

AstexViewer™@MSD-EBI:

Introduction

The viewer is designed to display coordinate and sequence information from macromolecules and small molecules. It can also display graphs of data in a number of different ways. The viewer is designed to allow coordination between the different views, any pick or highlight in one view will highlight the equivalent reference point in the other views.

- 1) Picking atoms, sequence residues or graph point will update all views
- 2) Multiple data can be brushed from each view
- 3) Colour can be queried from one view and dropped onto another view

The Viewer is run as an applet to display the results of a search from the MSD . It therefore has a number of properties that result from this use.

Groups of structures

The result of a search of the DB can result in a list of molecules, these molecules are grouped on whether they align both by structure and sequence. A group of structures are those that can be superposed by structure and sequence. A search of the DB can therefore result in one or more groups of structures to display, and the viewer has a drop down list of buttons to view each of these independently.

Visibility

The user has the ability to select structures from the result list of a search that they wish to see in the viewer. This visibility list is passed to the viewer and defines the default display of structures when each group is viewed. The show button in the viewer can be used to turn on and off visibility of structure within a group while in the viewer without referring back to the search results.

Load limit

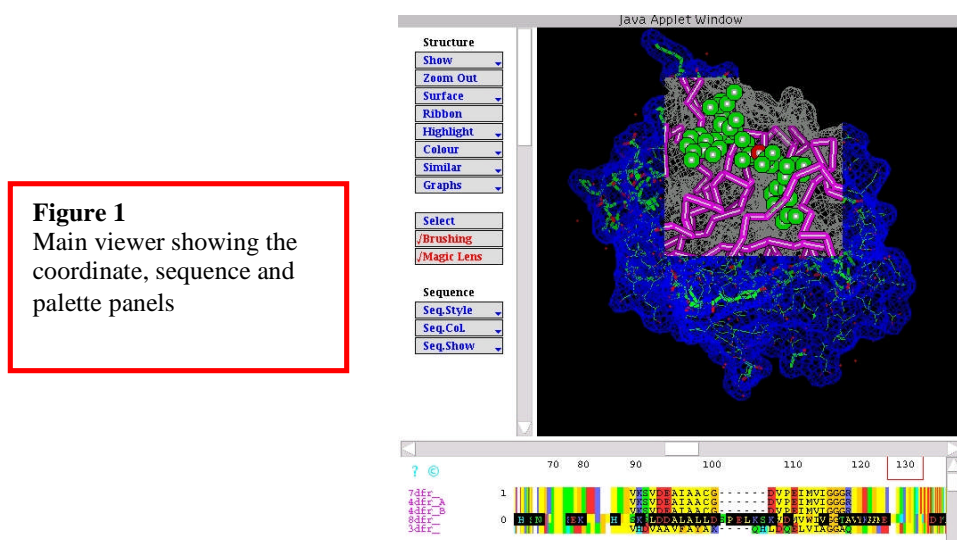
Since it not possible to know the resources available to a client process, the resources taken by the viewer is controlled by limiting the number of structures that can loaded at any one time. The viewer will only load a maximum of 8 coordinates sets at any one time, and loading an unloading of structure is automatically controlled by picking the list of names on the sequence panel.

General usage

The viewer

The main viewer consists of a single window that is divided into 3 panels (figure 1). The left-hand panel contains the palette of tools, the right-hand top region (black background) is the coordinate viewer, and the bottom is taken up with the sequence viewer. If the viewer size is adjusted then the palette width and the sequence height remains constant. Additional windows are drawn for graphs.

If the viewer is run as an applet it can only be removed by leaving/destroying the html page that started the applet; it can be hidden by using the [X] window tool on the viewer. The separate graphs can be removed using the [X] window tool on each window as required.



The palette.

The Palette panel consists of buttons and drop down buttons or when in “tree view” mode, consists of a foldable hierarchy view of buttons.

Button view:

A drop down button is shown within a triangle to the right of the button. If a drop down button is picked using <Left-mouse-click> then it remains dropped down until a second pick is made in the palette panel. It does NOT use a click and drag action. No action is performed on the palette or an open drop down if an area of the panel is clicked that is not a button. When a drop down button is open only the tools on the drop down are available and all other picks on the panel will close the drop down and result in no action. A <Shift><Left mouse click> selection will provide online help within a separate browser window for this tools.

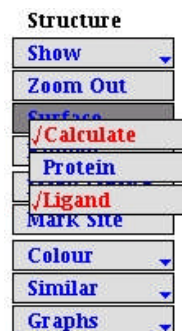
Tree view:

The tools are shown using a hierarchy view of the palette, where a tool name prefixed with [+] has a number of sub-tools that will be made visible when [+] is clicked. The same tool prefixed with [-] will

show the sub-tools, and can be folded closed by clicking the [-] prefix. The tool name with [+] or [-] prefix results in no action when clicked. A tool with no prefix is an action tool, and will result in a change in the viewer when clicked. These tools are referred to as buttons in subsequent text.

Any button that is picked that results in some functionality will change the cursor to the system dependent wait cursor until the function selected has completed. The only exception is Surface/Calculate that submits a surface calculation and does not wait for completion. Any button that controls a currently active function has red text with tick. Blue text is used for an inactive function or a button that does not have activity associated with it.

Figure 2:
Part of the palette from the viewer with an open drop down for the surface function. Calculate is active indicating that surfaces have already been calculated, and the ligand surface is currently displayed and the protein surface is not displayed.



The status bar

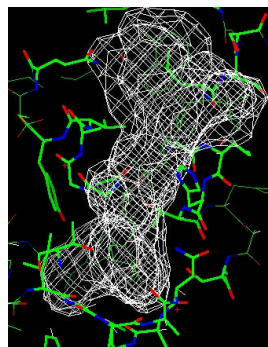
The viewer has a status bar that lies at the bottom of the main view window. This status bar is used to show a number of different pieces of information.

- 1) The number of structures that failed to load on start up
- 2) Tool-tips for each tool when the cursor is placed and left static over a tool
- 3) Text marks (as spheres) on the structure viewer
- 4) URL marks (as rotating sphere) on the structure viewer
- 5) Consensus sequence alignment value
- 6) Warnings and status remarks

The Structure Panel

The Structure view region uses functionality written by M. Hartshorn, AstexViewer™, Astex – although some additional features have been added. (Figure 3) The molecule is drawn as a bond-stick diagram, a ribbon or net surface. The view can be manipulated by a virtual track ball action (Left mouse button + drag of mouse); the molecule is rotated about $y=0$ and $x=0$ of the screen. The structure viewer is depth-cued and clipped to enhance the 3D effect, and perspective is not applied. The viewer can display objects by colour, text annotations and highlights as thick lines. The clipping planes can be adjusted using the “-“ and “+” keys.

Figure 3 : the structure viewer with ligand surface included



An atom/residue/chain/molecule is selected from the view using a left mouse button click close to any atom. An atom is any vertex at the end of one or more drawn bonds. The selected atom and parent objects are indicated using the status line at the bottom of the viewer window. Any selection from this or other panels will result in the structure viewer updating its view focus. The view focus can either be as whole molecule (zoom out), or as a single residue and ~8Å region about this residue. A selection using a mouse pick will either fly the molecule view or snap to the focus. A fly change has two modes. A change of focus to a region close by (less than 8Å) will result in a translation to this new region. A large change of focus will result in the view flying out to “zoom out”, then flying in to the new focus.

The structure viewer is designed to only load a maximum of 8 molecules at any one time. This is to prevent overload of both memory and CPU resources, as protein structure information is large. The loading and unloading of molecules is carried out using the sequence view panel. The left hand edge of the

Sequence Name	Load Number
seven_	9
eight_	10
nine_	11
ten_	12
eleven_	2
twelve_	1
thirteen_	0
fourteen_	-
fifteen_	-
sixteen_	-
seventeen_	-

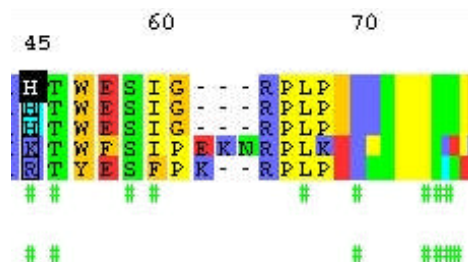
Figure 4
The structure loading data

the sequence panel contains the sequence name and structure number (figure 4). Any sequence entry where the structure number is indicated using a “-“ is not currently loaded into the structure viewer. The numerical value shown indicates the load order of the structures. Any sequence entry with no load number or “-“ has no associated structure, and selecting this name will no affect. Picking the name of an already loaded structure will update the load order, so that the picked name will become the newest one with a numerical designation of “0”. To load an unloaded structure into the structure viewer then the sequence file name (shown magenta in figure 4) is selected, a structure number will appear after the structure has been loaded. The structure with the oldest structure number (shown here as 12 for structure ten) will be unloaded from the viewer, and a “-“ will replace the number.

The Sequence Panel

The sequence panel displays sequence information as strings of letters on coloured boxes (figure 5). It is normal that upper case letters conform to the 1 letter code of the amino acid sequence, and lower case letters denote nucleic acid residues. A “*” character denotes an undefined residue such as a ligand and waters are not included within the view. The sequence viewer has no limit to the number of sequences that can be displayed, and all sequence within a single group will be displayed at any one time. Additionally, sequences not assigned to any group will also be displayed. The left-hand region of the sequence panel shows the molecule/chain designation name and the structure load number (figure 4). The viewer will also render point information (such as active site residues) and join information (such as disulphide bonds). The viewer can highlight a single residue, and a column of selected residues, the former as a black background letter, and the latter as black-outlined boxes. The viewer has two different display modes, flat scrolled and hyperbolic. With the former, the sequence is shown as a string that scrolls left and right and out of the viewer edge. Any information outside the viewer edge cannot be seen. A hyperbolic view show the sequence distorted using a *tanh* function, so that the central portion of the sequence and annotation is at full resolution, while the peripheral regions are progressively compressed. This view has the advantage that the entire sequences + annotations can be viewed all the time.

Figure 5, the sequence shown using the hypobolic view and active site marks.



The sequence view can be manipulated by a click and drag action (left mouse button), or using the scroll bars. A change of focus will place the selected residue at the centre of the display and highlighted it. A focus change can be carried out by selecting a residue from the structure viewer, sequence viewer or graph. Any change in the sequence focus will either snap to the new residue, or fly to the new residue by scrolling.

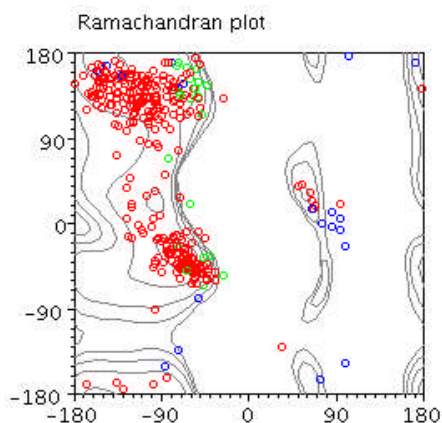
If a sequence entry is picked and this has an associated loaded structure, then the structure viewer and sequence viewer will update to this picked residue. If there is no associated structure then only the sequence viewer will update. If there is an associated structure, but it is not loaded within the structure viewer (as indicated by the load number “-“ see figure 4), then a warning will be issued, and only the sequence viewer will update. Unloaded structures are loaded by selecting the sequence name at the left of the view for the unloaded structure.

Sequence editing is performed using a <SHIFT><Left-mouse-button> click and drag action. To open a space within a sequence between residues XY (“----XY----“) then the letter “Y” should be clicked and dragged to the right. To close this gap again then the letter “Y” should be selected and dragged to the left. An insert action will insert “-“ characters into a sequence, and a gap deletion will remove these insert characters. A residue cannot be dragged beyond a position that results in no gap during gap closure. If multiple sequences are displayed then a normalized consensus score will be displayed within the status bar, and if the sequence is coloured using consensus colouring this will be updated during editing.

A Graph window

A user can draw a graph containing data from the current visible molecule using the graph button drop down and this will appear in a separate window. This window will contain the requested graph, and can be closed using the [X] window button on the top border. The user cannot control the graph display type at the current time. The graph display supports highlights, colour and text annotations. Any selection from the sequence, structure or graph will highlight the residue selected using a solid circle, and add an annotation to the base of the graph indicating the molecule, chain and residue selected. Graphs of data will be generated for all molecule VISIBLE at the time the graph was generated.

Figure 5. Ramachandran graph



The select palette

The select palette allows the user to make selections of atoms/residues/atoms and chains using a scrolling list of objects. Each object (atom/residue/chain/molecule) can be opened and closed using the [+] and [-] icon on the palette, while the selection toggle for an object can be turned on and off by clicking the object name. The selection will change the 3D coordinate view and allow the colour, display and highlighting to be changed. The palette tree of objects has the following properties :

- 1) An orange object is selected and a blue object is un-selected
- 2) Selecting an object (eg chain) will reset ALL sub-objects (ie residues and atoms) to the same state of selection as the object. To select all the residues within a chain then select the chain object name that contains these residues.
- 3) A range of objects (ie residues) is selected by selecting the first object in the range, then selecting the last object in the range while holding down the <SHIFT> keyboard button. This action is the same as that used within Microsoft explorer and icon views. Any object within this range will have its state toggled, so if selected then it will become un-selected and visa-versa.
- 4) Highlighted objects are shown selected on the 3D-atom display using yellow select boxes and on the sequence using magenta select boxes.

For example, to highlight and colour yellow all residues in the first chain except range 40 to 50 and then draw a Ramachandran plot with this colour scheme for just the first molecule – and then apply this to the sequence display.

- 1) Unfold the residue information from the first chain of the first molecule.
- 2) Pick the 1st chain object name – this will select (colour orange) all the residues in the 1st segment.
- 3) Pick name of the 40th residue – this will turn blue to indicate it is not selected.
- 4) Move the mouse to the 50th residue and <SHIFT> mouse click the name of this residue – this will select the range 40-50 and set there select colour to blue (un-selected)
- 5) Click the **highlight** button from the select palette.
- 6) Click the main palette **colour** drop down, and select the sub-button **yellow**.

Now all the residues in the first segment should be coloured yellow and highlighted except for those in the range 40 to 50. Now to draw a Ramachandran using this colour scheme for the first molecule, and change the sequence display to show this :

- 1) If there is more than 1 molecule displayed, use the **show** drop down to change the display of the molecules so that only the first is showing. Any graph produced will now only be for the first molecule.
- 2) Click the **graph** drop down button and select **Rama** to produce a graph – the colour scheme will be the same as the colour scheme of the 3D display.
- 3) Click the **Seq.Col** drop down on the main palette and select **Structure** and the sequence for the first molecule and chain will be coloured with the same colour scheme.

If the brushing is active set to on (**Brushing**) then the selection palette can be used to brush all the data views. This brush will result in the coordinate/sequence/graphs displays to be highlighted for any object that corresponds to the object row that the mouse traverses within the Selection palette. If the mouse cursor lies over a residue then all data for this residue will be highlighted; if the mouse is over a chain name then all the data in this chain will be highlighted, and similar for molecule.

This palette has three control buttons that act on the current selected objects within the scrolling panel.

Select/ Display On

The **Display On** button will set the current display status of the selected objects so that they are visible. This button only affects the atom display state, and has no affect on the molecule display state with is set using the **show** drop down button on the main palette.

Select/Display off

The **Display Off** button will set the current display status of the selected objects so that they are not visible. This button only affects the atom display state, and has no affect on the molecule display state with is set using the **show** drop down button on the main palette.

Select/Highlight

The **Highlight** will set the current selected atoms in the structure and sequence views so that they are drawn with highlights. The structure highlighting can be reset from the **highlight** drop down button on the main palette while selecting any sequence residue resets sequence highlights.

Buttons and their action

Tree View / Button View

This button changes between the drop down button palette and the hierarchy tree view palette.

Group

The group drop down button is only available if more than 1 group has been supplied to the viewer. The initial group displayed is the first group containing visible molecules. Selecting a different group will result in changing all the structures and sequences displayed to that designated by the grouping from the search. If a group is selected is that contains no visible molecules (selected from the search results) then a warning is added to the status line. The view of the structure and sequence is reset on picking a group drop down button and can be used to reset the view of the current displayed group by picking this current group button (for example, colour, visibility, centre and sequence edits. The user cannot change the group designation within the viewer as this is a server side designation.

Reset

The reset button is only available if there is only a single group of structures and will reset all colouring, annotation etc to the initial state – this includes any server side annotation and the sequence edits.

Show

The show drop down button will contain a list of buttons of all the currently loaded structure names and “none” + “all”. Picking a selected drop down button will toggle (hide/show) the structure view of that name. The all and none show drop down buttons will turn on/off respectively the visibility of all the structures. The names on the drop down buttons will be changed as structures are loaded/unloaded as the new-loaded structure is added to the show list and the unloaded structure is removed.

Zoom out

The zoom out button will scale the structure view so that all the molecules in the display will fit within the structure view panel. If the view is already zoomed out then the view is reset so that the displayed molecules fill the structure panel otherwise the view will fly/snap from a residue focus to the full molecule focus.

Surface

The surface drop down button is used to calculate and display surfaces for loaded molecules. The first button on the drop down list is calculate, this will submit a background process to calculate separate surfaces for each molecule and each ligand of each molecule. The status line at the bottom of the structure viewer indicates the progress of the calculation and on completion the status text is removed. Note, that since a surface is calculated separately for each ligand and macromolecule there will be two calculations per molecule. The ligand surfaces are calculated first. Additional use of this button will check if any additional surfaces need to be calculated, and only calculate those not yet available. This will be the case if the user has changed groups, or loaded additional structures within a large group.

The ligand drop down button will toggle on and off the display of ligand surfaces for the currently visible molecules, and the protein drop down button will toggle on and off macromolecule surfaces for the currently visible structure. It is not possible to display a surface if it has not yet been calculated.

Since the protein and ligand buttons will only display surfaces for visible molecules at the time of the button press. Hence, the display of particular surfaces should made using **surface/protein** , and then the molecule display altered to show the bonded structure of interest.

Ribbon

The ribbon button will generate ribbons for all loaded molecules and then display only those with visible coordinates. Additional use of the button will toggle on and off the ribbons. Each use of the **ribbon** button that turns on the display will check for uncalculated ribbons, this will be the case if the user has changed the group or loaded additional molecules within a group.

Highlight

The highlight drop-down button allows different pre-defined selections of atoms to be drawn with a thick line. Any previous highlighting with thick lines is superseded.

- 1) **None** will reset any thick bone line highlighting so that none is shown.
- 2) **Ligand** will highlight any non-protein, non-nucleic acid, non-water residues.
- 3) **Active Site** will highlight any whole residue with at least one atom within 4Å of a ligand residue.
- 4) **Backbone** will highlight the atoms CA, N, C and O.
- 5) **All** will highlight all the atoms.

Colour

The colour drop down button is used to change the colouring of the currently displayed molecules. If no atoms are selected (using the select palette) then all the atoms are coloured by the sub-button otherwise the colour is set for the current selected atoms. (note that if the colour change does not appear to work then it is likely that atoms are selected within the structure display. Clicking with the left mouse button in any background area of the structure display will clear this atom selection)

- 1) **Element** will colour atoms so that carbon = green, nitrogen = blue, oxygen = red, sulphur is yellow.
- 2) **Bvalue** will colour the atoms using a colour scale (red – white – blue) based on the Bvalue for each atom. Red is used for atoms with a large Bvalue, and blue for atoms of low Bvalue
- 3) **Chain** will colour each chain of each molecule a different colour.
- 4) **Sequence** will colour the all atoms of a residue the same colour as the equivalent residue in the sequence table. For example, if the sequence is currently coloured using sequence alignment consensus then the colour of each residue within the sequence will be used for all the atoms of a residue in the structure.
- 5) **Red, Blue...** will change the colour to the specified single colour. Of most use when used in conjunction with the select palette.

Similar

The similar drop down will carry out a similarity analysis for all overlaid molecules currently displayed.

- 1) **Join** will generate a white line segment to join all CA atoms of different molecules within 1.5Å of another molecule
- 2) **Col.Freq.** Will colour each molecule residue by a normalized frequency value. The frequency is defined as the proportion of the currently displayed molecules that residue is in close contact with (1.5Å). This is probably of most use when multiple (more than 2) molecules are overlaid. The colouring will give a consensus similarity of residue position over all the molecules. A blue residue is most similar over all molecules, and a red residue is most dissimilar.
- 3) **Col.Dist** will colour each residue of the displayed molecules with a colour proportional to the minimum distance to other equivalent residues in other molecules. This display is of most use when viewing two molecules, and blue atoms are close, and red atoms are 1.5Å or more away from an equivalent residue.

Graphs

The graphs drop down will contain a list of possible graphs that can be drawn based on the currently visible molecules.

- 1) **Rama** will display a Ramachandran plot of the currently visible molecules. The default colouring is for glycine residues to be drawn blue, proline residues to be drawn green, and all other residue type to be drawn red.
- 2) **Bvalue** will display a line graph of mean Bvalue/residue for each residue of a macromolecules; for all displayed macromolecules. Each line graph, for each macromolecular will be draw in a different colour.
- 3) **Omega** will display a graph of protein torsion angle omega for each residue of the protein. Each displayed molecule will result in a graph line of different colour.
- 4) **CAdist** will display a CA distance matrix for the first displayed molecule. The distance matrix will be contoured so that blue is a small separation and red is a large separation; the highest contour drawn represents a separation of 25A.
- 5) **Residue (n) / Atom (n)** If additional data have been supplied to the viewer using the object language then additional graphs will be available from this drop down.

Select

The select button is a toggle button the opens and closes the data tree view window. This window allows the user to select/highlight/colour/display individual atoms/residues/chains/molecules within in the structure display.

An [+] mark in the data tree will unfold the sub-object within this object; ie if residue, then this will unfold the atoms within this residue.

An [-] mark will close up a fold.

Picking the name of the object (ie ALA) will select all the atoms within this residue.

Picking the name of an object (ie ALA) will set the selection all members of this object (ie Atoms).

A range of objects can be selected as "first pick" + <SHIFT> "second pick"

Brushing

Brushing sets up the interactive highlighting in each of the views. Moving the cursor over each of the different views will highlight all the other views with the selected range. The selection is done with a rectangle in the sequence view, and an ellipse within the graph views, a cylinder in the structure view, and a single point in the select view.

The size of the selection area is changed using a <SHIFT>mouse-drag action. That is, while holding the <SHIFT> key down and the mouse button down the mouse should be moved. The highlight shape will be stretched in x/y as the mouse is moved away from the initial starting click point. The sequence rectangle size has discrete values of size defined by single sequence character width and height. The minimum size is a single sequence and residue. The ellipse in the graphs can have any value larger than 0.0001.

On leaving a graph/sequence view all highlighting is removed by default. The selection in a view can be locked using a <SHIFT> move mouse out of window option. Once the cursor is out of the window the selection will remain locked and visible within the other windows until the cursor is moved back into the original view used to make the selection. For example, the residue that lie in a helical conformation can be selected using the ellipse brush from the Ramachandran plot; holding the <shift> key down as the mouse is moved out the ramachandran plot will hold this selection and allow the manipulation of the structure/sequence views with highlights. Moving the mouse cursor back into the Ramachandran plot will reset the brush.

Magic Lens

The magic lens will place annotation on the 3D coordinate view within window. The default lens shows solid ribbons for proteins with HET groups shown as CPK. If any server side annotation is defined then this lens annotation will be different. The window will follow the mouse cursor position within the structure view. The key “<” will make the window larger and the “>” key will make the window smaller. Any view manipulation (ie clip plan adjustment and virtual track ball) will hide the window temporarily

Seq.Style

The sequence style drop down will change the sequence display type between 5 different modes. The Block and Group styles are toggle button pair and the Linear and Hyperbolic are a toggle button pair.

- 1) Group will arrange the sequence information so that all the sequence strings are displayed together, and the active site markers are arranged together in groups.
- 2) Block will arrange the sequence information so that sequence strings and site information pair is arranged together for each sequence.
- 3) Linear will display the sequence information so that a flat scrolling display is used.
- 4) Hyperbolic will display the sequence information using a tanh function so that the centre as if the sequence is drawn on a cylinder.

Seq.Col.

The sequence colour button allows the sequence to be coloured using a number of different colour schemes. If no residues are selected then all residues in all sequences are coloured as defined by the option. If a selection has been made with the selection palette then only these residues will be coloured.

- 1) **Residue** : A default residue colouring
- 2) **Charge** : ASP/GLU = red, ARG/LYS = blue, HIS = cyan, rest = grey
- 3) **Hydro** : A colour range (blue – white – red) based on a hydrophobicity scale. Red = hydrophilic, Blue = hydrophobic
- 4) **Consensus** : A colouring based on the consensus alignment over all the sequences displayed using a colour range from blue – white – red. Blue is used for columns with no residue type variance, and red is used for columns with highly variable residue types.
- 5) **Structure** : This changes the colour scheme so that the residue within the sequence have the same colour as the equivalent residue within and referenced structure. If the atoms have different colour then the residue is set as the most occurring colour atom within the residue.
- 6) **Red/blue...** changes the sequences/selection to the specified colour. Of most use when used in conjunction with the selection palette.

Seq.Show

Adds additional information to be added to the sequence display.

- 1) **Disulphide** will add lines between cystine residues that are connected by a disulphide bond within the coordinate structure
- 2) **Site** will add a # mark for each residue that is in contact with a ligand in the coordinate structure
- 3) **Cons. On Edit** will toggle the consensus alignment calculation during sequence editing. That is, when not active (default) then any sequence edit will not return the normalize consensus score and not change the colour as the sequence is dragged. If active, then the consensus score and colour is interactively recalculated as the sequence is dragged.