Skyline 现有和定量实验

本教程介绍了如何处理并非由 Skyline 创建的离子对列表采集的选择反应监测（简称 SRM；亦称为多反应监测，简称 MRM）数据。本教程还介绍了如何使用 Skyline 处理利用匹配同位素标记的内部标准肽进行的定量实验。

在本教程中，您将处理采用工具MRMer1 公布的数据，还可以处理由 Addona 及其合作研究者2开展跨实验室研究（以下简称“研究 7”）所获取的数据。研究 7 是美国国家癌症研究所 (NCI) 支持的临床蛋白质组学技术评估癌症评估 (CPTAC) 的一部分。研究 7 随即优先采用 Skyline 作为 CPTAC 验证工作组选择大型实验室间研究的工具。因此，其积累的数据文件还提供了完善的测试包，以确保Skyline 能支持验证工作组正进行的实验类型。

然而，即使您缺少 Skyline 之外创建的离子对列表，您仍然可以了解 Skyline 提供的用于处理 LC-MRM 采集的数据的特点，即利用稳定同位素标记的肽作为内标物进行定量分析所获取的数据。如果您在探索如何有效地进行色谱峰鉴别和定量测定，匹配参照多肽和 Skyline 对这些肽的支持，并可以在帮助您实现目标方面发挥至关重要的作用。

# 入门指南

如要开始本教程，请下载下列 ZIP 文件：

<https://skyline.gs.washington.edu/tutorials/ExistingQuant.zip>

将其中的文件解压到您的电脑文件夹，如：

C:\Users\brendanx\Documents

这将创建一个新文件夹：

C:\Users\brendanx\Documents\ExistingQuant

现在启动 Skyline，您将看到一个新的空白文档。

将现有离子对列表导入这个新文档之前，您应当尽可能多地向 Skyline 提供关于设计离子对列表和收集数据的实验背景方面的有用信息。要做到这一点，您可以在尝试将离子对列表插入到该文档之前，调整文档的设定值。

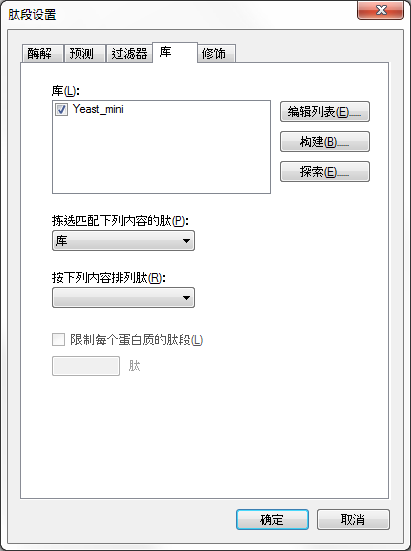
# 准备接受离子对列表的文档

本教程开始时首先要看看一款名为 MRMer1（其发音如英文单词“murmur”）的 MRM 分析软件工具提供的数据集。在查看和整合 MRM 色谱图方面，MRMer 是 Skyline 的早期前身，数据最初于 2008 年从其网页中下载(http://proteomics.fhcrc.org/CPL/MRMer.html)。在 MRMer 数据集中，所有肽均源自于酵母菌，并且拥有在由国家标准与技术研究所 (NIST) 提供的谱图库中的 MS/MS 谱图。这就意味着通过指定谱图库和背景蛋白质组，可以非常轻易地向 Skyline 提供实验中所监测的肽的相关有用信息。在“[靶向方法编辑](https://skyline.washington.edu/labkey/wiki/home/software/Skyline/page.view?name=tutorial_method_edit)”教程中详细介绍了创建您自己的谱图库和背景蛋白质组文件的方法。在本教程中您将使用现有文件，这些文件已经被压缩至完成本教程所需的最小信息量，以便将下载的 ZIP 文件大小控制在合理的范围之内。

设置 MRMer 文档的谱图库请按以下步骤操作：

* 在“设置”菜单上，单击“肽段设置”。
* 单击“库”标签。
* 单击“编辑列表”按钮。
* 单击“编辑库”表单中的“添加”按钮。
* 在“编辑库”表单的“名称”字段内输入“Yeast\_mini”。
* 单击“浏览”按钮。
* 导航至先前创建的 ExistingQuant 文件夹下的 MRMer 子文件夹。
* 选择仅包含本研究中使用的酵母肽谱图的“Yeast\_MRMer\_mini.blib”文件。
* 单击“打开”按钮。
* 单击“编辑库”表单中的“确定”按钮。
* 单击“编辑库”表单中的“确定”按钮。
* 选中“库”列表中新的“Yeast\_mini”库。

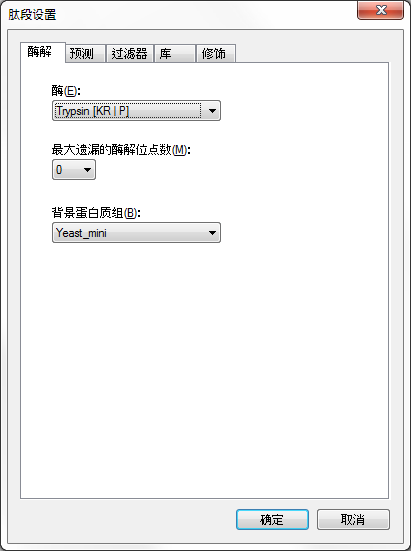
“肽段设置”表单将显示如下：



现在设置 MRMer 文档的背景蛋白质组，请按以下步骤操作：

* 在“肽段设置”表单中，单击“酶解”标签。
* 在“背景蛋白质组”下拉列表中，选择“<添加……>”。
* 在“编辑背景蛋白质组”表单的“名称”字段中输入“Yeast\_mini”。
* 单击“浏览”按钮。
* 导航至先前创建的 ExistingQuant 文件夹下的 MRMer 子文件夹。
* 选择“Yeast\_MRMer\_mini.protdb”文件。
* 单击“打开”按钮。
* 在“编辑背景蛋白质组”表单中单击“确定”按钮。

“肽段设置”表单将显示如下：

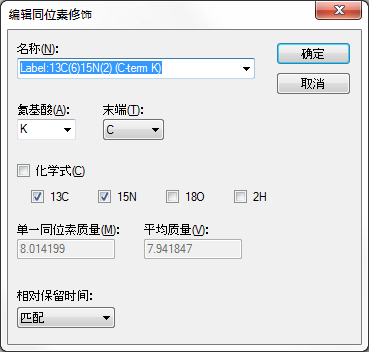


最后，在将 MRMer 实验的离子对列表插入当前文档之前，您还需要规定如何对其所包含的肽进行同位素修饰。MRMer 实验包括未标记轻肽段和匹配的重肽段，后者使用稳定同位素标记氨基酸残基 (SILAC) 对赖氨酸和精氨酸进行标记。如果事先没有指定正确的同位素修饰，就插入 MRMer 离子对列表，Skyline 将无法识别出离子对列表中重链肽的质核比值。

要在 Skyline 文档设置中指定 SILAC 标记，请执行下列步骤：

* 单击“修饰”标签。
* 单击“同位素修饰”列表旁边的“编辑列表”按钮。
* 单击“编辑同位素修饰”表单中的“添加”按钮。
* 在“编辑同位素修饰”表单的“名称”字段内输入“Label:13C(6)15N(2) (C-term K)”。
* 在“氨基酸”字段中输入“K”。
* 在“末端”下拉列表上，选择“C”。
* 选中“13C”和“15N”复选框，告诉 Skyline 对赖氨酸分子中的所有碳原子使用 13C，所有氮原子使用 15N，总质量偏移为 8 道尔顿 (6x 13C + 2x 15N)。

此时“编辑同位素修饰”表单将显示如下：

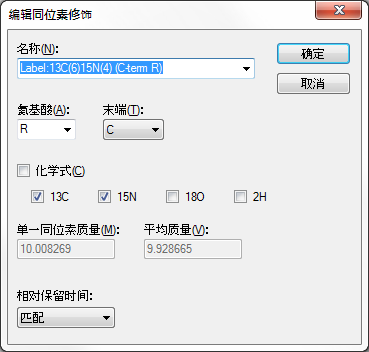


单击“确定”按钮，再通过以下步骤添加第二个同位素修饰：

* 单击“编辑同位素修饰”表单中的“添加”按钮。
* 在“编辑同位素修饰”表单的“名称”下拉列表中选择“Label:13C(6)15N(4) (C-term R)”。

自动选中“13C”和“15N”复选框，告诉 Skyline 对精氨酸分子中的所有碳原子使用 13C，所有氮原子使用 15N，总质量偏移为 10 道尔顿 (6x 13C + 4x 15N)。

此时“编辑同位素修饰”表单将显示如下：



Skyline 自动计算单一同位素质量和平均质量，赖氨酸 (K) 约为 8 道尔顿，精氨酸 (R) 约为10 道尔顿，这是由于在这些氨基酸残基中使用了 13C 和 15N。

要准备将 MRMer 公布的离子对列表插入到当前文档中，请执行下列步骤：

* 单击“编辑同位素修饰”表单中的“确定”按钮。
* 单击“编辑同位素修饰”表单中的“确定”按钮。
* 选中“同位素修饰”列表中您刚创建的“Label:13C(6)15N(2) (C-term K)”和“Label:13C(6)15N(4) (C-term R)”修饰对应的复选框。
* 还要确保选中“结构修饰”列表中的“脲甲基半胱氨酸”复选框，因为在默认情况下其并不存在于 Skyline 的新安装中。
* 在“肽段设置”表单中，单击“确定”按钮。

您将看到 Skyline 文档区域出现的空白谱图。现在您已准备好插入 MRMer 离子对列表。

# 插入具有相关蛋白质的离子对列表

将离子对列表插入 Skyline 的方法有两种：

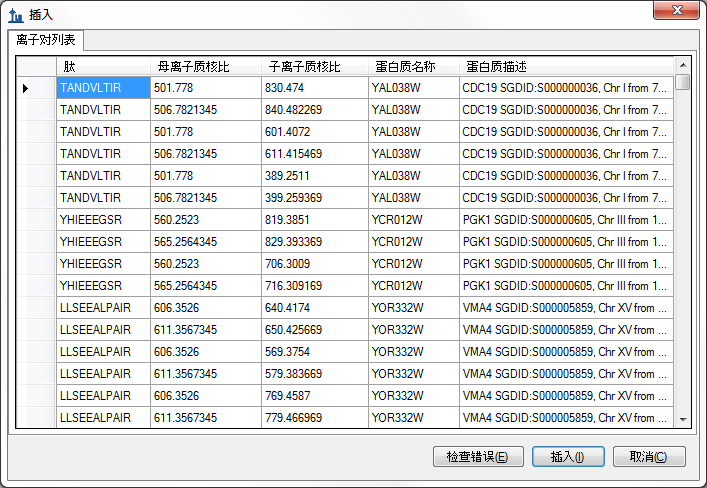
1. 单击“编辑”>“插入”>“离子对列表”，插入到所显示的表单。
2. 单击“编辑”>“粘贴”，直接插入到文档。

对于 MRMer 数据集，可使用第一种方法。对于研究 7 数据集，可使用第二种方法。如果您的文档包括背景蛋白质组，则采用“插入”表单的第一种方法的优势在于可以将肽与包含这些肽的蛋白质自动关联。当前仅用于在背景蛋白质组中出现在单一蛋白质中的肽，而Skyline 的未来版本将提供一些如何处理存在于多个蛋白质中肽的选择。对于本教程，已经删除出现在多个蛋白质中的 MRMer 离子对列表中的两个肽。

如要向当前文档添加离子对列表肽，请执行下列步骤：

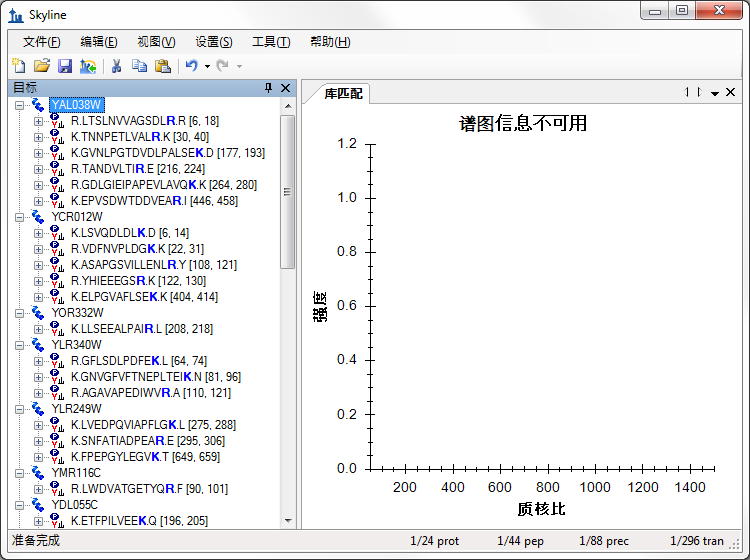
* 使用 Microsoft Excel 打开为本教程创建的 ExistingQuant 文件夹下面 MRMer 子文件夹中的“silac\_1\_to\_4.xls”文件。此 Excel 文件由来自于 MRMer 下载的初始“silac\_1\_to\_4.transition.tsv”（制表符分隔值）文件创建而来，该文件可用于涉及从 1 到 4、从轻到重的比值混合的实验。
* 确保在水平滚动条左侧标记有“固定”的电子表格中的页面处于活动状态。
* 选择所有 296 行的前 3 列单元格：肽段序列（A 列）、母离子质核比（B 列）、子离子/片段离子质核比（C 列）。
* 复制单元格 (ctrl-C)。
* 切换回 Skyline。
* 在“编辑”菜单上，选择“插入”，并单击“离子对列表”。
* 在键盘上按下 ctrl-V，粘贴复制的单元格。

“插入”表单现在将显示如下：



已添加复制的单元格，以及相关蛋白质的名称和描述。Skyline 通过匹配肽序列与背景蛋白质组文件中的蛋白质，自动添加蛋白质信息。如要向 Skyline 文档中插入这些肽，请单击“插入”按钮。

Skyline 主窗口将显示若干在其蛋白质内部分组的肽。肽图标应包含 3 条垂直线和一条基线 ()，就好像在其右下角有一个非常小的 MS/MS 谱图 ()。存在此图像即表明肽具有相关的 MS/MS 库谱图。肽标记应将 C-末端赖氨酸或精氨酸以蓝色突出出来，这显示出在重链标记形式中稳定同位素标记的氨基酸：

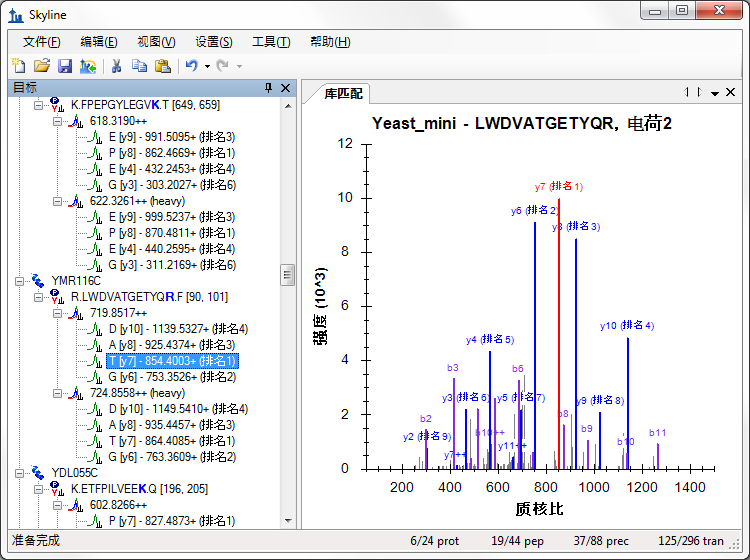


您还可以在窗口的右下角状态栏中看到指示符“296 tran”，确定 MRMer 离子对列表中的所有 296 个离子对已添加到此文档中。在状态栏的左侧，您可以看到此文档包含 24 个蛋白质、44 个肽段和 88 个母离子，或者每个肽对应 2 个母离子。

如要进一步查看在此文档中的母离子和离子对，请执行以下操作：

* 在“编辑”菜单上，选择“全部展开”并单击“母离子”。

花一些时间来选择肽视图中的各个肽和离子对。预览这些离子对，并如何对在 MS/MS 库谱图中的这些子离子峰按照强度大小排名。如要在肽树形视图中改变选项，则 MS/MS 谱图会作出更新，显示出与当前肽相匹配的谱图，同时，与已选离子对相匹配的谱峰还会以红色突出显示：



您将注意到并非所有的肽都具有最充足的子离子供离子对选择，正如在上述图像中显示的肽那样，同时也并非所有的谱图均具有可匹配性。

如未看到 b-离子或者突出显示的双电荷离子，您可以通过进行以下菜单选择使 Skyline 将其显示出来：

* 在“视图”菜单上，选择“离子类型”并单击“B”。
* 在“视图”菜单上，选择“电荷”并单击“2”。

您可能已经注意到 Skyline 对于轻母离子和重母离子都显示相同的谱图。此谱图库可能仅包含轻母离子的谱图，然而，虽然此库包含轻和重母离子的匹配谱图，但是 Skyline 仅使用其中一个（不用轻母离子的），以避免两个 MS/MS 谱图之间不同的强度等级出现差错。如果此库仅包含匹配肽段的重型、已标记形式的谱图，那么 Skyline 将使用该谱图为轻和重母离子的离子对排名。

# 导入数据

当然，像这样从现有离子对列表创建 Skyline 文档最有意义的原因是使用 Skyline 检验利用原始离子对列表在三重四级杆 MS 仪器上采集的数据。

如要将 MRMer 公布提供的数据导入到您已创建好的文档中，请执行下列步骤：

* 在“文件”菜单上，单击“另存为”。
* 导航至先前创建的 ExistingQuant 文件夹下的 MRMer 子文件夹。
* 在“文件名”字段内输入“MRMer”。
* 单击“保存”按钮。
* 在“文件”菜单上，选择“导入”并单击“结果”。
* 单击“导入结果”表单中的“确定”按钮。
* 选择文件列表中的“silac\_1\_to\_4.mzXML”文件。来自 Waters Quattro Premier 的最初原始文件（无法使用）被转换为 mzXML，是因为 MRMer 无法阅读仪器原始数据文件格式。
* 单击“打开”按钮。

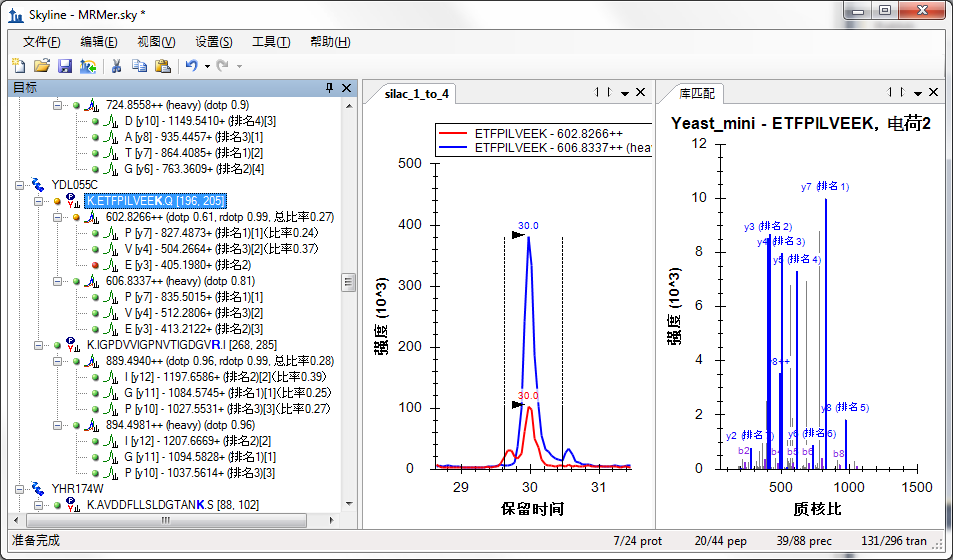
Skyline 开始把文件导入其高效数据高速缓冲存储器中，这时需要更少的磁盘空间，Skyline 可以从中非常快地检索到所需信息。Skyline 窗口底部的状态栏将显示进度。

导入完成后，在整合边界之间已测出信号的离子对（以黑色虚线显示）将会有绿色圆圈添加到离子对图标的左侧。未包含在已选峰值组中的峰值，其离子对会显示出红色圆圈。仅含有绿色圆圈的离子对，其母离子和肽段还显示出绿色圆圈。此为非常好的数据。您将主要看到绿色圆圈。

如要看到显示为红色圆圈的离子对，请执行以下步骤：

* 在“编辑”菜单上，单击“查找肽” (ctrl-F)。
* 在“序列片段”字段内输入“ETFP”。
* 单击“查找下一个”按钮。

此操作将使 Skyline 显示如下：

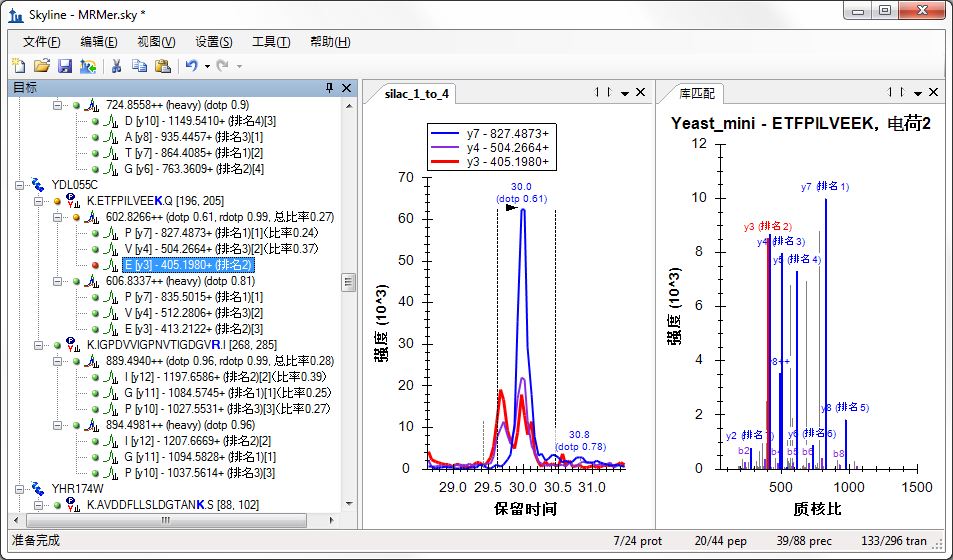


相反，假如在 Skyline 窗口中图表窗格按照计划相互之间加以叠放，您可以通过单击其中一个图表标签并拖动至新的位置上，重新安排这些窗格。您还可以通过单击并拖动所有窗格之间的分割控制条来改变分配到各自的比例。

如果不想让色谱图视图按照上述图像中所显示的那样放大，请执行以下步骤：

* 在“视图”菜单上，选择“自动缩放”并单击“最佳峰值”(F11)。

如要进一步了解此肽段的 y3 离子未包含在已选取谱峰组之内的原因，请在肽视图中选择该离子。Skyline 将发生变化并显示如下：



如果未看到如上所述的所有三个离子对的色谱图，请执行以下步骤：

* 在“视图”菜单上，选择“离子对”并单击“全部” (shift-F10)。

现在您可能会看到 y3 和 y4 似乎存在一些干扰，以至于它们显示出两个未完全分离的谱峰。这些很显然不是由于相同肽段造成的，是因为它们均不存在 y7，而且这两个离子对之间的相对强度随着两个峰之间的不同而不同。

如果您仍然在优化该方法的过程中，则您可以下次尝试测定 y5 和y8，因为 MS/MS 库谱图表明这两个离子可能具有可测性。然而，如果您确实至少想得到这些数据的初始量测，则您有两种选择：

1. 删除受到干扰影响的峰值。
2. 调整峰值边界，以排除产生干扰的区域。

在 MacCoss 实验室，我们倾向于第一种选择，是因为第二种选择根据让人们对未知干扰的界限的合理判断将会引入未知的差异。然而，在本教程中，您将尝试这两种选择。

## 移除存在干扰的离子对峰

如要从轻峰值组中除去 y4 离子对，请执行以下步骤：

* 右键单击色谱图。
* 在右键菜单上，选择“删除峰值”并单击“y4 - 504.2664+”。

在肽视图中，您会看到文本“（比率为 0.38）”从轻离子对的末端消失，同时母离子比值“（总比率为 0.27）”变为“（总比率为 0.24）”。

母离子总面积比值是加权平均值，以内标物峰面积作为权重值，采用代数法简化为：

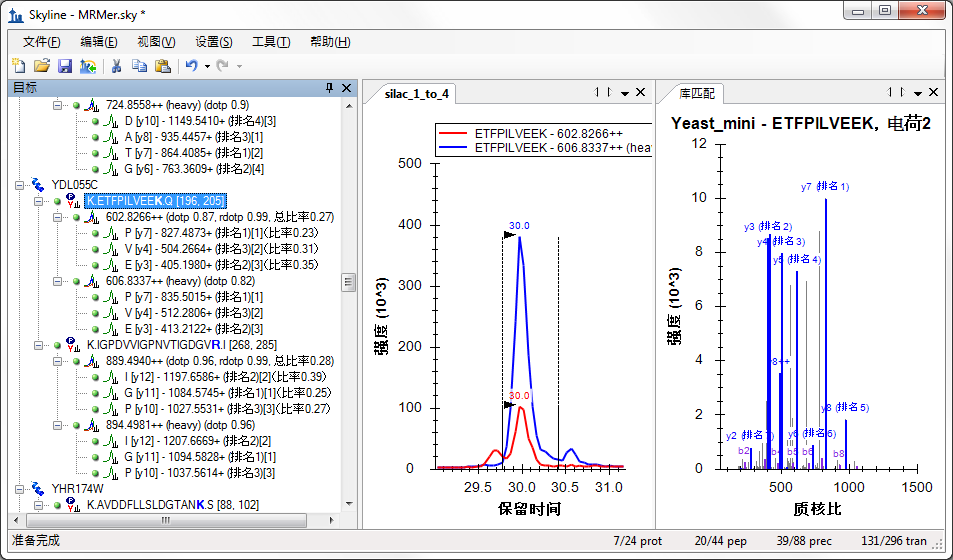
移除轻峰值或重峰值将移除与其相匹配的离子对的峰值。因此，仅有一对离子对保留时，正如本研究中的 y7，在离子对的轻峰值和重峰值之间的比值（此处为 0.24）也将成为母离子的总面积比值。

## 调整峰边界以排除干扰

如要通过调整峰的整合边界以达到删除定量测定中的干扰的目的，请执行下列步骤：

* 在“编辑”菜单上，单击“撤销” (ctrl-Z)，以撤销删除最后部分的峰值。
* 选择肽段“ETFPILVEEK”。
* 在 X-轴下面单击鼠标，也即在轻（红色）色谱图在预期峰值与干扰峰值（约 29.8 分钟）之间达到了最小值的右侧，并拖动至重（蓝色）色谱图中右边出现的第二个峰值（约 30.4 分钟）。

此操作将使 Skyline 显示如下：



在这种情况下，y4 和 y3 之间的比值显示为 0.31，y7 的显示为 0.23，产生的加权平均母离子总比率为 0.26。

采用两种不同的处理干扰作用的方法，获取的两个总比值 0.24 和 0.26，与 1 至 4（由轻至重）SILAC 混合的真实值 0.25 同样接近。然而，显示存在干扰作用（EFP 肽段的 y3 和 y4）的峰值比值更接近于 1:3，而不是 1:4，提出手动调节技术如何实际运作的疑问，并强调我们更倾向于从定量计算中将这样的离子对完全移除的原因。

对本文档中的数据进一步检查，将显示这些肽段的大多数比值非常接近于预期值 0.25。此外，对于具有 4 个或更多个离子对的母离子，Skyline 在峰面积和匹配 MS/MS 峰强度之间显示点积值 (dotp)。对于完全匹配，这些数值的大部分都非常接近于 1.0。

最后，在转向本教程中的第二个文档之前，您可能会注意到 2 个肽段（K.YVDPNVLPETESLALVID**R**.L 和 K.FPEPGYLEGV**K**.T）的离子对仅为空白区，而其他具有绿色或红色圆圈。这就意味着之前导入的 mzXML 文件未包含这些离子对的数据。您可以在文本编辑器中打开 mzXML 文件，查找母离子质核比，以亲自确认母离子缺少由原始离子对列表表示的离子对。对于处理手动创建的离子对列表，或通过比 Skyline 接收更少使用和测试的工具创建的离子对列表，这些异常现象都属于非常典型的情况。

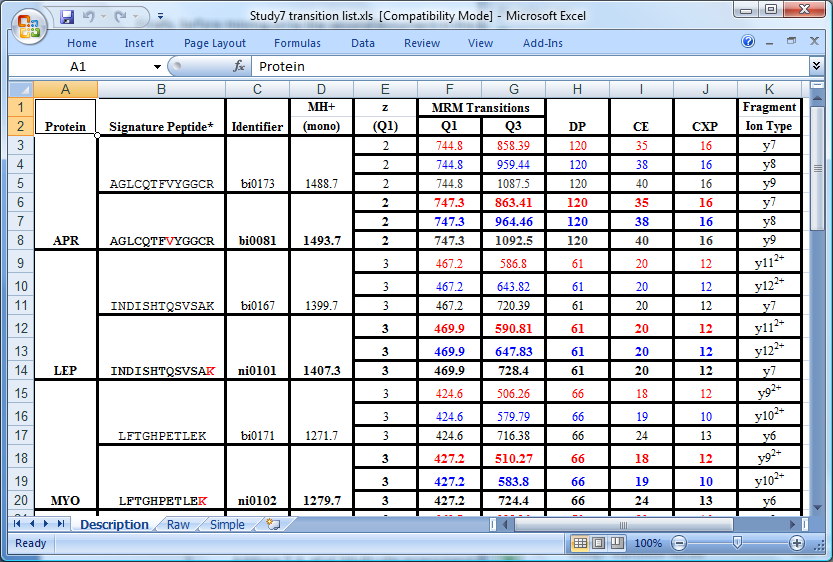
# 为 CPTAC 研究 7 准备文档

在下一个章节中，您将处理跨实验室 APTAC 研究 72，该项研究在 Skyline 第 0.1 版首次发行之前由验证工作组完成。利用电子表格创建方法，并利用供应商特定的软件分析所产生的质谱仪数据。

# 准备接受研究 7 离子对列表的文档

同样，首先要完成的第一项任务是从现有的离子对列表中创建 Skyline 文档。向 Skyline 中插入离子对列表的第一步即是检查离子对列表，了解 Skyline 要求完成哪些设置才能识别出列表中的质核比值。如要开始这项检查，使用 Microsoft Excel 打开之前在本教程中创建的 ExistingQuant 文件夹下研究 7 子文件夹中的“Study7 transition list.xls”文件。

您将看到如下所示的电子表格：



您可以判定这是一个由某人手动创建的电子表格，创建人尽最大努力、使用界限、合并单元格和突出显示，使这一系列离子对让其他人看起来也是清晰明了的，而 Skyline 现在可以自动处理这项工作。

这个列表中的每一肽段均具有轻形式和重形式。在每一个重肽的“Signature peptide”（署名肽）列中，单个稳定同位素标记的氨基酸残基以红色突出显示。向下滚动整个列表，您将看到有 4 种标记方案：

1. 6 个肽具有标记的 C-末端赖氨酸，增加了 8 道尔顿。
2. 2 个肽具有标记的 C-末端精氨酸，增加了 6 道尔顿。
3. 2 个肽具有标记的内部缬氨酸，增加了 5 道尔顿。
4. 1 个肽具有标记的内部亮氨酸，增加了 6 道尔顿。

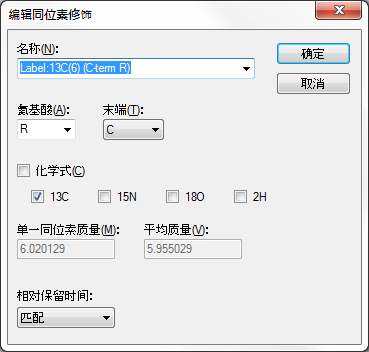
这个标记方案不可能像在 MRMer 文档中使用的一样仅代表总体修饰，因为有一些赖氨酸和精氨酸标记的肽段还包括内部缬氨酸和亮氨酸。处理这种现象最简单的方式就是对 C-末端赖氨酸和精氨酸再次使用总体修饰，正如您在 MRMer 文档中做的那样，然后再手动应用缬氨酸和亮氨酸修饰。

如要为研究 7 离子对列表准备具有合适修饰的新文档，请执行下列步骤：

* 在 Skyline 工具栏上，单击“保存”按钮，以保存对 MRMer 文档所做的改变。
* 在 Skyline 工具栏上，单击最左边的“新文档”按钮。
* 在“设置”菜单上，单击“肽段设置”。
* 单击“修饰”标签（若有必要）。
* 单击“同位素修饰”列表旁边的“编辑列表”按钮。
* 选择“Label:13C(6)15N(4) (C-term R)”修饰，以便在下方添加将您的新修饰。
* 单击“编辑同位素修饰”表单中的“添加”按钮。
* 从“编辑同位素修饰”表单中的“名称”下拉列表中选择“Label:13C(6) (C-term R)”。

自动选中“13C”复选框，告诉 Skyline 对赖氨酸分子中的所有碳原子使用 13C，总质量偏移为 6 道尔顿 (6x 13C)。

此时“编辑同位素修饰”表单将显示如下：



如要完成为研究 7 离子对列表准备文档，请执行以下步骤：

* 单击“编辑同位素修饰”表单中的“确定”按钮。
* 单击“编辑同位素修饰”表单中的“确定”按钮。
* 选中刚才创建的“Label:13C(6) (C-term R)”修饰。
* 取消选中为 MRMer 文档创建的“Label:13C(6)15N(4) (C-term R)”修饰，如果未取消选中的话。
* 确保仍然选中为 MRMer 文档创建的“Label:13C(6)15N(2) (C-term K)”修饰的复选框。
* 还要确保已选中“脲甲基半胱氨酸”的“结构修饰”列表中的复选框。
* 单击“库”标签。
* 取消选中在 MRMer 文档中使用的“Yeast\_mini”库，如果未取消选中的话。
* 单击“酶解”标签。
* 如果尚未选择“无”，则在“背景蛋白质组”中选择“无”。
* 在“肽段设置”表单中，单击“确定”按钮。

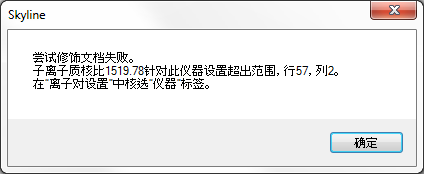
## 将离子对列表粘贴到文档

现在切换回到 Excel，并单击研究 7 电子表格中的“raw”（原始）标签。在本页面，您将看到具有 66 个离子对的原始离子对列表已导入至本研究所采用的 4000 QTRAP 仪器中。由于您知道 Skyline 仍然缺少处理缬氨酸和亮氨酸标记的肽段中重离子对设置方面的信息，现在在底部水平滚动条的左侧单击“Simple”（简单）电子表格标签，您将发现与移除的 9 个离子对具有相同页面的版本。

如要把这些离子对添加到新的空白 Skyline 文档，请执行下列步骤：

* 选择研究 7 电子表格“Simple”页面上的所有 6 列和 57 行。
* 复制单元格 (ctrl-C)。
* 切换回 Skyline。
* 在“编辑”菜单上，单击“粘贴” (ctrl-V)。

Skyline 将呈现出以下错误消息：



在处理 Skyline 之外创建的离子对列表时，出现此类消息并不少见。在本主题的 Skyline 教学视频中还列出了其他消息（[视频 3：现有实验](https://skyline.washington.edu/labkey/wiki/home/software/Skyline/page.view?name=video_0-5b)）。出现这种类型的错误最常见的原因有以下几个：

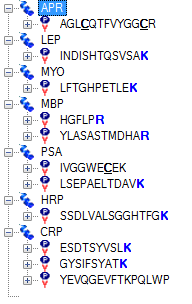
1. 忘记指定必要的修饰
2. 手动或利用其他工具创建的质核比值出现错误
3. 包括超出当前仪器设置的范围的质核比值

最后一个可能存在的原因是由于当前的离子对列表，而且您可以看到指导您到“离子对设置”表单中的“仪器”标签的消息。

如要纠正这一问题，请执行下列步骤：

* 在错误消息中单击“确定”按钮。
* 在“设置”菜单上，单击“离子对设置”。
* 单击“仪器”标签。
* 在“最大质核比”字段中输入“1800”。
* 单击“确定”按钮。
* 在“编辑”菜单上，再次单击“粘贴” (ctrl-V)。
* 在“编辑”菜单上，选择“全部折叠”，并单击“肽” (ctrl-shift-D)。

此操作将使 Skyline 肽视图显示如下：



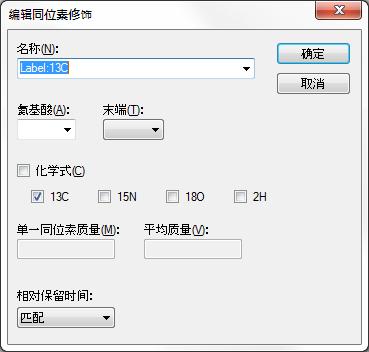
## 手动调整修饰

八个肽显示加粗、蓝色的 C-末端 K 或 R，但其他 3 个缺少标记的氨基酸。这些都是需要 V 或 L 明示标记的肽段，因为通过文档设置无法实现这些修饰。

如要指定第一肽段中 V 上的同位素修饰，请执行下列步骤：

* 选择“AGL**C**QTFVYGG**C**R”肽段。
* 在“编辑”菜单上，单击“修改肽段”。
* 从 C-末端 R 的“重同位素”下拉列表上选择空白条目，移除该精氨酸中稳定同位素标记。
* 在该肽段中 V 残基的“重同位素”下拉列表中选择“<添加…>”。
* 从“编辑同位素修饰”表单的“名称”下拉列表中选择“Label:13C”。

“编辑同位素修饰”表单将显示如下：



依据其包含的碳原子的数量，这种修饰将应用于所有氨基酸的各种质量偏移。单击“确定”按钮返回“编辑修饰”表单，表单现在显示如下：

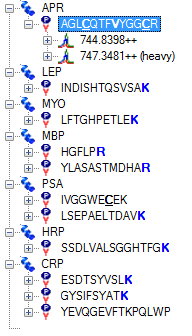


单击“编辑修饰”表单中的“确定”按钮，返回文档。

选中的肽段中的缬氨酸将不会加粗，因为此肽段尚未包含重母离子。如要现在添加重母离子，请执行以下步骤：

* 将鼠标悬停在此肽段上，并单击出现在其右侧的箭头。
* 选中弹出选择列表中的“747.3481++（重）”复选框。
* 单击左上角的绿色选中按钮（或者按 Enter 键）。

肽视图现在将如下所示：



现在展开“744.8398++”母离子和“747.3481++（重）”母离子，单击轻母离子和重母离子左侧的“+”符号，确认其包含匹配离子对，且如预期的那样它们之间相差 5 道尔顿。

如要对剩下两种肽段创建标记的母离子，请执行下列步骤：

* 右键单击肽段“IVGGWE**C**EK”，并单击“修饰”。
* 从 C-末端 K 的“重同位素”下拉列表上选择空白条目，从该赖氨酸中删除稳定同位素标记。
* 在该肽段中 V 残基的“重同位素”下拉列表中选择“Label:13C”修饰。
* 单击“确定”按钮。
* 将鼠标悬停在此肽段上，并单击出现在其右侧的箭头。
* 选中弹出选择列表中的“541.7637++（重）”复选框。
* 单击左上角的绿色选中按钮（或者按 Enter 键）。
* 右键单击肽段“YEVQGEVFTKPQLWP”，然后单击“修饰”。
* 在该肽段中 L 残基的“重同位素”下拉列表中选择“Label:13C”修饰。
* 单击“确定”按钮。
* 将鼠标悬停在此肽段上，并单击出现在其右侧的箭头。
* 选中弹出选择列表中的“913.9746++（重）”复选框。
* 单击左上角的绿色选中按钮（或者按 Enter 键）。

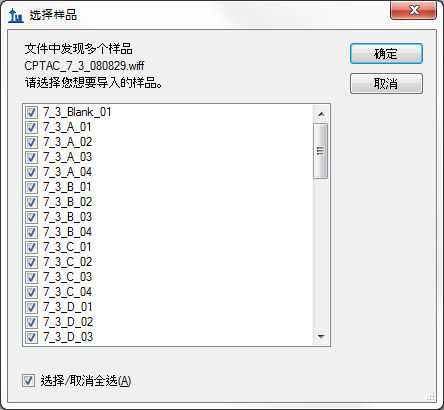
现在已创建好 Skyline 文档，准确地反映出在初始研究 7 离子对列表中的信息。将此文档保存至 ExistingQuant 文件夹下的 Study 7 子文件夹中的 “Study 7.sky”。

## 从多重样品 WIFF 文件导入数据

由于四个主要三重四级杆质谱仪的支持，现在 Skyline 安装了完全支持导入所有格式的文件，不需要进行任何转换。这就意味着您可以通过完成以下步骤从使用 4000 QTRAP 测定这些离子对的其中一个位点向此文档中导入数据：

* 在“文件”菜单上，选择“导入”并单击“结果”。
* 单击“导入结果”表单中的“确定”按钮，导入单次注射重复测定。
* 在研究 7 文件夹中选择“CPTAC\_7\_3\_080829.wiff”文件。
* 在“导入结果文件”表单中单击“打开”按钮。

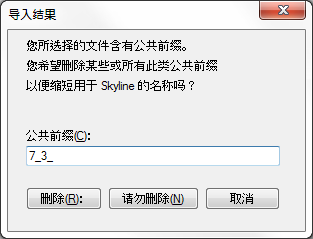
Skyline 可能会花一两秒的时间读取多样品 WIFF 文件中的样品名称列表，然而它将呈现出如下所示的表单：



对于本教程，为了缩短完成导入所需的时间：

* 取消选中包含文本“空白”（1 个样品）、“QC”（4 个样品）和“梯度洗脱”（4 个样品）的条目。
* 同时还取消选中在最后的“A2”（4 个样品）和“A3”（4 个样品）条目。
* 单击“确定”按钮。

Skyline 应呈现出以下信息，允许从所有名称中移除重复前缀“7\_3\_”，这些名称将用于显示这些重复测定的相关消息：



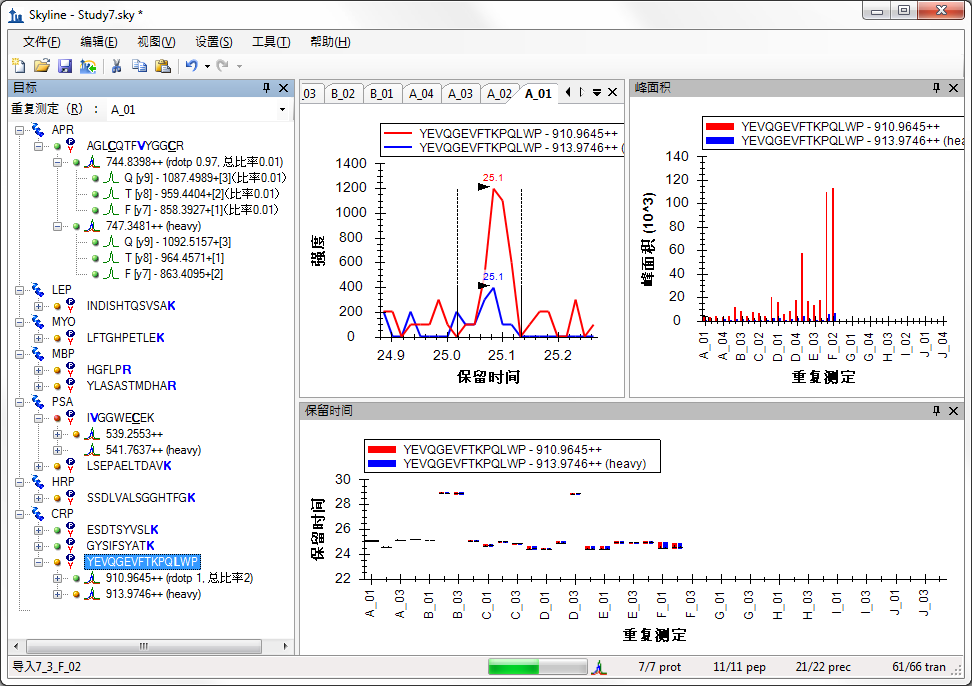
单击“删除”按钮，接受这项操作。

Skyline 开始将此 WIFF 文件中的数据导入其高效数据缓存文件 (Study 7.skyd) 中，此时数据存取更快，并将文档中导入的所有数据压缩成单一文件以便共享。

## 检查和调整峰值整合

Skyline 导入原始数据时，完成此项工作需要花费几分钟的时间，您可以开始重新安排图表视图，以达到最佳视图效果。如要显示并重新定位峰面积和保留时间重复测定比较视图，请执行以下操作：

* 在“视图”菜单上，选择“保留时间”并单击“重复测定比较” (F8)。
* 单击并拖动窗口至 Skyline 窗口的底部，将其停放在那里。
* 在“视图”菜单上，选择“峰面积”并单击“重复测定比较”(F7)。
* 单击并拖动窗口至 Skyline 窗口的右边，奖项停放在那里。
* 调整视图比例，使之如下图所示。



目前已选择的肽段“YEVQGEVFTKPQ**L**WP”（此为 CRP 蛋白质的C-末端上的肽段），事实上对于研究 7 中的验证工作组有些问题。在导入数据时，您可以看到在连续重复测定过程中 Skyline 有时仍无法整合峰值。仍然有几个保留时间异常值，甚至有一些保留时间非常接近于 25分钟的峰值没有显示为相同的肽段，因为大多数峰值是在约 24.7 分钟整合。一个清晰的问题是相对于轻、内源形式而言，重、加标形式的强度从未表现得非常强。

值得注意的是，在手标出现之前您可以将鼠标悬停在任何一个复制图表的状态栏上，然后单击导航至相应重复测定的色谱图。您可以使用这种方法导航至每一个很可能识别错的峰值所对应的色谱图，并利用与在 MRMer 文档中使用的相同方法（即单击并向下拖动 X 轴）对其纠正。但是对于本教程，您可以效仿验证工作组在其后续实验中的方式，仅删除该肽段。

您将看到 Skyline 对此文档中的其他肽段做了更好的整合。比起之前发行的 Skyline 版本而言，目前 Skyline 第 0.7 版在自动峰值整合方面得到了大大的改进。然而，在真实的数据集中，此文档还包含一些要求手动整合的离子对。

首先，在上述图像中，您可以看到肽段“I**V**GGWE**C**EK”的母离子“541.7637++（重）”似乎缺失数据。这是因为提供给质谱仪的离子对列表仅指定了母离子质核比的小数点数，并将其错误地四舍五入为“541.7”。您可以在 Excel 电子表格中“Raw”标签上检查这项内容。

如要正确地计算此文档中的母离子，与测定的数据相匹配，请执行以下步骤：

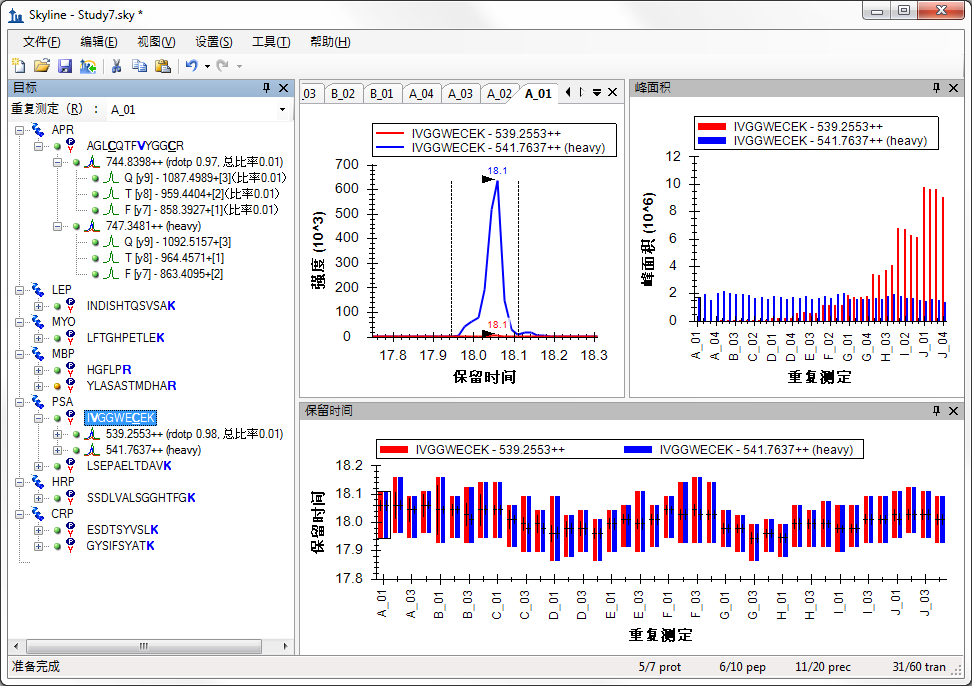
* 在“设置”菜单上，单击“离子对设置”。
* 在“仪器”标签的“质核比匹配耐受性”字段内输入“0.065”
* 单击“确定”按钮。

这将导致在“541.7637++（重）”母离子旁边出现一个绿色圆圈。

另一个整合问题是，在肽视图中有许多元素显示出橙色和红色圆圈，这表明离子对没有整合的峰值面积。在方法优化时这一点是非常有用的，正如在[目标方法调整](https://brendanx-uw1.gs.washington.edu/labkey/wiki/home/software/Skyline/page.view?name=tutorial_method_refine)教程中讨论的那样。但是对于像这样的高精密方法，这倾向于产生不太准确的定量测定。对于这一原因，已添加了这项选项，即在最强峰值的界限范围之间对所有离子对的峰面积进行整合。如要现在打开这项功能，请执行以下操作：

* 在“设置”菜单上，单击“全部合并”。

此时 Skyline 将显示如下：



## 用峰面积视图检验数据

上述图像中的“峰面积”视图明确告诉您“研究 7 实验是什么”的概念。它是以技术四联测量的各个浓度的校正曲线。您会看到，在恒定浓度下对重同位素标记的内标物进样，但经过 50 多次进样后，其峰面积稍微减小。在“保留时间”视图，您可以看到肽保留时间非常稳定，非常适合应用于预定的运行中。

在本节内容中，您将专注于“峰面积”视图以及其提供的用于检查多个重复测定数据集的许多选项。在屏幕上为“峰面积”视图提供更多的空间：

* 双击标记有“峰面积”的标题栏。

“峰面积”视图将弹出 Skyline 主窗口，并在其上浮动。重新调整其位置，使之免于遮挡肽视图。

开始选择此文档中剩下的 10 个肽段，您将看到，第一个肽段和最后 5 个肽段呈现出峰面积图表，与上述图像中的图表类似，而在这些肽段之间的 4 个肽段显得很不精密。

同时进样稳定同位素标记参照多肽的一个重要原因是，它们可用于对内源性未标记肽的峰面积进行归一化，应移除一些在测定结果中运行之间的差异。如要在“峰面积”视图中形象地看到这些，请执行以下操作：

* 右键单击“峰面积”图表，选择“归一化”并单击“重”。

再次检查之前表现正常的 6 个肽段。选择肽段 SSDLVALSGGHTFG**K**，您现在将看到如下所示的归一化图表：



确实，您将发现每一种浓度下的重复测定所获得结果之间的精密度得到了提高。

再看下其他 4 个肽段，您将发现它们仍然未显示出预期的模式。

查看此归一化图表的另一种有趣的方式是分别查看每一个离子对。在行为正常的肽段中，每一个轻离子对与其对应的重离子对之间的比值将较为相似。如要分别复查离子对的比值，请执行以下操作：

* 在“编辑”菜单上，选择“全部展开”，并单击“肽”(ctrl-D)。
* 选择每一个肽段的轻母离子。

对于 6 个行为正常的肽段，您会看到如下所示的图表，ESDTSYVSL**K** 肽段轻母离子 564.7746++ 被选中：



不出所料，这些比值非常相似。第二个和第三个肽段 (INDISHTQSVSA**K**和LFTGHPETLE**K**) 并不是非常干净，但在离子对比值中不是主要问题。然而，对于第四个肽段 (HGFLP**R**)，选择其轻 363.7059++ 母离子时，会产生如下所示的峰面积图表：



这看起来显示出在 y3 的离子对中存在干扰，因为在低浓度下其比值与此肽段的其他比值不相同。

“峰面积”视图提供了检查母离子中离子对相对强度的另一种方式。现在您可以通过执行以下操作来再次查看 HGFLP**R** 肽段：

* 右键单击“峰面积”图表，选择“归一化”，并单击“总数”。

图表将发生变化并显示如下：



此图表再次清晰地显示出 y3（棕色）的离子对中存在干扰，随着内源性肽的浓度增加（超出 E 重复测定时的浓度），其已变成更少的因素。如果您移动峰面积视图到您同时还可以看到色谱图表的位置，您这时可以单击各个标题条以查看存在干扰的峰值，在这种情况下，该峰值非常清晰，如下文中重复测定 E\_03 所示：



与 MRMer 文档中的情况一样，您可能为了排除干扰峰而尽力调整整合界线，但只是同时删除 y3 离子对很可能是一种更好的办法，尤其是因为该离子对从未对此肽段的总峰面积起到非常大的作用。如要深入了解这些测定结果的精密度，请执行以下操作：

* 在“视图”菜单上，选择“峰面积”，并单击“肽段比较” (ctrl-F7)。
* 在“视图”菜单上，选择“离子对”，并单击“总数” (ctrl-F10)。
* 右键单击“峰面积”图表，并单击“CV 值”。

“峰值面积”图表将发生变化并显示如下：



如果该肽段未按照这个顺序出现，请执行以下操作：

* 右键单击“峰面积”图表，选择“优化”并单击“文档”。

当然，对于被分析物多肽（以红色显示），有很大的变异系数 (CV) 数值，因为数据集包含被分析物蛋白质及其多肽的 10 个不同浓度水平。然而，重肽（以蓝色显示）的 CV 值提供了大量的信息，因为将其在恒定浓度下进样至所有样品中。鉴于到目前为止您所看到的，不会惊奇地发现，能清晰显示出其预期浓度变化的 6 个肽段也具有约为 10% 或更低的变异系数值，而其他 4 个肽段的变异系数值更接近于 40%，甚至 50%。

在 Skyline 中仅有几个简单的操作，您就可以了解到这一数据集的相关信息。此数据集要求数周时间以及学习原始研究的验证工作组中统计员和程序员的参与。如果您已经完成了[靶向方法优化](https://brendanx-uw1.gs.washington.edu/labkey/wiki/home/software/Skyline/page.view?name=tutorial_method_refine)教程，现在您可能在想为什么您要运行优化周期，这一优化涉及到预期安排好的多次重复测定的数据集，以便能更好地理解目标多肽在测定高浓度样品之前在这项设置中将会做出怎样的反应。

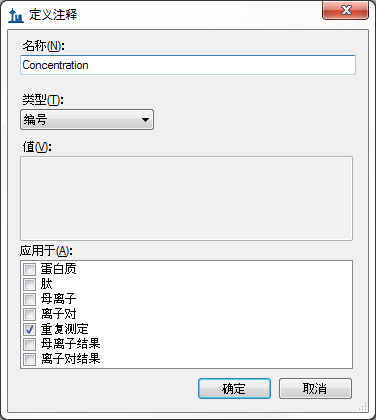
## 以浓度值注释重复测定

如果 X-轴显示样品浓度而不是重复测定的名称，则将使对图表中数据的解释变得更加容易。

Skyline 允许您通过定义“注释”的方式关联重复测定的其他信息。

* 在“设置”菜单上，单击“注释…”。
* 单击“注释设置”表单中的“编辑列表”按钮。
* 单击“定义注释”表单中的“添加”按钮。
* 在“定义注释”表单中的“名称”字段内输入“Concentration”（浓度）。
* 在“类型”字段内选择“编号”。
* 选中“应用于”列表中“重复测定”左侧的复选框。

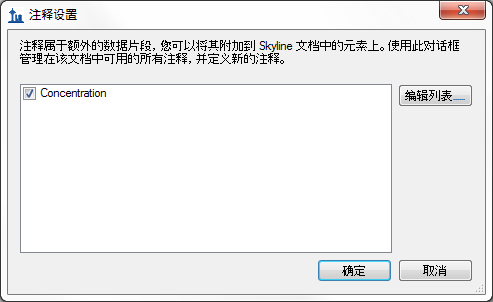
“定义注释”表单现在将如下所示：



现在完成对注释的定义，并通过完成以下内容将其添加到您的文档中：

* 单击“定义注释”表单中的“确定”按钮。
* 单击“定义注释”表单中的“确定”按钮。
* 选中新的“Concentration”注释的复选框。

“注释设置”表单现在将如下所示：



* 单击“确定”按钮。

可使用“结果网格”编辑这些注释数值。如要显示结果网格视图，请执行以下操作：

* 在“视图”菜单上，单击“结果网格” (Alt-2)。

“结果网格”将显示峰面积以及其他测定的值。为了确保“Concentration”列可见，请执行以下操作：

* 在肽树形视图中选择蛋白质“APR”。

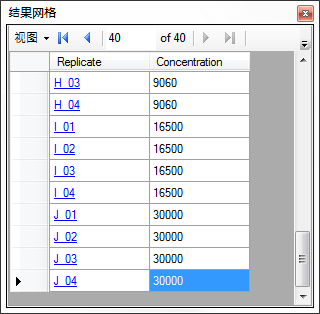
现在“结果网格”将仅显示出重复测定名称和浓度注释。配制样品的浓度为：

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 样品 | A | B | C | D | E | F | G | H | I | J |
| 浓度 (fmol/µL) | 0 | 60 | 175 | 513 | 1500 | 2760 | 4980 | 9060 | 16500 | 30000 |

利用上述表格，填写各个重复测定的“浓度”值。如果您在向“结果网格”中输入内容时看到 Skyline 主窗口在闪动，您可能需要通过完成以下内容允许 Skyline 改变网格中的行，而不是活动相应的重复测定色谱图：

* 右键单击“结果网格”，并单击“同步选择”。

在你正在输入这些数值时，“结果网格”的底部将如下所示：



可以使用这个新注释的地方之一是“峰面积重复测定比较”图表。

* 在肽树形视图中选择肽段“SSDLVALSGGHTFGK”。
* 在“视图”菜单上，选择“峰面积”并单击“重复测定比较”(F7)。
* 右键单击“峰面积”图表，选择“分组依据”并单击“浓度”。
* 右键单击“峰面积”图表，选择“归一化”并单击“重”。

“峰面积”图现在将如下所示：



该图显示直至最低浓度下的 CV 值非常小。通过执行以下操作，您可以轻易切换到显示平均比值和标准偏差的图表：

* 右键单击“峰面积”图表，单击“CV 值”。

图将发生变化并显示如下：



## 进一步探索

八个不同标签参与了研究 7，并且每一个标签均产生了与在本教程中复查的数据相类似的几个数据集。然而，没有任何标签以相同的方式经历了相同的问题。 本教程还包含来自不同位点和本研究中不同的小节的数据。在这种情况下，已创建好文档，并且已导入数据集。

如要打开含有研究 7-II 来自于位点 52 的数据的文档，请执行下列步骤：

* 在“文件”菜单，单击“打开” (ctrl-O)。
* 导航至本教程开始时创建的 ExistingQuant 文件夹下的 Study 7 子文件夹中的 Study II 子文件夹。
* 选择“Study 7ii (site 52).sky”文件。
* 单击“打开”按钮。

此文件会快速打开，并且您会看到进样同位素标记对照品的 CV 值不尽相同，如下述图表中所示：



肽段INDISHTQSVSA**K**仅是来自第一组数据集的 4 个肽段的其中之一，此数据集的变异系数仍然约为 40%，而现在肽段 LSEPAELTDAV**K** 的变异系数约为 25%。然而，如果您在“保留时间”视图中看到“重复测定比较”图表，您会看到 Skyline 从其中 3 次重复测定中选择错误的峰值（尽管这很可能会固定在将来的发行版本中）。



如果您通过单击并向下拖动 X-轴的方式纠正这些错误，LSEPAELTDAV**K** 的重母离子变异系数下降至其他大多数肽段那样的 10% 左右。

如要查看此数据集中的轻：重比值，请执行下列步骤：

* 在“视图”菜单上，选择“峰面积”并单击“重复测定比较”(F7)。
* 右键单击“峰面积”图表，选择“归一化”并单击“重”。

您将发现所有肽段的图表，即使 INDISHTQSVSA**K** 在此数据集中看上去非常好：



这似乎显示出内标物对此肽段测定结果的差异性起到补偿作用。查看未归一化的数据：

* 右键单击“肽面积”图表，选择“归一化”并单击“无”。

从这张图表中很难想象到归一化会如此有效：



最后，回到我们在第一组数据集中检测到的存在干扰的肽段上来：

* 选择肽段 HGFLP**R** 的轻母离子 363.7059++。
* 右键单击“峰面积”图表，选择“离子对”并单击“全部”。
* 右键单击“峰面积”图表，选择“归一化”并单击“重”。

另外在这里，您会看到 y3 离子对中存在干扰的明显证据：



然而，和第一组数据集不同，通过检查各个色谱图发现干扰作用显得更困难些。在这种情况下，由于稍微改变色谱条件，几乎不可能在 y3 的整合范围中找到两个峰值，如下面所示的重复测定 E\_03：



NCI CPTAC 研究 7 提供了更多的数据，现在这些数据均已公开。Skyline 可以帮助您迅速大量了解这些数据，即使知晓 Skyline 存在的参与者并没有参与完成本项研究。

# 结论

通过本教程的学习，您了解到 Skyline 可以大大简化来自未考虑 Skyline 而设计及执行的实验中的数据的处理过程，无论这些实验是您自己的实验，抑或是您开始使用 Skyline 之前进行的实验，抑或是您希望回顾或重复进行的他人曾经进行的实验。对于 MRMer 和 NCI CPTAC 研究 7 公布的数据集，您很快从离子对列表创建好了 Skyline 文档，甚至这些列表包含相当复杂的修饰方案，包括同位素标记内标物。

您还了解到一些 Skyline 的特征，即 Skyline 可以提供有效的定量靶向蛋白质组实验，包括同位素标记的肽母离子。从同位素标记修饰的定义到设计，Skyline 简化了这些实验创建仪器方法的工作。从精确的轻：重峰面积比值到强大的图表显示选项，Skyline 助您深入了解这些实验采集的数据。

# 参考文献列表

1. Martin,D.B. *et al.*MRMer, an interactive open source and cross-platform system for data extraction and visualization of multiple reaction monitoring experiments.*Mol.Cell Proteomics.***7**, 2270-2278 (2008).

2. Addona,T.A. *et al.*Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma.*Nat.Biotechnol.***27**, 633-641 (2009).