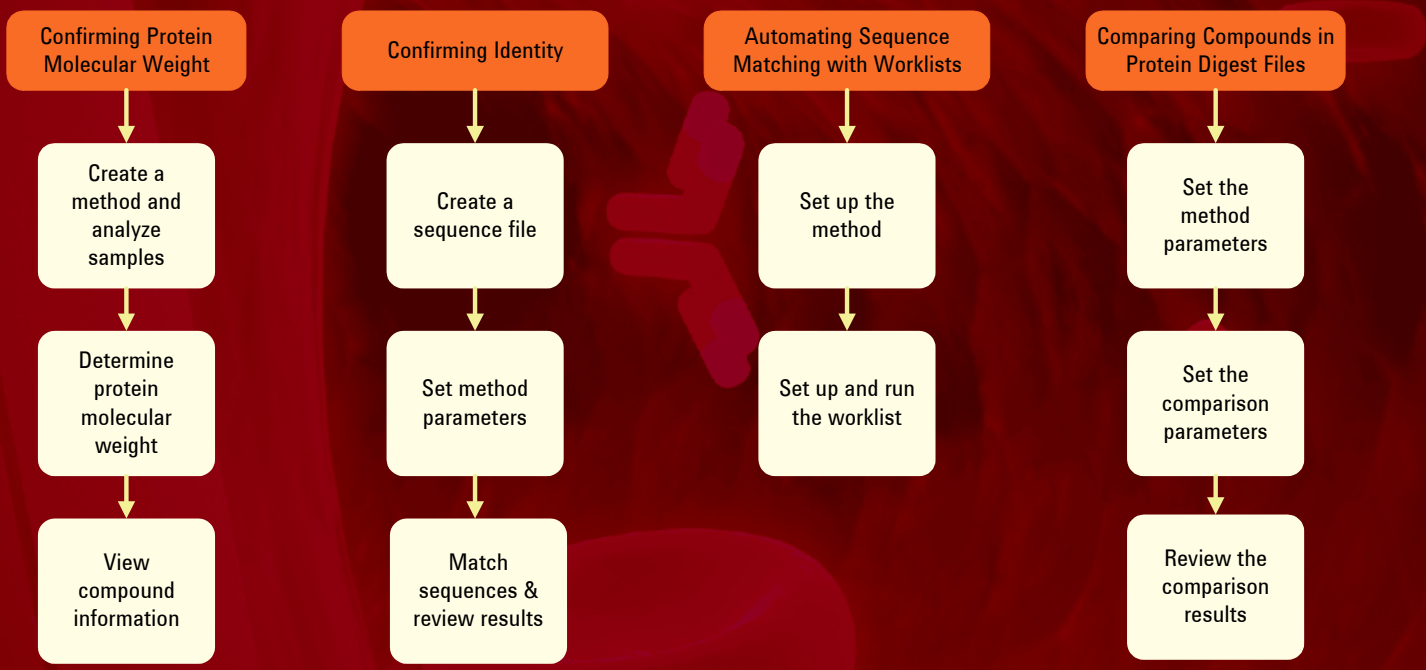
Agilent MassHunter BioPharma

Workflow Guide



Agilent Technologies

Notices

Manual Part Number

5990-7064EN

Edition

Revision B, January 2016

Printed in USA

Agilent Technologies, Inc.
5301 Stevens Creek Blvd.
Santa Clara, CA 95051

Acknowledgments

Warranty

The material contained in this document is provided "as is," and is subject to being changed, without notice, in future editions. Further, to the maximum extent permitted by applicable law, Agilent disclaims all warranties, either express or implied, with regard to this manual and any information contained herein, including but not limited to the implied warranties of merchantability and fitness for a particular purpose. Agilent shall not be liable for errors or for incidental or consequential damages in connection with the furnishing, use, or performance of this document or of any information contained herein. Should Agilent and the user have a separate written agreement with warranty terms covering the material in this document that conflict with these terms, the warranty terms in the separate agreement shall control.

Technology Licenses

The hardware and/or software described in this document are furnished under a license and may be used or copied only in accordance with the terms of such license.

Restricted Rights

If software is for use in the performance of a U.S. Government prime contract or subcontract, Software is delivered and licensed as "Commercial computer software" as defined in DFAR 252.227-7014 (June 1995), or as a "commercial item" as defined in FAR 2.101(a) or as "Restricted computer software" as defined in FAR 52.227-19 (June 1987) or any equivalent agency regulation or contract clause. Use, duplication or disclosure of Software is subject to Agilent Technologies' standard commercial license terms, and non-DOD Departments and Agencies of the U.S. Government will receive no greater than Restricted Rights as defined in FAR 52.227-19(c)(1-2) (June 1987). U.S. Government users will receive no greater than Limited Rights as defined in FAR 52.227-14 (June 1987) or DFAR 252.227-7015 (b)(2) (November 1995), as applicable in any technical data.

Safety Notices

CAUTION

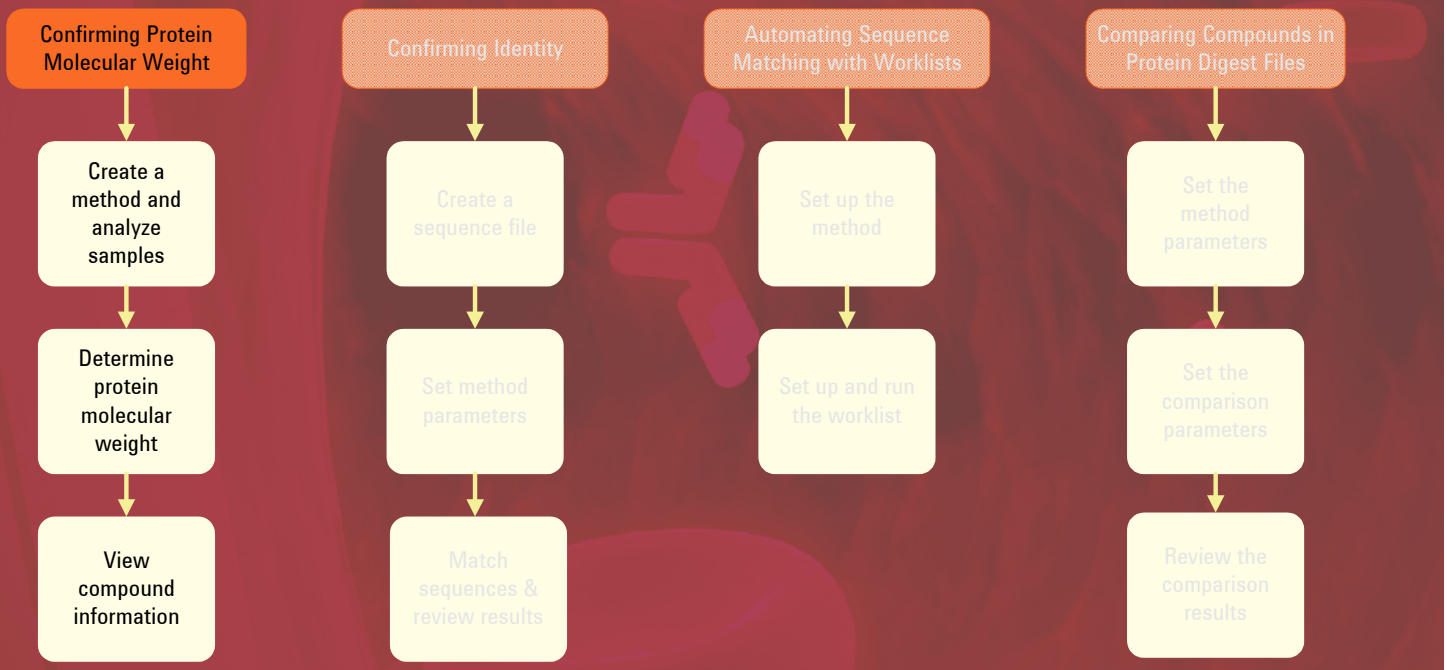
A **CAUTION** notice denotes a hazard. It calls attention to an operaDraft 2 Jan 2016ting procedure, practice, or the like that, if not correctly performed or adhered to, could result in damage to the product or loss of important data. Do not proceed beyond a **CAUTION** notice until the indicated conditions are fully understood and met.

WARNING

A WARNING notice denotes a hazard. It calls attention to an operating procedure, practice, or the like that, if not correctly performed or adhered to, could result in personal injury or death. Do not proceed beyond a WARNING notice until the indicated conditions are fully understood and met.

Contents

Confirming Protein Molecular Weight



Agilent Technologies

Before you begin

This workflow shows you how to confirm protein molecular weights.

Required products:

- LC/MS instrument, either 6200 Series Accurate-Mass TOF LC/MS or 6500 Series Accurate-Mass Q-TOF
- MassHunter Data Acquisition B.06.01 software
- MassHunter Qualitative Analysis B.07.00 software
- MassHunter BioConfirm B.07.00 program

Example files:

For use in the exercises in this chapter, copy the **myoblobin.d** data file from the Qualitative Analysis setup disk to your computer hard drive.

For more information:

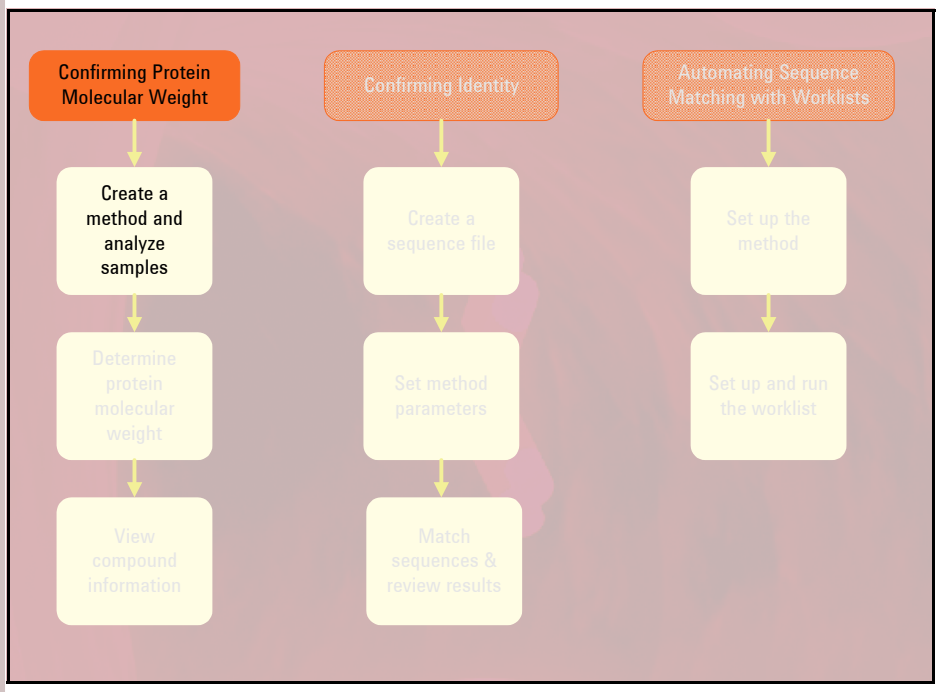
Use the following additional resources for more information on the steps in this workflow:

- *TOF/QTOF Quick Start Guide*
- *TOF/QTOF Familiarization Guide*
- *MassHunter LC/MS Data Acquisition online Help*
- *MassHunter Qualitative Analysis online Help*
- *MassHunter BioConfirm Quick Start Guide*

To access online Help, press **F1** or click **Contents**, **Index**, or **Search** on the Help menu.

Create an acquisition method and analyze samples

This section tells you how to create an acquisition method and analyze samples.



1. Create an acquisition method for your samples.

a Open the MassHunter Workstation Software LC/MS Data Acquisition program.

b (*optional*) To use an existing method as a starting point for a new method, click **Method > Open** and select the method of interest.

c Set LC method set points.

d Set MS method set points:

- On the **General** tab, set the Ion Polarity and other set points for MS data collection.
- For **Data Storage**, use either **Profile** mode or **Both** (Centroid and Profile data), since Maximum Entropy deconvolution requires profile spectra.
- On the **Source** tab, set the MS source set points, such as Gas Temp, VCap, and Fragmentor.
- On the **MS TOF** or **Q-TOF Acquisition** tab, set the Mass Range and Acquisition Rate/Time set points.

See "Tips for setting MS parameters for proteins" on the next page.

e Click **Method > Save As** and save the method with a new name.

Tips for setting MS parameters for proteins

Setting VCap voltage:

- To improve sensitivity for larger proteins, a value of 4500 to 5500 is recommended.

Setting Fragmentor voltage:

- For large molecules such as proteins, an increased fragmentor value will improve transmission and thus sensitivity and will also improve declustering of water molecules.
- For smaller proteins (< 100kDa), a value of 175 to 225 V is recommended.
- For large proteins (> 100kDa) such as IgG, a value of 350 to 420 V is recommended.
- If in-source fragmentation is observed, lower the fragmentor value.

Setting mode on the Q-TOF Acquisition Tab:

- For Q-TOF, select **MS Mode** for proteins.

2. To run a single sample:

a In the Method Editor pane, click the **Sample** tab.

b Under **Sample**, type the name and position.

c Under **Data File**, type the **Name** and enter the **Path**.

d Enter the **Run** parameters.

e Click **Apply**.

f Start the run in either of the following ways:

- Click **Sample > Run**.
- Click the  button on the toolbar.

3. To run multiple samples:

a If necessary, click the Worklist button on the toolbar to display the Worklist pane.

b (*optional*) To use an existing worklist as a starting point for a new worklist, click **Worklist > Open** and select the worklist of interest.

c Click **Worklist > Worklist Run Parameters**.

d On the Worklist Run Parameters dialog box:

- Set the **Data File Path**.
- Verify that the other parameters are set properly.
- Click **OK**.

e To add samples to the worklist, click **Worklist > Add Multiple Samples**.

f Fill out the Add Multiple Samples dialog box.

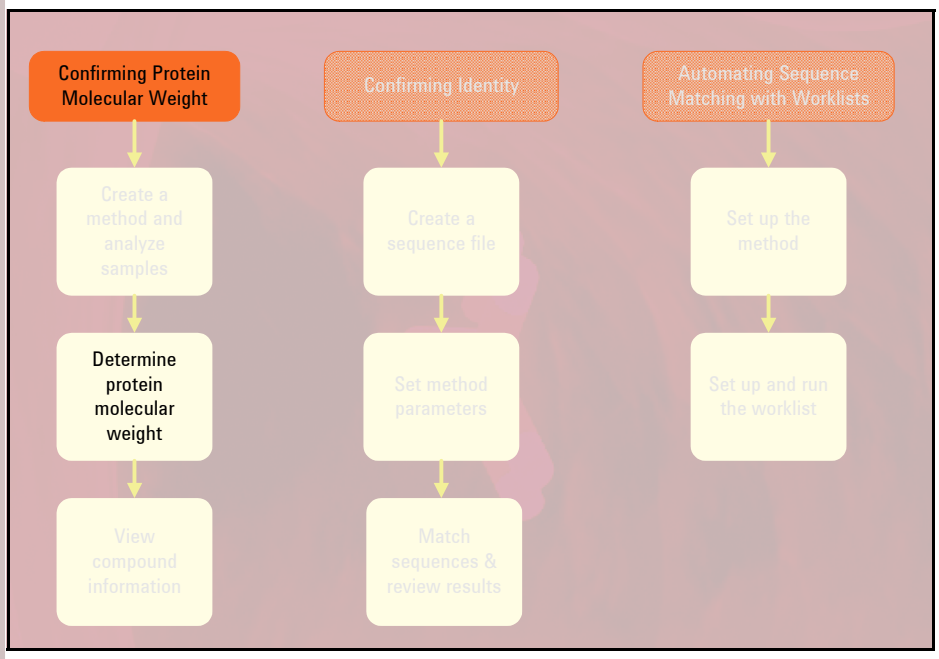
- Enter the paths for the method and data files.
- Specify the positions of the samples in the ALS or well-plate sampler on the **Sample Position** tab.
- Click **OK**.

g To start the run, click **Worklist > Run**.

Determine the protein molecular weight

This section tells you how to open a data file, integrate the chromatogram, extract spectra, deconvolute and view results. Deconvolution software does charge state deconvolution of mass spectra of large molecules with high charge states, such as proteins and large oligonucleotides.

You can use the **myoglobin.d** demo data file on the BioConfirm setup disk. Copy the data file from the Data folder on the BioConfirm setup disk to your computer hard drive.



1. Open the myoglobin.d data file in the MassHunter Qualitative Analysis software.

   a Open the MassHunter Workstation Software Qualitative Analysis program.

   b Click **File > Open Data File**.

   c Locate the data folder of interest, such as **myoglobin.d**, then click **Open**.

   The TIC is automatically displayed in the Chromatogram Results window.

2. Integrate and extract peak spectra.

   a Click **Method > Open**.

   b Select the appropriate default method, such as **BioConfirmIntactProtein-Default** for proteins such as myoglobin.

   c Use *either* of the following ways to integrate and extract peak spectra.
   - Right-click the TIC and select **Integrate and Extract Peak Spectra** from the shortcut menu, *or*
   - Click **Actions > Integrate and Extract Peak Spectra.**

   See Figure 1 on page 11.

3. Set the deconvolution range and mass step.

a Select **Deconvolute (MS): Protein** from the BioConfirm Workflow section of the Method Explorer to open the Deconvolute (MS) Method Editor section.

If the BioConfirm workflow is not available in Method Explorer, select **BioConfirm Intact Protein** from the **Configuration > Configure for Workflow** menu.

b Enter the following values on the **Deconvolution** tab of the Deconvolute (MS): Protein section of the Method Editor:
- 16000-18000 for **Mass range**
- 0.1 for **Mass step**

These values are appropriate for the myoglobin.d data file. You may need to set different values for your own data.

See Figure 2 on page 11.

4. Select the extracted MS peak and deconvolute the spectrum.

a Click the extracted MS peak spectrum in the MS Spectrum Results window.

b To start the deconvolution process, click the run button on the Method Editor toolbar.

Alternate method: To combine Steps 2 and 4 into one step, click **Chromatograms > Integrate and Deconvolute Peak Spectra**.

The results appear in the Deconvolution Results window.

See Figure 3 on page 11.

5. View peak information.

a Click the spectrum in the Deconvolution Results window to select it.

b Right-click the spectrum and select **View MS Spectrum Peak List 1** from the shortcut menu.

c Click twice on the **Height** column heading to sort results by spectral peak height. To see the column, you may need to scroll to the right or move the window divider bar. If the column is still not present, right-click the table and click **Add/Remove Columns**.

See Figure 4 on page 12.

d To close the peak list, click on the main toolbar.

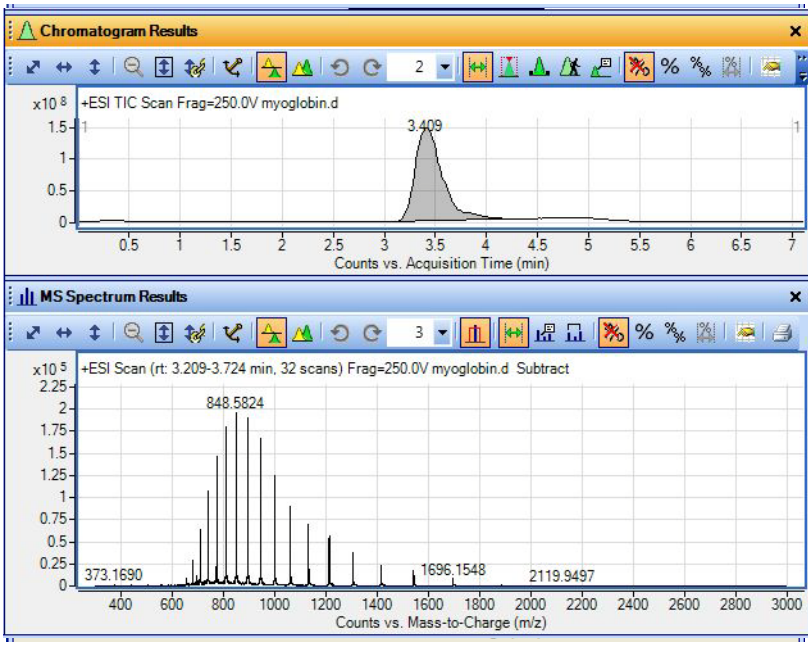


Figure 1 *Results of integration and spectra extraction for myoglobin.d*

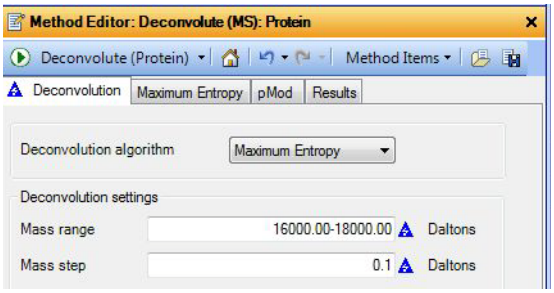


Figure 2 *Deconvolution parameters for myglobin.d*

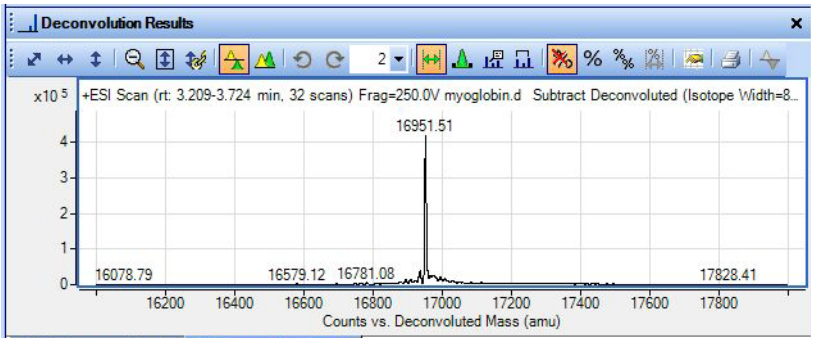


Figure 3 *Results of deconvolution for myglobin.d*

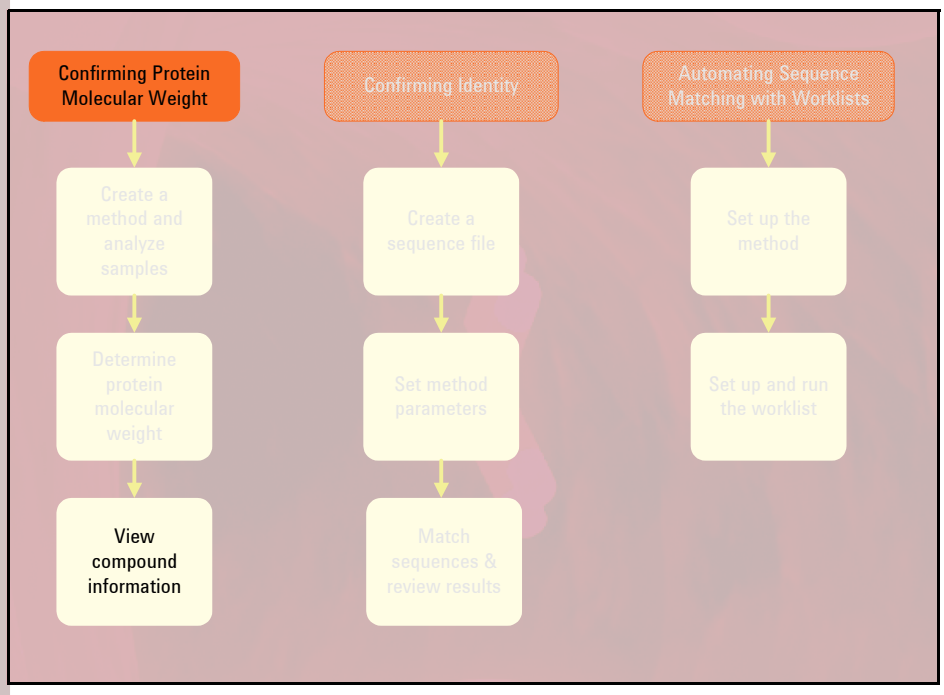
MS Peaks One: + Scan (rt: 3.209-3.724 min) Sub Deconv

Mass	Height	Height % (Norm)	Fit	m/z (prod.)	Z (prod.)	Ion	Mass (Start)	Mass (End)	Area	Peak Height (Max Er
16951.51	420791.07	0.01	10				16943.9	16959.1	24016307	420791.07
16934.65	37907.13	0	9				16926.6	16942.8	3084333	37907.13
16973.33	26429.23	0	8				16968.4	16987.7	4022554	26429.23
16965.78	23626.17	0	9				16960.9	16968.4	1670156	23626.17
16994.56	20926.96	0	9				16987.4	17000.9	2092438	20926.96
17006.64	16776.96	0	9				17000.9	17022.9	2566735	16776.96
16894.8	13612.81	0	8				16886.1	16901.6	1274045	13612.81
16908.89	13369.36	0	7				16901.1	16914.6	1283183	13369.36
17030.55	12115.51	0	7				17022.4	17044	1748294	12115.51
16920.6	12092.64	0	8				16914.6	16926.7	1184166	12092.64
16878.78	7665.62	0	6				16869.4	16887.1	964480	7665.62
17049.8	6891.84	0	9				17043.5	17055.6	720986	6891.84
17113.56	6600.45	0	3				17105.7	17123.2	779462	6600.45
16781.08	6036.82	0	8				16770.6	16791.9	567594	6036.82
17082.1	5925.5	0	7				17075.6	17090.3	721579	5925.5
17059.17	5678.58	0	8				17055.6	17067.9	634037	5678.58
17095.08	4924.12	0	7				17090.3	17106.8	650700	4924.12
17071.02	4828.62	0	5				17067.9	17075.6	355784	4828.62
16863.74	4649.39	0	4				16855.6	16870.2	544442	4649.39
17193.58	3974.26	0	9				17181.2	17203	644169	3974.26
17257.42	3541.09	0	7				17249.6	17263.7	402402	3541.09
17162.6	3456.8	0	7				17151.4	17171.1	584569	3456.8
17130.29	3429.78	0	7				17120.6	17141.2	645068	3429.78
16849.87	3335.9	0	6				16840.5	16855.6	444654	3335.9
17175.31	3126.96	0	7				17166.9	17184.9	494766	3126.96
17276.12	3023.79	0	8				17272	17285.1	350707	3023.79
17269.4	2976.1	0	8				17263.7	17272	226355	2976.1
16763.58	2947.08	0	9				16755.1	16773.7	353031	2947.08
17136.18	2940.08	0	8				17134.5	17151.4	497339	2940.08
17220.84	2747.39	0	7				17218	17224.6	172401	2747.39
17297.26	2692.27	0	6				17291.7	17304.2	314116	2692.27
17242.02	2677.02	0	7				17237.5	17249.6	290817	2677.02
17215.71	2671.56	0	7				17212	17218	155343	2671.56
17207.38	2670.8	0	6				17203	17212	225764	2670.8
17231.04	2611.12	0	8				17224.6	17237.5	323360	2611.12
17344.21	2596.34	0	8				17335.1	17349.5	322898	2596.34
17293.16	2561.13	0	3				17285.1	17297.6	308153	2561.13

Figure 4 *Peak information for the deconvoluted spectrum for myoglobin.d (partial list, sorted by Height)*

View compound information

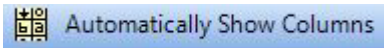
This section tells you how to view compound information for spectra deconvoluted in the previous section. The **myoglobin.d** data file is used as an example, but you can use your own data instead.



1. View the compound list.

The Compound List window should already be displayed. If it is not, click the  button on the main toolbar.

a On the Compound List toolbar, click the  button to Hide any currently empty columns.

b Then click the 

button to Automatically Show Columns.

See Figure 5 on page 14.

2. Select the mass of interest.

Click mass 16951.5 in the compound list for the **myoglobin.d** data file or an appropriate mass in your own data.

If linked navigation is turned on, associated data in the following windows are automatically displayed and selected:

- Deconvolution Results window
- A compound spectrum that displays all the charge states from the original *m/z* data for that specific protein mass in the MS Spectrum Results window.

3. Select the ion set spectrum of interest.

Select the ion set spectrum in the MS Spectrum Results window for the mass 16951.5.

4. View the charge states found for the protein along with their ppm error.

a Click the spectrum to select it.

b To open the MS Peak List 2 window, click the button on the main toolbar.

The following information is displayed for the ion set spectrum:
- Mass
- Abundance
- Charge state (Z)
- Diff (ppm)

See Figure 6 on page 15.

5. Switch from List mode to Overlay mode in the MS Spectrum Results window.

Click the button on the toolbar in the MS Spectrum Results window.

See Figure 7 on page 15.

6. Zoom in.

Right-drag to expand the area around the m/z 848.5 peak in the raw data for the myoglobin.d data file or an appropriate peak in your own data in the MS Spectrum Results window.

See Figure 8 on page 15.

7. Print a compound report.

a Display the Compound Report section in the Method Editor by selecting **Compound Report** from the BioConfirm Workflow section in the Method Explorer.

b Review the options in this section. Verify that the sections that you want to include in the report are marked.

c Display the Common Reporting Options section in the Method Editor by selecting **Common Reporting Options** from the BioConfirm Workflow section in the Method Explorer.

d Review the parameters in both the **Templates** and **Options** tabs.

e Click **Compound Report** from the File > Print menu to print the report.

Compound List

Automatically Show Columns

Show/Hide	Cpd	File	Name	Start	End	Mass	Fit Score	Min Z
☑	1	myoglobin.d	Compound 1	3.209	3.724	16951.505	10	
☑	2	myoglobin.d	Compound 2	3.209	3.724	16934.6536	9	
☑	3	myoglobin.d	Compound 3	3.209	3.724	16973.3253	8	
☑	4	myoglobin.d	Compound 4	3.209	3.724	16994.5588	9	
☑	5	myoglobin.d	Compound 5	3.209	3.724	17006.64	9	
☑	6	myoglobin.d	Compound 6	3.209	3.724	16894.8039	8	
☑	7	myoglobin.d	Compound 7	3.209	3.724	17049.8039	9	
☑	8	myoglobin.d	Compound 8	3.209	3.724	16781.0765	8	
☑	9	myoglobin.d	Compound 9	3.209	3.724	16763.5782	9	
☑	10	myoglobin.d	Compound 10	3.209	3.724	16748.1799	8	

Figure 5 *Left half of Compound List window for myoglobin.d*

MS Peaks Two: + Scan

Peak	Cpd Mass	Calc Mass	Diff (ppm)	Diff (Da)	m/z	m/z (Calc)	Diff (m/z)	Z	Abund	Abund %
1	16951.51	16951.65	-8.29	0.14	628.8457	628.8405	-0.0052	27	1473.48	0.77
2	16951.51	16951.34	9.98	-0.17	652.9814	652.988	0.0065	26	8753.23	4.55
3	16951.51	16951.56	-3.52	0.06	679.0696	679.0672	-0.0024	25	28119.4	14.61
4	16951.51	16951.48	1.51	-0.03	707.3186	707.3197	0.0011	24	61537.37	31.97
5	16951.51	16951.5	0.46	-0.01	738.0286	738.029	0.0003	23	103180.77	53.6
6	16951.51	16951.54	-1.95	0.03	771.5315	771.53	-0.0015	22	143109.5	74.34
7	16951.51	16951.57	-3.67	0.06	808.2245	808.2215	-0.003	21	175718.11	91.28
8	16951.51	16951.55	-2.77	0.05	848.5846	848.5823	-0.0024	20	192514.36	100
9	16951.51	16951.55	-2.57	0.04	893.1938	893.1915	-0.0023	19	188235.27	97.78
10	16951.51	16951.49	0.78	-0.01	942.7565	942.7573	0.0007	18	164246.91	85.32
11	16951.51	16951.46	2.94	-0.05	998.1514	998.1544	0.0029	17	122562.37	63.66
12	16951.51	16951.45	3.08	-0.05	1060.4728	1060.4761	0.0033	16	88387.08	45.91
13	16951.51	16951.45	3.35	-0.06	1131.1035	1131.1073	0.0038	15	69181.33	35.94
14	16951.51	16951.42	5.01	-0.08	1211.8227	1211.8288	0.0061	14	55564.74	28.86
15	16951.51	16951.43	4.4	-0.07	1304.9632	1304.9689	0.0057	13	37056.82	19.25
16	16951.51	16951.42	5.29	-0.09	1413.6249	1413.6324	0.0075	12	22214.49	11.54
17	16951.51	16951.51	-0.16	0	1542.0532	1542.0529	-0.0003	11	16660.98	8.65
18	16951.51	16951.51	-0.14	0	1696.1577	1696.1575	-0.0002	10	8065.46	4.19
19	16951.51	16951.65	-8.35	0.14	1884.5233	1884.5076	-0.0157	9	1449.6	0.75
20	16951.51	16951.59	-4.88	0.08	2119.9555	2119.9451	-0.0103	8	246.62	0.13

Figure 6 *MS Peaks Two window for myoglobin.d*

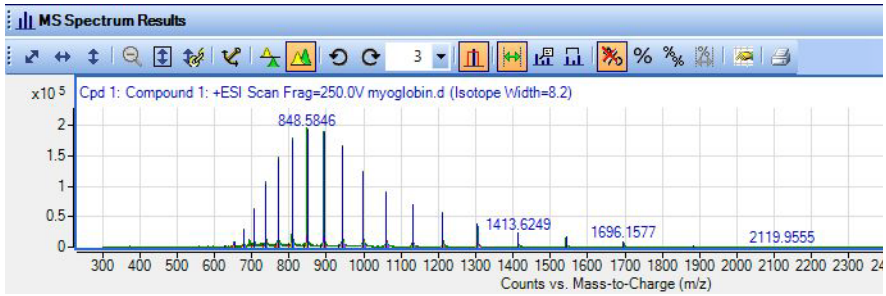


Figure 7 *MS Spectrum Results window for myoglobin.d (Overlaid Mode)*

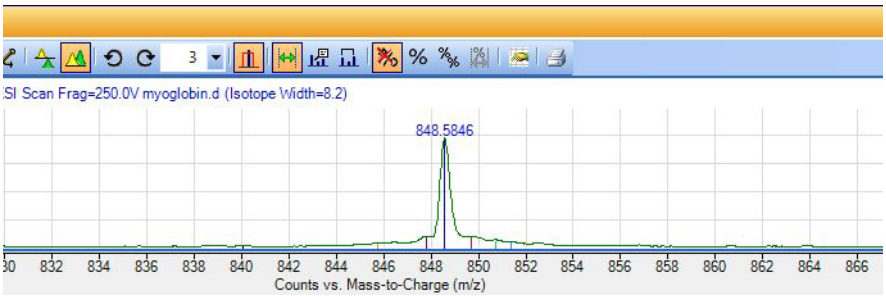
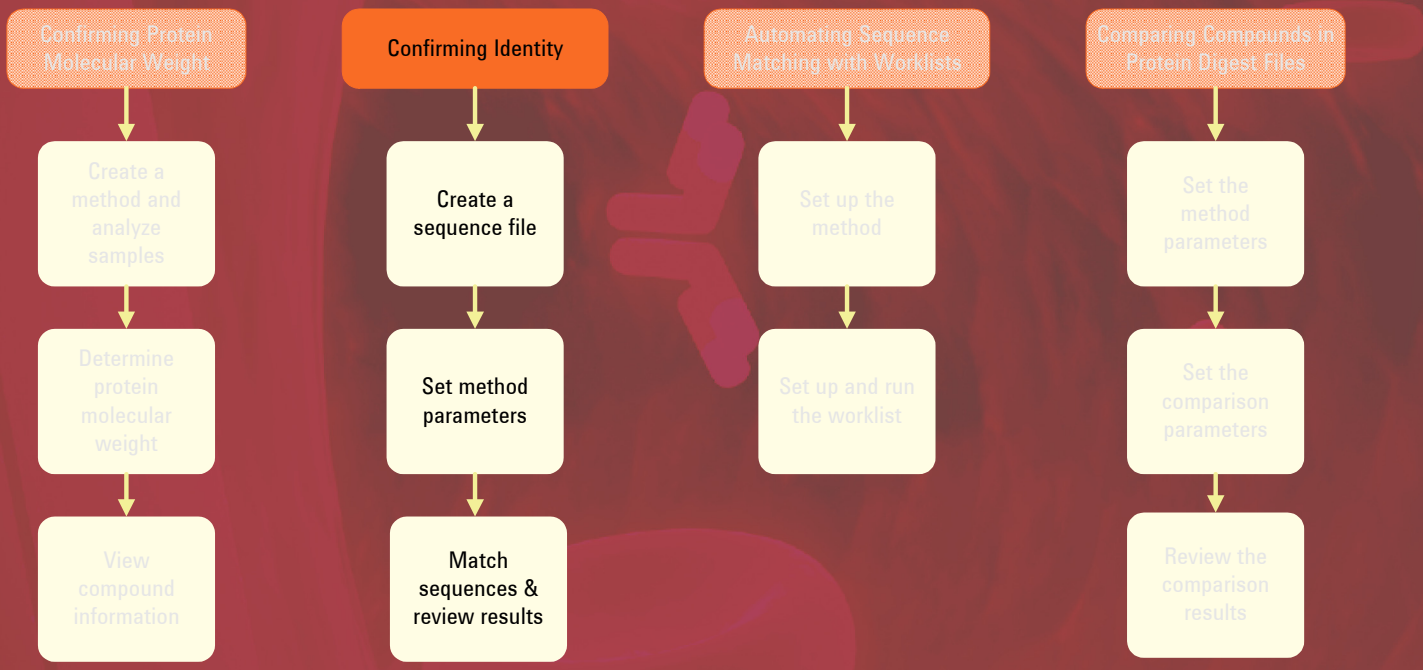


Figure 8 *MS Spectrum Results window for myoglobin.d (zoomed)*

Confirming Identity



Agilent Technologies

Before you begin

This workflow shows you how to create and match sequences.

Required products:

- MassHunter Qualitative Analysis B.07.00 software
- MassHunter BioConfirm B.07.00 program

Example files:

For use in the exercises in this chapter, copy the following sample sequence and data files from the BioConfirm setup disk to your computer hard drive.

For Proteins:

- **myoglobin.psq**
- **myoglobin.d**

For Protein Digests:

- **Serotransferrin.psq**
- **Serotransferrin.d**

For Synthetic Peptides:

- **SynPep3.psq**
- **SynPep3.d**

For Oligonucleotides:

- **21mer_oligo.psq**
- **DNA-2ug-r001.d**

For more information:

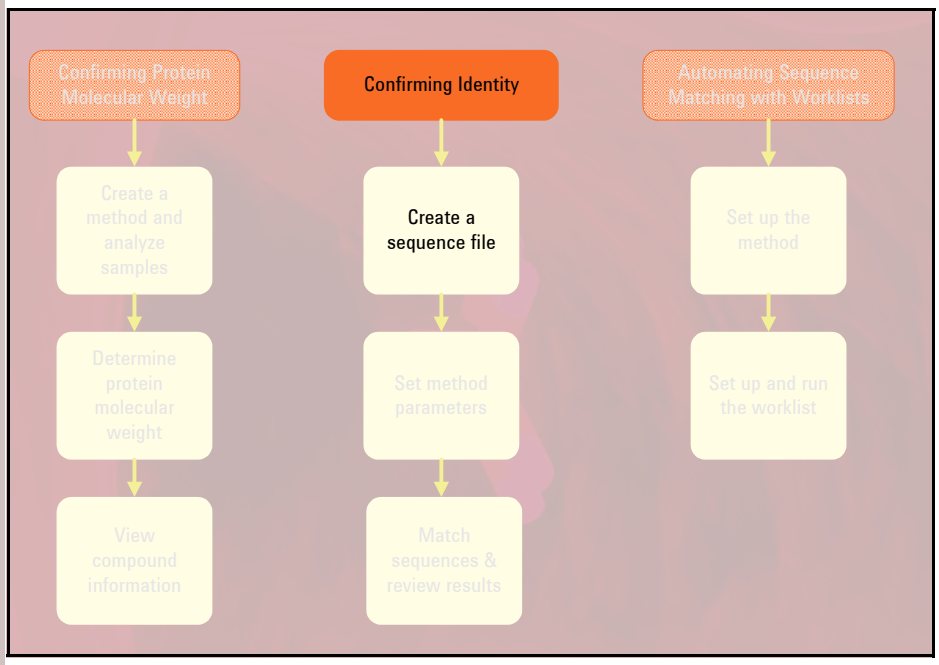
Use the following additional resources for more information on the steps in this workflow:

- *MassHunter Qualitative Analysis online Help*
- *MassHunter BioConfirm Quick Start Guide*

To access online Help, press **F1** or click **Contents**, **Index**, or **Search** on the Help menu.

Create a sequence file

This section tells you how to create a sequence file, using myoglobin as an example.



1. Create a new sequence.

a Click **BioConfirm Workflow > Define and Match Sequences** in the Method Explorer window to display the Define and Match Sequences section.

If the BioConfirm workflow is not available in Method Explorer, select it from the **Configuration > Configure for Workflow** menu.

b Click **Sequence > New Sequence**.

The Sequence Editor window opens automatically with a new sequence displayed for editing.

The new sequence is also added to the list of sequences in the Sequences tab of the Define and Match Sequences section of the Method Editor window.

c In the Sequence Editor window, type in Myoglobin in the **Sequence Name** box.

d In the Sequence Editor window, select the appropriate **Sequence Type** for the new sequence (one of the following):

- Protein
- Protein Digest
- Synthetic peptide
- Oligonucleotide

2. Enter the amino acid sequence.

Use *either* of the following techniques to enter the amino acid sequence of interest:

- Copy a sequence from a FASTA-formatted database or text file and paste it between the N-term and C-term symbols in the **Sequence** field.
- Type in individual amino acids one at a time between the N-term and C-term symbols. Use the single-character (letter) amino acids abbreviations, as shown in the Amino acid list on the left side of the Sequence Editor window.

Here is an amino acid sequence that will give a match for myoglobin. If you are reading this document as a PDF file on your computer, you can copy and paste the sequence into the Sequence Editor window.

GLSDGEWQQVLNVWGKVEADIAGHGQEVLIRLFTGHPETLEKFDKFKHLKTEAEMKASED
LKKHGTVVLTALGGILKKKGHHEAELKPLAQSHATKHKIPIKYLEFISDAIIHVLHSKHPGDFG
ADAQGAMTKALELFRNDIAAKYKELGFQG

Tips:

- The blue numbers to the left and right of the sequence are index numbers used to identify the positions of the amino acids for modifications and links.
- The number on the left is the number of the first amino acid or oligonucleotide base in the row.
- The number on the right is the number of the last amino acid or oligonucleotide base in the row.
- Blue arrows indicate digest information.

The myoglobin sequence shown above does not have any links or modifications, but some sequences do. In that case, add links and modifications as described in the *BioConfirm Quick Start Guide* or *Qualitative Analysis online Help*.

3. Save the sequence.

a Click the **Save as** button on the Sequences tab of the Define and Match Sequences section of the Method Editor window.

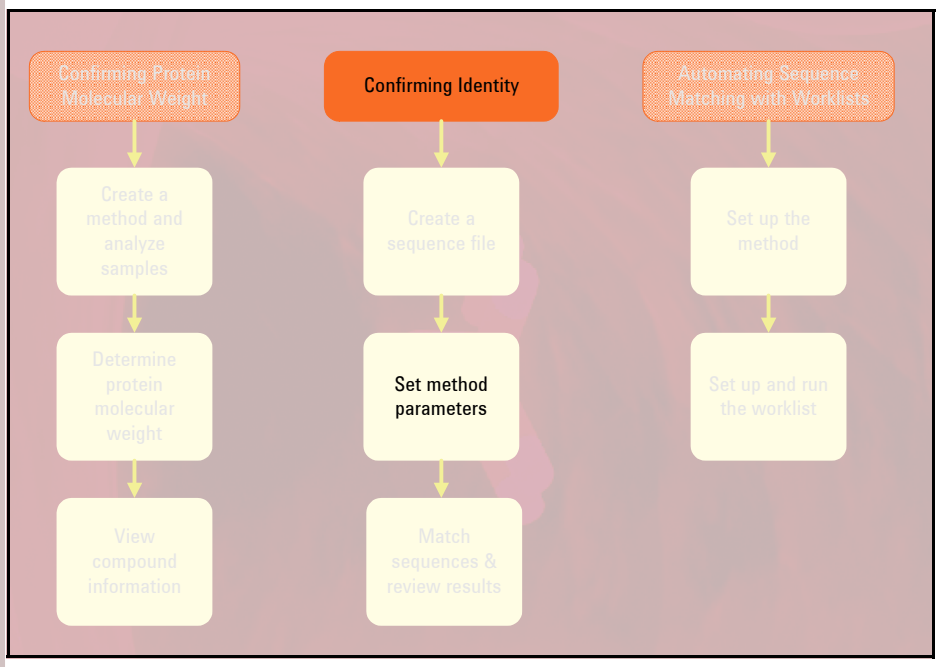
b Type ***iii*_myoglob** in the File name box, where *iii* represents your initials.

c Click the **Save** button.

A saved sequence file can be imported for use in other methods or referenced from worklists.

Set method parameters

This section tells you how to set method parameters, match an intact protein sequence, and view results. You can use your own data or the ***iii*_myoglob.psq** sequence file created in the previous section and the **myoglobin.d** data file.



1. Open the appropriate method.

a Click **Method > Open**.

b Select the appropriate default method from the following:

For Proteins:
- **BioConfirmIntactProtein-Default** (for Maximum Entropy or LMFE Deconvolution)
- **BioConfirmIntactProteinHighMass-Default** (for protein >25 kDa and LMFE Deconvolution)

For Protein Digests:
- **BioConfirmProteinDigest-Default** (with MFE Deconvolution)

For Synthetic Peptides:
- **BioConfirmSynthetic Peptide-Default** (with MFE Deconvolution)

For Oligonucleotides:
- **BioConfirmOligonucleotideSmall-Default** (for Oligonucleotides <7 kDa and MFE Deconvolution)
- **BioConfirmOligonucleotideLarge-Default** (for Oligonucleotides >7 kDa and LMFE or MaxEnt Deconvolution)

2. Open the data file of interest.

a Click **File > Open Data File**.

b Select the data folder of interest.

The following data files are provided on the BioConfirm setup disk. Copy the data file from the Data folder on the BioConfirm setup disk to your computer hard drive.

- For Proteins, use **myoglobin.d**
- For Protein Digests, use **enolase-Chip-final.d.** Note that you will process both enolase data files in the last chapter of this guide, so processing in the current chapter is optional.
- For Synthetic Peptides, use **SynPep3.d**
- For Oligonucleotides, use **DNA-2ug-r001.d**

c Click **Open**.

The TIC is automatically displayed in the Chromatogram Results window.

3. Find compounds.

a Click **BioConfirm Workflow > Find by Molecular Feature** in the Method Explorer window to display the Find Compounds by Molecular Feature section in the Method Editor window. If the BioConfirm workflow is not available in Method Explorer, select it from the **Configuration > Configure for Workflow** menu.

b Review the method settings and modify them if necessary.

If you are processing the demo data file **DNA-2ug-r001.d** for oligonucleotides, certain settings in the default method **BioConfirmOligonucleotideSmall-Default** need to be changed to avoid lengthy processing times. See Figure 9, Figure 10, and Figure 11 for appropriate settings for processing this data file.

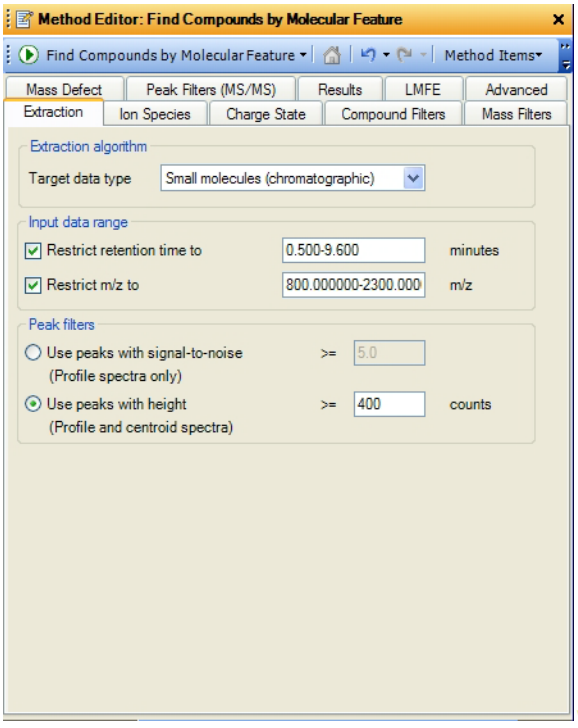


Figure 9 *Extraction Tab Parameter Settings for DNA-2ug-r001.d*

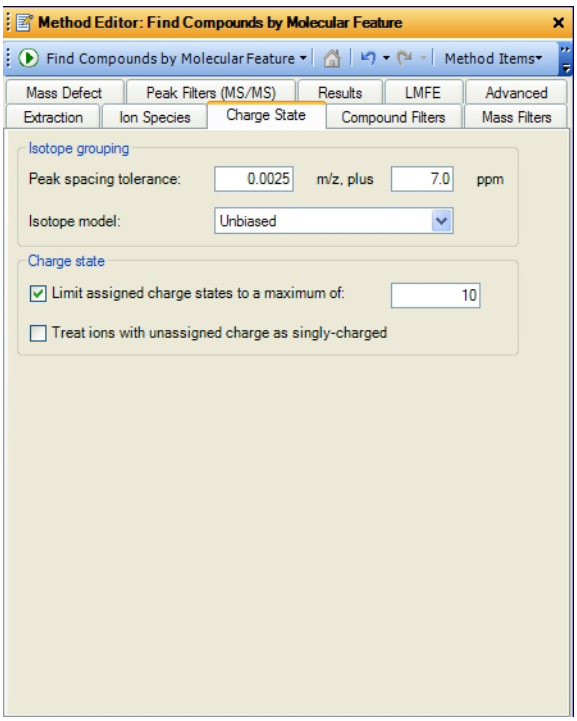


Figure 10 *Charge State Tab Parameter Settings for DNA-2ug-r001.d*

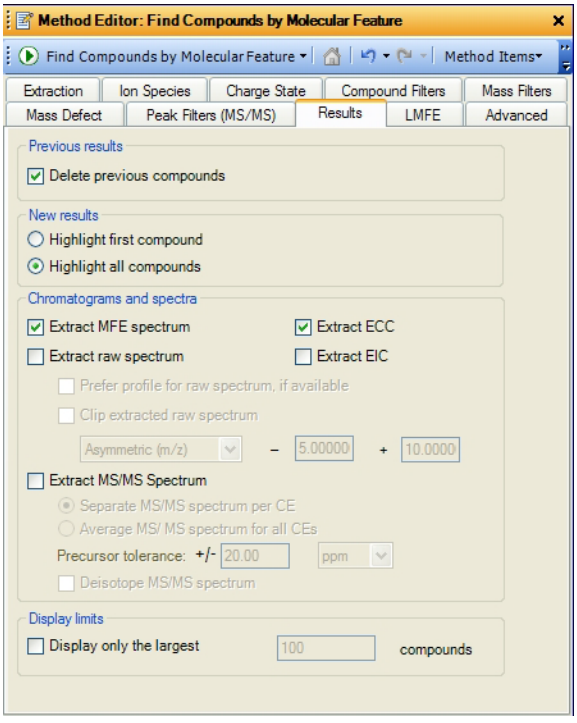


Figure 11 *Results Tab Parameter Settings for DNA-2ug-r001.d*

For more information on setting method parameters, see the *BioConfirm Quick Start Guide* or *Qualitative Analysis online Help*.

c To start the compound search, click the process button on the Method Editor toolbar.

d Review the results in the Compound List window.

4. Set the appropriate layout.

a Click **Load Layout** on the **Configuration > Window Layouts** menu.

b Select the appropriate layout from the following:

For Proteins:
- **BioConfirm-IntactProtein-MaximumEntropy-Default** (for Maximum Entropy Deconvolution)
- **BioConfirmIntactProtein-LMFE** (for LMFE Deconvolution)

For Protein Digests:
- **BioConfirm-ProteinDigest** (with MFE Deconvolution)

For Synthetic Peptides:
- **BioConfirm-Synthetic Peptide** (with MFE Deconvolution)

For Oligonucleotides:
- **BioConfirm-Oligonucleotide** (for MFE, LMFE, or MaxEnt Deconvolution)

c Click **Open**.

5. Import the sequence file.

a Click **BioConfirm Workflow > Define and Match Sequences** in the Method Explorer window.

If the BioConfirm workflow is not available in Method Explorer, select it from the **Configuration > Configure for Workflow** menu.

b Click the **Sequences** tab.

c Click **Import**.

d Select the appropriate sequence file.

The following sequence files are provided on the BioConfirm setup disk. Copy the files to your computer hard drive.
- For Proteins, use **myoglobin.psq** or the sequence file you created in the previous section.
- For Protein Digests, use **EnolaseDigest.psq**.
- For Synthetic Peptides, use **SynPep3.psq**.
- For Oligonucleotides, use **21mer_oligo.psq**.

e Click **Open**.

6. Assign digest reagents and view the digest list. (*For Protein Digests only*)

a Click the **Edit** button on the Sequences tab to open the Sequence Editor window.

b Right-click in the Sequence Editor window and click **Edit Digest Reagents** from the shortcut menu to open the Digest Reagents dialog box. Notice that **Trypsin** is already selected as the Reagent for Serotransferrin.

c Set the maximum number of **Missed Cleavages** to 2.

d Click **OK**.

Tip: You can customize the list of available reagents using the Chemical Data Dictionary; see *Qualitative Analysis online Help* for more information.

e Do *one* of the following to digest the sequence and display the results in the Digest List window:
- Right-click the sequence in the Sequence Editor window and click **Digest Current Sequence** from the shortcut menu.
- Click **Sequence > Digest Current Sequence**.

7. Select matching rules.

a To open the Rules dialog box, right-click in the Sequence Editor window and click **Edit Matching Rules** from the shortcut menu.

b Select the rules to use for matching the theoretical masses to those from MS data. Note that appropriate rules are set automatically by the method you opened in Step 1. Use **Ctrl+click** to select multiple rules from the list.

For Protein:

- Intact protein
- Predicted Modifications

For Protein Digests:

- Complete Digest
- Incomplete Digest
- Predicted Modifications

For Synthetic Peptides:

- Intact Peptide
- Extra Amino Acid
- Missing Amino Acid
- Fmoc blocking groups

For Oligonucleotides:

- Intact oligonucleotide
- Oligonucleotide truncation

c Click **OK** to close the Rules dialog box.

8. Select the match source.

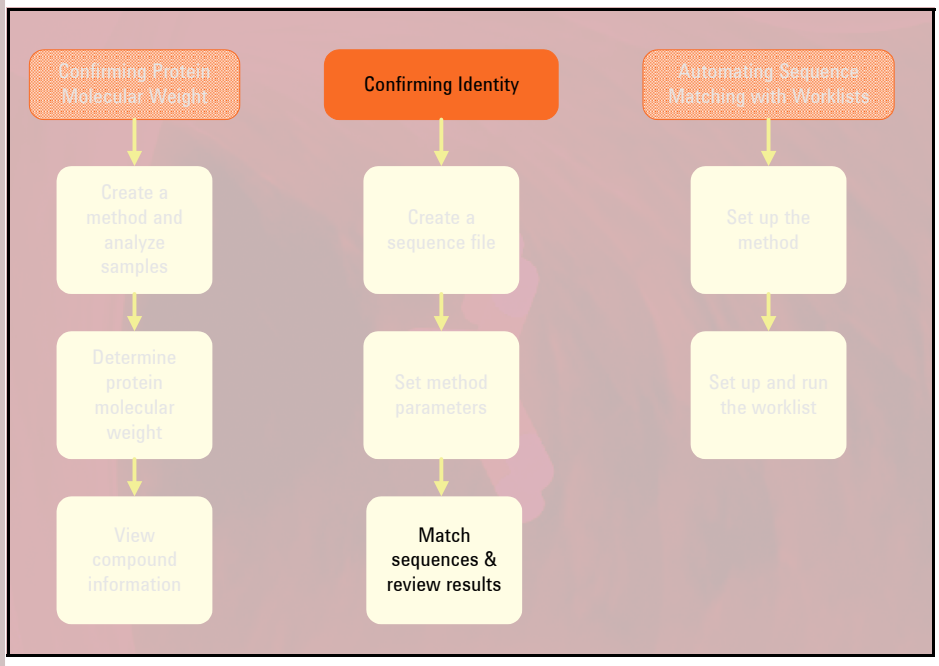
On the **Source** tab in the Define and Match Sequences section of the Method Editor window:

a Click **Qualitative method**.

b Mark the **Sequences** check box.

Match sequences and review the results

This section tells you how to run a match sequence and view the results.



1. Start the match sequence.

a Use the method parameters set in the previous section.

b Use *one* of the following methods to start the search:

- Click  on the Method Editor toolbar.
- Right-click in the Compound List window and click **Match Sequences**.
- Click **Sequence > Match Sequences**.
- Click **Match Sequences** on the Method Editor shortcut menu.

The results appear in the Compound List window.

2. Review the results.

a Display the Compound Identification Results window in *either* of the following ways if it is not currently visible:

- Click  on the main toolbar.
- Click **View > Compound Identification Results**.

When you open the window, the window displays the results for the first compound that is highlighted in the Data Navigator (that is marked to show). If you see a message such as "No data to display" or "No identification techniques applied to this compound," then click a row in the Compound List to reveal the data.

b View results for other compounds by clicking on them in the Data Navigator or Compound List or Compound Identification Results windows.

c To see other information for compounds in the list, right-click in the table and then click **Add/Remove Columns** from the shortcut menu.

d Evaluate/confirm the match quality by examining the scores. The specific scoring values you see vary and you may not see the ones described in the bullets below.:
- If the value in the **Score (Bio, MS)** column is good and the value in the Score **(Bio, MS/MS)** column is high, then it probably is a true match.
- If the value in the **Score (Bio, MS)** column is very good, but value in the Score **(Bio, MS/MS)** column is low, then it may be a false positive.

Note on scoring values: The scoring values that are reported depend on your data and on the parameter settings in Method Editor Window > Define and Match Sequences section Scoring tab. Scoring values range from 0-100. Using default method settings, if a compound has both MS and MS/MS data, a score value of 80 is considered high. For information on scoring settings, see Scoring Tab (Define and Match Sequence) in *Qualitative Analysis online Help*.

e To identify the location of predicted modification for proteins or protein digests, look for the entry in the **Pred Mods** column that has the highest value in the Score (Bio, MS/MS) column.

f You can also view sequence match results in the Compound List Window, which is a nested table of results.

3. Save the method for later use.

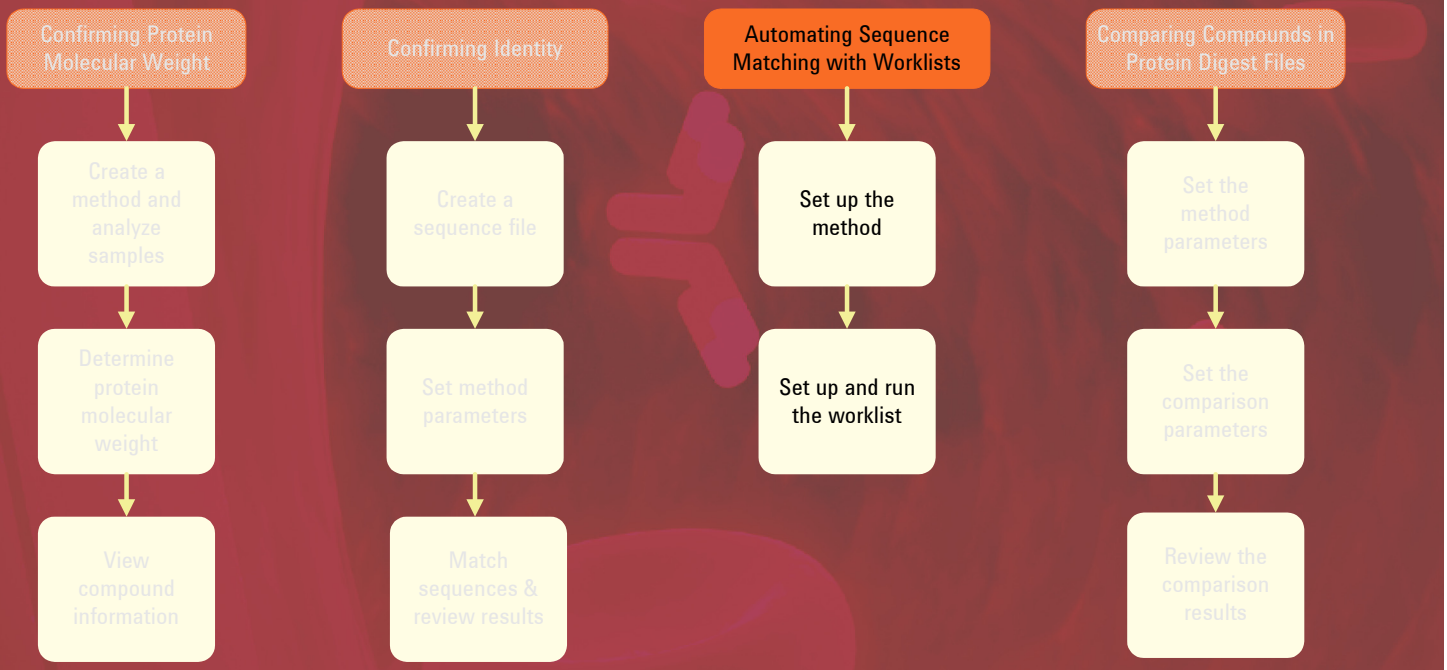
   You will use these methods to automate sequence matching in a worklist in the next chapter.

a Click **Method > Save As**.

b Type the File name, using one of the following names, where ***iii*** represents your initials:
- For Proteins, use ***iii*_myoglobin**.
- For Protein Digests, use ***iii*_EnolaseDigest.**
- For Synthetic Peptides, use ***iii*_SynPep3**.
- For Oligonucleotides, use ***iii*_21mer_oligo**.

c Click **Save**.

Automating Sequence Matching with Worklists



Agilent Technologies

Before you begin

This workflow shows you how to set up a worklist to automatically confirm the presence of a selected sequence in previously acquired samples.

Required products:

- MassHunter Data Acquisition B.06.01 software
- MassHunter Qualitative Analysis B.07.00 software
- MassHunter BioConfirm B.07.00 program

Example files:

For use in the exercises in this chapter, copy the following sample sequence and data files from the BioConfirm setup disk to your computer hard drive.

For Proteins:

- **myoglobin.psq**
- **myoglobin.d**

For Protein Digests:

- **EnolaseDigest.psq**
- **enolase-Chip-final.d**

For Synthetic Peptides:

- **SynPep3.psq**
- **SynPep3.d**

For Oligonucleotides:

- **21mer_oligo.psq**
- **DNA-2ug-r001.d**

For more information:

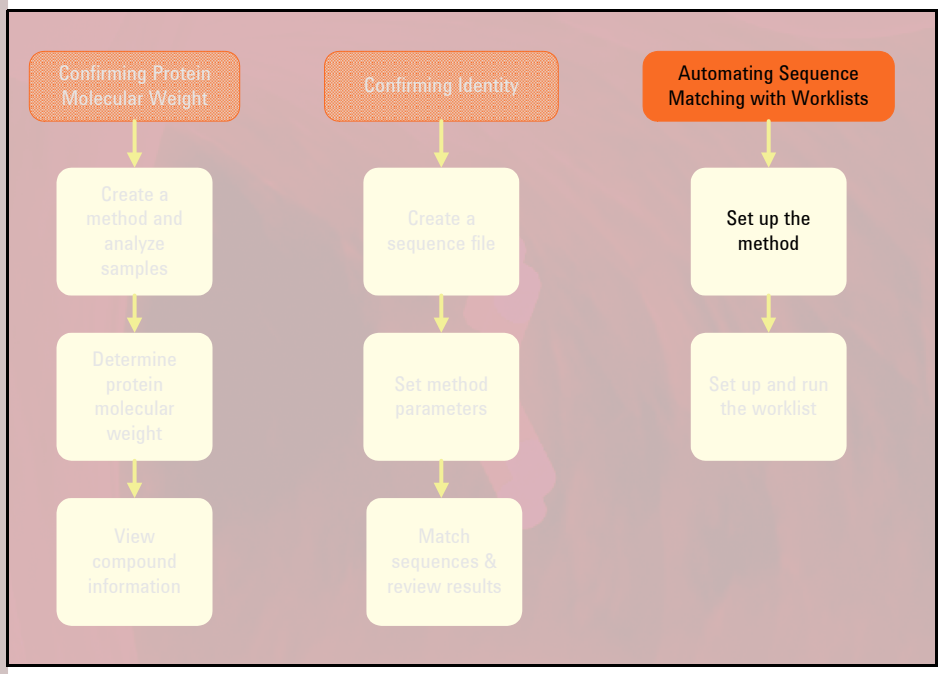
Use the following additional resources for more information on the steps in this workflow:

- *MassHunter LC/MS Data Acquisition online Help*
- *MassHunter Qualitative Analysis online Help*
- *MassHunter BioConfirm Quick Start Guide*

To access online Help, press **F1** or click **Contents**, **Index**, or **Search** on the Help menu.

Setting up the method in Qualitative Analysis

This section tells you how to set up a method in Qualitative Analysis that can be used to automatically confirm the presence of a known sequence in a samples acquired in Data Acquisition using a worklist.



1. Set up the method.

a Click **Method > Open** and select the method of interest.

Tip: The following default methods are provided as starting points:

For Proteins:
- **BioConfirmIntactProtein-Default**
  (for Maximum Entropy or LMFE Deconvolution)
- **BioConfirmIntactProteinHighMass-Default**
  (for protein >25 kDa and LMFE Deconvolution)

For Protein Digests:
- **BioConfirmProteinDigest-Default**
  (with MFE Deconvolution)

For Synthetic Peptides:
- **BioConfirmSynthetic Peptide-Default**
  (with MFE Deconvolution)

For Oligonucleotides:
- **BioConfirmOligonucleotideSmall-Default**
  (for Oligonucleotides <7 kDa and MFE Deconvolution)
- **BioConfirmOligonucleotideLarge-Default**
  (for Oligonucleotides >7 kDa and LMFE or MaxEnt Deconvolution)

b Click **BioConfirm Workflow > Define and Match Sequences** in the Method Explorer window.

If the BioConfirm workflow is not available in Method Explorer, select it from the **Configuration > Configure for Workflow** menu.

c Make the following selection on the **Source** tab in the Define and Match Sequences section of the Method Editor window:
- Click **Worklist**.

This selection causes the software to get the sequence from the worklist rather than the method.

2. Set the worklist automation options and save the method.

a Click **Worklist Automation > Worklist Actions** in the Method Explorer window. (You may need to scroll down to see that option.)

b Select **Find Compounds by Molecular Feature** in the Available actions list.

If you are using your own data, review the settings and modify them if necessary. For more information on setting parameters, see the *BioConfirm Quick Start Guide* or *Qualitative Analysis online Help*.

c If they are not already in the Actions to be run list, make the following selections in the Available actions list:
- Select **Extract Defined Chromatograms**
- Select **Integrate and Deconvolute**
- Select **Match Sequences**.
- Select **Generate Compound Report**.

d To add the selected actions to the Actions to be run list, click the  button.

The items are added to the end of the list.

e If other items are on the list, remove them.

f Because actions are executed in the order they appear in the list, use the Up and Down arrow buttons to the right of the list to set the desired order.

g Click **Method > Save As** to save the modified method with a new name.

h Type the File name, using one of the following names, depending on your data type:
- For Proteins, use ***iii*_myoglobin_auto**.
- For Protein Digests, use ***iii*_EnolaseDigest._auto**.
- For Synthetic Peptides, use ***iii*_SynPep3_auto**.
- For Oligonucleotides, use ***iii*_21mer_oligo_auto**.

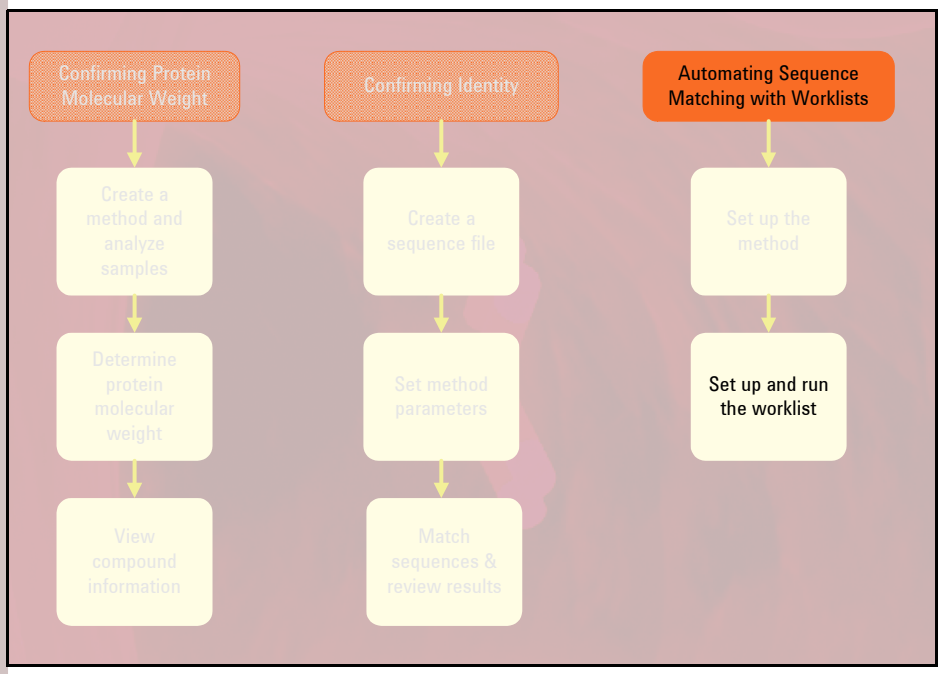
where ***iii*** represents your initials.

i Click **Save**.

You will use these methods to automate sequence matching in a worklist in the next section.

Setting up and running a worklist in Data Acquisition

This section tells you how to set up and run a worklist in Data Acquisition, using the Qualitative Analysis method created in the previous section to automatically confirm the presence of a known sequence. Myoglobin is used as an example.



1. Create a worklist in Data Acquisition.

   a Open the MassHunter Workstation Data Acquisition software window, if it is not already open.

   b Display the Worklist pane.

   c Click **Worklist > Add Sample**.

   A new sample row is added to the Worklist table.

2. Specify the sequence file in the worklist.

   a Click **Worklist > Add Column**.

   b When the Add Column dialog box appears:
   - Select **Protein** as the **ColumnType**.
   - Enter the **Column name** as **Sequence**.
   - Select ***iii*_myoglob.psq** as the **Value**.

   c Click **OK**.

3. Specify the method and data file in the worklist.

   a Enter ***iii*_myoglobin_auto.m** in the **Method** column of the worklist.

   b Enter **myoglobin.d** in the **Data File** column of the worklist.

4. Set up the worklist to run data analysis only.

   a Click **Worklist > Worklist Run Parameters**.

   b When the Worklist Run Parameters dialog box appears:
   - Select **DA Only** as **Part of method to run**.
   - Select the paths for the DA method and data file.

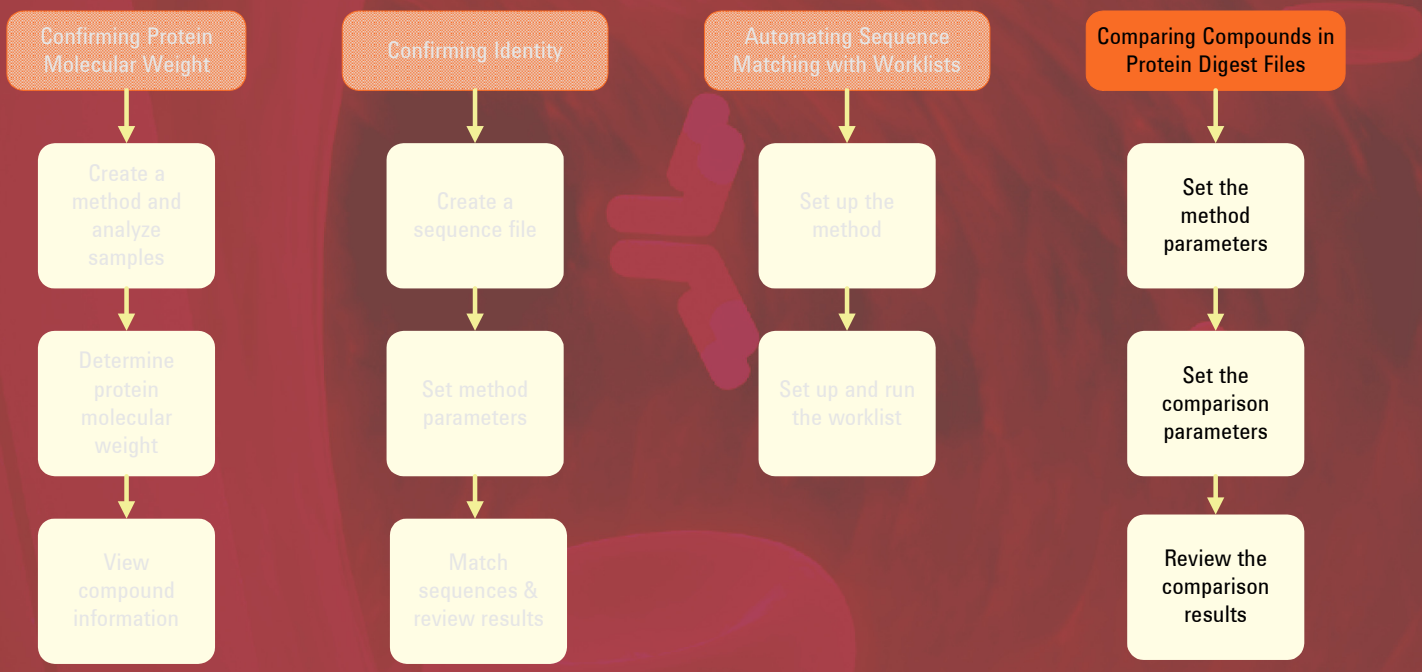
5. Run the worklist and review results.

c Click **OK**.

a Click **Worklist > Run**.

b Review the printed Compound reports.

Comparing Compounds in Protein Digest Files



Agilent Technologies

Before you begin

This workflow shows you how to compare compounds in two or more protein digest data files.

Required products:

- MassHunter Qualitative Analysis B.07.00 software
- MassHunter BioConfirm B.07.00 program
- MassHunter Comparative Analysis program (included with Qualitative Analysis software)

Example files:

For use in the exercises in this chapter, copy the following data files from the BioConfirm setup disk to your computer hard drive.

- **enolase-Chip-final.d**
- **enolase-oxidized-Chip-final.d**

Copy the following sequence file to **\MassHunter\ProteinSequences**:

- **EnolaseDigest.psq**

For more information:

Use the following additional resources for more information on the steps in this workflow:

- *MassHunter Qualitative Analysis online Help*
- *MassHunter BioConfirm Quick Start Guide*

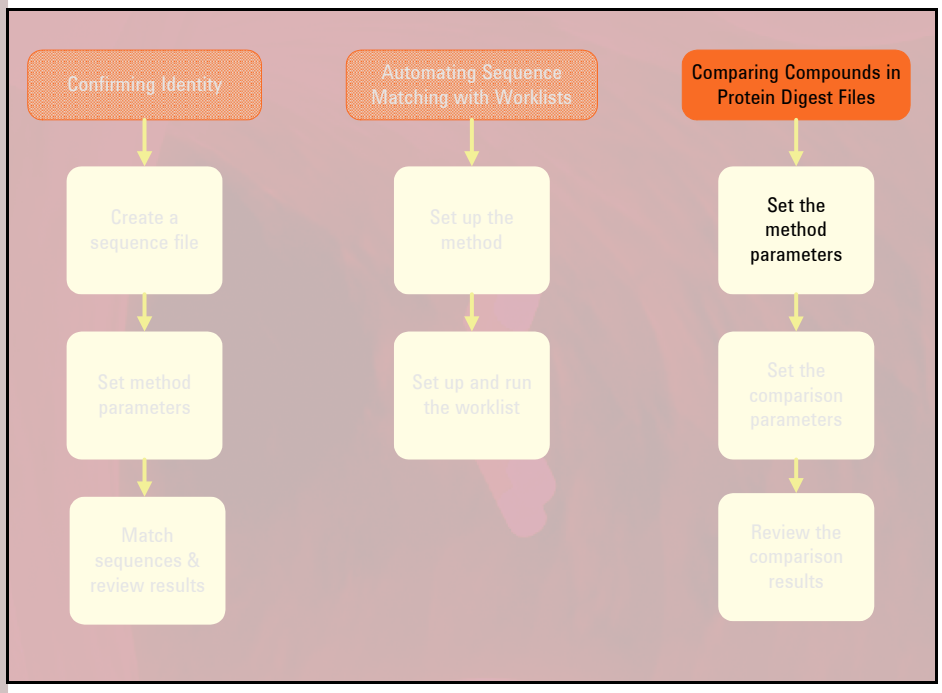
To access online Help, press **F1** or click **Contents**, **Index**, or **Search** on the Help menu.

Set the method parameters

This section tells you how to set parameters to compare compounds in two protein digest files. This workflow uses the **enolase-Chip-final.d** and **enolase-oxidized-Chip-final.d** data files, as well as the **EnolaseDigest.psq** sequence file.

Before you start:

If the BioConfirm workflow is not available in Method Explorer, select it from the Tools > Configure for Workflow menu.



1. Open the data files to compare.

   a Click **File > Open Data File**.

   b On the Open Data File dialog box, select the **enolase-Chip-final.d** folder.

   c Click **Open**.

   d Repeat *Steps a - c* to open **enolase-oxidized-Chip-final.d**.

2. View method parameters.

   a Open the **BioConfirmProteinDigest-Default** method.

   b To view the method parameters in this default method that will be used to find compounds, display the **BioConfirm Workflow > Find by Molecular Feature** sections of the Method Editor. For example parameters, see:

   - Figure 12, "Find Compounds by Molecular Feature: Extraction Tab," on page 38
   - Figure 13, "Find Compounds by Molecular Feature: Ion Species Tab," on page 38
   - Figure 14, "Find Compounds by Molecular Feature: Charge State Tab," on page 39
   - Figure 15, "Find Compounds by Molecular Feature: Results Tab," on page 39

   c To start the compound search, click the process button on the Method Editor toolbar, and then click **Find**. Under Operation in Progress, observe that both data files are processed.

d Review the results in the Compound List window.

e To view the method parameters in this default method that will be used to match sequences and identify compounds, display the **BioConfirm Workflow > Define and Match Sequences** section of the Method Editor. For example parameters, see:

- Figure 16, "Define and Match Sequences: Sequences Tab," on page 40
- Figure 17, "Define and Match Sequences: Mass Matching Tab," on page 40

f To start the match search, click the process button on the Method Editor toolbar, and then click **Match**.

g Review the results in the Compound Identification Results window. If you see no results, click a row in the Compound List to display the results for that row.

Tip: You can also use **Compare Protein Digest Files: Find Results, Identify and Compare** on the Wizards menu to have a wizard guide you through the process of finding and identifying compounds in your data files.

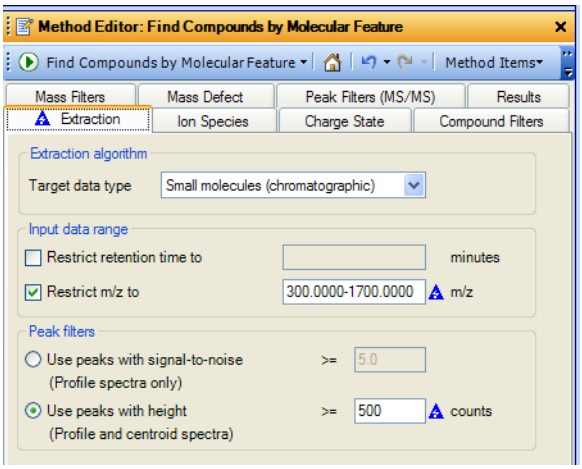


Figure 12 *Find Compounds by Molecular Feature: Extraction Tab*

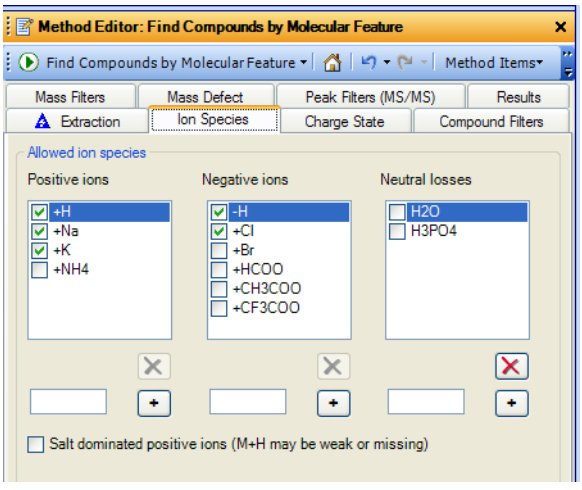


Figure 13 *Find Compounds by Molecular Feature: Ion Species Tab*

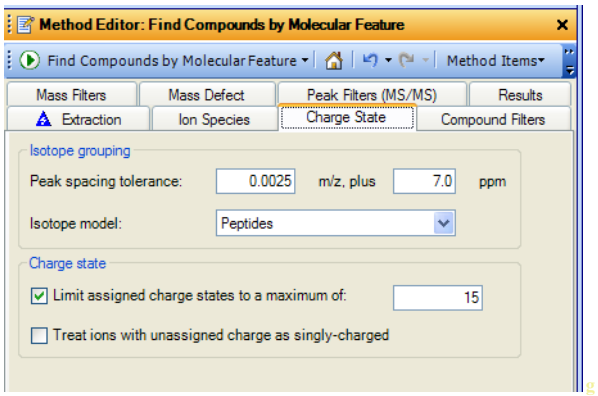


Figure 14 *Find Compounds by Molecular Feature: Charge State Tab*

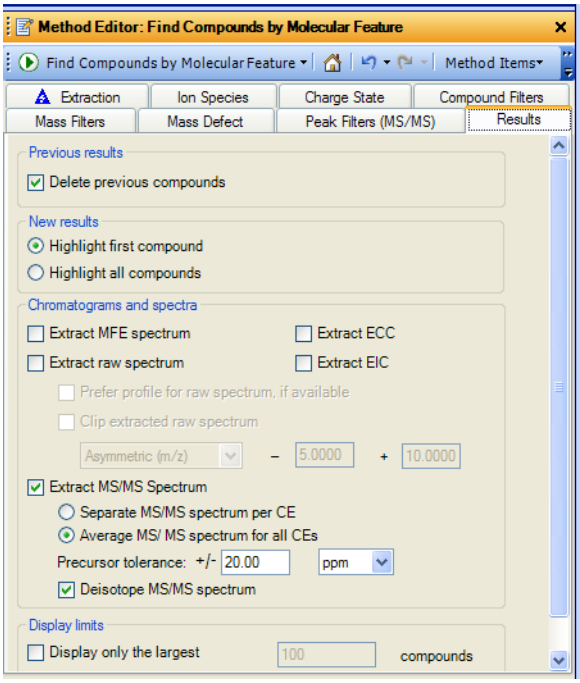


Figure 15 *Find Compounds by Molecular Feature: Results Tab*

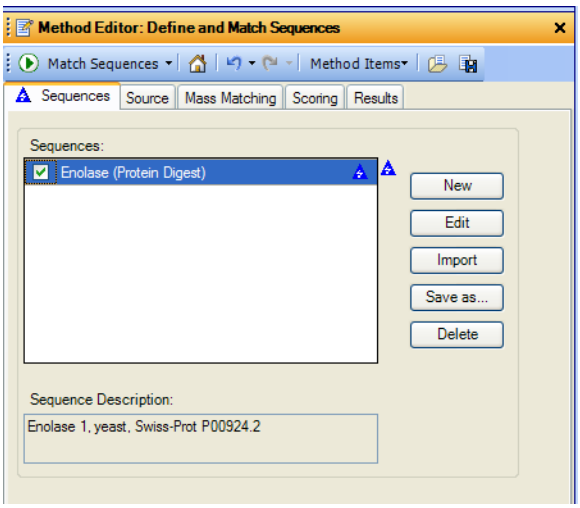


Figure 16 *Define and Match Sequences: Sequences Tab*

(If the enolase sequence does not automatically appear, click **Import** and select **EnolaseDigest.psg**. Then click **Open**.)

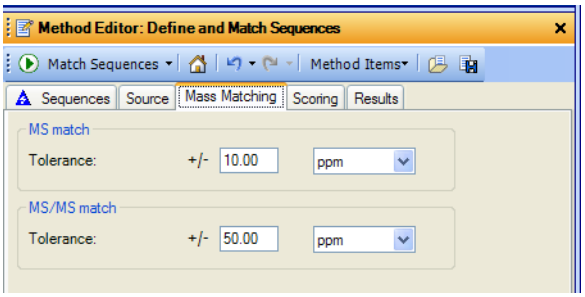


Figure 17 *Define and Match Sequences: Mass Matching Tab*

3. View the sequence coverage.

Note the information displayed in the Sequence Coverage Map window.

In this example, 93% coverage is obtained for the Enolase digest sequence as shown in Figure 18 below.

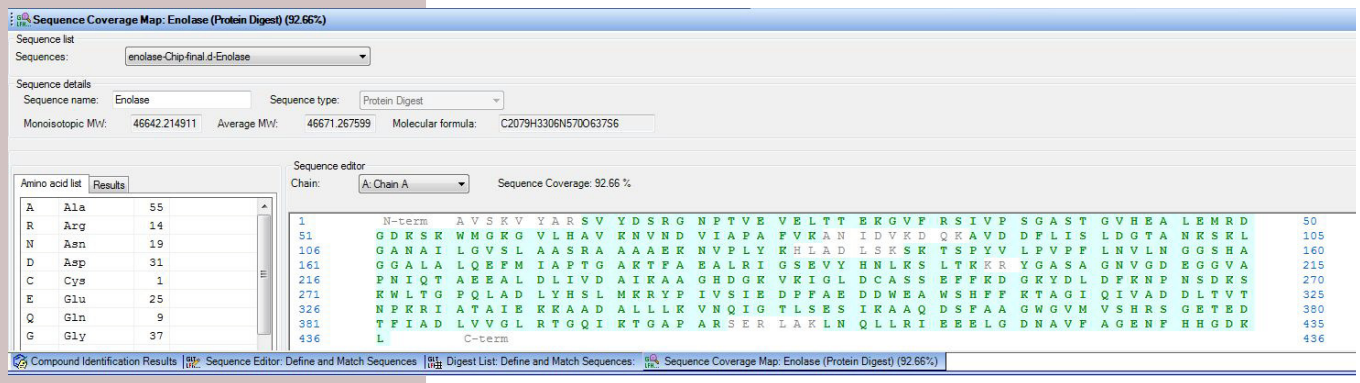
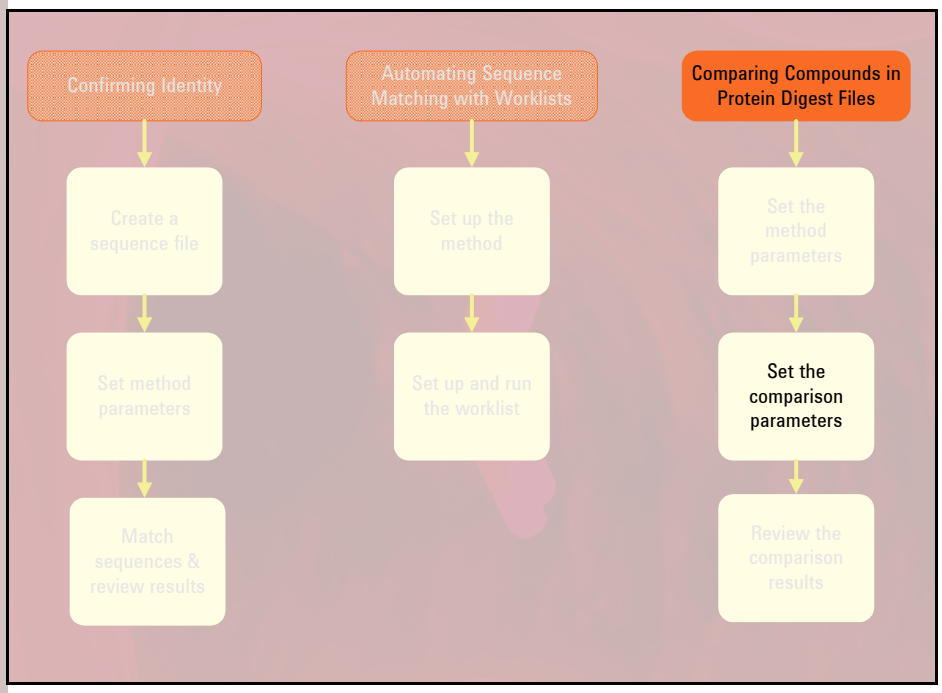


Figure 18 *Sequence Coverage Map for Enolase Digest Sequence*

Set the comparison parameters

This section tells you how to set the parameters used to compare compounds in two protein digest files.



1. Start the compare protein digest files wizard.

Click **Wizards > Compare Protein Digest Files: Compare Existing Results**.

The Select Reference file and Sample file(s) page is displayed.

2. Select the files to use as Reference and Sample files.

a From the list of opened files, select **enolase-Chip-final.d** and click **Select reference file**.
- The selected file name appears in the **Reference file** text box.
- The remaining file, **enolase-oxidized-Chip-final.d**, is automatically moved to the list of **Sample files**.

Tips for your own data files:
- To set the order that the sample files will be processed, use the Up and Down arrow buttons to the right of the list of samples.
- To open additional data files, click the Browse (...) button.
- To remove a selected Reference or Sample file, double-click the file or click the Up arrow button.
- To remove all Sample files, click the double Up arrow button.

b Click **Next** to display the next page of the wizard.

3. Set the compound correlation parameters on the Alignment Information page.

a Set the retention time **(RT) window** and **RT window tolerance** values.

b Set the **Mass window** and **mass window tolerance** values.

c When finished setting the parameters, click **Finish**.

A progress bar is displayed while the compounds are being compared.

This process may take several minutes to complete, depending on the number of samples and size of the data files. When completed, the results are displayed in the MassHunter Comparative Analysis program. See "Review the comparison results" on page 45.

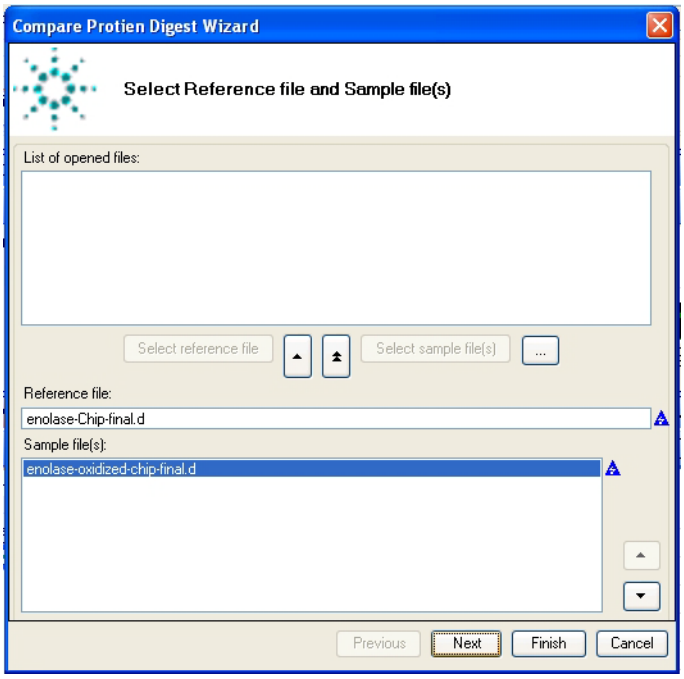


Figure 19 *Reference and Sample files selected*

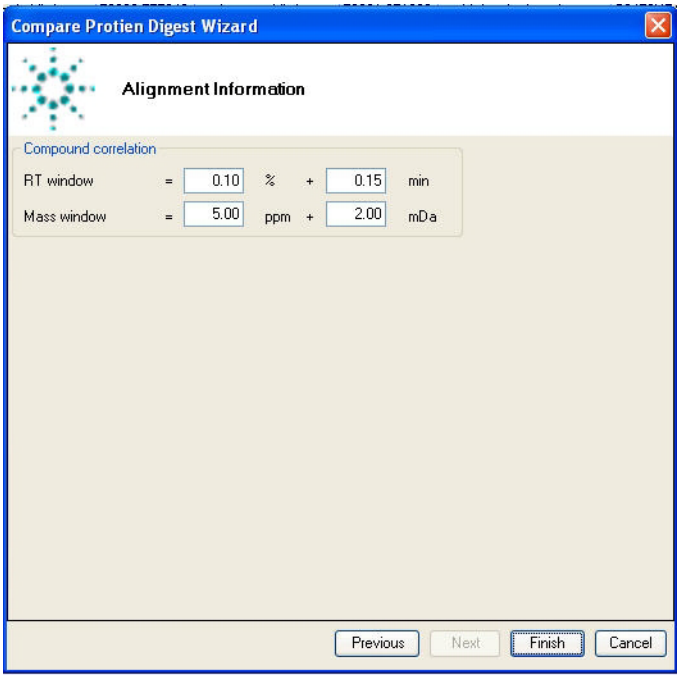
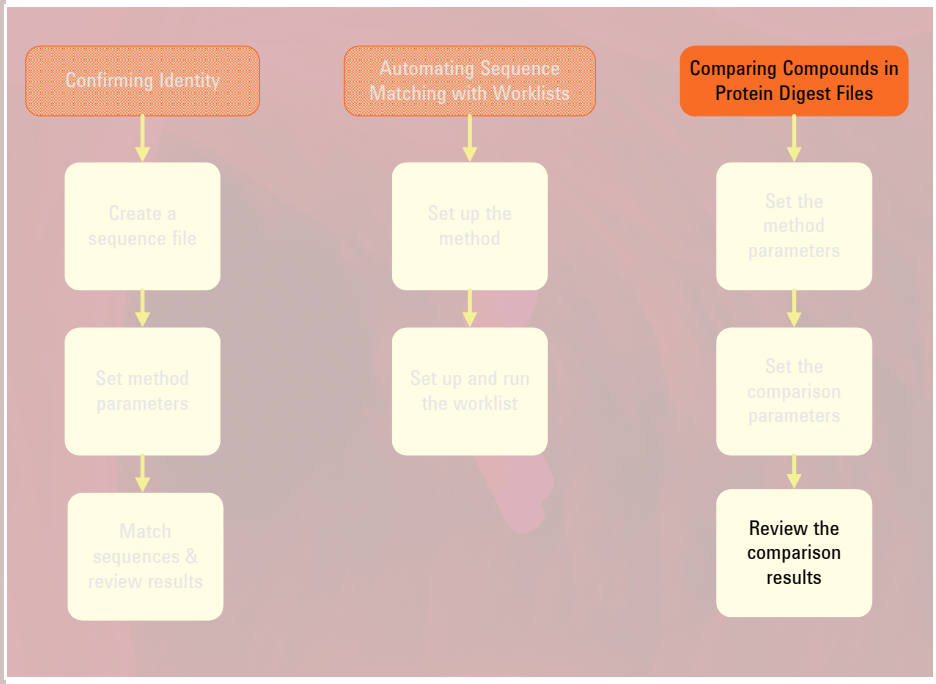


Figure 20 *Alignment Information page with default parameters*

Review the comparison results

This section tells you how to review the results of comparing compounds in two protein digest files, using the MassHunter Comparative Analysis program.



1. View results.

Results are displayed in the Comparative Analysis program window, as shown in the example in Figure 21 on page 46.

The figure is just an example and your results will differ depending on:

- The results windows you display
- The row you click
- Whether or not you display empty table columns
- The x-axis scale you choose for the chromatograms
- Whether you have previously clicked a row in the Compound Compare List.

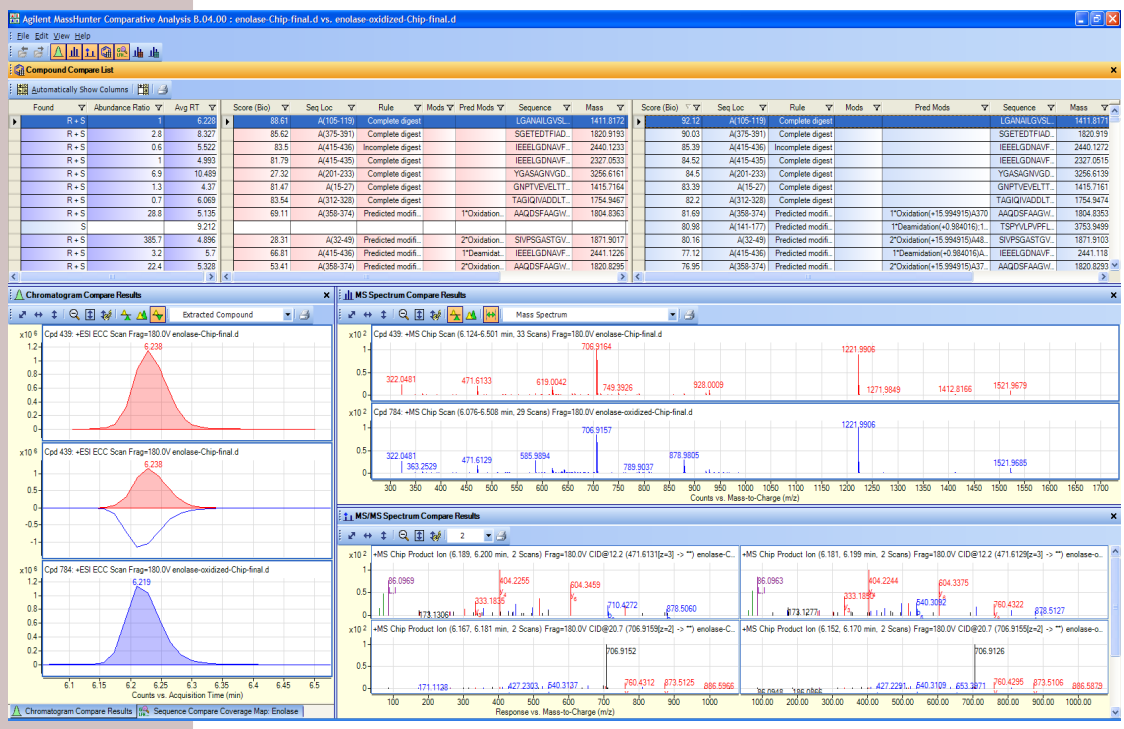


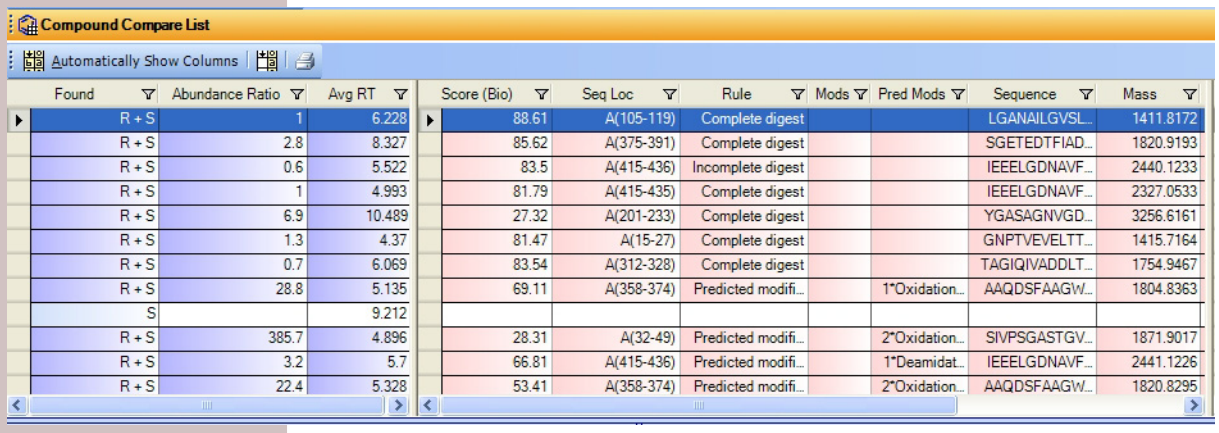
Figure 21 *MassHunter Comparative Analysis Program Window*

2. View compound comparison results.

a If the Compound Compare List window is not visible, open it by clicking **View > Compound Compare List**.

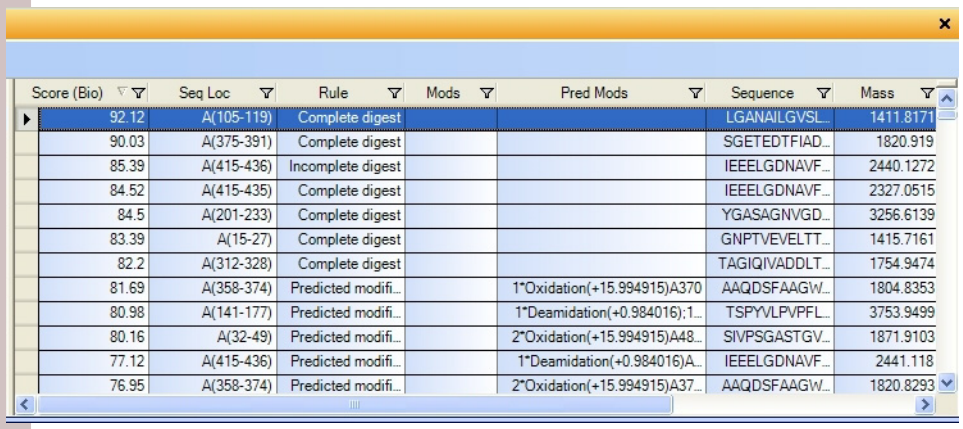
b Find the **Found** column in the first section of the table.
- Click a row that has the value **R** in the Found column.
  Compounds in these rows are found in the reference data file only.
- Click a row that has the value **S** in the Found column.
  Compounds in these rows are found in sample data files only.
- Click a row that has the value **R+S** in the Found column.
  Compounds in these rows are found in both reference and sample data files.

c Note that compound information is presented in three sections of the Compound Compare List window, as shown in Figure 22 and Figure 23 below.
- Section 1 (purple shading): correlated compound information
- Section 2 (red shading): reference file information
- Section 3 (blue shading): sample file information



Found	Abundance Ratio	Avg RT	Score (Bio)	Seq Loc	Rule	Mods	Pred Mods	Sequence	Mass
R + S	1	6.228	88.61	A(105-119)	Complete digest			LGANAILGVSL...	1411.8172
R + S	2.8	8.327	85.62	A(375-391)	Complete digest			SGETEDTFIAD...	1820.9193
R + S	0.6	5.522	83.5	A(415-436)	Incomplete digest			IEEELGDNAVF...	2440.1233
R + S	1	4.993	81.79	A(415-435)	Complete digest			IEEELGDNAVF...	2327.0533
R + S	6.9	10.489	27.32	A(201-233)	Complete digest			YGASAGNVGD...	3256.6161
R + S	1.3	4.37	81.47	A(15-27)	Complete digest			GNPTVEVELTT...	1415.7164
R + S	0.7	6.069	83.54	A(312-328)	Complete digest			TAGIQIVADDLT...	1754.9467
R + S	28.8	5.135	69.11	A(358-374)	Predicted modifi...		1*Oxidation...	AAQDSFAAGW...	1804.8363
S		9.212							
R + S	385.7	4.896	28.31	A(32-49)	Predicted modifi...		2*Oxidation...	SIVPSGASTGV...	1871.9017
R + S	3.2	5.7	66.81	A(415-436)	Predicted modifi...		1*Deamidat...	IEEELGDNAVF...	2441.1226
R + S	22.4	5.328	53.41	A(358-374)	Predicted modifi...		2*Oxidation...	AAQDSFAAGW...	1820.8295

Figure 22 *Expanded view of Sections 1 & 2 of the Compound Compare List*



Score (Bio)	Seq Loc	Rule	Mods	Pred Mods	Sequence	Mass
92.12	A(105-119)	Complete digest			LGANAILGVSL...	1411.8171
90.03	A(375-391)	Complete digest			SGETEDTFIAD...	1820.919
85.39	A(415-436)	Incomplete digest			IEEELGDNAVF...	2440.1272
84.52	A(415-435)	Complete digest			IEEELGDNAVF...	2327.0515
84.5	A(201-233)	Complete digest			YGASAGNVGD...	3256.6139
83.39	A(15-27)	Complete digest			GNPTVEVELTT...	1415.7161
82.2	A(312-328)	Complete digest			TAGIQIVADDLT...	1754.9474
81.69	A(358-374)	Predicted modifi...		1*Oxidation(+15.994915)A370	AAQDSFAAGW...	1804.8353
80.98	A(141-177)	Predicted modifi...		1*Deamidation(+0.984016);1...	TSPYVLPVPFL...	3753.9499
80.16	A(32-49)	Predicted modifi...		2*Oxidation(+15.994915)A48...	SIVPSGASTGV...	1871.9103
77.12	A(415-436)	Predicted modifi...		1*Deamidation(+0.984016)A...	IEEELGDNAVF...	2441.118
76.95	A(358-374)	Predicted modifi...		2*Oxidation(+15.994915)A37...	AAQDSFAAGW...	1820.8293

Figure 23 *Expanded view of Section 3 of the Compound Compare List*

3. View chromatogram comparison results.

An example of the Chromatogram Compare Results window is shown in Figure 24.

The content of this window is updated automatically when you select different compounds in the Compound Compare List window.

To see the middle chromatogram, in the Chromatogram Compare Results window, click the tool for Mirror mode. 

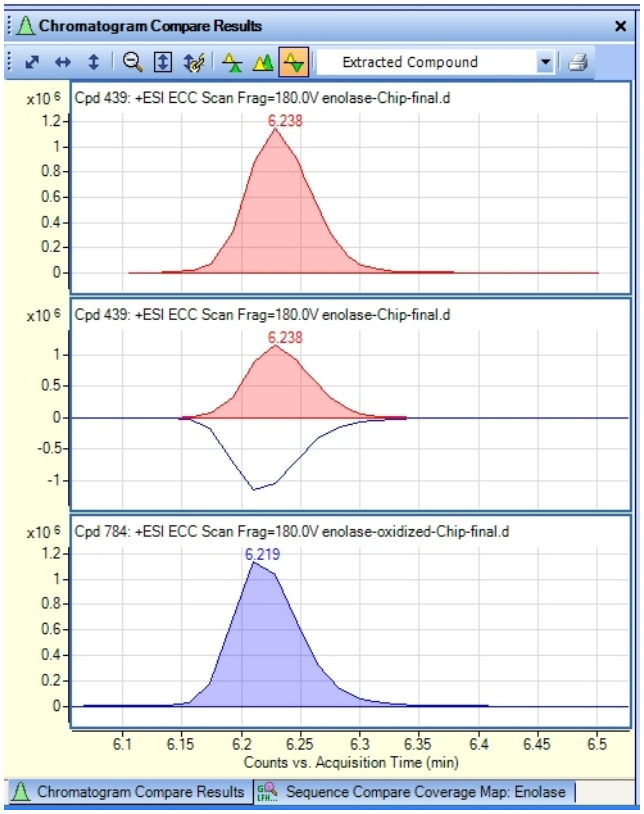


Figure 24 *Chromatogram Compare Results Window*

4. View spectrum comparison results.

Examples of the MS and MS/MS Spectrum Compare Results windows are shown in Figure 25.

The content of these windows is updated automatically when you select different compounds in the Compound Compare List window.

Figure 25 *MS and MS/MS Spectrum Compare Results Windows*

5. View results in other windows.

a Review comparison results in the following other windows:
- Sequence Compare Coverage Map window
- Reference MS Spectrum Peak List window
- Sample MS Spectrum Peak List window

b If the window of interest is not currently visible, select it from the **View** menu.

c See *Qualitative Analysis online Help* for more information on the content of these windows.

www.agilent.com

Printed in USA, January 2016

5990-7064EN

Agilent Technologies