCOVERY TECHNOLOGY

- Introduction of Drug Designing
- Science involved in disease target identification
- Virtual screening

Practical application will be done on 2-10 molecules and the software on which DEMONSTRATION & TRAINING will be given

- In-silico generation of ligands by ChemSketch
- Conversion of Mol files to Pdb files by Open Babel
- Protein optimization & Energy Minimization by SPDBV
- Molecular Docking by MGL Tools | Creation of Grid Parameter & Dock Parameter files by AutoDock Software
- Running the Docking Algorithm by Cygwin
- Selection of potent inhibitors on the basis of binding energies and Lipinski's Rule of 5
- Structure Analysis- Protein & ligand complex H-bond interaction by UCSF Chimera
- Prediction of Molecular Properties- Molinspiration
- Prediction of Bioactivity- Molinspiration & ACD iLabs
- Drug Likeness- Mol Soft
- Bioavailability & ADME- ACD iLabs
- Toxicity- OSIRIS Property Explorer & ACD iLabs

BioDiscovery Group
Operations
49, Pushpanjali Puram, Phase 4, Fatehabad Road, Agra- 282001, U.P India
www.biodiscoverygroup.com

WORKFLOW

IN-SILICO GENERATION OF LIGANDS

►CHEMSKETCH

In this software we are going to draw the structure of ligand/molecule i.e. Molecular Editing

- Draw the ligand/structure same as in handout
- AFTER DRAWING COMPLETE STRUCTURE just press F9 to clean the structure.
- Click on **FILE → SAVE AS**. *(GO TO YOUR FOLDER) (eg.1)(Molecule 1 will be saved in folder 1 and so on)*
- Change the save as type **"MDL Molfiles [V2000]"**
- CLICK **SAVE** (no need to give any filename, let it be the same as software has given the filename, eg, noname.)
- Now one file created in your own folder.

CONVERSION OF FILE FORMAT

►OPEN BABEL

This software is used for converting the format of the file. We will convert our **.mol** file to **.pdb** format.

- In the left hand side in the input format please select **mol--MDL MOL format** (from drop down arrow) and similarly on the right hand side in output format select **pdb--protein data bank format**.
- UNDER **INPUT FORMAT** (*In the left hand side)* CLICK "**...**" which is BROWSE button
- Go to your folder eg 1, click on the chemsketch file which you have just drawn and **OPEN**
- UNDER **OUTPUT FORMAT** CLICK "**...**"
- A DEFAULT WINDOW WILL OPEN, NOW TYPE "eg.**1.pdb**" Here we give the name to our file. For structure 1 type **1.pdb**, CLICK **SAVE.** *When doing this step for molecule/structure 2 type **2.pdb** and so on.*
- At last from the middle of the software CLICK **CONVERT**. So the left hand co-ordinates of MDL mol format will be converted to pdb format in right hand
- NOW **ONE** MORE FILE CREATED IN YOUR FOLDER.
- Right click on it and select open with and then select WORDPAD, not NOTEPAD/MS WORD but Wordpad.

PROTEIN OPTIMIZATION

►RCSB- Protein Data Bank

NOW WE ARE GOING TO PREPARE OUR PROTEIN FOR DOCKING.

- TYPE www.rcsb.org
- In the search column type **2fum,** for training purpose we have taken this protein id, i.e. 2fum, so please type 2fum in the search column and search
- See the title Catalytic domain of protein kinase PknB from Mycobacterium tuberculosis in complex with mitoxantrone
- **In complex with** means there is another molecule attached with this PKnB and the name of that molecule is Mitoxantrone.
- LOOK AT RIGHT HAND SIDE, THERE IS DOWNLOAD FILES, CLICK ON DROP DOWN ARROW ....**CLICK DOWNLOAD FILE → PDB File (Text)**
- YOUR PROTEIN IS NOW SAVED.
- Open it, it will open in Wordpad not notepad, not MS word but wordpad
- In the beginning of the file, you must be seeing the info of the protein, at the beginning like title, journal, author etc, when you will scroll down then you will see other info, when you will scroll more you will see ATOM, the first ATOM line would be this ATOM 1 N THR A 3 77.936 -5.559 -21.356 1.00 80.98 N
- These are the residue of protein.
- Now scroll till the last, at the end of file, you must be seeing END then MASTER and above that CONNECT, when you will scroll a little up just before CONNECT you will see MIX.
- These are co-ordinates of molecule mitoxantrone which is attached with this protein.
- For protein residue it was ATOM but for any other molecule which is not part of protein it is HETATM which is hetero atom
- Now we have to do optimization of protein, hence you have to select from the first line of HETATM

- till before END and delete with delete button or backspace button on keyboard
- Now press Control+s and save that file.
- Your last line of protein should look like this-

TER 7874 HIS D 277
END

ENERGY MINIMIZATION

►SPDBV- Swiss-PdbViewer

- Open spdbv
- File→ Open PDB File→ Open your 2fum.pdb file
- A General Communication will come, click OK
- Missing Atoms information will come→ Close that window.
- Now in main window of spdbv, click on Select→ All
- Now, click on Prefs→ Energy Minimization→ default parameters will come, Click OK
- Now click on File→ Save→ Current Layer→ Go to your folder **Molecules** and type 2fum.pdb and save.
- Now open this 2fum.pdb file, it will open in wordpad
- Go to the end of the file, SPDV has put its co-ordinates, select from just before end and delete.
- Your last line should look like this

ATOM 7984 NE2 HIS D 277 69.416 -68.245 -50.981 1.00 84.28
END

- NOW OUR PROTEIN IS READY.
- COPY THIS **2fum** PROTEIN FILE AND PASTE IT INTO ALL 10 FOLDERS.

MOLECULAR DOCKING (*Creation of .gpf- grid parameter file and .dpf dock parameter file)*

►AUTO DOCK 4-

Will come on desktop after installing mgl tools

Open the software AutoDock Tools from desktop(blue color icon)

- **At top on left hand side it is written AutoDock 4.2, double click over it and then minimize** it and again **maximize** it, a small window at top left corner will open, in that select **AutoDock 4.0** and **Dismiss**
- **FILE → READ MOLECULES** → Go to your folder **1**→ SELECT **2fum → OPEN**
- **EDIT → HYDROGEN → ADD→ YES→ → POLAR ONLY → OK**
- **EDIT → CHARGES → ADD KOLLMAN CHARGES**
- **EDIT → ATOMS → ASSIGN AD4 TYPE**(any one of it)
- **FILE → SAVE → WRITE PDBQT**. A WINDOW WILL OPEN
- **BROWSE** → Go to your folder 1and type the file name **2fum.pdbqt → SAVE**
- NOW SELECT FROM **"ATOM"** TO **"END" → ADD**
- Check boxes for **SORT NODES, SAVE TRANSFORM COORDS, Write all.....records → OK**
- **LIGAND → INPUT → OPEN→**Go to your folder **1**→ (CHANGE **FILES OF TYPE**) **ALL FILES →** Select **1.pdb→ OPEN-** A small window appeared giving the information regarding the ligand which we chose, i.e. 1
- **LIGAND → TORSION TREE → DETECT ROOT**
- **LIGAND → TORSION TREE → CHOOSE ROOT**
- **LIGAND → OUTPUT → SAVE AS PDBQT** → GO TO YOUR FOLDER 1 → (WRITE IN **FILE NAME**) **1.pdbqt- SAVE**
- **GRID → MACROMOLECULES → CHOOSE → 2fum → SELECT MOLECULE** → WARNING WILL COME → CLICK **OK** → GO TO YOUR FOLDER 1 → (WRITE IN **FILE NAME**) **2fum.pdbqt → SAVE** → REPLACE WINDOW WILL OPEN → **YES**
- **GRID → SET MAP TYPES →CHOOSE LIGAND** → CLICK ON **1 → SELECT LIGAND**
- Click on **Grid→Grid box→Center→Pick an atom-** and in **Center Grid Box:**
  in **X center:** type **62.467**
  in **Y center:** type **4.815** and in **Z center:** type **-32.164**
- **FILE → CLOSE SAVING CURRENT**
- **GRID → OUTPUT → SAVE GPF**... → GO TO YOUR FOLDER 1 → (WRITE IN **FILE NAME**) **2fum.gpf → SAVE**
- **DOCKING → MACROMOLECULES → SET RIGID FILE NAME →** GO TO YOUR FOLDER 1 → click **2fum.pdbqt → OPEN**

- **DOCKING → LIGAND → CHOOSE....→** CLICK **1 → SELECT LIGAND**. A WINDOW WITH DEFAULT PARAMETERS WILL OPEN → **ACCEPT**
- **DOCKING → SEARCH PARAMETERS → GENETIC ALGORITHM→** A WINDOW WITH DEFAULT PARAMETERS WILL OPEN→ **ACCEPT**
- **DOCKING → OUTPUT → LAMARCKIAN GA** → GO TO YOUR FOLDER 1 → (WRITE IN **FILE NAME**) **2fum.dpf → SAVE**
- NOW YOU SHOULD CLOSE THE AUTO DOCK SOFTWARE WINDOW AND AGAIN OPEN IT FOR DOCKING OF NEXT MOLECULE

RUNING THE DOCKING ALGORITHM

►CYGWIN-I

IN THIS SOFTWARE WE WILL CREAT **GLG** FILE AND **DLG** FILE, (*Run docking algorithm)*, COMMANDS ARE GIVEN BELOW.

Now in cygwin first you have to reach your folder, so for example if your folder 1 is in C drive, then you have one folder by the name of workshop and then Molecules and then you have all 10 folders in it then do the following.

- Open Cygwin terminal
- Type- cd c:/ and press enter
- Then type cd Workshop and press enter
- Then type cd Molecules and press enter
- Then type cd 1 and press enter
- Now you have reached your folder 1 in cygwin terminal, then type
- **./autogrid4.exe –p 2fum.gpf –l 2fum.glg &** and press enter
- Note: In above command -l is minus small L
- JUST WAIT **5 minutes**. THEN....
- **"AUTOGRID SUCCESSFUL COMPLETION"** WILL COME and press enter
- If autogrid successful completion has not come then type this command **tail –f 2fum.glg** and press enter
- If you see autogrid successful completion, close this cygwin window and open a new cygwin window and repeat the above process for molecule 2.
- REPEAT THIS STEP FOR ALL 10 MOLECULES
- Then, for AutoDock, repeat the process of going to your folder and when you reach then type command for autodock
- **./autodock4.exe –p 2fum.dpf –l 2fum.dlg &** and press enter
- Note: In above command -l is minus small L
- NOW WAIT FOR **5 minutes** you will get **"AUTODOCK SUCCESSFUL COMPLETION"**
- If it does not come then type **tail –f 2fum.dlg**
- If autodock successful completion has come then close cygwin window and repeat the process for molecule 2.
- Follow this cygwin procedure for all 10 molecules.

►AFTER CREATING DLG FILES FOR ALL 10 MOLECULES, OPEN first dlg file in WORDPAD and press **Ctrl+F → TYPE RMSD and press 3 times enter → You will reach** *RMSD TABLE.*

- Open a MS EXCEL window or you can note down on page also and make 3 columns**, S. No., Min. Binding Energy, RUN** and WRITE DOWN ACCORDING TO RMSD TABLE OF DLG FILE. The binding energy at the top of the table and the run in which it has come.
- When you do this step for all 10 molecules, make another table and from this 10 molecules select best 5, which has minimum binding energy. MAKE A SEPARATE LIST IN **ASCENDING ORDER** NOT ACCORDING TO MOLECULE NUMBER.
- NOW RUN THESE 2 COMMANDS FOR THESE TOP 5 MOLECULES.

►CYGWIN- II

- Open Cygwin terminal
- Type- cd c:/ and press enter
- Then type- cd Workshop and press enter
- Then type- cd Molecules and press enter
- Then type- cd (FOLDER Number of top 5 molecule which you selected) and press enter
- Now you have reached your folder in cygwin terminal, then type

- **grep '^DOCKED' 2fum.dlg | cut –c9– > 2fum_run.pdbqt** and press enter. Unlike autogrid and autodock when you give this command no successful completion will come just press enter and file will be created.
- Then type **cut –c–66 2fum_run.pdbqt > 2fum_run.pdb** and press enter, file will be created. Close Cygwin window.

►NOW GO TO FOLDER Number (of top 5 molecule which you selected) and you will find that 2 more files are there named **2fum_run.pdbqt** & **2fum_run.pdb**

- Open the **2fum_run.pdb**.
- Go the excel sheet which you have made and see the run in which you have got minimum binding energy. Then in this file scroll down to that run in which you have got minimum binding energy.
- COPY ALL THE **ATOMS** AND PASTE IT INTO THE ORIGINAL PROTEIN (which you used for DOCKING after optimization & Energy Minimization) BEFORE **END**.
- THEN REMOVE ALL THE LINES *(END BRANCH, END POINT, TORSDOF, etc.)* EXCEPT **ATOM**
- SAVE AS IT… Ctrl+ s

H-BOND ANALYSIS

►UCSF CHIMERA

WE USE THIS SOFTWARE FOR VISUALISATION & ANALYSIS OF RESULT.

- OPEN UCSF CHIMERA
- **FILE → OPEN → CHANGE DRIVE → C:/ → WORKSHOP→MOLECULES → 1 → 2fum.pdb → OPEN**. Please select the drive in which you have your molecules.
- After the protein is opened on UCSF Program and you see it, then,
- **SELECT → RESIDUE → UNK OR LIG**.
- **SELECT → ZONE → OK**
- **ACTION → LABEL → RESIDUE → NAME + SPECIFIER**.
- **TOOLS → STRUCTURE ANALYSIS → FIND H-BOND**
- A DEFAULT WINDOW WILL OPEN JUST check box "**colour H-bonds not meeting precise criteria differently**" and "**only find H-bonds with…selected"**
- **SELECT → CLEAR SELECTION.**
- ZOOM IT AND SEARCH FOR **BLUE/ORANGE** COLOUR line which is **H-BOND** If it is THERE or NOT
- IF YES THEN CHECK HOW MANY BONDS ARE THERE
- AND THEY ARE SHOWING THE H-BOND WITH **VAL 95** (*Our active site of protein)* OR NOT.
- IF YES THEN NOTE DOWN THE NO. OF H-BONDS AND MOLECULE NO.THIS IS YOUR FINAL RESULT
- AND PUT THE MOUSE over the line showing **H-BOND**, SHOWING WITH **VAL 95 .**
- PRESS THE **"PRINT SCREEN SYSRQ"** KEY.
- OPEN THE **PAINT** WINDOW PASTE IT & **SAVE** IT ANY WHERE IN YOUR NAMED FOLDER.
- NOW YOU MAKE A TABLE OF **BEST 3 H-BOND** SHOWING RESULT WITH ALL THE DETAILS OF MOLECULES WITH **PHOTOGRAPH.**

►ADMET

The molecules which have shown H-Bond with the active site residue or any other residue of the binding pocket note down those molecules and then run these molecules on the following online ADMET servers. Now these molecules are ready for DRUG LIKENESS SCORE, ADMET studies.

- For MOLECULAR PROPERTIES and DRUG LIKENESS SCORE website is http://molsoft.com/mprop/
  You can draw or import the ligand file from ChemSketch. Draw the molecule as you have drawn in ChemSketch software.
- For TOXICITY and BIOACTIVITY website is http://www.organic-chemistry.org/prog/peo/
  You have to draw the ligand.
- For ADMET STUDIES website is https://ilab.acdlabs.com/iLab2/ you can draw or import the ligand file from ChemSketch.
  In this you can only do 5 tasks at the time and then do it after 24 hours. So first do the Bioavailability, Absorption, LD50, Health Effects, AMES Test.

FOR ALL THESE ONLINE SOFTWARE JAVA HAS TO BE ENABLED WITH PLUGINS AND SHOULD BE UPDATED.