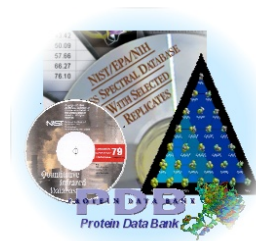


# MASS SPECTRAL REFERENCE LIBRARY OF PEPTIDES



**BACKGROUND** The NIST mass spectral library, NIST 2005, contains nearly 200,000 evaluated reference spectra and has long been routinely used for automated chemical identification using the method of GC/MS (gas chromatography/mass spectrometry). A current focus in the Physical and Chemical Properties Division of NIST is to develop a similar reference library for peptide identification using newly developed methods of LC/MS-MS (liquid chromatography/tandem mass spectrometry). Using a reference library of tandem mass spectra for peptide identification will have distinct advantages over current methods which use “sequence libraries.” These sequence libraries contain proteins with predicted sites of enzymatic cleavage and predicted m/z ratios of the peptide fragments. Because searching is done, in effect, against a very coarse theoretical spectrum, and proteins that cannot be present are also considered, use of sequence libraries tends to result in many false-positive identifications. Also, peptides identified in the sample can depend greatly on the specific search engine employed as well as post-processing steps. On the other hand, using a library of reliable, complete peptide spectra will improve the accuracy of identifications and speed of processing.

**PURPOSE** This project was initiated in 2004 to improve the reliability of peptides identified from complex mixtures using the NIST mass spectral library.

**DESCRIPTION** An infrastructure for automated processing and annotating peptide MS/MS spectra is being created to provide a library of high quality peptide spectra with reliable peptide assignments. Since spectral libraries are not currently used in peptide identification, software is being developed to process and match spectra. The reliability of peptide assignments in the library is being determined by studying peptide fragmentation patterns to distinguish between unusual fragmentation and incorrect assignments. Digests of pure proteins are also being studied to identify all enzymatic and chemical cleavages that occur. After developing measures of reliability of peptide assignments, further testing and refinement of the library will be done using complex mixtures of peptides obtained from collaboration with others at Hollings Marine Lab. Peptide spectra for protein standards, protein contaminants, and proteins of interest in marine organisms are being collected.

**Sources of spectra** are:

- Gaithersburg, MD (triple quadrupole)
- Hollings Marine Lab, Charleston, SC (linear ion trap)
- On-line spectral repositories (Peptide Atlas, The GPM, Open Proteomics Database)

Please contact us at the addresses given below if you would like to donate spectra.

**CONTACTS** **Lisa Kilpatrick**, NIST, HML, Charleston, SC; *tel:* 843-762-8835; *email:* [lisa.kilpatrick@nist.gov](mailto:lisa.kilpatrick@nist.gov)  
**Steve Stein**, NIST, Gaithersburg, MD; *tel:* 301-975-2505; *email:* [steve.stein@nist.gov](mailto:steve.stein@nist.gov)