Peptide and Protein Identification

The process of searching mass spectral data for the purpose of peptide and protein identification can roughly be divided into six steps:

* **Step 1: Convert the raw, typically binary, output from the MS instrument into open formats.**
* **Step 2: Process the MS/MS spectra into peak lists.**
* **Step 3: Retrieve the desired sequence database and adapt it to the identification strategy.**
* **Step 4: Search the peak lists against a sequence database using one or more search engines.**
* **Step 5: Identify the peptides and infer the proteins.**
* **Step 6: Validate the detected peptides and proteins.**

**(1) Convert   
Raw Files**

**(3) Retrieve Database**

**(2) Process MS/MS Spectra**

**(4) Match Peptides   
to Spectra**

**(5) Infer Peptides and  
Proteins**

**(6) Validate Peptides and Proteins**

This chapter will guide you through all these steps, separated into five sections:

**1.1 Database Generation**

**1.2 Peak List Generation**

**1.3 Peptide to Spectrum Matching**

**1.4 Browsing Identification Results**

**1.5 Peptide and Protein Validation**