

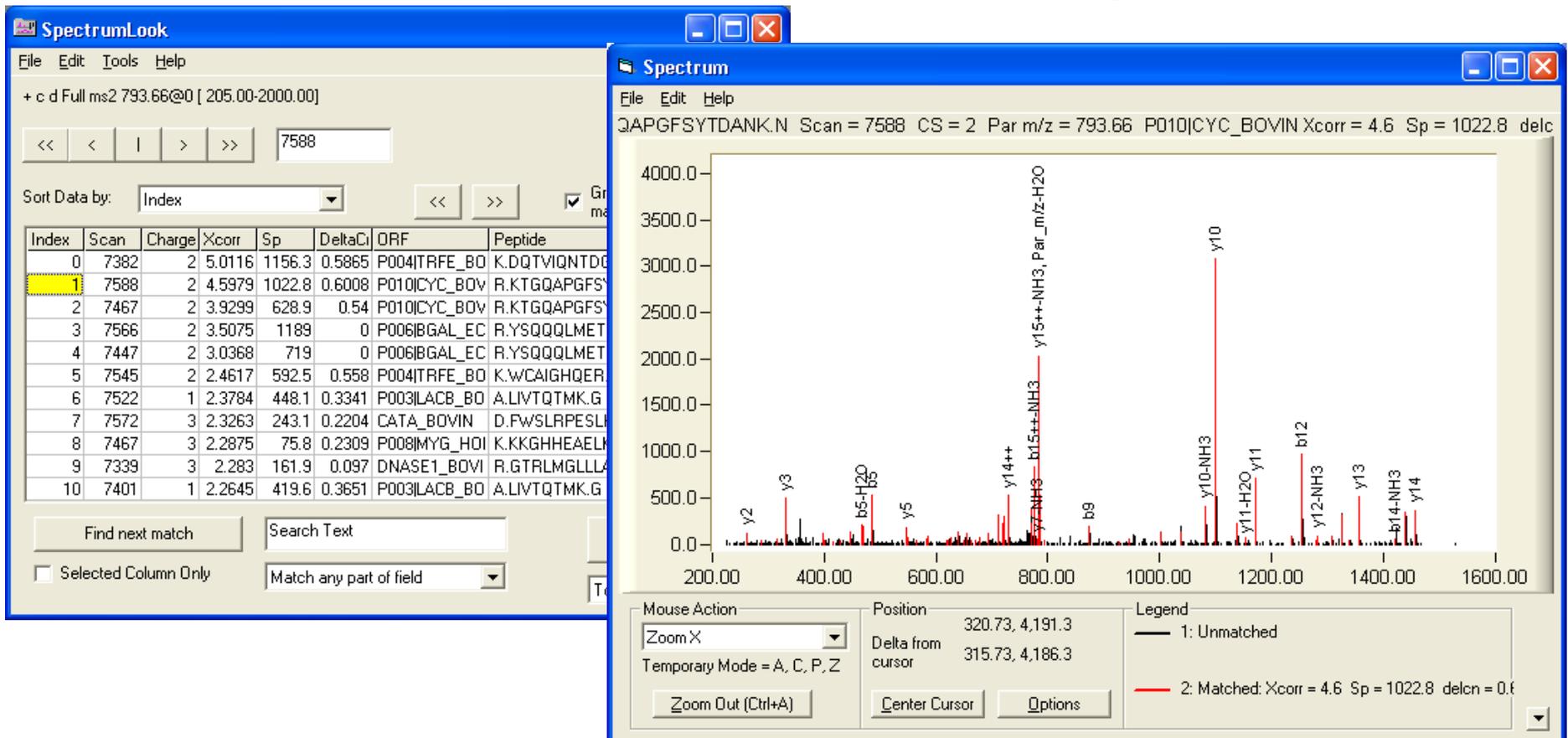
# SpectrumLook

Matthew Monroe  
Pacific Northwest National Lab  
Richland, WA

Installer available at:  
<http://omics.pnl.gov/software/SpectrumLook.php>

# SpectrumLook

- Purpose:
  - Browse MS/MS spectra and accompanying Sequest IDs
    - Labels b and y ions, plus several neutral loss masses
    - Also supports ETD data (c and z ions)
  - Distribute app with data files when submitting articles to journals



# SpectrumLook

- Windows-only program (VB6)
  - SpectrumLook.exe
  - MZDataFileReaderDLL.dll
  - MZXmlFileReaderDLL.dll
- History:
  - Written by Eric Strittmatter in 2005 to provide a browsing tool to the Phosphoproteomics researchers
  - Updated by Matthew Monroe in 2006
    - Improved user interface, adding many new features
    - Added support for mzXML and mzData files
  - Since 2006, have added support for:
    - ETD data (c and z ions)
    - Additional neutral losses
    - Additional mod symbols
    - Annotation of fragment ions up to charge state 50+

# SpectrumLook

- Input files required
  - MS Data file; supports:
    - Finnigan .Raw files (binary file)
    - mzXML files
    - mzData files
  - Tab-delimited MS/MS search result files
    - Sequest synopsis file ( \_syn.txt, PNNL-specific format)
      - Described at <http://omics.pnl.gov/software/SynopsisAndFirstHitsFiles.php>
      - Example at [http://omics.pnl.gov/software/PHRP\\_Pages/ExampleXTandemData\\_xt.txt](http://omics.pnl.gov/software/PHRP_Pages/ExampleXTandemData_xt.txt)
    - X!Tandem synopsis file ( \_xt.txt, PNNL-specific format)
      - Example at [http://omics.pnl.gov/software/PeptideFileExtractor\\_Pages/ExampleData\\_syn.txt](http://omics.pnl.gov/software/PeptideFileExtractor_Pages/ExampleData_syn.txt)
  - If available, will load a Settings file (.Ini file)

## [SpectrumLookSettings]

```
LastInputFileMode=1
DiscardMultipleORFNamesOnLoad=True
MaximumPointsPerSpectrumToKeep=750
NLBasePeakIntensityMinimum=3000
NLPeakCountMinimum=35
NLAbundanceThresholdFractionMax=0.3
MassTolerance=0.7
AbundanceThresholdFractionMax=0.01
MinMZEForDoubleCharged=500
```

# MS XML files

- Example mzXML file: text file in XML format

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<mzXML
  xmlns="http://sashimi.sourceforge.net/schema_revision/mzXML_2.0"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://sashimi.sourceforge.net/schema_revision/mzXML_2.0
http://sashimi.sourceforge.net/schema_revision/mzXML_2.0/mzXML_idx_2.0.xsd">
  <msRun scanCount="297"
    startTime="PT39.517S"
    endTime="PT100.046S">
    <parentFile fileName="file://QC_05_2_12Aug05_Firefly_0505-04.RAW"
      fileType="RAWData"
      fileSha1="b7d81c2ebc39c19c91ce54ae69d7de1258de3d61"/>
    <msInstrument>
    <dataProcessing centroided="1">
    <scan num="7301"
      msLevel="1"
      peaksCount="1574"
      polarity="+"
      retentionTime="PT2268.96S"
      lowMz="400"
      highMz="2000"
      basePeakMz="647.407"
      basePeakIntensity="94717.5"
      totIonCurrent="2.19454e+006">
    <peaks precision="32"
      byteOrder="network"
      pairOrder="m/z-
int">Q8grIkNmOPxDyKJvQ/QZrOPJBHFEHu4jQ8l4cUNWGw1DymZdQpXReEPK7IdDjFMNQ8va9kQj4t9DzFS1Q3i1MEPM8H1DtTo1
Q82NkERSFn9DzF8ZRBK4hEPOcjJC61UsQ87Z7EPm+hJDzy4EQ8V2VEPPjkREGvH+Q8/sdURLoqBD0Gy+Q8J+dkPQ7KBDv11JQ9Fz
EMPZQhD0gB3Q+/s+EPSWypDHanwQ9LCQUPZp95D00zLRDkTNEPUDB1DFrUAQ9R8hUPOyaVD1PGNRB5p80PVh0ZEAKroQ9YI2EPz0q
BD1xbBRCTF90PXsPlDznynQ9govOPbknVD2IXLR&eUWEPY85FDk8T6Q9m0OEQwpw9D2i+2RCDAXEPaxBRDBgxoQ9shNEMMRy9D257
nRD4UpUPeLXZEMNWwQ9ydsKl7h3ND3U07Q2GMkUPd2qhDjFi2Q95XJEO3eJxD3qcrQzt+xUPfH55Cx55VQ9+ELkOXS7VD4AXmREZM
OUPgiN5EUtW9Q+DjOEHDnZD4TNJQrvhRkPhrR9DbC43Q+IMPEOPa4BD4s3qQ2KvZEPj08BEL5VAQ+009U05eTVD5EcCR&GjrUPky
H9Cx2B3Q+VELkP+2IND5drsQ8t8HkPmTN5DZLpBQ+azGkQycRND5wahQxBQ3kPn8a5EDz9yQ+hrckNOu8pD6McXQ4rM3OPpSjRDkf
9iQ+u7BENrBVSD6iBQRB8HWOPgrRyD4YbZQ+tBUHLeQ1D65suQ6iJ7EPp6u1C72WcQ+u5NEPMWLLD7ME/Q2kIUQP+S75FPMWcQ+3
```

# MS XML files

## Example mzData file: text file in XML format

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- edited with XMLSPY v5 rel. 4 U (http://www.xmlspy.com) by Randy K Julian (Eli Lilly) -->
<mzData version="1.05" accessionNumber="psi-ms:100" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <cvLookup cvLabel="psi" fullName="The PSI Ontology" version="1.0.0"
  address="http://psidev.sourceforge.net/ontology/" />
  <description>
    <admin>
      <sampleName>myo 7/22 1/50</sampleName>
      <sourceFile>
        <nameOfFile>myo.RAW</nameOfFile>
        <pathToFile>http://psidev.sf.net</pathToFile>
        <fileType>XCalibur 1.3 SR1 SP2</fileType>
      </sourceFile>
      <contact>
      </contact>
    </admin>
    <instrument>
    </instrument>
    <dataProcessing>
    </dataProcessing>
  </description>
  <spectrumList count="220">
    <spectrum id="1">
      <spectrumDesc>
        <spectrumSettings>
          <acqSpecification spectrumType="discrete" methodOfCombination="sum" count="1">
            <acquisition acqNumber="1"/>
          </acqSpecification>
          <spectrumInstrument msLevel="1" mzRangeStart="300.000000" mzRangeStop="1500.000000">
            <cvParam cvLabel="psi" accession="PSI:1000036" name="ScanMode" value="MassScan"/>
            <cvParam cvLabel="psi" accession="PSI:1000037" name="Polarity" value="Positive"/>
            <cvParam cvLabel="psi" accession="PSI:1000038" name="TimeInMinutes" value="0.005833"/>
            <userParam name="ScanType" value="full"/>
          </spectrumInstrument>
        </spectrumSettings>
      </spectrumDesc>
      <mzArrayBinary>
        <data precision="32" endian="little"
        length="346">JhOWQ8b/10PMTJhd80+YQ6RVmUPK2p1DgsmaQ2BomOPa3JtDpm+cQ2z6nENWd51DstedQ1xAnOPm5Z9Dvna
        gQ7zwoEOg2qFDemOiq2S6okMur6NDHlekQwpwpUMAwaZdsuemQ3ZopOPABKhDDEcPQwSLqkMYbatDIOerQzBlrEPGTK1Deg6
        uQ5RurkPk565D/HuvQ9LNrOPygLFDihmyQ6ycskNc9bJD5GCzQxYUtEM2d7RDBEq1Q+BntkOyxzrZD3Ge3Q2pauEOIz7hDoFW
        5Qz4oukNQeLpD/G68Q2I5vUOKQb5DmJm+Q1RXvON8GMBDoMDAQ9BnwUPqBcJDgGfCQxz3wkP6VMNDPjDEQ2SGxE04d8Vdajb
        HXQDxyEP4Rc1DSA/LQzicyONMhMxD8NjMQ6agzUMQf85DdnjPQ1bWz0OyftBDMG7SQ+wN1EMc3dRD8irVQ6Kt1UMQHNZDkNH
```

# Synopsis Files

- Tab-delimited text file of Sequest or X!Tandem results
  - At PNNL, produced when DMS analyzes datasets by Sequest
- Example:

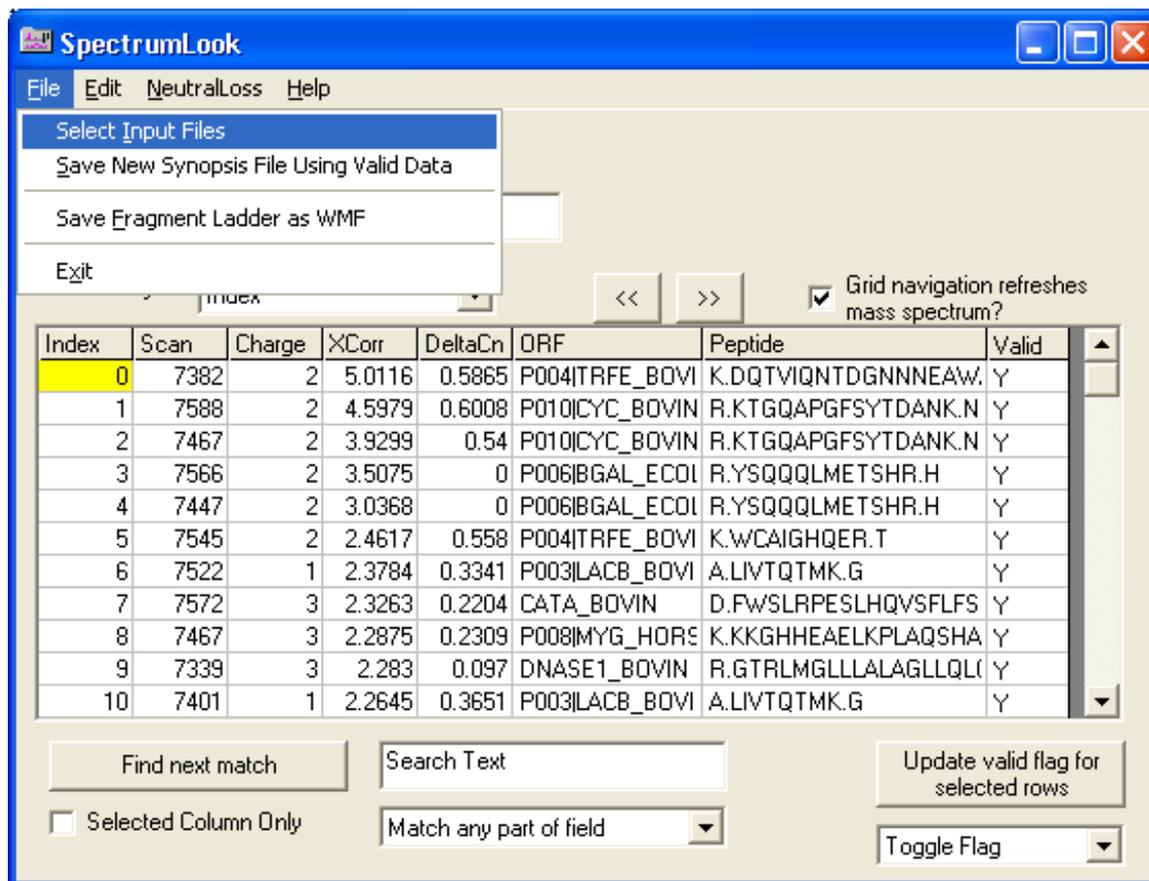
HitNum	ScanNum	ScanCount	ChargeState	MH	XCorr	DelCn	Sp	Reference	MultiProtein	Peptide
117	7382	1	2	2019.0753	5.0116	0	1156.3	P004 TRFE_BOVIN	0	K.DQTVIQNTDGNNNEAWAK.N
225	7588	2	2	1585.6997	4.5979	0	1022.8	P010 CYC_BOVIN	0	R.KTGQAPGFSYTDANK.N
482	7467	1	2	1585.6997	3.9299	0	628.9	P010 CYC_BOVIN	0	R.KTGQAPGFSYTDANK.N
686	7566	1	2	1508.6419	3.5075	0	1189	P006 BGAL_ECOLI	0	R.YSQQQLMETS.H
987	7447	1	2	1508.6419	3.0368	0	719	P006 BGAL_ECOLI	0	R.YSQQQLMETS.H
1515	7545	1	2	1100.236	2.4617	0	592.5	P004 TRFE_BOVIN	0	K.WCAIGHQER.T
1618	7522	2	1	934.1792	2.3784	0	448.1	P003 LACB_BOVIN	0	A.LIVTQTMK.G
1712	7572	1	3	2409.6867	2.3263	0	243.1	CATA_BOVIN	0	D.FWSLRPESLHQVSFLFSDRG
1780	7467	1	3	2376.7061	2.2875	0	75.8	P008 MYG_HORSE	0	K.KKGHHEAELKPLAQSHATKH
1791	7339	1	3	2769.4284	2.283	0	161.9	DNASE1_BOVIN	0	R.GTRLMGLLLALAGLLQLGLSLI
1835	7401	1	1	934.1792	2.2645	0	419.6	P003 LACB_BOVIN	0	A.LIVTQTMK.G
1976	7407	1	2	1143.3187	2.1996	0	698.7	P005 G3P_RABIT	0	K.AENGKLVINGK.A
2248	7405	1	1	934.1792	2.1158	0	401.5	P003 LACB_BOVIN	0	A.LIVTQTMK.G
2321	7339	1	2	1849.9829	2.0942	0	91.5	CATA_BOVIN	0	A.GESGSADTVRDPRGFAVK.F
2395	7382	1	2	2018.0888	2.0724	0.5865	211.9	K2C1_HUMAN	0	K.QISNLQQSISDAEQRGEN.A
2428	7337	1	3	2464.7189	2.0646	0	200.6	AMYG_ASPNG	0	S.FILANFDSSRSRGKDANTLLGSI
2438	7339	1	3	2769.1031	2.0615	0.097	277.7	P012 PHS2_RABIT	0	K.TIFKDFYELEPHKFQNKTNGIT
2447	7335	1	2	1585.7646	2.0602	0	342.7	DNASE1_BOVIN	0	T.TATSTNCAYDRIVVA.G
2499	7531	1	1	1585.6997	2.052	0	277.2	P010 CYC_BOVIN	0	R.KTGQAPGFSYTDANK.N

See also

<http://omics.pnl.gov/software/SynopsisAndFirstHitsFiles.php>

# SpectrumLook

- Loading data:
  - File->Select Input Files
  - Choose a .Raw file or .mzXml file
  - Choose the corresponding \_syn.txt file
    - Does not yet support X!Tandem files, but will be easy to add



The screenshot shows the SpectrumLook application window. The menu bar includes File, Edit, NeutralLoss, and Help. The File menu is open, showing options: Select Input Files, Save New Synopsis File Using Valid Data, Save Fragment Ladder as WMF, and Exit. Below the menu is a toolbar with navigation buttons (left and right arrows) and a checked checkbox for 'Grid navigation refreshes mass spectrum?'. The main area contains a table with 10 rows of search results. The first row is highlighted in yellow. Below the table are controls for finding matches, including a 'Find next match' button, a 'Search Text' input field, a 'Match any part of field' dropdown, and buttons for 'Update valid flag for selected rows' and 'Toggle Flag'.

Index	Scan	Charge	XCorr	DeltaCn	ORF	Peptide	Valid
0	7382	2	5.0116	0.5865	P004ITRFE_BOVI	K.DQTVIQNTDGNNNEAW	Y
1	7588	2	4.5979	0.6008	P010ICYC_BOVIN	R.KTGQAPGFSYTDANK.N	Y
2	7467	2	3.9299	0.54	P010ICYC_BOVIN	R.KTGQAPGFSYTDANK.N	Y
3	7566	2	3.5075	0	P006JBGAL_ECOI	R.YSQQQLMETS.H	Y
4	7447	2	3.0368	0	P006JBGAL_ECOI	R.YSQQQLMETS.H	Y
5	7545	2	2.4617	0.558	P004ITRFE_BOVI	K.WCAIGHQER.T	Y
6	7522	1	2.3784	0.3341	P003LACB_BOVI	A.LIVTQTMK.G	Y
7	7572	3	2.3263	0.2204	CATA_BOVIN	D.FWSLRPESLHQVSLFS	Y
8	7467	3	2.2875	0.2309	P008MYG_HORSE	K.KKGHHEAELKPLAQSHA	Y
9	7339	3	2.283	0.097	DNASE1_BOVIN	R.GTRLMGLLLALAGLLQL	Y
10	7401	1	2.2645	0.3651	P003LACB_BOVI	A.LIVTQTMK.G	Y

# SpectrumLook

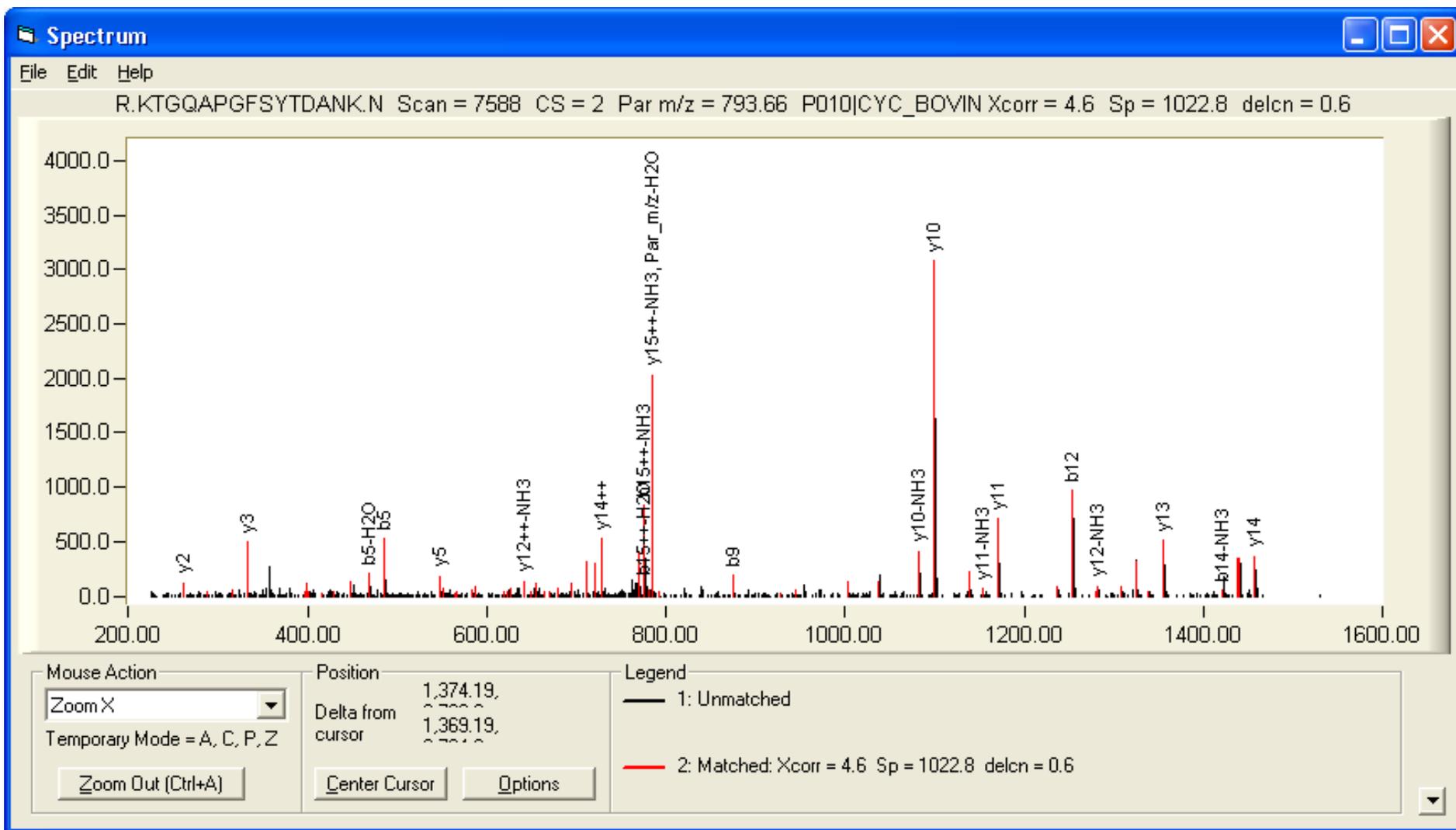
- Grid shows peptide IDs
  - Can change sort order
  - Changing rows updates plot
  - Can mark data as valid/invalid by double-clicking Valid column
  - Can save a new Synopsis file using only the Valid data
  - Can edit peptide sequences to adjust the residue containing a modification symbol

The screenshot shows the SpectrumLook application window. The title bar reads "SpectrumLook". The menu bar includes "File", "Edit", "NeutralLoss", and "Help". The main window displays a mass spectrum plot area with a search bar containing "7467". Below the search bar, the "Sort Data by:" dropdown is set to "Index". A checkbox labeled "Grid navigation refreshes mass spectrum?" is checked. The main data area is a table with the following columns: Index, Scan, Charge, XCorr, DeltaCn, DRF, Peptide, and Valid. Row 8 is highlighted in yellow. Below the table, there is a search box containing "kkggh" and a "Find next match" button. A checkbox labeled "Selected Column Only" is unchecked. A dropdown menu is set to "Match any part of field". A "Toggle Flag" dropdown is also visible. A button labeled "Update valid flag for selected rows" is present.

Index	Scan	Charge	XCorr	DeltaCn	DRF	Peptide	Valid
0	7382	2	5.0116	0.5865	P004ITRFE_BOVI	K.DQTVIQNTDGNNNEAW.	Y
1	7588	2	4.5979	0.6008	P010ICYC_BOVIN	R.KTGQAPGFSYTDANK.N	Y
2	7467	2	3.9299	0.54	P010ICYC_BOVIN	R.KTGQAPGFSYTDANK.N	Y
3	7566	2	3.5075	0	P006IBGAL_ECOL	R.YSQQQLMETS.H	Y
4	7447	2	3.0368	0	P006IBGAL_ECOL	R.YSQQQLMETS.H	Y
5	7545	2	2.4617	0.558	P004ITRFE_BOVI	K.WCAIGHQER.T	Y
6	7522	1	2.3784	0.3341	P003LACB_BOVI	A.LIVTQTMK.G	Y
7	7572	3	2.3263	0.2204	CATA_BOVIN	D.FWSLRPESLHQVSFLFS	Y
8	7467	3	2.2875	0.2309	P008MYG_HORSE	K.KKGHHHEELKPLAQSHA	Y
9	7339	3	2.283	0.097	DNASE1_BOVIN	R.GTRLMGLLLLALAGLLQLI	Y
10	7401	1	2.2645	0.3651	P003LACB_BOVI	A.LIVTQTMK.G	Y

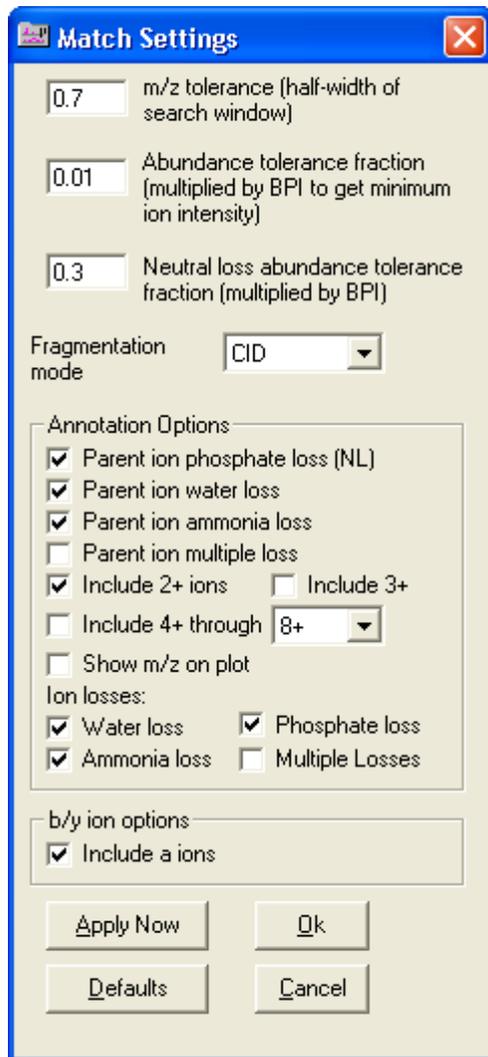
# SpectrumLook

- Spectrum shows MS/MS data with black bars and matched data with red bars



# SpectrumLook

- Can update the match settings, neutral losses to find, and masses associated with modification symbols



**Match Settings**

m/z tolerance (half-width of search window)

Abundance tolerance fraction (multiplied by BPI to get minimum ion intensity)

Neutral loss abundance tolerance fraction (multiplied by BPI)

Fragmentation mode:

Annotation Options:

- Parent ion phosphate loss (NL)
- Parent ion water loss
- Parent ion ammonia loss
- Parent ion multiple loss
- Include 2+ ions  Include 3+
- Include 4+ through
- Show m/z on plot

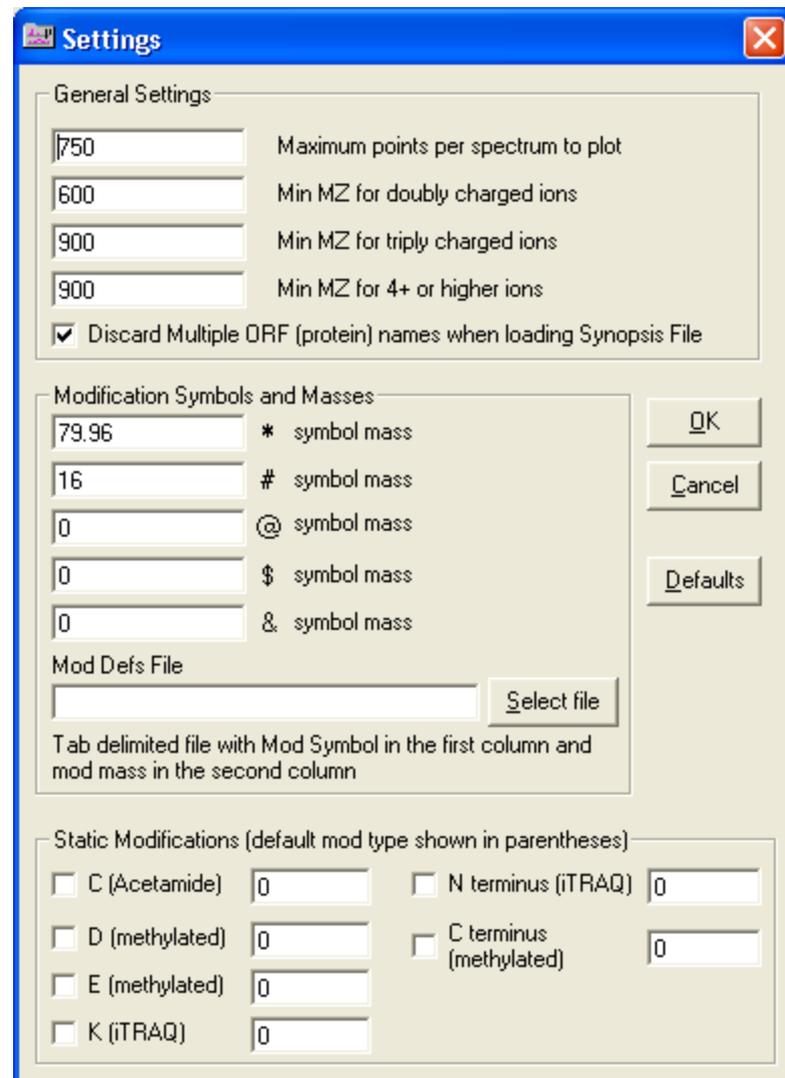
Ion losses:

- Water loss  Phosphate loss
- Ammonia loss  Multiple Losses

b/y ion options:

- Include a ions

Buttons:



**Settings**

General Settings:

- Maximum points per spectrum to plot
- Min MZ for doubly charged ions
- Min MZ for triply charged ions
- Min MZ for 4+ or higher ions
- Discard Multiple ORF (protein) names when loading Synopsis File

Modification Symbols and Masses:

- \* symbol mass
- # symbol mass
- @ symbol mass
- \$ symbol mass
- & symbol mass

Buttons:

Mod Defs File:

Tab delimited file with Mod Symbol in the first column and mod mass in the second column

Static Modifications (default mod type shown in parentheses):

- C (Acetamide)
- D (methylated)
- E (methylated)
- K (iTRAQ)
- N terminus (iTRAQ)
- C terminus (methylated)



# SpectrumLook

- Able to provide some support and bug fixes, but no major updates
- Need to add support for mzML files
- Would like to migrate program to VB.NET (or even Java), but it is not an active development project