PROTMAP 1.2.2 Tutorial

This tutorial introduces new users to PROTMAP2D and its capabilities. Detailed instructions are provided to demonstrate the most typical analyses.

All the file structures mentioned below are to be downloaded from the program's ftp directory: ftp://ftp.genesilico.pl/pub/software/protmap2d/. However, many types of input files can be used in different jobs (e.g. a molecular dynamics trajectory can be analysed as a trajectory or as an ensemble etc.).

Getting started: run the executable protmap2d file.

1. Calculate a contact map for a single model

• Select **single model** from the job list in the main menu

• Click the **Browse...** button and select the file **pdb1acv.ent**

• Select **residue numbers like in the input file(s)** option from the corresponding residues' assignment list to keep the numbering of residues corresponding to that in the PDB file (*selecting the other option, **renumber residues from 1 (N-terminus)** would result in renumbering the residues starting from 1. Renumbering ignores the residues with missing coordinates*)

• Click the **Load** button (the structure will NOT load until you actually click the **Load** button)

After uploading the coordinates, a new menu appears, which allows you to define the way how the distance is calculated and which features of the model to display (*see more details in the PROTMAP2D-manual document*).

• Select **pdb** option from the **secondary structure** list to enable displaying helices and strands, according to the assignment that should be present in the original input file. (*If you load a different file without the helices and strands assigned, this option will result in the apparent "absence" of secondary structure. In such cases the user may use third-party programs STRIDE or DSSP that have to be downloaded separately – see below. If STRIDE and DSSP are installed, corresponding options will appear in the **secondary structure** selection box*)

• Click the **Go** button to run the calculation of intramolecular distances. A contact map similar to the one shown below should appear in the PROTMAP2D window in about 2 seconds. (*If you changed the distance metric, secondary structure representation or the matrix type, the map will be different*)
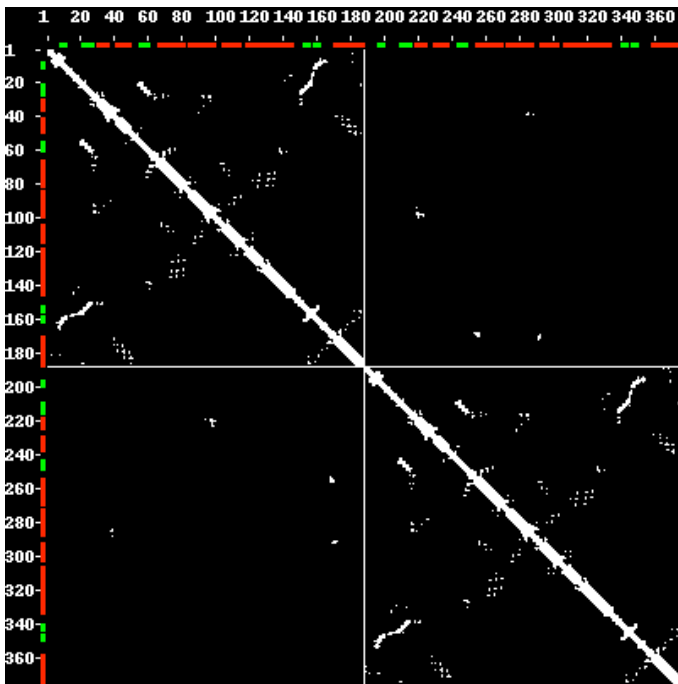
The uploaded structure is dimeric. Perpendicular white lines, crossing in the middle of the picture, separate the C-terminal residue of the first chain and the N-terminal residue of the second chain. The resulting four squares delineate contacts present within each chain and between the chains.

Color bars on the X/Y axes denote secondary structure elements (red- helices, green-strands. If white color appears (*not for this example file*), it means there's a structural gap, e.g. missing residues in the PDB file.

• To zoom on any part of contact map, select the starting point by pressing left mouse button and draw a rectangle, then release the button – a new pop-up window appears (example shown below). It can be resized to visualize details of the map.


• Additionally, it is possible to zoom and resize the whole map - to do this, press the **Fullscreen** button

2. Calculate a distance map for a single model

• Go back to the main menu, and choose **single model** from the job list

• Click the **Browse...** button and select the input file **pdb1aca.ent.** It is an NMR file with 20 models, only the first model will be analyzed. (*do not confuse it with a dimeric structure pdb1acv.ent used in the previous example – it is much larger and the calculation of a distance map would take about 10 minutes!!!*).

• Click the **Load** button

• In the new menu choose **distance map** option from the matrix type list (*notice that upon changing selection from contact map to the distance map, the distance threshold disappears from the menu*)

• Click the **Go** button to run the calculations

A distance map similar to the one shown below should appear in a few seconds
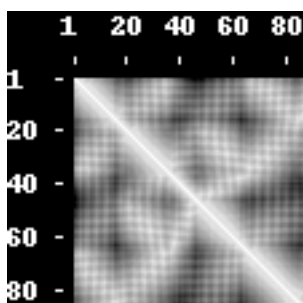
• To switch a view to contact map, press the **Contact** button (distance cutoff is set to 8 Ångströms by default). Result shown below:
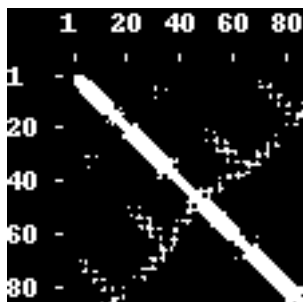

• Now it is possible to change the cutoff value ("floating cutoff") by rolling mouse wheel up and down and see the result on the fly (actual value is shown below the picture). Result shown below with cutoff **11.0** Å
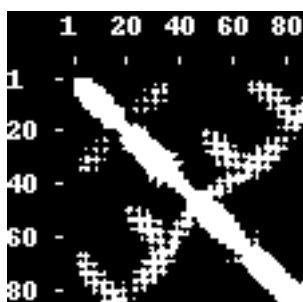
• To switch a view back to distance map, press the **Distance** button

• To crop distance map by a preset distance cutoff, input value to **lower limit** and **upper limit** fields, e.g. **7.0** and **20.0**, respectively (to do this, click on the field first, then type the value). Result shown below
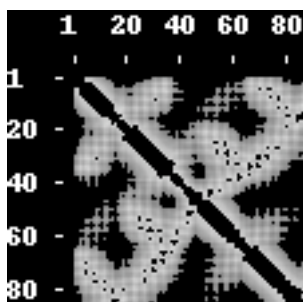

• Set the **lower limit** value back to **0.0**

• To cut off close sequential neighboring contacts, set the **separation** value, e.g. to 4 which results in excluding up to three nearest neighbors in sequence. Result shown below:
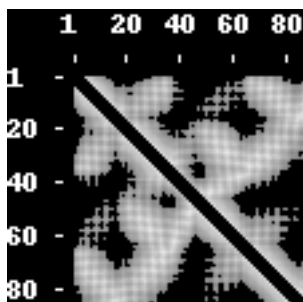
3. Model ensemble statistics

• Go to the main menu, choose **model ensemble** from the job list

• Click the **Browse...** button, select the file **pdb1aca.ent** (NMR file with 20 models), and click the **Load** button

• In the new menu, select **pdb** option from **secondary structure** select list

• Click the **Go** button to run the calculations

A map of contact frequency should appear in a few seconds.
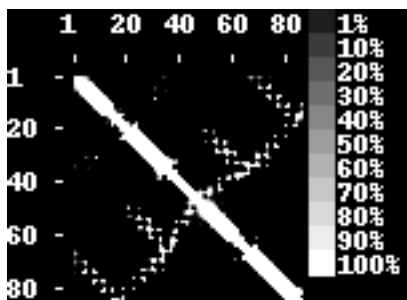
Shades of grey indicate the percentage of the contact occurrence in the whole model set (legend on the right side of the picture).

4. Model ensemble rough set

• Go to the main menu, choose **model ensemble** from the job list

• Click the **Browse...** button, select the file with multiple alternative models (e.g.: **pdb1aca.ent**), and click the **Load** button

• In the new menu, select **rough set** option from the task list

• Click the **Go** button to run the calculations

• A black&white&grey contact map similar to the one shown below should appear in a few seconds. White color indicates contact that occur in all models from the file. Grey color indicates contacts that occur in some models, but not all.
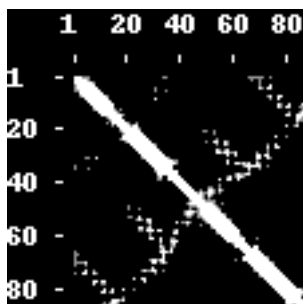
5. Model vs. model (structures used: example-model.pdb, example-xray.pdb)

• Go to the main menu, choose **model vs. model** from the job list

• Click the first (upper) **Browse...** button and select file *example-model.pdb*

• Click the second (lower) **Browse...** button and select file *example-xray.pdb*

• Click the **Load** button

• In the new menu select **Cb** option from the **metric** select list (*this time we will measure the distances according to C-beta atoms, as they better reflect interactions between the amino acid side chains; see the Manual for more information about the distance metrics*)

• Click the **Go** button to run the calculations

A black&white&grey contact map similar to the one shown below should appear in a few seconds.
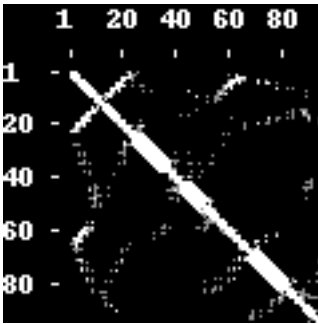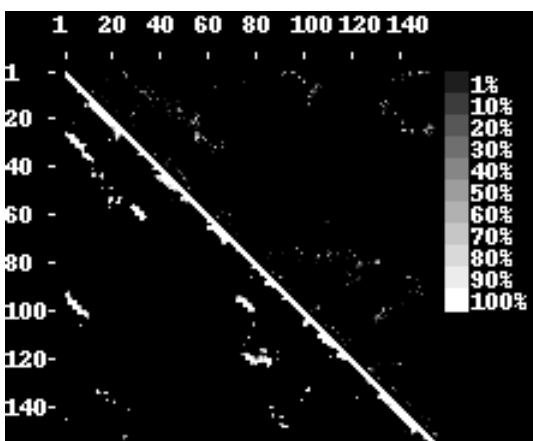
The lower-left section of the map indicates contacts in the first model (example-model.pdb) while those the upper-right part indicates contacts in the second model (example-xray.pdb). White dots present in both triangles indicate contacts common to both models, while grey dots specific to each triangle indicate the contacts that appear either in one model or the other.

6. Trajectory statistics

• Go to the main menu, choose **trajectory** from the job list

• Click the **Browse...** button and select the file **500_16.pdb** (output of a molecular dynamics simulation of protein unfolding, comprising an ordered series of structures, starting with the native structure. This file is a 'minimal' version of the **500.pdb** file with 200 structures, which would require more time for all calculations)

• Click the **Load** button to upload the structure

• In the new menu, select **statistics**

• Click the **Go** button to run the calculations (about 10 seconds)

A map of contact frequency appears.

This map is analogous to the one for ensemble (p.3), but focuses on comparison of the first structure (assumed to be the starting point of the simulation) with all others. The lower left section of the map indicates contacts present in the first model in the file (assumed to represent the native structure). The upper right section indicates all other contacts (non-

native, i.e. absent from the native structure). Shades of grey indicate the percentage of the contact occurrence in the whole set of models (legend on the right side of the picture).

• To merge the native and non-native contacts (and make the map equivalent to one that could be obtained with the 'ensemble' option), click the **Merge** button.

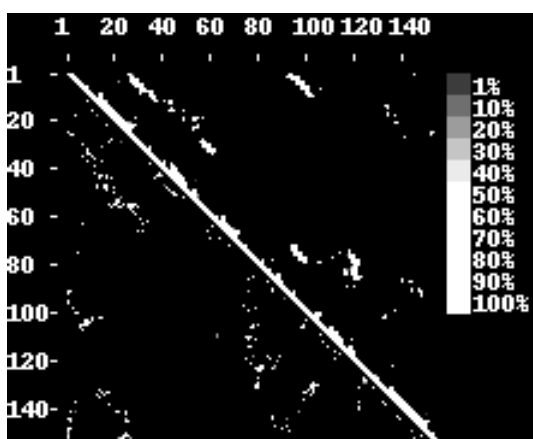• To undo this, click the **Unmerge** button

7. Trajectory history

• Go to the main menu, choose **trajectory** from the job list

• Click the **Browse...** button and select the file **500_16.pdb**

• Click the **Load** button

• In the new menu, select **history** option from task list

• Click the **Go** button to run the calculations (about 10 seconds)

A contact map with shades of grey similar to the one shown below appears.

• Click the **Merge** button.

In this example, shades of grey correlate with the index of the model in which the given contact was observed for the last time in the simulation. Thus, white dots correspond to contacts present in the last model, and dark grey dots correspond to contacts that appeared at the very begin of the analysis and never reappeared in further models. In other words, the appearance of a contact "whitened" the corresponding spot on the map, which then faded to dark with the progress of the simulation. This image shows the situation from the perspective of the last frame. As in the previous example, the lower left section of the map indicates contacts present in the first model (native) while the upper right section indicates contacts absent in the first model (non-native).

8. Trajectory contact map persistence

• Go to the main menu, choose **trajectory** from the job list

• Click the **Browse...** button and select the file **500_16.pdb**

• Click the **Load** button

• In the new menu, select **contact persistence** option from the task list

• Input value **5** in the **contact persistence** field

• Click the **Go** button (calculations will take about 10 seconds)
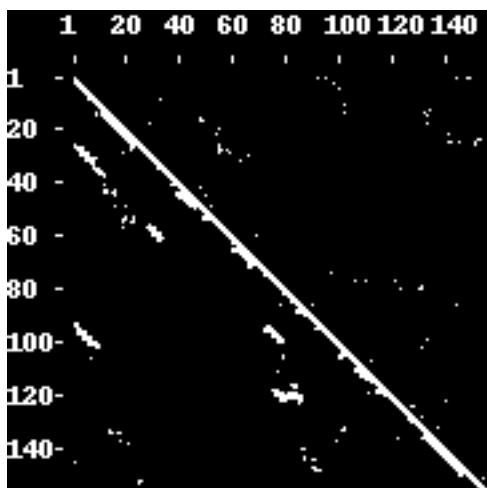
The resulting map highlights contacts that are present in at least 5 (or any other number given by the user) successive models in the trajectory file. This feature is designed to search for stable contact groups formed during the simulation, e.g. involved in stabilization of secondary structure elements. Again, native and non-native contacts are shown separately and can be merged and unmerged.

9. Trajectory contact map comparison

• Being at the main menu, choose **trajectory contact map comparison** from the job list

• Click the first (upper) **Browse...** button and select file **300_16.pdb**

• Click the second (lower) **Browse...** button and select file **500_16.pdb**

• Click the **Load** button

• In the new menu, input value **7.0** Å to the **distance** field

• Select **speedy** option from the computing type list (*should work on most computers with reasonable RAM memory – if you experience any problems – use the slower memory-safe option*)

• Click the **Go** button to run the calculations (which will take about 20 seconds)

A contact map similarity matrix similar to the one shown below should appear.
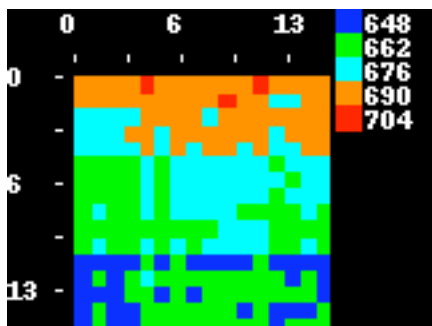

Colors symbolize the degree of similarity (defined as the absolute number of common contacts) between contact maps from the two files (values at the axes indicate the model index). See the color legend for the value range.

10. Contact map clustering

• Go to the main menu, choose **contact map clustering** from the job list

• Click the **Browse...** button and select the file **500_16.pdb**

• Click the **Load** button

• Select **Cb** option from the **metric** list

• Select **speedy** option from computing type list

• Click the **Go** button to run the calculations

A contact map similarity matrix similar to the one shown below should appear.
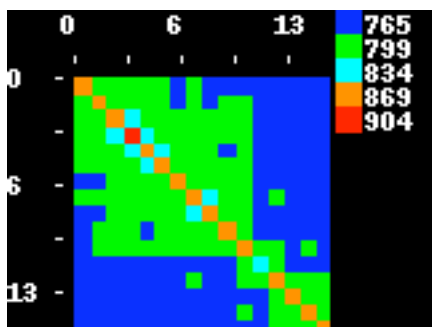
The picture shows a matrix similar to the one in the previous example. This time it illustrates (dis)similarity between contact maps of different structures in the same ensemble (*here, all models in the trajectory are treated as equivalent, no special attention is paid to the first or last model*). PROTMAP2D provides a very simple routine for extraction of subsets of structures with similar contact maps.

You can use a naïve PROTMAP2D clustering routine by imputting the percentage of similarity that defines a cluster and obtaining a text file with model ids grouped in clusters.

Alternatively, you can save the matrix in one of the many supported formats (MS Excel, PHYLIP, CLANS) and import it to some 3rd party clustering software to carry out clustering using a different algorithm (see PROTMAP2D-manual for examples).

11. Contact map from file

• Go to the main menu, choose **contact map from file** from the job list

• Click the uppermost **Browse...** button and select one of the following files: **spreadsheet_1hg3_allchains.xls** (a <u>contact</u> map in the MS Excel format) or **1aca_distancemap_phylip.txt** (a <u>distance</u> map in the PHYLIP format). To select MS Excel format, you must first enable "*MS Excel Workbooks*" file filter.

• Click the **Load** button

• Compact / distance map, depending on the input, should appear in 1-3 seconds

12. Trajectory - movie (file used: 500.pdb)

Warning: This exanmple is time-consuming. Be patient.

• Go to the main menu, choose **trajectory - movie** from the job list

• Click the **Browse...** button and select the file

• Select **renumber residues from 1 (N-terminus) option** from the corresponding residues' assignment list

• Click the **Load** button. Wait about 20 seconds for file pre-processing

• Click the **Go** button. First frame of a movie should appear in about 3 minutes.

• Click the **fadeout** box and set the value e.g. to 8

• To run the animation manually: press the **Movie** button. Now roll mouse wheel up and down to see animation

• To exit the animation, press the **Stop** button

• To run the movie automatically: press the **Animation** button.

• To exit the animation, press the **Stop** button