REDCAT LABORATORY EXERCISE
BCMB/CHEM 8190, 2014

Installation note for instructors:

REDCAT may be downloaded from: http://ifestos.cse.sc.edu/software.php#redcat

You will also find further instructions and a manual accessible on the pages that appear. Versions exist for Linux and Mac OS X. REDCAT uses gnuplot and Tcl/Tk 8.0 or later.

Input files may be obtained from: http://tesla.ccrc.uga.edu/courses/bionmr/labs/

The tarball (redcat-lab-2010.tar) includes:

- REDCAT-lab-2010.pdf (this file)
- 1bq8.pdb, 1bq8-frag1-1-25.pdb, 1bq8-frag2-27-54.pdb (the protein structure files)
- 1bq8-RDCs-2-53.txt, 1bq8-RDCs-27-53.txt (the residual dipolar coupling files)
- map.dat, map3D.dat (files used in plotting results in REDCAT)

Laboratory exercise:

Introduction. REDCAT (REsidual Dipolar Coupling Analysis Tool), is a program that aids the interpretation of residual dipolar couplings (RDCs) in terms of structural models. It is primarily designed to test RDC data for consistency with structural models, but it can also aid in the assembly of structures from fragments by orienting each fragment in a common principal alignment frame. It incorporates a number of other useful tools such as graphical display of principal axis directions and prediction of averaging effects on RDCs for comparison with experimental data. It operates by taking in coordinates for various dipole interaction vectors (H-N bonds for example) and experimental RDCs, and then solving for an order tensor by singular value decomposition. It does this thousands of times using a random sampling of RDCs within defined error limits to produce a picture of reliability of results.

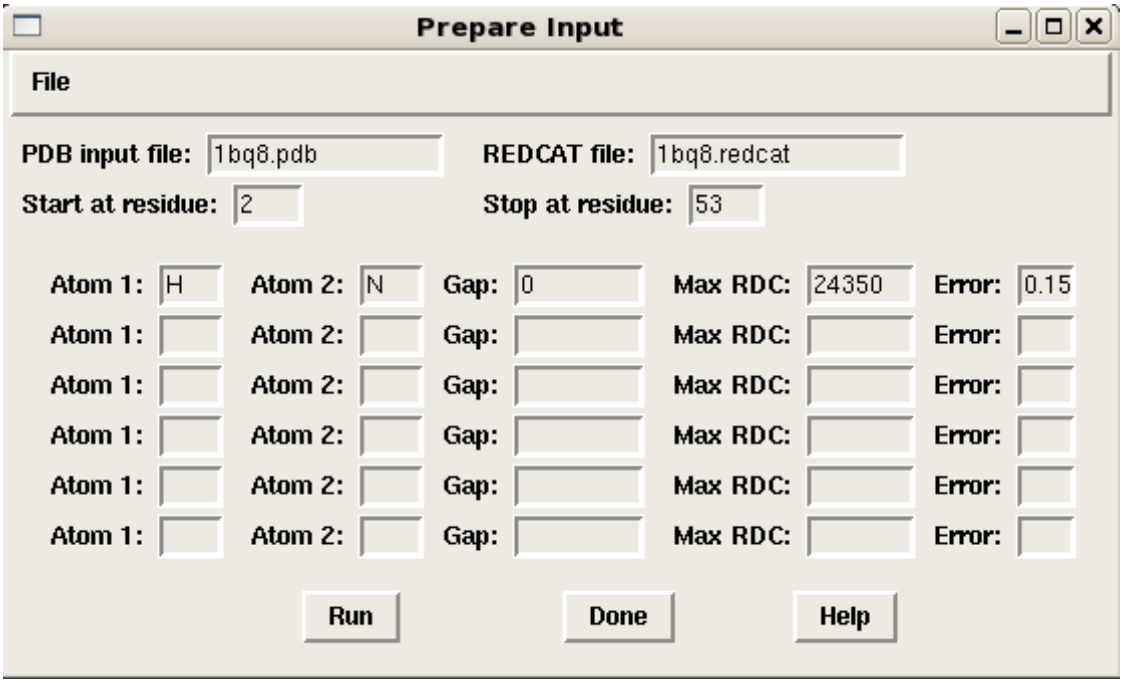
REDCAT has been installed on the computers you are using (or is accessible by opening a remote shell for a server having REDCAT installed). The instructions that follow are intended to supplement the manual by leading you through an exercise that manipulates data on a small paramagnetic protein, rubredoxin. The RDCs are only for H-N vectors. They are small in magnitude because they were collected using only natural field induced orientation of this paramagnetic protein. However, the data should suffice to illustrate principles of the program. You will first use data for the intact protein, and use REDCAT tools to evaluate consistency between RDC data and a crystal structure. You will then do a fragment assembly project in which the protein has been divided into two parts with one rotated from its original orientation. You will use REDCAT tools to reestablish the proper orientation.

Step 1. Creating an input file

- Open a terminal window and verify that there is a REDCAT directory with required files present (**ls**). Change to the that directory (**cd**).

- Start REDCAT by typing **REDCAT.tcl** or **redcat** (aliased on our systems)

- Under the **File** menu select **Prepare Input**



- Enter **1bq8.pdb** as the pdb input file

- Enter a file name (1bq8.redcat) for the REDCAT file to be created

- Tell the program what residue to start and end at (**2** and **53** for this example).

- Enter the type of data you intend to use in the first line of the table. For our example this will be "**H**, **N**, **0**, **24350**, **0.15**". The atom names tell the program what coordinates to extract from the PDB file (these must be exactly as they appear in the PDB file), the **0** tells the program to find these in the same residue, the **24350** is the coupling that would be seen for a pair of atoms separated by 1Å and oriented rigidly along the magnetic field. The **0.15** is an estimated error for the RDC data that will be entered in the next step.

- Select **Run**. A message should tell you that the input file was written to disk.

- Select **OK,** then **Done in the input window**.

Step 2. Loading the input file and RDC data

- Select **Load** from the **File** menu. Navigate to the directory you are in and select the file you just created. Click **Open**. The **999** entries in the "**Dipol"** column indicate that no valid RDCs are yet in the file. The gray (unselected) boxes at the left indicate that none of these data will be used in the calculations.

- Select **Import RDC** from the **File** menu. Use **1bq8-RDCs-2-53.txt** as the input for this section. Note that the entries in this text file must start at the starting residue indicated in Step 1 and continue in order for all coordinate pairs selected. Most of the 999s should be replaced by actual data. The checkboxes of these entries are now selected to indicate that this data will be use in calculations.

- Use the default values for number of trials (10,000), number of null space values (10), and error search range (1).

Step 3. Validating data and the model

- Press **Run** and after a brief wait you should see a **Message!** window appear. This displays success of the calculation. The second line tells you that 10,000 out of 10,000 tries were rejected. Subsequent lines list how many times a particular RDC was the cause of a rejection. If you scroll down you will see a red line for equation 32 and 47. Each of these data points caused all possible solutions to be rejected. It might be a good idea to check this piece of data, or for the demo, just eliminate it. This can be done by deselecting the equation number in the main window. First click on **Clear all** then **OK** to return to the main window. In the Main window, find the equation number you wish to deselect and click on the box next to the number. It will turn gray (or be unchecked) indicating that the data point will not be used. A few other equations (ones with 999 for RDC values) were already eliminated. Most are for prolines which have no H-N vector. Press **Run** again. New results will be appended to the bottom of the **Message!** screen (if not cleared from the previous run).

- This run may still show 10,000 out of 10,000 rejections, but this cannot be traced to any one problem piece of data. Go into the **Tools** menu on the main screen and select **Error Analysis > Perform Error Analysis**. Each equation highlighted in red has an error equal to or higher than the 0.15 error limit set. Change the error limits by **selecting Tools > Error Analysis > Get Estimated Errors for All**. This will increase the error for those RDCs in the output which have estimated errors equal to or higher than the error set in the **Main window**. Select **Run**. You should now get more than a thousand accepted solutions.

Step 4. Displaying and plotting results

- Go to the **Tools** menu and select **Get Solutions** under **Solution**. You should see the principal order tensor solutions (S_{xx}, S_{yy} and S_{zz}), Euler angles (denoted in this program as **a**, **b** and **c**) as well as Eta (asymmetry) and GDO (the general degree of order) for each possible solution. Euler angles relate the molecule frame to the principal alignment frame. This large number of solutions is not easy to digest, but you can also **Get Best Solution**.

- Plot the solutions to show the directions of the principal alignment axes in the molecular frame. This is done with the Sauson-Flamsteed plot (**Tools > Plot > 2D SF Plot**).

Step 5. Working with fragments

- Repeat steps 1 through 4 using the same protein, rubredoxin (1bq8), this time broken into two fragments. The fragment files are: **1bq8-frag1-1-25.pdb** and **1bq8-frag2-27-54.pdb**. For fragment 1 use the same RDC input file as before (**1bq8-RDCs-2-54.txt**). For this first fragment you should start at residue 2 and stop at residue 25. For fragment 2 use **1bq8-RDCs-27-53.txt** for the input file, start at residue 27 and stop at residue 53. The PDB file for the second fragment has been rotated into the principal alignment frame. When you look at the axes plot for this fragment you should see solutions fall on the axes of the globe.

- Rotate the first fragment into its principal axis frame. To do this you need to record the Euler angles for the best solution. You'll find the option under **Tools > Rotate > Rotate PDB**. Enter the PDB file name, output file name (make it different from 1bq8-frag1-1-25.pdb**)**, and Euler angles (a = Alpha, b = Beta, c = Gamma; from **Tools > Solution > Get Best Solution**) into the window that opens. This allows you to create a new PDB file in the principal alignment frame (PAF). You can recalculate order tensor solutions with this new PDB and view its Sauson-Flamsteed plot to verify this operation.

Step 6. Examining results

- Both PyMOL (command line: pymol) and Chimera (command line: chimera) are available on the UGA Linux systems. Open your favorite molecular graphics program and select the PDB files for fragments 1 (1bq8-frag1-1-25.pdb) and 2 (1bq8-frag2-27-54.pdb). In Chimera pull down the list under **file** and select **open**. You can select the appropriate file using the directory structure shown. The default display is an all atom stick diagram display. You may find it easier to look at a ribbon diagram. Look under the **Actions** pull down list; select **Atoms and Bonds**; select **hide**; go back to **Actions**; select **ribbon** and **show.** Note that you can change color and other display options. Look at the **Favorites** pull down list and select **Command Line.** You should see an "active Models" line with a check mark by model 0. Open your second pdb file in the same way. You should now

see a check mark by model 1. You can select and unselect the models by clicking on the checked boxes. You can use mouse buttons to rotate, zoom in and out, and translate the model selected. The Chimera command **reset** is useful. Typing it in the command line followed by a return restores views to those originally loaded.

- You will see that the models as loaded do not connect as a single protein. You can **Close Session** from the **File** menu or simple hide the ribbon diagrams working from the **Actions** menu to eliminate the current display. Now open your new PAF 1-25 fragment (created in step 5) along with fragment 2 (1bq8-frag2-27-54.pdb). See if you can make an intact molecule just using translations (select one model and use the mouse with the translation button depressed). Remember that RDC based orientations can be 4 fold degenerate (turn 180 about x, about y or about z). You may have to experiment with these alternate orientations to produce an intact molecule. The rotation operations can be executed on the selected model by typing command like **turn x 180** in the command line.