Agilent Protein and Peptide Software

Quick Start Guide

This guide has instructions for installing the Agilent Protein and Peptide Software. It then guides you through the following steps of verifying the primary sequence of chicken lysozyme to illustrate how the software may be used:

- Set up the theoretical data file, which includes entering the expected sequence, adding disulfide bonds, simulating chemical reduction and alkylation, and simulating the tryptic digest.
- Compare (mass match) this theoretical data file with a results file from a previous run and review the results.

When you are done with this example, you will be more familiar with the software and some of its capabilities.

Where to find more information

Use the online help for in-depth information not given in this *Quick Start Guide*. Display online help in any of the following ways:

- Click the **Help** button in the toolbar.
- Select **Help Topics** from the **Help** menu.
- Click the **Help** button on most dialog boxes to display task-specific help.

Agilent Technologies

Introduction to Agilent Protein and Peptide Software

The protein and peptide software (Peptide Tools) is part of the Agilent G2720 Bioanalysis software. It is used with Agilent 1100 Series LC/MSD Systems, Agilent 1200 Series LC/MSD Systems, and Agilent 6100 Series Single Quad LC/MS Systems.

The protein and peptide software performs sophisticated biochemical calculations. It can automatically calculate most abundant and average molecular weights of proteins, peptides, and other biomolecules.

You can enter the amino-acid sequence of any protein or peptide in the main screen and save as a peptide tools data (.PTD) file for future use. This file can then be used to process many sample runs. The software also allows you to create amino-acid modifications and the different types and locations of links and bridges in a sequence.

Agilent Protein and Peptide Software can also be used to simulate protein digestion. This simulation creates theoretical data that can then be compared to the results of an actual analysis. For example, the software can rapidly correlate mass spectral data from a peptide map analysis with theoretical peptide digest fragments, and show differences resulting from posttranslational modifications, DNA point mutations, or incomplete digestion.

The comparison of theoretical data against actual experimental data is called mass matching. Mass matching can be performed interactively or it can be automated as part of a ChemStation method.

The Isotope Distribution view of the software can be used to graphically display the theoretical distribution of the isotopes of a sequence chain or fragment. It will also predict the possible charge states and m/z values based on the building block composition.

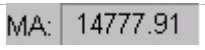
The CID (collision-induced dissociation) view in the software helps you confirm the presence and identities of peptides analyzed by mass spectrometry. The CID view can be used to predict the fragment ions that may be produced by CID of a selected digest fragment or sequence chain in the mass spectrometer. You can then compare the predicted fragment ions with the fragment ions actually observed. The CID view works well as a tool for confirmation.

Menus and Toolbar

The following menus options are available in Agilent Protein and Peptide Software:

File Menu	Edit Menu	Operations Menu	Options Menu	View Menu
New Open Save Save As --- Print Print Setup Page Setup --- Load LCMS Results --- Exit --- Recently Opened File List	Undo Cut Copy Paste Delete --- Links Modifications Building Blocks Chain Names Sample Info --- Find --- Discard Chain New Chain	Digest Global Modifications Find Matches Search for Mass	Sequence Display --- Automatic Calculations Calculate Now --- Insert from Comp Table	Sequence Chain Digest Fragments Summary of All Chains --- Show MS Peak Data Isotope Distribution CID

The following items are available on the Agilent Protein and Peptide Software toolbar:

Button	Equivalent Menu Item	Button or Information	Equiv. Menu Item or Description
	File > Open		Edit > Find
	File > Save		Subsequence Search
	File > Print		Operations > Digest
	File > Load LCMS Results		View > CID
	Edit > Undo		Avergage molecular weight
	Edit > Cut		Options > Calculate Now
	Edit > Copy		Most abundant molecular weight
	Edit > Paste		Basis for mass calculation*
	*S = Sequence chain, BB = Building block composition table, MF = Molecular formula, T = Total of all sequence chains, H = Highlight section of sequence chain		

To install Agilent Protein and Peptide Software

If the Agilent G2720 Bioanalysis software is not already installed on your system, install it as follows.

1. Verify that ChemStation software is installed before you installing Agilent Protein and Peptide Software.
2. Insert the CD-ROM labeled G2720 into the CDROM drive.
3. Click the **Start** button and select **Run**.
4. Type **x:\setup**, where **x** is the letter of the CDROM drive, and click **OK**.
5. Follow the instructions on the screen to install the software.

Start Agilent Protein and Peptide Software

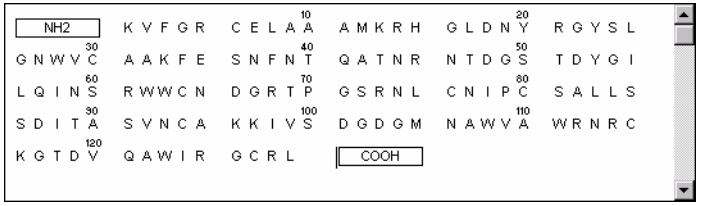
1. Switch to the **Data Analysis** view in ChemStation software.
2. Select **Peptide Tools** from the Spectra menu.

Verifying the primary sequence of chicken lysozyme

The example of verifying the primary sequence of chicken lysozyme contains the following steps, which are described on the indicated pages of this guide:

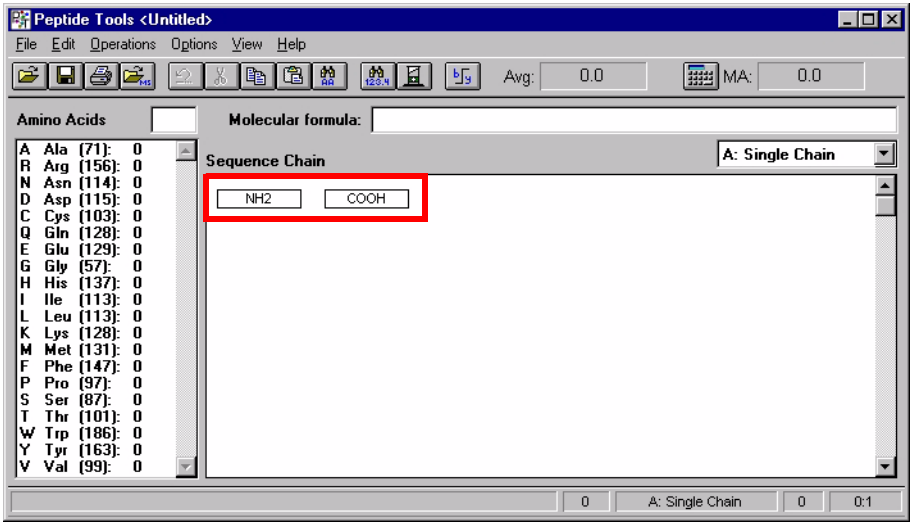
Enter the sequence

The expected sequence of amino acids for chicken lysozyme is as follows:



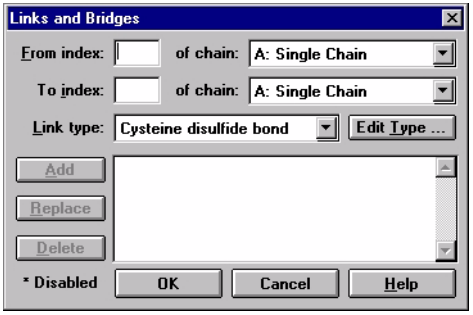
Enter the sequence as follows:

1. Turn **Caps Lock** on (the sequence must be entered in capital letters).
2. Click between the two boxes in the Sequence Chain text box and begin to type each letter of the sequence.



Add disulfide bonds

1 Select **Links** from the Edit menu to open the Links and Bridges dialog box.



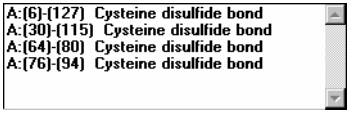
2 Enter **6** in the **From index** field.

3 Enter **127** in the **To index** field.

4 Select **Cysteine disulfide bond** for the Link type.

5 Click the **Add** button.

6 Repeat steps 2 through 5 for each link in chicken lysozyme (refer to the figure shown below).



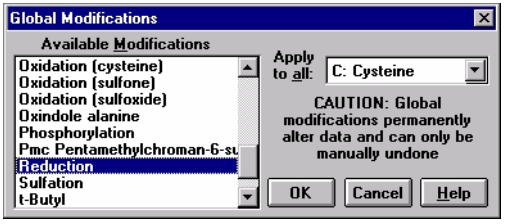
7 Click **OK**.

Tip Use the **Edit Type** button to create a custom link or bridge.

Simulate chemical reduction of disulfide bonds

1 Select **Global Modifications** from the Operations menu to open the Global Modifications dialog box.

2 Select **Reduction** from the list of available modifications, as shown below.

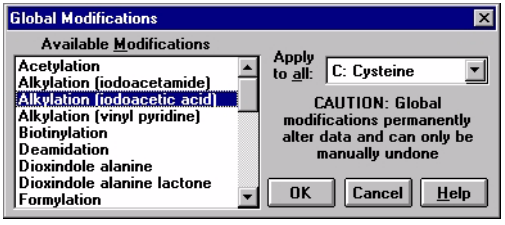


3 Click **OK**.

Simulate alkylation

1 Select **Global Modifications** from the Operations menu to open the Global Modifications dialog box.

2 Select **Alkylation (iodoacetic acid)** from the list of available modifications, as shown below.

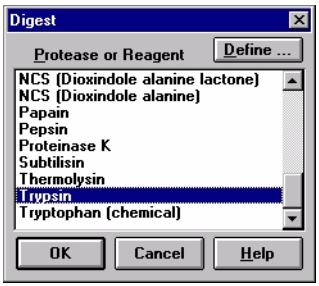


3 Click **OK**.

Simulate the tryptic digest and save the file

1 Select **Digest** from the Operations menu to open the Digest dialog box.

2 Select **Trypsin** from the list as shown below.



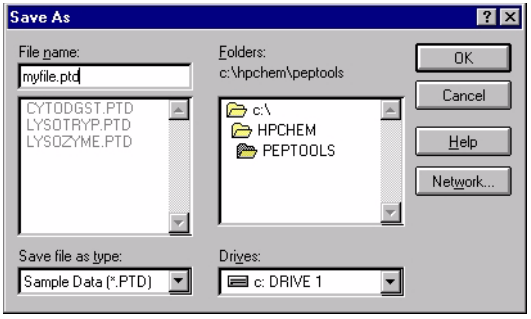
3 Click **OK**.

Tip Use the **Define** button to create a custom digest.

Save the file

1 Select **Save** from the File menu.

2 Type **myfile.ptd** for **File Name** as shown below. Be careful not to overwrite the example files shipped with the software.



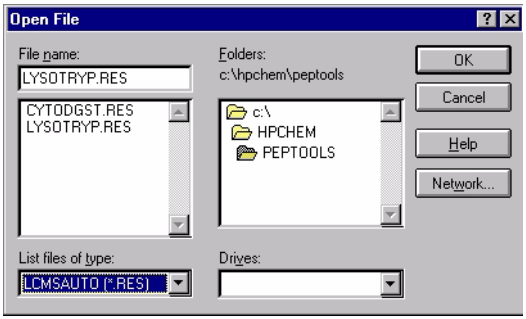
3 Click **OK**.

Load the results file

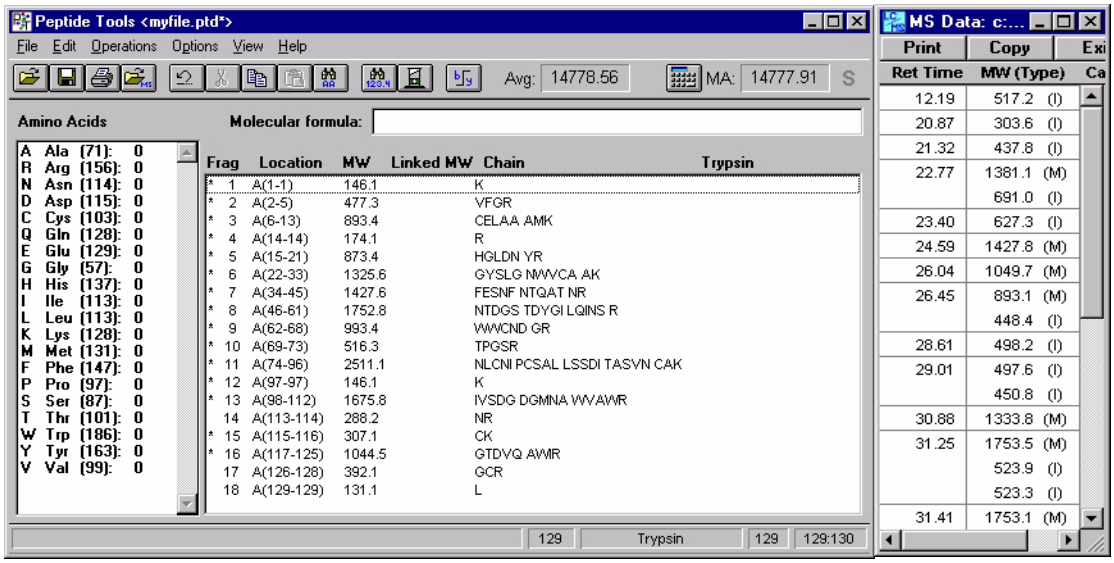
In this example, you will be using a results file that was shipped with the software. In your own experiment, you would run the actual analysis now and then generate the results file.

1 Select **Load LCMS Results** from the File menu.

2 Select the file **lysotryp.res** as shown below.



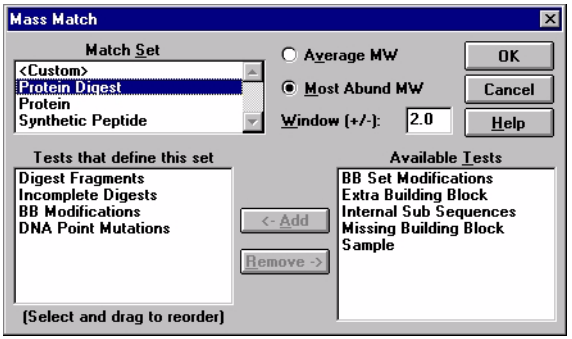
3 Click **OK**. Your screen should look similar to the example below. The theoretical data file you created is displayed on the left; the experimental results file on the right.


Peptide Tools <myfile.ptd*>
File Edit Operations Options View Help
Avg: 14778.56
MA: 14777.91
Amino Acids
Molecular formula:
A Ala (71): 0
R Arg (156): 0
N Asn (114): 0
D Asp (115): 0
C Cys (103): 0
Q Gln (128): 0
E Glu (129): 0
G Gly (57): 0
H His (137): 0
I Ile (113): 0
L Leu (113): 0
K Lys (128): 0
M Met (131): 0
F Phe (147): 0
P Pro (97): 0
S Ser (87): 0
T Thr (101): 0
W Trp (186): 0
Y Tyr (163): 0
V Val (99): 0
Frag Location MW Linked MW Chain Trypsin
* 1 A(1-1) 146.1 K
* 2 A(2-5) 477.3 VFGR
* 3 A(6-13) 893.4 CELAA AMK
* 4 A(14-14) 174.1 R
* 5 A(15-21) 873.4 HGLDN YR
* 6 A(22-33) 1325.6 GYSLG NWVCA AK
* 7 A(34-45) 1427.6 FESNF NTQAT NR
* 8 A(46-61) 1752.8 NTDGS TDYGI LQINS R
* 9 A(62-68) 993.4 WWCND GR
* 10 A(69-73) 516.3 TPGSR
* 11 A(74-96) 2511.1 NLCNI PCSAL LSSDI TASVN CAK
* 12 A(97-97) 146.1 K
* 13 A(98-112) 1675.8 IVSDG DGMNA WVAWR
14 A(113-114) 288.2 NR
* 15 A(115-116) 307.1 CK
* 16 A(117-125) 1044.5 GTDVQ AWIR
17 A(126-128) 392.1 GCR
18 A(129-129) 131.1 L
129 Trypsin 129 129:130
MS Data: c:...
Print Copy Exi
Ret Time MW (Type) Ca
12.19 517.2 (I)
20.87 303.6 (I)
21.32 437.8 (I)
22.77 1381.1 (M)
691.0 (I)
23.40 627.3 (I)
24.59 1427.8 (M)
26.04 1049.7 (M)
26.45 893.1 (M)
448.4 (I)
28.61 498.2 (I)
29.01 497.6 (I)
450.8 (I)
30.88 1333.8 (M)
31.25 1753.5 (M)
523.9 (I)
523.3 (I)
31.41 1753.1 (M)

Find the matches

1 Select **Find Matches** from the Operations menu.

2 Select **Protein Digest** from the **Match Set** list as shown below.



These are the default tests and parameters used for Protein Digest matching:

Digest Fragments Masses are compared to the theoretical digest fragments in the .PTD file.

Incomplete Digest Masses are compared to fragment masses, assuming that the digest was not complete and that contiguous fragments might still exist.

DNA Point Mutations Masses of existing fragments are adjusted, assuming the possibility that single DNA point mutations could have occurred. For example, an error in the DNA triplet that encodes for methionine (AUG) could result in arginine (AGG) being encoded instead.

BB Modifications Masses of existing fragments are adjusted by any modifications that may have been previously assigned to specific building blocks (amino acids), and are then compared. In this example, there were no modifications assigned.

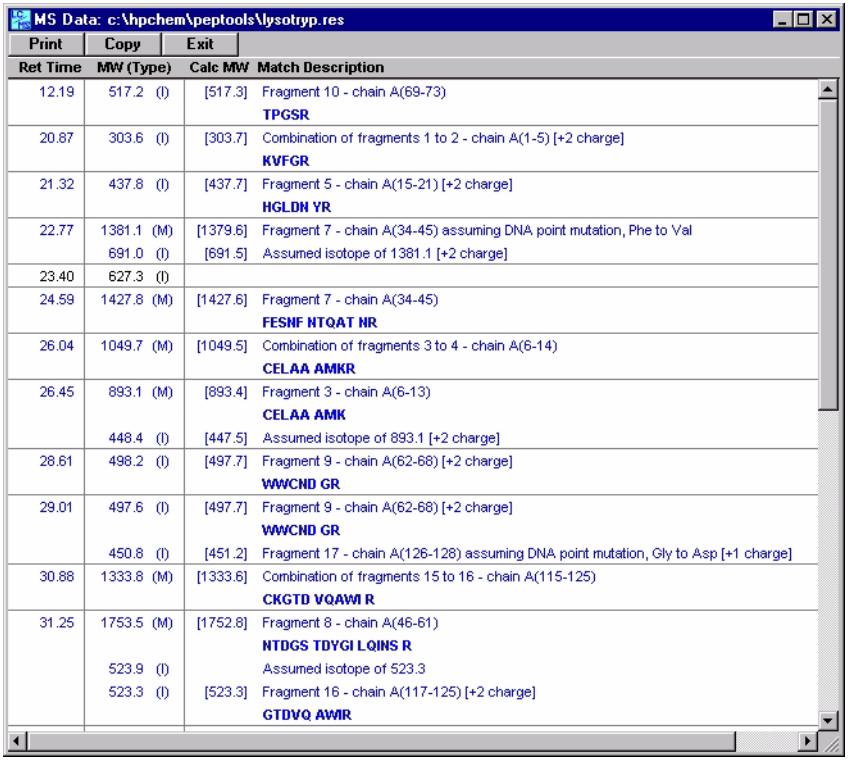
Tip You can customize these parameters and then resave them in the .PTD file for future use.

3 Click **OK**. A mass match is performed and the data displayed in the results window is updated.

- If multiple matches are found within a given test, the closest match (the one with the minimum mass difference) is used.
- Masses that have been matched are not considered in subsequent tests.

Review the report

1. Resize the results window as necessary to view the data.
2. Review the results of the mass matching as shown in the example below. The first two columns display the retention times and molecular weights of the experimental data. The last column displays the calculated molecular weights of the theoretical data and the proposed matches. If no match is found, this field is left blank.

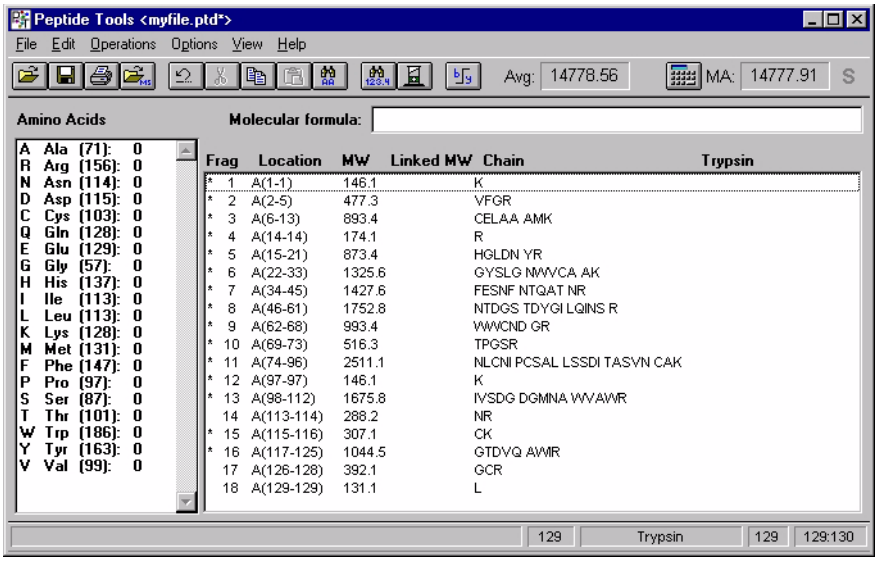


Ret Time	MW (Type)	Calc MW	Match Description
12.19	517.2 (I)	[517.3]	Fragment 10 - chain A(69-73)
			TPGSR
20.87	303.6 (I)	[303.7]	Combination of fragments 1 to 2 - chain A(1-5) [+2 charge]
			KVFGR
21.32	437.8 (I)	[437.7]	Fragment 5 - chain A(15-21) [+2 charge]
			HGLDN YR
22.77	1381.1 (M)	[1379.6]	Fragment 7 - chain A(34-45) assuming DNA point mutation, Phe to Val
	691.0 (I)	[691.5]	Assumed isotope of 1381.1 [+2 charge]
23.40	627.3 (I)		
24.59	1427.8 (M)	[1427.6]	Fragment 7 - chain A(34-45)
			FESNF NTQAT NR
26.04	1049.7 (M)	[1049.5]	Combination of fragments 3 to 4 - chain A(6-14)
			CELAA AMKR
26.45	893.1 (M)	[893.4]	Fragment 3 - chain A(6-13)
			CELAA AMK
	448.4 (I)	[447.5]	Assumed isotope of 893.1 [+2 charge]
28.61	498.2 (I)	[497.7]	Fragment 9 - chain A(62-68) [+2 charge]
			WWCND GR
29.01	497.6 (I)	[497.7]	Fragment 9 - chain A(62-68) [+2 charge]
			WWCND GR
	450.8 (I)	[451.2]	Fragment 17 - chain A(126-128) assuming DNA point mutation, Gly to Asp [+1 charge]
30.88	1333.8 (M)	[1333.6]	Combination of fragments 15 to 16 - chain A(115-125)
			CKGTD VQAWI R
31.25	1753.5 (M)	[1752.8]	Fragment 8 - chain A(46-61)
			NTDGS TDYGI LQINS R
	523.9 (I)		Assumed isotope of 523.3
	523.3 (I)	[523.3]	Fragment 16 - chain A(117-125) [+2 charge]
			GTDVQ AWIR

3. (optional) Use the **Print** button to print a copy of the report.
4. (optional) Use the **Copy** button to copy the report to the clipboard.

Review digest fragments

1 Select **Digest Fragments** from the View menu.



An asterisk (*) next to a fragment in the list indicates that the fragment was explained (i.e. a match was found). If *no* asterisk appears next to a fragment, then no data was found for that fragment (i.e. the data is unexplained).

A few possible reasons no match was found are:

- Data is outside the scan range
- Missing peptide eluted in the dead volume
- It's not in the protein due to mutations or modifications

2 Resolve any unexplained data or missing peptides. For example, you might want to redigest using an alternative enzyme.

www.agilent.com

In this book

This guide has instructions for installing the Agilent Protein and Peptide Software. It then guides you through the steps of verifying the primary sequence of chicken lysozyme to illustrate how the software may be used.

Printed in USA
Second Edition, July 2007

G2720-90004

Agilent Technologies