

Skyline iRT 保留时间预测

长久以来，肽段保留时间的预测一直是靶向蛋白质组学感兴趣的研究内容。早在 0.2 版时，Skyline 就整合了基于未修饰的肽段序列计算疏水性的 SSRCal 算法¹作为一种预测肽保留时间的方法，在“[靶向方法优化](#)”教程中有详细的描述。到了 0.5 版，Skyline 就开始支持目前最先进的“预定采集”技术。“[靶向方法优化](#)”教程中对这种技术进行了介绍：首先在正常色谱条件下，对关注的全部目标肽段进行非预定采集质谱检测，只要该平台的色谱条件未发生变化，这些肽段的色谱保留时间信息就可以被用于预定后续重复实验中目标肽段的保留时间。

基于非预定质谱采集获取肽段保留时间有一项缺点，即可能需要进行多次质谱实验，当色谱条件发生任何变化都需要重新开始一次质谱的非预定采集实验，无论这样的色谱条件变化是来源于不同实验室之间的方法共享还是实验室内部不同仪器，又或者是实验过程中同一台仪器上仅仅是色谱分离柱的替换。例如，在 MacCoss 实验室的一次实验中，为了完成在一次实验中对 780 个离子对进行预定采集的 45 次重复测定，总共进行了 5 次未预定采集。在 NCI-CPTAC 验证研究小组最近进行的一次实验中，为了完成在一次实验中对 750 个离子对进行预定采集的 150-200 次重复进样检测，总共进行了 6 次未预定分析。这项研究跨 11 个实验室，采用了 14 种仪器，并且一些实验室还采用了不同的色谱柱条件来完成进样检测。显然，如果一项技术能够利用已经测定的肽保留时间，跨实验室、仪器平台和不同的色谱分离梯度使用，将极大地简化靶向实验。这样，只需要进行一次靶标肽段的校正实验，就可以在后续所有的实验分析中使用肽段的色谱保留时间信息。

更准确保留时间的预测也能够用于峰识别验证。例如，在一般的实验中，假设色谱保留时间预测的 2 倍标准差 (2σ) 距离平均值是 5 分钟，如果能够精确到 1 分钟，将会筛选出更多可信的候选峰。

iRT 标准²由我们在 Biognosys (<http://www.biognosys.ch/>)的合作者提出，在 Skyline 1.2 版就已经支持，该标准可实现高精度以及具有很好可移植性的肽段色谱保留时间预测。在本教程中，您会学习到如何将 30 分钟色谱梯度上测定得到的色谱保留时间存储为 iRT 值，然后用这些 iRT 值来为 90 分钟色谱梯度的实验分析设计靶向方法。您还将了解更小的 iRT 预测误差会提高峰鉴定的可靠性。此外，您还将学会如何把数据依赖型采集 (DDA) 实验图谱库中的肽段保留时间转化为 iRT 值。这样，可以直接利用探索性实验分析中的肽段保留时间信息，来设计排靶向实验。当然，这仅仅需要一次校正进样实验。

入门指南

为了开始本教程，请首先下载以下 ZIP 文件：

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

将其中的文件解压到您的电脑上的某个文件夹，如：

C:\Users\brendanx\Documents

这将创建一个新的文件夹：

C:\Users\brendanx\Documents\iRT

该文件夹包含了本教程所需要的全部文件。在 Windows Explorer 中双击此文件夹中的“iRT-C18 Standard.sky”文件，或通过单击 Skyline “文件”菜单中的“打开”，均可打开该文件。

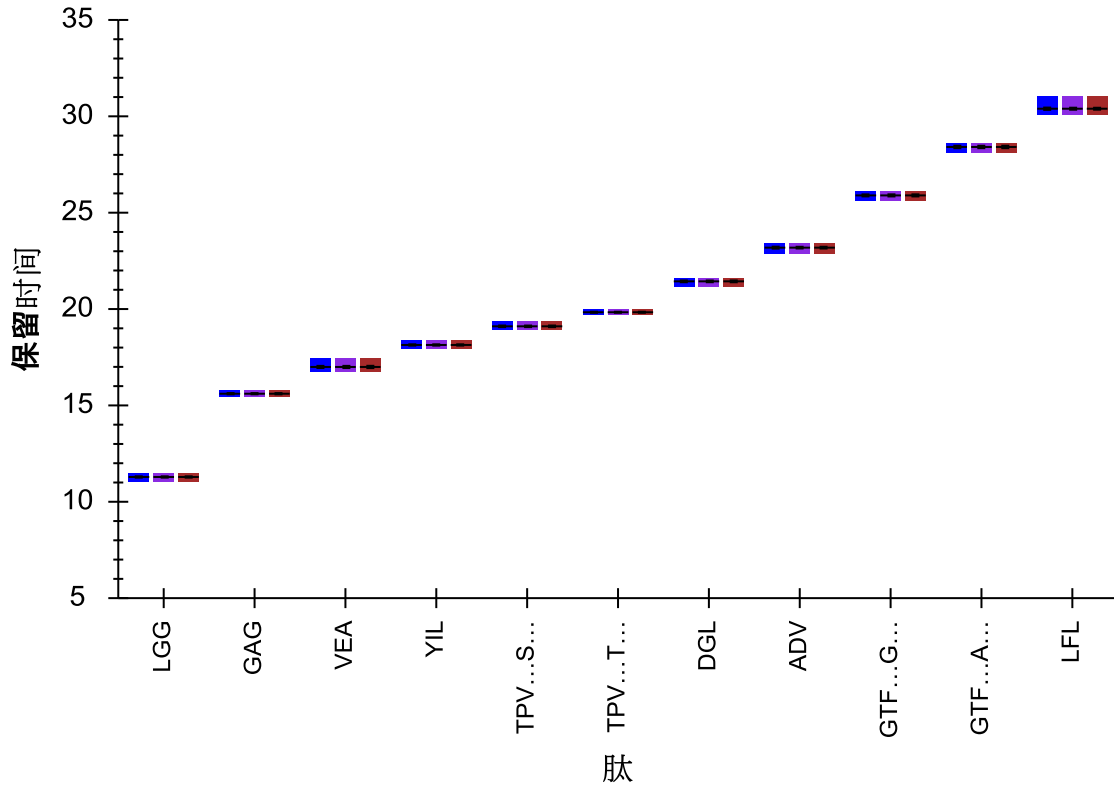
校准 iRT 计算器

虽然在本教程中，您将仅仅使用 Biognosys 定义的 iRT-C18 标准以及标准肽段混合物 (<http://www.biognosys.ch/products/rt-kit.html>)，但 iRT 本身是一个通用概念，可以用于任意易于检测并且覆盖您所使用的色谱分离梯度的标准肽段混合物。在对已打开的文档做出任何修改之前，单击“文件”菜单中的“另存为”，在 iRT 文件夹中保存一个新的副本，命名为“iRT-C18 Calibration.sky”。

此时，为了开始本教程，我们假定您已经在自己的仪器上完成了一批合适的标准肽段检测。执行以下操作，准备校准新的 iRT 计算器：

- 在“文件”菜单上，选择“导入”并单击“结果”。
- 单击“导入结果”对话框中的“确定”按钮。
- 选择 iRT 文件夹中列出的、您为本教程创建的前两个.raw 文件。
 - A_D110907_SiRT_HELA_11_nsMRM_150selected_1_30min-5-35.raw
 - A_D110907_SiRT_HELA_11_nsMRM_150selected_2_30min-5-35.raw
- 单击“打开”按钮。
- Skyline 要求删除公共前缀时，单击“删除”按钮。
- 在“视图”菜单中，选择“排列图”并单击“平铺”(Ctrl-T)。
- 在“视图”菜单中，选择“保留时间”并单击“肽段比较”(Ctrl-F8)。

此时 Skyline 将显示如下“保留时间”视图：



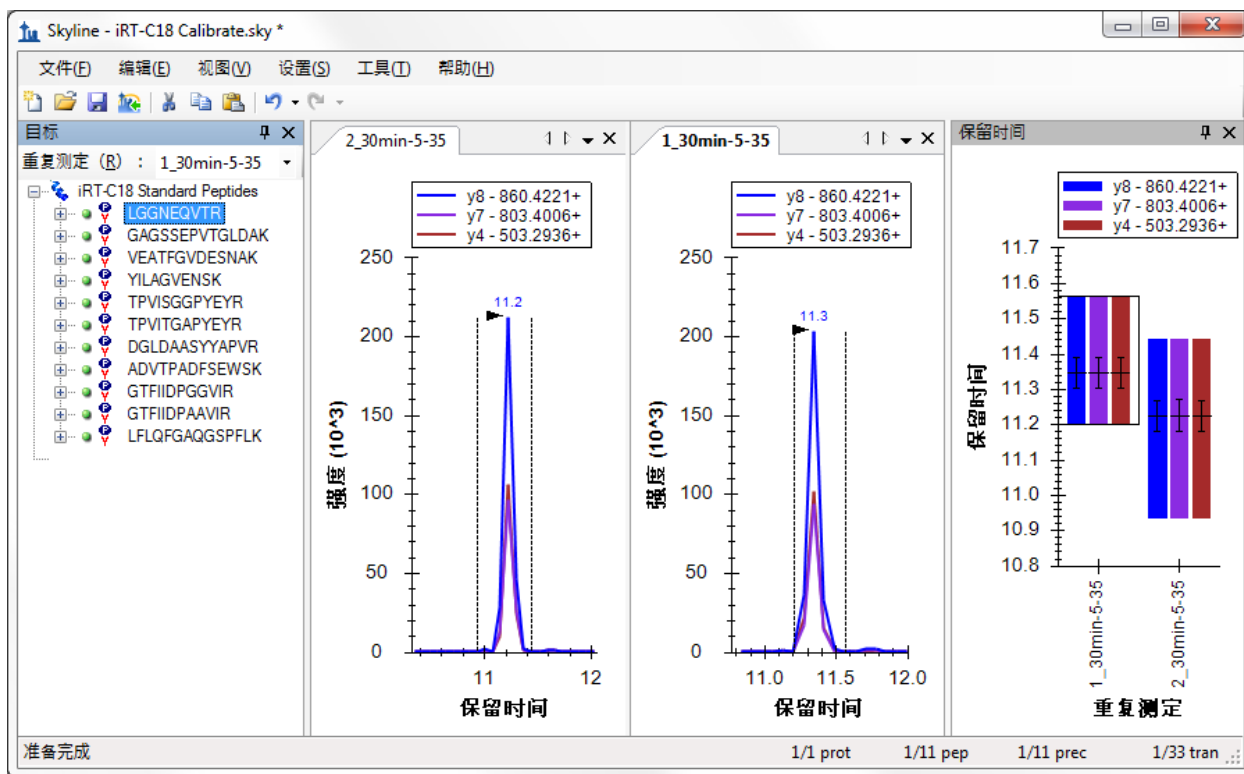
上图呈现了 30 分钟梯度条件下，每条肽段平均洗脱时间的顶层视图。执行以下操作继续查看数据：

- 在“视图”菜单上，选择“保留时间”并单击“重复测定比较”。
- 单击“保留时间”视图标题栏，将“保留时间”视图拖至 Skyline 窗口的边缘，直到鼠标光标位于图标内侧，右侧边缘附近，如图所示：




- 选择文档中的首个肽段 LGGNEQVTR。

此时，Skyline 将显示如下：

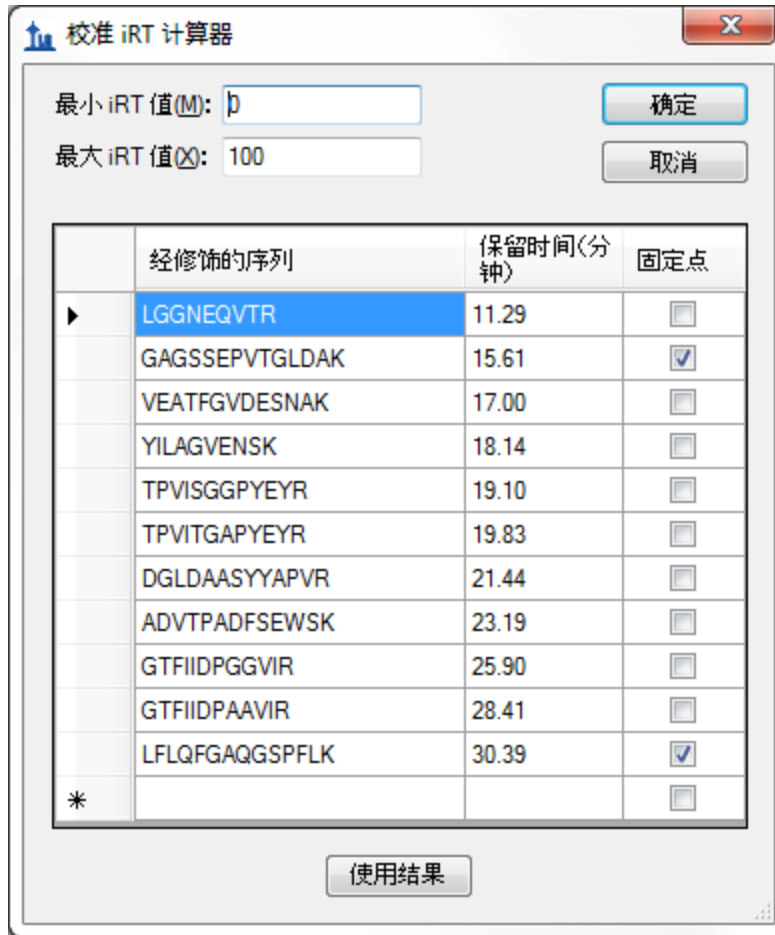


使用向下箭头键分别查看 Biognosys 标准肽段混合物中的 11 个肽段，确保色谱图积分边界的确定正确无误，并且两次重复检测均在相似的保留时间出现色谱峰。其实，自动积分非常适用于这些肽，您无需执行任何手动修改。本教程仅包括两次重复检测，如果您是实际校准新的 iRT 计算器，则可能需要使用更多的重复实验以改善对肽段平均保留时间的估计。在利用搜库结果校准保留时间计算器时，Skyline 将使用每个肽段多次检测的平均值。使用平均值估计保留时间，其精度与测量次数的平方根成正比。

验证校准数据质量后，执行以下步骤，以创建新的 iRT 计算器并对其进行校准：

- 在“设置”菜单上，单击“肽段设置”。
- 单击“预测”标签。
- 单击  (“保留时间计算器”) 按钮，并单击弹出菜单中的“添加”。
- 在“编辑 iRT 计算器”表单中的“名称”字段中，输入“iRT-C18”。
- 单击“创建”按钮。
- 在“创建 iRT 数据库”表单中，导航至您为本教程创建的 iRT 文件夹（如有需要）。
- 在“文件名”字段中，输入“iRT-C18”。
- 单击“保存”按钮。
- 单击“编辑 iRT 计算器”表单中的“校准”按钮。
- 单击“校准 iRT 计算器”表单中的“使用结果”按钮。
- 选中肽 GAGSSEPVTGLDAK 所在行中的“固定点”复选框。

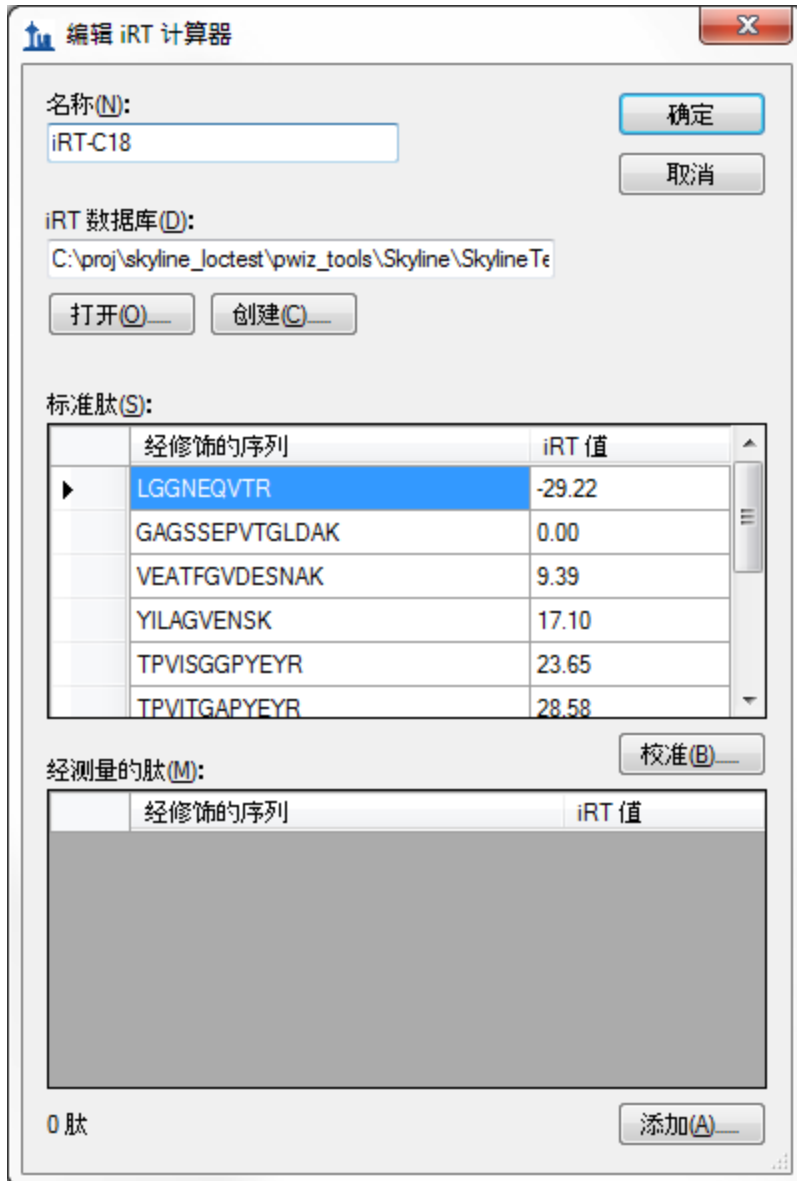
此时，“校准 iRT 计算器”表单将显示如下：



- 单击“确定”按钮。

注意：这仅仅是 Biognosys 定义 iRT-C18 标尺的方法。在您定义自己的标尺时，可以将固定点保留为第一个和最后一个洗脱肽段，也可以使用您所选择的任意其他肽段。也就是说，固定点和标尺的选择具有一定的随意性。因此，您可以定义与任意时间都无关的保留时间标尺，然后将其余 iRT 值的映射映射到该标尺上。

此时，“编辑 iRT 计算器”表单显示如下：



整个过程就是这样。此时，您已经利用实验数据校准了新的 iRT 计算器。实际上，在本案例中，想必 Biognosys 团队已经使用了 2 次以上重复测定，来校准其标准肽段混合物。因此，在本教程中，您可以简单地利用 Biognosys 提供的标准校准结果来替代这次的校准。为了更好地了解两者之间的接近程度，请执行以下步骤，：

- 单击“标准肽”表格。
- 按下 Ctrl-A 选择所有行。
- 按下 Ctrl-C 复制选定单元格中的文本。
- 在 Windows Explorer 中，导航至您为本教程创建的 iRT 文件夹。
- 双击 iRT 文件夹中的“iRT definition.xlsx”文件。
- 选择电子表格中的单元格 C2。
- 按下 Ctrl-V 粘贴复制的单元格。

- 按下 Ctrl 键并释放，然后按下 M 键以匹配目标格式。

此时，电子表格显示如下：

	A	B	C	D	E
1	sequence	iRT			
2	LGGNEQVTR	-24.92		LGGNEQV	-29.06
3	GAGSSEPVTGLDAK	0.00		GAGSSEPV	0
4	VEATFGVDESNAK	12.39		VEATFGVI	10.03
5	YILAGVENS	19.79		YILAGVEN	16.82
6	TPVISGGPYEYR	28.71		TPVISGGP	23.54
7	TPVITGAPYEYR	33.38		TPVITGAP	28.18
8	DGLDAASYAPVR	42.26		DGLDAASV	38.67
9	ADVTPADFSEWSK	54.62		ADVTPAD	50.27
10	GTFIIDPGGVIR	70.52		GTFIIDPGV	68.66
11	GTFIIDPAAVIR	87.23		GTFIIDPAV	85.38
12	LFLQFGAQGSPFLK	100.00		LFLQFGAC	100

正如您所看到的那样，新计算出的 iRT 值相当接近定义值，但您应使用定义中提供的值，以得到最佳结果。为此，请执行下列步骤：

- 选择 Excel 中的 A2 到 B12 单元格。
- 按下 Ctrl-C 复制这些单元格。
- 切换回“编辑 iRT 计算器”表单。
- 按下 Ctrl-V 粘贴定义值。

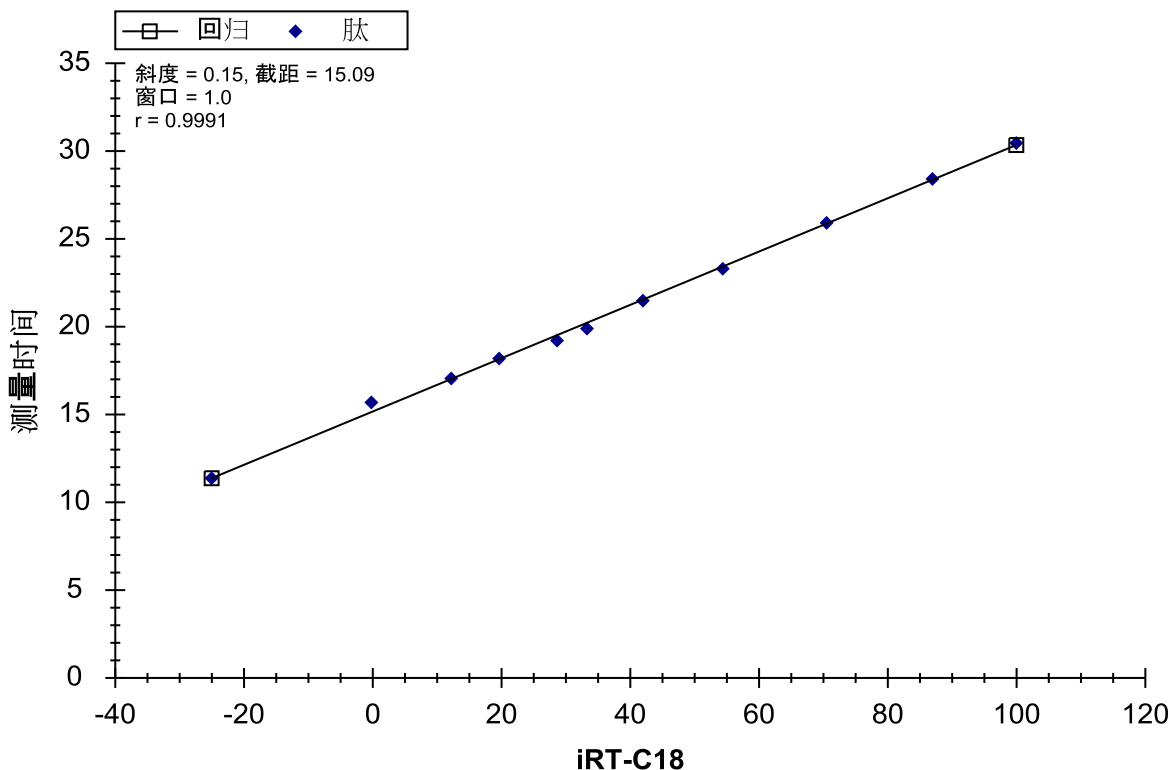
您将看肽段的色谱保留时间变成了定义值。

- 单击“编辑 iRT 计算器”表单中的“确定”按钮。
- 单击“肽段设置”表单中的“确定”按钮。

要查看肽段色谱保留时间定义值和实验测定值之间的相关性，进行下列步骤：

- 双击“保留时间”视图的标题栏，将其解锁。
- 在“视图”菜单中，选择“保留时间”并单击“线性回归”(Shift-F8)。

Skyline 将显示如下图:



在图形的左上角, 可见 Pearson 相关系数为 0.9991。如果您在 x 轴下方未看到 iRT-C18, 则需要执行下列操作:

- 右键单击“保留时间”图形, 选择“计算器”并单击“iRT-C18”。

如果您要分别查看两个已导入重复实验分析的回归图, 请执行下列操作:

- 右键单击“保留时间”图形, 选择“重复检测”并单击“单一”。

此时, 如果单击两个重复检测的色谱图形, 将看到截距值从 1_30min-5-35 的 15.15 变为 2_30min-5-35 的 15.04。差异非常微小, 但对更复杂数据集执行这类检查是十分必要的。

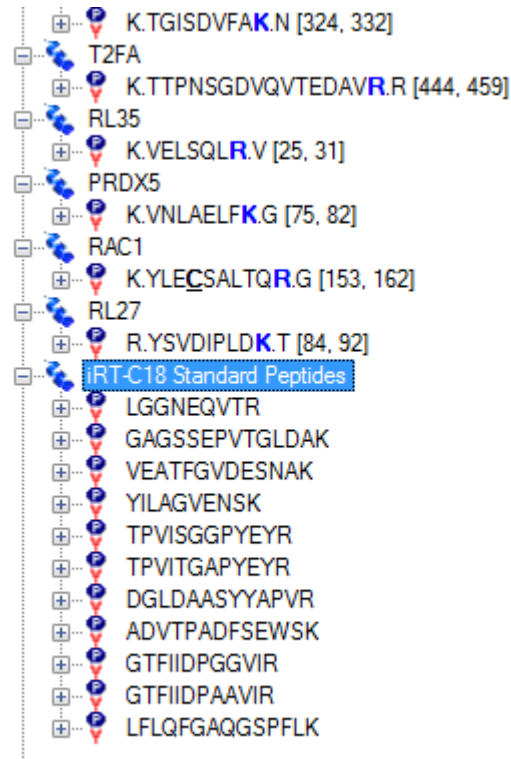
为新的靶向肽段添加 iRT 值

您现在已经拥有完全校准的 iRT-C18 计算器, 但如果不用来计算除了标准肽段以外的其它肽段的 iRT 值, 它对您用处不大。在这里, 您将从 SRM 实验得到的结果中, 把首个目标肽段添加至计算器。开始处理新肽段之前, 保存当前文件, 然后执行下列步骤, 以创建计算新目标肽段 iRT 值的数据文档:

- 在“文件”菜单, 单击“打开”。
- 双击您创建的 iRT 文件夹中的“iRT Human.sky”文件。

- 按下 END 键选择位于“肽视图”底部的插入元素。
- 在“文件”菜单，选择“导入”并单击“文档”。
- 双击“iRT-C18 Standard.sky”文件。
- 如果遗漏上述步骤，并试图保存此文档，则 Skyline 可能询问如何处理这些结果。在这种情况下，您应选择“删除结果信息”并单击“确定”按钮。
- 向下滚动查看并确认“iRT-C18 标准肽”列表已添加至文档末端。

“肽视图”显示如下：



将此文档另存为“iRT Human+Standard.sky”，然后再次将其另存为“iRT Human+Standard Calibrate.sky”。

创建 iRT 采集方法

如果您正在使用自己仪器为计算新的 iRT 值而采集数据，则可能需要获得数据的仪器采集方法。查看 Skyline 窗口的右下角，您将看到此文档当前包含 1231 个离子对，可能需要几次未预定检测来完成分析，但您可利用以下两个事实，减少工作量：

1. 该文档是为检测稳定同位素标记的参考肽段而设置。
2. 您只对保留时间感兴趣，而对定量分析并不感兴趣。

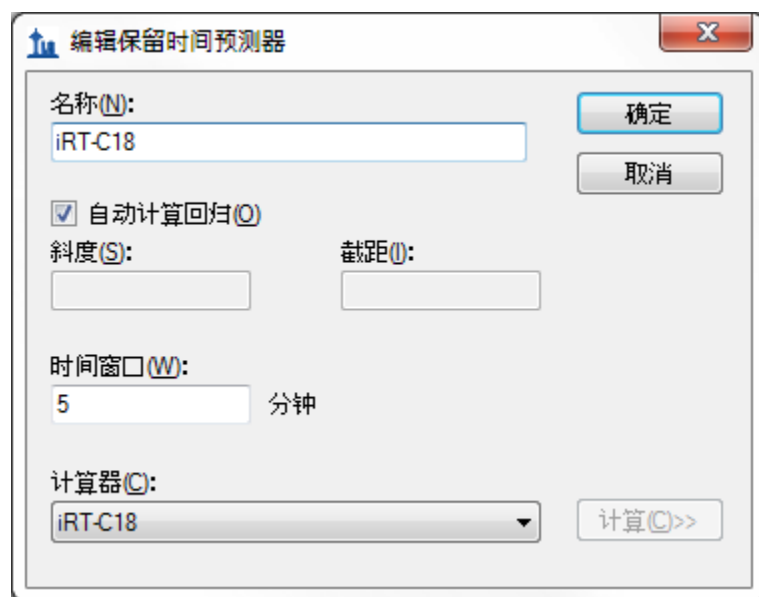
因此，实际需要处理的离子对数量可以减半，请执行下列步骤，以删除此文档中重同位素标记的母离子：

- 在“编辑”菜单中，选择“调整”并单击“高级”。
- 在“删除标记类型”字段中，选择“heavy”。
- 单击“确定”按钮。

您将看到离子对数量减少至 632。对于这些新的肽段，需要设计新的实验来检测色谱保留时间。在导出实验方法之前，您需要指明正在使用新的 iRT-C18 计算器，以便 Skyline 将标准肽段的离子对包括在所有方法之中。可以通过执行下列步骤完成此操作：

- 在“设置”菜单上，单击“肽段设置”。
- 单击“预测”标签（如果其尚未处于活动状态）。
- 在“保留时间预测器”字段中，选择“<添加.....>”。
- 在“编辑保留时间预测器”表单的“名称”字段中，输入“iRT-C18”。
- 在“计算器”字段中，选择新的计算器“iRT-C18”。
- 选中“自动计算回归”复选框。
- 在“时间窗口”字段中，输入“5”。

表单显示如下：



- 单击“确定”按钮。
- 确保“保留时间预测器”字段包含新的预测器“iRT-C18”。
- 单击“肽段设置”表单上的“确定”按钮。

为检测新的目标肽段，要导出未预定的离子对列表，以便在新的实验中使用。请执行下列步骤：

- 在“文件”菜单中，选择“导出”并单击“离子对列表”。

- 单击“多个方法”选项。
- 选中“忽视蛋白质”复选框。
- 在“每次进样允许的最多离子对数”字段，输入“335”。

此时，“导出离子对列表”表单将显示如下：

- 单击“确定”按钮。
- 在“导出离子对列表”保存表单的“文件名”字段中，输入“iRT Human+Standard Calibrate”。
- 单击“保存”按钮。

刚才的操作导出了两个用于新目标肽段色谱保留时间测定的离子对列表，均包括了 Biognosys 标准肽段混合物，这两个列表中的肽段完成实验测定色谱保留时间后，就可以用于计算新的 iRT 值。重要的是，在每次实验中，Biognosys 标准肽段的色谱保留时间都需要获得，Skyline 将自动处理此问题，即使每次重复实验中进行了多次进样分析。

对于自己的实验，您可以选择直接导出仪器方法，从而避免手动加载离子对列表。您可能想要设定更小的最多离子对数。的确，您只是为了挑选出一批可识别的色谱峰，用来获得可靠的肽段保留时间，但 335 这个数值已经过于激进了。请相信，生成此数据的 Biognosys 团队非常熟悉这些目标肽。因此，更常见的数值是 100-150，正如本教程前言部分提到的那样。

在 Excel 中打开生成的 CSV 文件“iRT Human+Standard Calibrate_0001.csv”和“iRT Human+Standard Calibrate_0002.csv”，您将看到适用于 Thermo TSQ 仪器的离子对列表。在每个文件的底部，您将看到 iRT 计算器中出现的标准肽段的离子对。

导入并审核数据

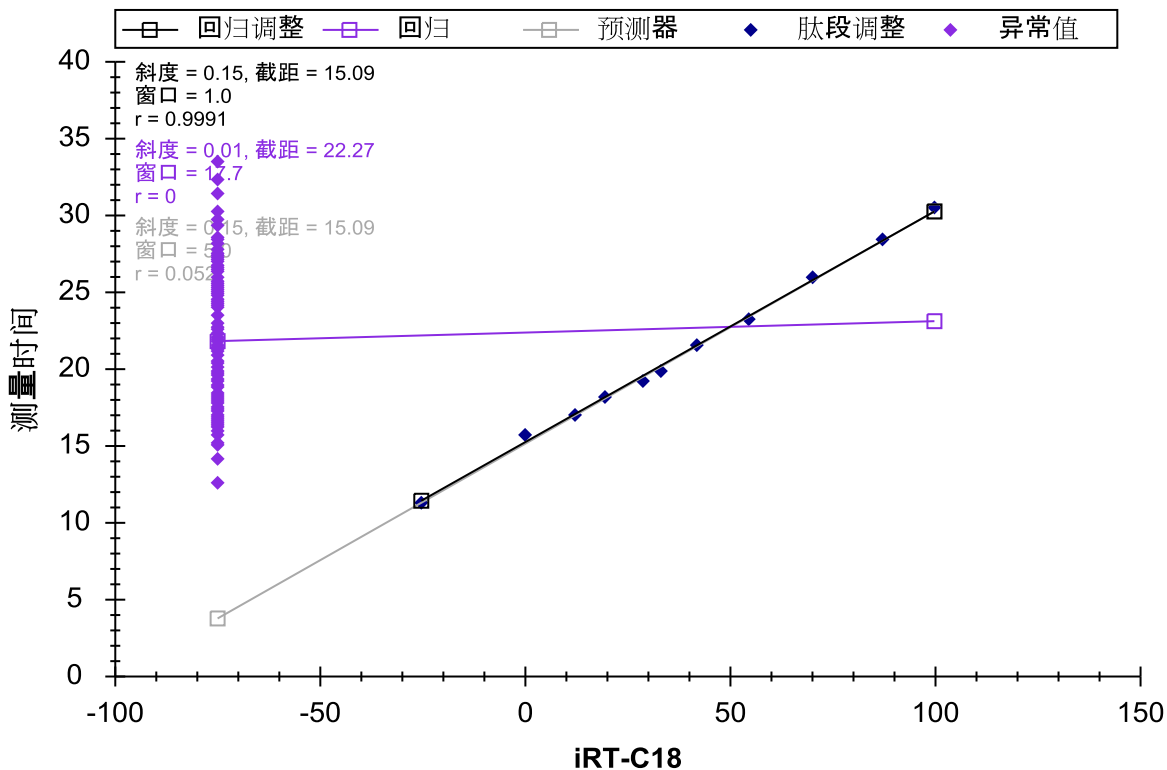
本教程文件夹中包含了您刚才创建的离子对列表的数据采集文件。实际上，您已在本教程的 iRT 校准部分导入了这些文件。只不过，您选择忽略了这些人类肽段的色谱图。要将文件导入当前文档，请执行下列步骤：

- 在“文件”菜单上，选择“导入”并单击“结果”。
- 单击“添加一个新的重复测定”选项。
- 单击“导入结果”表单中的“确定”按钮。
- 选择您为本教程创建的 iRT 文件夹中列出的前两个 .raw 文件。
 - A_D110907_SiRT_HELA_11_nsMRM_150selected_1_30min-5-35.raw
 - A_D110907_SiRT_HELA_11_nsMRM_150selected_2_30min-5-35.raw
- 单击“打开”按钮。

要查看 iRT-C18 计算器如何计算这些新肽段的分数，请执行下列操作：

- 在之前解除停靠的“保留时间”图形中单击右键（已经设置为显示线性回归），选择“计算器”并单击“iRT-C18”。

完成数据导入之后，您将看到如下图形：



图形左侧的紫色散点表示这些肽段尚无校准的 iRT 值。然而，计算 iRT 值之前，最好对峰面积积分边界进行审查。如果您真正希望对自己的 iRT 值进行准确校准，最好非常仔细地对所有肽段执行此操作。

在真正创建 iRT 之前，您可能希望利用一次未预定实验，来为多个重复的预定实验创建实验方法，然后利用平均值计算每个肽段的色谱保留时间，以提高精度。如果仅仅进行一次实验分析，基本的统计学知识告诉我们，大约有 5% 的肽段将会出现在 2 被标准差之外。

然而，本教程将使用一次实验的数据，对基本边界也仅仅进行粗略的检查。要审查 Skyline 在其中发现积分边界问题的肽段，请执行下列步骤：

- 在“编辑”菜单，单击“查找”(Ctrl-F)。
- 确保“查找项”字段空白。
- 单击“显示高级”按钮。
- 在“同时搜索”列表，选中“未整合的离子对”复选框。

“查找”表单将显示如下：



- 单击“查找全部”按钮。
- 单击“关闭”按钮。

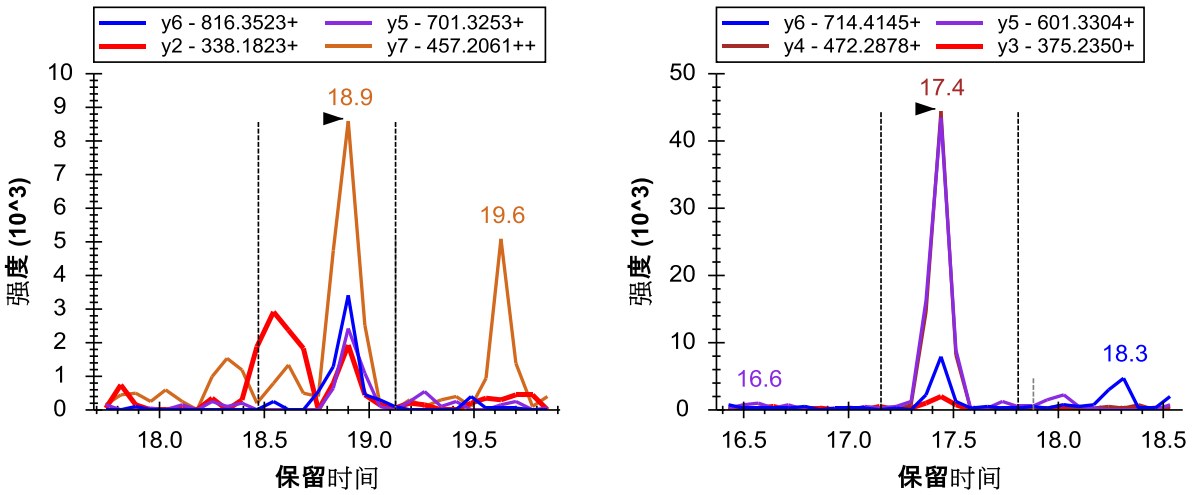
在 Skyline 窗口底部，您将看到“查找结果”视图，显示 8 个未完成积分的离子对：

位置	类型	文本
Y [y3] - 451.2663+ (色谱图)	离子对 结果	未整合的离子对
G [y6] - 642.3821+ (色谱图)	离子对 结果	未整合的离子对
Y [y2] - 338.1823+ (色谱图)	离子对 结果	未整合的离子对
A [y4] - 432.2817+ (色谱图)	离子对 结果	未整合的离子对
S [y3] - 375.2350+ (色谱图)	离子对 结果	未整合的离子对
I [y8] - 959.5156+ (色谱图)	离子对 结果	未整合的离子对
D [y5] - 575.3148+ (色谱图)	离子对 结果	未整合的离子对
Y [y3] - 409.2445+ (色谱图)	离子对 结果	未整合的离子对

要检查这些肽段的色谱峰，请执行下列操作：

- 在“视图”菜单，选择“离子对”并单击“全部”(Shift-F10)。
- 在“视图”菜单，选择“自动缩放”并单击“最佳峰值”(F11)。
- 双击“查找结果”视图中的每一行。

您将看到若干离子对干扰峰，但是其信号强度低于最强离子对色谱峰的 1%（按区域面积计算）。



Skyline 将排除这类离子对，从而帮助您确定保留哪些离子对，用于最终定量。但是，如果您已做出了决定，最好强制 Skyline 连续整合所有包含的离子对。为此，请进行以下操作：

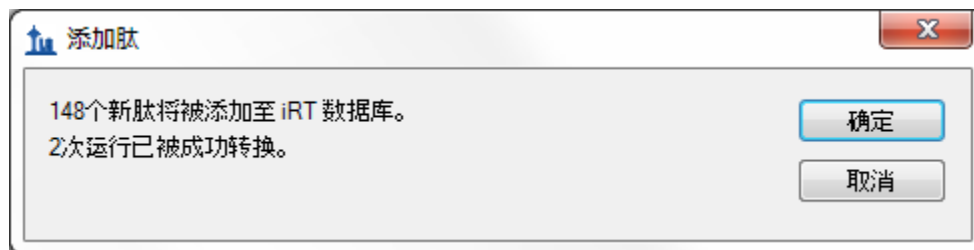
- 在“设置”菜单，单击“全部合并”。

计算 iRT 值

此时若要计算本档中目标肽段的 iRT 值，请执行下列步骤：

- 右键单击“保留时间”视图中的线性回归图，选择“计算器”并单击“编辑当前”。
- “编辑 iRT 计算器”表单中将显示 iRT-C18 计算器，单击“添加”按钮，并单击按钮下方显示菜单中的“添加结果”。

Skyline 将显示以下消息：



请注意，在为两次实验检测生成 iRT 值时，Skyline 会为每次实验分别进行线性回归，然后利用各自的回归模型计算该次实验检测肽段的 iRT 值。如果多次实验包含相同的肽段，则 Skyline 将采用这些最终计算所得 iRT 值的平均值。与直接计算物理上保留时间平均值的方法不同，这种方法允许跨多次实验，并且允许这些实验的色谱分离梯度不同。

- 单击“确定”按钮。

此时，“编辑 iRT 计算器”表单将显示如下：

名称(N): iRT-C18 确定 取消

iRT 数据库(D): C:\proj\skyline_loctest\pwiz_tools\Skyline\Skyline Te 打开(O)... 创建(C)...

标准肽(S):

经修饰的序列	iRT 值
LGGNEQVTR	-24.92
GAGSSEPVTGLDAK	0.00
VEATFGVDESNAK	12.39
YILAGVENS K	19.79
TPVISGGPYEYR	28.71
TPVITGAPYEYR	33.38

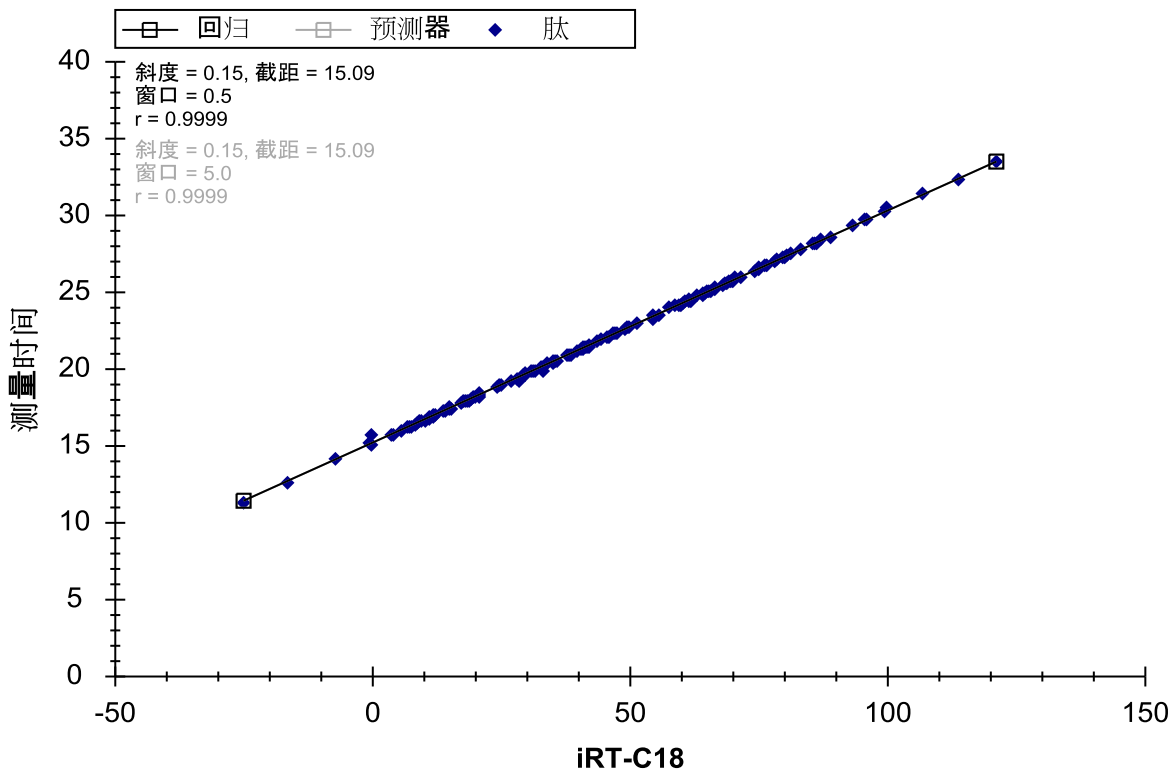
经测量的肽(M): 肽... 重新校准

经修饰的序列	iRT 值
NSAQGNVYVK	-16.31
ASTEGVAIQGQQGTR	-6.99
FGQGGAGPVGQQGPR	-0.35
LSSEMNTSTVNSAR	-0.28
SVTEQGAELSNEER	-0.28
AQYEDIANR	3.81

148肽 添加(A)...

- 单击“确定”按钮。

“保留时间”视图将发生变化并显示如下：



您刚刚使用 30 分钟梯度采集的数据校准了 147 个新的人类肽段的 iRT-C18 值。

使用 iRT 预定新的数据采集

接下来，您将探索 iRT 如何利用现有方法支持新的色谱条件设置，即使改变洗脱梯度，iRT 也能在仅进行一次校准后，提供相对较小的时间窗口开始预定采集。

如果您在自己的实验室执行此操作，则需要打开原始“iRT Standard.sky”文件，为其导出实验方法，然后在准备好的样品中加入标准肽段混合物，实现数据采集。本教程文件中夹包含了一个严格按照这种方法创建的原始数据文件。与上述同样的样品再次进行了质谱检测，但是采用不同的洗脱梯度（90 分钟）和一个新的色谱柱，不过这次仅仅分析标准肽段。

继续之前，请执行下列操作：

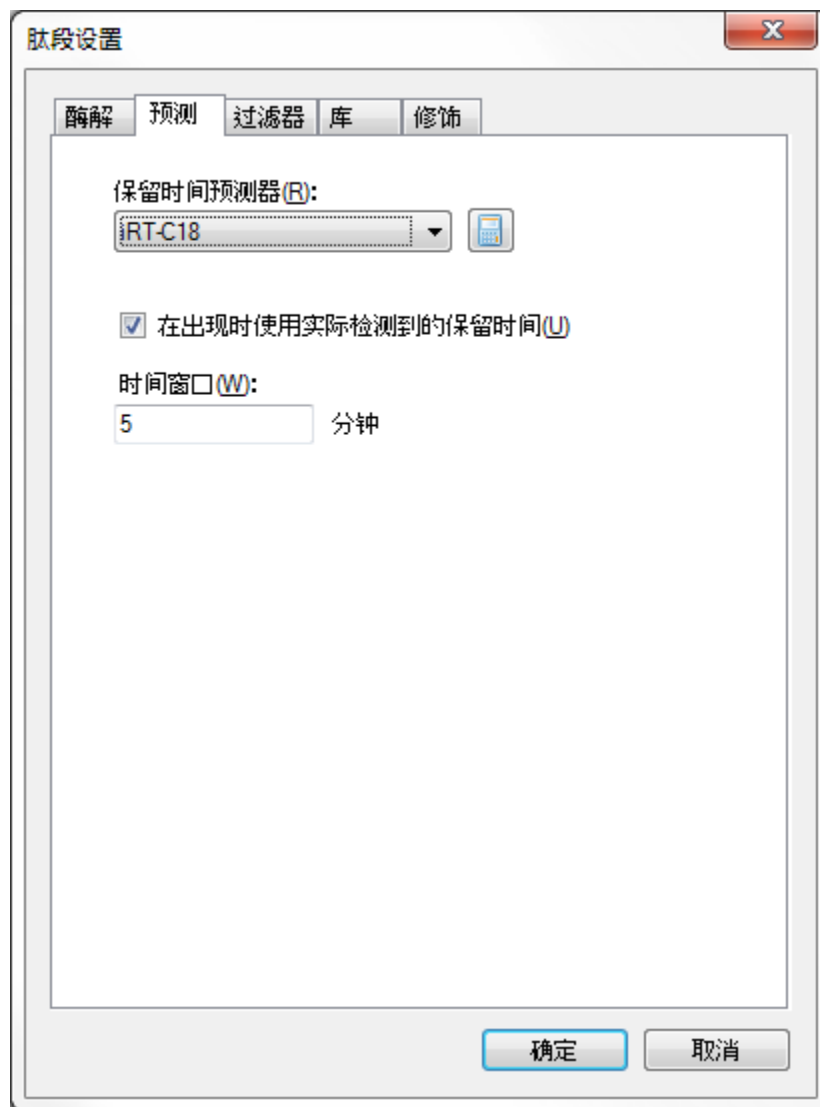
- 在“文件”菜单中，单击“保存”(Ctrl-S)。
- 关闭“查找结果”视图。

为了适应新的色谱柱和 90 分钟梯度，需要重新校准方法，执行下列操作：

- 在“文件”菜单，单击之前保存的“iRT Human+Standard.sky”文件 (Alt-F, 2)。
- 在“编辑”菜单中，单击“查找”(Ctrl-F)。
- 单击“隐藏高级”按钮。

- 在“查找项”字段，输入“NSAQ”。
- 单击“查找下一个”按钮。
- 单击“关闭”按钮。
- 按下 Delete 键删除您在另一个文档中已经删除的肽段。
- 在“设置”菜单上，单击“肽段设置”。
- 单击“预测”标签。
- 在“保留时间预测器”下拉列表中，选择您创建的“iRT-C18”预测器。
- 选中“在出现时使用实际检测到的保留时间”复选框。
- 在“时间窗口”字段中，输入“5”。

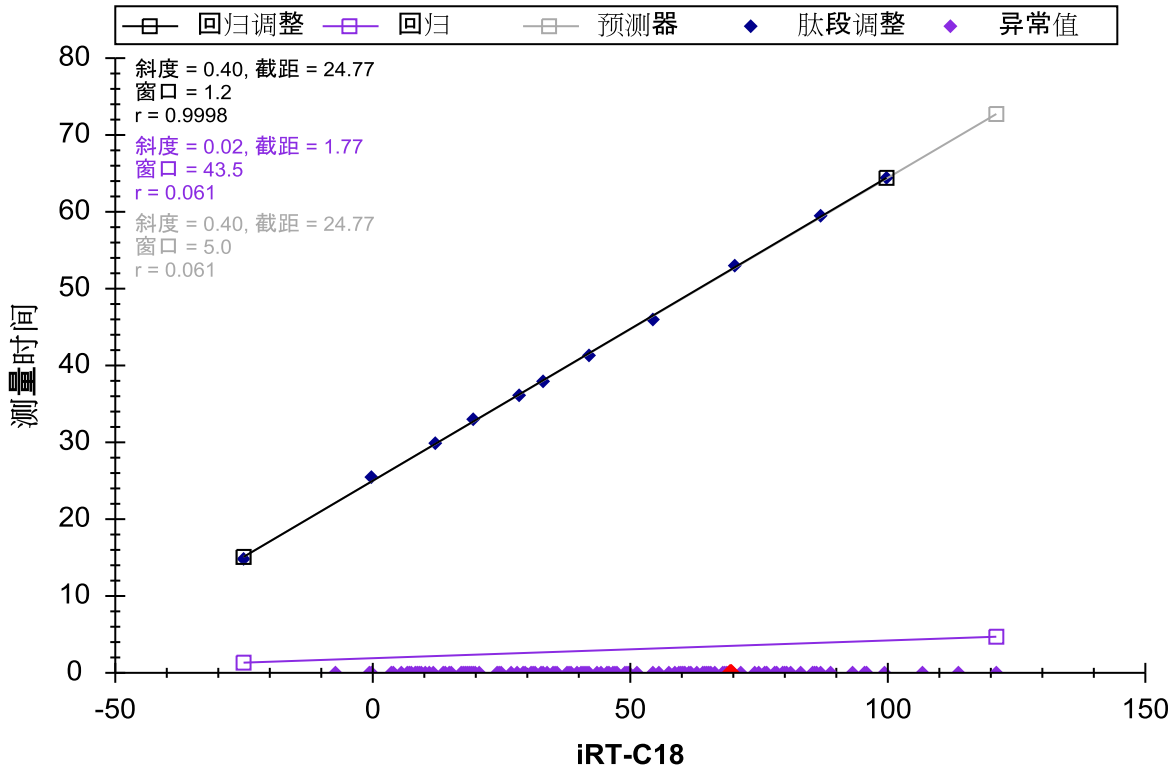
“肽段设置”表单将显示如下：



- 单击“确定”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。
- 双击您为本教程创建的 iRT 文件夹中的
“A_D110913_SiRT_HELA_11_nsMRM_150selected_90min-5-40_TRID2215_01.raw”文件。

完成数据导入后，“保留时间”回归图将显示如下：



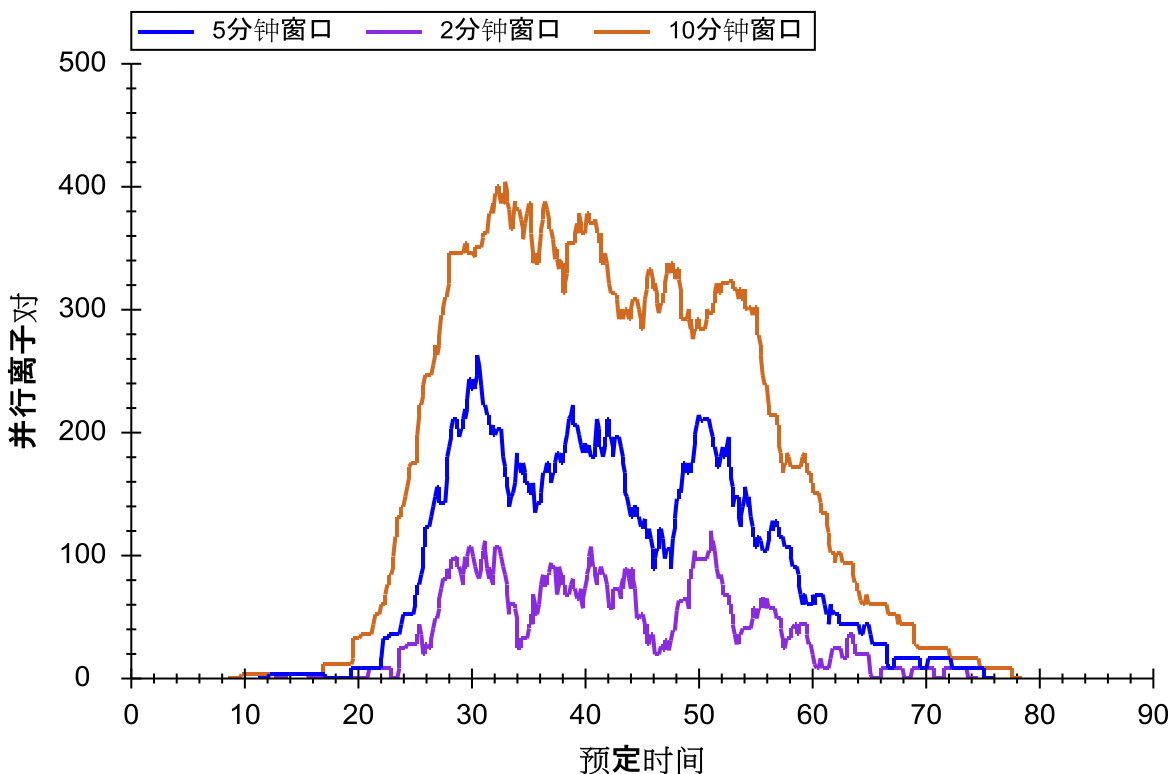
此时，您可看到标准肽段的洗脱时间范围约介于 15 至 65 分钟之间，此次运行中未对额外的目标肽段进行检测。然而，此时您已准备好在新梯度上对肽段进行检测。

使用 iRT 预测器导出预定方法

创建预定方法之前，您可以通过执行以下操作，更好地了解在某种预定参数下离子对是如何被检测：

- 在“视图”菜单，选择“保留时间”并单击“时序安排”。

“保留时间”视图将发生变化，显示类似于下图的图形：



如果您未能在上图中看到所有的三条线，则可执行下列操作：

- 右键单击图形，并单击“属性”。
- 在“时序安排图形属性”的“时窗”字段，输入“2, 5, 10”。
- 单击“确定”按钮。

从此图可以看到窗口尺寸对于时序安排的影响。时间窗口越小，则在给定时间内允许同时测定的离子对越少，这使得能够在特定驻留时间内，对单次质谱进样实验中的更多离子对进行测定。窗口尺寸大小对于能够捕获到整个色谱峰的肽段百分比，可以通过以下公式进行近似估计：

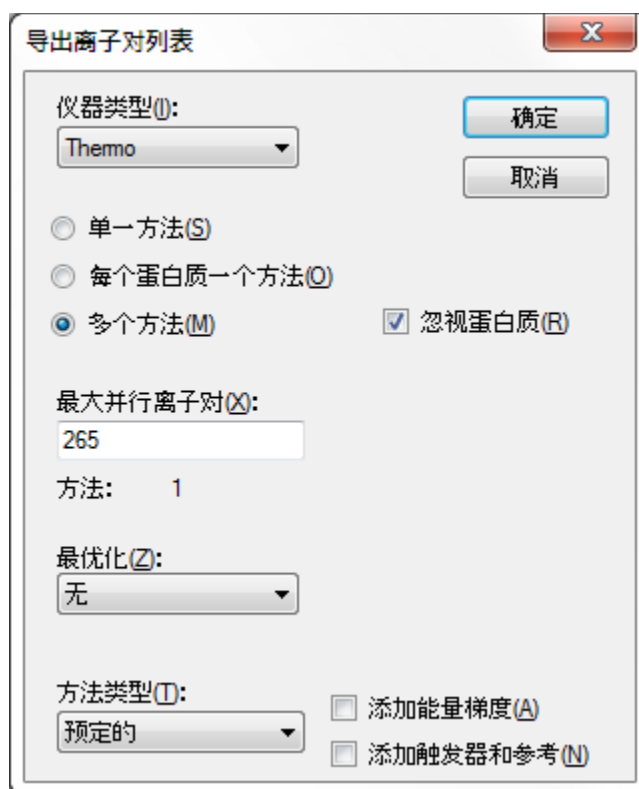
$$\begin{aligned} \text{窗口尺寸} = & \text{平均(色谱峰宽度)} + \text{平均(预测误差)} \\ & + z * ((SD(RT) + SD(\text{预测误差})) * 2 + SD(\text{色谱峰宽度})) \end{aligned}$$

其中“z”为在正态分布条件下，一定百分比范围覆盖时的标准差临界值，如 $z(95\%) = 1.96$ 。在拥有完美预测并且不考虑峰宽和保留时间偏差的条件下，所需窗口的大小正好为色谱峰宽。即使能够进行完美预测，峰宽以及保留时间的偏差也会扩大所需窗口尺寸。最后，预测误差也会进一步扩大窗口尺寸。值得注意的是，即使采用最先进的保留时间预测方法，在预定采集之前仅仅进行一次未预定实验也是不够的。因为，您试图预测平均保留时间，但是在一次实验中，大约 5% 的肽段会偏离均值 2 被标准差之外。

本教程中，iRT 方法允许在 90 分钟的洗脱梯度运行中，采用 5 分钟的时间窗口对目标肽进行检测。上图表示在一个检测周期中可以检测不超过 265 离子对。在 10 毫秒的驻留时间条件下，最多将产生 2.6 秒的检测周期。为了创建一个能够对本文中涉及的 1223 个离子对进行预定检测的方法，请执行下列步骤：

- 在“文件”菜单中，选择“导出”并单击“离子对列表”。
- 在“方法类型”字段中，选择下拉列表中的“预定的”。
- 在“最大并行离子对”字段中，输入“265”。

“导出离子对列表”表单将显示如下：



您可以简单地选择“单一方法”选项，但表单实际显示出“方法：1”，确认了在 5 分钟检测窗口以及同时测定的离子对数量不会超过 265 的条件下，所有的离子对能够在一次进样实验中被检测。对于定量分析而言，这个数量可能仍然稍微偏高，但却好于只能检测一半离子对而采用 2 次进样，各检测 335 个离子对的方法要好。您还可以根据自己的喜好将此数字降至 135，此时您将看到 Skyline 显示实现这个目标需要进行两次进样。如果需要将此数字降至 90，则显示需要进行 3 次进样。但是，继续下一步之前，需确保将此数字恢复为 265。

- 单击“确定”按钮。
- 在“文件名”字段，输入“iRT Human+Standard”。
- 单击“保存”按钮。

在 Windows Explorer 中，您将发现此操作将在 iRT 文件夹中创建“iRT Human+Standard_0001.csv”文件。在 Excel 中，您将确认此文件包含全部 1223 个离子对，预定起始和终止时间间隔为 5 分钟。

审查预定数据

要审查采用刚刚创建的方法采集的数据，首先通过执行下列操作删除 90 分钟梯度校准数据：

- 在“编辑”菜单上，单击“管理结果”(Ctrl-R)。
- 单击“删除所有”按钮。
- 单击“确定”按钮。

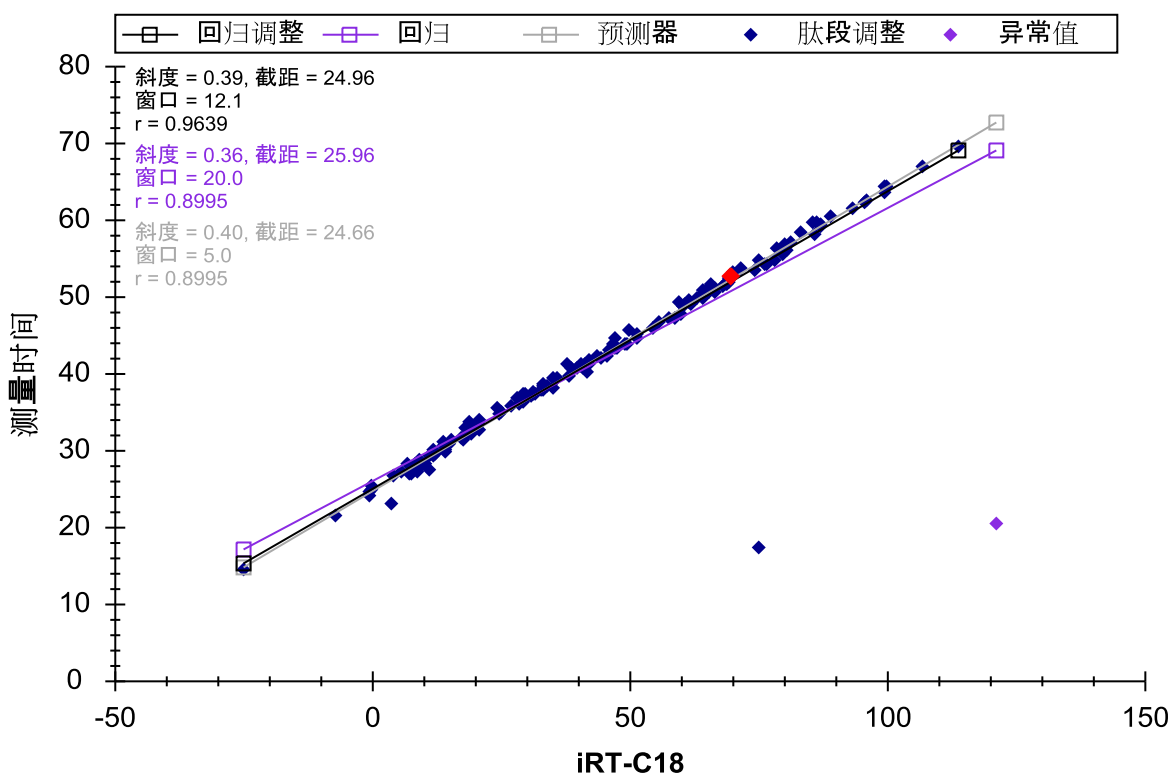
此时，执行下列操作导入预定方法采集的数据：

- 在“文件”菜单中，选择“导入”并单击“结果”。
- 单击“导入结果”表单上的“确定”按钮。
- 双击本教程 iRT 文件夹中的“A_D110913_SiRT_HELA_11_sMRM_150selected_90min-5-40_SIMPLE.raw”文件。

加载数据时，您可通过执行下列操作将“保留时间”视图切换回“线性回归”：

- 在“视图”菜单中，选择“保留时间”并单击“线性回归”(Shift-F8)。

数据加载完成后，“保留时间”视图将显示如下：

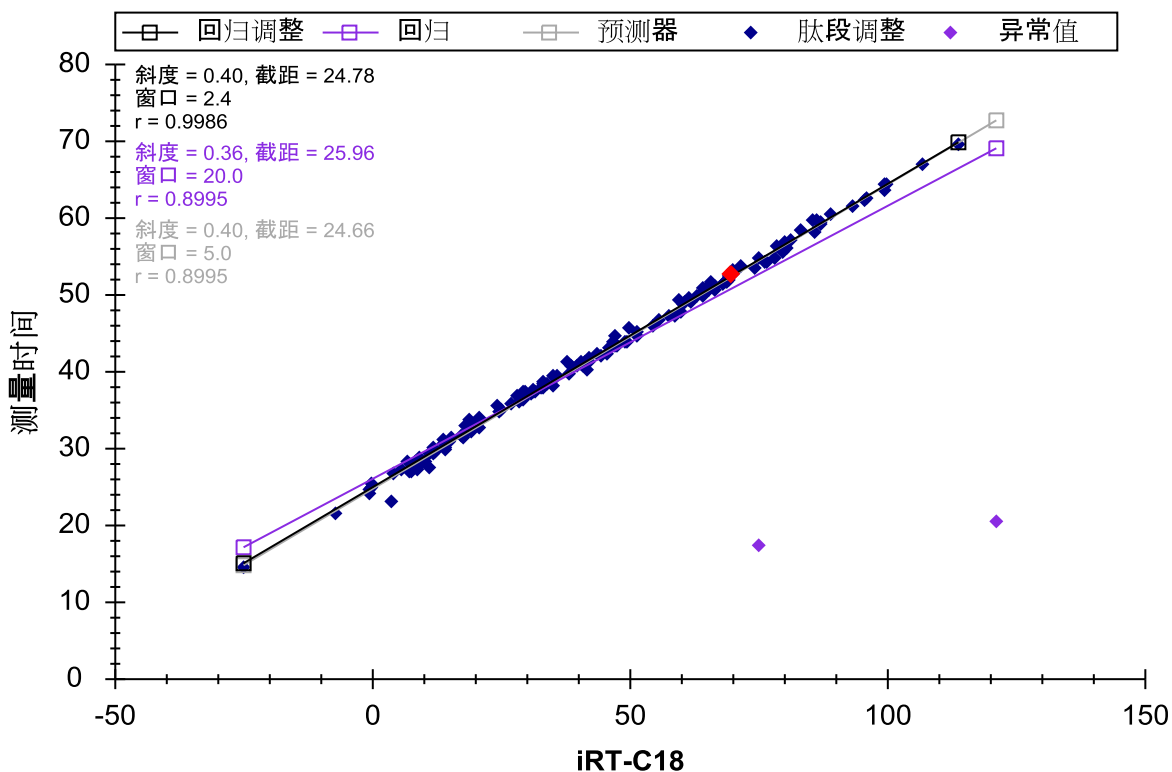


从此图形中，显而易见有 2 个异常肽段。可能的原因有当前数据中的色谱峰积分边界错误，或用于计算 iRT 值的校准数据中的色谱峰积分边界错误。本案例中，问题出现在 iRT 在 30 分钟梯度校准期间，Skyline 自动选择了错误的色谱峰。实际上，您正在查看的数据并非通过上述生成的预定方法所收集，了解这一点非常重要。如果是根据上述生成预定方法收集数据，则异常肽段的色谱图甚至不会包括此处检测到的几个峰。此数据是在基于更加彻底地审查校准数据（本教程跳过了该步）创建的预定方法收集的。

如果您想知道图例中的“异常值”为何只有一个点显示为紫色（这是因为所设置的相关系数阈值尚不适用于相关性如此高的计算器），您可执行下列操作，更改相关性阈值：

- 右键单击“保留时间”图形，并单击“设置阈值”。
- 在“阈值”字段，输入“0.998”。
- 单击“确定”按钮。

此时，“保留时间”图形将显示如下：



现在您可以单击每个异常点，Skyline 将在肽段视图中对其进行显示。然后，按下 Esc 恢复主窗口，并按下 Ctrl-C 复制肽段标记。您也可将这些内容收集到单独的编辑器中，供以后审查使用，或者为之前创建的“iRT Human+Standard Calibrate.sky”文件打开另一个 Skyline 实例。然后，您可使用“查找”表单审查这两个肽：

EVVEEAENGR

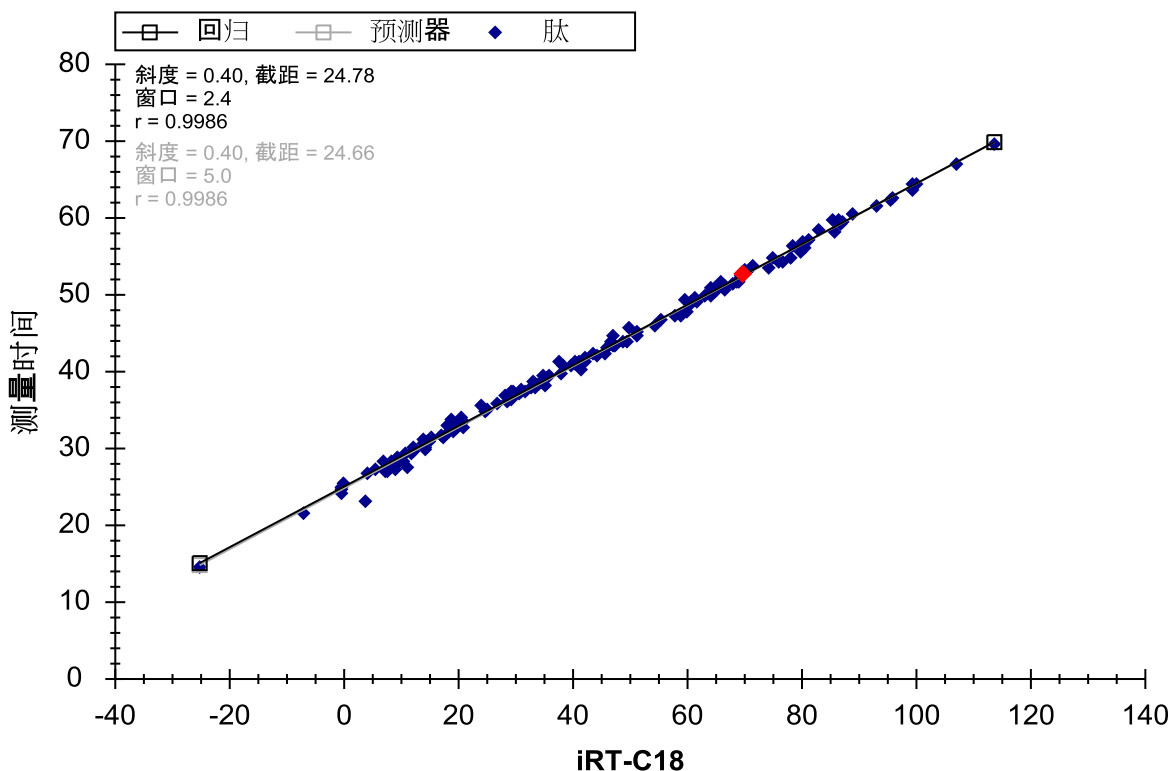
LLADQAEAR

对这两个肽段，你会发现很难在校准数据中看到匹配很好的色谱峰。如果您在校准实验时仔细审查，很有可能会发现这些异常肽段，从而进行删除。这也就解释了为何在校准实验中，您应该尽可能小心。

现在您可以根据这些更准确的数据，重新计算本文档中所有肽段的 iRT 值，这些数据通过使用同位素标记肽段，确保了色谱峰选择的正确性。重复上述校准步骤，如有要求，选择“替换现有值”。然而在本教程中，您可以直接执行下列步骤除去错误校准的肽段：

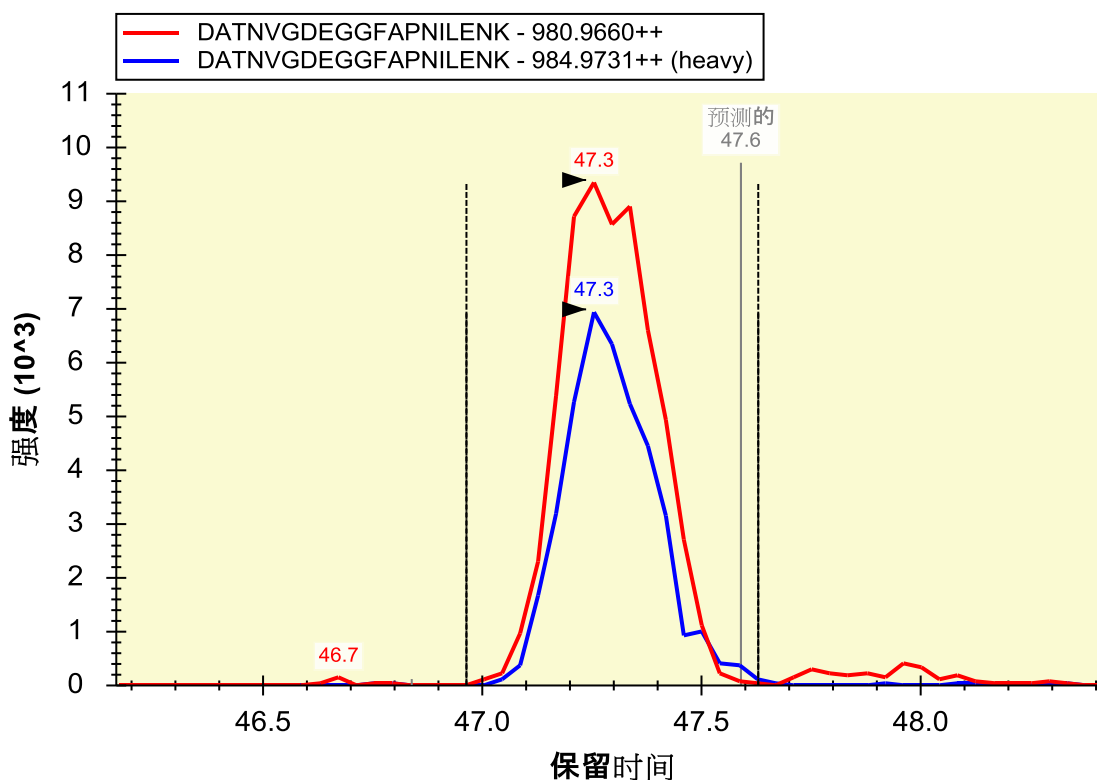
- 右键单击“保留时间”图形，单击“删除异常值”。

删除图形中的 2 个异常值，肽数量应减去 2，降至 156。



请注意，上图中名为“预测器”的线性方程是 Skyline 使用本文档中标准肽段的测定时间和 iRT 值进行回归，自动计算出的。这是因为选中了“编辑保留时间预测器”表单中的“自动计算回归”复选框。

现在单击肽段视图，并使用向下箭头键审查肽段色谱图。 Skyline 将显示如下图形：



您将看到除标准肽段以外的所有肽段都具有轻母离子对和重母离子对，并且每个色谱峰通常比未预定数据具有更多的点。您还将看到 Skyline 在标记“预测的”下方显示了色谱峰的预测时间。

本案例中，数据源自单一进样。但是，因为设置了“自动计算回归”，Skyline 将为每次进样（甚至包括那些需要多次预定进样才能测定所有肽段的文档）分别计算回归，这样做可以让定量更加精确。在为此类文档导出方法时，Skyline 将在每个方法中自动包含标准肽段的离子对。自动回归为每次进样计算一个单独的回归方程，比对所有进样采用一个回归方程更加精确，这也使“预测”注释成为更加强大的肽段身份验证工具。

根据 MS/MS 谱图采集时间计算 iRT 值

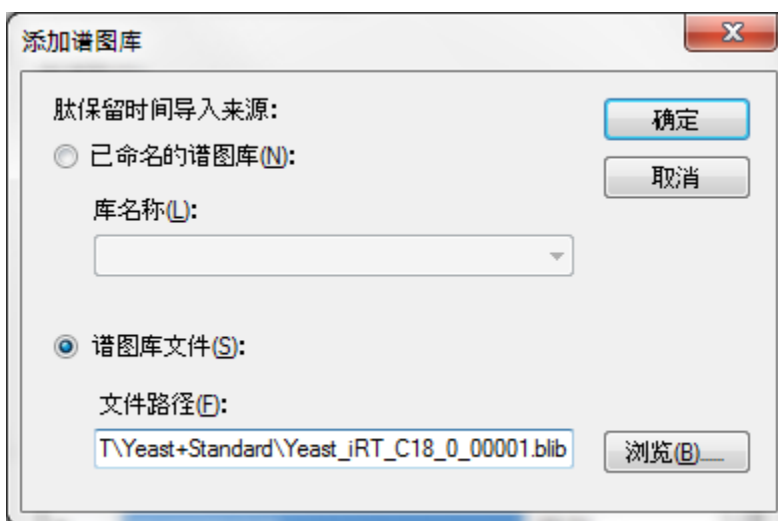
如果您收集的是数据依赖实验（DDA）的数据，其中也参入了足够高浓度的用于构建 iRT 计算器的标准肽（因而这些标准肽采样密度足够高以及被可靠鉴定到），则可以采用和处理 SRM 数据相同的方法计算 iRT 值。一般来说，这些 iRT 值精确度较低，因为是基于肽段色谱洗脱峰上随机采样获得的 MS/MS 扫描时间，并不是真正的肽段保留时间。然而，如果直接采用探索性数据分析实验的 iRT 数值指导构建 SRM 预定实验，可以节省大量的仪器时间。在数据无关（DIA）分析中，由于所有肽段的离子全部碎裂产生 MS/MS 图谱，MS/MS 图谱的采样频率足够高，采用 MS/MS 图谱中的碎片离子信号来构建离子流色谱峰，利用离子流色谱峰的最大值来建立 iRT 模型也可以得到十分精确的结果。

在本教程的 iRT 文件夹中，您将发现包含谱图库“Yeast_iRT_C18_0_00001.blib”的子文件夹“Yeast+Standard”。此谱图库来自酵母蛋白质酶切肽段混合物的两次 DDA 实验，采用了 SEQUEST 搜库，添加了 Biognosys 标准肽段混合物。如同下面所示，只要 iRT 数据库中的肽段足够多，导入的数据中不一定需要包括标准肽段。然而，您必须使用 Skyline 建立的 BiblioSpec 图谱库格式。因为其它图谱库格式中，不会为每个肽段分别存储每次进样实验的保留时间，从而导致无法对同样色谱条件下的不同实验进行回归。

您可通过执行下列操作，以在此库中添加肽段谱图匹配（PSM）的 iRT 值：

- 右键单击“保留时间”图，选择“计算器”并单击“编辑当前”。
- 单击“编辑 iRT 计算器”表单中的“添加”按钮，并单击“添加谱图库”。
- 单击“添加谱图库”表单中的“谱图库文件”选项。
- 单击“浏览”按钮。
- 双击“iRT”文件夹中的“Yeast+Standard”子文件夹。
- 双击“Yeast_iRT_C18_0_00001.blib”文件。

“添加谱图库”表单将显示如下：



- 单击“确定”按钮。

Skyline 将显示如下表单：



该表单显示 Skyline 可以为库中的两次 DDA LC-Run 计算有效回归。并且利用这些回归模型为 558 个新肽段计算了 iRT 值。当然，每次 LC-Run 的 iRT 值是分别计算的。在两次 LC-Run 中同时出现的肽段将产生两个 iRT 值，最终结果为这两个值的平均值。Skyline 还发现有 3 个肽段的 iRT 值已经根据色谱峰获得，因此将跳过。

- 单击“确定”按钮。

此时，“编辑 iRT 计算器”表单将显示“经测量的肽”列表中有 706 个肽。

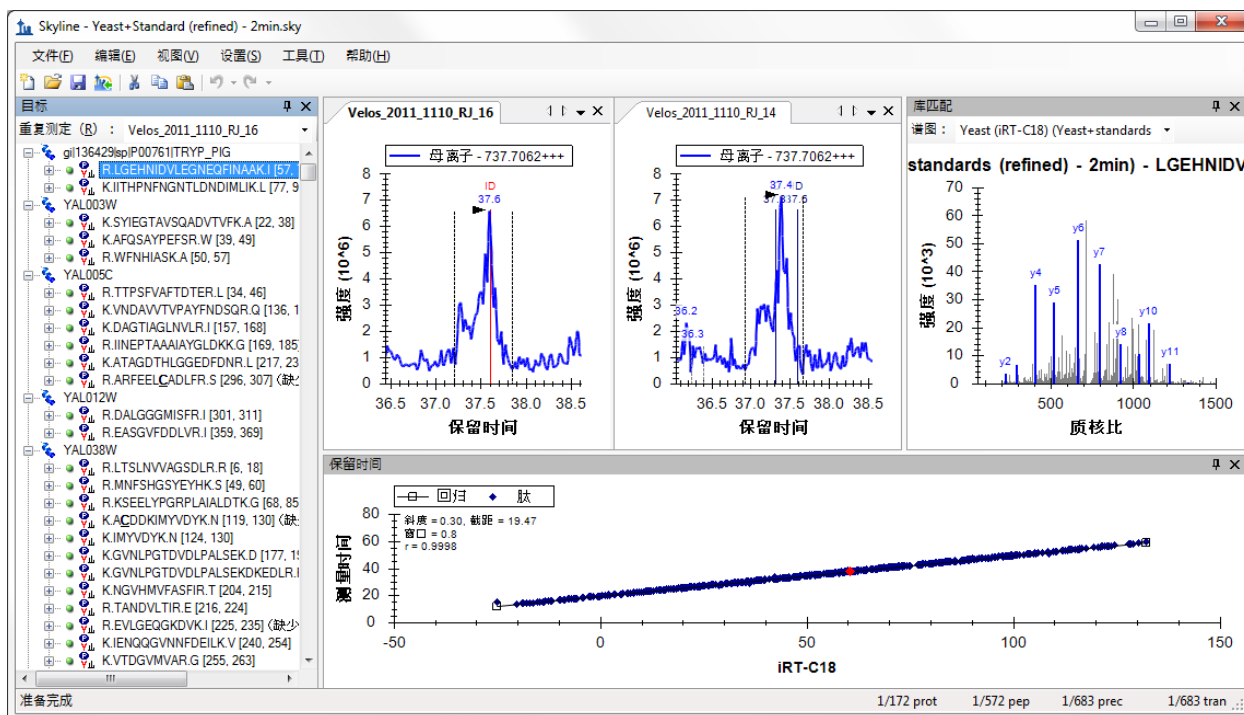
- 单击“确定”按钮。

将 MS/MS 扫描时间转换为色谱峰时间

您可以使用刚刚根据 MS/MS 扫描时间计算所得的 iRT 值来预定 SRM 采集，然后使用 SRM 数据根据色谱峰时间获取更加精确的 iRT 值。但是，您也可以使用 Skyline MS1 过滤，直接从原始 DDA LC-Run 提取色谱峰时间。可在 [MS1 过滤](#) 教程中查看如何设置并将数据导入 MS1 过滤文档的详细信息。在本文档中，您可通过执行下列操作快速浏览已创建并导入数据的文档，包括用于创建谱图库的两次 DDA LC-Run:

- 在“文件”菜单中，单击“打开”(Ctrl-O)。
- 导航至“iRT”文件夹中的“Yeast+Standard”子文件夹。
- 双击文件“Yeast+Standard (refined) - 2min.sky”。
- 选择肽段视图中的第一个肽段。

Skyline 窗口将显示如下：



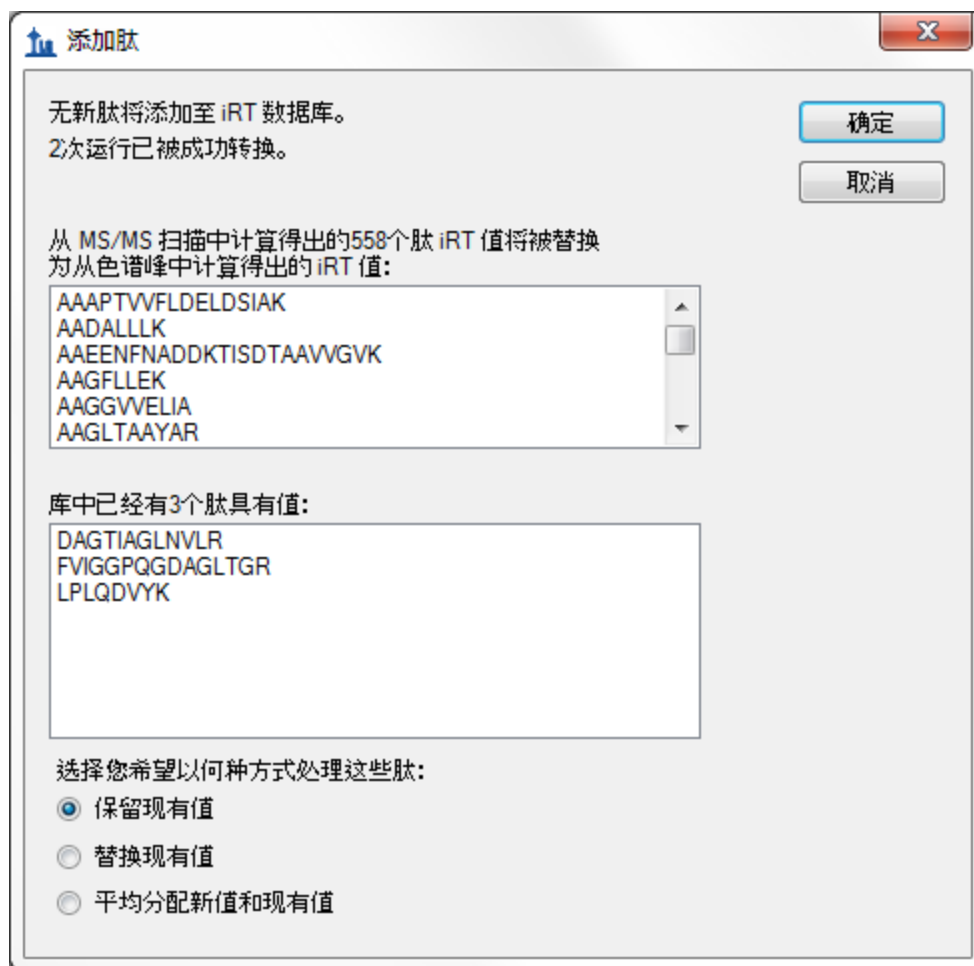
您可以双击“保留时间”视图中的标题栏，以更好地查看该图形，并了解到库谱图方法计算的 iRT 值与检测时间的相关系数为 0.9998。因此相对于采用 MS/MS 扫描时间获得模型来说，使用色谱峰获得的模型精度可能低于预期。但另一方面需要说明的是，该数据进行了手工调整，仅仅采用了那些在两次 LC-Run 中都出现的，并且具有很好色谱峰的肽段。从这里可以看出，使用谱图库数据计算肽段的初始 iRT 值时，您可能还需要附加一些检验条件，以筛选出更加可靠的数据。

您在此文件中看到的色谱图提取自 LC-Run 中的 MS1 扫描。您还可以查看记录以确认 MS/MS 扫描的时间，这些时间在色谱图中利用“ID”进行标记。此外，您还可以在 [MS1 过滤](#) 教程中学习如何使用这一十分有用的数据处理方法。

要将使用 MS/MS 扫描时间计算所得的 iRT 值转化为本档中使用色谱峰时间的值，请执行下列步骤：

- 右键单击“保留时间”图，选择“计算器”并单击“编辑当前”。
- 单击“编辑 iRT 计算器”表单中的“添加”按钮，并单击“添加结果”。

Skyline 将显示如下所示表单：



这样做，系统将替换您在上一部分添加的 558 个 iRT 值。由于正在使用色谱峰时间，您还可以选择更换酵母和人类样品共享的 3 个肽段，或使用两者的平均值。

- 单击“确定”按钮接受更改。

更多 iRT 计算器编辑选项

现在您拥有 705 个具有相当良好 iRT 值的肽段，虽然这些值仅基于不超过两次重复实验计算得出。在这些例子中，您已使用了包含 iRT-C18 计算器时采用的标准肽段混合物的数据集。然而实际中并不一定需要这样。现在，您可以利用与现有 iRT 数据库具有足够多共享肽段的数据集来计算新的 iRT 值。只要存在至少 20 个共同肽段，Skyline 就将使用所有这些共同肽段来构建相关性为 0.99 或更高的回归模型。

测试 Skyline 中的 iRT 功能时，利用了公共数据 PeptideAtlas³ 中的图谱库来创建 Skyline 文档。此数据集包含 20 多次重复检测，产生了 1000 多个 iRT 值，但由于显然太大而无法涵盖在本教程中。

在使用“编辑 iRT 计算器”表单的“添加”按钮时，您可能还注意有一个“添加 iRT 数据库”的操作。此菜单项可用于将现有 iRT 数据库合并至当前计算器。如果数据库使用相同的标准肽段，则它们将用于完成从一个数据库转换到另一个数据库的回归。否则，与其他数据源一样，只要存在至少 20 个共同肽段，Skyline 就将使用所有这些共同肽段来构建相关性为 0.99 或更高的回归模型

“编辑 iRT 计算器”表单中的“打开”按钮允许您使用现有的 iRT 数据库，该数据库可能是您从他人那里接收的数据库。

您也可使用“编辑 iRT 计算器”表单中的“肽段”按钮，将标准肽更改为数据库中包含的任何肽段集合，并且您可使用“重新校准”按钮更改 iRT 标尺。

结论

在本教程中，您学习了如何使用 Skyline 的 iRT 功能，这是存储肽段保留时间实际检测值的标准方法，以便用于 SRM 采集的实验设计和采集后的肽段身份验证。只要您为测定的肽段存储了 iRT 值，则通常只需要进行一次校准即可预定任意数量的离子对。更加精确的保留时间预测也使 iRT 预测器成为一种比基于序列的预测更强大的肽段鉴定结果验证工具。Skyline 的 iRT 方法易于使用，且易于生成 iRT 值。您可根据任何标尺和任何标准肽段集计算 iRT 值。您甚至可以将内生于特定实验的肽段集合作为标准，只要它们可进行持续检测并跨越您试图测定的大多数色谱梯度范围。此外，数据库具有共同肽段时，Skyline 能够很方便的完成 iRT 数据库合并。您还学习了关于 iRT-C18 的内容，这是最初 Biognosys 团队使用 Biognosys 试剂盒定义的 iRT 标尺。您可以将此试剂盒用于自己的实验，Skyline 使得 iRT-C18 标尺易于使用，但可以将标准肽段更改为根据该标尺校准的任何肽段集合，正如此时您已完成对数百个常用人类和酵母肽段的操作。

参考文献

1. Krokhin, O. V. *et al.* An improved model for prediction of retention times of tryptic peptides in ion pair reversed-phase HPLC: its application to protein peptide mapping by off-line HPLC-MALDI MS. *Mol. Cell Proteomics* **3**, 908-919 (2004).
2. Escher, C. *et al.* Using iRT, a normalized retention time for more targeted measurement of peptides. *Proteomics (accepted)* (2012).
3. Deutsch, E. W., Lam, H. & Aebersold, R. PeptideAtlas: a resource for target selection for emerging targeted proteomics workflows. *EMBO Rep* **9**, 429-434 (2008).