GUI for Omega pseudo torsion parameter-calculation tool

ITWS 3 PROJECT REPORT

Submitted by:
Raina Arora
(200663002)

Goal:

To build a graphical-user interface for an in-house tool (which takes atomic coordinates of RNA molecules as input and computes parameters defining certain structural features) with enhanced interactivity and visualization. The GUI should have the facility for uploading a new RNA file and doing the computations, and should also be able to parse the already computed results according to given parameters for enhanced visualization.

Faculty Advisor:

Dr.Abhijit Mitra

Project Mentor/Teaching Assistant:

Swati Jain

IIIT HYDERABAD,
October 2007

About the Project:

RNA molecules consist of a connected chain of nucleotide bases of four different types. Each nucleotide is associated with structure, sequential connectivity and interaction data. An in house tool defines and calculates five parameters which capture the base-pair geometry and the local strand orientation simultaneously. The parameters currently available are based on either of the factors taken one at a time.

The tool built by the Centre of Computational Natural Sciences and Bioinformatics, IIIT-H, first detects potential base pairs within the cylinder, the height and radius of which are provided as input by the user, and thereafter calculates these parameters known as, omega-eta, omega-theta, omega 1, omega 2 (pseudo-torsion angles), and omega distance (pseudo distance) for the detected base pairs. As part of the further analysis being carried out using omega-eta and omega-theta values, it was required that queries of different kinds (such as those based on base pair type, edge orientation type etc) be executed on all the potential base pairs and then values for these be plotted in order to analyze the base pairs behaving in a similar fashion.

The GUI allows the user to detect base pairs by taking a protein data bank file (which provides the coordinates of various atoms) as input and storing these in a database. Further, it allows the user to plot the omega-eta, omega-theta and omega 1, omega 2 values for base pairs of a particular type among all those available in the database.

Technologies used:

Qt3.3, C++, MySQL 5.0.22 or higher
Best viewed in 1024x768 resolution

Note:

Before using the GUI, the user must create a database named 'RRBPF' and a table 'base_pairs' by writing the following queries in MySql

1. CREATE DATABASE RRBPF;
2. USE RRBPF;
3. CREATE TABLE base_pairs(PDB_name varchar(20),Chain1 varchar(5),Residue_id1 varchar(20), Base1 varchar(5),Base2 varchar(5),Residue_id2 varchar(20),Chain2 varchar(5), Edge1 char,Edge2

char,Orient varchar(20),Omega_eta float, Omega_theta float, Omega_dist float, Omega1 float, Omega2 float);

Platform:

Linux

Source Files :

The files which I got as the source code which were incorporated in the GUI are:

atom.cpp
point.cpp
vector.cpp
Residue.cpp
model.cpp
extract_new_coord.pl
BPF_pre.pl

The files which I made in order to make the GUI are:

main.cpp
basepair.ui.h
basepair.ui and Makefile generated by Qt Designer.

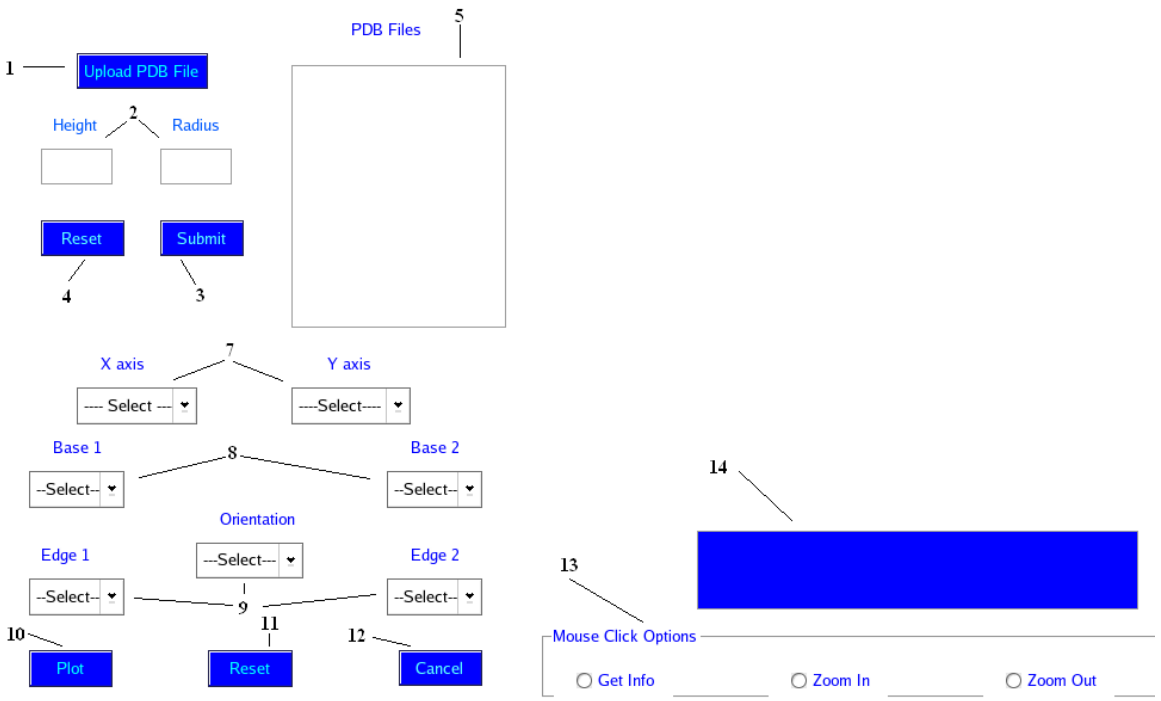
Number of lines of code written :

main.cpp – 14
basepair.ui.h - 458
Lines of code generated by Qt Designer - 3062

Number of functions: 8

User's Manual

This is how the GUI looks:

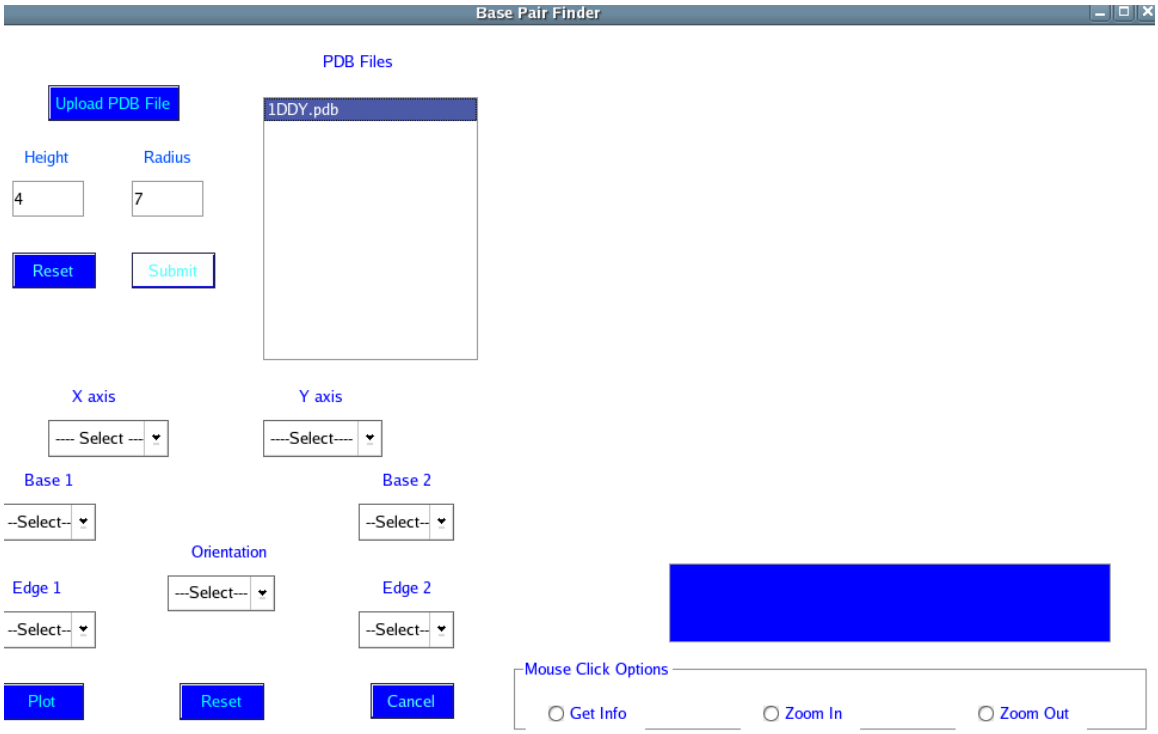


The various buttons in the GUI are as follows:

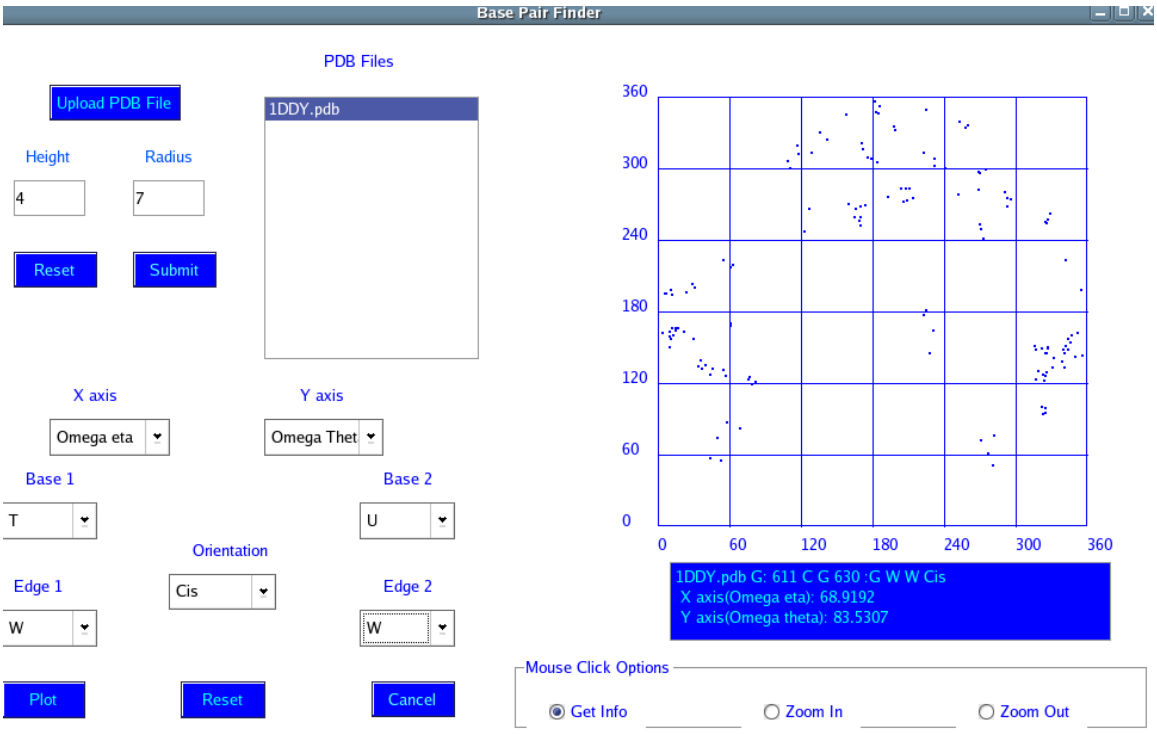
1. 'Upload PDB File' allows the user to browse and then upload a PDB file. More than one PDB files can be uploaded.
2. In the text boxes corresponding to height and radius, user has to enter the values of height and radius in Angstorms. (Reset button description)
3. When the 'Submit' button is pressed, it detects all the base pairs within the height and radius specified by the user and calculates the parameter values (Omega eta, Omega theta, Omega distance, Omega 1, Omega 2) for them.
4. 'Reset' button allows resetting of the values of the height and the radius.
5. The names of the uploaded PDB files appear in the text box next to these buttons.
6. Now the PDB file for which the graph has to be plotted is selected in the text box by clicking on it.
7. The parameters to be plotted on X and Y axes are to be selected in the drop down menus labeled 'X-axis' and 'Y-axis'
8. Base types are selected in the drop down menus labeled 'Base1' and 'Base2'.
9. Interacting edge types and their orientation (cis/trans) are similarly selected.
10. When the 'Plot' button is clicked, all the points corresponding to the input given are plotted in the form of a graph.
    (Plot button also refreshes the graph.)
11. 'Reset' button allows the user to reset values of all the above drop down menus.
12. 'Cancel' button allows the user to exit from the GUI.

13. A group of radio buttons (GetInfo, ZoomIn, ZoomOut) gives user an option to extract residue information by clicking the residue on the graph (Get Info.),or to zoom into a particular area of the graph by clicking on it (Zoom in) or to zoom out of it (Zoom out).
    Magnification is limited to four times the original size.

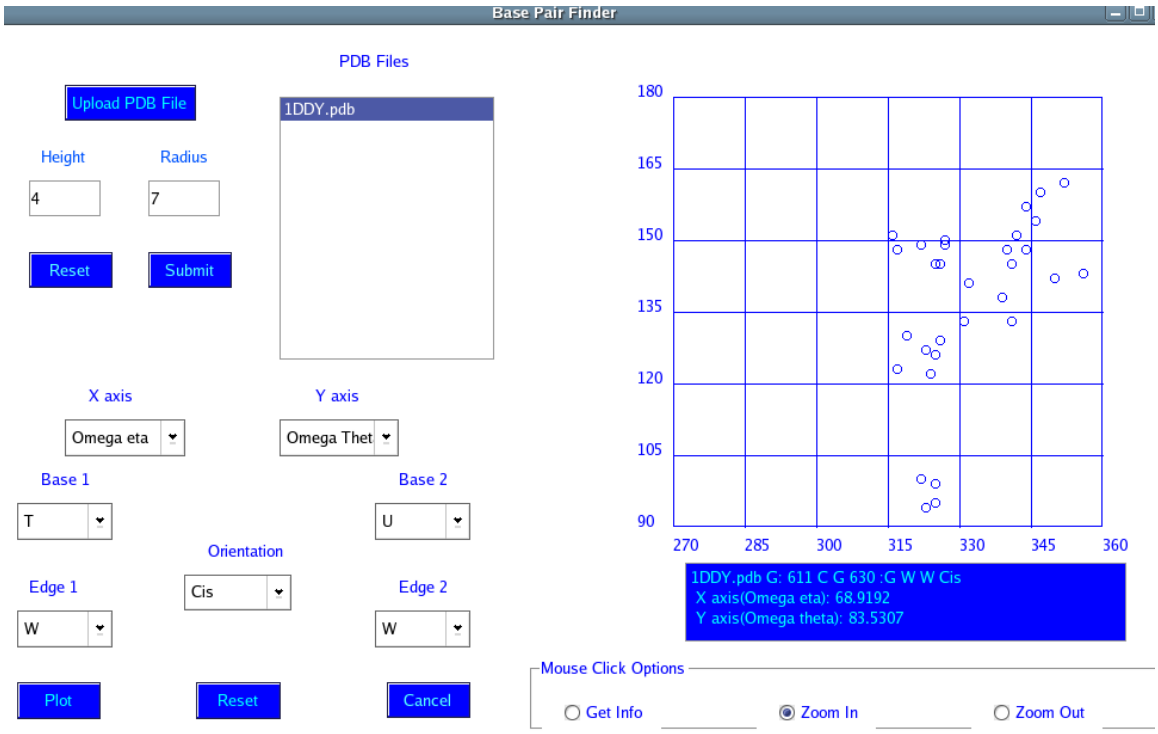
After the ‘Submit’ button is pressed:



After various options are selected, and the ‘Plot’ button is clicked:



After ‘Zoom In’ radio button is checked:



Programmer's Manual

void MainForm::BrowseFile():

This function accepts the file name entered by the user and inserts this name in a list box. All the subsequent files selected by the user are appended in this list box. The Perl script extract_new_coord.pl (provided as source code) is run on the uploaded PDB file and the result is stored in the file called 'model_1.out'.

void MainForm::SubmitValues():

In this function, a function called 'get_base_pairs' is invoked, pseudo torsion parameters are calculated for the base pairs and a file called output_model_1.txt containing this information is created. From this file, all the entries are written in the table named 'base_pairs' in the database RRBPF.

void MainForm::Plot_clicked():

This function runs an SQL query on the entries in the table (base_pairs) to select all those entries that contain the values corresponding to the selections made by the user in the various drop down menus. The corresponding resulting entries are stored in a 'view' table called temp_result.

This view is created for the first time and replaced if it already exists.

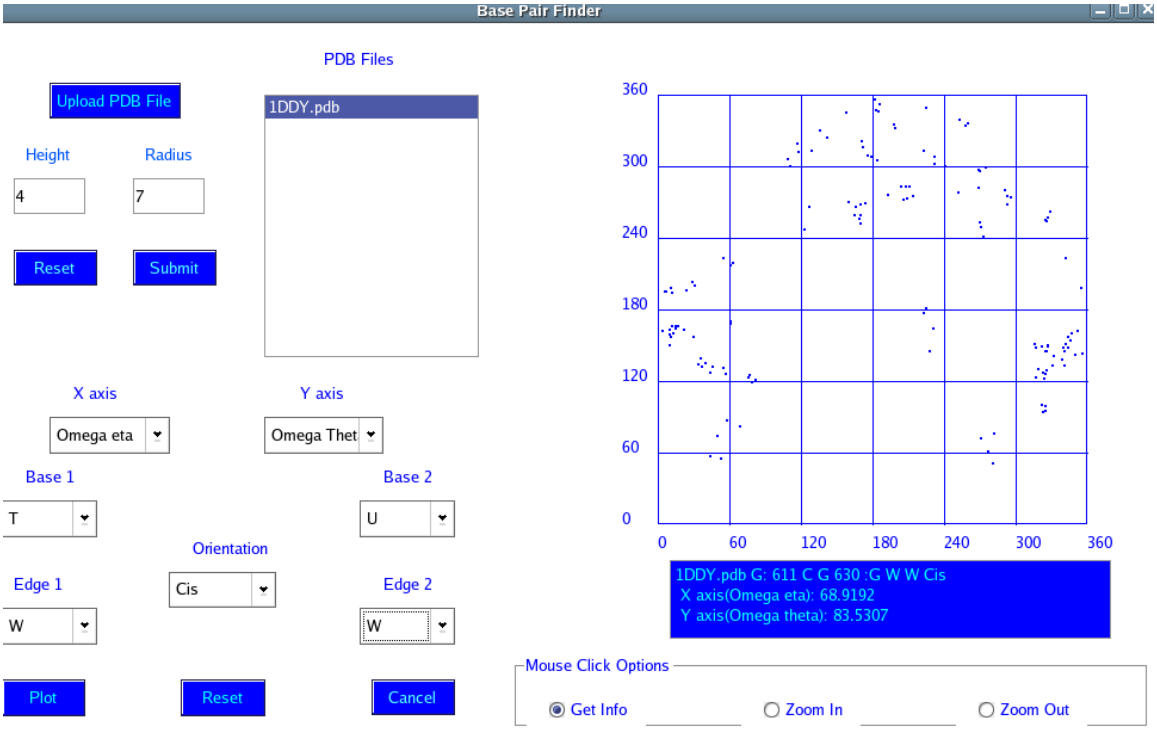
void MainForm::Plot_graph():

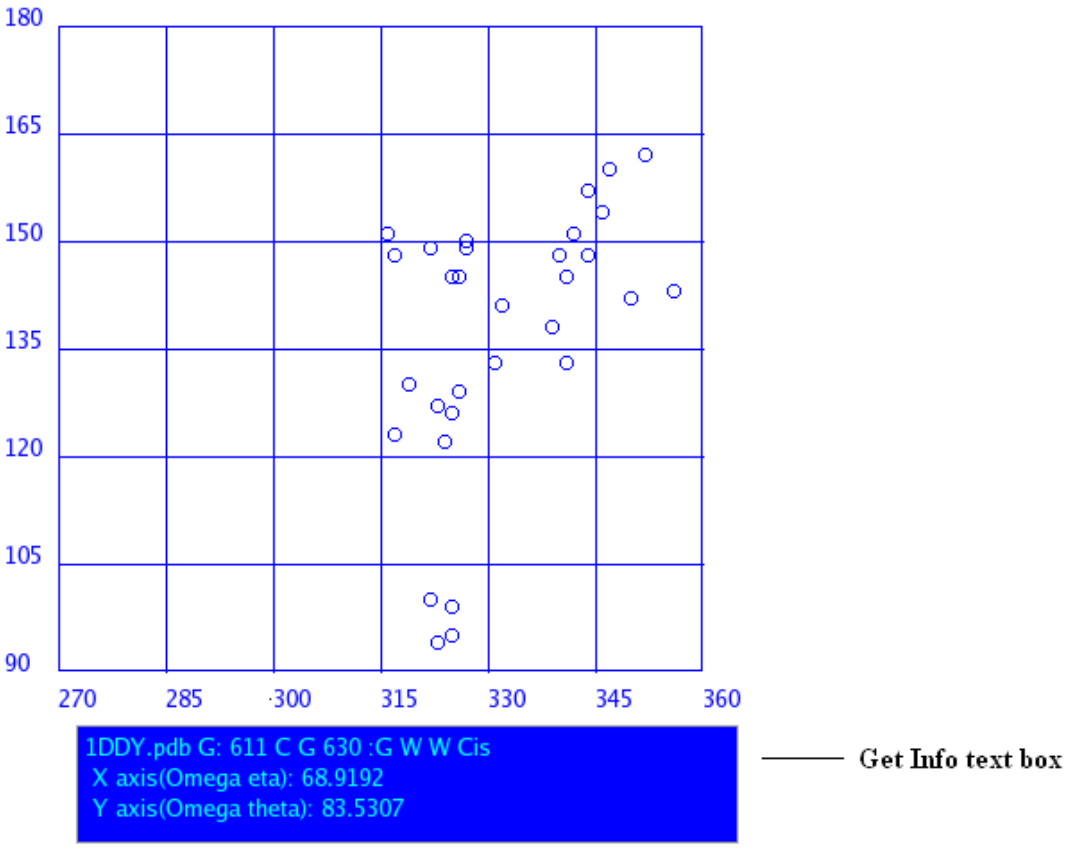
Values of omega-eta, omega-theta/omega1, omega2 for each of these entries in the view are taken as coordinates and plotted as points on the graph.

void MainForm::mousePressEvent():

This function handles the mouse press event in the group of radio buttons (GetInfo, ZoomIn, ZoomOut).

The option GetInfo is invoked when the user is desirous of obtaining detailed information about a residue, he clicks on a point on the graph after it has been plotted. The function accepts the pixel coordinates of mouse click on the graph and then proceeds to determine which molecule the user referred to by his click. An approximation to two pixels is assumed. The form widget recognizes residues within a range of two pixels from the actual mouse click. Thereafter the residue list is scanned in a loop to determine the residue the angles of which best satisfy the clicked coordinate. Failure leads to no action. A successful search displays the information about the residue in the text area. The information made available comprises of the PDB file name, Residue id, chain and type for both the bases, interacting edges and orientation and pseudo torsion parameters plotted.


Base Pair Finder
PDB Files
Upload PDB File
1DDY.pdb
Height
Radius
4
7
Reset
Submit
X axis
Omega eta
Y axis
Omega Thet
Base 1
T
Base 2
U
Orientation
Edge 1
Cis
Edge 2
W
W
Plot
Reset
Cancel
360
300
240
180
120
60
0
0
60
120
180
240
300
360
1DDY.pdb G: 611 C G 630 :G W W Cis
X axis(Omega eta): 68.9192
Y axis(Omega theta): 83.5307
Mouse Click Options
Get Info
Zoom In
Zoom Out



'ZoomIn' option is responsible for magnification of the graph. Magnification is based on the principle that whenever a certain point on the graph is clicked with radio button set for "Zoom In", the eight square grids surrounding the grid clicked on the graph are brought, to the center of the graph. The factor that determines the scale of the x and y axis is made double of its previous value. This factor is called 'mag' is bounded to a value of four, beyond which magnification cannot take place. This limitation is caused because of the fact that pixels can be integral values only. Hence we cannot allow angle values to be expressed in floating point values.

First case magnification causes scale (separation between two grid lines) on axis to become 30° then 15° for second.

The 'mag' value is used to determine the starting angle values on the axis for labeling purposes using linear algebraic equations. These calculations alter the values of global variables "rangex" and "rangey" that are used for labeling the x and y axis on the graph.

The ‘ZoomOut’ option is invoked when the user clicks on the plot with radio button set for “Zoom out”. This restores the graph to its previous state before it had been zoomed into last. The magnifying factor is made half its current value. Using algebraic linear equations the values of “rangex” and “rangey” are altered and the graph is repainted with these fresh values. In case the user attempts to zoom out without zooming in even once, mag value is set to one and “rangex” and “rangey” values are set to initial values of 0° and 360° start and end values respectively.

Known Bugs

A known bug in the Plotter tool is the lack of precision of mouse click events on the graph. As pixels on the form widget can be addressed only as integral values, there are chances that attempts at retrieving information about a particular residue using the Get Info option does not provide the information about the residue which the user wanted. Such a case would arise when residues are co-joined or are very near each other.

Future Developments

The GUl still does not provide the user with an option of plotting a specific type of base pairs for more than one file simultaneously, with points from different files being indicated in different color.