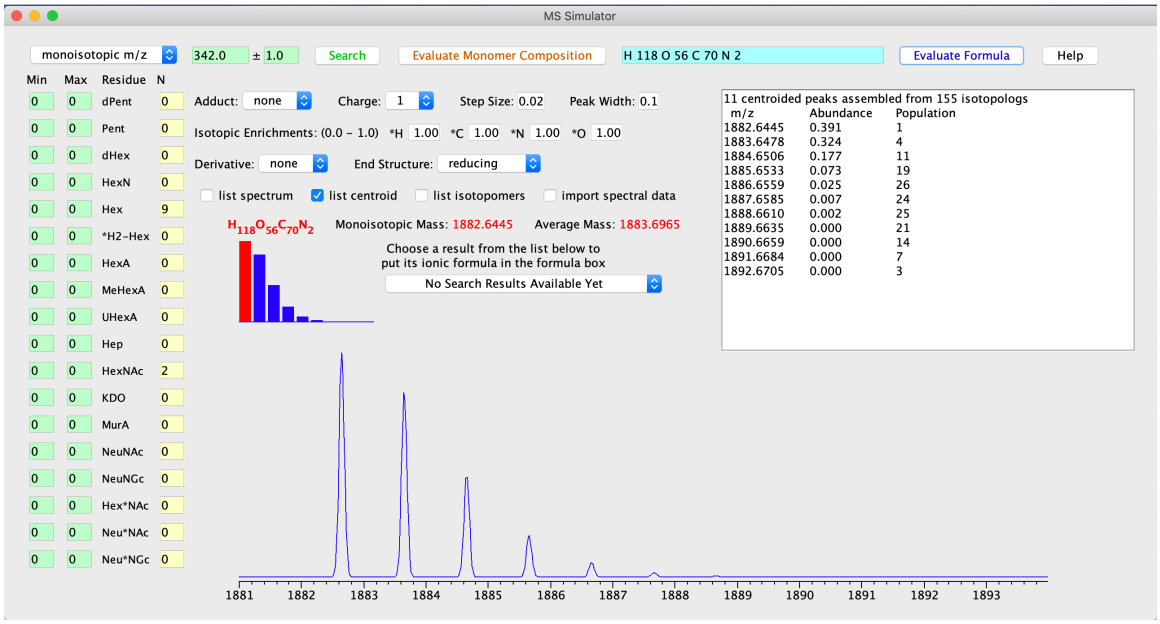
Brief User Guide for JMassApp

The JMassApp class (copyright 2009-2013 William S. York and the University of Georgia) implements a software application that utilizes the JMass API. By default, the current version of JMassApp uses the following data/configuration files:

```
monomers.xml
msconfig.xml
isotopes-5.xml
```



Screenshot of JMassApp

The user can search for an oligomer structure that matches a specified (monoisotopic or average) *m/z* value, given a range for the stoichiometric coefficient for each monomer. The user can also specify a monomer composition of an oligomer by entering the number of copies of each monomer type from a list on the left side of the application.

These files must be in the same directory as the executable jar file. Experienced programmers can modify these files and (using the API) generate executable code that uses the modified files or other files with appropriate content.

The chemical properties of these monomers are specified in an (XML format) input file. This specifies the name, abbreviation, number of etherification sites and number or acylation sites that is added to a polymer when each monomer residue is added to a polymer. The number of atoms of each element of the residue that is added to the polymer upon addition of the residue is also specified. Note that this usually corresponds to the elemental composition of the free monosaccharide, amino acid, etc MINUS a water molecule.

For example,

```
<monomer name='hexose' abbrev='Hex' e_sites='3' a_sites='3'>
      <component symbol='H' number='10'/>
      <component symbol='C' number='6'/>
      <component symbol='O' number='5'/>
</monomer>
```

Compare this composition ($C_6H_{10}O_5$) to the composition of a free hexose ($C_6H_{12}O_6$). The number of acylation sites and the number of etherification sites are both three in this case. For example, the disaccharide maltose, composed of two hexose residues, has 8 acylation sites. The trisaccharide maltotriose, composed of three hexose residues, has 11 acylation sites. So, one hexose residue adds 3 acylation sites.

Molecules can be chemically modified by gobally replacing all of the exchangeable H atoms (e.g., in hydroxyl groups and amino groups, defned as etherification sites or acylation sites) with a "derivative". Each possible derivative is also defined in the xml file. In this case, the name , abbreviation and type of the derivative is specified. The number of atoms of each element added to the polymer when a single copy of the derivative is also specified, as follows.

```
<derivative name='acetyl' abbrev='Ac' type='a'>
      <component symbol='H' number='3'/>
      <component symbol='C' number='2'/>
      <component symbol='O' number='1'/>
</derivative>
```

In the above example, derivitization corresponds to replacing the proton of the OH group with an acetyl group. For derivatives, the complete elemental composition of the derivative is specified (C_2H_3O in the case of an acetyl substituent).

Users of JMassApp can set parameters that select chemical modifications and define the ionization of the oligomer during mass spectrometry. In addition, the user can specify the ionizing species (e.g., Na^+) the charge, the step size (in *m/z* units) of the spectral simulation, the peak width (in *m/z* units) of the spectral simulation, and the isotopic enrichment of any monomers or derivatives that are defined in the monomer XML file. Isotopic enrichment is defined using the concept of a pseudoelement, which correspond to a natural abundance element, except the isotopic composition is altered. For example, the pseudoelement "*C" might be composed of 97% ^{13}C and 3% ^{12}C. The default isotopic compositions of all natural abundance elements and pseudoelements are specified in the isotopes XML file. By convention, the symbol of a pseudoelement (e.g., "*C") is the same as the corresponding natural abundance element (e.g., "C") with a leading asterisk. The user can modify the isotopic composition of a pseudoelement by entering values (from 0.0 to 1.0) for the fractional abundance of the most abundant isotope of the pseudoelement in the row labeled "Isotope Enrichments". This is reflected in the isotopic compositions of all isotopically enriched oligomers (i.e., those that contain momomers or derivitives that themselves contain pseudoelements).

The user can also select the derivative that replaces exchangeable hydrogens at all acylation or etherification sites. One can also specify the structure of the "reducing end". The options here are "reducing" (i.e., -OH, with ***no derivative***), "derivatized" (i.e., -OR, where R is the selected derivative) or "alditol" (i.e., the straight-chain, reduced form, with all acylation and etherification sites occupied by the selected derivative).

The monomers defined in the XML file are shown along the left side of the application. To calculate the composition, mass, etc. of an ion of known monomer composition, set up the appropriate structural parameters, enter the stoichiometric coefficients for the composition in the yellow boxes and click the "Evaluate Monomer Composition" button. The elemental composition, monoisotopic mass, chemical mass of the ion are then displayed next to a histogram of the isotopolog distribution in the center of the application. In the histogram, the bar corresponding to the monoisotopic species is hightlighted in red.

Alternatively, the user can specify the elemental composition of the oligomer or ion by entering a molecular formula directly into the light blue box. The entry ***must*** be in the following format: each element or pseudoelement *symbol* followed by its stoichiometric *coefficient*. These parameters are separated by spaces. For example the elemental composition of the doubly charged ion "$C_6H_{12}O_6Na_2$" is specified as "C 6 H 12 O 6 Na 2". Click on the "Evaluate Formula" button to display the elemental composition, monoisotopic mass, chemical mass, etc, as above.

One can search for ions whose *m/z* values match a given value, which is entered in the green box at the top of the application between the button that specifies the format (monoisotopic or average) and the text box for the *m/z* tolerance. The stoichiometric ranges for each monomer are entered in the green boxes on the left side of the application (e.g., limiting the search range to oligomers containing from 3 to 5 hexose residues). Clicking the "Search" button (green text) will evaluate the *m/z* for all combinations of monomers within the specified ranges and generate a list of mathes, which is displayed in the text panel on the right side. Many matches may be found by a search, so it is necessary to choose one before its mass and isotopolog distribution can be displayed. This is accomplished by clicking on the appropriate value in the pull-down selector in the center of the application. Doing so puts the elemental composition in the elemental composition box at the top, and then clicking on the "Evaluate Formula" button calculates the mass, etc. of the selected ion.

The text box also displays other types of information, as appropriate. Very brief help text is displayed here when the help button is clicked. The contents of the text box are chosen by checking boxes in the row just above the histogram. Thus, checking the"list spectrum" box shows a list (*m/z* - *abundance* pairs) corresponding the ion (including the selected ionizing species and derivative groups) using the specified step size and peak width. Checking the "list centroid" box shows a list of all of the centroided peaks composed of the isotoplogs that are calculated. Each

centroided peak can contain many isotopologs with slightly different masses, but the iosopologs in each peak are isobaric, that is, *they contain the same number of baryons* (protons and neutrons). Checking the "list isotopomers" box shows a list of isotopologs that were calculated, in order of abundance. By default, the most abundant isotopologs that comprise 99% of the total isotopolog population are listed, although many more (with insignificant abundances) may be calculated. All calculated isotopologs are combined to generate the centroided peak list.

One can also import experimental data for comparison to the simulated data using the following steps.
1) Clear all text from the text window
2) Paste in (tab-separated) two-column data, as: (*m/z abundance*)
3) Check the "import spectral data" and "list spectrum" boxes
4) Run the simulation.

A simulated spectrum will then be generated using the same *m/z* values as the imported spectrum, and the imported and simulated abundances will displayed side-by-side. The graphical representation at the bottom of the application will include both data sets.