

Statistical Process Control in Proteomics SProCoP

This tutorial will guide you through the installation of SProCoP and using it to perform statistical analysis on a sample Skyline file.

Getting Started

To start this tutorial, download **SProCoPTutorial.zip** from this page:

<https://skyline.gs.washington.edu/labkey/skyts/home/software/Skyline/tools/toolDetails.view?name=SProCoP>

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

C:\Users\brendanx\Documents

This will create a new folder:

C:\Users\brendanx\Documents\SProCoPTutorial

The zip file contains the following files:

`Skyline_SProCoP.sky`: A Skyline document containing sample data for SProCoP.

`Skyline_SProCoP.skyd`: Contains extracted chromatogram data for the Skyline document.

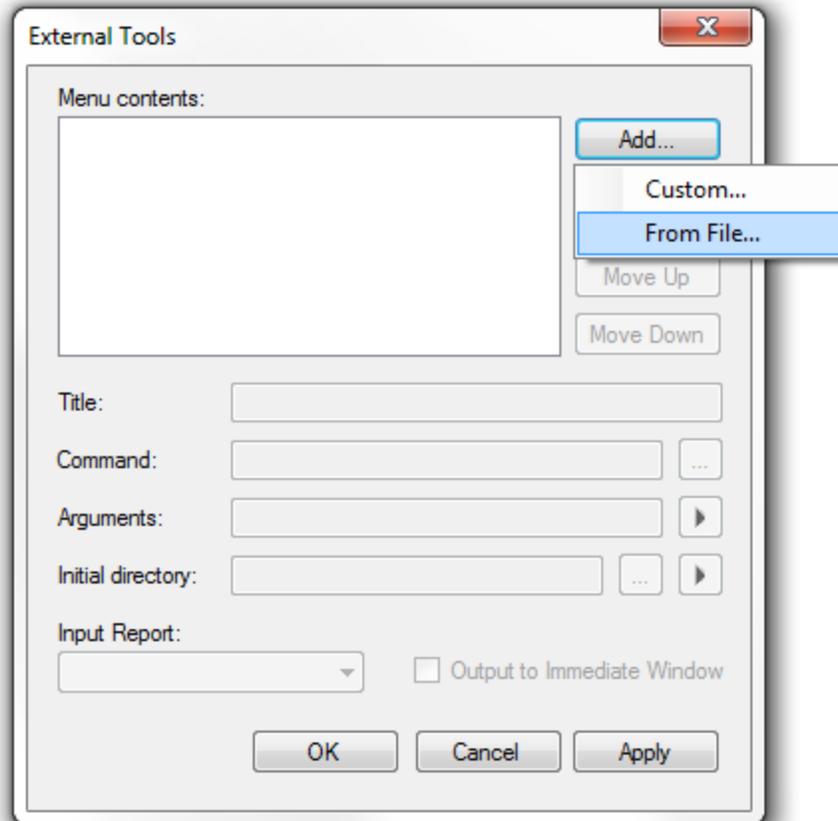
Open the provided file `Skyline_SProCoP.sky` in Skyline. The peptides measured in the document along with the data acquisition parameters are described in **Figure 1** at the end of this tutorial. The data in the document are Quality controls that were monitored within the context of a larger study assessing the time course tryptic digestion of proteins from blood. All experimental data was collected on a Q-Exactive tandem mass spectrometer (ThermoFisher, Bremen, Germany) coupled to a nano Acquity UPLC (Waters). The bath gas in the C trap of the Q Exactive was low which led to decreasing ion abundance across the runs. The gas was then replenished and the targeted peak areas significantly increased.

Installing SProCoP as an External Tool

The Skyline external tools framework supports the direct integration of statistical analysis tools that process the data in a Skyline document.

To add SProCoP as an external tool, perform the following steps:

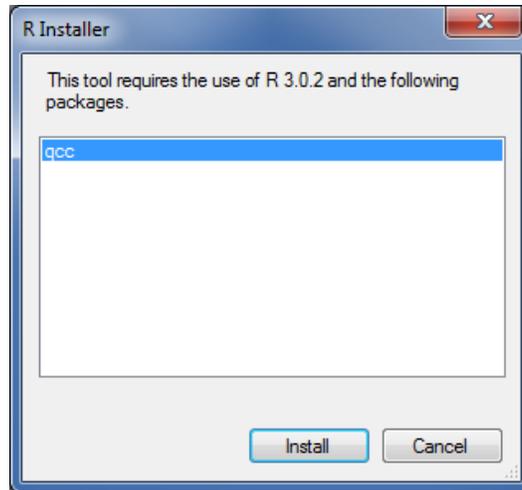
- On the **Tools** menu, click **External Tools** to bring up the **External Tools** form.



Note: Your External Tools menu may already contain some Tools

- Click the **Add** button, and select **From File** from the context menu.
- Navigate to the location of the `SProCoP-1_0.zip` file in the folder you created and select it.
- Click the **Open** button.

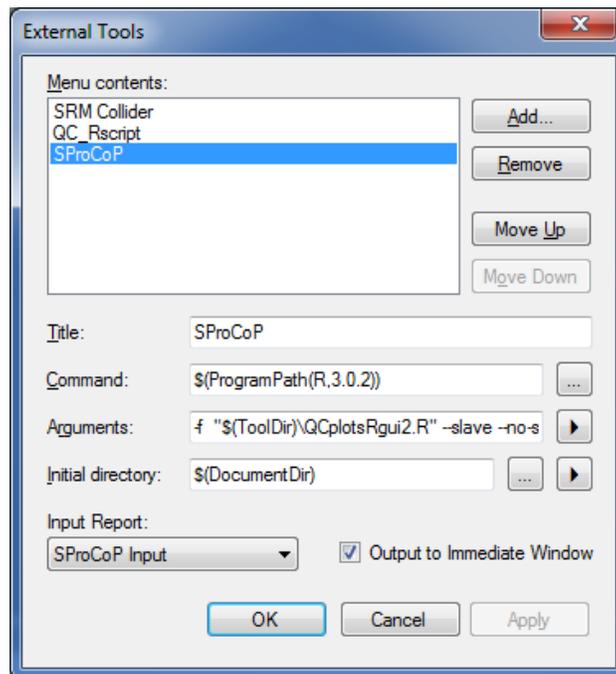
Skyline will now guide you through the process of installing SProCoP as an external tool. This may require you to install the R statistical programming environment, as well as a number of R packages which are used by SProCoP to perform its statistical analysis. For example, if you do not have the appropriate version of R installed, the following form will be displayed:



To complete the installation, performing the following steps:

- On the **R Installer** form, click **Install**.
- If you do not have the specified version of R installed on your machine, wait for R to download. When the download is completed, you will be asked to install R onto your computer. Complete the R installation process by working through the R installer wizard.
- Once the R installation is completed, Skyline will download and install the necessary R packages for SProCoP. When prompted to run the “SkylineProcessRunner”, click the **Yes** button to install packages.

Once the installation is complete, the SProCoP tool will appear in the **External Tools** form:

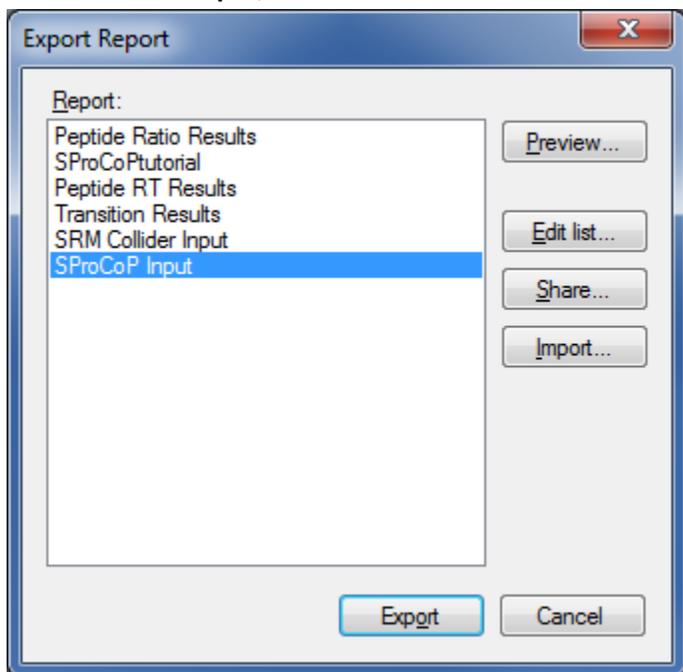


- Click the **OK** button on the **External Tools** form.

Checking the Skyline Report

SProCoP will automatically generate a **SProCoP Input** report based on the .sky file. Although this step is not necessary we recommend that you check the Skyline Report to ensure that the report is rendered correctly. The demo `Skyline_SProCoP.sky` file report should look like the figure on the next page.

- On the **File** menu, click **Export > Report**.
- Select **SProCoP Input**, and click the **Preview** button on the right (see figure below)



Once **Preview** is selected, the **Preview Report** page will appear:

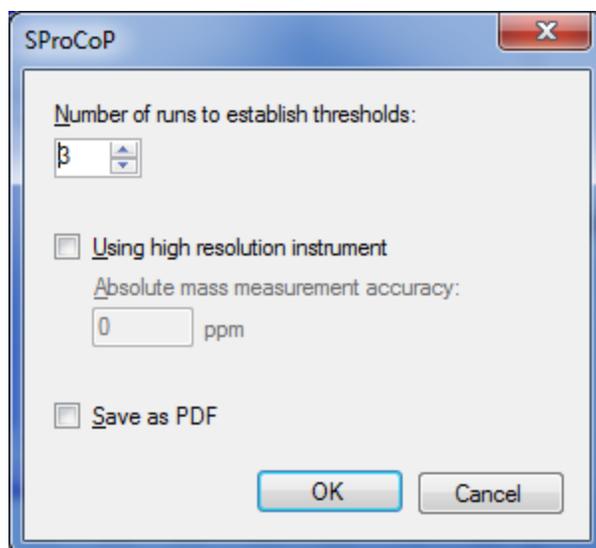
ReplicateName	PeptideModifiedSeq	TotalArea	BestRetentionTime	MaxFwhm	MinStartTime	MaxEndTime	AverageMassErrorPPM
Q_Exactive_08_...	VYVEELKPTPE...	110572224512	29.55	0.1	29.41	29.76	0.8
Q_Exactive_08_...	VYVEELKPTPE...	160706428928	29.5	0.14	29.41	29.78	1.1
Q_Exactive_08_...	VYVEELKPTPE...	100025868288	46.42	0.19	46.29	46.93	1
Q_Exactive_08_...	VYVEELKPTPE...	173198966784	29.39	0.18	29.31	29.68	1.8
Q_Exactive_08_...	VYVEELKPTPE...	153556402176	29.36	0.16	29.26	29.66	1
Q_Exactive_08_...	VYVEELKPTPE...	176071475200	29.32	0.22	29.24	29.63	1.2
Q_Exactive_08_...	VYVEELKPTPE...	148374831104	29.39	0.15	29.26	29.65	1.3
Q_Exactive_08_...	VYVEELKPTPE...	137785720832	29.42	0.1	29.34	29.73	0.8
Q_Exactive_08_...	VYVEELKPTPE...	166037176320	29.49	0.17	29.31	29.68	1.6
Q_Exactive_08_...	VYVEELKPTPE...	145370038272	29.42	0.17	29.29	29.73	0.8
Q_Exactive_08_...	VYVEELKPTPE...	142750416896	29.36	0.17	29.28	29.68	0.9
Q_Exactive_08_...	VYVEELKPTPE...	155523088384	29.44	0.13	29.31	29.68	0.5
Q_Exactive_08_...	VYVEELKPTPE...	115039272960	29.41	0.16	29.32	29.71	0.9
Q_Exactive_08_...	VYVEELKPTPE...	112991633408	29.41	0.16	29.31	29.7	1.6
Q_Exactive_08_...	VYVEELKPTPE...	107816869888	29.47	0.18	29.34	29.7	1.2
Q_Exactive_08_...	VYVEELKPTPE...	128046080000	29.36	0.11	29.28	29.7	1.1
Q_Exactive_08_...	VYVEELKPTPE...	130433499136	29.37	0.1	29.26	29.65	1.2
Q_Exactive_08_...	VYVEELKPTPE...	116482187264	29.45	0.15	29.29	29.66	2.2
Q_Exactive_08_...	VYVEELKPTPE...	127296757760	29.37	0.13	29.28	29.6	1.9
Q_Exactive_08_...	VYVEELKPTPE...	106591420416	29.44	0.11	29.34	29.7	2.6
Q_Exactive_08_...	VYVEELKPTPE...	96447987712	29.39	0.11	29.29	29.63	2.5
Q_Exactive_08_...	VYVEELKPTPE...	91348934656	29.44	0.13	29.29	29.65	2.4
Q_Exactive_08_...	VYVEELKPTPE...	82229624832	29.34	0.17	29.26	29.65	3.4
Q_Exactive_08_...	VYVFFI KPTPF	72177934336	29.39	0.07	29.29	29.62	3.5

Using SProCoP

To run SProCoP, perform the following steps:

- On the **Tools** menu, click **SProCoP**.

Skyline will begin exporting the SProCoP input report. This may take a few minutes. You can track the progress of the report export with the green progress bar in the bottom of the Skyline Window. Once the report is exported, Skyline will display the **SProCoP** form:



The “Using high resolution instrument” checkbox should be checked if your data was acquired on a high resolution instrument. The reports generated by SProCoP can be saved to a PDF file by checking the **Save as PDF** checkbox. High quality charts are saved to a single PDF file with one chart per page.

- For this tutorial you will leave the options in their default settings.
- Click on the **OK** button.

SProCoP will generate two windows (pictured below). The larger window represents a control chart matrix for all the targeted peptides across 4 different metrics: Peak areas, RT, FWHM and Peak asymmetry. The first three are all reported by Skyline directly from the raw data. Peak asymmetry is calculated by $P_{asym} = (a+b)/2a$ and determined from the min start, RT at apex, and max end. Parameters a and b are the width at base from the start of the peak to the apex and from the apex to the end of the peak, respectively. For calculation of MMA, a MMA value is calculated and stored for every time-intensity pair in the chromatograms it extracts from high resolution data. The mass error at each chromatogram point is calculated by taking a weighted average of the m/z-intensity pairs found within the m/z extraction window, and subtracting the expected m/z from that average. When Skyline performs linear interpolation of extracted chromatogram points onto a scale with a consistent interval, it applies the same interpolation to the mass error values. Finally, to calculate the mass error for any integrated peak, Skyline calculates a weighted average of the mass errors for chromatogram points within the integration boundaries, weighted by their background-subtracted intensities. For more information on the control chart see **Figure 2** at the end of the tutorial.

The smaller window represents a Pareto Chart which is a combination of a bar and line graph. The bar graph represents the frequency of times a peptide, across all the runs, is outside the $\pm 3sd$ threshold determined empirically for the 4 metrics. The line graph is a cumulative percentage chart and represents the proportion of total “nonconformers” (i.e., outside $\pm 3sd$) that a particular metric constitutes.

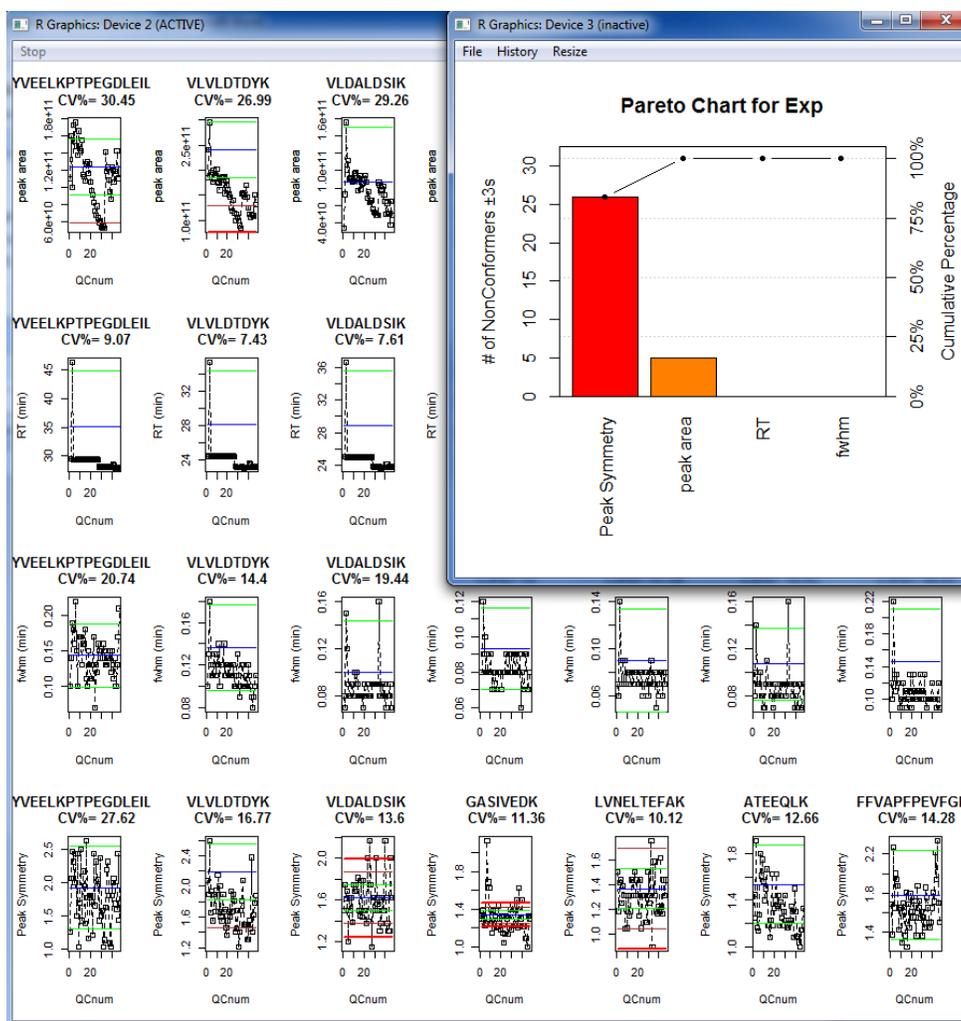


Figure 1

Parameter	Setting
Source	Nano-ESI
Polarity	Positive
Capillary temperature (°C)	325°
S-Lens RF Voltage	50%
Source Voltage (kV)	2.2

Full-scan MS parameter	Setting
Mass range (<i>m/z</i>)	400-2000
Resolution (FWHM at <i>m/z</i> 200)	70,000
AGC Target	1×10^6
Spectrum data type	Profile
Maximum injection time (msec)	10

Targeted MS/MS parameters	Setting
Isolation window (<i>m/z</i>)	2
Resolution (FWHM at <i>m/z</i> 200)	17,500
AGC Target	1 x 10 ⁶
Spectrum data type	Profile
Collision energy (NCE)	25
Default charge state	+2
Maximum injection time (msec)	55

Peptide from bovine QC mix	Mono-isotopic mass for 2+ precursor ion	Retention Time
VLVLDTDYK	533.29500	23.2
VLDALDSIK	487.28190	23.8
ATEEQLK	409.71630	14.5
DDGSWEVIEGYR	713.31770	26.0
YSTDVSVDEVK	621.29840	20.0
FFVAPFPEVFGK	692.86860	30.8
VYVEELKPTPEGDLEILLQK	1157.13300	28.0
LVNELTEFAK	582.31900	24.6

Figure 2

A detailed description of the control chart used in the studies. The blue line represents the empirically defined mean of the process and the control lines are 1s, 2s, and 3s from the mean. The control lines are determined empirically from the user specified number of QC standards. The CV% represents the coefficient of variation across all runs. It's the standard deviation of all QC runs divided by the mean of all QC runs. In this particular example 46 standards were used to calculate the CV shown.

