A HANDBOOK ON PROTEIN-LIGAND DOCKING TOOL: AUTODOCK4

LOKESH RAVI, KANNABIRAN K*

Department of Bio-Medical Sciences, School of Biosciences and Technology, VIT University, Vellore - 632 014, Tamil Nadu, India.
Email: kkb@vit.ac.in

INTRODUCTION

AutoDock is an automated suite of protein-ligand docking tools. AutoDock tools are abbreviated as ADT. It is designed to predict the protein interactions with small molecules such as drug molecule and substrate. The application of this tool is immense, ranging from structure-based drug design, lead molecule optimisation, protein-ligand docking, protein-protein docking, analysis and validation of mechanism of action of drug molecules, etc., AutoDock has two versions, namely, AutoDock4 and AutoDock Vina. In this tutorial, the focus is on AutoDock4.

AutoDock4 analyzes the interactions of ligand molecules at the specified target site of the protein. The users can define this specific target sites with the use of GridBox. AutoDock4 has two key programs to be executed, i.e., Autogrid4 and AutoDock4. Autogrid4 prepares a grid map of the amino acids presents within the GridBox defined by the user. AutoDock4 then analyzes the interactions of those amino acids with the ligand molecule.

AutoDock4 has options to set the protein molecule to be rigid or flexible. While, the ligand molecule is automatically annotated for its rotatable and non-rotatable bonds and determine the flexibility. In this tutorial, the docking procedure is displayed for rigid macromolecule. Rigid macromolecule docking is faster than that of the flexible macromolecule docking.

AutoDock4 predicts the free binding energy with a scoring function based on the AMBER force field and linear regression analysis, additionally by referring to a large set of library data of known protein-ligand interactions with their inhibition constants that were used in AutoDock3. It is found that the standard error in free binding energy in AutoDock4 is approximately 2.5 kcal/mol.

In this tutorial, a basic protocol to use AutoDock4 to analyze interactions of selected protein and ligand molecule is discussed. This protocol is presented to aid students and research scholars those who are interested in using AutoDock4 for learning and or research purpose. This protocol is considered to be adequate and effective by the authors, to use AutoDock4 for protein-ligand docking analysis.

The discussed protocol can be used to study the interactions of selected ligand molecule (drug molecule) with chosen protein targets. Users, however, have to do a little preparative work on the protein molecule pdb file. Downloaded 3D structure of proteins from RCSB website has to be edited before docking in AutoDock4. Although a small discussion on the preparation of protein molecules is presented in this tutorial, users are advised to do further reading in the preparation of protein structures.

INSTALLING AUTODOCK

Download the installation files "autodocksuite-4.2.5.1-i86Windows.exe and mgltools_win32_1.5.6_Setup.exe" from AutoDock website (http://autodock.scripps.edu/). Run "mgltools_win32_1.5.6_Setup.exe" file and leave al the settings default and the installation folder should not be altered. Finish the installation with the default settings. Once the installation is finished, "AutoDockTools-1.5.6.exe" will be executed by default. Close the application and start the "AutoDockTools-1.5.6.exe" once again. Once the application is opened for the second time, a new folder ".mgltools" will be created automatically in the users folder (C:\Users\Guest\.mgltools).

To successfully run the AutoDock software, the protein and ligand structures in their ".pdb" format have to be present within the ".mgltools" folder. The ".mgltools" folder has a subfolder "1.5.6." Create a new folder within "1.5.6" folder, to use as a working folder (e.g.: "WorkFolder") as given in Fig. 1A. The ".pdb" files of the protein and ligand to be studied must be pasted within this working fold er. For easier use of the tool, the default startup directory for AutoDock has to be changed to a working folder.

To setup the default folder:

File » Preferences » Set

↪ **Startup Directory » "C:\Users\Guest\.mgltools\1.5.6\ WorkFolder"**

↪ **Make Default » Dismiss**

Next, run the "autodocksuite-4.2.5.1-i86Windows.exe" file and select a custom folder for this. At the end of the installation, there will be two execution files created in the selected custom folder (i.e., AutoDock4. exe & autogrid4.exe). These two files have to be moved into "1.5.6" folder within the ".mgltools" folder [C:\Users\Guest\.mgltools\1.5.6] highlighted in Fig. 1B. At this stage, the software is ready to be executed.

PREPARING THE PROTEIN MOLECULE

Protein molecules can be downloaded from Protein Data Bank website (www.rcsb.org). The downloaded protein structure in their ".pdb" format has to be edited to remove the non-aminoacid residues, such as water molecules, ions, ligands that are in the complex. These can be removed using either PyMol software or WordPad. This has to be done, since, these molecules will interfere with the interaction between the target molecule and protein in AutoDock.

Once the ".pdb" file is downloaded, right-click the file and open it with WordPad. Scroll down to the end of the document and delete all the lines that begin with "HETATM" and "CONNECT," these are the atoms that do not belong to amino acids and their interaction with the amino acids, respectively. The last line of the document should contain a like begins with "TER" (e.g.: "TER 8910 SER B 557") this indicates the end of the protein chain. Once this is done, save the document, now it is ready for AutoDock analysis.
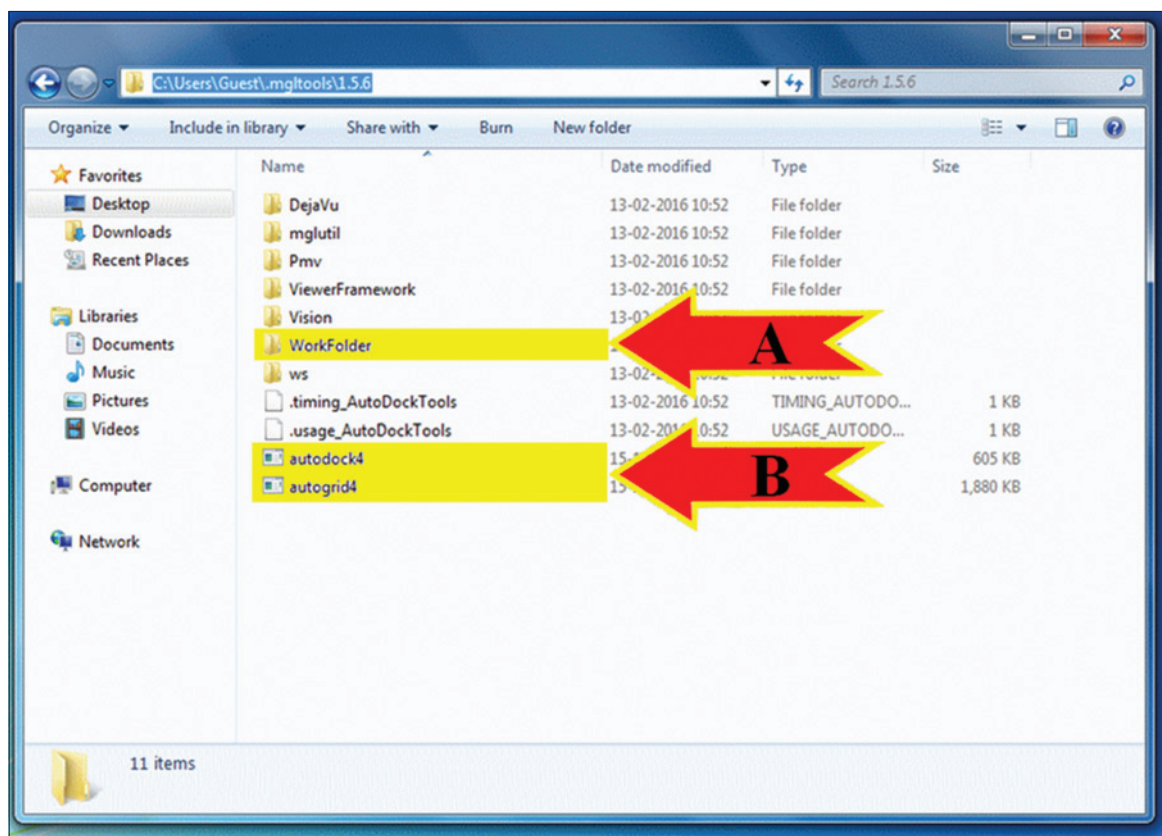
Fig. 1: Contents of the folder C:\Ussers\Guest\.mgltools\1.5.6

EXECUTING AUTODOCK

AutoDock software calculates and predicts the interaction between the ligand molecule and protein molecule based on predefined parameters. To be precise, the interactions between the molecules will be calculated at a user specified region in the protein. This region can be defined by users, using the Grip map option. Ultimately, the software predicts the interaction and binding energy of the ligand molecule and the amino acids present within the GridBox only. Thus setting, the GridBox at the binding site or active site or other essential regions of the protein is important. Before executing the AutoDock, the ".pdb" files of the protein and ligand have to be moved into the "WorkFolder." Fig. 2 gives a graphical representation of contents of "WorkFolder" before and after executing AutoDock. The "WorkFolder" should contain only the ".pdb" files of protein and ligand at the beginning.

Analysis in AutoDock can be divided into following categories, (1) Initializing molecules; (2) Running AutoGrid; (3) Running AutoDock; (4) Analyzing Interaction energy.

Initializing molecules

Initializing the molecule mainly includes addition of hydrogen atoms and addition of Kolman charge to the protein molecule. While for the ligand molecule, addition of Gasteiger charge, identifying aromatic carbons, detecting rotatable bonds, and setting TORSDOF value. The protein has to be initialized manually, while the ligand is automatically initialized when opened in the tool.

Once the protein molecule is opened, it is important to change the view of the protein. It will be in "line view" by default. Enable the "Ribbon view" and "Surface view" as given in Fig. 3A. It is essential that the protein is in "Ribbon view," failing to do so, might result in errors in the later steps. The surface view makes it user-friendly to set the GridBox in the later steps. Fig. 3 highlights all the important menu options that will be used in the given protocol.

To initialize molecules

File » Read Molecule

   ↳ **Select Protein File** (".pdb" file)

     Set the protein view to Ribbon and Surface view   (Fig. 3A)

Edit » Hydrogens » Add

   ↳ **Polar Only » OK**

Edit » Hydrogens

   ↳ **Merge Non Polar » Continue**   *[Click 'CONTINUE' for the Warning]*

Edit » Charges » Add Kolman Charges

   ↳ **OK**

File » Save » Write PDB

   ↳ **Sort Nodes** (Check) **» OK »** (Overwrite) **YES**

Ligand » Input » Open

   ↳ (Change view to all files) **Select Ligand File** (".pdb" file) **» OK**

Ligand » Output » Save as PDBQT

   ↳ **OK**

Executing AutoGrid4

AutoGrid has to be executed, to define the region/area in the protein to be analyzed for the interaction with the ligand molecule. In general,
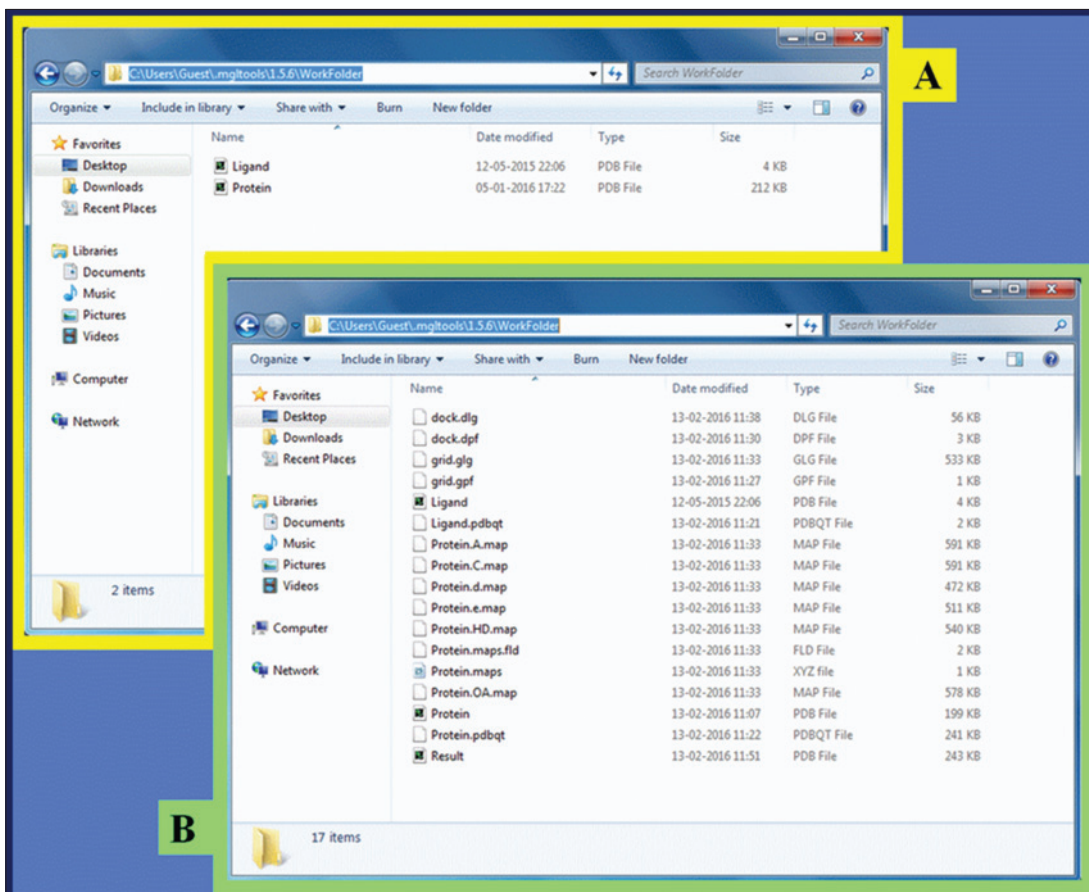
Fig. 2: Contents of WorkFolder before and after executing AutoDock (A: Before executing AutoDock, B: After executing AutoDock)
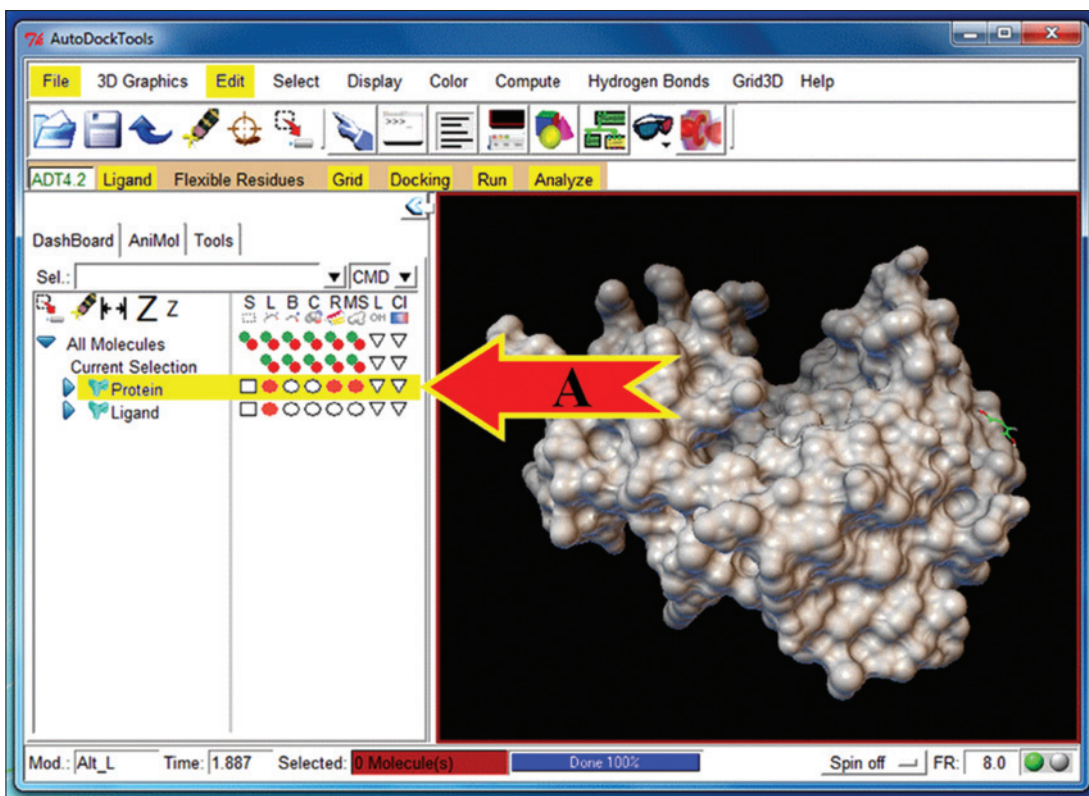
Fig. 3: Key options to be used in the AutoDock window

the region of interaction could be identified using prediction tools such as Q-Site finder and MetaPocket to identify the binding pockets on the surface of the proteins. Then, the GridBox is set in AutoDock to cover the identified binding sites.

AutoDock only analyzes the interactions of ligand molecule and the amino acids that are present within the GridBox. So setting up, the GridBox is a crucial step.

The grid options menu has two fields as shown in Fig. 4. The size of the GridBox can be increased or decreased using the number of points in X/Y/Z dimension scroll option (Fig. 4A). The position of the GridBox can be adjusted to cover the binding site or binding residues, using the Center GridBox field, that moves the GridBox in the X/Y/Z axis (Fig. 4B).

To Run AutoGrid

Grid » Macromole » Choose

    ↳ **Click (Protein) » Select Molecule » OK » SAVE**

Grid » Set Map types » Choose Ligand

    ↳ **Click (Ligand) » Select Ligand**

Grid » GridBox

    ↳ **SET the BOX**

        ↳ **File » Close Saving Current**

Grid » Output » Save GPF

    ↳ **File name "grid.gpf" » SAVE**

Run » Run AutoGrid   (Fig. 5A)

    ↳ **Select Program Pathname: 'autogrid4.exe'**

    ↳ **Select Parameter Filename: 'grid.gpf'**

    ↳ **AutoFill for Log Filename: 'grid.glg'**

        ↳ **LAUNCH**

Executing AutoDock4
Once AutoGrid is successfully completed, AutoDock can be executed. AutoDock calculates the interactions of the amino acid within the GridBox and ligand molecule.

To Run AutoDock

Docking » Macromolecule » Set Rigid filename

    ↳ **Select 'Protein.pdbqt' » Open**

Docking » Ligand » Choose

    ↳ **Click 'Ligand' » Select Ligand**

        ↳ **Accept**

Docking » Search Parameters » Genetic Algorithm

    ↳ **Accept**

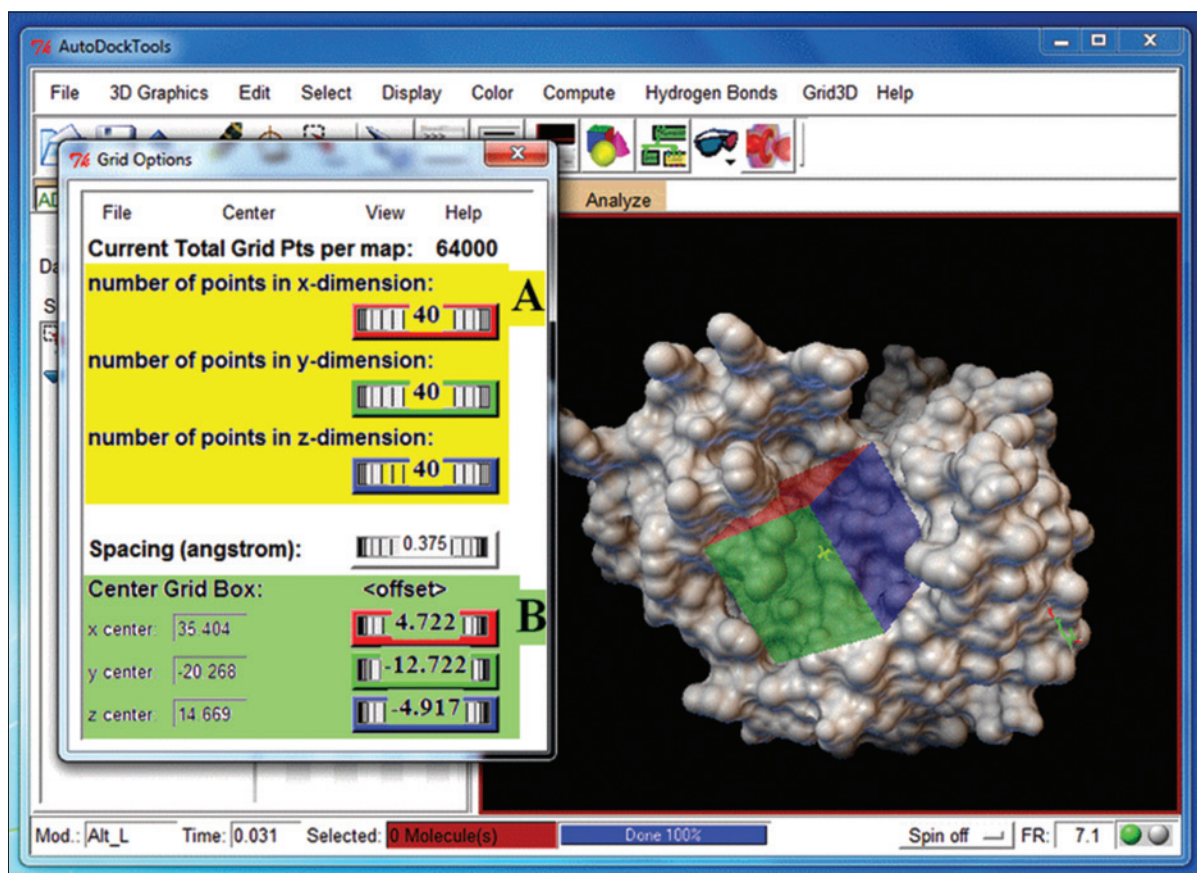**Docking » Output » Lamarckian GA**

Fig. 4: Grid options in AutoDock. (A: To adjust the size of the GridBox, B: To adjust the position of the GridBox)
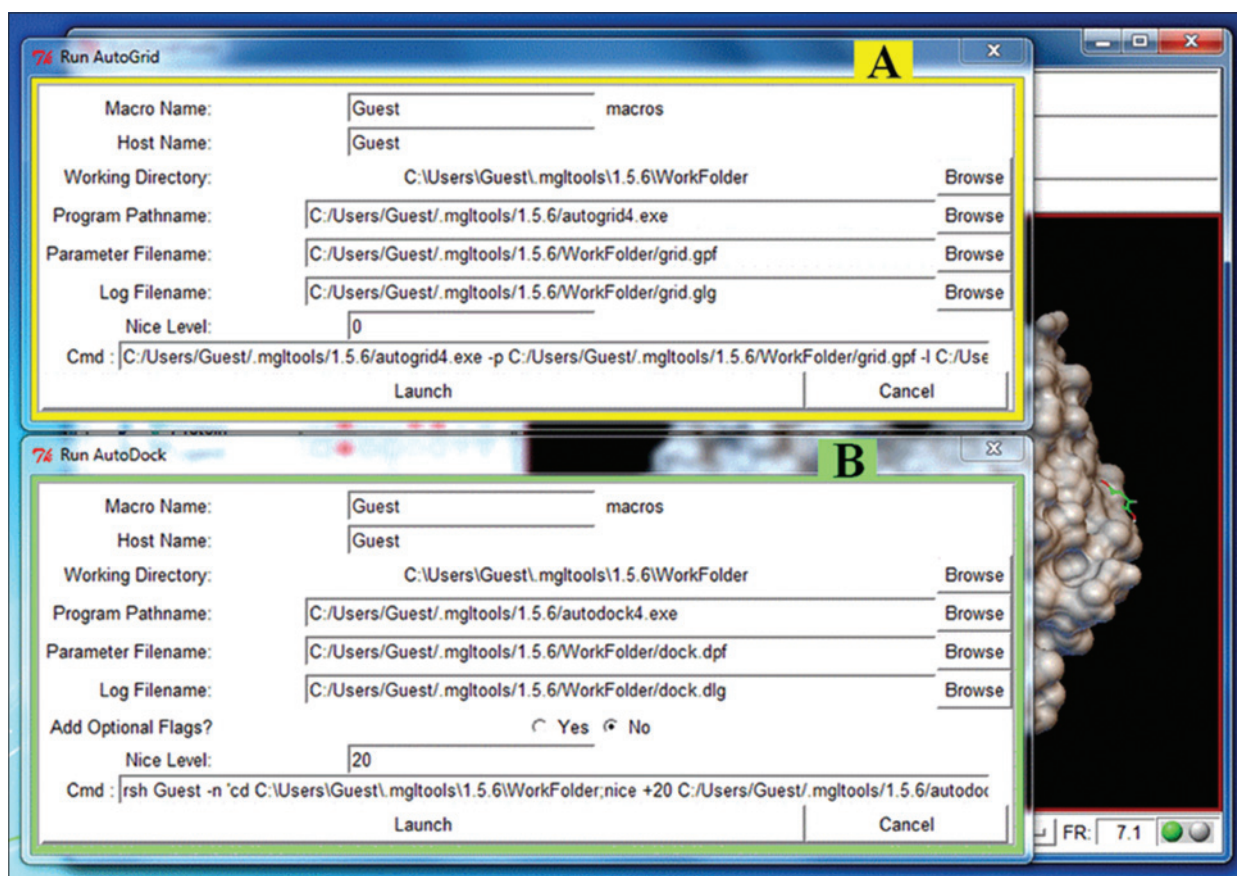
Fig. 5: Panels to execute autogrid4.exe and autodock4.exe (a: Panel to execute autogrid4.exe, b: Panel to execute autodock4.exe)

↳ **Save file as 'dock.dpf'**

Run » Run AutoDock (Fig.5B)

↳ **Select Program Pathname: 'autodock4.exe'**

↳ **Select Parameter Filename: 'dock.dpf'**

↳ **AutoFill for Log Filename: 'dock.dlg'**

↳ **LAUNCH**

Analyzing interaction energy

Once AutoDock4.exe is successfully executed. The result will be given as the ten best confirmations. These can be viewed in the analyze options. The confirmations can be viewed in the order of their free energy binding, by choosing the "Play, ranked by energy" option.

Analyzing the result takes a few steps that is given as a graphical representation in Fig. 6. The panels a, b, c and d will open one after the other, in the same order as given. The ten conformations can be viewed by changing the conformations number in the panel A (Fig. 6A). The interaction energy of the given conformation can be viewed in the panel D (Fig. 6D). The number of hydrogen bonds formed between the ligand and protein can be viewed in the panel C (Fig. 6C).

To Analyze Interaction energy

Analyze » Docking » Open

↳ **Select 'dock.dlg' » Open**

↳ **Assign Ligand New Name » OK**

Analyze » Macromolecule » Choose

↳ **Click 'Protein' » Select Macromolecule**

Analyze » Conformations » Play, Ranked By Energy (Fig. 6A)

↳ **Click on the '&' Button** (Open panel to change play options)

Set Play Options (Fig.6B)

↳ **Check 'Build H-Bonds'**

↳ **View the hydrogen bonds formed** (Fig.6C)

↳ **Check 'Show Info'**

↳ **View the Interaction Energy** (Fig. 6D)

↳ **Build Current Write Complex**

↳ **Save as 'Result.pdb' " Save**

CONCLUSION

The presented protocol to use AutoDock4 is one of the many ways to use ADT. The authors find this to be a simple and beginners protocol to learn AutoDock4. This protocol is designed based on the authors experience in handling of the ADT and no troubleshooting is required, if the protocol is followed as such. If error messages do appear, then kindly start over the protocol, by editing or changing the protein molecule, as it is most likely that, the error could originate from the protein molecule structure. Thus, it is advised to the users, to do further reading in editing and/or preparation of the protein molecule. This protocol is best suited for teaching and learning of AutoDock4. Furthermore, it is a beginner's
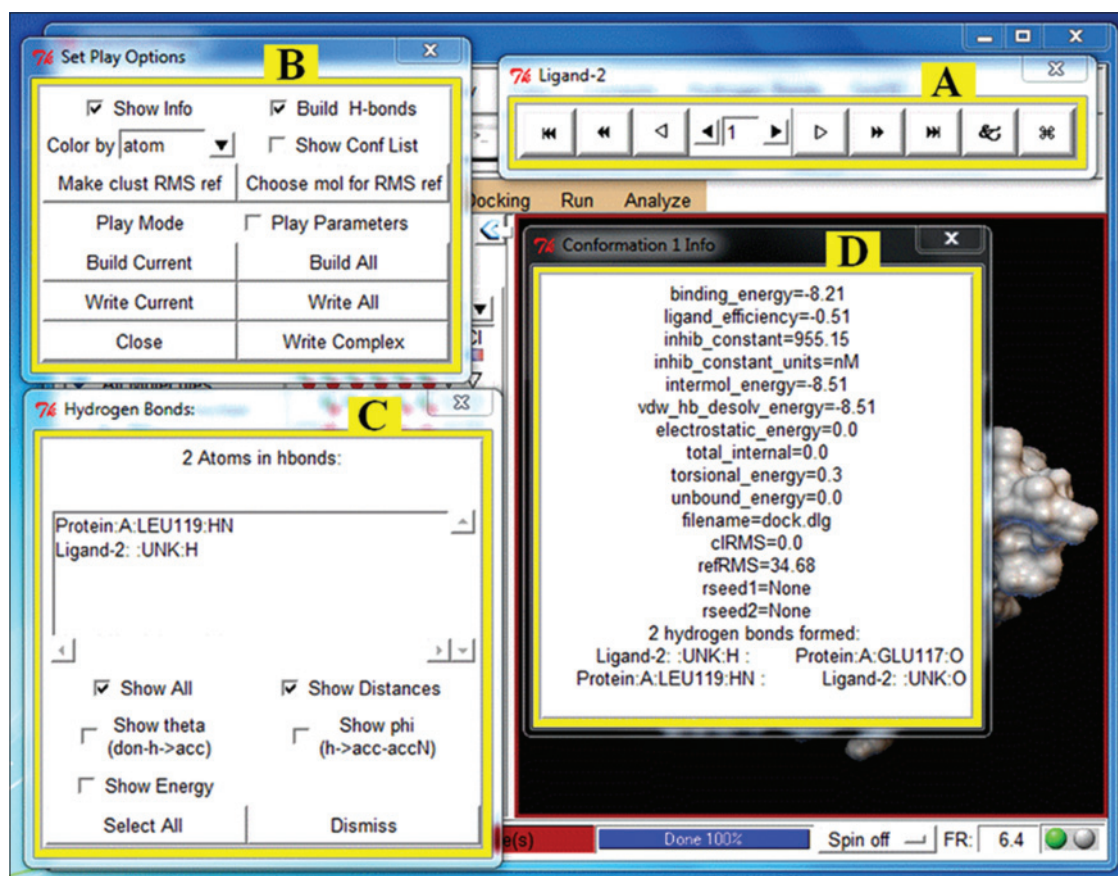
Fig. 6: Panel to view the interaction energy of the top ten conformations

protocol for researchers to analyze simple interactions of ligand molecule and target protein. Since AutoDock4 is only suitable for selective site docking, if the users do not know the region of interest or binding site, the results might be irrelevant. Hence, the users are advised to identify the region of interest in the given protein using binding pocket identifiers, active site prediction tools, etc. Researchers have to identify the region of binding or docking, with help of several online tools that predicts the binding site, active site, binding pockets, etc. Once the given protocol is completed, the energy values and inhibition constant can be obtained from the AutoDock output. The users are recommended to use PyMol software to view the final result.pdb file since it is easier to visualize.