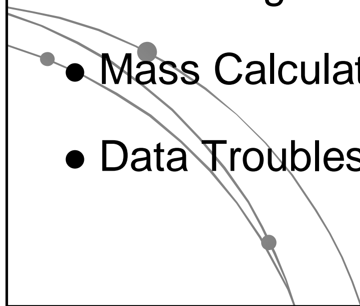
Electrospray Ionization Time-of-Flight MS for Large Molecules

Jonathan A. Karty, Ph.D.

August 4, 2008

Topics Covered

- Sample Requirements
- Instrument Overview (technical)
- Collecting Data
- Mass Calculations
- Data Troubleshooting



Some Important Caveats

- Switching to self-serve is an evolutionary process
  - We are adapting to issues as they come up
  - Procedures are likely to change in the first few months
- Much of this information is a generalization
  - Not all procedures will apply to every sample
  - However, we are open to try new things

LCT System Description

- Electrospray Ionization/Time-of-Flight (ESI-TOF) mass spectrometer
  - 75-4,000 m/z range
  - ~6,000 resolving power
  - 50 ppm mass accuracy
  - APCI source also available
- Inlet is a Waters CapLC System
  - Ternary gradient system (3 solvents)
  - 1-40 μL/min flow rate (0.25-1.0 mm i.d. columns)
  - Autosampler 0.2-5 μL injection volume on KC-379
  - 0.2-10 μL injections on on KC-366
- Waters 2487 Dual λ Detector (KC-379 only)
  - 190-600 nm UV-VIS detector
  - Can alternate between 2 wavelengths during a run

LCT Picture

What Samples Can Be Analyzed by ESI-TOF-MS?

- Non-volatile organic molecules
  - Should be soluble in polar, volatile solvents
  - Molecule must be ionizable (MS detects ions)
    - $R\text{-}NH_2$, $R\text{-}CO_2H$, $R\text{-}HSO_3$, R-OH, $R\text{-}H_2PO_3$ work best
- Biopolymers
  - Peptides, proteins, DNA, RNA
    - K, R, H residues and unblocked N-termini make positive ions
    - D, E, residues, unblocked C-termini, and phosphate backbone of nucleotides make negative ions
    - *Sequence determines which ionization mode works best*
- Organometallic complexes
  - Organometallic salts work especially well
- HPLC-MS
  - With appropriate mobile phases

What Samples Are Inappropriate for ESI-TOF-MS analysis?

- Samples in non-polar or non-volatile solvent:
  - Hexane, benzene, CH_2Cl_2, DMSO, etc.
- Buffer systems incompatible with ESI
  - 6M urea, 10% glycerol, 0.1 M NaH_2PO_4 (involatile)
  - 100 mM HCl solution (too conductive)
  - detergents
- Molecules that have no ionizable groups
- Complex mixtures yield complex spectra

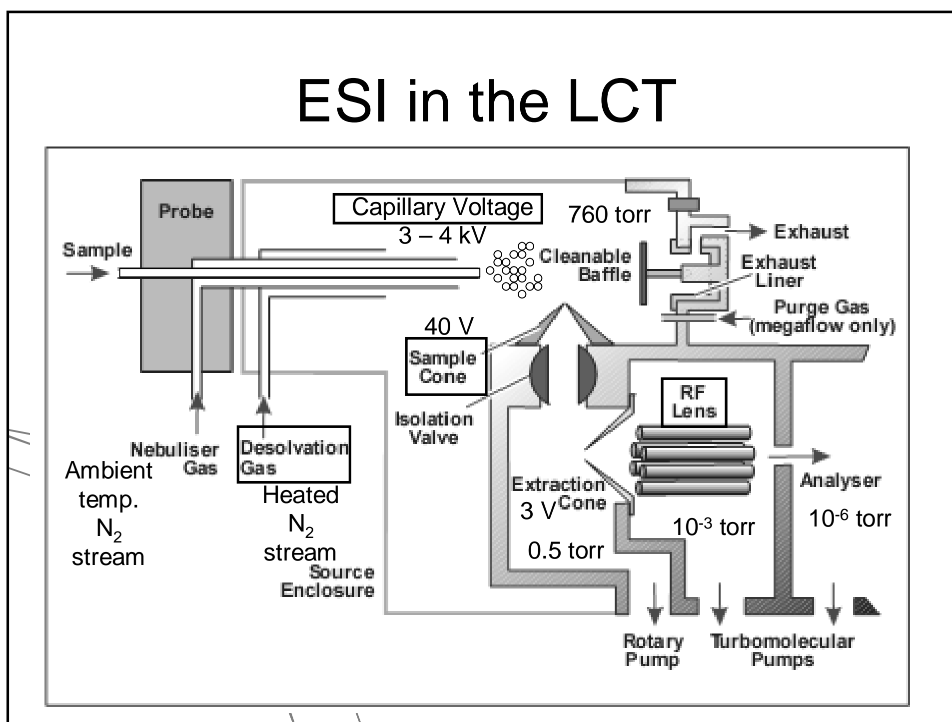
General Sample Guidelines

- **Purify analyte if possible**
  - Analyte should be 5 – 100 μM in concentration
  - ***Sample must have no particulates or precipitates***
- **Use only volatile solvents/buffers**
  - MeOH, H_2O, EtOH, acetone, CH_3CN, THF, etc.
  - HOAc, HCOOH, NH_4OAc, NH_3, etc. (weak acids/bases)
    - ***Trifluoracetic acid suppresses ionization***
  - Ionic strength < 20 mM is best (e.g. 0.1% v/v HOAc)
  - Include 20-80% volatile, polar organic solvent if possible
  - *1 – 5 μM NaOAc for acid/base labile samples*
- **Need at least 15 μL for flow injection analysis**
- **Non-aqueous samples can be analyzed by ESI**
  - THF, 1:1 $CHCl_3$:MeOH, acetone, EtOH:$CHCl_3$

Characteristics of ESI Ions

- ESI is a thermal process (1 atm in source)
  - Little fragmentation due to ionization process (cf EI)
- Solution conditions affect ESI
  - Species often observed with same charge distribution as exists in solution (e.g. organometallic salts)
  - Sample pH drastically affects charge state distribution
    - Think about pH and pKa/pI relationship
- ESI ions are mostly generated by ion transfer
  - $(M+H)^+$, $(M+Na)^+$, or $(M-H)^-$, rarely $M^{+\cdot}$ or $M^{-\cdot}$
- ESI often generates multiply charged ions
  - $(M+2H)^{2+}$ or $(M+10H)^{10+}$
  - Most biomolecules make ions between 500-1,500 m/z
- ESI molecular weights listed as mass/charge (m/z or Th, not amu or Da) 1 Th = 1 Da/z

ESI in the LCT

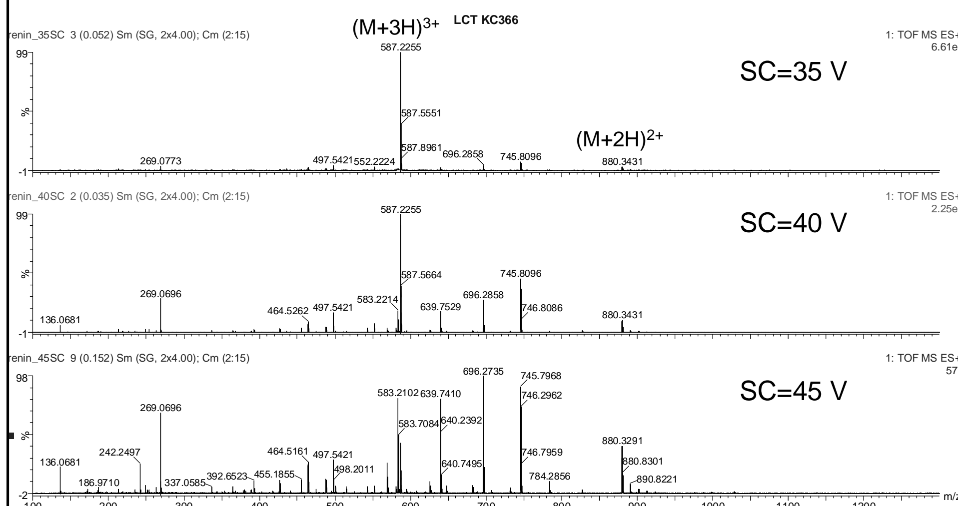


Sample Cone Voltage

- Sample Cone (SC) region at moderate vacuum
  - Typical SC value 25-85 V (default = 40 V)
  - Typical extraction cone value 3 V
  - Collisional cooling reversed in partial vacuum
    - Ion attains significant velocities between collisions
    - Collision-induced heating can cause ion to fragment
- SC voltage difference affects observed mass spectrum
  - High SC can cause fragmentation
  - High SC limits ion/solvent clusters
  - High SC can improve sensitivity (counteracts diffusion)
  - High SC alters charge state distribution to favor lower charge states (protein/DNA)

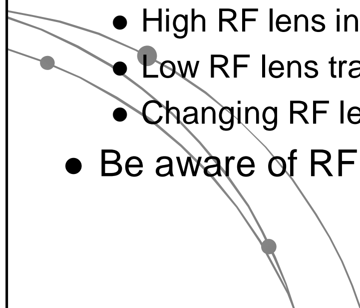
Sample Cone Example

Sequence: DRVYIHPFHLLVYS

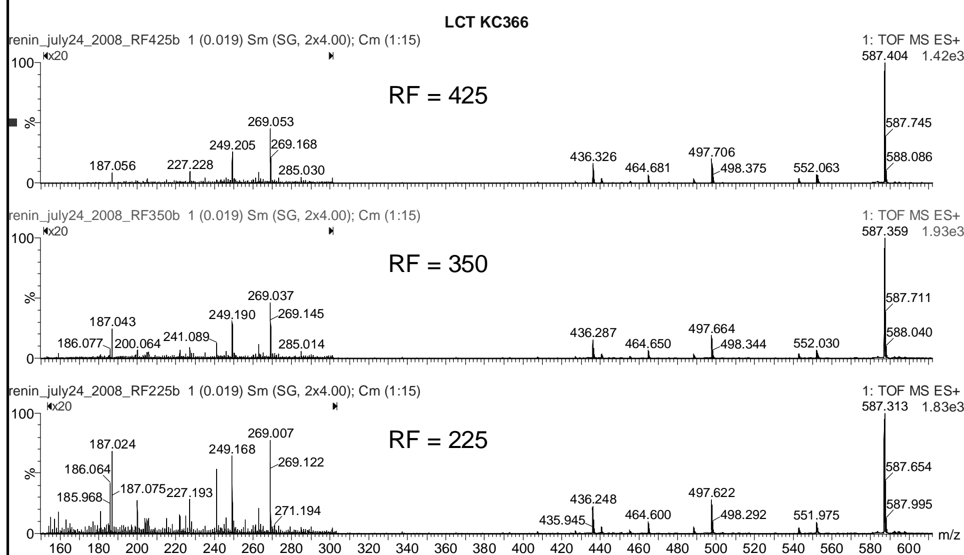


RF Lens Voltage

- Hexapole lens focuses ions into TOF
  - RF lens is amplitude of the radio frequency signal applied to the hexapole
- Although operated in RF only mode, it does induce some mass discrimination
  - High RF lens inhibits transmission of low m/z ions
  - Low RF lens transmits high m/z ions poorly
  - Changing RF lens slightly alters m/z calibration
- Be aware of RF lens voltages in your analyses



RF Lens Example



What This All Means

- Tune files are named as follows
  - Mode_polarity_CAPVOLTS_RF#_CV#_T#.ipr
  - ES_POS_3600_RF225_CV40_T125.ipr is electrospray, positive ion, 3600 V on capillary, 225 V on RF lens, 40 V on sample cone, 125 ℃ desolvation gas temperature
- MS experiment files are named:
  - Mode_polarity_RF#_time_lo#_hi#.exp
  - ES_POS_RF225_3min_150_1300.exp is electrospray, positive ion, 225 V on RF, recording from m/z 150-1300 for 3 minutes
- Be aware of the RF value for both tune files and MS experiment files
- See Jon or Angie to make new tune and/or calibration files

About the HPLC

- KC-379 is set up for reversed-phase HPLC-UV-MS
  - Solvent A is 95% H_2O, 5% CH_3CN, 0.1% HCOOH
  - Solvent B is 95% CH_3CN, 5% H_2O, 0.1% HCOOH
  - Solvent C is 1:1 CH_3CN:H_2O
  - Column is 0.5 mm i.d.x150 mm 5 µm Zorbax C18-SB300
  - UV-VIS detector has a 3 µL flow cell + tubing
    - At 15 µL/min, ~0.6 minute delay introduced by UV-VIS
- LC methods are named as follows
  - Time_grad_lambda_flow.clc
  - 20min_10-50-85_216_15.clc means a 20 minute gradient starting at 10% B, linear to 50%B, then linear to 85% B, monitoring 216 nm with a 15 µL/min flow rate

More HPLC Info

- KC-366 is set up for flow-injection analysis
  - no column, 3 min run time
  - Solvent C is 100% methanol, A and B are same as KC-379 at present
- Liquid must be at least 5 mm deep
  - Just below the 0.5 mL line on a 12 mm vial
  - Low volume insert requires only 15 μL
- It is strongly suggested that all biological samples be analyzed by ballistic LC-MS
- See Jon or Angie if you need to make up new methods, change the eluents, or try a different column

Quick Sample Prep Instructions

- Place 25 μL of water into an autosampler vial with a low volume insert
  - Vials, inserts, and caps are in blue container on the middle shelf by the balance
  - Solution is methanol with 0.5% v/v acetic acid and 5 μM sodium acetate
- Add 5 μL of your protein/peptide analyte solution
- Mix with pipette and seal
  - Label vial with sample name and/or online #
- Tap vial against table to get bubbles out of insert
- Place vial into autosampler; keep track of where you put it
- For organics/inorganics, see Small Molecule talk

About the Autosampler

- Both LCTs are configured to use HPLC pump and autosampler as sample inlet
- Solvents in HPLC and autosampler can be changed if 95:5 H_2O:CH_3CN is not compatible
- Source region is purged with N_2 gas from a liquid N_2 dewar
- Talk to Jon or Angie if you need to change HPLC solvents

Sample Entry Page

- All samples are analyzed using the HPLC
- Pull up the MassLynx window
  - Note project name to find files later
- Fill out all columns
  - File name must follow Windows XP rules
  - File text can be anything including &%' etc.
  - Sample location is always column,row (1,C)
  - Injection volume should be 2 µL
- Up to 48 samples may be analyzed at one time
- Sample ID is the online sample number
  - You can have up to 39 samples under a single number

Sample Page View

MassLynx july222008 - jul252008.SPL

File View Run Help

Shortcut Queue Status

Queue Is Empty

Instrument Tools BioLynx OpenLynx QuanLynx SDMS

Instrument

Inlet Method

MS Method

MS Tune

Edit Shutdown or Startup

Shutdown

Startup

Options

Spectrum Chromatogram Map Edit Samples

	File Name	File Text	MS File	Inlet File	Bottle	Inject...	Sample ID	MS Tune File
1	pep + chloesterol e undiluted	pep + chloest...	c18_20min_15_6...	20_min_test	1,a	5.000	15684	protein_pos
2	water blank	water blank	c18_ballistic_15_bio	C8_C18_15_ballistic	1:1,B	2.000		protein_pos
3	1-11-2 A IM	1-11-2 A IM	c18_20min_15_L...	20_min_test	1:1,C	2.000	15689	ESI_pos_lo
4	1-11-2 B IM	1-11-2 B IM	c18_20min_15_L...	20_min_test	1:1,D	2.000	15689	ESI_pos_lo
5	1-11-2 C IM	1-11-2 C IM	c18_20min_15_L...	20_min_test	1:1,E	2.000	15689	ESI_pos_lo
6	1-11-2 D IM	1-11-2 D IM	c18_20min_15_L...	20_min_test	1:1,F	2.000	15689	ESI_pos_lo
7	1-11-2 E IM	1-11-2 E IM	c18_20min_15_L...	20_min_test	1:2,A	2.000	15689	ESI_pos_lo
8	1-11-2 F IM	1-11-2 F IM	c18_20min_15_L...	20_min_test	1:2,B	2.000	15689	ESI_pos_lo
9	1-11-2 G IM	1-11-2 G IM	c18_20min_15_L...	20_min_test	1:2,C	2.000	15689	ESI_pos_lo
10	1-11-2 H IM	1-11-2 H IM	c18_20min_15_L...	20_min_test	1:2,D	2.000	15689	ESI_pos_lo
11	1-11-2 I IM	1-11-2 I IM	c18_20min_15_L...	20_min_test	1:2,E	2.000	15689	ESI_pos_lo
12	1-12-1 IM A	1-12-1 IM A	c18_20min_15_L...	20_min_test	1:2,F	2.000	15689	ESI_pos_lo
13	1-12-1 IM B	1-12-1 IM B	c18_20min_15_L...	20_min_test	1:3,A	2.000	15689	ESI_pos_lo
14	1-12-1 IM C	1-12-1 IM C	c18_20min_15_L...	20_min_test	1:3,B	2.000	15689	ESI_pos_lo
15	1-12-1 IM D	1-12-1 IM D	c18_20min_15_L...	20_min_test	1:3,C	2.000	15689	ESI_pos_lo
16	kh 1-24-1 MQ water Jon_vial	1-12-1 IM D	c18_20min_15_L...	20_min_test	1:3,E	2.000	15694	ESI_pos_lo
17	kh 1-24-1 DI water Jon_vial	1-12-1 IM D	c18_20min_15_L...	20_min_test	1:3,F	2.000	15694	ESI_pos_lo
18	kh 1-24-1 PBS Jon_vial	1-12-1 IM D	c18_20min_15_L...	20_min_test	1:4,A	2.000	15694	ESI_pos_lo
19	kh 1-24-2 LysRA off Polystyri...	kh 1-24-2 Lys...	c18_20min_15_L...	20_min_test	1:4,B	2.000	15694	ESI_pos_lo

Ready Not Scanning 0:0 Only Batch Shutdown Enabled

start LCT - c:\masslynx\... MassLynx - july22... Experiment Setup ... 20_min_testB.clc, ... 9:39 AM

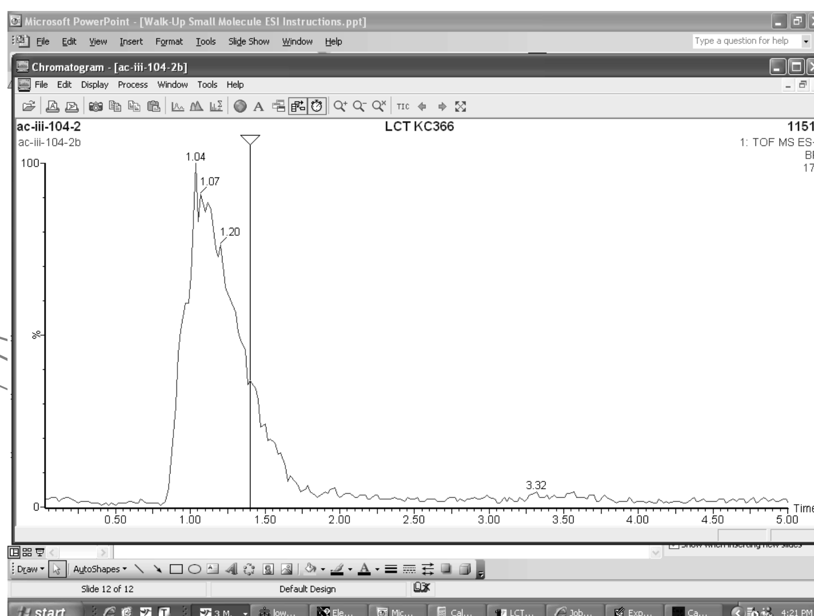
Starting the Analysis

- Save the completed sample table
- Highlight the rows you wish to run
  - Entire row should turn black when selected
- Click the “Start Analysis” icon (looks like a play button)
- Check and make sure the sample numbers in the popup are correct
- Press OK to add runs to the queue
- More samples can be added to the queue while instrument is in use

Analyzing Data

- Highlight the row containing the sample you wish to view
- Click on the “Chromatogram” link along the top of the sample list
  - A chromatogram will appear
  - One can open also files from the “Chromatogram” window
    - Files are in Project/Data directory
- The left mouse button is used for zooming in on data
  - Drawing a horizontal line expands the X-axis
  - Drawing a vertical line expands the Y-axis
  - Drawing a box makes whatever is in the box fill the screen
- The right mouse button is used to extract mass spectra
  - Dragging across a region sums the mass spectra
  - Double right-clicking extracts a single mass spectrum

Chromatogram Window



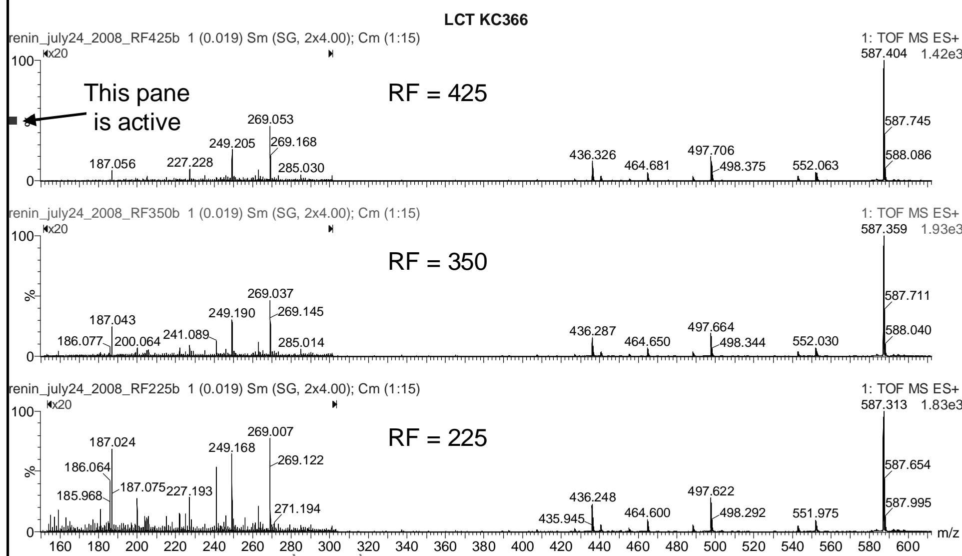
Analyzing Data 2

- Zooming in Spectrum app works just like in Chromatogram app
  - Right clicking generates a plot of the selected m/z ratio (or range) as a function of time
- Smooth the data by clicking on the smooth icon
  - 2x4 point Savitzky-Golay is adequate for most data
- Find the centers of the peaks by clicking on the centroid icon
  - 6 points wide for things <800 m/z, 10 pts for things 800-1800 m/z; 15 pts for things <1800 m/z
- Mousing over any button will cause a description of its function to appear along the bottom left border of the window

Analyzing the Data 3

- Sample text appears in upper-left corner
- Colored text below sample text describes the processes performed on data
- Camera icon copies current view to clipboard for pasting into other programs
- Overwrite icon determines whether a process creates a new pane
  - Colored dot on left indicates the active pane
  - Clicking in a pane makes it active
- Intensities can be added from the Display Menu
  - Display → Peak Annotation…

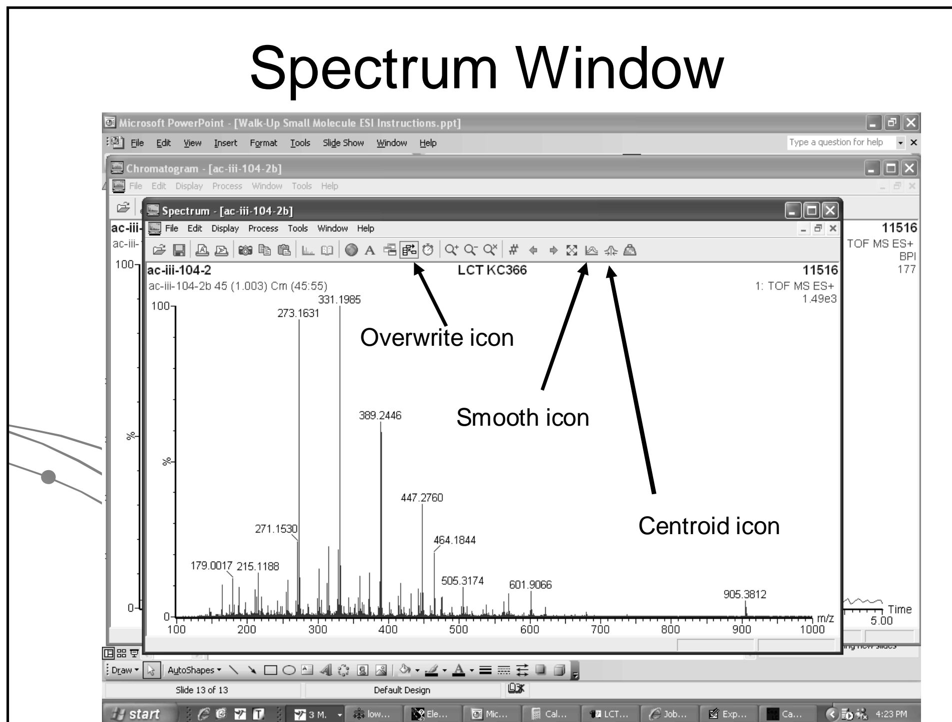
Active Pane Dot



Labels and Printing

- Clicking the “A” icon allows one to create a text box to place anywhere in a data pane
- Data can be printed or sent to a pdf
  - Select the printer using File→Printer Setup…
    - MSF-2420 is printer on MALDI instrument
    - DeskPDF creates pdfs
  - Printer icons determine portrait or landscape orientation
- Both of the LCT computers are on the IU network so you can transfer pdfs to your PC or slashtmp
  - Due to security concerns, flash drives are **NOT** permitted on MSF computers

Spectrum Window



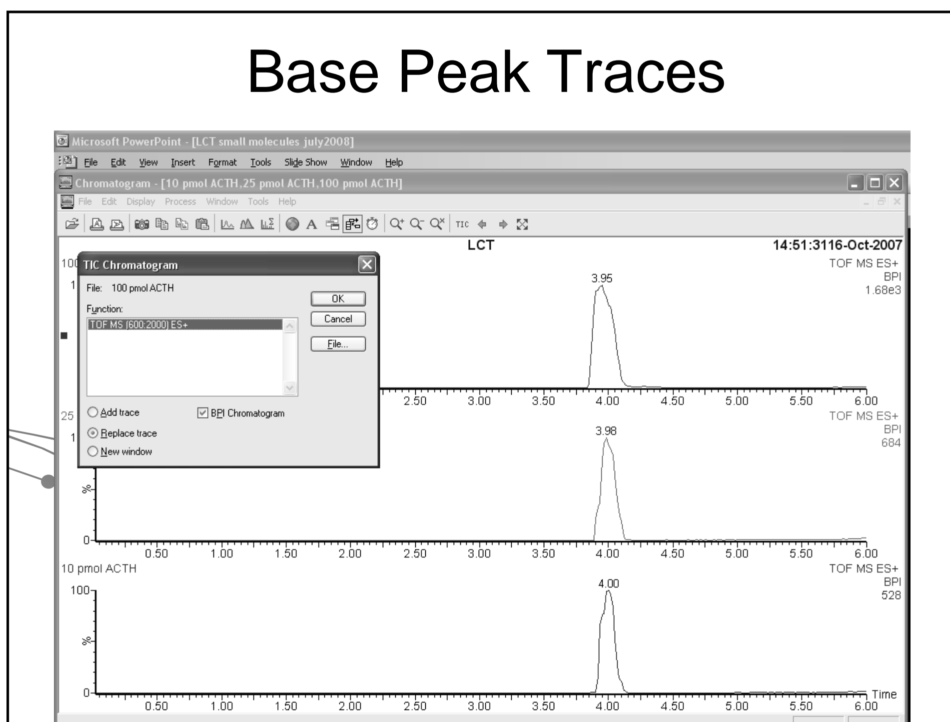
Interpreting Data Overview

- Is LCT overloaded?
- Molecular ion mass
  - $(M+H)^+$ or $(M+Na)^+$ or other
- Isotope Pattern
  - Are isotopes even resolved?
  - Deconvolution algorithm for multiply charged peaks
- Mass Defect
- Fragmentation/Multimers

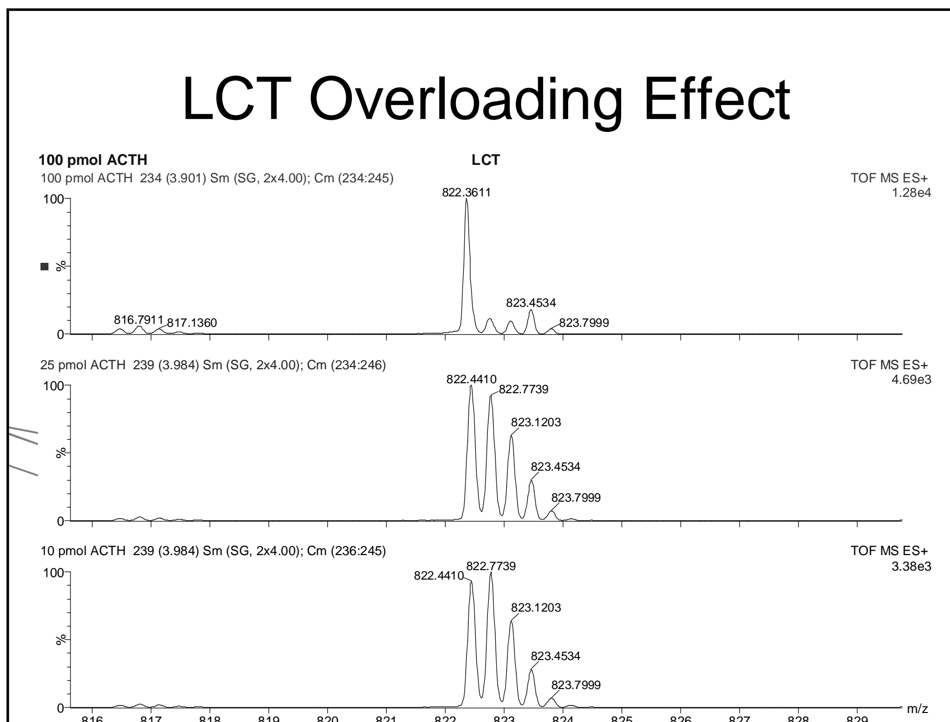
Overloading and TDC Dead Time

- Digitizer in LCT only records ion arrival times
- If 2 ions arrive simultaneously, they are counted as 1
  - Can cause peak to appear at a lower m/z
  - High ion currents also suppress sensitivity for adjacent isotope peaks
- Must have fewer than 600 counts per second in any single peak in a single scan
- Use Base Peak (BPI) trace to check for detector overloading
  - BPI tracks intensity of most intense ion in any scan as a function of time
  - Total ion chromatogram (TIC) tracks sum of all ion current in any single scan
  - Traces can be added by clicking Display→TIC

Base Peak Traces



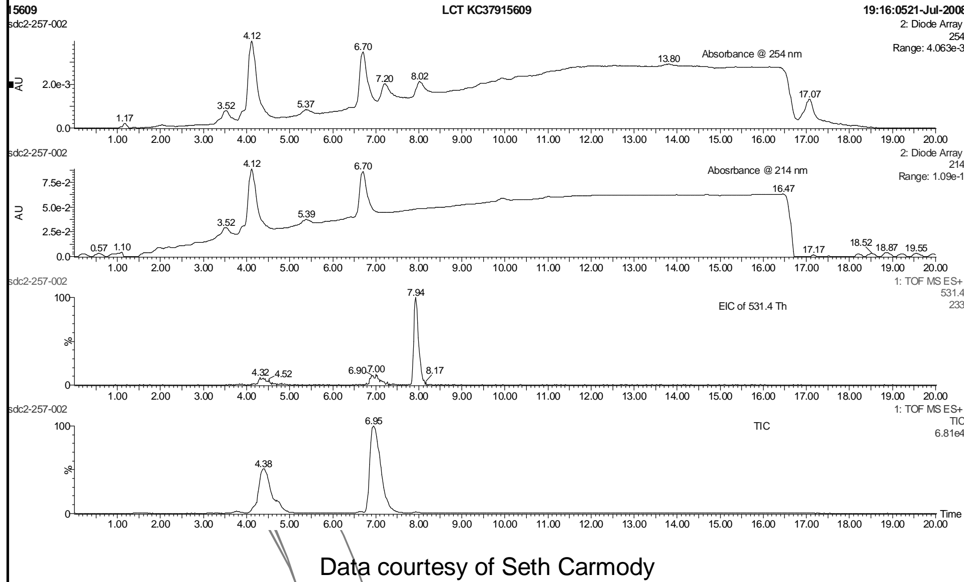
LCT Overloading Effect



Adding Other Chromatograms

- 3 other chromatogram types exist beyond TIC and BPI
- Extracted ion chromatograms (EIC)
  - EIC plots intensity of a particular m/z as a function of time
  - Create by selecting Display→Mass
  - A TIC, BPI, or EIC trace must be active to create an EIC
- Diode array trace shows sum of absorbance at both UV wavelengths
  - Create by selecting Display→TIC→Diode Array
  - Highlighting BPI plots most intense absobance
- Wavelength trace can extract individual absorbance chromatograms
  - Create by selecting Display→Wavelength
  - A Diode array or wavelength trace must be highlighted first

Alternate Chromatogram Examples



Data courtesy of Seth Carmody

Isotopes and Charge States

- Many elements have more than one naturally occurring isotope (Cl has 2, Se has 6, C has 2)
- Mass spectra show distribution of all isotopes of all atoms
  - Isotope pattern can be useful when interpreting data (indicate presence/absence of a particular element)
- Isotope peaks are spaced by ~1 Da/charge
  - Spacing can be used to discern charge state of an ion
  - e.g. @ +4, ^{13}C isotope peaks are 0.25 Th apart
- Proteins tend to make ions between 500 Th and 1500 Th
  - Modifications and/or sequence oddities will invalidate this rule

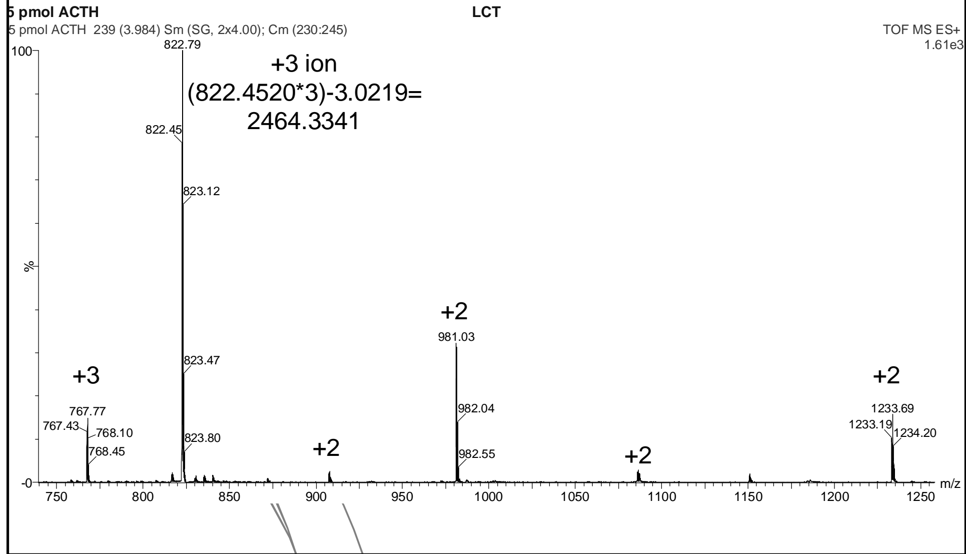
Ionization by Successive Protonation

- Most biological molecules are ionized by multiple proton exchange
  - For positive ions, m/z = $(MW+nH)^{n+}$
  - For negative ions, m/z $(MW-nH)^{n-}$
- Knowing this, we can set up the following equation to compute a protein's mass
  - $m/z_{obs} = [MW_{prot} + (n*1.0073)] / n$
    - You can do this with two successive peaks if you have no idea of charge state (n and n+1 cases)
    - Remember, you are transferring a naked proton, thus its mass is 1.0073, NOT 1.0078 (electron has mass)

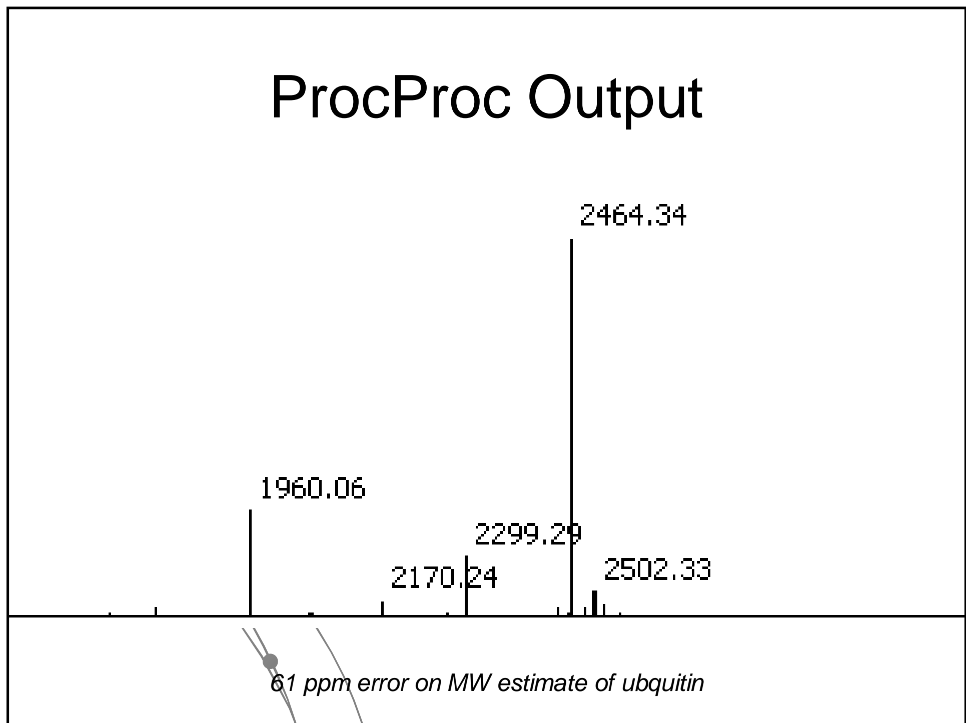
Deconvolution: ProcProc

- ProcProc is a program to deconvolute isotopic distributions from MassLynx data
- Proc files are saved spectra generated from raw MassLynx data
- ProcProc can interpret EITHER isotopically resolved data or protein charge state distributions
- In the future, it should be able to handle both data types simultaneously
- ProcProc will be handled in individual training

Resolved Isotopes and Charge States

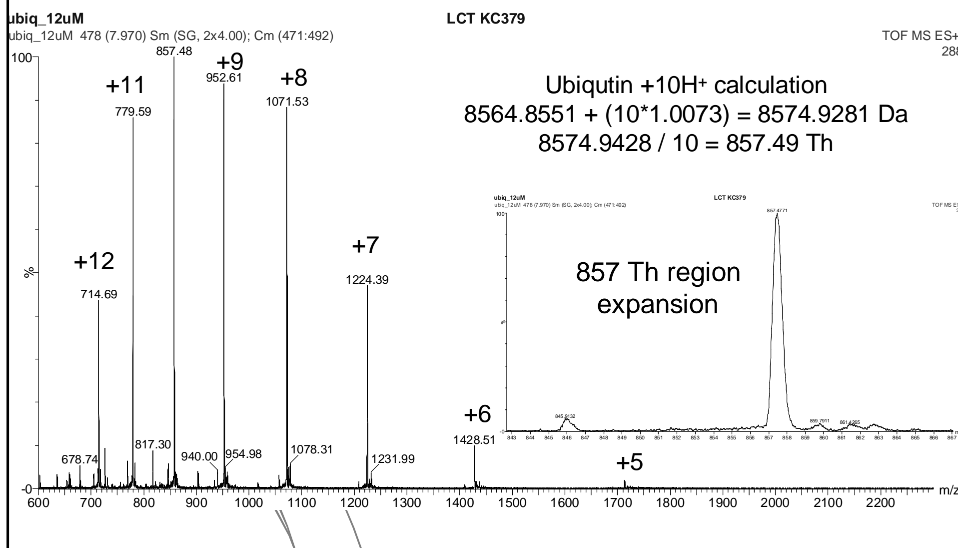


ProcProc Output

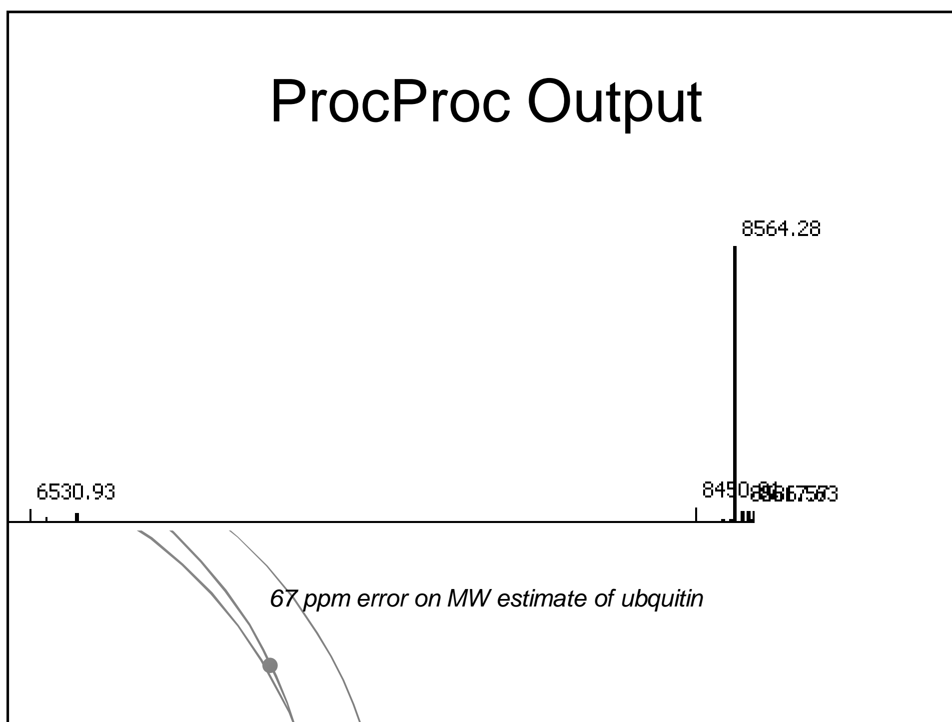


61 ppm error on MW estimate of ubquitin

Non-Resolved Isotopes



ProcProc Output



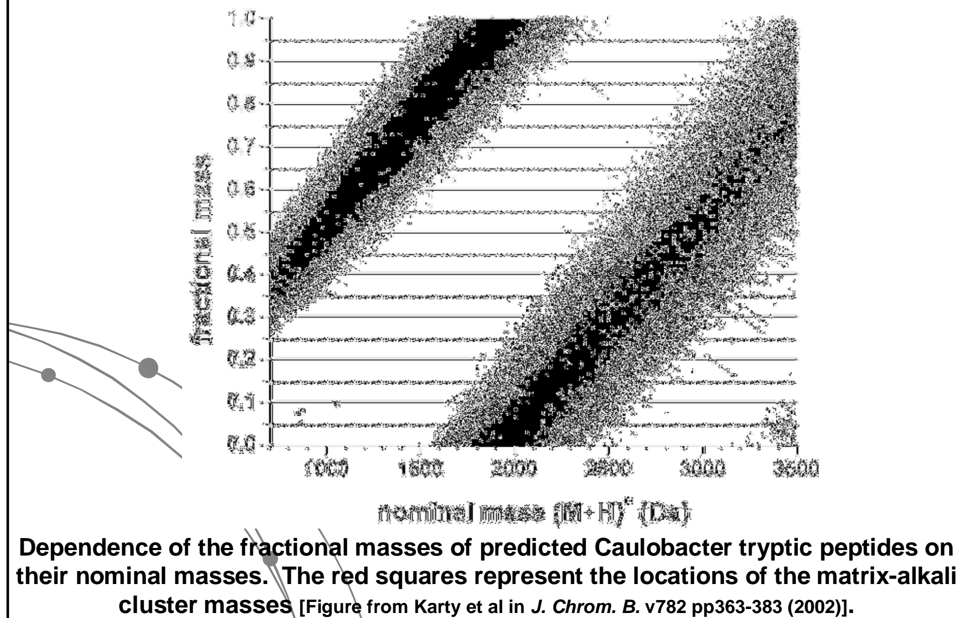
Deconvolution: ProTrawler

- ProTrawler is similar to ProcProc except that it handles entire chromatograms
- Only one data model can be used at a time (e.g. resolved isotopes)
- Ouput is a table of protein masses, intensities, and retention times
- Training will be handled individually as needed

Interpreting Data: Mass Defect

- **Atomic weights are not integers (except ^{12}C)**
  - ^{14}N = 14.0031 Da; ^{11}B = 11.0093 Da; ^{1}H = 1.0078 Da
  - ^{16}O = 15.9949 Da; ^{19}F = 18.9984 Da; ^{56}Fe = 55.9349 Da
- Difference from integer mass is called “mass defect” or “fractional mass”
  - Related to nuclear binding energy
- **Sum of the mass defects depends on composition**
  - H, N increase mass defect
    - Hydrogen-rich molecules have high mass defects
    - Eicosane ($C_{20}H_{42}$)= 282.3286
  - O, Cl, F, Na decrease it
    - Hydrogen deficient species have low mass defects
    - Morphine, ($C_{17}H_{19}NO_3$) = 285.1365

Peptide Mass Defect Figure



Dependence of the fractional masses of predicted Caulobacter tryptic peptides on their nominal masses. The red squares represent the locations of the matrix-alkali cluster masses [Figure from Karty et al in *J. Chrom. B.* v782 pp363-383 (2002)].

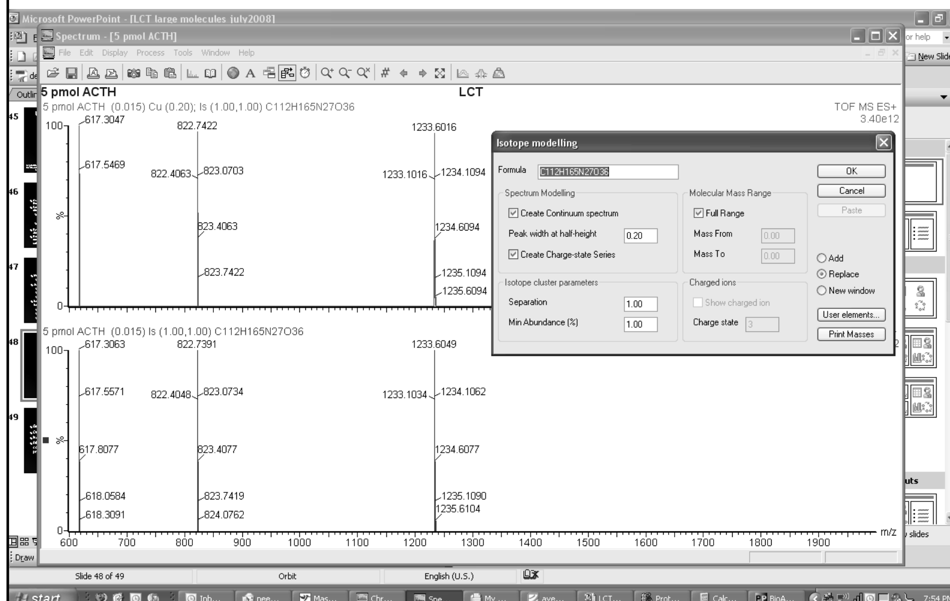
Multimers and Fragments

- As seen previously, labile ions can fragment in the sample cone region
- Conversely, strongly associating ions can make dimers and trimers
  - Gas phase multimers of protein ions usually appear as very small features exactly half-way between adjacent charge states
  - Some times, ProcProc and/or Trawler can over-estimate the intensity of gas-phase multimers
- If you suspect the presence of a multimer, talk to Jon
- If multimers are desired, sample cone voltage, desolvation temperature, capillary voltage, and solvent composition may need to be optimized

Isotope Model Module

- MassLynx has an isotopic prediction module
  - Found under Tools → Isotope Model…
- Just type in chemical formula for molecule
  - For ACTH 18-39 $(M+nH)^{n+}$, $C_{112}H_{165}N_{27}O_{36}$
  - Check status of Add or Replace buttons
- Isotope model can predict multiply charged ion isotope patterns
  - Isotope model always adds or removes protons automatically to generate charges
  - Remember, always enter in zero charge chemical composition for proteins/peptides
  - Other species, see small molecule talk

Isotope Model Dialog



Hands-on Training

- Groups of no more than three
- One hour or so to complete
- No charge for first session
- After training, students must demonstrate competency by running their own samples prior to being granted after-hours access