

# User Guide for the HELM Editor II

Last updated: January 18<sup>th</sup> 2018

## Contents

|  |    |
|--|----|
| Contents.....                                  | 1  |
| Intended Audience.....                         | 3  |
| Quick reference guide to the main screen.....  | 4  |
| Monomer Browser.....                           | 5  |
| Filtering the monomer set.....                 | 5  |
| Using Monomer Favorites.....                   | 6  |
| Browsing the different monomer categories..... | 7  |
| Rules.....                                     | 10 |
| Canvas.....                                    | 11 |
| Viewing area.....                              | 12 |
| Sequence viewing.....                          | 12 |
| Adding monomers via the sequence tab.....      | 13 |
| HELM notation.....                             | 14 |
| Properties.....                                | 14 |
| Structure View.....                            | 15 |
| Creating a new molecule.....                   | 16 |
| Using the drawing Canvas.....                  | 16 |
| Nucleotides.....                               | 17 |
| Using unknown monomers (Peptide and RNA).....  | 19 |
| Alternative monomers and mixtures.....         | 20 |
| Alternative polymers and mixtures.....         | 23 |
| Monomer Annotations.....                       | 25 |
| Unknown simple polymers.....                   | 25 |
| Connection ambiguity.....                      | 27 |
| Repeat sequences.....                          | 28 |
| Conjugates.....                                | 29 |

|   |    |
|---|----|
| R-Groups .....                              | 29 |
| Working with structures .....               | 32 |
| Building a molecule using the sequence..... | 33 |
| Cyclizing a sequence .....                  | 34 |
| Saving and loading monomers.....            | 35 |
| Helm notation .....                         | 35 |
| Exchangeable HELM notation (xHELM) .....    | 35 |
| Managing monomers.....                      | 36 |
| Using the Monomer Library .....             | 36 |
| Monomer List Menu .....                     | 36 |
| Monomer List.....                           | 37 |
| Monomer information .....                   | 38 |
| Managing Rulesets.....                      | 39 |
| Using the Ruleset Manager.....              | 39 |
| Rule Set .....                              | 39 |

## Intended Audience

This document explains the essential functions of the HELM editor.

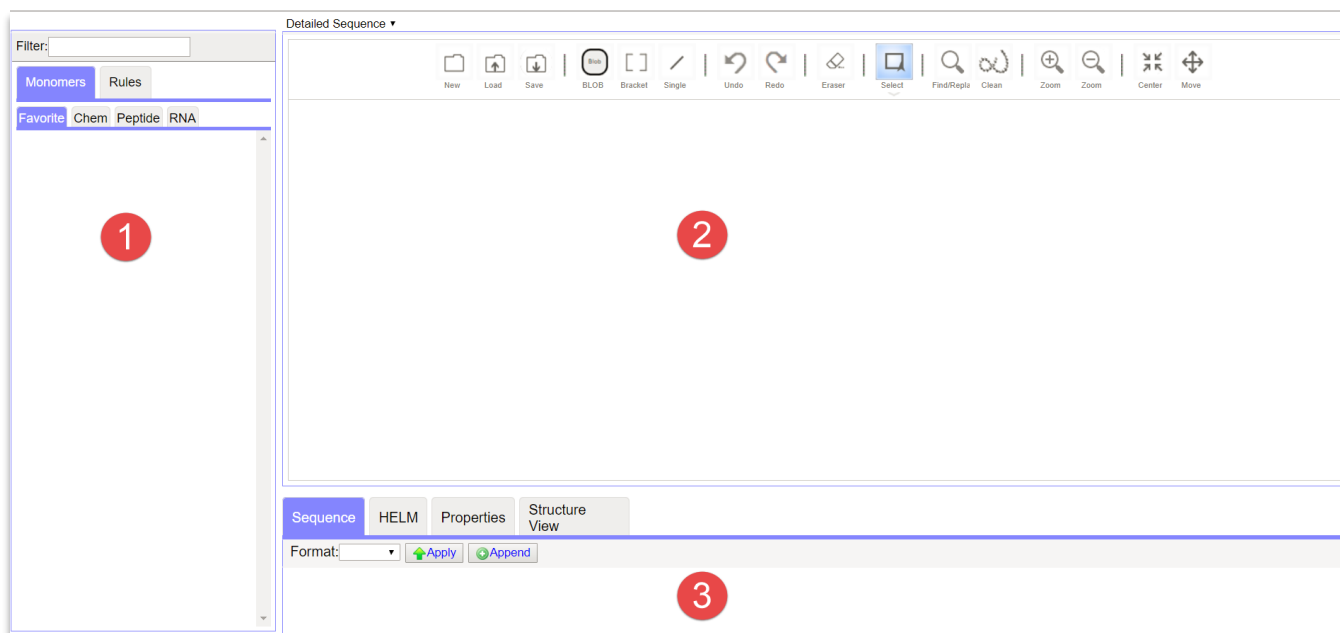
It assumes familiarity with the basics of the HELM notation and common types of macro-molecules, such as, nucleotides and peptides.

For more information on the HELM notation, please see the HELM website:

[www.openHELM.org](http://www.openHELM.org)

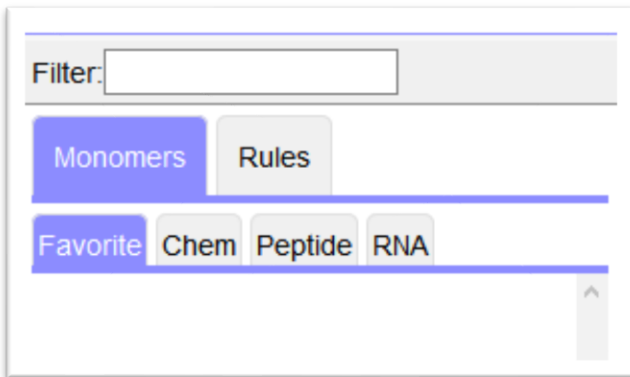


## Quick reference guide to the main screen



1. Monomer browser
  - a. Provides access to monomers required to build macromolecules.
2. Canvas
  - a. Drawing area where you can place monomers and build macromolecules.
3. Viewing area
  - a. Provides alternative views of the molecule such as natural analog sequence, HELM, properties such as mol wt and the atom-bond structure.

## Monomer Browser

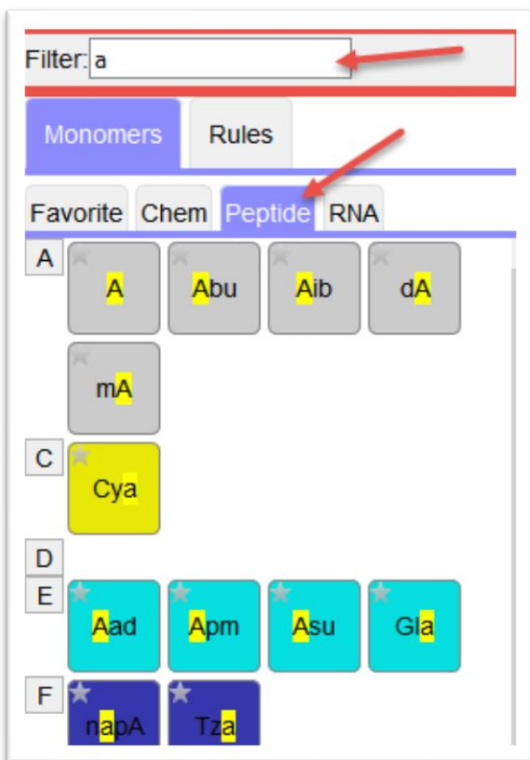


The monomer browser provides access to the monomer library, which is divided into chemical monomers, peptide and nucleotide monomers. The tabs are used to browse the library.

Frequently used monomers can be displayed in the favorites tab and frequently used functions can be saved under the Rules tab.

### Filtering the monomer set

The Filter box can be used to search for monomers.

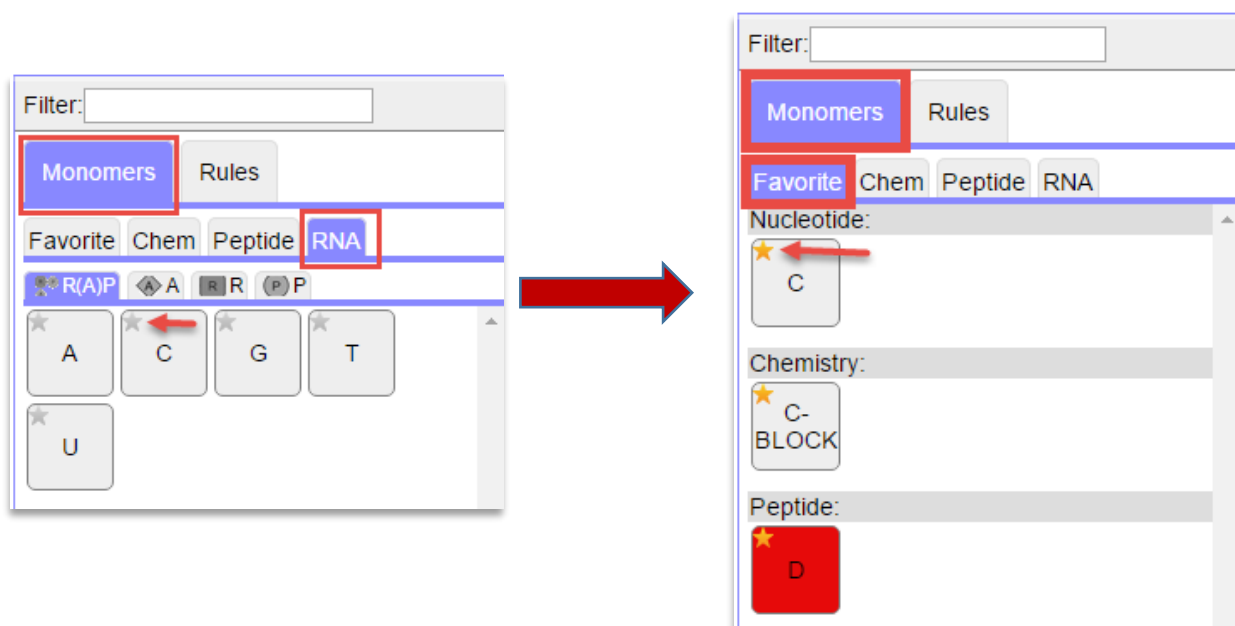


To search for a monomer, choose the tab for the type of monomer, then, in the Filter box, type letters or numbers from the monomer symbol.

**Note:** Entering lower case letters will display items with either upper or lower-case letters anywhere in the code.

## Using Monomer Favorites

Frequently used monomers can be saved under the Favorite tab.



To create a favorite:

- Select the desired monomer under the Chem, Peptide or RNA tabs.
- Click on the star symbol in the upper left corner of the monomer icon.
- The star will turn from grey to yellow and the monomer icon will be saved in the Favorite tab.

Click the yellow star to deselect a Favorite.

## Browsing the different monomer categories

### *Chem and peptides*

Choose the Chem or peptide tab under the Monomer tab to display the different lists. The monomers will be displayed in the area below. Peptide structures are categorized according to their natural analog.

The screenshot shows a web interface for selecting monomers. At the top, there is a 'Filter:' input field and a 'Detailed Sequence' dropdown menu. Below these are two tabs: 'Monomers' (highlighted with a red box) and 'Rules'. Under the 'Monomers' tab, there are four sub-tabs: 'Favorite', 'Chem' (highlighted with a red box), 'Peptide', and 'RNA'. A grid of monomer icons is displayed, each with a star icon in the top-left corner. The icons are labeled: R, A6OH, Alexa, Az, C-BLOCK, Cys-BLOCK, EG, hxy, MCC, N-BLOCK, PEG2, sDBL, SMCC, SMPEG, and SS3. A red arrow points to the 'Az' icon. A tooltip window titled 'Azide' is open over the 'Az' icon, showing the chemical structure of an azide group attached to a chain, with a red oxygen atom and a blue nitrogen atom. Below the structure, it says 'R1= OH'.

A tooltip of the chemical structure will appear by mouse-hovering over the chemical icon.


Filter:

Monomers Rules

Favorite Chem Peptide RNA

|   |   |    |     |     |
|---|---|----|-----|-----|
| A | A | dA | mA  | meA |
| C | C | dC | meC | seC |
| D | D | dD | meD |     |
| E | E | dE | meE |     |
| F | F | dF | meF |     |

**N-Methyl-Alanine**



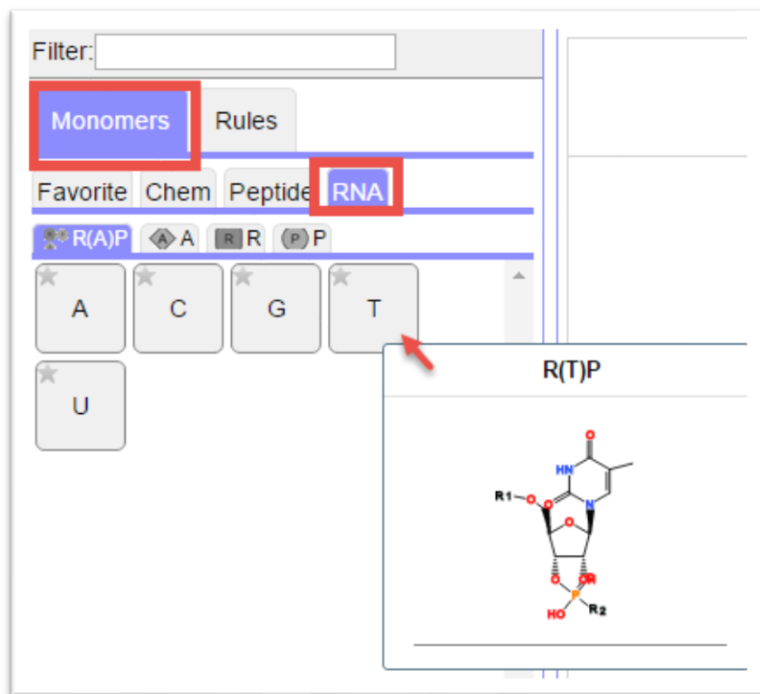
R2= OH  
R1= H

The image shows a software interface for selecting nucleotides. A grid of buttons is organized by base (A, C, D, E, F) and modification (A, d, m, se). The 'Peptide' tab is active, and a tooltip for 'N-Methyl-Alanine' is displayed over the 'meA' button. The tooltip shows the chemical structure of N-methyl-alanine with R1 and R2 labels, and specifies R2=OH and R1=H.

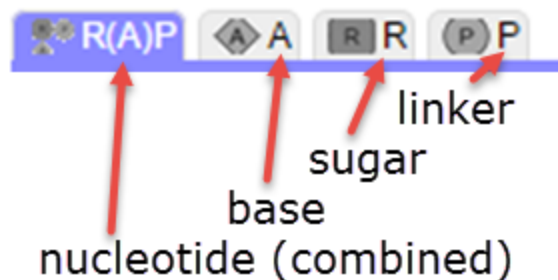
RNA

Nucleotides have subcategories as well as the main tab.





Choose the RNA tab under the Monomer tab to display the nucleotide list.



There are four tabs under the RNA tab:

- 1) R(A)P: Choose this tab to display the nucleotide (combined) options.
- 2) A: Choose this tab to display all the bases.
- 3) R: Choose this tab to display all the sugars.
- 4) P: Choose this tab to display all the linkers.

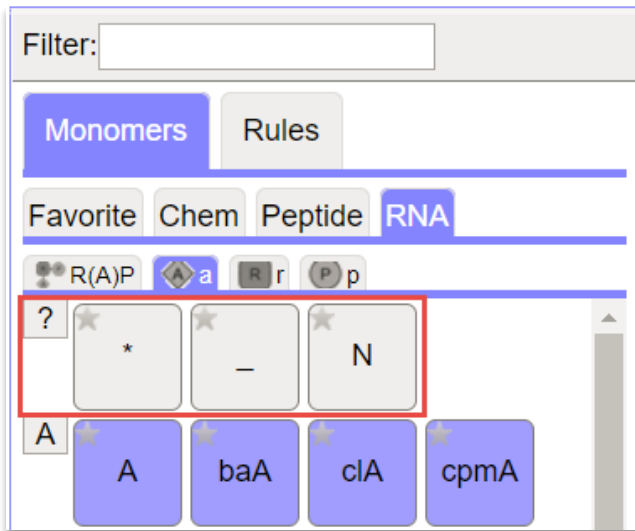
### *Monomer Ambiguity*

You can add ambiguous monomers where the structure is not known. There are 4 different types:

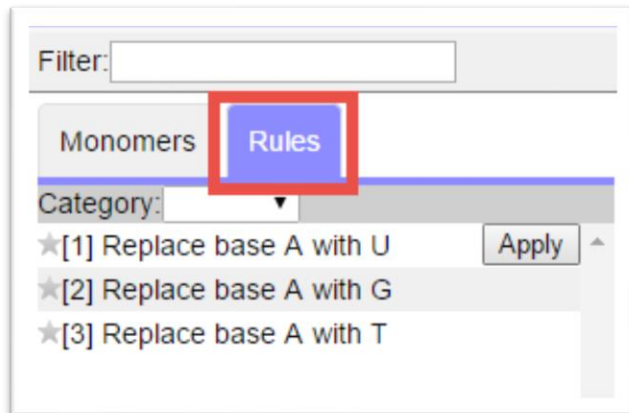
- The character "\*" represents 1 or more unknown monomers
- The character "X" represents one single unknown amino acid in a PEPTIDE polymer.

- The character “N” represents one single unknown base in a RNA polymer.
- The character “\_” represents a deleted or missing single monomer. This is typically used for list elements

These monomers can be found in favorites section of the monomer palette (only the appropriate monomers for the polymer type). E.g.



## Rules



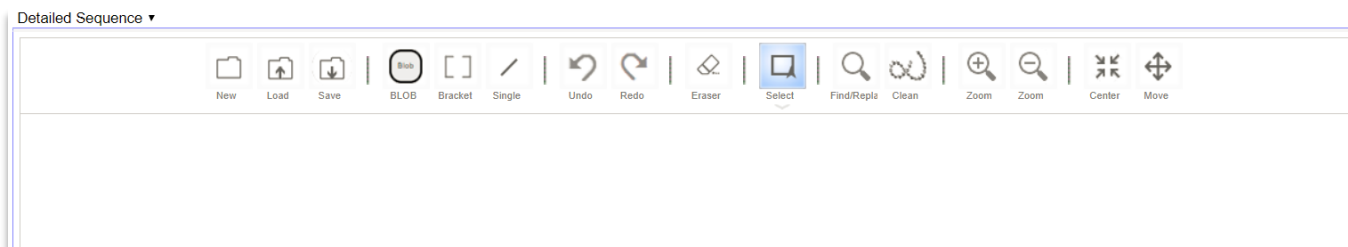
Choose the Rules tab to display preset rules.

Select a Category of preset rules from the Category dropdown menu.

Click the Apply button to apply the rule.

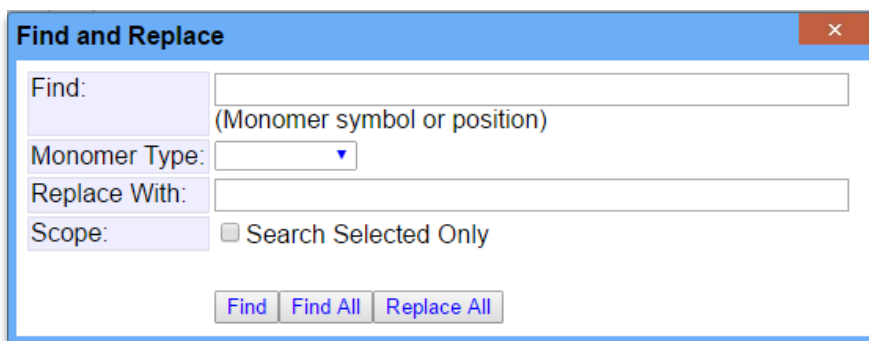
## Canvas

The canvas is the main drawing area of the editor and contains a number of buttons to help the user construct a macromolecule.



### Buttons:

- New: Clears the contents in the canvas.
- Load: Loads a HELM or xHELM file.
- Save: Save contents as a HELM or xHELM file.
- BLOB: Used as unknown simple polymer type.
- Bracket: To select a monomer or fragment and specify repeating units.
- Single: Single bond connect.
- Undo/Redo: To undo or redo a change on the canvas.
- Eraser: Erases a selected monomer.
- Select: Dropdown menu with Select, Lasso, Fragment or All tools to capture some or all of the monomers displayed on the canvas.
- Find/Replace: To find a monomer or to replace a monomer with another.

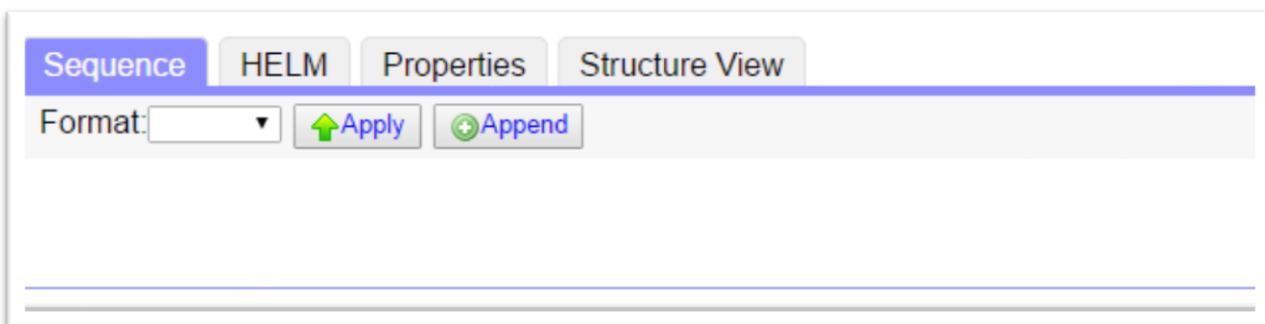


- Layout: Cleans up monomer display.
- Zoom +/-: Makes selected monomers larger or smaller.
- Center: Centers monomer display on canvas and displays zoomed monomers at default size.

- Move: Moves all monomers around the canvas. Also used to enlarge an area of a structure. Select move and use the mouse-wheel to zoom in and out on a structure. Mouse-click on area on which to focus.

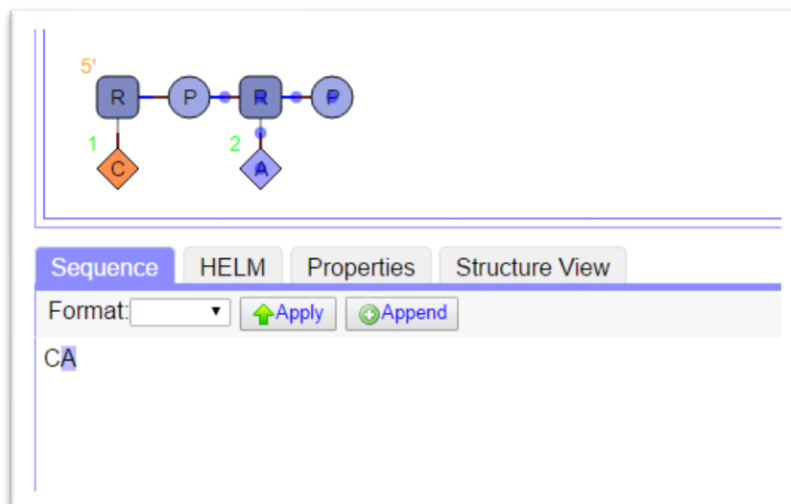
Choose Detailed Sequence/Sequence from the dropdown menu to switch the display of the Canvas and Viewing Area.

## Viewing area



See a sequence view of a molecule or the component view that provides details such as Mol. Wt., Mol. Formula etc.

## Sequence viewing

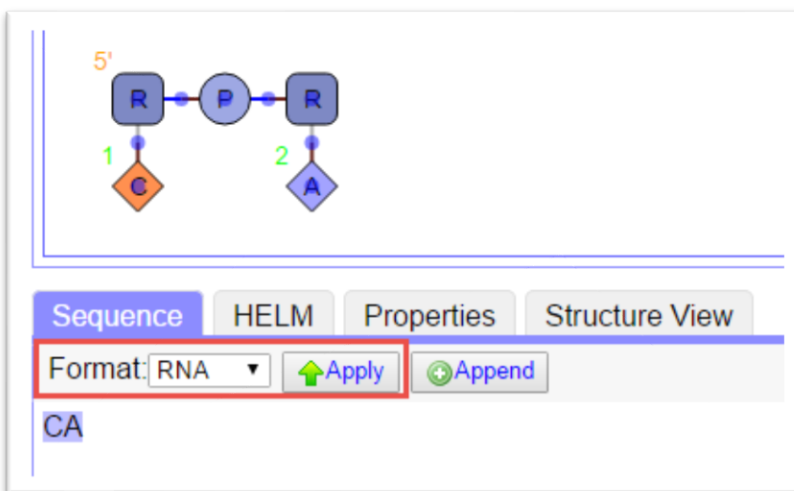


Choose the Sequence tab to display the sequence of monomers drawn on the canvas.

In the example sequence view the Nucleotide 'A' has been selected on the canvas and the corresponding nucleotide in the sequence display is highlighted.

### Adding monomers via the sequence tab

A sequence can also be imputed directly by writing the sequence in the Sequence Viewing area or copying and pasting a sequence into the Viewing area and clicking Apply. Choose the format before applying!



OR



You can add to an existing structure using the append function. This will add the new sequence to the canvas and you can then attach them together using normal drawing functionality.

The image shows a software interface for editing a chemical structure. At the top, there are two identical diagrams of a dinucleotide. Each diagram consists of a 5' phosphate group (p) connected to a ribose sugar (r) which is bonded to a cytosine base (C), and another phosphate group (p) connected to a ribose sugar (r) which is bonded to an adenine base (A). The bases are labeled '1' and '2' respectively. Below the diagrams is a control panel with four tabs: 'Sequence', 'HELM', 'Properties', and 'Structure View'. The 'Sequence' tab is currently selected. Below the tabs, there is a 'Format' dropdown menu set to 'RNA', and two buttons: 'Apply' and 'Append'. The 'Append' button is highlighted with a red rectangular box. Below the buttons, the sequence 'CA CA' is displayed.

## HELM notation

Choose the HELM tab to display the HELM notation. The notation for the sequence 'CA' is displayed with 'A' highlighted because this is highlighted on the canvas.

The image shows the same software interface as above, but with the 'HELM' tab selected. The 'Sequence' tab is now greyed out. The 'Apply' and 'Append' buttons are still visible. Below the buttons, the HELM notation for the sequence 'CA CA' is displayed: `RNA1{R(C)P.R(A)P}$$$$V2.0`. The 'A' in the notation is highlighted in blue.

## Properties

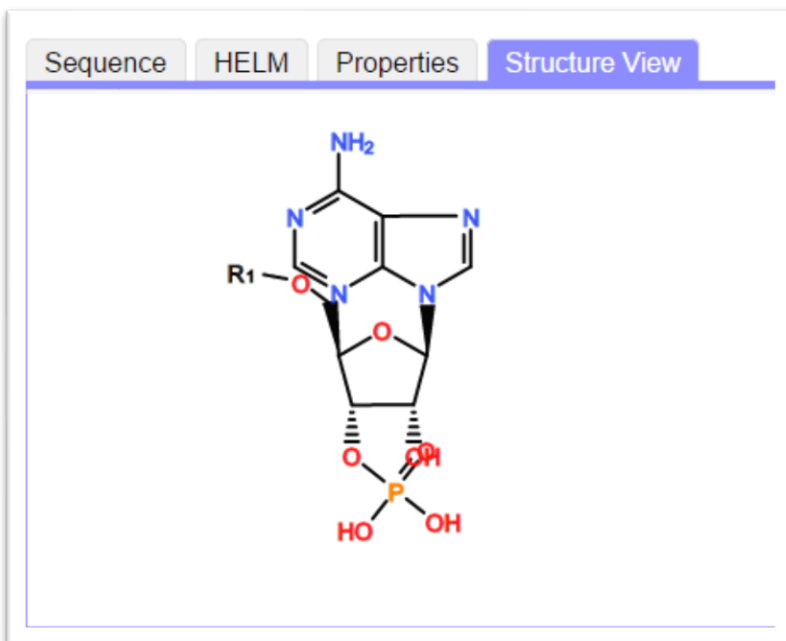
Choose the Properties tab to display the Molecular Weight, Molecular Formula and Extinction Coefficient for the entire sequence.

Sequence HELM Properties Structure View

Molecular Weight: 652.4  
Molecular Formula:  $C_{19}H_{26}N_8O_{14}P_2$   
Extinction Coefficient: 22.94

## Structure View

Choose Structure View to see the structure of the selected monomer (in this case A).



The Viewing area can be enlarged by dragging the line separating the Canvas and Viewing area up or down.

All or part of a structure can be viewed by using the Select button to choose the area to view.

## Creating a new molecule

### Using the drawing Canvas

Explore all the available monomers under the 'Chem', 'Peptide' and 'RNA' tabs.

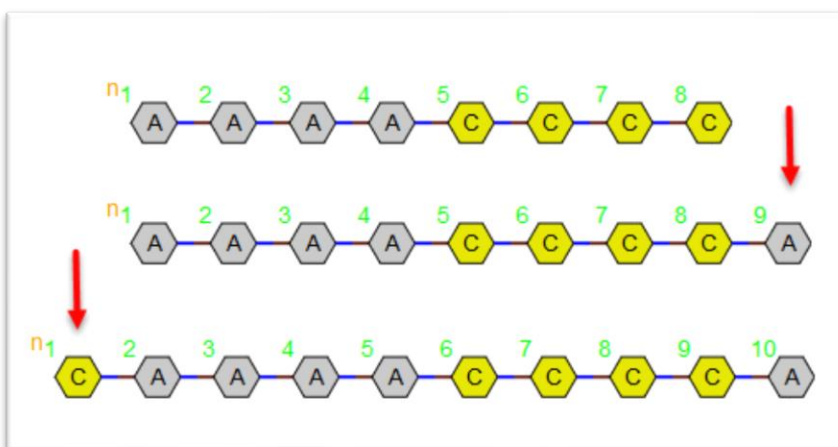
There are three ways to add a monomer to the Canvas:

- 1) Click on the monomer in the Monomer Browser, and drag to the drawing Canvas.
- 2) Click on the monomer then click in the Canvas.
- 3) Write a sequence in the Viewing area.

After adding the first monomer, the next can be added in the same way.

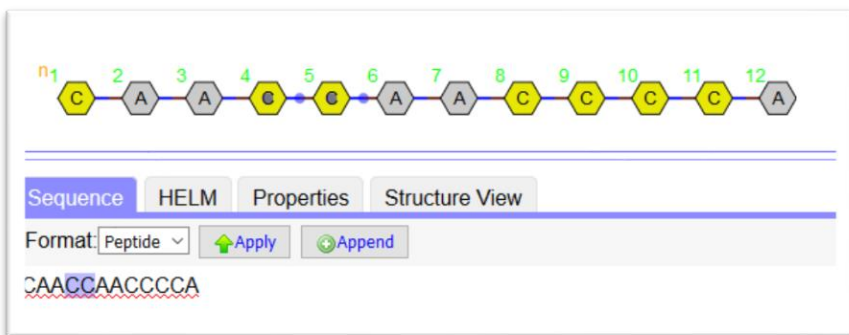
A monomer on the Canvas can be changed to another monomer of the same type in three ways:

1. Highlight the replacement monomer in the Monomer Browser, hover over an existing monomer on the Canvas and click on the monomer,
2. Hover over an existing monomer on the Canvas and using the keyboard type the letter representing the monomer,
3. Highlight the replacement monomer in the Monomer Browser, hover over an existing monomer on the Canvas, left-mouse-click and drag left or right away from the monomer. In the example, an 'A' is added to the right, and then, a 'C' is added to the left. Note that the numbering changes when a monomer is added to the left.



Monomers can also be added to any position in the sequence by adding them into the sequence in the Sequence Viewing area.

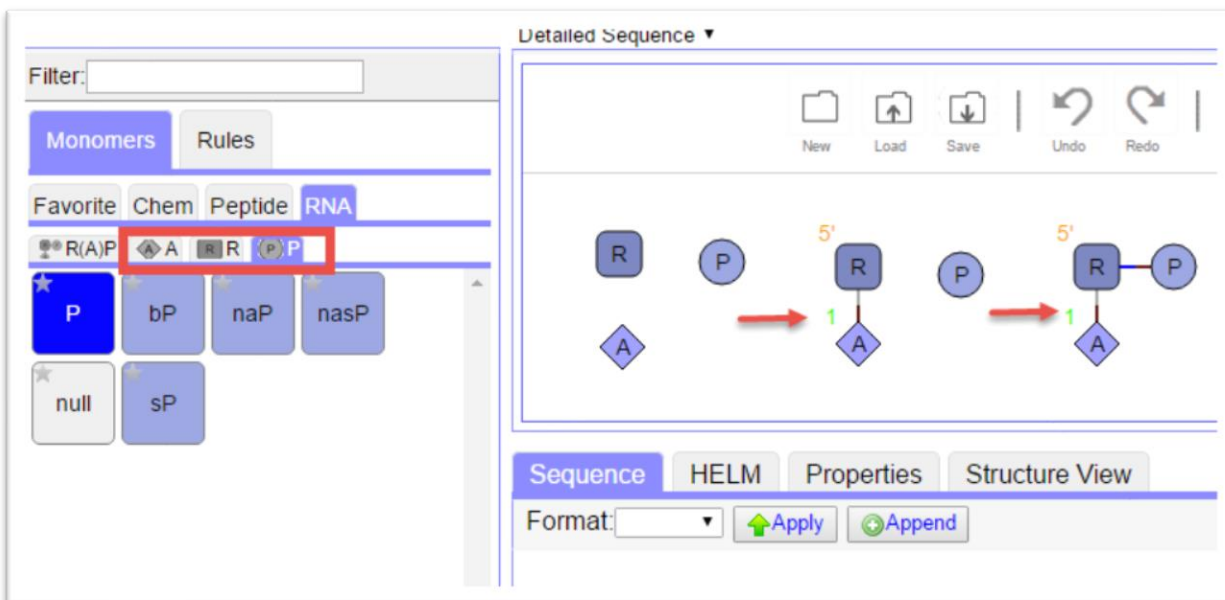




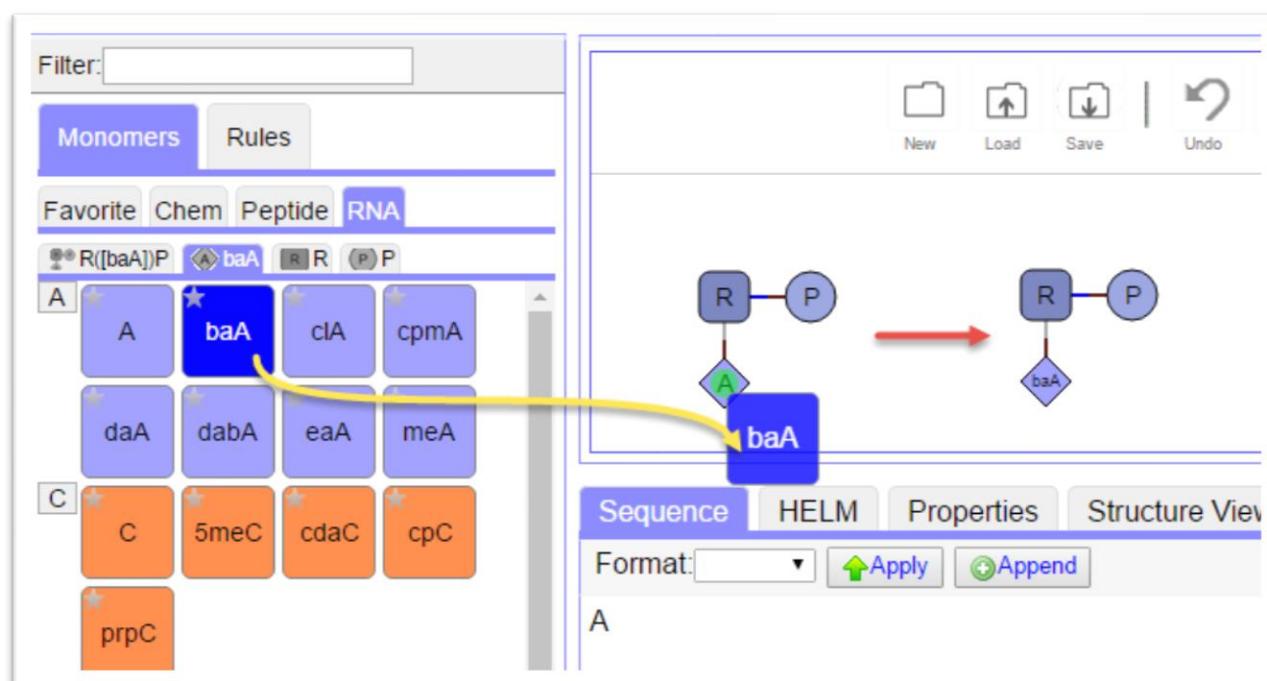
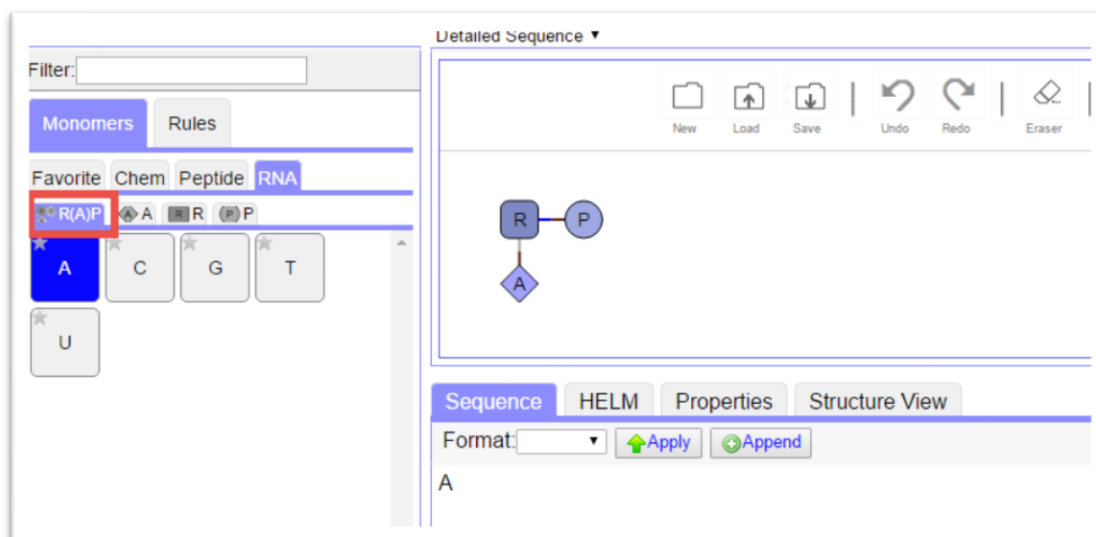
## Nucleotides

Nucleotides can be added to the Canvas in two ways:

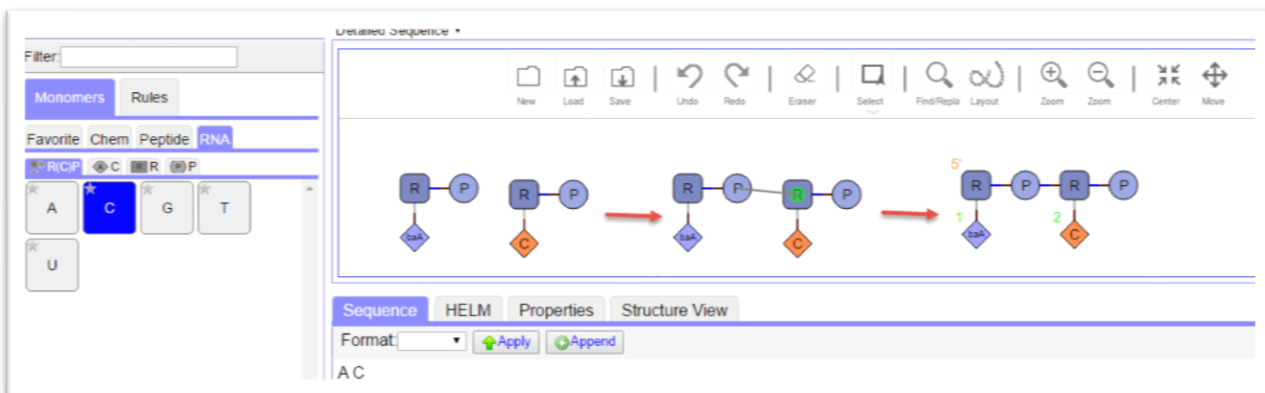
- 1) Linkers, Bases, and Sugars from the RNA tab can be placed directly onto the canvas and individually linked. Note that the 5' end of the sequence is indicated.



- 2) The Nucleotide can be added from under the Nucleotide (Combined) tab and then the individual linkers, bases and sugars can be changed by clicking on the desired linker, base or sugar in the monomer list then dragging to the desired position or clicking on the desired position.



Once there is more than one monomer on the screen, a bond can be drawn between them by selecting one monomer, left-mouse-click, hold and drag to the second monomer.

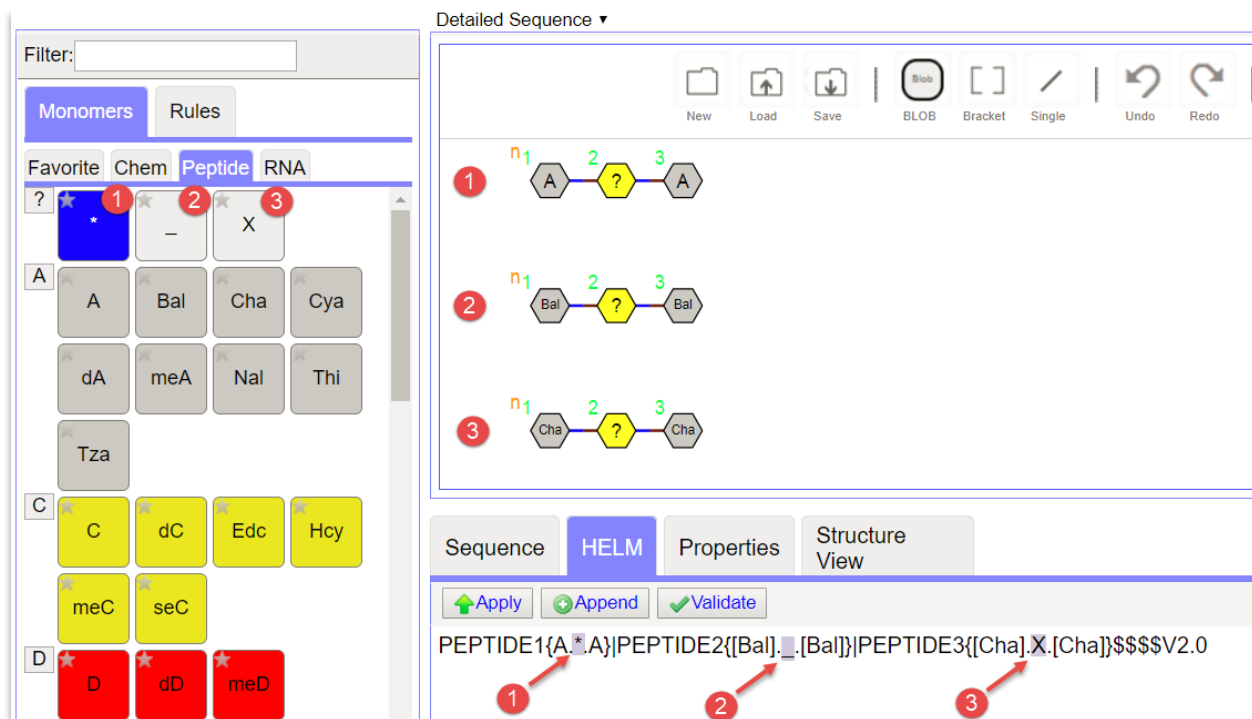


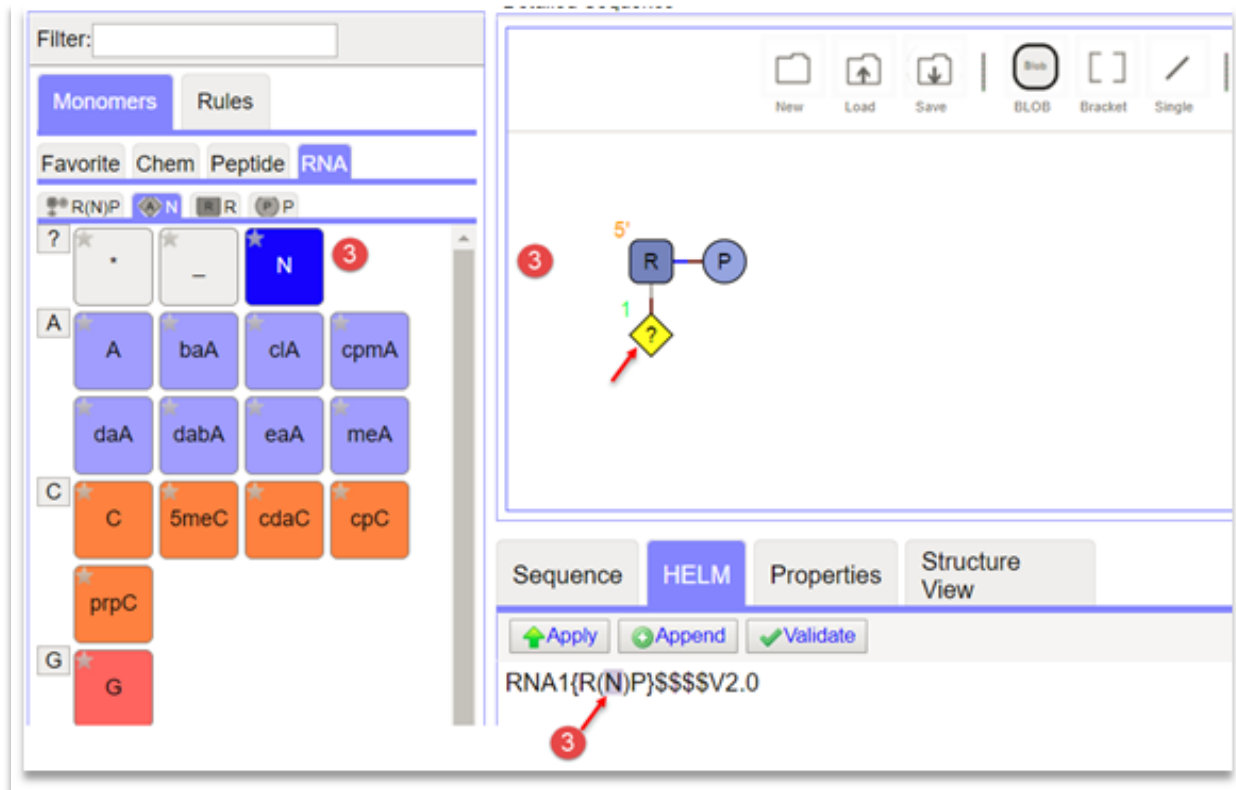
Please note the following principles:

- Connections are drawn from N to C terminus for peptides and from 5'→3' for the nucleotides.
- If monomers cannot be connected a dialogue box will appear with a message.
- If two sugars are connected, the phosphate linker is automatically inserted.

### Using unknown monomers (Peptide and RNA)

Unknown monomers can be selected from the monomer palette in the same way as other monomers. The monomers are shown in yellow with a "?" to indicate that they are not specific.





You can see the monomer details in the HELM string, or by hovering over the monomer on the canvas.

There are 4 types of unknown monomer:

- "\*" represents 0..n unknown monomers.
- "X" represents one single unknown amino acid in a PEPTIDE polymer.
- "N" represents one single unknown base in a RNA polymer.
- "\_" represents a deleted or missing single monomer. This is typically used in monomer lists (see Alternative monomers and mixtures)

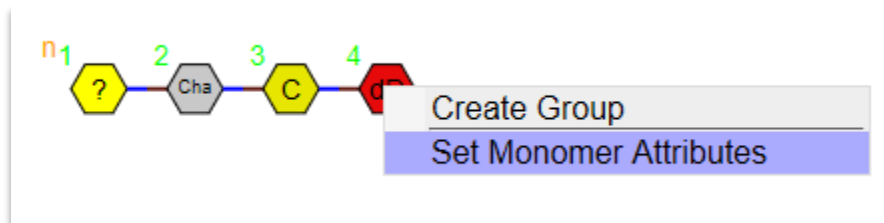
## Alternative monomers and mixtures

There may be situations where you have synthesized a molecule, but you either know that it is not pure, or are uncertain what the structure is. In these circumstances you may want to define a list of monomers at one or more positions in the molecule.

HELM provides the option to list monomers and specify whether you have

- monomer A OR monomer B.
- or monomer A AND monomer B

You can specify alternative monomers and mixtures by right-clicking on the monomer and choosing 'set monomer attributes'.



You can then type alternatives into the monomer symbol box

**Monomer Attributes** [X]

Monomer Symbol:

Annotation:

A comma between monomer symbols = alternatives

A plus sign = mixture.

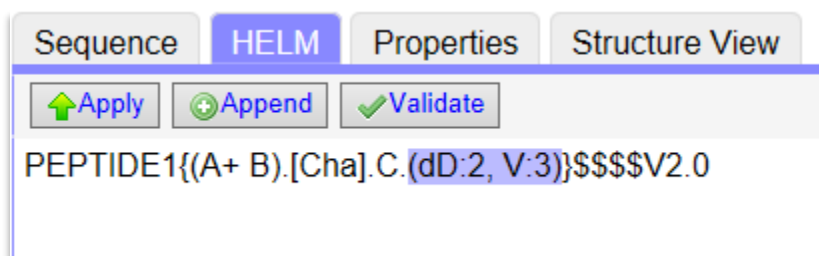
If you know the relative proportions of each monomer, you also can add this to the monomer symbol box. For example:

**Monomer Attributes** [X]

Monomer Symbol:

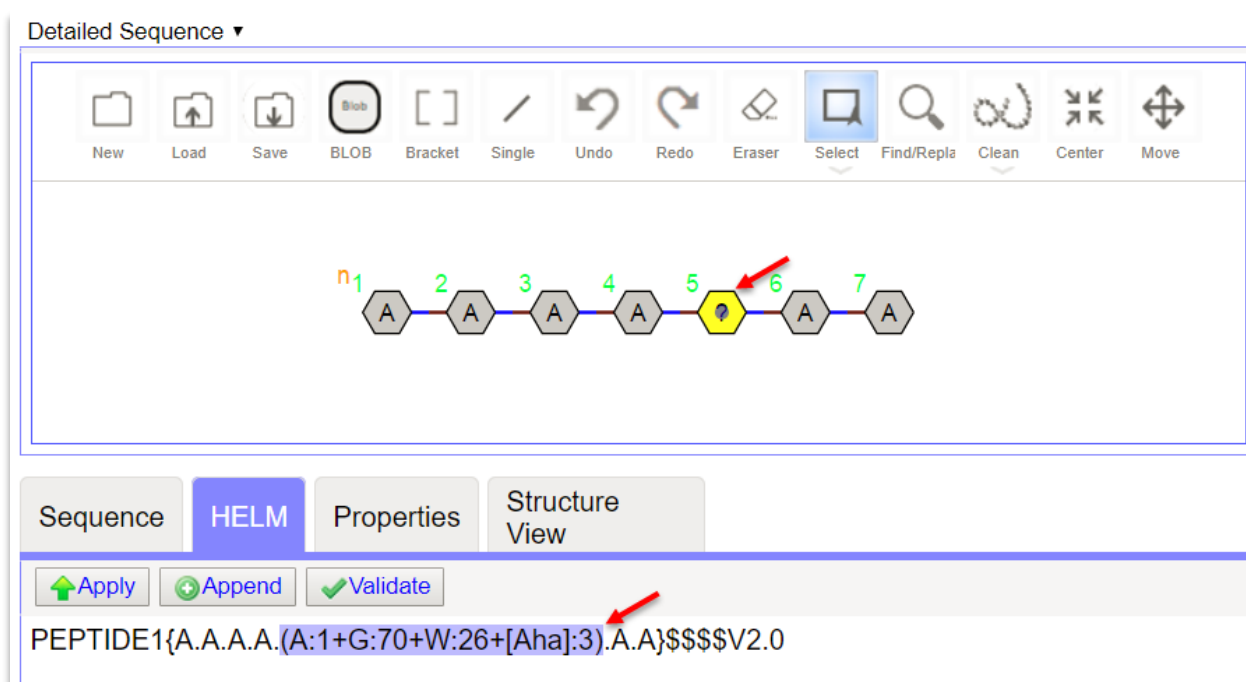
Annotation:

This is reflected in the HELM for this molecule



The screenshot shows a software interface with four tabs: "Sequence", "HELM", "Properties", and "Structure View". The "HELM" tab is active. Below the tabs are three buttons: "Apply" (with an upward arrow), "Append" (with a plus sign), and "Validate" (with a checkmark). The main text area contains the HELM string: PEPTIDE1{(A+ B).[Cha].C.(dD:2, V:3)}\$\$\$\$V2.0. The portion (dD:2, V:3) is highlighted in blue.

A more complex example is shown below:



The screenshot shows a software interface with four tabs: "Sequence", "HELM", "Properties", and "Structure View". The "HELM" tab is active. Below the tabs are three buttons: "Apply", "Append", and "Validate". The main text area contains the HELM string: PEPTIDE1{A.A.A.A.(A:1+G:70+W:26+[Aha]:3).A.A}\$\$\$\$V2.0. A red arrow points to the [Aha]:3 portion. Above the text area is a "Detailed Sequence" section with a toolbar containing icons for New, Load, Save, BLOB, Bracket, Single, Undo, Redo, Eraser, Select, Find/Repla, Clean, Center, and Move. Below the toolbar is a structural diagram of a peptide chain with seven amino acid residues represented by hexagons labeled A. The residues are numbered 1 through 7. Residue 5 is highlighted in yellow and contains a question mark, with a red arrow pointing to it from above.

### Missing monomers

In some cases, a position can be occupied by a monomer or not. Example: A trailing Lysine might be present at the C-terminus of an antibody or not. An underscore is used to denote a potential blank position

Detailed Sequence ▾

Sequence HELM Properties Structure View

↑ Apply ↕ Append ✓ Validate

PEPTIDE1{A.A.C.D.D.E.E.(,K)}\$\$\$\$V2.0

## Alternative polymers and mixtures

It is also possible to define a mixture or set of alternatives at a polymer level. Take this example of a mixture of 4 proteins with differing monomers at positions 2 and 5 and

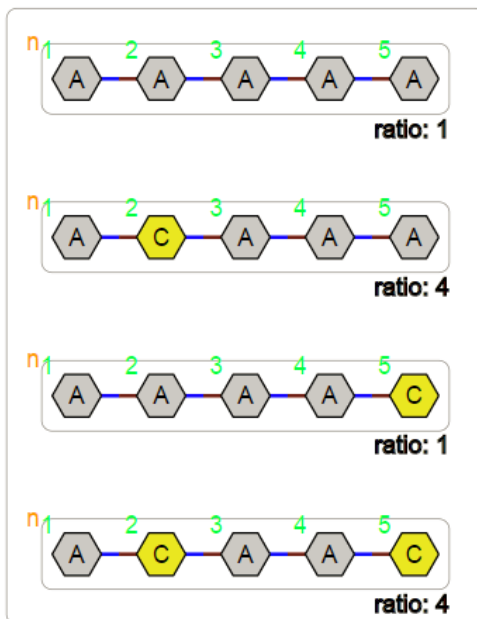
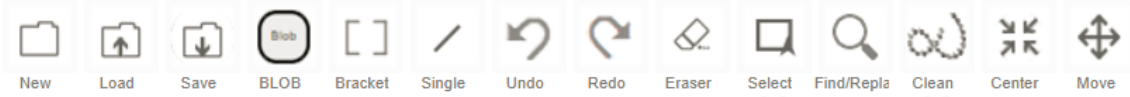
```
10% PEPTIDE1{A.A.A.A.A}
40% PEPTIDE1{A.C.A.A.A}
10% PEPTIDE1{A.A.A.A.C}
40% PEPTIDE1{A.C.A.A.C}
```

To do this in the webeditor, you must:

- 1) Draw each peptide.
- 2) Make each peptide a group.
- 3) Right-click the border of each group, **Set Attributes**, enter ratio and click **Save**.
- 4) Select all four peptides and make a group.

Once you have done this for all groups you will see the values reflected in the HELM string.

Detailed Sequence ▾



Sequence

HELM

Properties

Structure View

Apply

Append

Validate

PEPTIDE1{A.A.A.A}|PEPTIDE2{A.C.A.A.A}|PEPTIDE3{A.A.A.A.C}|PEPTIDE4{A.C.A.A.C}\$\$G5  
 (PEPTIDE1:1+PEPTIDE2:4+PEPTIDE3:1+PEPTIDE4:4)\$\$V2.0



## Monomer Annotations

The screenshot illustrates the process of adding an annotation to a monomer in a software interface. The left panel shows a grid of monomers categorized by A, C, D, and E. The right panel, titled "Detailed Sequence", shows a sequence of three monomers (1, 2, 3) on a canvas. Monomer 2 is highlighted with a context menu showing "Create Group" and "Set Monomer Attributes". A "Monomer Attributes" dialog box is open, showing "Monomer Symbol" and "Annotation: IL6". Below the canvas, the "HELM" tab is active, showing the sequence "PEPTIDE1{\*}IL6\$\$\$\$V2.0" with a red arrow pointing to the annotation. The bottom panel has "Apply", "Append", and "Validate" buttons.

You can add an annotation to a monomer by:

- 1) Placing a monomer on the canvas.
- 2) Right-click the monomer, right-click **Set Monomer Attributes**, add annotation, then click **Save**.
- 3) Mouse-hover over the monomer on the canvas to view the annotation
- 4) View HELM notation.
  - PEPTIDE1{\*}IL6\$\$\$\$ represents a peptide chain with the name IL6.

## Unknown simple polymers

Blobs represent simple polymers with an unknown structure/sequence.

Example

- BLOB1{Bead}Aminated Polystyrene\$\$\$\$ represents aminated polystyrene bead. Simple polymer type is "Bead", and simple polymer name is "Aminated Polystyrene".)

To add a blob:

New Load Save **BLOB** Bracket Single Undo Redo

2 Blob

3 Blob Type > Bead  
Set Monomer Attributes Gold Particle

4 Blob Type >  
Set Monomer Attributes

**Monomer Attributes**

Monomer Symbol:

Annotation:

Save

**Aminated Polystyrene**

Sequence **HELM** Properties Structure View

5

BLOB1{Bead}"Aminated Polystyrene"\$\$\$\$V2.0

- 1) Select **BLOB** button.
- 2) Left-click in the drawing canvas to place the BLOB.
- 3) Right-click the BLOB and select **Blob Type**.
- 4) Left-click the BLOB, right-click **Set Monomer Attributes**, add annotation/change monomer symbol, then click **Save**.
- 5) Click **Apply** button in the viewing area to apply HELM notation.

## Connection ambiguity

Sometimes it is not possible to specify where or how many times components are connected to each other. You can represent this in the editor by using the bond attributes.

You can only change the bond attributes on bonds between groups or blobs.

### Example: Connecting BLOB to Group and Setting Bond Attributes

The screenshot shows a chemical editor interface with a monomer palette on the left and a main canvas on the right. The canvas displays a sequence of four monomers (A, D, A, D) with positions 1, 2, 3, and 4. The process is illustrated in five steps:

- 1) A red arrow points to the 'Select' button in the toolbar.
- 2) A red arrow points to the 'Create Group' context menu option.
- 3) A red arrow points to a 'Blob' icon being placed on the canvas.
- 4) A red arrow points to the 'Single' button in the toolbar.
- 5) A red arrow points to the 'Set Bond Attributes' context menu option.

The 'Bond Attributes' dialog box is open, showing the following fields:

| Bond Attributes                     |       |
|-------------------------------------|-------|
| Monomer #1:                         | Blob  |
| Position:                           |       |
| R#:                                 | ?     |
| Ratio:                              | 1     |
| Monomer #2:                         | Group |
| Position:                           |       |
| R#:                                 | ?     |
| Ratio:                              | 1     |
| <input type="button" value="Save"/> |       |

- 1) Left-click the **Select** button and select the sequence on the canvas.
- 2) Right-click on the selected sequence and select **Create Group**.
  - a. A border will now be around the sequence, indicating a group.
- 3) Place a BLOB on the canvas.
- 4) Left-click on the **Single** button and draw a bond connecting the Blob to the Group (Connect to any green circle on the Group border).
- 5) Left-click the **Select** button, right-click on the bond, select **Set Bond Attributes** (Position, R#, Ratio), and click **Save**.

## Repeat sequences

One or more monomers may be repeated in a polymer. You can represent this in a compact way, by using brackets.

Filter:

Monomers Rules

Favorite Chem Peptide RNA

? \* - X

A A Bal Cha Cya  
dA meA Nal Thi  
Tza

C C dC Edc Hcy  
meC seC

D D dD meD

E E dE Ggu Gla  
meE

F F dF meF

Detailed Sequence ▾

New Load Save BLOB Bracket Single

1

2

Change subscript

Number

Number range

Ambiguous

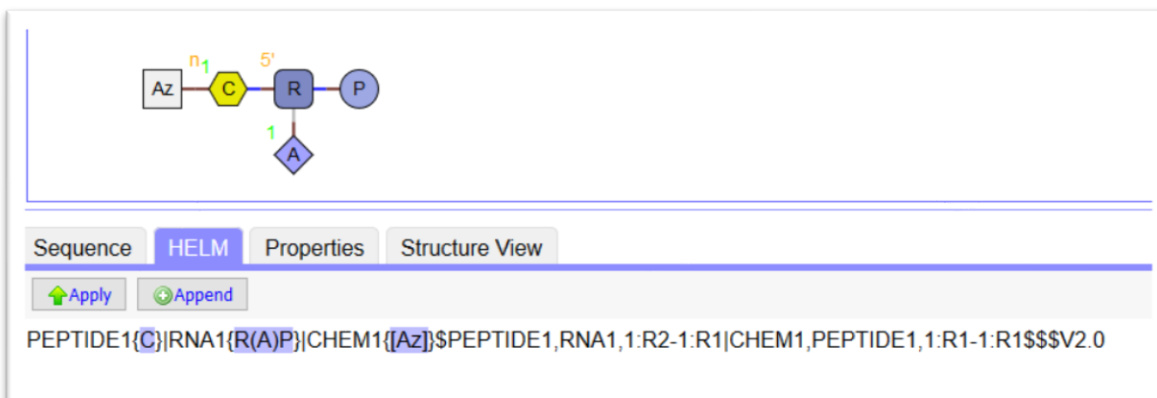
Sequence HELM Properties Structure View

Apply Append Validate

PEPTIDE1{A.(S.D)3.F}\$\$\$\$V2.0

- 1) Select the **Bracket** button; click and drag to put brackets around desired sequence.
- 2) In the field, type a number, number range, or a "\*" for ambiguity (e.g. 3, 3-6, \*), then click on the canvas. To change the subscript, click the subscript and make changes in the field.

## Conjugates



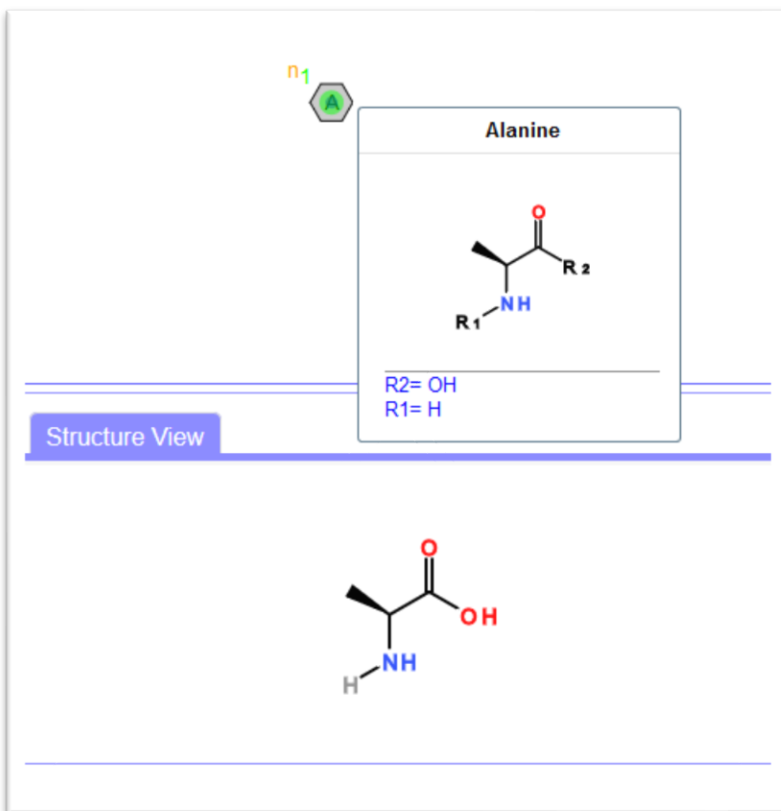
Conjugates can be drawn by selecting chemicals, peptides and nucleotides from the Monomer Browser and linking the monomers together.

If monomers cannot be connected a dialogue box will appear with a message.

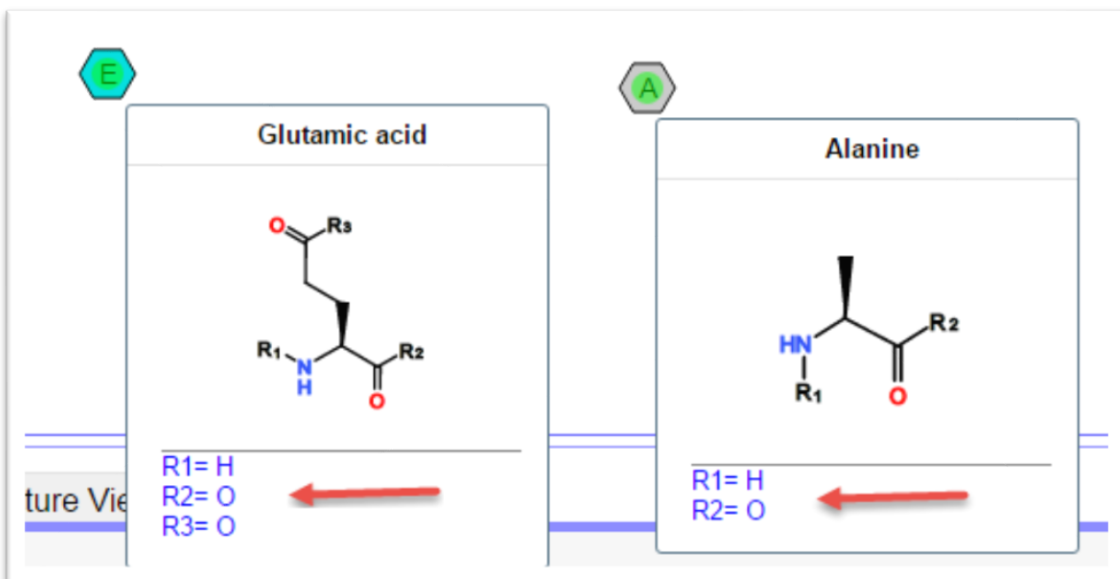
## R-Groups

It is important to understand the r-groups available in a monomer. The tooltip structure will show you where they are in the structure itself and the list at the bottom of the box shows the capping groups for each R group.

R groups are shown as capped in the Viewing area\Structure View tab. In this Alanine example R1 is H-capped.



Monomers in the middle of a chain are typically linked to two other monomers, however certain monomers can form additional bonds. For example, Alanine (A) has two R groups (R<sub>1</sub> and R<sub>2</sub>) with which to connect to other monomers, while Glutamic Acid (E) has three R groups (R<sub>1</sub>, R<sub>2</sub> and R<sub>3</sub>).



In this example, 'E' is attached to 'C'. When drawing the connecting bond between 'C' and 'E' a pop-up message will appear. The available connecting points for the two monomers will be displayed in the dropdown menu. There is only one available connecting point for C2 (R3). There are three available connecting point for E (R1, R2 and R3).

Detailed Sequence

Filter:

Monomers Rules

Favorite Chem Peptide RNA

A A dA mA meA

C C dC meC seC

D D dD meD

E E dE meE

F F dF meF

G G meG

H

Sequence HELM

Format: Apt

E ACD

Choose Connecting Points

E

R3

R1-NH

R2

R1

R2

R3

C2

R3-S

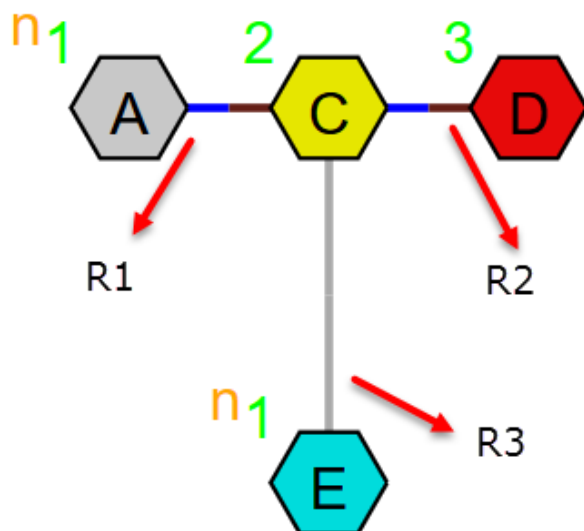
R1-NH

R2

R3

OK

The number of attachment points of a monomer defines the number of covalent bonds in which the monomer can be involved.



The Connection points are indicated by color: blue for R1, orange for R2, and gray for R3. The n-terminal is indicated by 'n'.

## Working with structures

Monomers or bonds can be removed by mouse-hovering over them and hitting delete. Also, use the Select button to select the monomers to delete and hit delete.

To replace any of the monomers, drag a replacement monomer from the monomer browser and drop it onto the current monomer (as described previously), so long as their type and attachment points allow the replacement.

### *Replacing monomers*

If a sequence contains a consistently modified monomer there are two options for updating it.

1. Use the Rules tab from the monomer browser (as described previously) to perform a search and/or replace the monomer.
2. Use the Find/Replace button (as described previously) in the Canvas menu to find a monomer or to replace a monomer with another.



Using the Find/replace box:

Find:

Enter the monomer in the Find box, choose the Monomer Type from the dropdown menu, and click Find to find the first monomer in a sequence or click the Find All button to find all monomers in a sequence. A pop-up message will display the number of monomers found. The monomers found will be highlighted in the Sequence or HELM Viewing area.

Replace:

Enter the monomer in the Find box, choose the Monomer Type from the dropdown menu, enter the replacement monomer in the Replace With box and click the Replace All button. A pop-up message will display the number of monomers replaced.

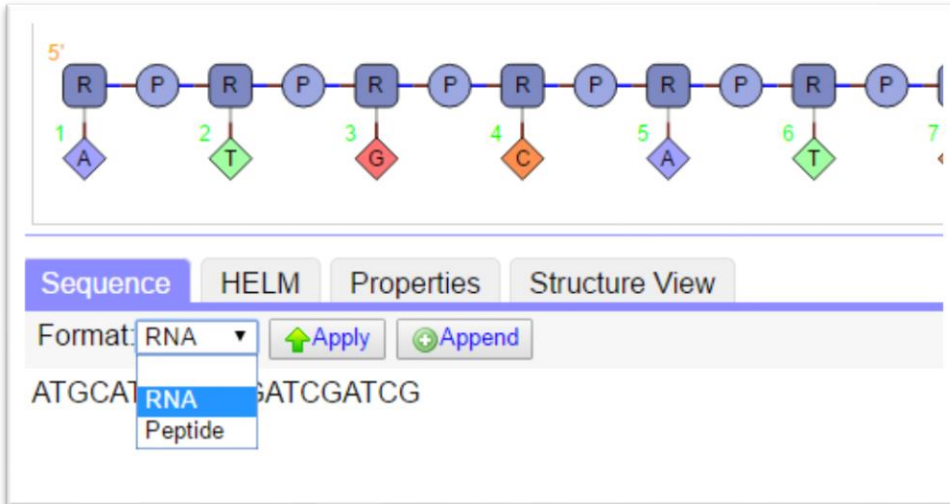
If only finding/replacing in a selected area, select that area using the Select button and check the Search Selected Only box.

The screenshot displays a software interface for sequence manipulation. At the top, a sequence of 16 amino acids is shown as a chain of colored hexagons: A (grey), E (teal), A (grey), C (yellow), A (grey), D (red), A (grey), C (yellow), A (grey), C (yellow), A (grey), E (teal), D (red), A (grey), C (yellow), A (grey). Below this, a 'Find and Replace' dialog box is open. It contains the following fields: 'Find:' with 'A', '(Monomer symbol or position)'; 'Monomer Type:' with a dropdown menu set to 'Amino Acid'; 'Replace With:' with 'I'; and 'Scope:' with an unchecked checkbox for 'Search Selected Only'. At the bottom of the dialog are three buttons: 'Find', 'Find All', and 'Replace All'. Below the dialog, the 'Sequence' tab is active, showing the text 'AEACADACACAEDACA'. A red arrow points from the first 'A' in this sequence to the 'Find and Replace' dialog box. Below that, the 'Sequence' tab is still active, but the text has changed to 'IEICIDICIEDICI', with the first 'I' highlighted.

## Building a molecule using the sequence

Build new molecules by entering sequence information directly in the Viewing area, rather than dragging and dropping individual monomers, a process which is slow for larger polymers.

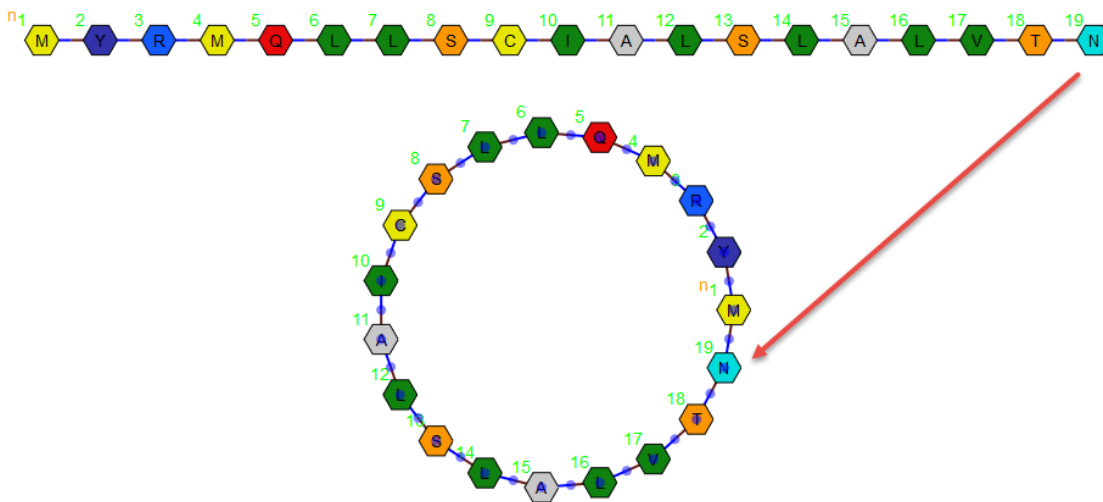
There are two sequence loading options: nucleotide sequence (RNA) and peptide sequence (Peptide).



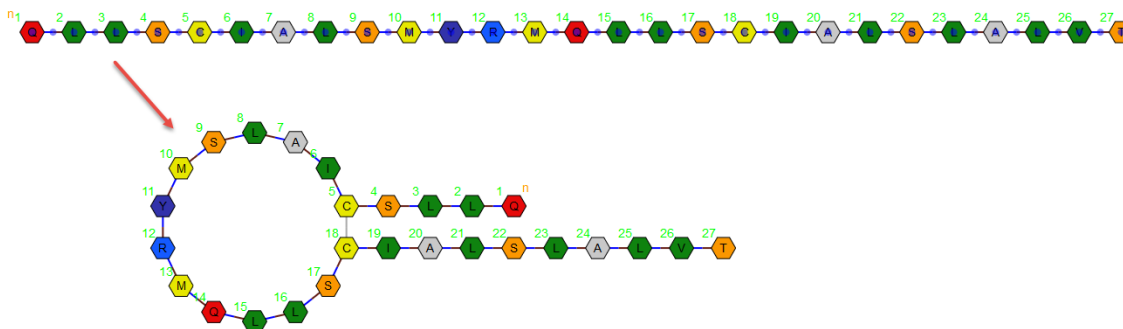
## Cyclizing a sequence

To form a cyclic sequence, add a bond between two monomers that have suitable attachment points.

To create a backbone cycle, connect the first and the last monomers. In this example connect 'M', the first peptide, with 'N', the 19<sup>th</sup> peptide.



To create an intra-chain bridge, connect two monomers in the same sequence, that have more than two attachment points. Once the bond is added, the molecule will rearrange to form a cycle, and show the link as appropriately as possible.



## Saving and loading monomers

Once a new molecule is created or an existing monomer is modified, there are two formats to save and load the molecule data (using the Save and Load buttons in the Canvas), the HELM notation and the xHELM notation.

### Helm notation

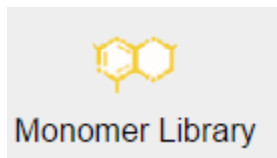
The HELM notation is the native data format of the HELM editor. The format is described extensively on the OpenHELM.org website. If you want to use your molecules mainly within your organization, HELM is the format of choice. It is a very compact and clear but it relies on a common set of monomers for the molecules.

### Exchangeable HELM notation (xHELM)

The xHELM format was developed as the “portable” version of HELM. Like the HELM notation, xHELM is a unique notation. It contains the normal HELM notation in addition to the full set of monomers contained within the molecule. This is useful if you want to exchange your molecule data with other organizations, especially if you use non-standard monomers.

## Managing monomers

HELM is a hierarchical notation where polymers are built from a predefined set of monomers. In the open source editor, these are stored in an XML file, but in an organization, these can be held in a central database.



Add to the monomer set using the Monomer Library.

## Using the Monomer Library

The Monomer Library is used to view the current monomer collection and register new monomers.

The screenshot shows the Monomer Library interface. At the top, there are search and filter options: Symbol/Name, Polymer Type, Monomer Type, and Count. Below this is a table of monomers. The 'Az' monomer is highlighted in orange. Below the table, there is a detailed view for the 'Az' monomer, showing its symbol, name, natural analog, polymer type, monomer type, and structure. The structure is a chemical diagram of an azide group attached to a chain with a terminal carbonyl group.

| Symbol    | Name   | Natural Analog | Polymer Type | Monomer Type | R1 | R2 | R3 | Author | Created Date |
|-----------|--|----------------|--------------|--------------|----|----|----|--------|--------------|
| ACC       | 4-(N-maleimidomethyl)cyclohexane-1-carboxylate |                | CHEM         |              | OH |    |    |        |              |
| SOH       | 6-amino-hexanol                                |                | CHEM         |              | H  | H  |    |        |              |
| Az        | Azide  |                | CHEM         | Backbone     | OH |    |    | Fred   |              |
| Cys-BLOCK | Cys-BLOCK                                      |                | CHEM         |              | X  |    |    |        |              |
| SS3       | Dipropionol-disulfide                          |                | CHEM         |              | H  | H  |    |        |              |
| EG        | Ethylene Glycol                                |                | CHEM         |              | H  | OH |    |        |              |
| Z         | Hydraz   |                | CHEM         | Backbone     | H  | OH |    | Claire |              |
| Is        | glycyl   |                | CHEM         | Branch       | H  |    |    |        |              |
| xy        | Hexyl alcohol                                  |                | CHEM         |              | H  |    |    |        |              |
| N-BLOCK   | N-BLOCK  |                | CHEM         |              | OH |    |    |        |              |

**Monomer** 3

Symbol: Az  
Name: Azide  
Natural Analog:  
Polymer Type: CHEM  
Monomer Type: Backbone  
Author: Fred  
Structure:

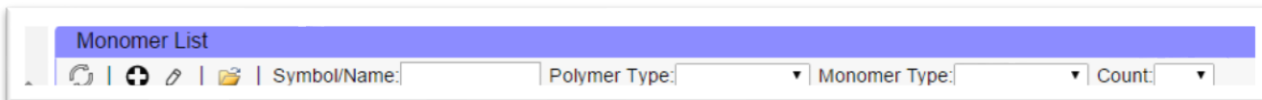
R1C(=O)CCCC[N+]=[N-]

R1: OH  
R2:


The Monomer library:

1. Monomer List Menu
2. Monomer List
3. Monomer information

## Monomer List Menu



Hover over icons for identity key.

Refresh  to refresh the page.

New  to add a new monomer to the list.

Edit  to edit an existing monomer in the list.

Import Monomer XML Library  to import an XML file.

Use the Symbol/Name box to search for an existing monomer via the dropdown list.

Use the Polymer Type dropdown list to filter monomers listed as RNA, CHEM or PEPTIDE.

Use the Monomer Type dropdown list to filter monomers listed as Backbone, Branch or Undefined.

Use the Count dropdown list to change the number of items displayed on a page.

## Monomer List

Previous Page 1 2 3 4 5 6 7 8 9 10 11 ... 15 Next Page

| Symbol    | Name   | Natural Analog | Polymer Type | Monomer Type | R1 | R2 | R3 | Author | Created Date |
|-----------|--|----------------|--------------|--------------|----|----|----|--------|--------------|
| MCC       | 4-(N-maleimidomethyl)cyclohexane-1-carboxylate |                | CHEM         |              | OH |    |    |        |              |
| ABOH      | 6-amino-hexanol                                |                | CHEM         |              | H  | H  |    |        |              |
| Az        | Azide  |                | CHEM         | Backbone     | OH |    |    | Fred   |              |
| Cys-BLOCK | Cys-BLOCK                                      |                | CHEM         |              | X  |    |    |        |              |
| SS3       | Dipropionyl-disulfide                          |                | CHEM         |              | H  | H  |    |        |              |
| EG        | Ethylene Glycol                                |                | CHEM         |              | H  | OH |    |        |              |
| Z         | vdasg  |                | CHEM         | Backbone     | H  | OH |    | Claire |              |
| ds        | gresye   |                | CHEM         | Branch       | H  |    |    |        |              |
| hxy       | Hexynyl alcohol                                |                | CHEM         |              | H  |    |    |        |              |
| N-BLOCK   | N-BLOCK  |                | CHEM         |              | OH |    |    |        |              |

To create a new monomer

1. Click on the 'New' button in the menu and a structure drawing editor will pop up.

Monomer

Symbol:\*

Name:\*

Natural Analog:\*

Polymer Type:\*

Monomer Type:\*

Author:

Structure:

R1:

R2:

R3:

Save  Delete

2. Fill in the required information. Required information is indicated with a red asterisk.
3. Draw the molecule in the structure editor.
4. Make sure to choose the attachment points as R groups in the way recommended for that monomer type.
5. Click Save
6. The new monomer should now be available in the list.

Note: Both the monomer name and the symbol must be unique for the monomer to be registered. Please check carefully before registering.

Alternatively, an existing monomer can be used as a starting point.

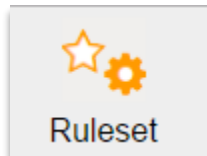
1. Select the monomer to use as a template.
2. Click on the "Edit" button in the menu.
3. Change the structure to the new monomer.
4. Edit the monomer symbol and name.
5. Click Save.

### Monomer information

Select a monomer in the list to display the available information in the "Monomer" window.

## Managing Rulesets

Add to the Ruleset using the Ruleset icon in the HELM 2.0 top menu bar.



## Using the Ruleset Manager

The Ruleset Manager is used to view the current ruleset and add new rules. These rules are then available to apply in the HELM canvas (see Rules under the Monomer Browser section).

| ID | Category | Name                  | Description | Author | Created Date |
|----|----------|-----------------------|-------------|--------|--------------|
| 2  | Demo     | Replace base A with G | null        | Nobody |              |
| 3  | Test     | Replace base A with T | null        | Nobody |              |
| 1  | Demo     | Replace base A with U | null        | Nobody |              |

Rule 2

ID:  
Category:  
Name:  
Description:  
Author:  
Javascript:


Main page:


1. Rule Set: Window to add or edit rules and display existing rules.
2. Rule: Window to view selected rules from the Rule Set.


## Rule Set

| ID | Category | Name                  | Description | Author | Created Date |
|----|----------|-----------------------|-------------|--------|--------------|
| 2  | Demo     | Replace base A with G | null        | Nobody |              |
| 3  | Test     | Replace base A with T | null        | Nobody |              |
| 1  | Demo     | Replace base A with U | null        | Nobody |              |

Hover over icons for identity key.

Refresh  to refresh the page.

New  to add a new rule to the list.

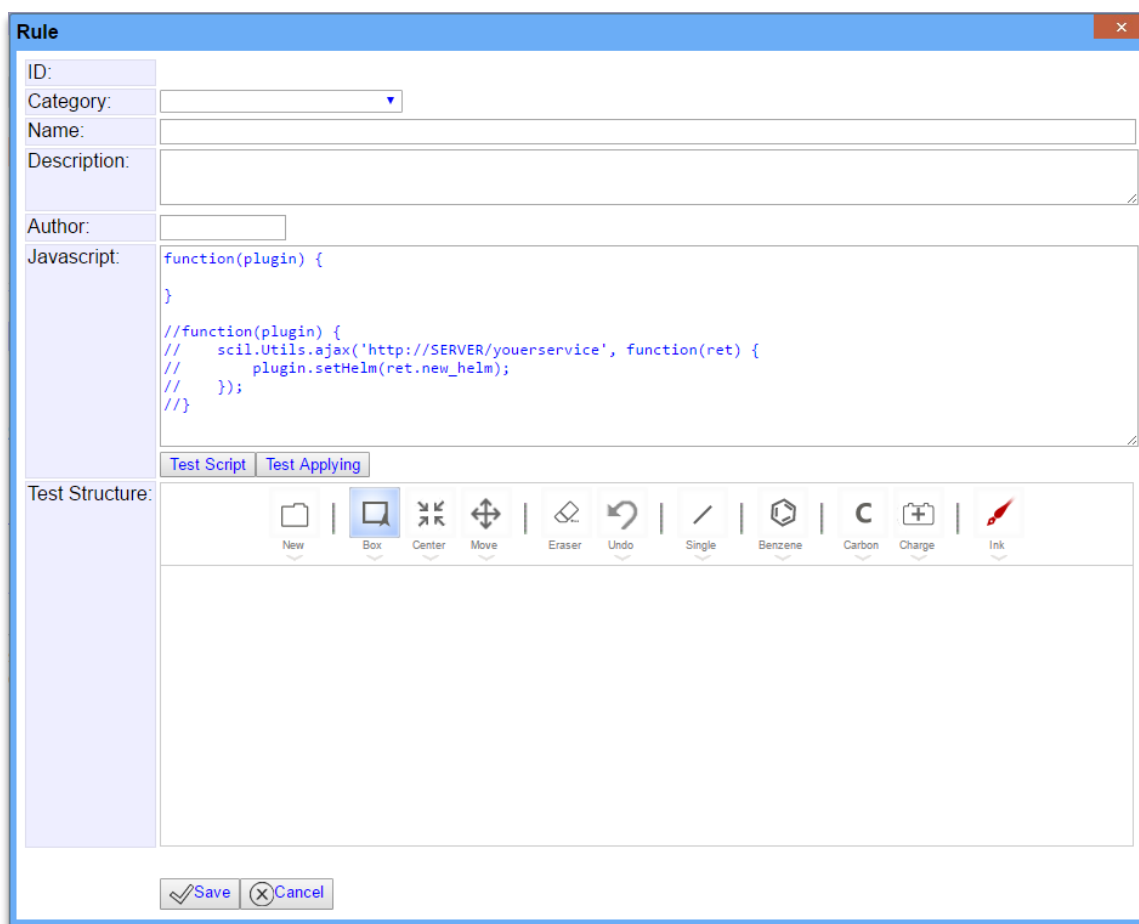
Edit  to edit an existing rule in the list.

Use the Category box to choose category via the dropdown list or leave blank to display all items.

Use the Count dropdown list to change the number of items displayed on a page.

To add a rule to the Rule Set list:

1. Click on the 'New' button in the menu and a rule editor will pop up.



**Rule**

ID:

Category:












Name:

Description:

Author:

Javascript: 

```
function(plugin) {  
  
}  
  
//function(plugin) {  
//  scil.Utils.ajax('http://SERVER/youerservice', function(ret) {  
//    plugin.setHelm(ret.new_helm);  
//  });  
//}
```

Test Structure:           

2. An ID will be automatically generated for the Rule.
3. Fill in the desired information.
  - a. Category: Choose from the dropdown list or leave blank.
  - b. Name: Enter the rule name.
  - c. Description: Enter a rule description.
  - d. Author: Enter the rule creators name.
4. JavaScript: Using the script provided, enter the JavaScript to implement the rule between the { }.



|             |  |
|-------------|--|
| Javascript: | <pre>function(plugin) {<br/><br/>}</pre> |
|-------------|--|

- a. Test Script: Click Test Script for a check of the entered JavaScript.
  - b. Test Applying: Click Test Applying to test the rule application.
5. If desired draw a molecule in the Test Structure editor.
6. Click Save
7. The new rule should now be available in the Rule Set list.