

# MS & Proteomics Resource

Yale School of Medicine  
Keck Biotechnology Resource Laboratory



## Application Note 2: Interpretation of SILAC (Stable Isotope Labeling by Amino Acids in Cell Culture) Data

SILAC quantitation is performed using the Matrix Science Mascot Distiller Quantitation tool box (<http://www.matrix-science.com/home.html>). Quantitation is based on the relative intensities of the extracted ion chromatograms (XIC's) of the peptide precursors. There are many criteria involved in correct quantitation. Hence, our typical analysis is quite stringent. Results with quantitation are reported in the Yale Protein Expression Database (YPED). In addition, excel tables in several different formats are provided:

- a list of the proteins identified with determined either H/L or L/H ratios ( or if a medium tag is used, M/L as well).
- A list of peptides identified with the corresponding ratios (note that this only contains heavy peptides)
- A table of the Mascot matches which will include both heavy and light peptides
- And a separate table listing of the proteins and the peptides

Each of these formats contain similar information but you may find one format more amenable to your use.

### Table of Proteins

Mascot uses a Protein Family Summary which groups proteins into families based on shared peptides. A hierarchical clustering is then used to show which proteins are closely related and which are distant. In this excel table, the protHit (or protein hit) is listed in column A with 1 being the most significant or highest Mascot scoring match. If a family member was also identified, then column B will have the number of the family member with 1 being the most closely related. Column C contains the identified protein accession number, Column D the Mascot score, and Column E the intact protein mass. Column F shows the indicated SILAC ratio (e.g. L/H or H/L) and Column H the number of peptides that were used for the ratio. The higher the number of peptides used, the better the ratio. Column G is the geometric standard deviation which means that the confidence interval was obtained by dividing and multiplying the value by the standard deviation and this will never be less than 1. Column J gives the L/H ratios with columns K, L and M, the same as G, H and I for the H/L ratios. An "X" in the I Column (**BOLD**) indicates that this match and ratio is most significant, and that it is significantly different from unity using the comparison test. Matches/ Ratios that are not bold did not pass this test. Hence, the **BOLD** proteins/peptides are the best.

The **Protein Excel Table** is very similar but includes more information on the matched protein's accession number.

### Table of peptides

The Table of the peptides contains all of the same information as in the Table of Proteins but also shows for each peptide if it was included in the ratio determination (Column Q for H/L and column T for L/H), the ion intensity (Column Y) and any peptide modifications (but not the heavy SILAC tags- Column Z). There are 3

thresholds that Distiller uses to determine whether or not to accept a ratio. These are fraction (shown in column W), correlation (Column X), and standard error (Column V).

- Fraction is the percentage of the peak area that is accounted for by the expected precursor peaks. If this is less than 25%, the ratio is rejected (no ratio determined)
- Correlation coefficient tests the shape of the peak distribution. This is to enable eliminating cases where there are interfering peaks. Again, if this does not pass, the ratio is rejected (no ratio determined)
- Standard error for the calculated ratio. The ratio comes from making a least squares fit to the component intensities from the XIC scans. The fitted line is the best estimate of the ratio and is a good measure of the reliability. If the Std err is too high, no ratio will be calculated
- In order to eliminate false positives, ratios may also not be given for the following reasons
  - the initial run will only give a ratio for identifications with 2 or more peptides. Identifications based on only 1 peptide will not be quantitated
  - The Std Err used for initial passes is a high stringency of 0.1. If no ratio is determined, it is possible to re-process the data using 0.5 but the accuracy of the ratios will be greatly diminished

**NOTE** that you do not have to have database matches for all SILAC components in a ratio. As long as there is a match for one component, the “partner” can be inferred from its mass and co-elution.

The **Peptides Excel table** contains the same/similar information but breaks it out by protein identified and then peptides. The D column or Incl. shows what peptides were used (by an “X”) in the ratio calculation. If there is an “O” there, this means it is an outlier and was not included in the protein ratio determination. You can also see a ratio for proteins that were not given a ratio in the protein list. This is because they only have 1 peptide identified. You can see the ratio for this single peptide in this table.

### **Table of Mascot Matches**

This table contains similar information as found in both the Protein and Peptide tables but with the following added information

- Column O is the “z” value which is the charge state of the identified peptide
- Column AA or Query is a Distiller Query number
- Mr(calc) is the calculated molecular weight of the matched peptide and observed is the mass to charge ratio (m/z) as seen in the mass spectrometer
- Score is the peptide (not the protein) score
- And component shows you what peptide was identified as light and what as heavy

### **SILAC Phosphoproteome Analysis:**

- For SILAC phosphoproteome analysis, the digested samples are subjected to a Titanium dioxide enrichment (TiO<sub>2</sub>) step. This will enrich for phosphopeptides as well as some acidic peptides and those containing oxidized methionine
- With SILAC phosphoproteome analysis, you have a protein level change (which has to be determined from the flow through SILAC ratios – not the TiO<sub>2</sub> enriched fraction) and then the enriched fractions will give you the phosphopeptide change.
- **FOR THIS DATA SET** the flow through or FT and enriched or EN runs were searched as MERGED. Hence, to get at the phosphopeptide expression changes, you need to sort the Peptide table and sort by the modification to locate phosphopeptides.

Information on the SILAC Quantitation algorithm is based on Matrix Science’s website and Distiller help menus  
<http://www.matrix-science.com/home.html>