Virtual-Lab

|  |  |  |  |
| --- | --- | --- | --- |
| **Sl.**  **No.** | **Practical title** | **Page no** | **Signature (Internal**  **Supervisor)** |
| **1** | Calculating the distance between the ligand and a particular amino acid [https://vlab.amrita.edu/index.php?sub=3&brch=275](https://vlab.amrita.edu/index.php?sub=3&brch=275&sim=1446&cnt=1)  [&sim=1446&cnt=1](https://vlab.amrita.edu/index.php?sub=3&brch=275&sim=1446&cnt=1) | 07 |  |
| **2** | Secondary structure analysis of a protein using SOPMA .  [https://vlab.amrita.edu/index.php?sub=3&brch=275](https://vlab.amrita.edu/index.php?sub=3&brch=275&sim=1454&cnt=1)  [&sim=1454&cnt=1](https://vlab.amrita.edu/index.php?sub=3&brch=275&sim=1454&cnt=1) | 09 |  |
| **3** | Visualize the secondary structure of a protein  [**https://vlab.amrita.edu/index.php?sub=3&brch=**](https://vlab.amrita.edu/index.php?sub=3&brch=275&sim=1427&cnt=1)[**275&sim=1427&cnt=1**](https://vlab.amrita.edu/index.php?sub=3&brch=275&sim=1427&cnt=1) | 111 |  |
| **4** | Primary Structure Analysis of a Protein Using ProtParam https://vlab.amrita.edu/index.php?sub=3&brch=275  &sim=1455&cnt=1\ | 13 |  |
| **5** | Finding the Active Site Pockets of a given Protein Molecule.\ [https://vlab.amrita.edu/index.php?sub=3&brch=275](https://vlab.amrita.edu/index.php?sub=3&brch=275&sim=1443&cnt=1)  [&sim=1443&cnt=1](https://vlab.amrita.edu/index.php?sub=3&brch=275&sim=1443&cnt=1) | 15 |  |
| 6 | Surface Analysis of a Protein Using CASTp  [**https://vlab.amrita.edu/index.php?sub=3&brch=**](https://vlab.amrita.edu/index.php?sub=3&brch=275&sim=1479&cnt=1)[**275&sim=1479&cnt=1**](https://vlab.amrita.edu/index.php?sub=3&brch=275&sim=1479&cnt=1) | 17 |  |
| 7 | Retrieving details of drug molecules [https://vlab.amrita.edu/index.php?sub=3&brch=275](https://vlab.amrita.edu/index.php?sub=3&brch=275&sim=1502&cnt=1)  [&sim=1502&cnt=1](https://vlab.amrita.edu/index.php?sub=3&brch=275&sim=1502&cnt=1) | 19 |  |
| **8** | Protein-Ligand Interaction  [https://vlab.amrita.edu/index.php?sub=3&brch=275](https://vlab.amrita.edu/index.php?sub=3&brch=275&sim=1495&cnt=1) [&sim=1495&cnt=1](https://vlab.amrita.edu/index.php?sub=3&brch=275&sim=1495&cnt=1) | 21 |  |
| **9** | Absorption and Distribution Property Prediction in Drug Designing Process  [https://vlab.amrita.edu/index.php?sub=3&brch=277](https://vlab.amrita.edu/index.php?sub=3&brch=277&sim=1497&cnt=1) [&sim=1497&cnt=1](https://vlab.amrita.edu/index.php?sub=3&brch=277&sim=1497&cnt=1) | 23 |  |
| **10** | Toxicity prediction of a Molecule. [https://vlab.amrita.edu/index.php?sub=3&brch=277](https://vlab.amrita.edu/index.php?sub=3&brch=277&sim=1494&cnt=1) [&sim=1494&cnt=1](https://vlab.amrita.edu/index.php?sub=3&brch=277&sim=1494&cnt=1) | 25 |  |

Experiment no: 01

# AIM:

Calculating the Distance between the Ligand and a Particular Amino acid.

# OBJECTIVE:

To calculate the distances between the ligand and a particular amino acid of a given protein using PyMol.

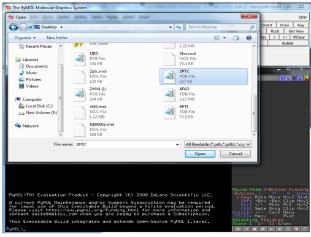
# Theory:

Proteins are important biological macromolecules, usually considered as fundamental units of a cell, play a vitally important role in different cell functions. Protein functions are very specific and mostly depend on the molecule to which it binds. All proteins are made up of long chain of amino acids that fold into 3D shape. Protein structure is classified into primary, secondary, tertiary and quaternary.

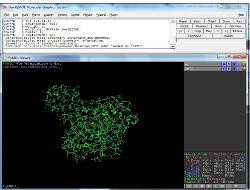
Primary structure is linear arrangment of amino acids in the form of a chain which fold into secondary, tertiary and quaternary structures. Amino acids are organic compounds that contain hydrogen atom, alpha carbon, two functional groups and a side chain R group. Nearly 20 amino acids were found in human body which only vary in their R groups. Amino acids are linked to each other with the help of peptide bond. Peptide bonds are formed when the carboxyl group of one amino acid linked to the amino group of another molecule through a covalent bond. Proteins exhibit their biological functions by interacting with other molecules called Ligands.

# PROCEDURE:

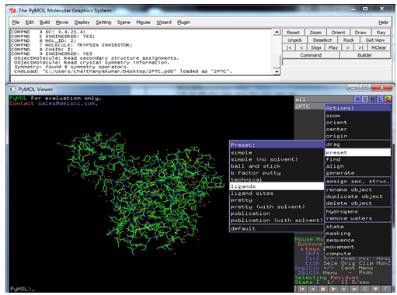
**Step 1: Open the PDB file in pymol**



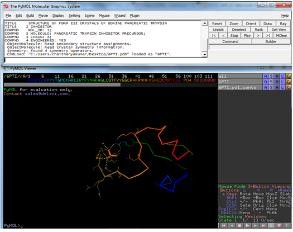
Do the Structure visualization in pymol



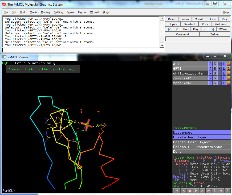
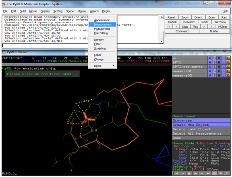
Now View ligands through pymol



The dotted line represents the ligand molecule.



Step 3: Calculate the distance between Ligand and an amino acid using pymol



Conclusion:

The Distance between the Ligand and a Particular Amino acid has been calculated. Primary structure is linear arrangment of amino acids in the form of a chain which fold into secondary, tertiary and quaternary structures. Amino acids are organic compounds that contain hydrogen atom, alpha carbon, two functional groups and a side chain

Experiment no: 02

**Aim:** Secondary structure analysis of a protein using SOPMA

**Objective**: Secondary structure analysis of a protein using SOPMA

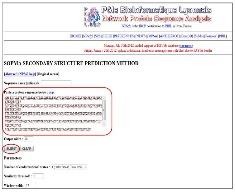
**Theory**: In this exericise one can learn how to analyze the secondary structure of a protein using SOPMA. The structure of a protein has a very important role in its function. The binding of a protein with other molecules is very specific to carry out its function properly. For this reason every protein has a particular structure. Protein structures are classified into primary, secondary, tertiary, and quaternary. The proteins are synthesized as primary sequence and then it fold to form secondary, tertiary and quaternary structure.

# Procedure:

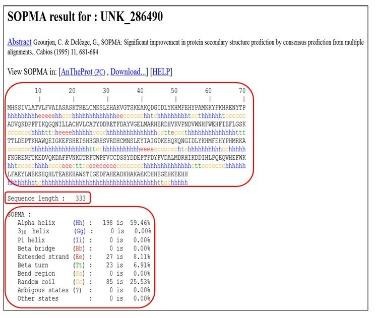
Go to homepage of SOPMA



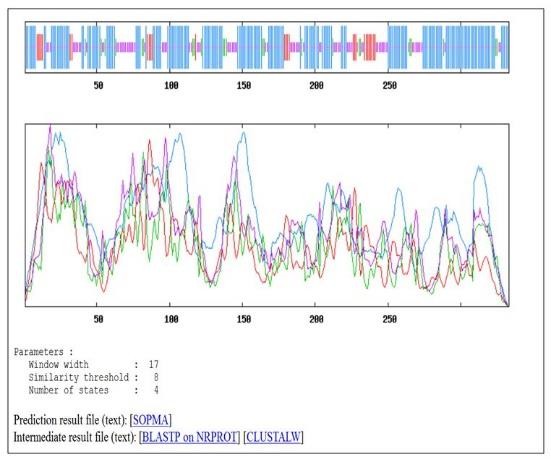
Now paste the computerized protein sequence in the text box provided. By default the output width is 70. It means that in the output shows up to 70 amino acids in each line. We can change the output width if we want. In the parameters there are options like ‘Number of conformational states’, ‘Similarity threshold’ and ‘Window width’. The user can select ‘Number of conformational states’ as either ‘(3Helix, Sheet, Coil)’ or ‘4(Helix, Sheet, Turn, Coil). The former predicts the percentage of helix, sheet and coil structure while the latter predicts percentage of helix, coil, turn and sheet.



The following figures shows the SOPMA results.



Since the output width we set as 70, here it shows 70 amino acids and corresponding predicted structures in each line. The sequence length is also displayed in the output (333 amino acids in this case). The percentage of each structure is also listed in this page. For example, for Alpha helix it is 59.46%.



There are two graphs shown in the result page of SOPMA. First one is to visualize the prediction. The second contains score curves for all predicted states.It also shows the parameters such as window width, number of states etc. that are used for the prediction. It provides a link on prediction result file which gives the result in a text format. There are links to find the intermediate result files also.

Conclusion: The Secondary structure analysis of a protein using SOPMA is done successfully. . The structure of a protein has a very important role in its function. The binding of a protein with other molecules is very specific to carry out its function properly. For this reason every protein has a particular structure. Protein structures are classified into primary, secondary, tertiary, and quaternary.

Experiment no: 03

Aim: Visualizing the Secondary Structure of a Protein Objective: To visualize the secondary structure of a protein.

Theory: Proteins are important biological macromolecules, usually considered as fundamental units of a cell which play a vitally important role in different cell functions. The function of a protein is very specific and dependent on the molecule to which it binds. The protein structure is classified into primary, secondary, tertiary and quaternary. These molecules, form a linear chain of amino acids initially, and then fold into secondary, tertiary and quaternary structures. The different secondary structures of a protein are alpha helices, beta pleated sheets and loops.

During the due course of evolution, some regions of the protein remain conserved which are regarded as motifs, play a key role in determining the function of that particular protein. In simple terms, researchers say that the molecules having similar structure will have a similar function.

# Procedure:

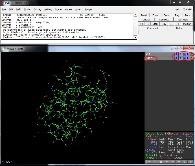
I**nstallation and the links to download the software or the data can be obtained through the simulator tab**



File →open →choose the PDB file where it is actually located. User can also load the file through command line using the command load <file path>

E.g., “load C: Downloads1UBQ.pdb”

By default, the loaded PDB file structure will be shown with line representation in PyMol

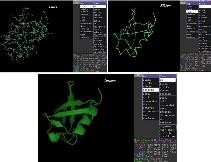


We can see two rows on the right side of Figure 8, labeled as ‘all’ and ‘1UBQ’ where we can change individual appearance of a molecule as well as all molecules together. If there are multiple molecules, user can also select every molecule or object at same time.

User can also see A,S,H,L and C labels in the two rows as different columns, these alphabets implies, Action , Show , Hide , Label and Color respectively.



Representing protein structure in lines is not interpretable, to view or highlight secondary structures of a protein. There are other representations to view the model as stick, ribbon, and cartoon.



Conclusion: Visualizing the Secondary Structure of a Protein done successfully. . The function of a protein is very specific and dependent on the molecule to which it binds. The protein structure is classified into primary, secondary, tertiary and quaternary. These molecules, form a linear chain of amino acids initially, and then fold into secondary, tertiary and quaternary structures.

Experiment no: 04

Aim: Primary Structure Analysis of a Protein Using ProtParam Objective:

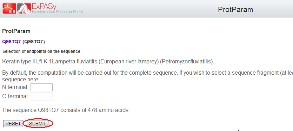
* To compute the various physical and chemical parameters of a protein.
  + To perform primary structure analysis of proteins.
  + To introduce a protein analysis software that is available through the ExPASy server.

Theory: P**roteins** are one of the important fundamental units of all living cells. Proteins have a wide range of functions within all the living beings. Some of the important functions such as DNA replication, catalysis of metabolic reactions, transportation of molecules from one location to another etc. are performed with the help of proteins.

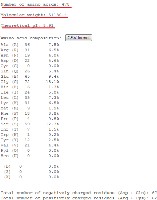
Procedure: **Analyze the protein using ProtParam**

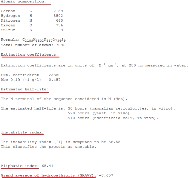


If you provide the accession number of a Swiss-Prot/TrEMBL entry, you will be prompted with an intermediary page that allows you to select the portion of the sequence on which you would like to perform the analysis.



The result page is as shown below:





This experiment uses the Protparam tool, available through the ExPASy server: SIB bioinformatics resource portal.

Conclusion: Primary Structure Analysis of a Protein Using ProtParam done successfully. Some of the important functions such as DNA replication, catalysis of metabolic reactions, transportation of molecules from one location to another etc. are performed with the help of proteins.

Experiment no: 05

Aim: Finding the Active Site Pockets of a given Protein Molecule.

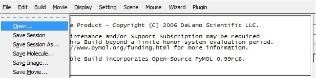
Objective: To view the active site (where the site of a protein binds to ligand molecule) pockets of the protein

Theory: Protein molecules are the fundamental units of all living cells. These macromolecules have a vital role in various cellular functions. Each protein has specific function in our body. The structure of the protein has a very important role in its function. The binding of a protein with other molecules is very specific to carry out its function properly.

Procedure: **Working of PyMol and Loading the .pdb file**

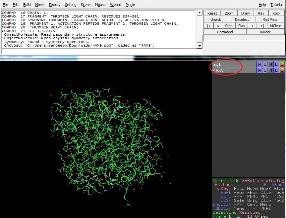
# Installation instructions and the links to download the software or the data can be obtained through the simulator tab

Load the downloaded PDB file into PyMol



File → open→ choose the PDB file where it is actually located. User can also load the file through command line using the command load <file path> E.g., “load C: Downloads1UBQ.pdb”

By default, the loaded PDB file structure will be shown with line representation in PyMol



Screenshot of PyMol with a loaded molecule

We can see two rows on the right side of Figure 8, labeled as ‘all’ and ‘4AX9’ where we can change individual appearance of a molecule as well as all molecules together.

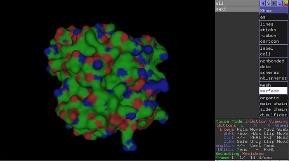
User can also see A,S,H,L and C labels in the two rows as different columns, these alphabets implies, Action , Show , Hide , Label and Color respectively. We can use these GUI elements or the user can give commands to get performed actions (Figure 3).



A protein named ‘thrombin’ which forms a complex with “Napsagatran” ligand molecule, has been downloaded from the protein data bank. Through the detailed annotation of the cited paper or journal article in protein data bank, it has been reported that thrombin has different amino acid residues which act as active sites like Serine (ser 195th residue), Glycine( Gly 219th and 216th residue), Aspartate(Asp 189th residue), Tryptophan ( Trp 215th residue) and Tyrosine (Tyr 60th residue). These residues bind to the ligand molecule “Napsagatran” (N5N).

To view these active sites, hide all the objects loaded into PyMol by using the command “hide”.

Represent entire protein with surface representation, setting with a 50% transparency. Select the object protein molecule, show  surface turns the entire protein molecule into surface representation. Alternatively use the command “show surface, 4AX9”.



User can set up the Transparency by choosing settings in the standard menu bar , thus choose submenu transparency , surface with 50% , alternatively by the command “set transparency 0.5”



The protein databank will provide the information about the ligand molecule bound to the protein. User can know the ligand molecule name. Here the ligand molecule name is “N5N”



Conclusion: Finding the Active Site Pockets of a given Protein Molecule done successfully, . Each protein has specific function in our body. The structure of the protein has a very important role in its function. The binding of a protein with other molecules is very specific to carry out its function properly.

Experiment no: 06

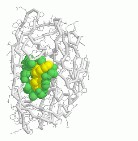
Aim: Surface Analysis of a Protein Using CASTp

Objective: To get a detailed and complete quantitative characterization of surface pockets and interior voids of proteins using CASTp (Computed Atlas of Surface Topography of proteins).

Theory: **Proteins** are one of the important fundamental units of all living cells. Proteins have a wide range of functions within all the living beings. Some of the important functions such as DNA replication, catalysis of metabolic reactions, transportation of molecules from one location to another etc. are performed with the help of proteins.

# CASTp

The CASTp (**C**omputed **A**tlas of **S**urface **T**opography of **p**roteins) web server is an online tool that locates, measure and characterizes the pockets on the protein surfaces and the voids in the interior of proteins.



Procedure :

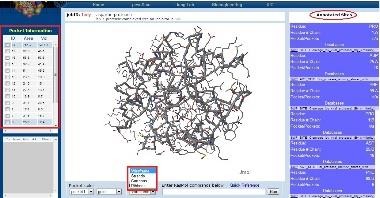
A graphic user interface allows users to request a CASTp calculation. The user can provide the input in either of the two ways. One can either upload the structure of the protein or can type the four letter PDB code of a protein structure if it is available from the Protein Data Bank. In this case the CAST server will fetch that structure.

# Analysis of the protein using CASTp

Home page of CASTP is shown in Figure 1.Give the protein ID (PDB ID) in the box provided and clicks on the ‘search’ button (figure 1).

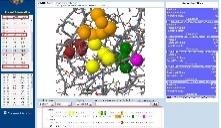


# Click on the search option to get the result



**Pocket information:**

By clicking each pocket from the pocket information (which is shown in left of the window), the pocket residues would be represent in different colors based on their ids. It will shows the pocket information like area, volume, and the aminoacid sequence which is highlighted by the color regarding to the pocket color.



The result page gives different options interactively visualize the surface pocket. This is a very good method to explore the surface features of a protein. For example, each time you can check and see if a specific residue is in a pocket and what that pocket looks like.

Conclusion: Surface Analysis of a Protein Using CASTp done successfully. So basically The CASTp (**C**omputed **A**tlas of **S**urface **T**opography of **p**roteins) web server is an online tool that locates, measure and characterizes the pockets on the protein surfaces and the voids in the interior of proteins.

Experiment no: 07

Aim: Retrieving details of drug molecules

Objective: To retrieve more information about the drug molecules, drug targets, enzymes and pathways related to drugs.

Theory: Drug is a chemical molecule that can interact and bind to a target and control the function biological receptors to control the disease. Proteins are the important biological receptors which interact with the other biological molecules to maintain various cellular functions. Drugs have a specific activity that goes and binds to the target and can activate/inhibit the reactions in a proper manner

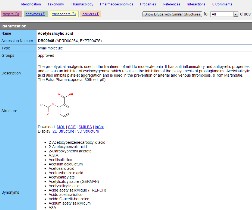
Procedure: Go to Drug bank home page and type the query with suitable name and suitable information in the query box provided. By clicking search, it will redirects to the drug with full details as we mentioned in theory.



**Step 2**: For example give the query name as “ASPIRIN” and click on search. Hits will appear related to the query as shown below.



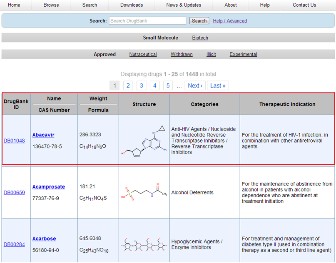
**Step 3**: Each hit gives an account of the drug activity and its structure along with their drug targets, enzymes, transporters and carriers followed by the drug details as explained in theory.



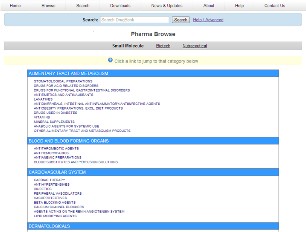
# Using different types of browsing options in Drug bank.

Select “Drug browse” from the browse option in the home page. This will allows one to search the drugs by its name. Each drug will have the characteristics like specific ID, Name, Molecular

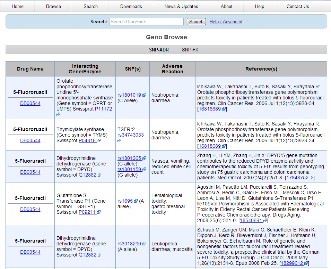
weight, Structure of the drug molecule, In which categeory it belongs, and the therapeutic indicaton of the drug molecule. The screen shot is shown below.



Select "Pharma browse" from the browse option will allow one to search the drugs by their drug category like small molecule, biotech and nutraceutical. User can select the drugs depending on their functional activities.



Geno browse from the drug bank allows the user to get the information about the Genes and SNPs related to the drug molecules. It provides the information about the Drug Name, Drug interacting enzymes, SNPs, Adverse reaction and the references related to the Drugs.



Conclusion: Retrieving details of drug molecules done successfully. Drugs have a specific activity that goes and binds to the target and can activate/inhibit the reactions in a proper manner

Experiment no: 08

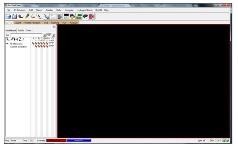
Aim: Protein-Ligand Interaction

Objective: To find the interaction between the protein and a ligand molecule by performing docking studies.

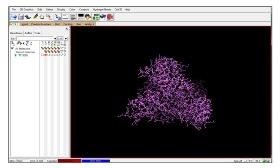
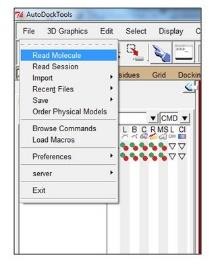
Theory: A molecule is a small chemical element that is made up of two or more atoms held together by chemical bonds. A molecule can be composed of either single kind of element (e.g. H2) or different kinds of elements (e.g. CO2). Molecules can be found in both living things and non living things. A drug is a small molecule that can interact, bind and control the function of biological receptors that helps to cure a disease. Receptors are proteins that interact with other biological molecules to maintain various cellular functions in body. Enzymes, hormone receptors, cell signaling receptors, neurotransmitter receptors etc. are some important receptors in our body.

Procedure: For more information about the software and other pre-requisites refer simulation tab.

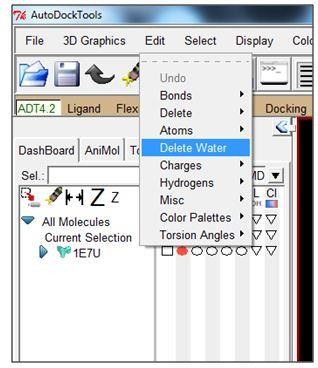
1. Open the Autodock software by clicking on Autodock icon from your desktop.



1. Read the downloaded PDB molecule 1E7U in the work space panel by clicking on the tab “File” and then select “Read molecules”.



1. PDB files can have errors such as missing atoms, chain breaks, water molecules etc. which is needed to be corrected. Select all water molecules which obstruct the accuracy of docking procedure.
2. Click on the “Edit” tab and select “Delete Water” to delete the water molecules from the receptor molecule.



Conclusion: Protein-Ligand Interaction studied successfully. A molecule is a small chemical element that is made up of two or more atoms held together by chemical bonds. A molecule can be composed of either single kind of element (e.g. H2) or different kinds of elements (e.g. CO2). Molecules can be found in both living things and non living things.

Experiment no: 09

Aim: Absorption and Distribution Property Prediction in Drug Designing Process

Objective: To predict the absorption and distribution properties of small chemical drug like compounds.

Theory: A molecule is formed when two or more atoms are combined together chemically (e.g. H2). A compound is said to be a molecule when it contains at least two different elements (e.g. H2O). All compounds can be called as molecules but all molecules are not compounds. In terms of pharmacology or in biochemistry, a small molecule is an organic compound which has low molecular weight that may act as a substrate or inhibitor. In medical field, the term is restricted to the molecule that binds to a biopolymer and act as an effector. Most of the small molecules are drug molecules.

Procedure: **For information on how to get the data, to access PreADMET with registration instructions, go to simulator tab**

Click on the link ADME Prediction in the home page, when it redirects to login page.



Login into the server to get redirected to the page for ADME prediction



One can draw the chemical structure of the drug molecule using the tool bar in the white window as shown in the above screenshot



One can also import the structure of a molecule into the tool for predicting ADME. The structure of the molecule should be loaded in “mol” format.



Once the molecule is open and loaded, click on the calculate button seen at bottom right side of the window.

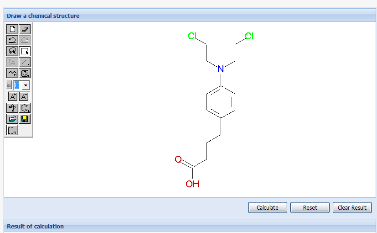


Figure 5: Screenshot of the uploaded molecule

Conclusion: Absorption and Distribution Property Prediction in Drug Designing Process

Studied successfully. A molecule is formed when two or more atoms are combined together chemically (e.g. H2). A compound is said to be a molecule when it contains at least two different elements (e.g. H2O). All compounds can be called as molecules but all molecules are not compounds

Experiment no: 10

Aim: Toxicity prediction of a Molecule

Objective: To predict the toxicity of a potential lead molecule.

Theory: Molecule is formed when two or more atoms combine together chemically (ex: H2). A compound is said to be a molecule when it contains at least two different elements (ex: H2O). All compounds can be called as molecule but all molecules are not compounds. In terms of pharmacology or in biochemistry, a small molecule is an organic compound which has low molecular weight and may act as a substrate or inhibitor. In medical field, the term is restricted to the molecule that binds to a biopolymer and act as an effector. Most of the small molecules are drug molecules.Drug molecules are potential lead molecules which act as therapeutic agents and gives beneficiary effects. To come up with single potential lead molecule it takes 12 -16 years.

Besides the beneficiary aspects, there may be adverse effects also when using drug/potential lead molecules. It has been known that, most well known drugs are poisonous substances. All useful drugs produce unwanted effects due to complex nature of human body. Some drugs are more adverse and can produce dangerous effects. So, toxicity is more important measurement during the synthesis of a molecule. One knows that it is difficult to synthesize a potential lead molecule in a shorter time period by undergoing all types of tests.

Procedure

Click on the link Toxicity Prediction in the home page, when it redirects to login page.



Figure 2 shows the interface for Toxicity prediction.



Figure 2: Screenshot of interface for toxicity prediction.

One can draw the chemical structure of the drug molecule using the tool bar in the white window shown in the above screenshot (Figure 2).



Figure 3: Screenshot of toolbar to draw the compound strucutre

One can also import the structure of a molecule into the tool for predicting toxicity; the structure should be loaded in “mol” format.



Figure 4: Screenshot to import the molecule

Once the molecule is open and loaded, click on the calculate button seen at bottom right side of the window.



Figure 5: Screenshot of the uploaded molecule

The results of the particular molecule will be shown as in Figure 6.



Figure 6: Screenshot showing the results of toxicity prediction

# Interpretation of results

PreADMET predicts mutagenicity and carcinogenicity of compounds. It conducts AMES test , which is a method to test mutagenicity. The test uses several strains of bacterium *Salmonella typhimurium* that contains several mutations in genes involved in histidine synthesis.

PreADMET predicts the toxicity to TA98, TA100, TA1535 strains, which are often used and result is considered both with the metabolite and without it. The result of prediction says whether it is positive or negative.

PreADMET also predicts carcinogenicity results from the model of mice or rat, which is built from the data of NTP (National Toxicology Program). Carcinogenicity test results can also be used as a classifier since result would be either positive or negative.

Conclusion: Prediction of the toxicity of a potential lead molecule done successfully. Molecule is formed when two or more atoms combine together chemically. A compound is said to be a molecule when it contains at least two different elements . All compounds can be called as molecule but all molecules are not compounds.