
CMView 1.1 User's Manual

Table of Contents

Overview	1
Contact map definition	1
Contact types	1
Screen elements	2
Function Reference	3
File	3
Select	5
Color	7
Action	7
Compare	8
Help	10
Context menu	10
Experimental features (unsupported)	10
Overlays	11
Command line parameters	12
Configuration	12
Overview	12
Available configuration parameters	12

Overview

Contact map definition

CMView is a software tool for visualizing and analyzing the network of contacts between amino acid residues in a protein structure. Formally, this network is a graph where each residue corresponds to a node and two nodes are connected by an edge if and only if the two residues are in contact. Two residues are considered to be in contact if they are spatially close in the three dimensional structure. In CMView the definition of contact is specified by two parameters: the contact type and the distance cutoff. The contact type defines a subset of atoms of the residue. Two residues i and j are then in contact if two atoms out of this subset, one from i and one from j , are not further apart than the distance cutoff. The depiction of the contact network as a matrix (formally, the adjacency matrix of the contact graph) is called a contact map.

Contact types

The supported contact types are:

AL	All atoms
BB	All backbone atoms
SC	All side-chain atoms
C	The C atom of the backbone
Ca	The C-alpha atom of the backbone
Cb	The C-beta atom of the side-chain (C-alpha for glycine)

Screen elements

Contact map window

CMView shows the contact map of a single polypeptide chain in its main window. A position (i,j) in the contact map corresponds to residues i and j in the sequence of the displayed chain. The contact map is symmetrical with respect to the main diagonal. In the lower right corner information is displayed about the residue pair (or pairs in compare mode) at the current mouse position:

- The residue numbers of i and j in the sequence
- The residue types (in three letter notation)
- The secondary structure types (alpha, beta or none)
- The PDB residue numbers. These are shown only if a tertiary structure file was loaded. By default, CMView numbers residues by their position in the sequence. This does not necessarily match the numbers assigned to residues in the ATOM lines of PDB files. To compare with other data where the PDB residue numbering is used, the residue numbers as found in the PDB file are printed here. (Note: This function has to be enabled with `SHOW_PDB_RES_NUMS=true` in the config file).

Secondary structure rulers

The rulers shown above and to the left of the contact map are used for:

- Displaying secondary structure elements
- Selecting residues (see selection modes)
- Selecting secondary structure elements (see selection modes)

The color scheme for secondary structure elements is:

- Blue: Alpha-helix
- Red: Beta-sheet
- Green: Turn

Side bar

To the right of the contact map is the side bar which contains further elements:

- The overlay menu: Different background overlays can be selected which display additional information in the contact map window. The overlays can be selected separately for the upper left and lower right half of the (symmetrical) contact map. See [Overlays](#).
- Further context-dependent tools for experimental features (if enabled)
- At the bottom: the coordinates display (information about the residue pair at the current mouse position in the contact map)

The following overlays can be selected from the overlay menu:

- Distance map (shows distances between C-alpha atoms if 3D coordinates are available)
- Contact density (helps to identify domain- and subdomain structures)
- Common neighbours (triangle inequality relationships)

Function Reference

File

Info

Shows information about the currently loaded contact map (or contact maps in compare mode):

Name	An identifier for the contact map (usually the PDB four-letter plus chain code if applicable)
Contact type	The contact type used when creating the contact map, e.g. Ca, Cb, SC, BB
Distance cutoff	The distance cutoff in angstroms used when creating the contact map
Min Seq Sep	The minimum sequence separation filter used when creating the contact map
Max Seq Sep	The maximum sequence separation filter used when creating the contact map
Sequence	The protein sequence. To see the full sequence, hover mouse over the first few letters, or use click the Show sequence button at the bottom.
Sequence length	The number of residues in the full protein sequence
Unobserved residues	The number of residues for which no 3D coordinates are available
Secondary structure source	The source of the secondary structure annotation (author, DSSP or none)
Number of contacts	The number of contacts in the contact map
Number of selected contacts	The number of contacts currently selected

Only in compare mode:

Number of unique contacts	The number of contacts unique to this structure
Number of common contacts	The number of contacts present in both structures
Contact map overlap	The overlap between the two contact maps measured by the Tanimoto coefficient [1]: $T(A,B) = \frac{\text{common contacts}}{\text{contacts}(A) + \text{contacts}(B) - \text{common contacts}}$

[1] Tanimoto, T.T. (1957) IBM Internal Report 17th Nov. 1957.

Loupe

Opens a window with a magnified view of the contact map around the current mouse position. This is helpful when working with large contact maps. The size of the loupe window and magnification can be set with the variables `LOUPE_WINDOW_SIZE` and `LOUPE_CONTACT_SIZE` in the config file.

Load from

CMView can import two general classes of data files:

1. Tertiary structure files (3D coordinates)
2. Contact map files

For tertiary structure files (e.g. PDB files), a contact map will be calculated from the 3D coordinates using the user specified contact type and distance cutoff, the structure will be automatically loaded into PyMol and the 3D visualization features will be available.

For contact map files (e.g. Casp RR files), only the main contact map window will be available.

Contact map files saved by CMView in its native text format are special in the sense that they store the PDB code of the structure they were derived from (if applicable). When opening such a .cm file, the application tries to retrieve the 3D coordinates from the online PDB.

Load from online PDB

Loads a structure from the PDB ftp site. The file is downloaded in mmCIF format and cached locally.

Load from PDB file / Load from Casp TS file

Loads a structure from a local PDB file. The file needs to contain an initial HEADER line and one or more ATOM lines following the PDB file specification [1]. Alternatively, files may also follow the CASP tertiary structure prediction format [2].

[1] <http://www.wwpdb.org/docs.html>

[2] <http://predictioncenter.org/casp7/doc/casp7-format.html#TS>

Load from contact map file

Loads a file in the native CMView contact map file format. A file saved in this format contains some metadata including the PDB-four-letter code of the structure it was derived from (if applicable). The corresponding structure will then be retrieved from the PDB when the file is loaded again.

Load from CASP RR file

Loads a contact map in CASP residue-residue contact prediction format.

See <http://predictioncenter.org/casp7/doc/casp7-format.html#RR> for the format specification.

Save to

Save to contact map file

Saves the current contact map in CMView contact map file format. The sequence, the residue-residue contacts and some meta data (original PDB four-letter code and chain code, contact type, distance cutoff) will be stored in the file.

Save to CASP RR file

Saves the current contact map in CASP residue-residue contact prediction format.

See <http://predictioncenter.org/casp7/doc/casp7-format.html#RR> for the format specification.

Save to PNG file

Exports the current contents of the main contact map window as a PNG image file.

Save Alignment to file

Saves the current alignment to a text file in FASTA alignment format. This function is only available in comparison mode.

Quit

Closes all windows and exits CMView.

Select

Selection modes

The selection mode defines the mouse behavior for selecting contacts in the main contact map window and in the secondary structure rulers at the left and at the top of the contact map window.

Rectangular selection mode

In contact map

- Drag to select rectangular region of contacts
- Drag to select rectangular region of contacts
- Click on contact to select single contact
- Ctrl+click to add to current selection
- Click on white space to reset selection

In ruler

- Click on secondary structure element to select the corresponding residues
- Drag to select range of residues
- Intersection of horizontal and vertical residue selections selects contacts

Fill selection mode

In contact map

- Click on contact to select cluster of contacts
- Ctrl+click to add to current selection
- Click on white space to reset selection

In ruler

- Same behavior as in rectangular selection mode

Diagonal selection mode

In contact map

- Click to select contacts on current diagonal
- Drag to select contacts in several diagonals
- Ctrl+click to add to current selection

- Click on white space to reset selection

In ruler

- Same behavior as in rectangular selection mode

Neighborhood selection mode

In contact map

- Click on residue pair (i,j) to select the neighborhoods of i and j, i.e. all contacts made by residue i or residue j.
- Click on diagonal to select neighborhood of single residue (This can also be done by clicking in the ruler. See below.)
- Ctrl+click to add to current selection
- Click on lower left half to reset selection

In ruler

- Click on residue to select its neighborhood
- Ctrl+click to add to current selection

Color selection mode

Select contacts which have been previously marked with the color function.

In contact map

- Click on contact to select all contacts having the same color
- Ctrl+click to add to current selection
- Click on white space to reset selection

In ruler

- Same behavior as in rectangular selection mode

Add/remove contacts mode

Add or remove individual contacts. Note that this will modify the contact map and possibly make it inconsistent with the loaded 3D structure. To exit this mode, choose one of the other selection modes.

In contact map

- Click on an existing contact to remove it from the contact map
- Click on a white cell to add the respective contact to the contact map

In ruler

- Same behavior as in rectangular selection mode

Select all contacts

Selects all currently visible contacts.

Select by residue number

Selects all contacts between elements from a user selected set of residues. The residue set is specified by a selection string. A selection string is a comma separated list of residue numbers and/or residue ranges where a residue range r1-r2 specifies all residues between and including r1 and r2.

Example: 1,3,5-7,10-12

Select helix-helix contacts

Selects all contacts between distinct alpha helices.

Select sheet-sheet contacts

Selects all contacts between distinct beta strands.

Select contacts between secondary structure elements

Selects all contacts between distinct secondary structure elements.

Select contacts within secondary structure elements

Selects all contacts between two residues from the same secondary structure element.

Color

Different regions of interest in the contact map can be assigned user-defined colors, e.g. to highlight different domains or a particular functional site.

See also: Select by color

Choose color

Chooses the active color for coloring contacts using the Paint selection contacts function.

Color selected contacts

Assigns the active color to the currently selected contacts

Reset contact colors to black

Resets all contacts to the default color (black)

Action

Toggle Real-Time Displaying of Contacts

Toggles the real time mode on and off. In real time mode, the current position in the contact map as well as the currently selected contacts are immediately shown as edges in the 3D viewer. This can be switched off to increase performance.

Show selected contacts in 3D

Visualizes the currently selected contacts as edges in the 3D viewer. The edges are drawn between the C-alpha atoms of the residues in contact.

In single contact map mode, two selection objects are created in the 3D viewer:

1. The set of contacts (drawn as solid yellow lines)
2. The set of residues participating in one of the contacts

These selection objects can be manipulated directly in the 3D viewer, for example to change colors or to display side-chain atoms.

In comparison mode, six contact objects and the corresponding residue selections are created:

1. Contacts in structure A which are present in both structures – drawn as solid yellow lines
2. Contacts in structure B which are present in both structures – drawn as solid yellow lines
3. Contacts in structure A which are present only in structure A – drawn as solid pink lines
4. Contacts in structure B which are present only in structure B – drawn as solid green lines
5. Pseudo-contacts in structure A (present in B but absent in A) – drawn as dashed pink lines
6. Pseudo-contacts in structure B (present in A but absent in B) – drawn as dashed green lines

Pseudo-contacts means that the distance between the two residues is above the defined distance cutoff. They are shown to allow easier analysis of conformational changes, where contacts have been lost or new contacts have been established.

Note: In PyMol 1.00 and above, above selections will be grouped together for a better overview. Earlier versions of PyMol do not support this feature.

Show threshold spheres for selected contacts in 3D

For selected contacts (i,j) draws spheres around residues i and j in the 3D viewer, where the sphere radius is the current contact threshold. This shows the sphere of influence of the individual residues under the chosen contact definition. Threshold spheres can also be drawn for residue pairs not in contact by right-clicking on the residue pair in the contact map and choosing Show Threshold Spheres For Residue Pair from the context menu.

Delete selected contacts

Permanently deletes the currently selected contacts from the contact map.

Run Cone Peeling Algorithm

Runs the Cone Peeling algorithm described in

Sathyapriya et al., Defining an Essence of Structure Determining Residue Contacts in Proteins. PLoS Computational Biology 5(12): e1000584 (2009).

This algorithm attempts to calculate a minimal subset of contacts which are sufficient to reconstruct the fold with distance geometry. See the reference for more information. This will modify the current contact map and show the minimal subset and the original contact map in compare mode.

Compare

Load second contact map

Loads a second contact map (in the following called B) to compare it to the currently open one (A). For the different input options see Load from

After a second contact map has been loaded, contacts are shown with the following color coding:

1. Black: contacts present in both structure A and structure B

2. Pink: contacts present in contact map A but absent in B
3. Green: contacts present in structure B but absent in A

To overlay the contact maps, an alignment between the residue of A and B needs to be defined. CMView offers the following options to obtain an alignment:

1. Needleman-Wunsch sequence alignment

Calculates a global sequence alignment using the classic Needleman-Wunsch-Gotoh algorithm with standard parameters (Matrix: BLOSSUM50, Gap-open: 10, Gap-extend: 0.5). The JAligner package [1] is used for the calculation. Use this for sequence identical structures (e.g. different conformations of the same protein).

2. Maximum contact map overlap structural alignment

Calculates a structure based alignment using the SADP algorithm [2]. It applies a very fast heuristic to maximize the contact map overlap between two structures. The algorithm is particularly well suited for aligning contact maps since it can be applied even if no exact 3D coordinates are available.

3. DALI structural alignment

Computes a structural alignment using DALI[3]. DALI needs to be installed locally and the parameter `DALI_EXECUTABLE` has to be set in the `cmview.cfg` config file. See installation for instructions on how to install DALI. This feature can only be selected if 3D coordinates for both structures to be compared are available.

4. Load alignment from FASTA file

Loads an alignment from a text file in FASTA alignment format. This format can be exported by many standard sequence alignment programs (e.g. Muscle, T-Coffee, EMBOSS, Bioperl). The sequences in the alignment file have to match the ones of the structures to be compared.

References:

[1] Ahmed Moustafa, JAligner: Open source Java implementation of Smith-Waterman, <http://jaligner.sourceforge.net> (2006/03/23).

[2] Jain, B.J. and M. Lappe (2007). Joining Softassign and Dynamic Programming for the Contact Map Overlap Problem; Springer Lecture Notes in Computer Science, S. Hochreiter and R. Wagner (Eds.): BIRD 2007, LNBI 4414, pp. 410-423.

[3] Holm, L. and Park, J. (1999). DaliLite workbench for protein structure comparison; *Bioinformatics* 16, 566-567.

Show common contacts

Shows/hides contacts present in both structure A and structure B.

Show contacts unique to first structure

Shows/hides contacts present in contact map A but absent in B (displayed in pink).

Show contacts unique to second structure

Shows/hides contacts present in structure B but absent in A (displayed in green).

Show difference map

Displays/hides the difference distance map of the two structures.

The difference distance map is a powerful tool to visualize conformational changes between two protein structures. For a pair of residues (i,j) the distance map visualizes how much the distance between i and j has changed from one structure to the other. This highlights regions of conformational change as red hotspots while regions which remain unchanged appear in blue.

For each pair of residues, the absolute difference between the distance in structure A and the distance in structure B.

The difference map can only be shown if 3D coordinates for both structures are available.

Superimpose from selection

Superimposes the structures in the 3D viewer based on the currently selected contacts.

When two structure are loaded for comparison, they are initially superimposed using an all-residue C-alpha minimum RMSD fit. In such a rigid-body superimposition, some regions may not align well in 3D. This feature allows to select a region of interest and to perform a minimum-RMSD fit only based on the residues participating in the contact selection.

Note: When changing the orientation of structures in PyMol (e.g. using the custom superposition function), earlier contact objects are not moved with the original structure. There is currently no workaround for this.

Show corresponding residues from selection

Visualizes aligned residues in the 3D viewer.

To highlights the correspondence between residues in structure A and structure B, lines are drawn between corresponding residues. This gives a visual three dimensional impression of the underlying alignment.

Help

Help

Shows the inline help.

About

Shows version and licensing information.

Context menu

Available by right-clicking on the main contact map window.

Show residue pair as edge in 3D

Draws a line in the 3D viewer labeled with the distance between the respective residues in Angstroms. This feature is independent of whether the two residues are in contact or not.

Show threshold spheres for residue pair in 3D

Draws the threshold spheres for the selected pair of residues in the 3D viewer. This feature is independent of whether the two residues are in contact or not.

Experimental features (unsupported)

Experimental features can be enabled by setting `USE_EXPERIMENTAL_FEATURES = true` in the config file.

Load structure from relational database

Lloads PDB structures from a relational database in PDBase[1] or MSD[2] format.

[1] http://openmms.sdsc.edu/OpenMMS-1.5.1_Std/openmms/docs/guides/PDBase.html

[2] <http://www.ebi.ac.uk/msd-srv/docs/dbdoc/>

Load from/save to graph database

Lloads and saves contact maps from/to a simple relational database with tables for graphs, nodes and edges.

Show common neighbor triangles in contact map

Shows the common neighborhood for a pair of residues as triangles in the contact map.

Show common neighbor triangles in 3D

Shows the common neighborhood for a pair of residues as triangles in the 3D viewer.

Show common neighborhood size map

Shows the common neighborhood sizes for all residue pairs as colors in the contact map.

Overlays

Overlays can be selected from the overlay menu in the side bar on the right. They display additional information in the contact map windows.

Contact density

Displays/hides the contact density map.

The contact density map shows for every residue pair (i,j) the density of contacts within the backbone fragment ranging from i to j. The density is defined as the number of contacts within the fragment normalized by the average number of contacts over all fragments of the same size. High densities are shown in red, low densities are shown in blue.

High-density regions often coincide with structural units like domains, secondary- and super-secondary structure elements.

Distance map

Displays/hides a distance map showing all pairwise distances between residues.

Distances are visualized as colors. The color coding depends on the distance cutoff of the current contact map. Distances above the cutoff are shown in blue, distances around the cutoff are shown in green and distances below the cutoff are shown in red, where the darkest blue is the longest distance and the darkest red is the shortest distance. Distances are calculated between C-alpha atoms.

The distance map can only be calculated if 3D coordinates for the structure are available.

Common neighbours

Displays/hides common neighbor relationships.

If two residues *i* and *j* are each in contact with a third residue *k* they are said to share the common neighbour *k*. Then by the triangle inequality one can infer that *i* and *j* can be no further apart than twice the contact distance cutoff (even if *i* and *j* are not in contact themselves). For predicted or incomplete contact maps this gives information that is not immediately visible in the raw contact map. For example, short range contacts (those, close to the diagonal) can often be inferred from long range contacts by common neighbor relationships.

The darker the shade of green in the overlay, the more common neighbours the particular residues share.

Command line parameters

The following command line parameters are available when calling CMView from the command line, help can be obtained by calling **cmview -h**:

- **-p <pdb code>**: a 4-letter PDB code to be loaded in CMView from Online PDB (Internet connection required).
- **-f <pdb file>**: a local PDB file to be loaded in CMView.
- **-c <chain code>**: a PDB chain code. This will be the chain selected either from the PDB file or for the given PDB code. If none is given the first chain in the file will be loaded.
- **-t <contact type>**: the contact type for the contact map to be loaded in CMView. See contact types.
- **-d <distance cutoff>**: the contact distance cutoff in Angstrom for the contact map to be loaded in CMView.
- **-o <config file>**: a file to read configuration parameters from. See Configuration.
- **-v**: print version and exit.
- **-h**: print help on command line parameters and exit.

Configuration

Overview

CMView can be customized using key-value pairs in a configuration file (`cmview.cfg`). The example file contains some commonly used settings. CMView is looking for configuration files:

1. In the current directory
2. In the user's home directory
3. At the location specified with the command line parameter **-o**

If the same variable is set in more than one file, the latest value takes precedence according to the above order.

Available configuration parameters

DEFAULT_CONTACT_TYPE The contact type shown in the load dialog by default and used for graph generation when loading structures with command line parameters **-p** or **-f**

Default: Cb

<i>DEFAULT_DISTANCE_CUTOFF</i>	The distance cutoff shown in the load dialog by default and used for graph generation when loading structures with command line parameters -p or -f Default: 8.0
<i>DEFAULT_MIN_SEQSEP</i>	The minimum sequence separation shown in the load dialog by default, negative values meaning no restraint Default: -1
<i>DEFAULT_MAX_SEQSEP</i>	The maximum sequence separation shown in the load dialog by default, negative values meaning no restraint Default: -1
<i>CONFIG_FILE_NAME</i>	The name of the configuration file Default: cmview.cfg
<i>DIST_MAP_CONTACT_TYPE</i>	The contact type being used to calculate the distance map Allowed values: Ca, Cb Default: Ca
<i>PDB_FTP_URL</i>	A URL where PDB structures are being loaded from. (See Load from Online PDB) This can be used to load structures from a local mirrored copy of the PDB Default: ftp://ftp.wwpdb.org/pub/pdb/data/structures/all/mmCIF/
<i>TEMP_DIR</i>	The directory for storing temporary files Default: The default system temp dir (e.g. /tmp or C:\Windows\Temp)
<i>INITIAL_SCREEN_SIZE</i>	The initial size in pixels of the main window on startup Default: 650
<i>LOUPE_WINDOW_SIZE</i>	The initial size in pixels of the loupe window Default: 200
<i>LOUPE_CONTACT_SIZE</i>	The size of each contact in the loupe window Default: 15
<i>SHOW_ICON_BAR</i>	Whether the toolbar with command shortcut icons is shown Default: true
<i>SHOW_PDB_RES_NUMS</i>	Whether legacy PDB residue numbers are shown in the lower left corner of the contact map below the secondary structure Default: true
<i>SHOW_RULERS</i>	Whether secondary structure rulers are shown (above and to the left of the contact map) Default: true

<i>SHOW_ALIGNMENT_COORDS</i>	Whether residue numbers in the alignment are shown in the lower left corner of the contact map (in brackets after the default residue numbers) Default: false
<i>USE_PYMOL</i>	Whether PyMol specific features should be shown in the menus. See also: <i>PYMOL_EXECUTABLE</i> Default: true
<i>USE_DSSP</i>	Whether DSSP is used for secondary structure assignment from 3D coordinates. See also: <i>DSSP_EXECUTABLE</i> Default: true
<i>DSSP_EXECUTABLE</i>	The path to the local DSSP executable
<i>DSSP_PARAMETERS</i>	Parameters passed to DSSP when called by CMView Default: --
<i>DALI_EXECUTABLE</i>	The path to the locally installed DALI executable to be used for structural alignments.
<i>PYMOL_EXECUTABLE</i>	The path to the local PyMol executable
<i>PYMOL_SHUTDOWN_ON_EXIT</i>	Whether PyMol will be shutdown when CMView is closed Default: true
<i>PYMOL_LOAD_ON_START</i>	Whether CMView will try to start PyMol on startup Default: true
<i>PYMOL_CONN_TIMEOUT</i>	The time in milliseconds that CMView will wait for PyMol to startup before reporting a time out error Default: 15000
<i>SHOW_CONTACTS_IN_REALTIME</i>	Whether currently selected contacts will be shown in PyMol in real time. This can be switched off to increase performance. In this case, contacts can be shown in PyMol manually. Default: true

Parameters affecting experimental features

<i>USE_EXPERIMENTAL_FEATURES</i>	Whether experimental features will be enabled. Experimental features are unsupported, probably not tested and are to be used on your own risk. Default: false
<i>SHOW_WEIGHTED_CONTACTS</i>	Whether contacts with weights between 0 and 1 are shown as shades of grey in the contact map Default: true

<i>USE_DATABASE</i>	Whether database specific features should appear in the menus Default: true
<i>DB_HOST</i>	The database server for structure/graph databases
<i>DB_PWD</i>	The password for accessing the database
<i>DB_USER</i>	The username for accessing the database
<i>DEFAULT_PDB_DB</i>	The default database for loading structures in PDBASE format
<i>DEFAULT_GRAPH_DB</i>	The default database for loading contact maps