Skyline Small Molecule Quantification

The Skyline Targeted Proteomics Environment provides informative visual displays of the raw mass spectrometer data you import into your Skyline documents. Originally developed for proteomics use, Skyline has been extended to work with non-proteomic molecules. This tutorial explores a relatively straightforward example of using Skyline for targeted quantification of a single small molecule using an external calibration curve and stable-isotope labeled internal standard.

In this tutorial, you will learn about Targeted Quantification based on TQ-MS (in this example, out of crashed plasma) starting from a method you may already be running (e.g. a pharmacokinetic assay). In the analysis of this dataset you will learn:

* Insertion of simple set of known transitions
* Data Analysis and peak integration for non-proteomic molecules
* Small Molecule Quantification workflow in Skyline

You may also wish to view the first half of Skyline webinar 16, on which this tutorial is based: <https://skyline.ms/project/home/software/Skyline/events/2017%20Webinars/Webinar%2016/begin.view?>

Skyline aims to provide a vendor-neutral platform for targeted quantitative mass spec research. It can import raw data from the instrument vendors Agilent, SCIEX, Bruker, Shimadzu, Thermo-Scientific and Waters. The ability to import data across various instrument platforms greatly facilitates cross-instrument comparisons and large multi-site studies. This remains equally true in using it to target small molecules, as it has been for years in the field of proteomics.

If you have not already looked at the “Skyline Small Molecule Targets” tutorial you should do so now, in order to pick up a few basics about how Skyline works with small molecule descriptions including chemical formulas and adducts.

# Getting Started

To start this tutorial, download the following ZIP file:

https://skyline.gs.washington.edu/tutorials/SmallMoleculeQuantification.zip

Extract the files in it to a folder on your computer, like:

C:\Users\bspratt\Documents

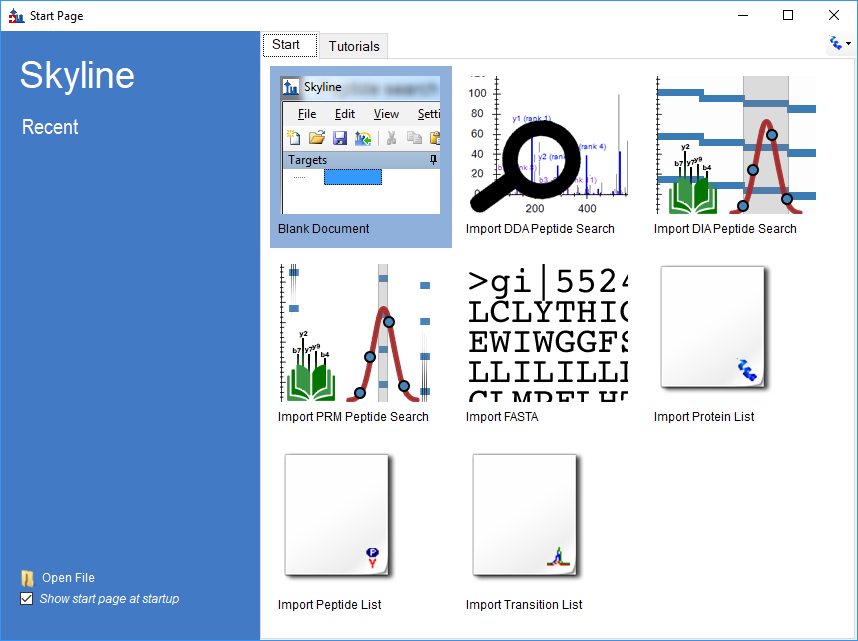
This will create a new folder:

C:\Users\bspratt\Documents\SmallMoleculeQuantification

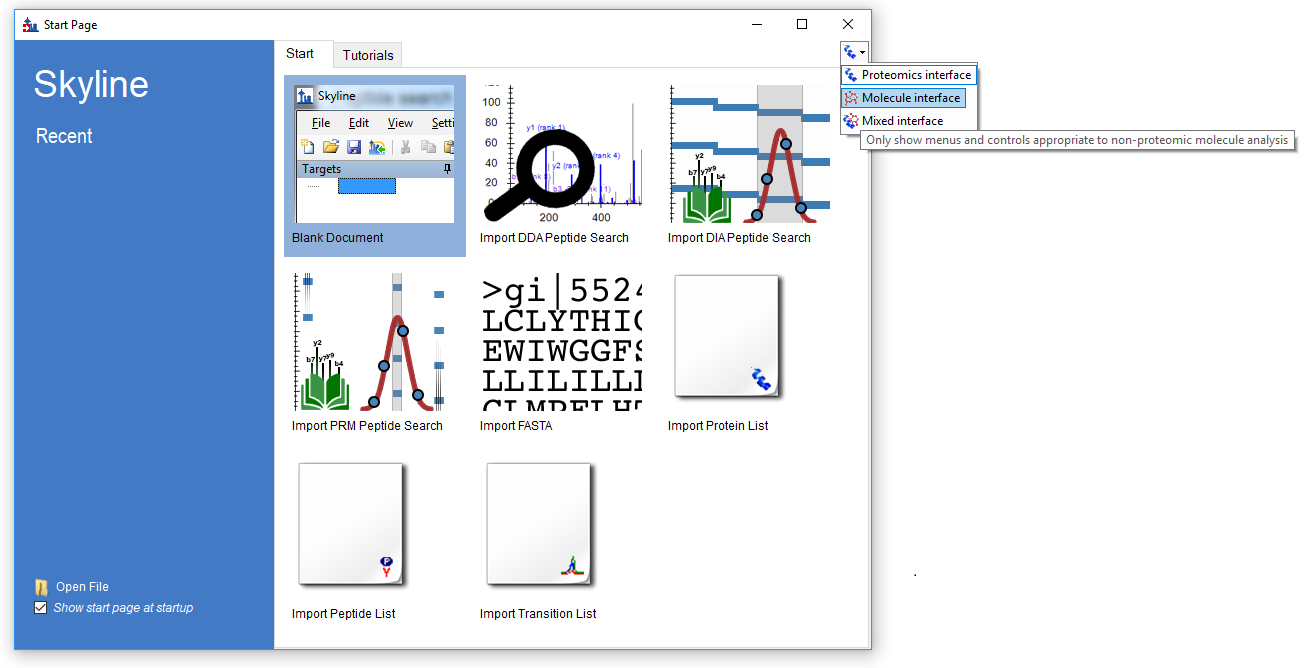
It will contain all the files necessary for this tutorial.

Now:

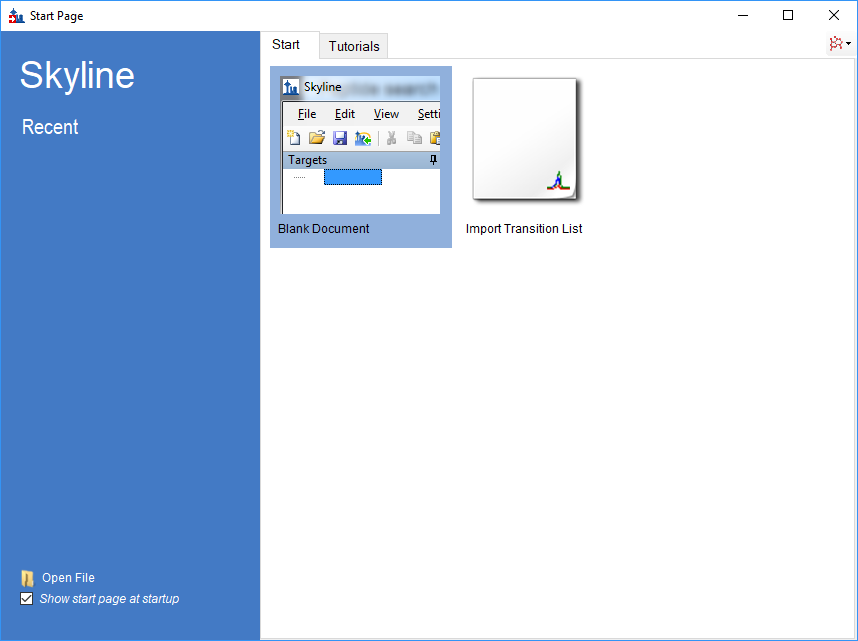
* Start Skyline. You should see the Start page. It will look something like this:



* Set the user interface control in the upper right corner of the Start Page to “Molecule interface”.



Now the start page is not cluttered with proteomics-related controls that we don’t need.



* Click on “Blank Document”.

## Experimental Layout

This experiment was designed according to the FDA Guidance on Bioanalytical Method Validation, and as such contains much more than just the study samples. A full description of the plate layout and run order typically utilized for such a study has been published (<https://www.ncbi.nlm.nih.gov/pubmed/29039849>). Briefly, the samples for this dataset were laid out in a 96 well plate as follows:



Blanks, or ‘zero’ standards, contain only the internal standard, double blanks contain no standard at all.

Calibration curve samples are a dilution series for calibration.

QC samples are “known unknowns”. These are quality control samples which we will treat as unknowns in our study. In actuality, we know what the results should be and thus they can be used to check the accuracy of our measurements.

Serum SPQC is a serum pooled QC, a pooling of all study samples which is run at several points at the start, middle and end of the experiment to verify that quantitative reproducibility is constant across the study.

The NIST SRM 1950 is a pooled plasma standard from the National Institute for Standards and Technology, which is available to all researchers as a reference standard for ‘normal’ plasma metabolite measurements. It serves as a reference between studies in different laboratories.

Injections were performed in this order:



In all, 113 injections were used in collecting the mass spec data for these samples.

## Internal Standards

For this study we have just two targets: a molecule and an internal standard, which is an isotopically labeled variant of the molecule and thus co-elutes. It is also possible to establish a relationship between unrelated molecules by declaring one of the molecules as a surrogate standard. The surrogate standard method is covered in the “Skyline High Resolution Metabolomics” tutorial.

## Importing a Small Molecule Transition List into a Skyline Document

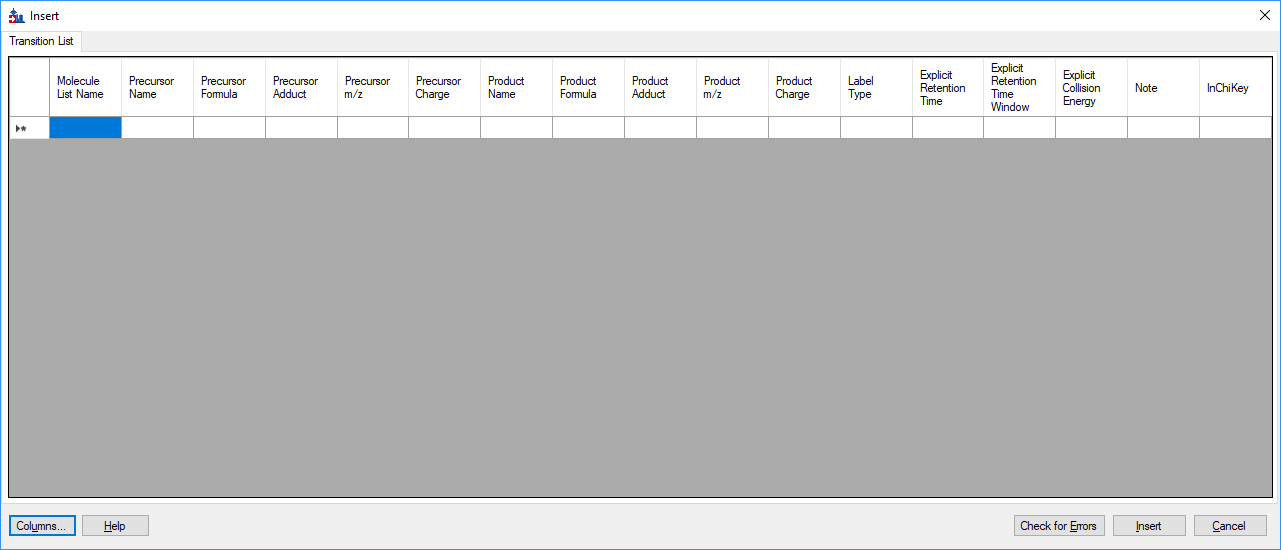
The easiest way to get a small molecule transition list into a Skyline document is to start with an empty document and use the **Edit > Insert > Transition List** menu item.

## Transition list insert

To begin, do the following:

* On the Skyline **Edit** menu, choose **Insert** and click **Transition List**.

Skyline will show the **Insert** form. You may have a different column selection and order from previous uses of Skyline:

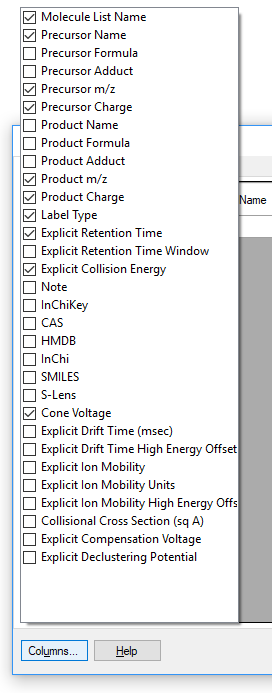


Normally you would copy and paste a transition list from Excel or some other external source, but in this case we have a small enough transition list that we can enter it by hand.

You can see that there are some extra column headers in the **Insert** form, and the column order is not the desired order. Both issues are easy to correct:

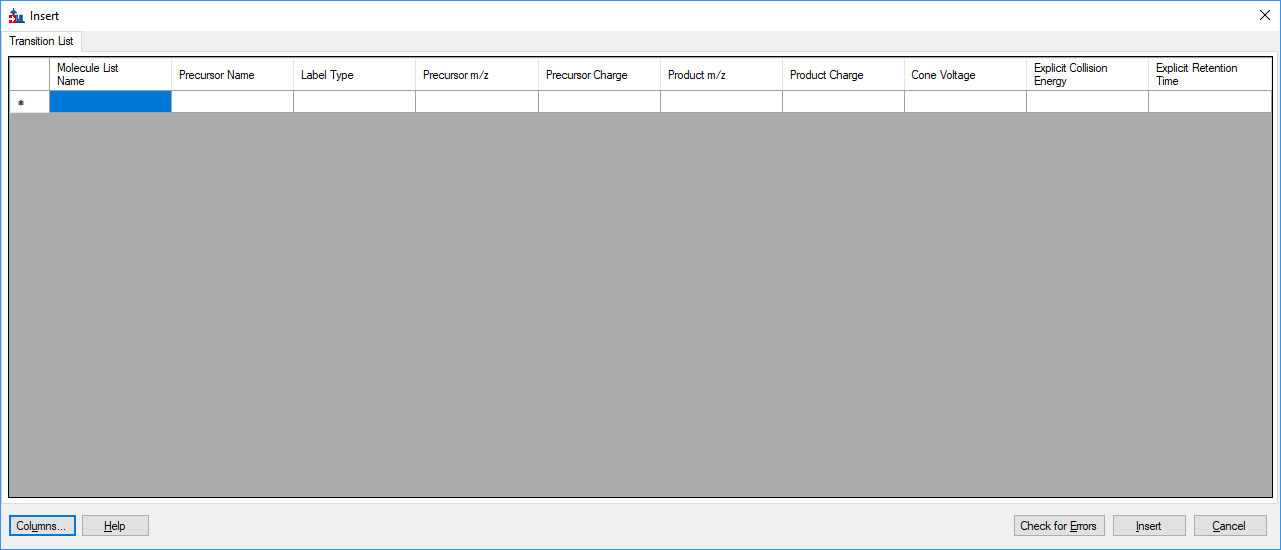
* Click the **Columns** button and uncheck the columns that do not appear in the spreadsheet.

This should result in a column picking menu like the one shown below:



Next do the following to reorder the columns in the **Insert** form:

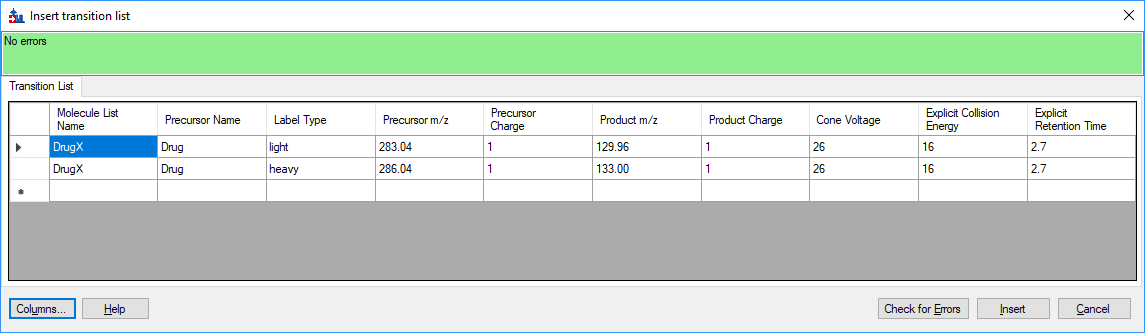
* Click and drag each column header you want to move to the order matching the spreadsheet. Once you have selected and arranged your columns, the insert form should now appear as shown below:



Enter the following values in the **Insert** form (or better yet, copy and paste from this PDF):

DrugX,Drug,light,283.04,1,129.96,1,26,16,2.7  
DrugX,Drug,heavy,286.04,1,133.00,1,26,16,2.7

If you accidentally got the column order wrong, then you will see an error at this point. Otherwise, the **Insert** form should look like this:

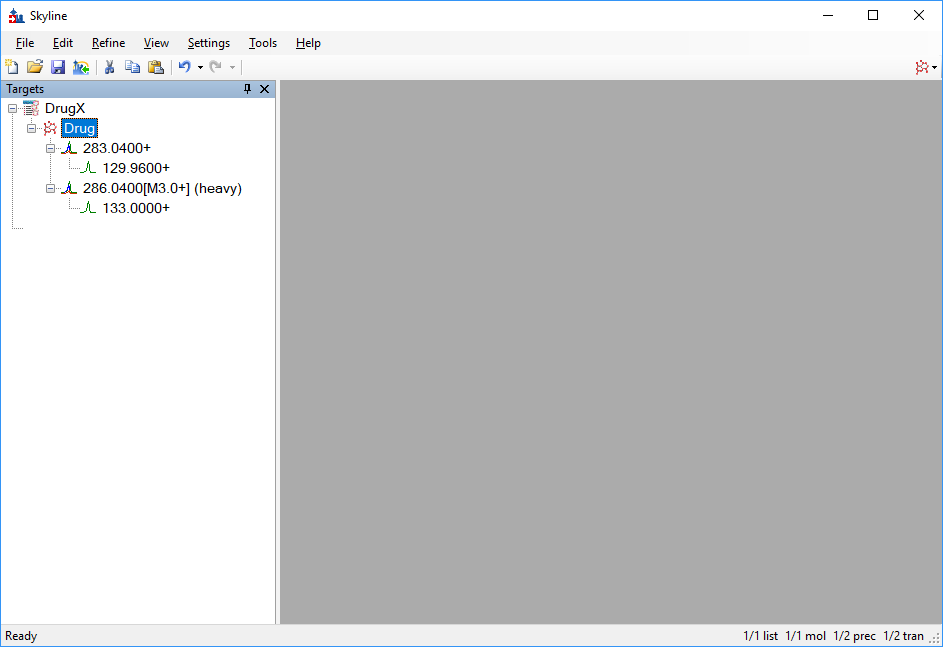


* Click the **Insert** button.

NOTE: In this instance, the targets are described with m/z and charge values. Skyline can accept higher level descriptions including chemical formulas and heavy isotope labels etc. Having the chemical formula is especially useful when working with full scan data as it allows Skyline to calculate isotopic distributions, but for SRM data such as this m/z and charge are perfectly adequate.

To see the newly imported targets in full detail:

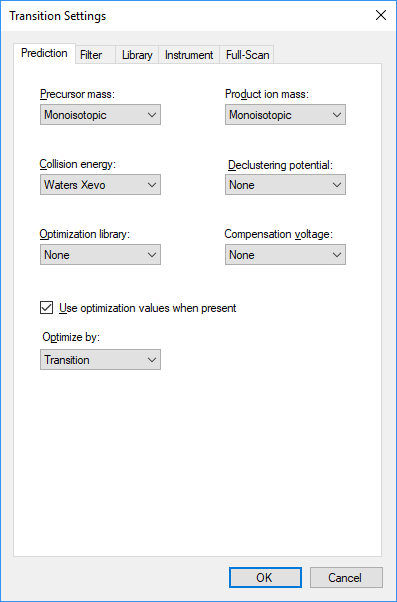
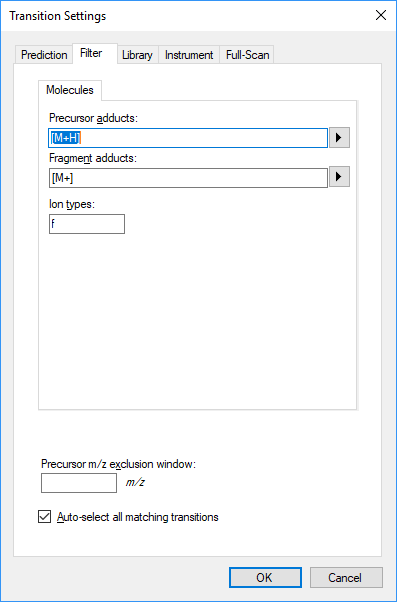
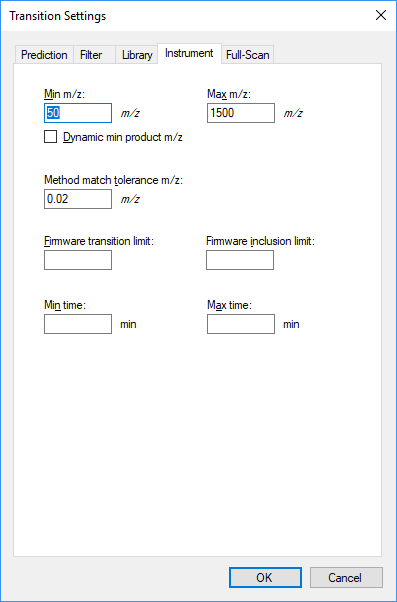
* From the **Edit** menu, click on **Expand All**
* Then click on **Precursors**.

Your Skyline window should now look like this: 

## Transition settings

Next we have to make sure Skyline’s Transition Settings are correctly set for importing the experimental mass spectrometer results. To do this, perform the following steps:

* On the **Settings** menu, click **Transition Settings**.
* Change settings as necessary to match the following:

Note: the “f” setting in Ion Types on the Filter tab means we are only interested in fragment ion transitions. If you wanted to measure precursor ions as well, it would be “f,p”.

In the Instrument tab, make sure that the minimum and maximum *m/z* values make sense for your actual instrument. The purpose of these settings is to alert you when you propose to monitor transitions that your mass spectrometer is not actually able to measure.

Method match tolerance is another important setting in the Instrument tab – it determines how closely a measured *m/z* value must match the transition value to be considered a match.

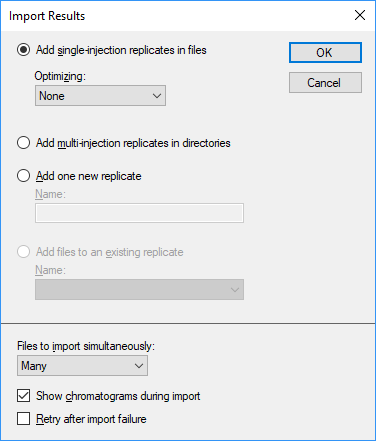
Now we are ready to import the experimental mass spectrometer results.

## Importing mass spectrometer runs

This experiment has 113 mass spec result files associated with it. In cases like this, it can be useful to initially import just a handful of the unknowns along with all of the Calibration Curve and Quality Control (QC) files. This lets us start out with a less complicated document while we verify chromatography and calibration curves etc.

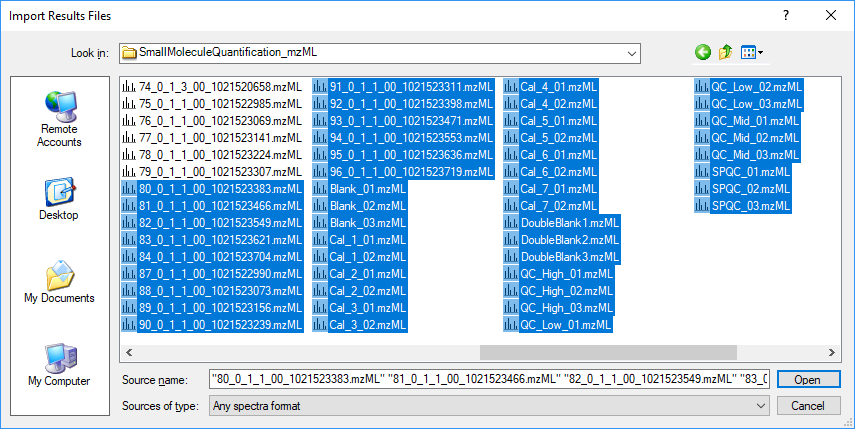
Perform the following steps.

* On the **File** menu, click **Save**. (Ctrl-S)
* Save this document as “SMQuant\_v1.sky” in the tutorial folder you created.
* On the **File** menu, select **Import** and click on **Results**.
* In the **Import Results** form, choose to import single-injection replicates. For best performance, be sure to select “Many” in the “Files to import simultaneously” control at the bottom of the form.



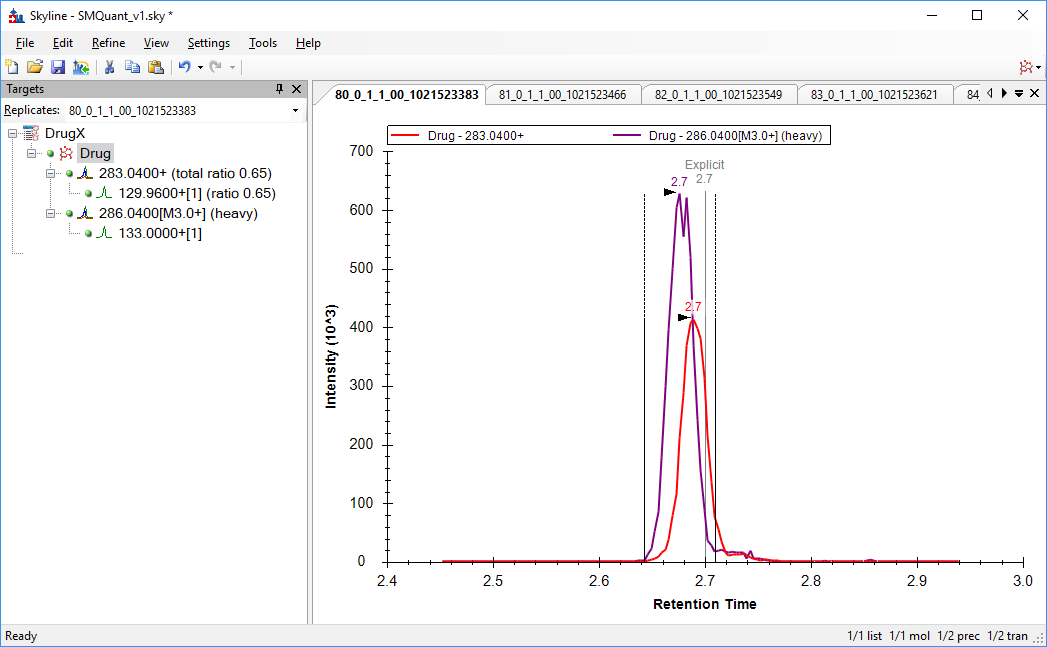
* Now click the **OK** button.
* You should see a file selection dialog with all the raw data folders in the tutorial folder. Select the last 16 unknowns samples and all the QC samples by clicking the “80\_0\_1\_1\_00\_1021523383.raw” file and then holding down the Shift key and clicking the last file in the list.

The **Import Results Files** form should look like:



* Click the **Open** button.

The files should import within 30 seconds or so, leaving your Skyline window looking something like this:



To take advantage of the Skyline summary graphs for viewing individual targets, do the following:

* On the **View** menu, choose **Peak Areas** and click **Replicate Comparison**.
* On the **View** menu, choose **Retention Times** and click **Replicate Comparison**.
* Click and drag these views to the arrows to dock them above the chromatogram graphs.
* Select the first target “Drug” in the **Targets** view.

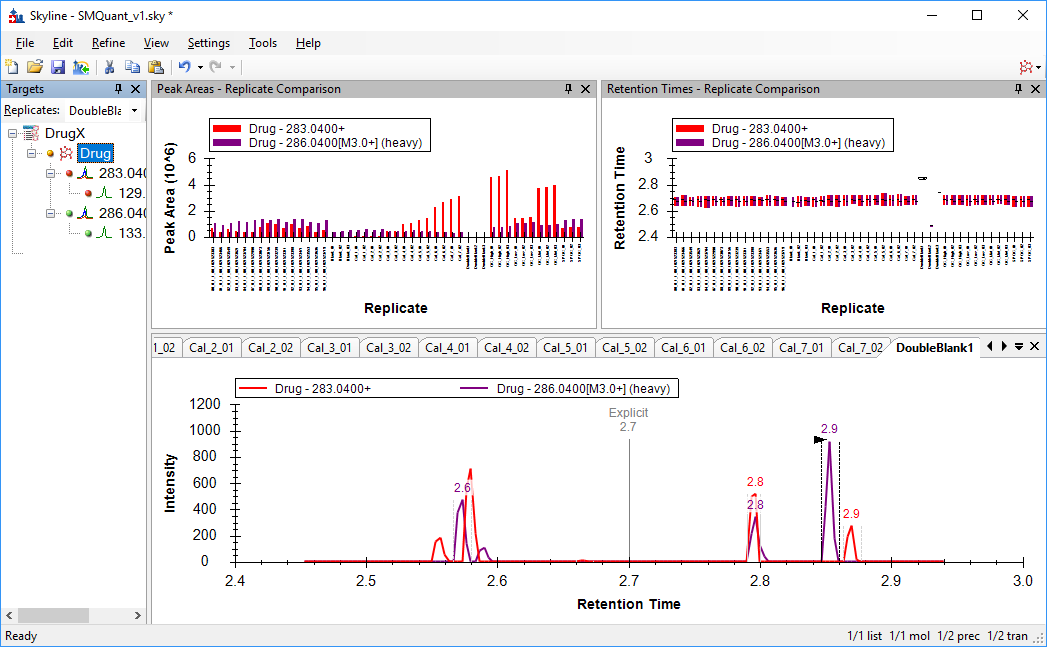
The Skyline window should now look something like this:

# 

# Checking Peak Integration

Looking at the **Retention Times – Replicate Comparison** window we can see by the outliers that in a few replicates Skyline has not chosen peaks at the retention times we might have expected.

Clicking on the first outlier in the **Retention Times – Replicate Comparison** window changes the selected replicate to DoubleBlank1, where we would not actually expect Skyline to find a good peak for either the light or heavy form of the drug.



Clicking on the other two outliers reveals that DoubleBlank2 and DoubleBlank3 are the other replicates with apparently poor chromatography. But, of course, as these are double blanks we don’t expect any actual peaks in these replicates, so we will manually adjust the integration for each of the double blank replicates to center on the low signal area at 2.7 minutes.

# Adjusting Peak Integration

To adjust peak integration, select the DoubleBlank1 replicate then position the mouse below the Retention Time axis (the cursor will change its shape to this:). Click below the Retention Time axis at about 2.65 minutes and drag to about 2.75 minutes. The peak boundaries will change to these new values, and the original boundaries will be marked by a shaded area.

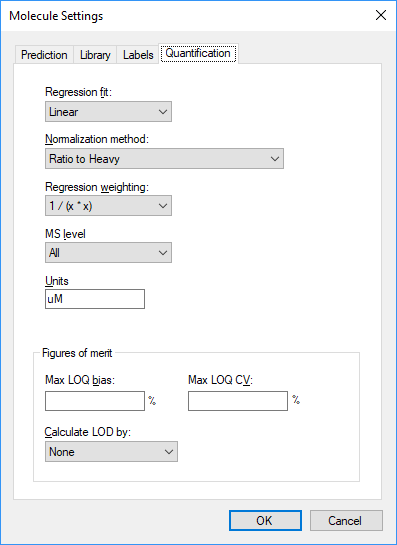
Repeat for the other two DoubleBlank replicates. This should leave Skyline showing something like:

# 

# Preparing for Quantitation

Now we can proceed to quantitation. To do this, perform the following steps:

* On the **Settings** menu, click **Molecule Settings**.
* Select the **Quantitation** Tab, and change settings as necessary to match the following:



We will use a linear regression fit, normalizing to the heavy labeled drug. Skyline offers options for weighting across the curve as a function of x: None, 1/x, and 1/(x\*x). We will use a Regression Weighting of **1/(x\*x)**. The Units setting is for display purposes, and can be set to any value that makes sense for your experiment. The data experiment was calibrated in micromolar, so we set this to **uM**.

* Click **OK**.

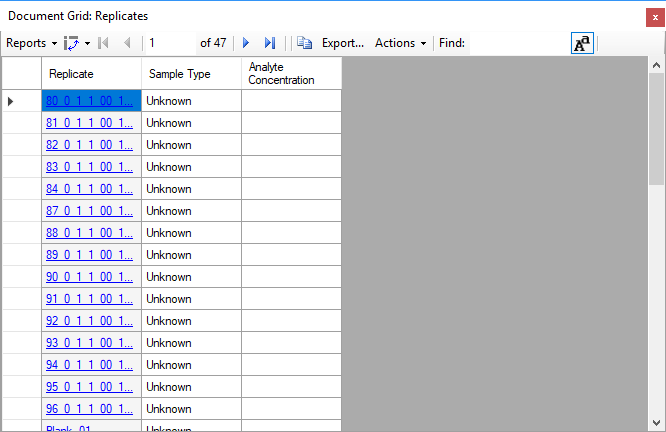
We are still not quite ready to view a calibration curve, however, as we have not yet declared the sample types of the various replicates for Skyline.

## Declaring Sample Types for Calibration Curve Display

We will use the Document Grid to examine and add information about the various replicates. The Document Grid is a highly useful tool in Skyline, providing spreadsheet-like views of many document details, much of which can be edited right there in the grid. In this case, we are interested in details of the various replicates, so:

* On the **View** menu, select **Document Grid**
* Click on the **Reports** control in the upper left corner of the grid, and select **Replicates**

The document grid should look like this:



* Expand the Document Grid if needed so that you can see all the replicates at one.
* If needed, alphabetize the list by clicking on the Replicate column header and choose “Sort Ascending”.

Now we can set the various sample types. The replicates with the unknowns are ready to go by default.

* Click on the Sample Type column for “Blank\_01”. Change the sample type from “Unknown” to “Blank”.
* Now shift+click on the Sample Type column for “Blank\_03” to select all three Blank replicates at once.
* Right-click and select **Fill Down**. Everything in the multiple selection now has the same value as the first item in the selection.

Repeat as needed (or, skip ahead to the table below):

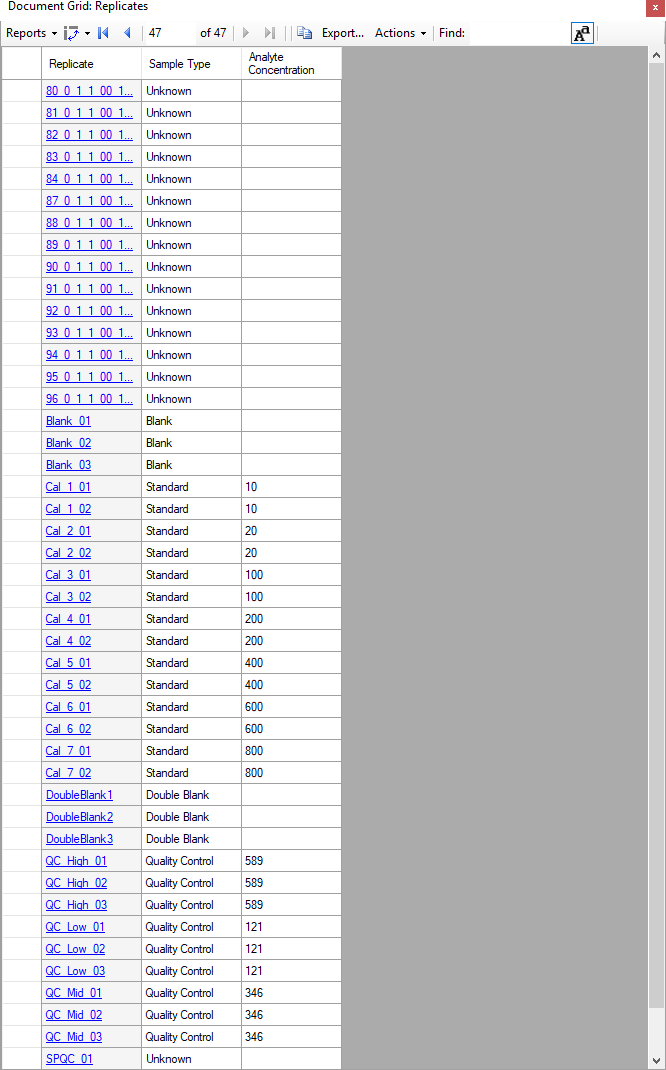
* Set the “Cal\_” replicates as **Standard**
* Set the “DoubleBlank\_” replicates as **Double Blank**
* Set the “QC\_” replicates as **Quality Control**

Recall that the SPCQC\_ replicates are quality control in a different sense (a pooling of all study samples), so we leave those as **Unknown**.

Analyte concentrations can be entered by hand, but it is much easier to copy and paste them into the grid. Locate the **Concentrations.xlsx** file in the directory you created for this tutorial and open it in Excel. It should look something like this:

|  |  |  |
| --- | --- | --- |
| Blank\_01 | Blank |  |
| Blank\_02 | Blank |  |
| Blank\_03 | Blank |  |
| Cal\_1\_01 | Standard | 10 |
| Cal\_1\_02 | Standard | 10 |
| Cal\_2\_01 | Standard | 20 |
| Cal\_2\_02 | Standard | 20 |
| Cal\_3\_01 | Standard | 100 |
| Cal\_3\_02 | Standard | 100 |
| Cal\_4\_01 | Standard | 200 |
| Cal\_4\_02 | Standard | 200 |
| Cal\_5\_01 | Standard | 400 |
| Cal\_5\_02 | Standard | 400 |
| Cal\_6\_01 | Standard | 600 |
| Cal\_6\_02 | Standard | 600 |
| Cal\_7\_01 | Standard | 800 |
| Cal\_7\_02 | Standard | 800 |
| DoubleBlank1 | Double Blank |  |
| DoubleBlank2 | Double Blank |  |
| DoubleBlank3 | Double Blank |  |
| QC\_High\_01 | Quality Control | 589 |
| QC\_High\_02 | Quality Control | 589 |
| QC\_High\_03 | Quality Control | 589 |
| QC\_Low\_01 | Quality Control | 121 |
| QC\_Low\_02 | Quality Control | 121 |
| QC\_Low\_03 | Quality Control | 121 |
| QC\_Mid\_01 | Quality Control | 346 |
| QC\_Mid\_02 | Quality Control | 346 |
| QC\_Mid\_03 | Quality Control | 346 |
| SPQC\_01 | Unknown |  |
| SPQC\_02 | Unknown |  |
| SPQC\_03 | Unknown |  |

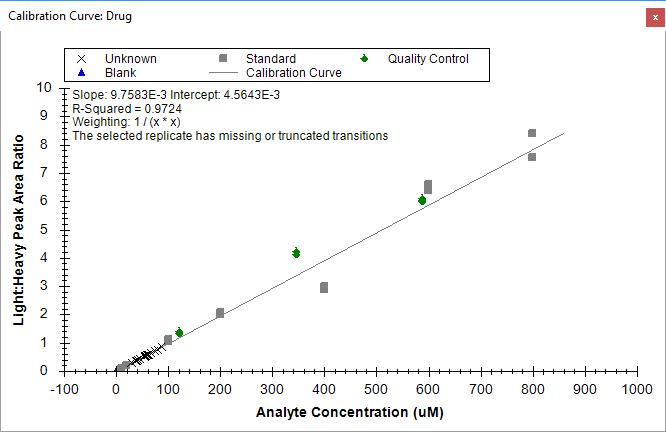
* Select the contents of the table (cells A1 to C32)
* Make sure the order matches the Document Grid
* Click on the “Blank\_01” cell and paste

The document grid should look like this when you are done: 

# Inspecting the Calibration Curve

Now we are ready to examine the calibration curve.

* Close the **Document Grid** view.
* From the **View** menu, select **Calibration Curve**. You should see this:

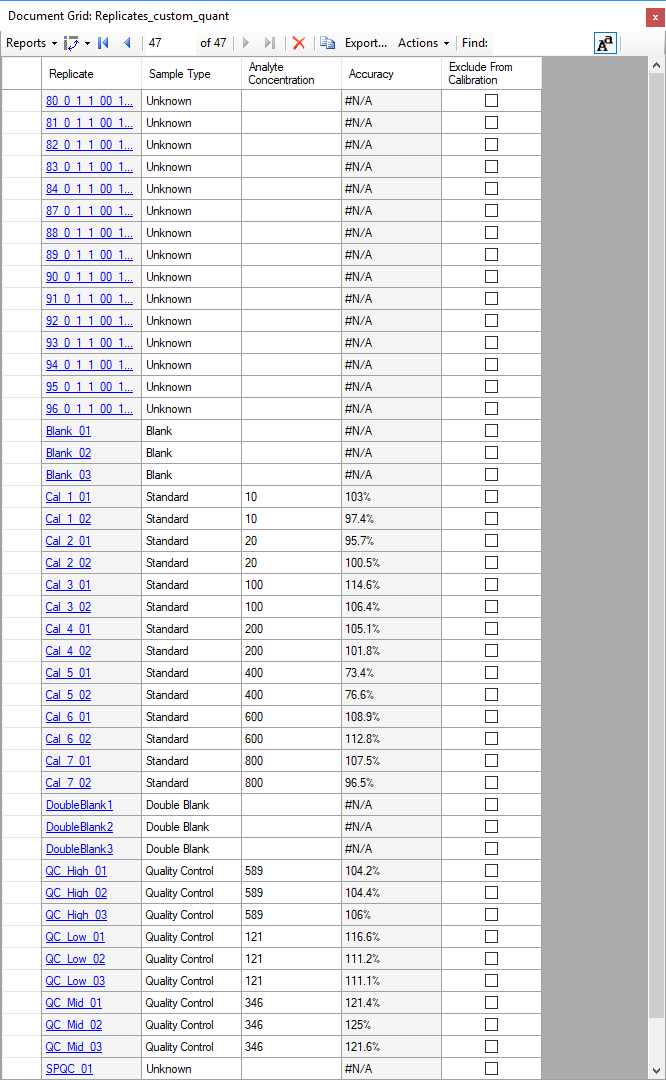


The note about the selected replicate missing transitions is expected, as the currently selected replicate is a double blank.

Looking at the graph, we note that the Unknowns loaded so far have relatively low dynamic range.

We also note that some of the calibration samples look like they might be outliers. We can use the Document Grid to get a qualitative sense of that, and to exclude any samples that are not suitable:

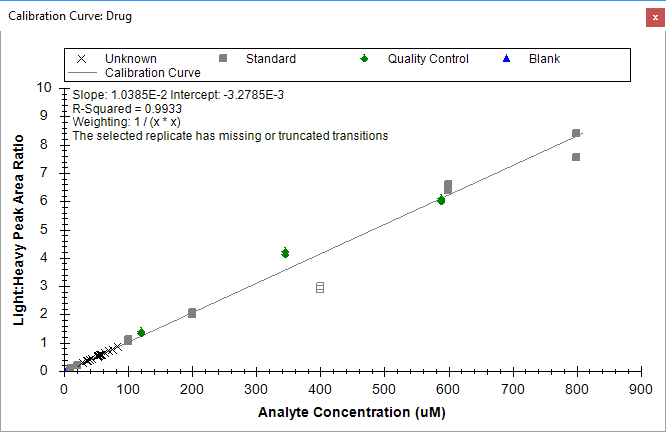
* On the **View** menu, select **Document Grid**
* Click on the **Reports** control in the upper left corner of the grid, and select **Replicates**
* Click again on the **Reports** control in the upper left corner of the grid, and select **Customize Report**
* Click on the search button  and type **accuracy** into the **Find What** field
* Click **Find Next**
* In the **Customize Report** form, **Accuracy** should be highlighted under Quantification
* Click on the checkbox to enable **Accuracy**
* Just above Quantification, click on the checkbox to enable **Exclude from Calibration**
* In the **View Name** field at the top of the **Customize View** form, type **Replicates\_custom\_quant**
* Click **OK**

The Document Grid view should now look like this: 

The FDA Guidance on which this assay was based states that calibration points should have bias < 15% (accuracy between 85% and 115%) between the known concentration and the backcalculated concentration from the calibration curve, and from the Accuracy column we can see that Cal\_5 does not meet that test. These replicates can be removed from consideration using the checkbox in the **Exclude from Calibration** column in the Document Grid, or by right-clicking on the outlier in the Calibration Curve view and selecting **Exclude from Calibration.**

* In the Document Grid, click on the checkbox in the **Exclude from Calibration** column for the **Cal5\_01** replicate, then press the **down-arrow key**
* Repeat for **Cal5\_02**

The Calibration Curve should now look like the image below. Note that the R-squared value improves from 0.97 to >0.99 by excluding the outlier Cal\_5.

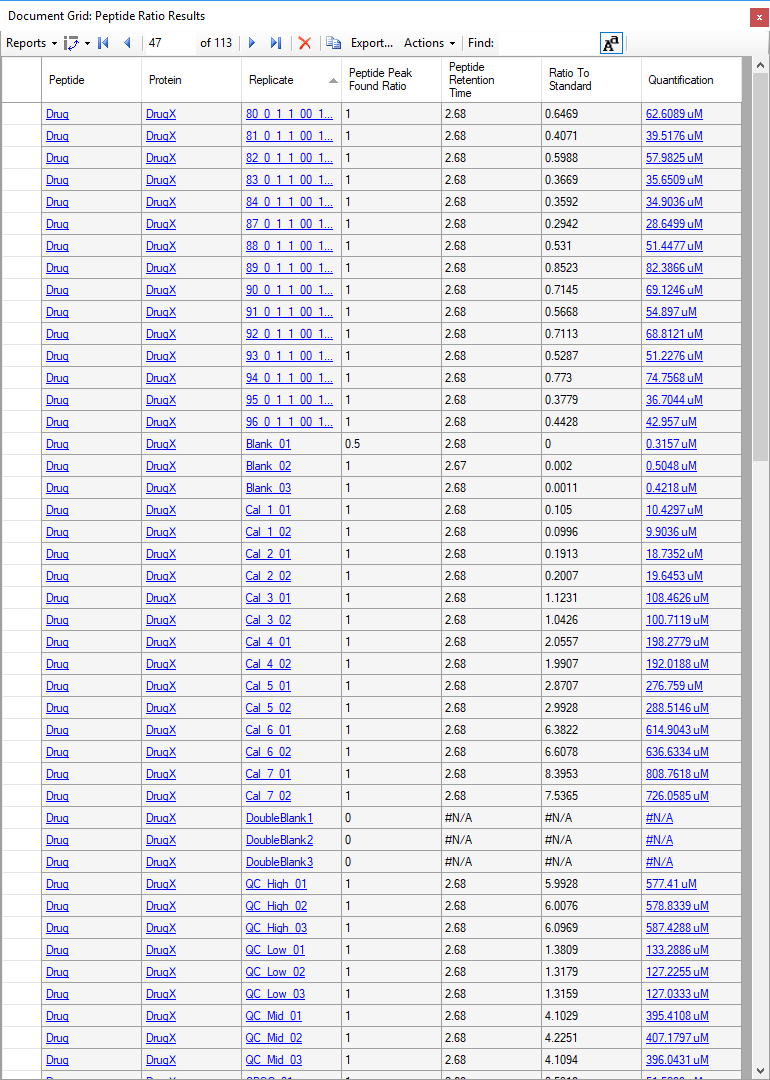


Now we can import the rest of our unknowns:

* On the **File** menu, select **Import** and click on **Results**.
* In the **Import Results** form, choose to import single-injection replicates. For best performance, be sure to select “Many” in the “Files to import simultaneously” control at the bottom of the form.
* Now click the **OK** button.
* You should see a file selection dialog with all the raw data folders in the tutorial folder. Select the as-yet-unloaded unknowns and click **OK**.

A convenient way to view quantitation data is to once again use the Document Grid, this time with the **Peptide Ratio Results** view.

* In the **Document Grid** window, click on the **View** control and select **Peptide Ratio Results**.
* Click on the **Replicate** column header and select **Sort Ascending**

The Document Grid should look something like this: 

After removal of the two Cal\_5 datapoints, further exploration of the data reveals that one of the Cal\_7 points has accuracy <85% and should likely be removed. This will have little effect on the measurement of the samples, since there are no samples above the level of Cal\_6, and only four samples have levels between Cal 4 and Cal 6.

To enable easier visualization of the dynamic range of the samples along the calibration curve:

* Right-Click in the calibration curve window and select **Log Plot**.

The calibration curve should look like this:

****

This presents a view which allows the user to easily observe that the samples mostly fall between Cal\_2 (20 uM) and Cal\_3 (100 uM), and well within the linear calibration range of the assay. The Quality Control samples (known unknowns, green diamonds above) all have accuracy measured between 85 and 115%, which meets the FDA Guidance criteria.

From here, the next steps would be to export the data for external statistical processing, or establish biological grouping within this document and utilize some of the statistical analysis tools or plugins within skyline. These options are covered in other tutorials.

# Conclusion

In this tutorial, you have learned how to create a Skyline document that targets small molecules specified as precursor ion chemical formulas and adducts, and product ion m/z values. You imported a multi-replicate data set collected using LC-MS/MS on a triple quadrupole, and saw how many existing Skyline features created initially for targeted proteomics use can now be applied to small molecule data. Non-proteomic molecule support is still a relatively new feature area for Skyline. As such, you can expect it to continue improving rapidly.

# Bibliography