Skyline Importing Integration Boundaries from Other Tools

An important step in SRM data processing, and to an even greater extent for chromatograms extracted from data independent mass spectra (MS1 in DDA or MS/MS in DIA), is selecting the correct time range over which to integrate chromatogram peak areas for a desired peptide. In data independent methods especially, which can involve interrogating thousands or even tens of thousands of peptides per file, it is increasingly important to do this peak selection with tools that produce well calibrated probabilities, rather than relying on manual inspection to detect errors. Development of better tools for fully-automated peptide chromatogram peak assignment is an active area of current research[1, 2]. To assist in development, testing, and use of these tools, Skyline now supports the ability to import chromatogram peak boundaries (i.e. the start and end times used in peak area integration) determined by other tools (e.g. OpenSWATH or mProphet). This feature can be used to benchmark or visualize the performance of other tools, or simply to incorporate their results into a Skyline-based workflow.

To use the Import Peak Boundary feature in Skyline, you must first prepare a text file containing a list of the peak boundaries to be imported and the data file and peptide corresponding to each. The first line of the file should contain column names, which Skyline uses to figure out which column is which. The columns can be in any order as long as they are correctly named, and extraneous columns are not a problem (Skyline simply ignores them). The table below shows the columns Skyline expects to see, as well as what each column means and the allowed names for each column. Some columns are required, while others are optional.

Column	Acceptable Names	Meaning	Optional or Required?
Starting Peak Time	MinStartTime / leftWidth	Time, in minutes or seconds, at which the start of the peak occurs.	Required
Ending Peak Time	MaxEndTime / rightWidth	Time, in minutes or seconds, at which the end of the peak occurs	Required
File Name	FileName/ filename	File corresponding to this peak group. The file must match the name of the file imported into Skyline.	Required

Peptide	ModifiedPeptideSequence / FullPeptideName	Modified sequence of the peptide. Can be in name, mass, or unimod notation (see below).	Required
Precursor Charge State	PrecursorCharge / Charge	For which precursor charge state should the peak boundaries be adjusted? If this column is not specified, all precursor charge states of the peptide are adjusted.	Optional
Decoy Status	PrecursorIsDecoy / IsDecoy	For documents where the same peptide occurs twice as a normal and decoy peptide. 0/False means decoy, 1/True means non-decoy.	Optional
Sample Identifier	SampleName	For specifying the sample, in the case where one Skyline "file" actually represents multiple injections and it is important to distinguish them.	Optional

Notice that in the table above, each column can be called by multiple names, any of which Skyline will accept. This is to allow the results of many different tools (which give their output columns different names) to be easily imported. The following is a typical text file for importing, and is based on a text file output by the OpenSWATH tool:

```
OpenSWATH_scratch.csv - Notepad
File  Edit  Format  View  Help
decoy   filename            Charge  leftwidth           rightwidth      FullPeptideName
FALSE   napedro_L120417_001_Sw.mzml    2    1003.9199829102 1061.9499511719 GSMADVPK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    2513.8898925781 2537.7900390625 APIPTALDTDSSK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    4920.6899414063 4924.1000976563 DITAFDETLFR(UniMod:267)
FALSE   napedro_L120417_001_Sw.mzml    2    2384.1599121094 2411.4699707031 LNTIYQNDLTK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    4214.009765625  4227.669921875  GDSSLLLAVTEVK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    1045.9000244141 1062.9699707031 ITVDDSDQGANAK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    3261.5400390625 3268.3701171875 LDLSEGEVAATVK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    4524.6801757813 4551.990234375  LGGWDITGPWAK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    2032.5300292969 2097.3898925781 TGDGSDVTSDFTK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    1742.3499755859 1745.7600097656 WTGHDVTVVQR(UniMod:267)
FALSE   napedro_L120417_001_Sw.mzml    2    5702.4702148438 5716.1298828125 LDASLPALLLIR(UniMod:267)
FALSE   napedro_L120417_001_Sw.mzml    2    2189.5700683594 2223.7099609375 EDNTGVVSQTFK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    2606.0700683594 2612.8898925781 EDQTLHYDFR(UniMod:267)
FALSE   napedro_L120417_001_Sw.mzml    2    2404.6398925781 2435.3701171875 LGAIDSTLSDESK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    1971.0799560547 1974.4899902344 LGAIDSTLSDQSK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    3200.0900878906 3213.7399902344 YYDYTLSINGK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    3135.2199707031 3145.4699707031 DSTIEAANLAGLK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    3244.4699707031 3264.9499511719 DSTIQAANLAGLK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    2759.6899414063 2797.25 ILDETLYENAK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    1479.4699707031 1499.9499511719 TYEDGSSVEYR(UniMod:267)
FALSE   napedro_L120417_001_Sw.mzml    2    4063.8000488281 4074.0500488281 GFAFVQYVNER(UniMod:267)
FALSE   napedro_L120417_001_Sw.mzml    2    3647.3100585938 3681.4499511719 GLPITITESDIR(UniMod:267)
FALSE   napedro_L120417_001_Sw.mzml    2    2145.1899414063 2192.9799804688 VGEPVALSEEER(UniMod:267)
FALSE   napedro_L120417_001_Sw.mzml    2    3097.669921875  3124.9799804688 FATSDLNDLYR(UniMod:267)
FALSE   napedro_L120417_001_Sw.mzml    2    4272.0498046875 4316.4301757813 IAEVFEELTGSK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    3654.1298828125 3708.7600097656 IISDFEEDLEK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    2947.4599609375 2950.8701171875 TGGDEFDEAIIK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    3995.5200195313 4012.5900878906 TSIALDTILNQK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    1776.4899902344 1834.5200195313 GDVQAVAYEEPK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    2128.1201171875 2145.1899414063 MDVSGQETDWR(UniMod:267)
FALSE   napedro_L120417_001_Sw.mzml    2    4005.7700195313 4036.4899902344 SGIHTLQELLSK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    3101.0900878906 3111.330078125  FGTEIITETVSK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    4592.9599609375 4610.0297851563 GSVNLSLPATLIK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    3070.3601074219 3097.669921875  GTAVVNGAFEEIK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    2899.6599121094 2920.1398925781 SNSYFFQNPPK(UniMod:259)
FALSE   napedro_L120417_001_Sw.mzml    2    4043.3200683594 4084.2900390625 EAQLLFDPVTR(UniMod:267)
FALSE   napedro_L120417_001_Sw.mzml    2    2158.8400878906 2203.2199707031 IDTPGEADYYR(UniMod:267)
```

This text file is tab-separated, but Skyline will also accept space, comma, or (if European language settings are enabled) semicolon as a separator. Peak Boundary times are specified in seconds here, but can also be specified in minutes (Skyline is able to detect which units were used). The columns can be in any order, and additional (extraneous) columns will be ignored. As noted above, peptide modified sequences can be expressed in many different ways, so GSMAFVPK(UniMod:259), GSMAFVPK[+8.0], and GSMAFVPK[Label (K) 6] are all acceptable.

Peak Boundary times of "#N/A" may also be entered; this will cause Skyline to remove peak information for that peptide.

To import a text file specifying peak integration, simply do the following in Skyline:

- On the **File** menu, choose **Import** and click **Peak Boundaries**
- Select the file you want to import in the directory tree, and click **Open**.

A progress bar should appear, and when it is done, the document's peak boundaries will be adjusted to match those specified in the file:
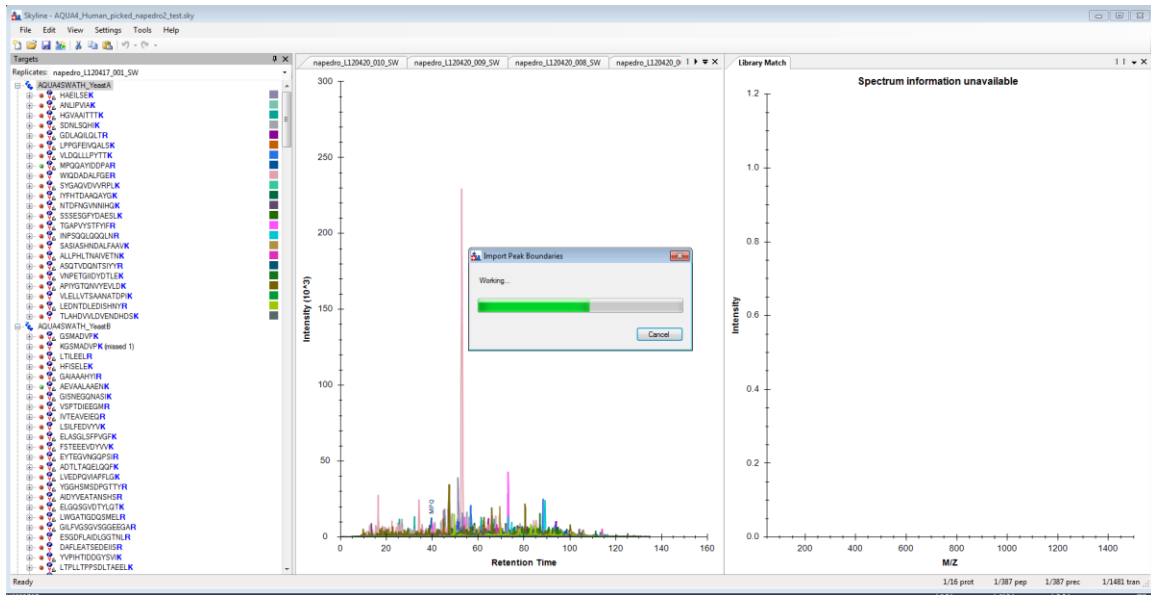
At present, only peak boundaries exported from OpenSWATH or Skyline itself can be imported without any change to the file. Other tools currently require some re-formatting of the text file, although the Skyline team is working to make external import easier for all peak picking tools. If there is a peak picking tool you would like us to support, please contact the Skyline team.

Bibliography

1. Reiter, L. *et al.* mProphet: automated data processing and statistical validation for large-scale SRM experiments. *Nat. Methods* **8,** 430–435 (2011).
2. Hannes L. Röst *et al.* OpenSWATH: A tool for the automated, targeted analysis of data-independent acquisition (DIA) MS-data. *Nat. Biotechnol.* doi:accepted