

Libraries in MSWS

There are a variety of libraries available for different applications. SCION Instruments offers various libraries including NIST, Wiley and Pfleger/Maurer/Weber (PMW) with customisable libraries and the ability to automatically search in multiple libraries. The most commonly used library is NIST. For the purpose of this technical note, NIST will be used as the example but the principles of all libraries is the same, the main difference is that the compounds available varies between libraries.

National Institute of Standards and Technology (NIST) Library

The National Institute of Standards and Technology (NIST) is one of the oldest physical science laboratories. The organization collected Electron Ionization (EI) Gas Chromatography (GC) mass spectral data of known standards from various sources to create a mass spectral reference library of compounds. The NIST library is used all over the world for the identification of unknown compounds from GC-MS chromatograms. SCION Instruments' software MS Workstation (MSWS), which is used for analysis using MS, is also compatible with NIST.

Identifying the unknowns

The NIST library will examine the peaks in the chromatogram. The NIST library works as follows; open the chromatogram in the software (MSWS), then click on the peak of interest in the chromatogram, click on the NIST button on top of the data review tab. The NIST library will now open and search for corresponding mass spectra. An example is shown below from the compound acenaphthylene (Figure 1). The red spectrum is from the sample run on the MS and the blue is from acenaphthylene from the NIST library.

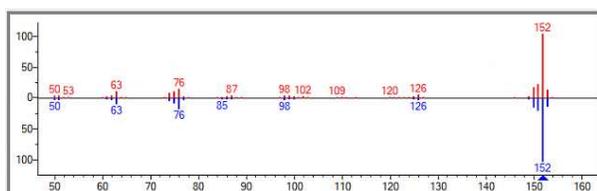


Figure 1 NIST mass spectrum comparison

The software calculates several numbers associated with each compound. The three mainly used are: Match Factor (Match), Reverse Match Factor (R. Match) and Probability (%). A list of different compounds will be displayed and they are ranked based on how well their reference mass spectra matches the unknown spectrum. The compound with the highest spectrum similarity is at top of the list and deemed most likely to be the correct identification of the unknown compound. In this case that is correct (Figure 2).

Name	R.Mat...	F.Match	Prob.
Acenaphthylene	927	927	32.49
Acenaphthalene	918	918	32.49
1-acenaphthenol, trifluoroacetate ester	905	905	12.84

Figure 2 list of matches

Match Factor (Match)

The "Match Factor" or "direct match" compares the unknown mass spectrum peaks to the peaks in the known library spectrum. The NIST library uses a scale of 0-1000 to score the similarity, no score of any compound can exceed 999. The score of 999 means a perfect match to all the peaks within the spectra. A score of 0 indicates that there are no peaks that match. The guidelines for the NIST Match Factor score are: >900 excellent match, 800-900 good match, 700-800 a fair match and <600 is a poor match.

Reverse Match Factor (R. Match)

The "R. Match" is the match score between the unknown compound spectrum and the library spectrum but it ignores any peaks that are in the unknown spectrum that are not in the library spectrum. This number is particularly useful when two compounds are co-eluting. When the Match Factor is low then the R. Match is leading.

Probability

The "Probability" first assumes that the unknown spectrum is present in the library. With this assumption, NIST then looks at the hit list which is created from the match factors. The hit list ranks the possible matches by calculating the probability. Probability calculated by comparing the hits to each other. NIST pays attention to the differences between these hits and from those differences it estimates how likely it is that any of the hits in the list is the correct identification.

Low Numbers but Correct Compound

The Match Factor scores are dependent on the number of peaks in the unknown spectra. The more peaks there are in the unknown spectrum, the more peaks that must align with the reference spectrum to achieve a high match factor. High noise and co-eluting of compounds can cause this effect. In these situations the reverse match factor may still remain high because it ignores peaks that do not match.

With probability, if the unknown spectrum closely resembles only a few spectra in the library, the probability will be higher and if it resembles many different spectra the probability will be lower.

If there is still uncertainty over that the correct compound has been identified then an analytical standard should be purchased and analyzed by GC-MS.