

**Agilent G1676AA
Fiehn GC/MS
Metabolomics RTL
Library**

User Guide



Agilent Technologies

Notices

© Agilent Technologies, Inc. 2013

No part of this manual may be reproduced in any form or by any means (including electronic storage and retrieval or translation into a foreign language) without prior agreement and written consent from Agilent Technologies, Inc. as governed by United States and international copyright laws.

Manual Part Number

G1676-90001

Edition

Revision A, July 2013

Printed in USA

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

Microsoft[®] is a U.S. registered trademark of Microsoft Corporation.

Software Revision

This guide is valid for A.02.xx revisions of the Agilent G1676AA Agilent Fiehn GC/MS Metabolomics RTL Library software, where xx refers to minor revisions of the software that do not affect the technical accuracy of this guide.

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 Legend

U.S. Government Restricted Rights. Software and technical data rights granted to the federal government include only those rights customarily provided to end user customers. Agilent provides this customary commercial license in Software and technical data pursuant to FAR 12.211 (Technical Data) and 12.212 (Computer Software) and, for the Department of Defense, DFARS 252.227-7015 (Technical Data - Commercial Items) and DFARS 227.7202-3 (Rights in Commercial Computer Software or Computer Software Documentation).

Safety Notices

CAUTION

A **CAUTION** 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 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.

In This Guide...

This guide contains information on the use of the Agilent Fiehn GC/MS Metabolomics RTL Library.

1 Introduction

Describes the scope and intended use of the library; the contents of the library, including the file names and location on the user's hard drive; and a discussion on the workflow of metabolic profiles is examined.

2 Sample Preparation and Derivatization

Reviews the preparation of the analytical sample, starting with a biological extract, tagged with internal standards and markers, processed to be compatible with GC analysis, and ending with a vial of the sample ready for GC/MS analysis.

3 Experimental and Instrument Method

Lists the analytical method to follow when setting up your Agilent GC/MS instrument system for successful use of the library.

4 Peak Identification

Explores how peaks are identified in the library, the compound naming convention that allows RT identification, and how to run an analysis using the Agilent GC/MS ChemStation and AMDIS.

5 Analyzing your Data

Guides you through the use of Agilent GC/MS ChemStation, AMDIS, Agilent MassHunter Qualitative Analysis, and Agilent Mass Profiler Professional in performing metabolomics GC/MS experiments.

6 AMDIS Getting Started

Introduces AMDIS and is taken directly from Chapter 2 of the AMDIS online help and is presented here with the

author's permission. Agilent has found that many AMDIS users who were having trouble with the program's concepts were able to master the program once they became aware of this help. Agilent thinks that this material is important enough to repeat it in this manual.

7 Using CAS and PubChem Numbers

Reviews investigative data found on popular Web sites for identifying metabolites using the CAS and PubChem numbers found in the library.

8 Reference Information

This chapter consists of definitions, derivatization instructions, and references. The definitions section includes a list of terms and their definitions as used in this workflow. The derivatization section contains instructions for derivatizing d₂₇-myristic acid to update the GC/MS retention time locking. The references section includes citations to Agilent publications that help you use Agilent products and perform your metabolomics analyses.

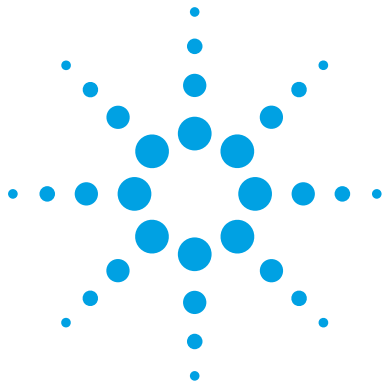
Contents

1	Introduction	9
	Scope and Intended Use	10
	Library Content	11
	User guide	11
	AMDIS program	12
	Agilent Fiehn Library metabolite list	12
	GC/MS ChemStation files	12
	AMDIS files	13
	Application Notes	14
	Backup files for the library	14
	Metabolic Profiling Workflow	15
	GC/MS analysis	16
	Peak identification	16
	Metabolite identification	17
	Statistical analysis	17
	Systems biology analysis	18
2	Sample Preparation and Derivatization	21
	Biological sample extracts	22
	Metabolites	22
	Derivatization	22
3	Experimental and Instrument Method	31
	Experimental	32
	Acquisition Method	35
	GC/MS Setup	38
	Quality Control	42
4	Peak Identification	45
	Peak Identifications from Metabolite Profiles	46

Compound Naming in the Agilent Fiehn Library	48
Metabolite Derivatizations Resulting in Multiple Compounds	50
5 Analyzing your Data	53
Using Agilent GC/MS ChemStation	54
PBM Quick Search	54
Library search report	56
Screen report	57
Search basics	58
Using AMDIS	59
Data analysis	60
Using a different analytical method	62
Using Agilent MassHunter Qualitative Analysis	63
Install the Fiehn library to use with MassHunter Qualitative Analysis	63
Search Unit Mass Library on spectra	64
Search Unit Mass Library with Retention Time Matching on compounds	68
Export results to Mass Profiler Professional using CEF file	73
Using Agilent MassHunter Mass Profiler Professional	75
Overview of a metabolomics experiment using MPP	75
Elements to consider in planning your experiment	77
Create a new project and experiment	79
Import and organize your data	84
Perform your initial analysis	99
Save your project	113
Functional areas of MPP	114
6 AMDIS Getting Started	115
Running AMDIS	116
To simultaneously view AMDIS and its Help window	116
To put AMDIS into an initialized state	116
Confirm Window and Results Window	118

To move between windows	118
Confirm Window Appearance	119
Results Window Appearance	126
Analysis Types	128
Tutorials Overview	130
Analysis Type - Simple Analysis Type	132
Analysis Type - Use Retention Time	141
Analysis Type - RI Calibration/Performance	148
Analysis Type - Use Retention Index Data	156
Analysis Type - Use Internal Standards for RI	163
Analysis Type - Use RI Calibr. Data + Internal Std.	170
Analysis Type - Performance Check for RI	176
Recap	183
Recommended Settings	183
7 Using CAS and PubChem Numbers	191
KEGG LIGAND Database	193
MetaCyc Database	196
ChemSpider Database	199
SciFinder Substance Identifier Database	202
PubChem Compound Database	203
8 Reference Information	207
Definitions	208
Derivatization Instructions	220
References	221

Contents



1 Introduction

Scope and Intended Use	10
Library Content	11
Metabolic Profiling Workflow	15



Scope and Intended Use

The Agilent Fiehn 2013 GC/MS Metabolomics RTL Library is intended to be part of a solution for laboratories doing metabolomics research using gas chromatography-mass spectrometry (GC/MS).

The primary purpose of the library is to identify metabolites commonly found in metabolomic studies. The library of electron impact (EI) spectra and retention data of pure chemical standards is built for the identification of compounds in biological matrices. It is compiled mainly from a collection of commercially available metabolites that are comprised in a variety of biochemical databases. The library intends to aid chemists and biologists in identifying as many metabolites as possible from GC/MS metabolite profiling screens.

This library is for use as a primary entry tool to metabolite profiling by GC/MS. It is not meant to comprehensively cover any specific part of known metabolic pathways or any specific organism. It does include key metabolites from a diverse set of biochemical pathways that are known to be conserved across organisms. The metabolites included are structurally diverse, and allow detection by GC/MS. Large and thermolabile metabolites such as acetyl-CoA, NADH, or ATP are excluded.

The derivatization procedure for most primary metabolites in this library involves drying samples. This drying process excludes volatile metabolites such as aroma compounds.

Library Content

The Agilent Fiehn 2013 GC/MS Metabolomics RTL Library includes files required by the Agilent GC/MS ChemStation, AMDIS, and Agilent MassHunter Qualitative Analysis for finding the metabolites in a GC/MS data file. This manual and useful application notes are also included. The files and programs listed in the following sections are installed on your computer in the specified directories.

Data processed by AMDIS using the Agilent Fiehn Library results in two output files that can be used for statistical analysis: elute data file (.elu) and find compound data file (.fin). The elute file contains a detailed ion and abundance data listing for all of the peaks (features) found by AMDIS. The find compound file includes all of the information contained in the elute file and additionally contains the Agilent Fiehn Library search assignment for each peak (feature).

Data files from single quadrupole MSDs (mass selective detectors) may also be processed using Agilent MassHunter Qualitative Analysis software. Qualitative Analysis generates a single output compound exchange format (CEF) file (.cef) for each data file.

After you have found the features (metabolites) in your GC/MS data, you can perform statistical and integrated biology analyses using Mass Profiler Professional. Mass Profiler Professional imports the files created by AMDIS and Qualitative Analysis.

User guide

Its purpose is to guide the user of this library application to successfully identify metabolites using either the Agilent GC/MS ChemStation and the NIST AMDIS programs or the Agilent MassHunter Qualitative Analysis software. Then, you can use the Agilent Mass Profiler Professional program with results from either AMDIS or Agilent MassHunter Qualitative Analysis software. The user is shown how to use the library, set up the GC/MS system for retention time repeatability, and run metabolite

identification reports using these programs. This user guide is installed in the **Program Files (x86)\Agilent\Fiehn GCMS Metabolomics RTL Database\App Notes** directory on the application CD.

AMDIS program

The AMDIS program is installed in the **\Program Files (x86)\AMDIS32** directory unless you change the default directory during the installation procedure.

Agilent Fiehn Library metabolite list

A listing of the metabolites contained in the library files is included on the CD application disk in the **\Program Files (x86)\Agilent\Fiehn GCMS Metabolomics RTL Database\App Notes** directory.

GC/MS ChemStation files

The library locates the following GC/MS ChemStation method file in the **MSDCHEM\1\Methods** directory.

Fiehn.m	ChemStation RTL Chromatographic Method
---------	--

The library locates the following GC/MS ChemStation files in the **Database** directory.

Fiehn.l	ChemStation PBM Library
Fiehn.scd	ChemStation screener file

To use this library with the MassHunter Qualitative Analysis software, see “Install the Fiehn library to use with MassHunter Qualitative Analysis” on page 63.

The library locates the following GC/MS ChemStation data files in the **MSDCHEM\1\DATA** directory.

RI-calibration.d	Contains all retention index standards.
L-Histidine.d	Example data file containing derivatized L-Histidine
FAMES-2COL-SPLIT01.d	Example data file containing FAME markers but using a nonstandard method

To use these data files with the MassHunter Qualitative Analysis software, see “Install the Fiehn library to use with MassHunter Qualitative Analysis” on page 63.

AMDIS files

The library locates the following AMDIS files in the **AMDIS32\LIB** directory.

Fiehn.msl	Library file for scan analyses
Fiehn_SIM.msl	Library file for SIM analyses, contains the 4 largest m/z for each compound
Fiehn.isl	Internal standards file
Fiehn.cal	Retention index calibration file
Fiehn.csl	Retention index standards file
Fiehn.msp	Mass spectral transfer file
Fiehn.cid	Compound information file

The library locates the following GC/MS data files used for the demonstration of the AMDIS program in the **AMDIS32\DATA\HP** directory.

RI-calibration.d	Contains all retention index standards; also contains a small amount of benzoic acid
L-Histidine.d	Example data file containing derivatized L-Histidine
FAMES-2COL-SPLIT01.d	Example data file containing FAME markers but using a nonstandard method

To use these data files with the MassHunter Qualitative Analysis software, see “Install the Fiehn library to use with MassHunter Qualitative Analysis” on page 63.

Application Notes

The library locates the following application notes in the **Program Files (x86)\Agilent\Fiehn Metabolomics RTL Database (G1676AA)\App Notes** directory.

MSD RTL 5968-3433.pdf	Description of RTL method creation and maintenance
NIST-MS_AMDIS.pdf	Discussion of the algorithms employed by the NIST and AMDIS search programs

Backup files for the library

All of the files mentioned above are also located in **Program Files (x86)\Agilent\Fiehn GCMS Metabolomics RTL Database** directory. If you need to replace one of the working files in AMDIS or the GC/MS ChemStation with an original file delivered with this library, copy the file from this location to your working directory.

Metabolic Profiling Workflow

Metabolomics is the systematic study of the small-molecule chemicals (metabolites) used in the biological processes of an organism. A study is typically done by looking at the chemical profile of an organism using either nuclear magnetic resonance (NMR) or mass spectrometry (MS) for the chemical analysis. This study is often referred to as metabolic profiling.

The metabolome is the collection of all metabolites in a biological organism, and the metabolites are the end products of its gene expression. Metabolic profiling can give a snapshot of the physiology of that organism.

Metabolic profiling by GC/MS follows a workflow that is graphically represented in Figure 1. Experiment design and sample preparation are very important to the analysis but are outside the discussion of this workflow analysis.

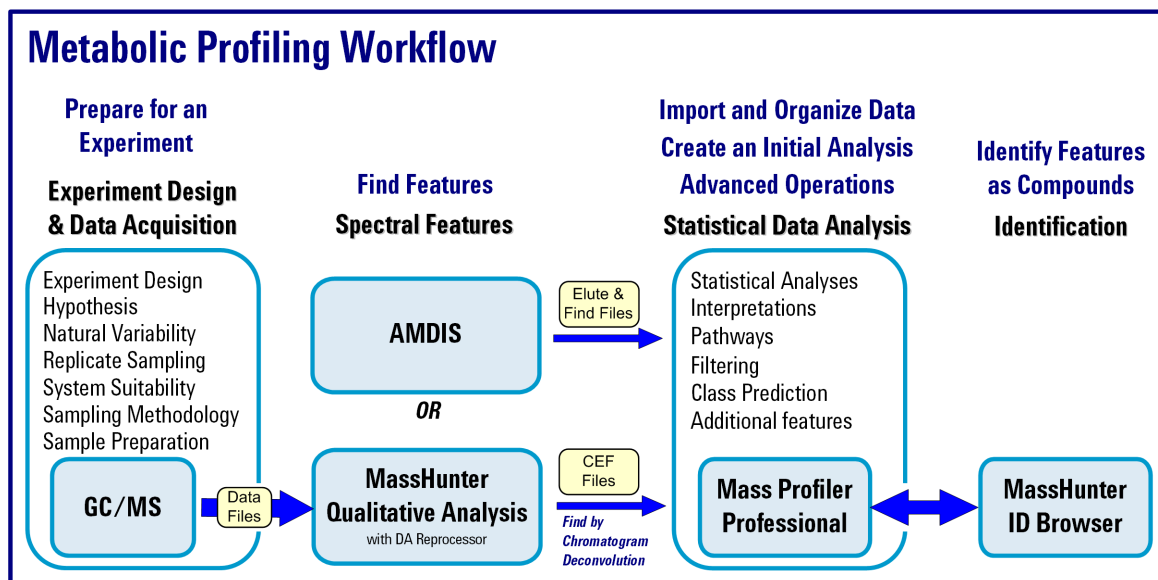


Figure 1 Metabolic profiling workflow for GC/MS data

GC/MS analysis

The metabolite sample is first derivatized to allow volatile compound separation by the GC. The derivatized metabolites are then analyzed by the MS.

Derivatization

See “[Derivatization](#)” on page 22 for a detailed discussion of the chemical derivatization used to create the Agilent Fiehn Library.

Sample analysis

The Agilent Fiehn Library is specifically built for use with an Agilent GC/MS system running the included Fiehn method. If the method you are using differs from the Fiehn method, see “[To update the AMDIS *.cal file](#)” on page 58.

Peak identification

GC/MS metabolomics analysis produces three-dimensional data that is normally displayed as a two-dimensional total ion chromatogram (TIC). The higher abundance ions obscure lower level metabolites. To find the lower level metabolites, the three-dimensional data is matched using either a forward or reverse peak finding approach. Forward peak finding first attempts to find all chromatographic peaks in the TIC and then attempts to identify the found peaks to translate chromatographic peaks into identified compounds. Reverse peak finding uses a compound library containing spectral and chromatographic information for specific compounds and searches the TIC data for the presence of each of those library compounds.

Each peak finding approach has its advantages and disadvantages. Reverse peak finding is very good at finding library compounds at low levels but only finds compounds in the library. Forward peak finding find peaks that are not contained in a library but tends not to find low-level compounds (false negatives) and find nonexistent compounds (false positives).

The Agilent Fiehn Library works with both forward and reverse peak finding approaches. The NIST software package AMDIS is used for forward peak finding. For reverse peak finding you can choose between two algorithms in the Agilent ChemStation software: Quant and Screener.

Metabolite identification

Using the Agilent Fiehn Library, metabolite identification is accomplished by EI spectral library matching in combination with either retention time or retention index matching. EI spectral matching alone is not sufficient for a confident identification in the presence of many biologically important isomers. The use of chromatographic retention as an additional identifying parameter results in much higher confidence for EI spectral matching. It does not resolve all identification issues, such as compound chirality. Like most analytical methods, the chromatography was developed with a compromise between complete separation and analysis time.

Statistical analysis

Metabolomics studies are comparative studies of biological organisms. The comparison can be simple, such as the metabolic difference between a cell culture under normal and stressed conditions, or a very complex study of metabolic differences of a human population with many variables such as diet, age, sex, race, etc. In both cases, the Agilent Fiehn Library and methodology can be used to investigate metabolic differences.

Data processed by AMDIS using the Agilent Fiehn Library results in two output files that can be used for statistical analysis: elute data file (.elu) and find compound data file (.fin).

The elute data file contains a list of all of the **peaks** found including ion and abundance. This is a file containing unidentified compounds (components).

The find data file contains a list of all of the **compounds** found including the compound information and assignments. This is a file containing identified compounds (targets).

The Agilent statistical analysis software package for analyzing GC/MS data is called Agilent MassHunter Mass Profiler Professional. Mass Profiler Professional imports the files generated using AMDIS.

The first step in the analysis of multiple data files is to align the found peaks in each data file across all the runs in the study. The software determines which peaks in each chromatographic run are the same compound and which are different. Compound identification prior to peak alignment makes this task much easier for the identified compounds. Identification of compounds using the Agilent Fiehn Library is very helpful in the statistical analysis of data sets.

Once data from all the GC/MS analyses are aligned and normalized, a number of interpretation algorithms can be applied to find the metabolite relationships to the problem being studied. The use of statistics to analyze complex data sets is outside the scope of this guide.

For more information on importing GC/MS data into the Agilent MassHunter Mass Profiler Professional software, please refer to [“Using Agilent MassHunter Mass Profiler Professional”](#) on page 75.

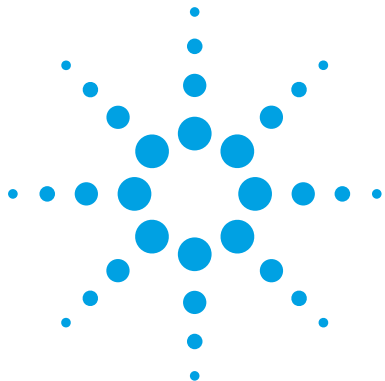
Systems biology analysis

The goal of any metabolomics study is to make sense of the metabolic differences discovered in the analysis. There are many ways to do this. One such approach is to use pathway analysis software to visualize and compare currently known metabolite pathways to highlight the relevant metabolites discovered in the statistical analysis process.

Agilent MassHunter Mass Profiler Professional (MPP) software is a powerful chemometrics platform designed to exploit the high information content of mass spectra data and can be used in any MS-based differential analysis to determine relationships among two or more sample groups and variables. MPP also provides advanced statistical analysis and visualization tools for your experiments and is the only platform that provides integrated identification/annotation of compounds and integrated pathway analysis for metabolomic and proteomic studies. MPP's pathway analysis tools allow you to analyze and compare entity list data collected using different technologies, enabling systems biology analyses.

MPP software supports the export of comma separated variable (CSV) formatted files that contain analysis results and compound identifiers. The compound identifier comes from compound identifications made using the Agilent Fiehn Library. The Agilent Fiehn Library supports CAS numbers in one field and the name and PubChem number in a combined field. A compound identifier is important to pathway analysis as the software needs a means to identify the metabolite to assign it to a pathway.

1 Introduction



2 Sample Preparation and Derivatization

Biological sample extracts 22

Metabolites 22

Derivatization 22



Biological sample extracts

The preparation of optimal extracts from biological samples is very much dependent on the type of matrix analyzed and the class of compound that is of interest for the study. The exact procedures have to be carefully examined and are not part of this guide.

Metabolites

Metabolites are generally small molecules, and most primary metabolites bear hydrophilic functional groups such as carboxyl, hydroxyl or amino groups. According to the functional groups, these molecules are often classified into amino acids, carbohydrates, fatty acids and organic acids. The presence of a variety of hydrophilic functional groups enables cells to use metabolites for a variety of cellular purposes, including transport in the aqueous cellular environment or between compartments and organs. The presence of these functional groups in extracts of biological samples causes a significant rise in boiling points, rendering most primary metabolites unsuitable for GC separations.

Derivatization

The hydrophilic functional groups must be derivatized to remove hydrogen bond formations to increase volatility. This also reduces interaction with the column phase that can cause tailing peaks, poor sensitivity, and poor chromatographic separation. The Agilent Fiehn 2013 GC/MS Metabolomics RTL Library uses a two step derivatization procedure routinely used in most published literature in metabolite profiling by GC/MS.

Trimethylsilylation

Trimethylsilylation is used to exchange acidic protons under mild, universal, and sensitive conditions. These protons are present in carboxyl, hydroxyl, amino, imino, or sulfuryl groups. In rare cases, like malonic acid with two carboxyl groups at 1,3-positions, hydrogen atoms at the

alpha-carbon are acidic enough to undergo trimethylsilylation. Trimethylsilylation dramatically reduces boiling points, improves thermal stability, and enhances chromatographic separation of metabolites.

Trimethylsilyl derivatives also have some disadvantages, such as a dominant neutral loss of hydroxytrimethylsilyl under 70 eV electron impact conditions, which often leads to very low intensities of the molecular ions. In addition, some compound derivatives prove to be unstable under unsuitable inlet conditions, especially select amino acids and amides, such as asparagine. Primary amines are generally more readily derivatized by trimethylsilylation than secondary amines.

Primary amines can be completely derivatized, modified by only one trimethylsilyl group, or left underivatized. A complete derivatization involves the exchange of both acidic protons with trimethylsilyl groups. A common incomplete derivatization results in the exchange of a single trimethylsilyl group that often is the major derivative peak. An underivatized amine is rarely found.

If more than one amine group is present in a metabolite, an even higher number of derivative peaks can potentially be formed. Amine groups are the primary reason why the Agilent Fiehn 2013 GC/MS Metabolomics RTL Library contains more spectra than metabolites. Hydroxyl- and carboxyl groups are fully derivatized in all cases.

Alternatives for trimethylsilylation have been proposed by using tertiary butyldimethylsilyl derivatives or ethyl chloroformate (see O. Fiehn, Extending the breadth of metabolite profiling by gas chromatography coupled to mass spectrometry. *TrAC - Trends Anal. Chem.* 2008: 27, 261-269). These or other derivatization methods are less universal than trimethylsilylations and are not included in the Fiehn 2013 GC/MS Metabolomics RTL Library.

Methoxyamination

Prior to trimethylsilylation, ketones and aldehydes are protected by derivatization to methoxyamino groups (methoxyamination). Two reasons underscore the rationale behind this derivatization. For one, alpha-keto acids tend to undergo chemical loss of carboxyl groups as carbon dioxide, if the keto group is left unprotected. Since many metabolites comprise alpha-keto acid groups, such as 2-oxoglutarate and pyruvate, derivatization is a precautionary step for quantitative metabolite profiling. More importantly, many carbohydrates are present in cyclic and in open

chain (linear) form, using the keto- or aldehyde group as a center for forming hemiketals or hemiacetals with the carbohydrate hydroxyl groups. The anomeric hydroxyl groups can then be present in alpha- or beta position. Each of the various forms would lead to different peaks in gas chromatography. However, without protection of the keto- or aldehyde groups, the ratio of the different peaks cannot be controlled, making quantifications extremely challenging.

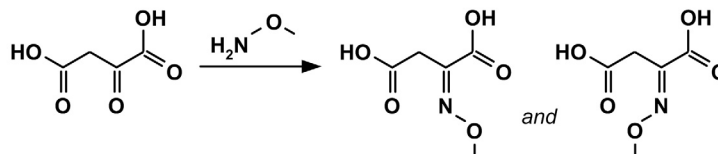
Cyclizations are inhibited when the carbonyl groups are methoximated, leading to two distinct peaks in GC/MS in most instances. The two peaks are the syn- and the anti-form of the methoxyamine group due to the inhibition of rotation across the carbon-nitrogen double bond. The ratio of syn- and anti-form is stable for each metabolite and is only dependent on the internal energy of the molecule. This is the second reason why more than one peak can be detected per metabolite in many instances of the metabolite library.

The examples below illustrate step by step how derivatization leads to modifications of the molecular structure and hence to improved stability for GC/MS screening.

Derivatization of oxaloacetic acid

Oxaloacetic acid is a dicarboxylic acid member of the citric acid cycle (TCA or Krebs cycle) with an alpha carbonyl (keto group). The methoxyamination reaction step replaces the oxygen atom of the carbonyl group by methoxyamine ($=N-OCH_3$). The second derivatization step, silylation using N-methyl-N-trimethylsilylation (MSTFA), introduces trimethylsilyl groups to the two carboxyl groups, replacing both acidic hydrogens as shown in [Figure 2](#). The reactions have to be carried out with polar aprotic, basic solvents. Pyridine is normally used as the solvent to capture protons and increase the nucleophilicity of methoxyamine. The use of dimethylformamide as the solvent has also been reported. Agents other than MSTFA can be used for trimethylsilylation, such as bis(trimethylsilyl)trifluoroacetamide (BSTFA) or variants that include trimethylchlorosilane as a further catalyst.

Step I. Methoxyamination to protect carbonyl groups



Step II. Trimethylsilylation to decrease the boiling point

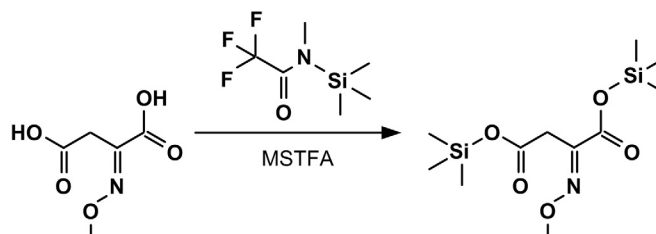


Figure 2 Two-step derivatization of oxaloacetate by methoxyamination and trimethylsilylation for GC/MS

The result is a more volatile and more stable modified derivative of oxaloacetic acid that is easily detectable by GC/MS. Methoxyamination always leads to syn- and anti-forms of the derivatized products, but these cannot always be separated into two distinct peaks under the GC/MS parameters given in the experimental section. For oxaloacetate, only one peak is observed at 12.45 minutes retention time in [Figure 3](#). The peak found at 13.239 minutes is an internal retention index marker compound (dodecanoic acid methyl ester).

The molecular mass of oxaloacetate (MW 132) increases by 29 Da due to the methoxyamination ($R_2C=N-OCH_3$ replacing $R_2C=O$) and twice by 72 Da due to the trimethylsilylation ($-Si(CH_3)_3$ is 73 Da, minus the replaced acidic proton). The derivatized oxaloacetate has a molecular mass of 305 Da, but its molecular ion is not represented in the spectrum. The presence of multiple methyl groups due to trimethylsilylation leads to detection of pseudo molecular ions of M-15, as seen at m/z 290 in [Figure 4](#) on page 27.

2 Sample Preparation and Derivatization

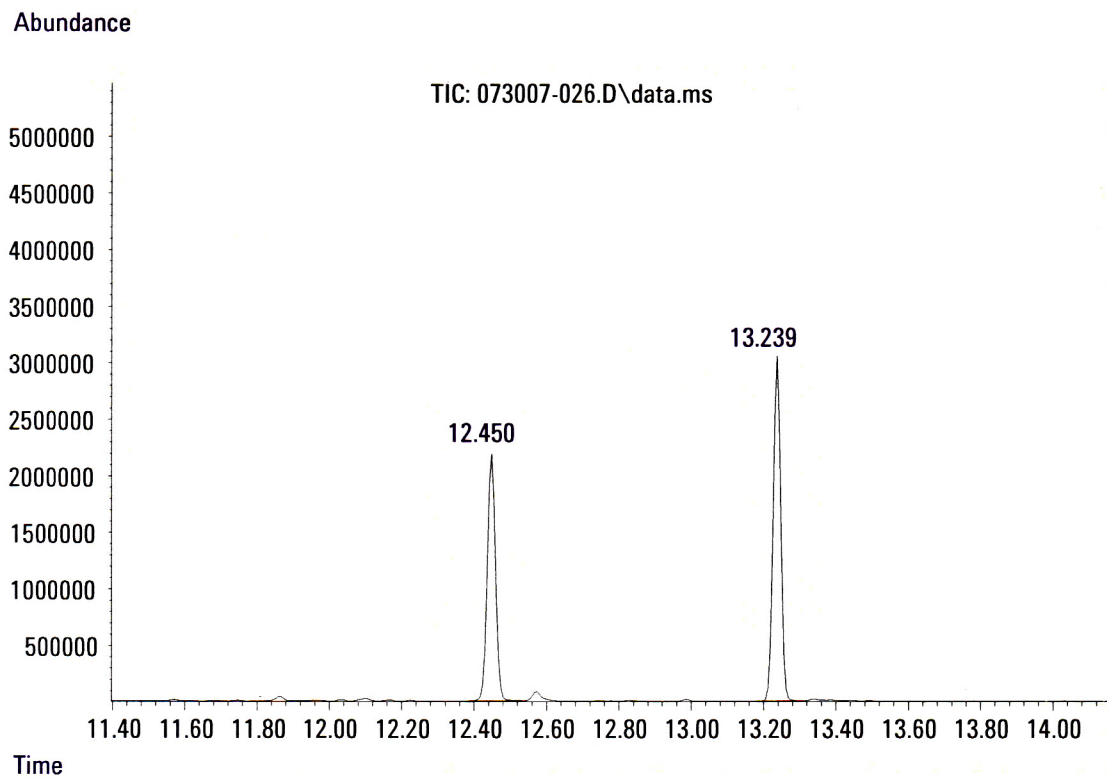


Figure 3 The total ion chromatogram of oxaloacetate detects the derivatized compound at 12.450 minutes retention time. The retention index marker methyl laurate elutes at 13.239 min.

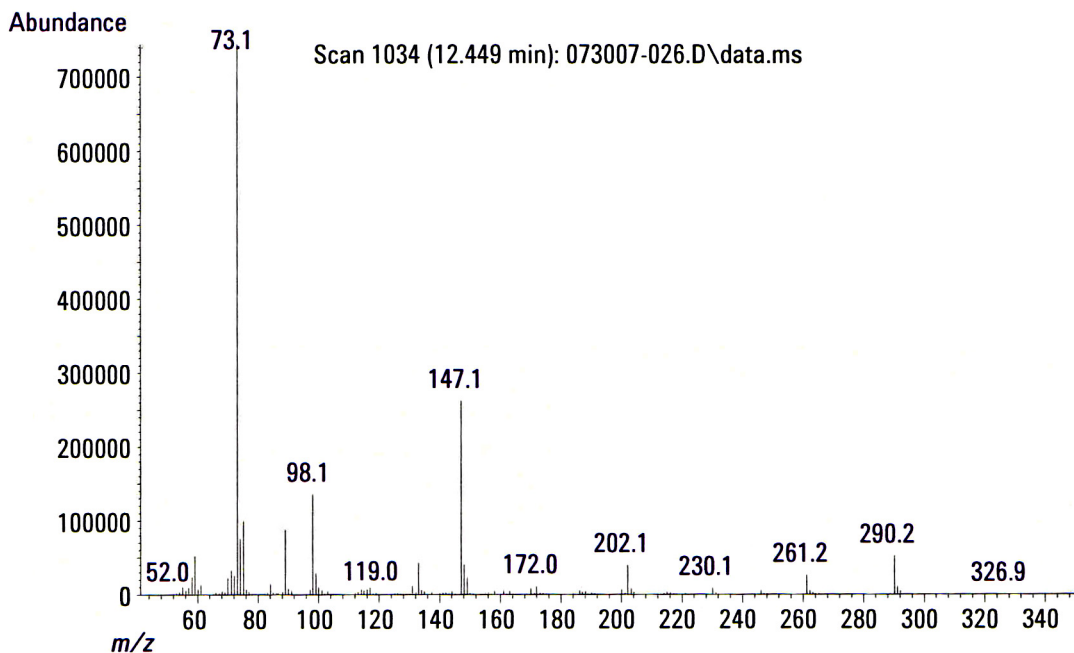


Figure 4 The electron impact mass spectrum of derivatized oxaloacetate (MW 305). The pseudomolecular ion at M-15 is clearly detectable at m/z 290. In most trimethylsilylated GC/MS spectra, the trimethylsilyl fragment is detected as base peak (m/z 73) whereas m/z 147 represents a rearrangement product of trimethylsilyl-O-dimethylsilyl, indicating the presence of at least one hydroxyl group in the compound substructure.

Derivatization of N-Acetylaspartate

N-acetyl aspartic acid comprises two hydroxyl groups and a secondary amine. Carbonyl groups adjacent to heteroatoms, such as nitrogen in the case of the N-acetyl group or oxygen in the case of ester groups, are not electropositive enough to become target of nucleophilic attack by the methoxyamine reagent. Therefore, the methoximation step does not alter the molecule.

The trimethylsilylation step yields two products. The first product replaces the protons at both carboxyl groups. The second product replaces the two carboxylic group and the less acidic secondary amine (Figure 5).

2 Sample Preparation and Derivatization

N-acetylaspartate 2TMS is the more abundant peak, whereas the completely derivatized molecule N-acetylaspartate 3TMS is detected as a minor peak (Figure 6 on page 29).

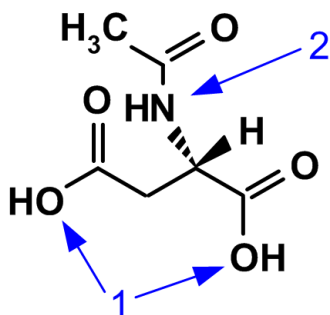


Figure 5 Molecular structure of N-acetyl aspartic acid. (1) Indicates the modification of N-acetyl aspartic acid at the two carboxyl groups, which always occurs instantly and completely. (2) Secondary amines are less reactive, so that two peaks are detected in GC/MS, one for the completely derivatized tri-TMS molecule and another one for the incompletely derivatized di-TMS compound at the two carboxyl groups without replacing the proton at the amine group.

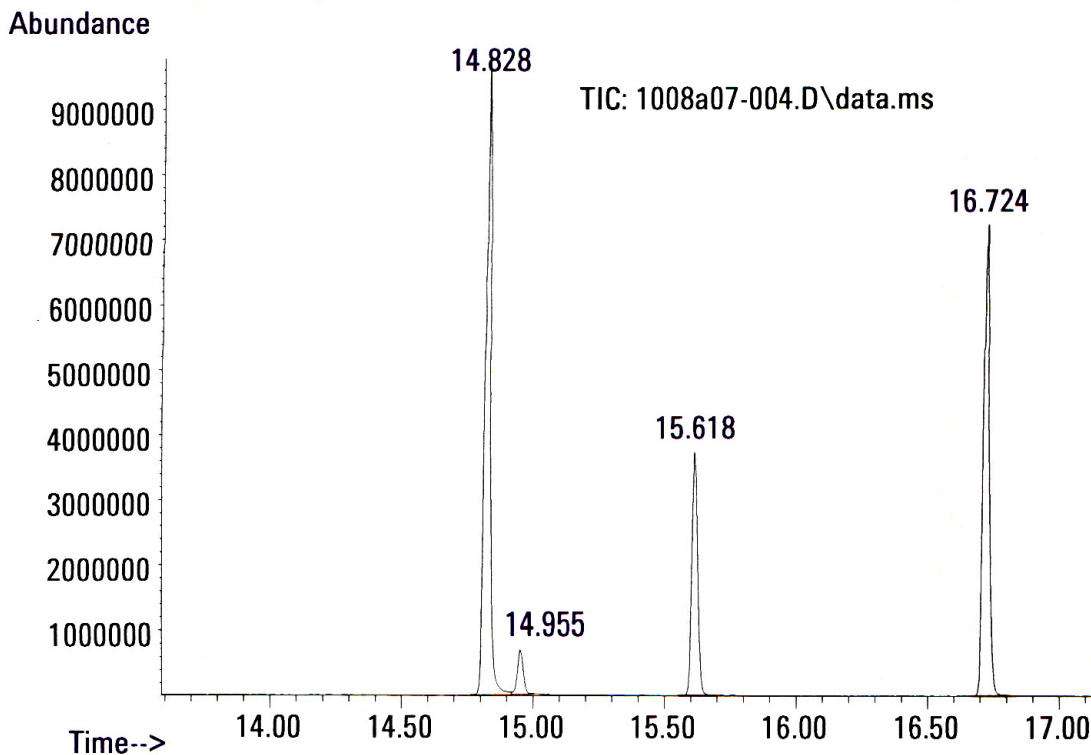


Figure 6 Total ion chromatogram of derivatized N-acetyl aspartic acid. The 2x trimethylsilylated molecule forms the major peak at 14.828 min, and the 3x TMS derivative elutes at 14.955 minutes. The internal locking compound trimethylsilylated d27-myristate is detected at 16.724 min. The retention index marker, methyl tetradecanoate, is found at 15.618 min.

2 Sample Preparation and Derivatization

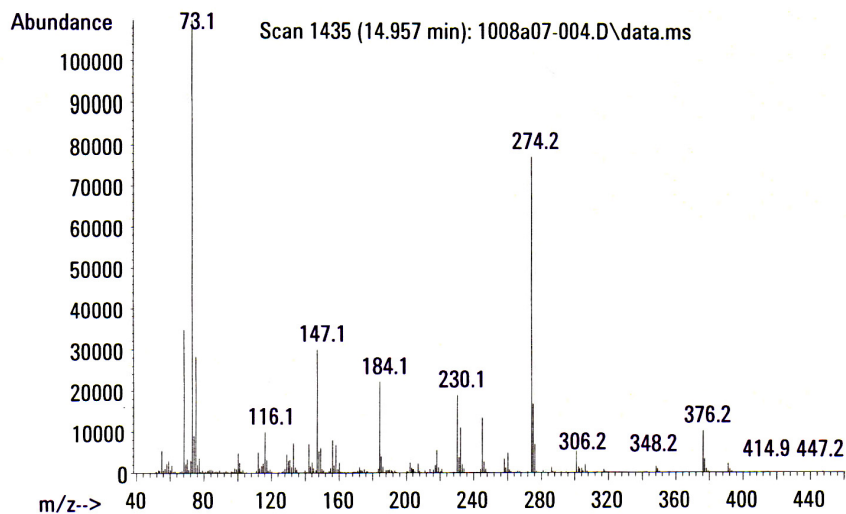
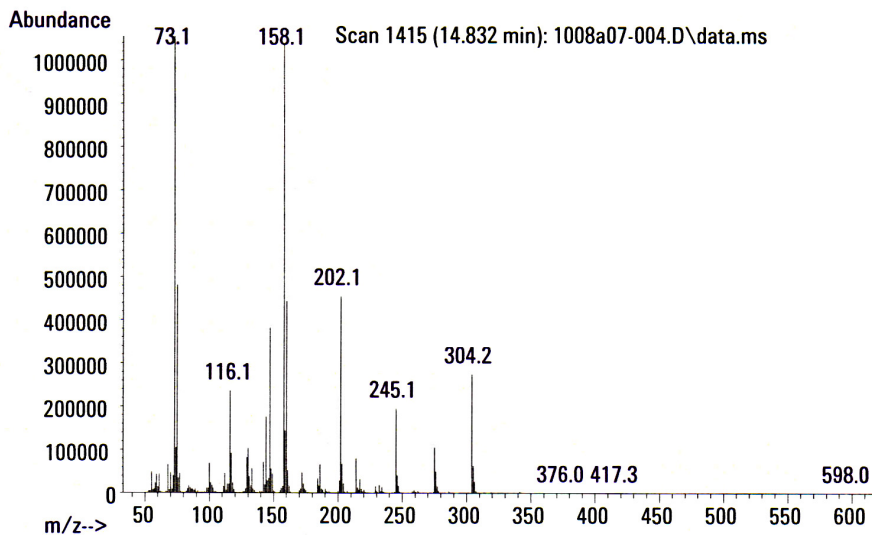
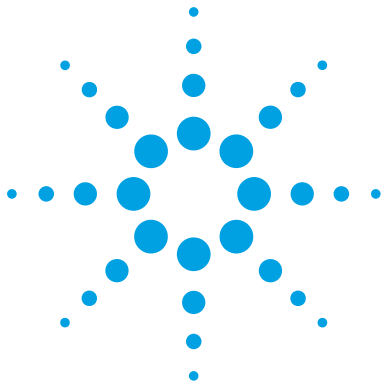


Figure 7 Upper spectrum: O,O-bistrimethylsilyl- N-acetyl aspartic acid (MW 319). The pseudo molecular ion m/z 304 is present.
Lower spectrum: N,O,O-trimethylsilyl-N-acetyl aspartic acid (MW 391). Both the molecular ion at m/z 391 and its pseudomolecular ion after methyl loss are detected at m/z 376.



3 Experimental and Instrument Method

Experimental	32
Acquisition Method	35
GC/MS Setup	38
Quality Control	42

This section specifies the parameters that were used for establishing the spectra and RT or computed RI for the metabolites included in the Agilent Fiehn 2013 GC/MS Metabolomics RTL Library.

An RT match with the library is necessary when using the Agilent GC/MS ChemStation to find metabolites in biological samples. The parameters specified in “[Acquisition Method](#)” on page 35 should be faithfully followed in addition to using an RTL method run on the GC/MS ChemStation software for acquiring data.

The AMDIS program’s library stores an RI calculated from the actual RT of each metabolite. The AMDIS program uses a calibration file to translate RT to RI values. If the exact acquisition method presented here was not followed, a good match can be obtained as long as the relative retention order of the compounds is the same in both acquisition methods.

The FAME markers used for calculating retention indices are not required in a sample. It is required that a FAME marker calibration sample is run and correctly associated with a sample. Once this association is established, the FAMEs markers can be left out of each sample.

The inclusion of FAMEs markers in one of the daily samples can guarantee a correct RI calibration can be performed. The RTL locking compound should be included in every sample. Examining the locking compound in a data file from an RTL method determines if a re-locking needs to be performed. Examining the RTs of the FAME markers determines if a



retention index calibration needs to be performed. The retention time or RI is a very important part of the analysis. This measurement is used in combination with the EI spectra to identify a metabolite.

Experimental

Internal standard

RTL locking compound: Myristic acid d27 (Product # 366889; Sigma-Aldrich, St. Louis, MO). A stock solution of 3 mg/mL is prepared in the solvent mixture water:methanol:isopropanol; 2:5:2 (v/v/v). RT of the locking standard is 16.752 minutes.

Retention index markers

Fatty acid methyl esters (FAME) (Sigma-Aldrich, St. Louis, MO) C₈, C₉, C₁₀, C₁₂, C₁₄, C₁₆, C₁₈, C₂₀, C₂₂, C₂₄, C₂₆, C₂₈, and C₃₀ linear chain length are dissolved in chloroform at concentrations of 5 mg/mL (C₈-C₁₆) and 2.5 mg/mL (C₁₈-C₃₀). The inclusion of these markers in a sample chromatogram can be used as a QC check. If identical chromatographic conditions as those employed to generate the Agilent Fiehn Library were used, the RT values match.

The *Agilent Fiehn GC/MS Metabolomics Standards Kit* (p/n 5182-0723) contains the RTL locking compounds and the FAMES.

Table 1 FAME markers

Name	Concentration	RT (min)
Methyl caprylate C ₈	100 µg/mL	7.812
Methyl perlargonate C ₉	100 µg/mL	9.248
Methyl caprate C ₁₀	100 µg/mL	10.647
Methyl laurate C ₁₂	100 µg/mL	13.25
Methyl myristate C ₁₄	100 µg/mL	15.597
Methyl palmitate C ₁₆	100 µg/mL	17.723
Methyl stearate C ₁₈	50 µg/mL	19.663

Table 1 FAME markers (continued)

Name	Concentration	RT (min)
Methyl eicosanoate C ₂₀	50 µg/mL	21.441
Methyl docosanoate C ₂₂	50 µg/mL	23.082
Methyl linocerate C ₂₄	50 µg/mL	24.603
Methyl hexacosanoate C ₂₆	50 µg/mL	26.023
Methyl octacosanoate C ₂₈	50 µg/mL	27.349
Methyl triacontanoate C ₃₀	50 µg/mL	28.723

Derivatization

Key points to consider:

- Thoroughly dry metabolite samples before derivatization. Reagents are not compatible with water or protic solvents.
- Use enough reagent to completely derivatize all metabolites.
- Derivatized samples have a 24-hour shelf life.

Before doing a large study, conduct a test to determine the maximum sample size that is completely derivatized with the recommended protocol.

Preparing the biological extract

An aliquot of 5 µL of myristic acid d27 stock solution is added to the biological extracts. The sample is then evaporated to dryness and 2 µL of the FAME RI marker solution is added to the dried extracts.

Methoxyamination

Add 10 µL of a 40 mg/mL solution of methoxyamine hydrochloride (Product # 67546; Sigma-Aldrich; St Louis, MO) in pyridine (Product # TS-27530; Thermo; Rockford, IL). This mixture is gently shaken at 30 °C for 90 minutes.

Trimethylsilylation

Add 90 μL of N-Methyl-N-trimethylsilyltrifluoroacetamide with 1% Trimethylchlorosilane (MSTFA +1% TMCS) (Product # 48915; Thermo; Rockford, IL) to the methoxyaminated samples. The mixture is incubated at 37 °C for 30 minutes. The derivatized samples are cooled to room temperature before being transferred into GC vials.

The vials are injected for GC/MS analysis under the conditions given in the “[Acquisition Method](#)” section.

When analyzing multiple samples, randomize the injection order to reduce the affects of variable reaction times on statistical analysis.

Acquisition Method

The following method was used to acquire the retention and spectral data used in the library. This GC/MS ChemStation method is also installed in your ChemStation's method directory. See "GC/MS ChemStation files" on page 12.

Equipment

Agilent 6890GC/5975MSD or Agilent 7890A/5975MSD system

Inlet

Type	Split/Splitless
Mode:	Split
Initial temp:	250 °C (On)
Pressure	9.02 psi (On) This pressure is not an absolute setting. The actual setting for your system is determined during the RTL procedure. The field-run RTL-acquired pressure is then saved with your method.
Split ratio:	10:1
Split flow:	11.1 mL/min
Total flow:	15.2 mL/min
Gas saver:	On
Saver flow:	20.0 mL/min
Saver time:	3.00 min
Gas type:	Helium

Column

Agilent ZORBAX DB5-MS +10m Duragard Capillary Column (Part number: 122-5532G; Santa Clara, CA) 30 m X 250 µm X 0.25 µm; max temperature: 325 °C; conditioned before use following the manufacturer's guidelines.

Oven

Initial temp:	60 °C		
Initial time:	1.00 min		
Ramps:			
Rate		Final temp	Final time
10.00		325 °C	10.00
Post temp:	60 °C		
Post time:	1.00 min		
Run time:	37.50 min		
Mode:	Constant flow		
Initial flow:	1.1 mL/min		
MSD Transfer Line temperature:	290 °C		

MS acquisition parameters

Tune the MSD (make sure to select atune) before each batch of samples run and do not exceed 12 hours between each tune. Save the tune parameters as atune.u. Keep a record of atune parameters to monitor MS performance.

Tune File:	atune.u
Acquisition mode:	Scan

MS information

Solvent delay:	5.90 min
EM offset:	188

Scan Parameters

Low mass:	50.0
High mass:	600.0

Threshold: 150
Changing the Threshold can improve results depending on the sample matrix.

Sample #: 2

A/D samples 4

MSZones

MS quad: 150 °C

MS source: 250 °C

GC/MS Setup

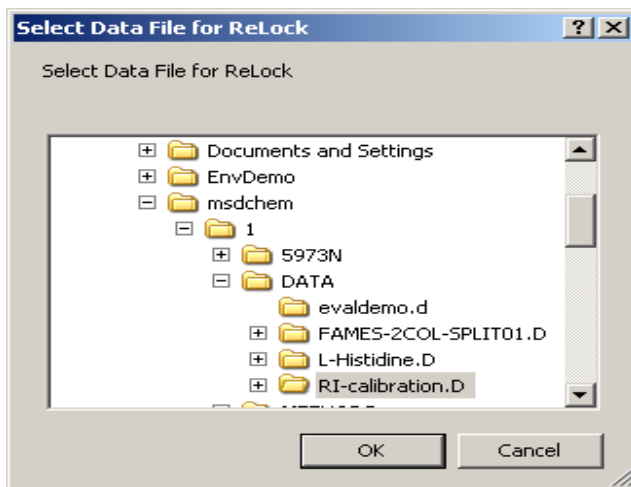
Before analyzing samples for metabolites:

- Verify “Quality Control” checks
- “Re-lock the Fiehn.M method”
- “Update the AMDIS .cal file”

Re-lock the Fiehn.M method

This procedure changes the Fiehn.M method’s column pressure so that the RT of the locking compound run on your GC/MS system is exactly the same as its RT on the GC/MS system used to analyze metabolites for creating the Fiehn.L library.

- 1 Prepare a clean RTL sample using the locking compound.
- 2 Using the GC/MS ChemStation, load the supplied Agilent method Fiehn.M.
- 3 Run the RTL sample prepared in [step 1](#) on your GC/MS system.
- 4 From the **Data Analysis** view select **RTLock Setup** from the **View** menu to enter the **RTLock** view.
- 5 Select **Relock Method...** from the **RTLock** menu.



- 6 Select data file of the RTL sample run in [step 3](#).

- 7 Click and drag using the right-click to select the peak used for locking in this data file.
- 8 Perform a cross-correlation of the spectrum (below 0.8 is not allowed).
- 9 Calculate a suggested pressure based on the retention time of the peak in the selected file, the pressure it was collected with, and the curve stored in the method.
- 10 Select a peak for locking from the list box.
- 11 Use the input box displayed to confirm or change the retention time in the relock data file.
- 12 Update the method with the new pressure, when you are prompted to do so.
- 13 Run a representative sample of metabolites containing the FAMES markers and locking compound using this relocked method on the GC\MS system.
- 14 Verify that the retention time of the new run matches the retention time stored in the supplied method RT curve.

AMDIS RI use

An RI is assigned to all compounds in the AMDIS Retention Index Standards file (fiehn.csl). This file contains the 12 FAME compounds used to assign an RI to unknown compounds found in an AMDIS analysis. AMDIS uses the RI difference between that calculated from the unknown's RT and that stored in a spectral library to assign a penalty to a spectral match. The RT of an unknown compound found in the GC/MS ChemStation data files analyzed by AMDIS must be converted to an RI for AMDIS analysis types that assign this RI penalty.

The AMDIS calculation of an unknown compound's RI is based on the RI of the nearest FAME marker with a higher RT and the nearest FAME marker with a lower RT than the unknown. The unknown compound is assigned an RI between these two marker RI values that is proportionate to their RT differences.

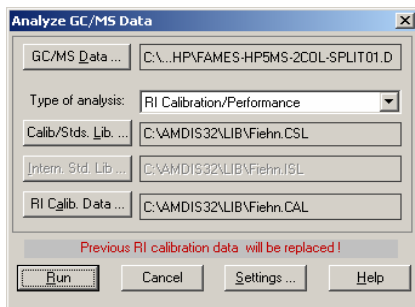
If your method and GC/MS system is not identical to those used to create the Agilent Fiehn Library, your retention times can be outside an acceptable RT window. This would result in an excessive penalty to the match fit, moving a potential good match further down the list. If these method differences do not change the compound elution order, a new conversion table can be created to assign the correct retention times.

An AMDIS analysis type of RI Calibration/Performance is used to create a conversion table of RT to RI values for each of the 12 FAME compounds. This conversion table is stored in the RI calibration file (.cal). The Fiehn.cal file supplied with this application stores these values for the specific GC/MS system used to acquire the data to create the Fiehn libraries (.msl). See “Update the AMDIS .cal file” on page 40 for instructions to create this conversion table file.

Update the AMDIS .cal file

This AMDIS analysis procedure integrates the calibration data file (.d) acquired by the GC/MS ChemStation and matches the found peaks with the FAME compounds in the provided Calibration Standards Library (.csl). It then takes the RTs of these found FAME markers and matches them to the RI stored in the provided Calibration Standards Library. This pairing of RIs with RTs is stored in the RI Calibration file (.cal), overwriting the file's previous matchings.

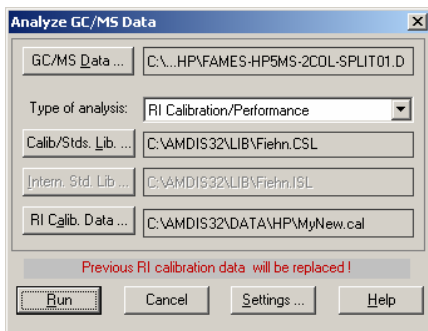
- 1 Prepare a clean calibration sample containing the 12 FAME markers and the locking compound.
- 2 Using an RTL method, run this sample on the GC\MS system.
- 3 Copy the resulting FAME calibration data file to the **AMDIS32\DATA\HP** directory.
- 4 Start AMDIS and select **Open** from the **File** menu. Navigate to the **AMDIS32\DATA\HP** directory and select the GC/MS ChemStation data file used for the FAME calibration run.
- 5 Select **GC/MS Data** from the **Analyze** menu.
- 6 Change the **Type of Analysis** to **RI Calibration/Performance**.



7 Change the **Calib/Std. Lib...** to the supplied Fiehn.csl file.

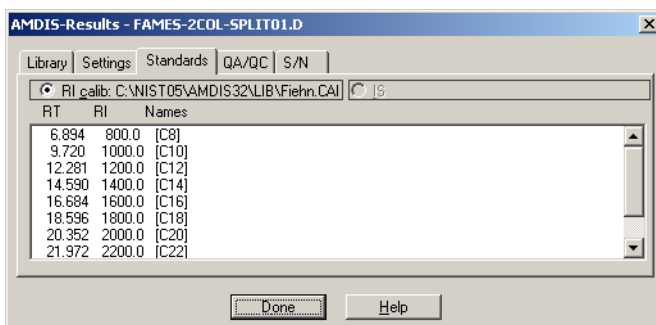
8 Change the **RI Calib. Data...** to Fiehn.cal.

This file is overwritten in this analysis. You can use a new file name if you wish to keep the original supplied file.



9 Click **Run** to update the RI Calibration Data file.

10 Click the **Info...** button above the chromatogram and click the **Standards** tab.



The results of the updated RI Calibration Data file are shown. Examine the results to determine if all 12 FAME compounds were detected and show the new RTs.

11 Click **Done**, then exit AMDIS.

It is important to use the appropriate RI calibration file that corresponds to the FAMES markers included in the analytical method. If multiple calibration files are maintained, it is critical to use the correct .cal file.

Quality Control

Tune the MSD daily Tune the mass spectrometer daily before starting an analysis. Follow the procedures in the Agilent user manual. Perfluorotributylamine (PFTBA) is the tune compound.

Ion source cleanliness The repeller values and the m/z 502 abundance must be checked to ensure that the ion source is clean.

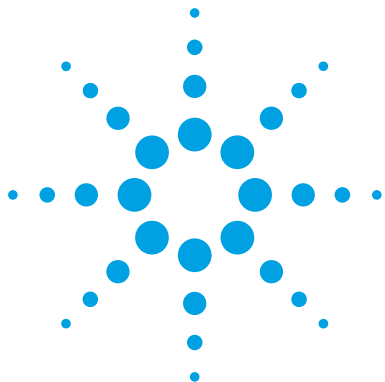
Air and water limits It is recommend that air (m/z 28) and water (m/z 18) abundances be below 5% of base peak (base peak m/z 69).

Derivatization reagent blank Each analysis sequence must be preceded by a derivatization reagent method blank to check for reaction products. Reaction products found are caused by equipment contamination or impurities in the reagents. This test confirms the purity and conformity of reagents as supplied by the chemical vendors.

Metabolite QC standards To check the overall instrument sensitivity as well as the compliance of the injector system, it is recommended to inject mixtures of metabolites as quality control (QC) standards. These QC samples are analyzed at a minimum of four different concentrations before and between sequences of biological samples. Specifically, low- and high-boiling metabolites have to be present (for example, pyruvate and raffinose) as well as representatives of various metabolite classes (aromatic compounds such as benzoic acid; carbohydrates, such as fructose; hydroxyl acids, such as malic acid; organic phosphates, such as glucose-6-phosphate; steroids, such as cholesterol; and a variety of amino acids, such as glycine, alanine, glutamate, asparagine, and serine). Peak shapes, sensitivities and peak ratios of these compounds have to be monitored and have to be compliant to lower and upper QC intervention limits.

GC maintenance In particular, the injector system has to be maintained regularly, depending on the type of matrix and number of injections. For example, liners should be exchanged at least every 40 samples. The gold-plated seal needs to be changed if too much matrix deposits lead to noncompliant QC conditions. After every liner exchange, the new liner must be cleaned and deactivated by reagent blank injections. Peaks that are generated from the liner deactivation are not considered genuine reagent peaks but represent dirt from various chemical sources.

3 Experimental and Instrument Method



4 Peak Identification

Peak Identifications from Metabolite Profiles	46
Compound Naming in the Agilent Fiehn Library	48
Metabolite Derivatizations Resulting in Multiple Compounds	50



Peak Identifications from Metabolite Profiles

Identifying isomers

Metabolites come in many isomeric forms. For example, there are eight D-aldohexoses, of which D-glucose is certainly the most well known metabolite. However, D-mannose and D-galactose frequently occur at low concentrations in biological samples. The presence of allose, altrose, gulose, idose, and talose also cannot be excluded without further knowledge of the sample. Electron impact mass spectra of the derivatized products of these metabolite isomers are virtually identical. The correct identification of the metabolite isomer requires the use of retention time as the qualifying factor. The use of retention information is one of the primary features of the Agilent Fiehn Library.

Retention time matching

All library spectra have been acquired using the Agilent retention time locking (RTL) feature.

The retention time included for each metabolite spectrum is a required part of peak identification routines when using this library.

Along with electron impact mass spectra, the library holds retention times of all the metabolite derivatives.

The RTL software has been used to generate all the retention times which enables universal retention times, as long as the same GC/MS method and chromatographic column are used. Therefore, any user laboratory can reproduce the results by locking the retention times to the mass spectral library.

Metabolite identification requires using the same GC/MS conditions listed in “[Acquisition Method](#)” on page 35. These parameters include using the Agilent ZORBAX DB-5MS column and trimethylsilylated myristic acid d27 as internal standard for RTL.

Criteria for successful compound identification have been frequently discussed in the scientific literature over the past five decades. It is evident today that a single parameter such as a mass spectral similarity threshold is insufficient to correctly confirm the presence of a particular

compound in a sample. The use of an absolute retention time, or alternatively, a retention index, is a necessary part of the identification strategy.

GC/MS data is acquired using a RTL method for retention time repeatability across all compounds in the library. The GC/MS library stores retention times that are used directly in identifications. The AMDIS program uses a retention index (RI) calculated from a compound's actual retention time (RT) relative to the RT of adjacent FAME markers included in an external calibrated sample.

Compound Naming in the Agilent Fiehn Library

Spectra of chemical derivatives of 669 metabolites have been obtained by individual analyses of GC/MS chromatograms. In many cases, more than one derivative peak per metabolite standard was observed, and accordingly, the Agilent Fiehn 2013 GC/MS Metabolomics RTL Library contains a total of 1,068 mass spectra. Since one of the primary ideas of the intended use of the library is to aid communication between analytical chemists and biologists to foster biological interpretations, the library naming refrains from using specific chemical structures or the chemical derivatives that are associated with each peak. Instead, naming refers to commonly used biological names wherever these were available. For example, the chemical 1,4-diaminobutane which is derivatized using four trimethylsilyl groups to increase volatility is referred to by the name *putrescine* instead of the chemical name *1,4-N,N-tetrakis(trimethylsilyl)diaminobutane*.

Since either way of naming metabolites, chemical structure or common biological, may lead to confusion, the Agilent Fiehn 2013 GC/MS Metabolomics RTL Library supports two different database identifiers. The CAS number and the PubChem identifier help you find further information on these compounds on the Internet. These identifier values are an important feature of the library to help you gain knowledge about the chemical and biological properties and relevance of detected GC/MS peaks.

CAS numbers

The Chemical Abstracts Service (CAS) index is a chemical database and a division of the American Chemical Society (ACS). CAS numbers can be copied from the NIST Search results and used either to gather biochemical information from freely available Web resources such as ChemSpider (www.chemspider.com) or publicly supported and curated databases such as KEGG (www.genome.jp/kegg/) or MetaCyc (www.metacyc.org/). Examples using these databases with CAS identifiers are given in [Chapter 7](#), “Using CAS and PubChem Numbers”. In addition to such pathway links, CAS entry numbers can be used in conjunction with the SciFinder database (<https://scifinder.cas.org/scifinder>) which is accessible for a fee. The SciFinder database contains chemical information on compounds and specifically the chemistry of metabolites.

PubChem identifiers

The Agilent Fiehn 2013 GC/MS Metabolomics RTL Library supports PubChem identifiers. The library includes them for every spectral entry, which is given in brackets ahead of the metabolite name. Therefore, the PubChem numbers can be retrieved either from AMDIS or from an Agilent PBM Quick Search (Figure 8 on page 55). PubChem identifiers refer to the publicly available and free library of chemical compounds hosted and curated by the U.S. National Center of Biotechnology Information (NCBI) (www.ncbi.nlm.nih.gov/pccompound) from which links to other NCBI databases can be found, such as PubMed scientific literature, NCBI's protein 3D database, and outlinks to biochemical databases such as the KEGG ligand database. PubChem identifiers are linked and could be used in other repositories as well, so that users of the Agilent Fiehn Library are enabled to query Internet resources using either the CAS numbers or the PubChem entries. An example of use of the PubChem identifier is given in Chapter 7, "Using CAS and PubChem Numbers".

Metabolite Derivatizations Resulting in Multiple Compounds

Often, GC/MS peak is obtained for a single unique chemical compound. As much as possible, the spectra included in the Fiehn library refer to actual chemical derivatives of a unique metabolite and not to an unwanted impurity in the commercial standard.

Because molecular ions are frequently absent in electron impact spectra, the correct chemical derivatization status can not always be positively established. This method would have been the optimal way to verify that the GC/MS spectra were correctly annotated to metabolite names. The Fiehn library was manually curated. Each individual spectrum was investigated and compared against available GC/MS libraries. However, some false spectra annotations can still exist. Contact Agilent if you find any such false spectra.

When more than one peak was detected for a single compound, usually one peak was more abundant than the compound's other derivatives. For N-acetylaspartic acid, an example of a minor 3TMS derivative peak and a major 2TMS derivative peak are included. In the Agilent Fiehn Library, the different derivatization products of each metabolite are listed by retention time order. See “[Compound Naming in the Agilent Fiehn Library](#)” on page 48 for RT information. Identification of a single peak is sufficient to identify this compound in a particular sample.

Derivatization products are included if peak abundance exceeds 1% intensity of the major compound derivative. However, we have not included the information of *major* or *minor* derivatives in the library. The relative ratio of two (or more) derivatives of a single metabolite sometimes depends on the status of the GC/MS instrument. The cleanliness of the injector system, its geometry, the syringe, the inertness and type of liner, the sample matrix (to a lesser extent), and the parameters used during the derivatization procedure influences this ratio.

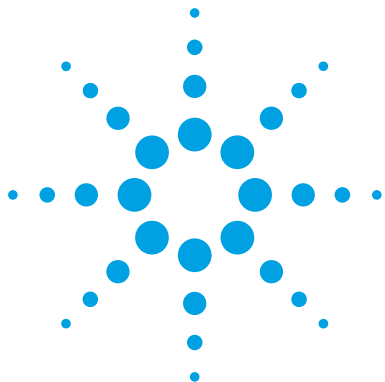
Examples and precautions to be taken are established in the scientific literature: Fiehn O, Wohlgemuth G, Scholz M, Kind T, Lee DY, Lu Y, Moon S, Nikolau BJ, Quality control for plant metabolomics: Reporting MSI-compliant studies. *Plant Journal* 2008: 53, 691-704.

Peak abundance information are not included in the peak names. A library user might be interested in a different intensity ratio for a metabolite's derivatives than was detected when establishing the library.

Make sure you establish constant peak ratios between metabolite derivatives when using their GC/MS profiles for quantitative purposes for their particular biological samples. In many cases, the monitoring of both high and low-abundance derivatives is done to watch for a peak ratio change over the course of an analytical sequence. Usually only the major peaks are used for statistical assessments and biological interpretations.

When these metabolite derivative's peak ratios change across samples of a specific biological study, serious concerns arise over the validity of quantitative results. After determining the cause of the derivative ratio changes, the analysis must be repeated.

4 Peak Identification



5 Analyzing your Data

Using Agilent GC/MS ChemStation 54

Using AMDIS 59

Using Agilent MassHunter Qualitative Analysis 63

Using Agilent MassHunter Mass Profiler Professional 75



Using Agilent GC/MS ChemStation

The Agilent GC/MS ChemStation software supports three ways to use the Agilent Fiehn Library for peak identifications;

- PBM Quick Search – interactive single spectrum search
- Library Search Report – all hits for every peak
- Screen Report – examine every compound stored in an RTL screen database to see if any exist in the sample

PBM Quick Search

A selected spectrum can be searched against a library in the Agilent GC/MS ChemStation using PBM Quick Search. This ChemStation search benefits from the Agilent Fiehn Library compound naming convention, which includes a bracketed RT value as part of the library name. See “Compound Naming in the Agilent Fiehn Library” on page 48 for more details.

This example assumes the included Fiehn.M method is already loaded. This method uses the Fiehn.L library stored in the **Database** directory. To confirm this library use, from the **Spectrum** menu select **Select Library...** before proceeding.

- 1 Select the peak's scans used for library searching from the TIC by using the mouse.

The averaged spectra for the selected scans is displayed in the spectrum window. In this example a peak at 17.191 minutes retention time was chosen.

- 2 Select **PBM Quick Search** from the **Spectrum** menu.

The PBM Search Results dialog box is shown in [Figure 8](#). The selected spectrum is displayed graphically at the top of the window with the currently selected PBM Search Results dialog box library compound hit graphically displayed directly below.

- 3 Move the selection in the PBM Search Results dialog box through the various hits by clicking them. As you select each hit, the graphic display's lower spectrum is replaced with the selected hit's spectrum.

The first entry listed is the one that has the most probable spectral match. This is not sufficient to positively identify a hit. You need to further confirm a library spectral match with retention time.

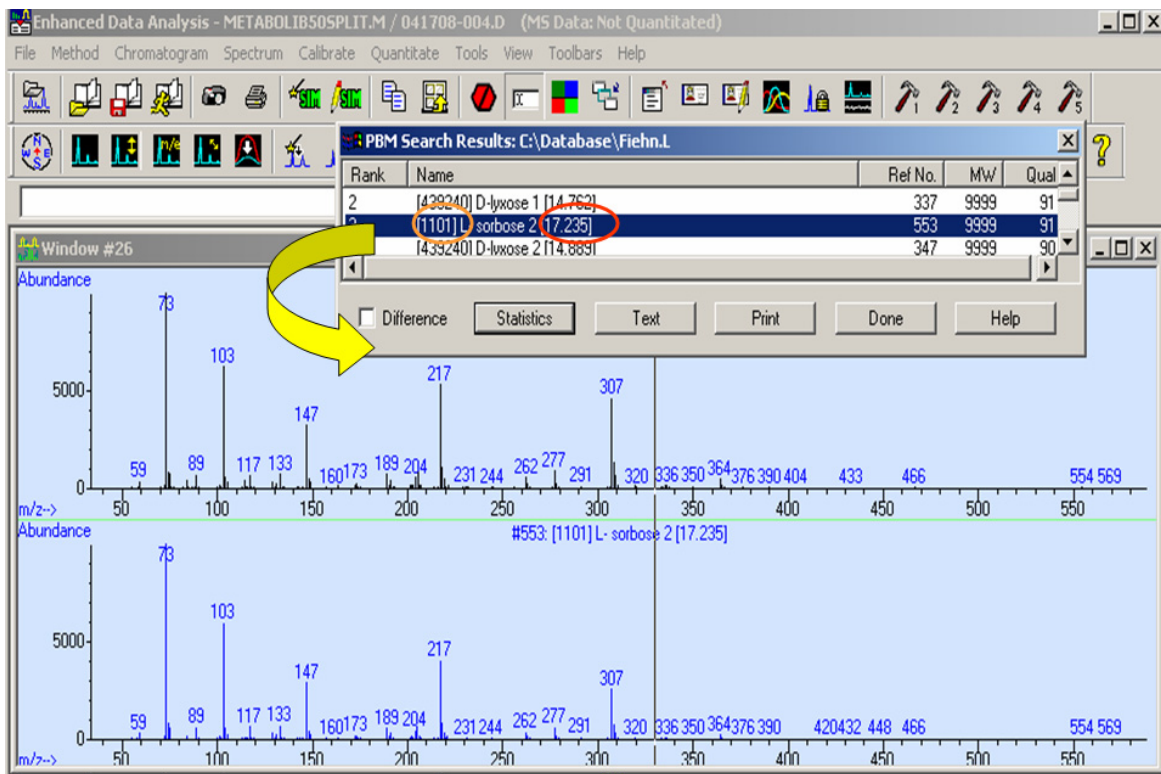


Figure 8 Identification of sorbose 2 by mass spectral similarity and RT comparison using the PBM Quick Search.

- 4 From the PBM Search Results dialog box of similar spectra, use the displayed bracketed RT values, shown circled in red in Figure 8, to make a correct identification of the compound.

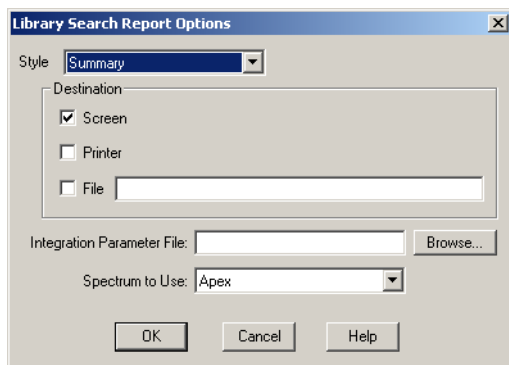
The correct identification is **sorbose 2** since its retention time of 17.235 minutes matches that of the selected peak. Other ketohexoses and ketopentoses with almost identical spectra such as the first PBM hit lyxose, do not have close RT matches. Only by using the retention time included in the compound name of this library was a correct identification possible.

The PubChem identifier number is shown circled in orange in [Figure 8](#).

Library search report

To generate this report from Data Analysis View after doing a manual integration proceed as follows:

- 1 From the **Spectrum** menu select **Library Search Report...** to display the Library Search Report Options dialog box.



- 2 Select **Summary** to generate a report listing all PBM matches of the current loaded data file to entries in the Fiehn.L library.

This report lists the retention time, area, Agilent Fiehn Library compound name, library compound number, and CAS number for each peak's spectral hits required for manual positive identification of all peak matching compounds.

- 3 Select the report destination(s): Screen, Printer, or File.
- 4 Select how the peak's spectrum is extracted for searching.

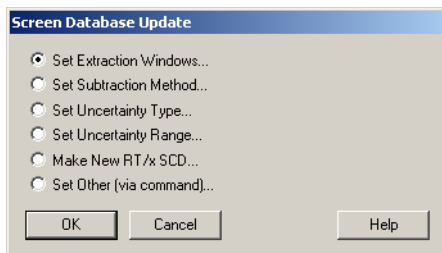
The spectrum can be taken from each peak's apex, the apex minus the spectrum at the start of the peak, the apex minus the background at tune, or the peak average.

- 5 Click **OK** to exit the dialog box and print the report to the selected destination.

Screen report

This is a very useful report for metabolite analysis using the GC/MS ChemStation. It looks at each RTL compound in a screen database provided with the Agilent Fiehn Library to determine if the screen compound exists in the sample. If found in the sample it reports the expected retention time and the difference between this time and the RT of the sample. It also reports the target ion response and any out-of-range qualifiers. To generate this report from Data Analysis View with the data file and Fiehn.M method loaded, proceed as follows:

- 1 From the **Tools** menu select **Specify Method Screen Database...** to check that Database\Fiehn.scd is the current screen database.
- 2 From the **Tools** menu select **Change Screen Database Parameters...**



Use the Screen database Update dialog box to select the parameters to change that are used for peak identification. For more information click **Help**.

- 3 From the **Tools** menu select **Create Screen Results for Current File...** to create the screener results file.
- 4 From the **Tools** menu select **Generate/Print Screen Report for Current File...** to send the report to a printer or select **Generate Screen Report for Current File...** to display the report in Wordpad. A screen results file must exist or the report is not generated.

The Screen Report uses an 'X' and '?' in the status column to indicate probable and possible identification, respectively. The '?' indicates that one or more qualifiers fail.

Expected RT and delta RT are reported, along with the target ion. Qualifier ions that are out of range are noted in the Qualifiers Out of Range columns. The last column reports the cross-correlation with the reference spectrum.

For probable matches with out-of-range qualifiers marked "?," use AMDIS to clean up and deconvolute spectra to help identify the compound.

Search basics

When using the Agilent Fiehn Library for metabolite identifications, a three-tiered identification approach is recommended. The PBM Quick Search and Library Search Report use mass spectral similarity but do not support use of retention times as a constraint. The correct hit may not be found at the top of the list. This it is recommended to review the list and select the correct metabolite using retention time, mass spectral similarity, and characteristic ions.

- (a) Use small RT differences between the sample peak RT and the library RT (for example, allow deviations of < 0.2 min).
- (b) Use very high mass spectral similarity thresholds for abundant peaks that are not compromised by coeluting peaks (for example, forward similarity matches > 850), but allow lower mass spectral similarity thresholds for low abundant peaks, especially when these are compromised by several co-eluting peaks (for example, forward similarity matches > 600).
- (c) Establish the presence and intensity ratios of qualifier ions at the high m/z values, which are characteristic for a specific metabolite. For example, spectra for many oligosaccharides would pass forward similarity thresholds (and often also coelute at close retention times), but can be easily distinguished using characteristic ions or characteristic ion ratios.

Using AMDIS

Before using AMDIS for analyzing your biological data, perform the following:

- **New AMDIS Users**

Please read [Chapter 6](#), “AMDIS Getting Started” if you are not familiar with the AMDIS program.

- **Set Up defaults**

Before using AMDIS for the first time with the Agilent Fiehn Library, it is necessary to set up program defaults. Start the AMDIS program and set these defaults as specified in “[Recommended Settings](#)” on page 183

- **Update the AMDIS .cal File**

This procedure is covered in “[Update the AMDIS .cal file](#)” on page 40 and must be run before using AMDIS for an analysis using RI data.

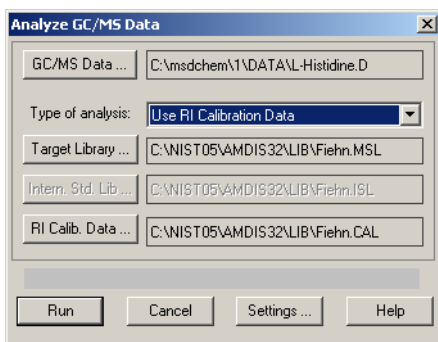
Data analysis

The example that follows is the most common AMDIS analysis used for finding metabolites.

- 1 Run the metabolite sample with the Fiehn.M method loaded in the GC/MS ChemStation.

The method might be set to print a report in ChemStation. This example is for AMDIS, and you can change the Fiehn.M method to omit printing the report.

- 2 At the end of the sample run, from GC/MS ChemStation's Analysis View, select **AMDIS** from the **Spectrum** menu to start AMDIS with the sample run data file loaded.
- 3 Select **Analyze GC/MS Data** from the **Analyze** menu.



- 4 For the **Type of Analysis**, select **Use RI Calibration Data**.

If you have followed the instructions in “Recommended Settings” on page 183, the correct Target Library and RI Calib. Data files specified should be as shown above.

- 5 Click **Run** to analyze the data file. See [Figure 9](#).

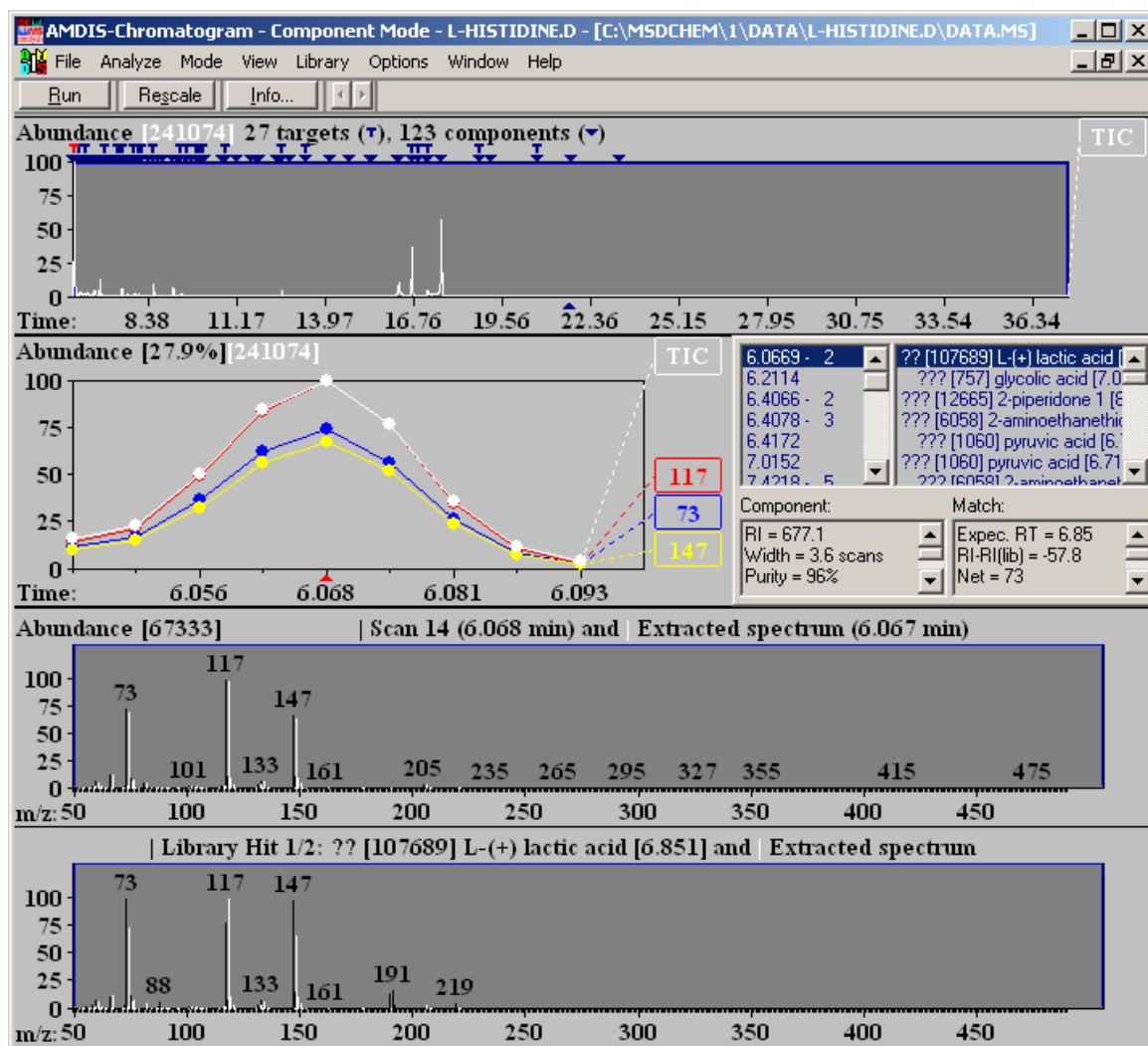


Figure 9 AMDIS after performing a Use RI Calibration Data run

Using a different analytical method

Sometimes it is necessary to use an analytical method that is different from the Fiehn.M method for GC/MS analysis of a metabolite sample. If these method differences do not change the compound elution order, it is possible to use AMDIS (but not GC/MS ChemStation) to identify the metabolites using the concepts presented in this guide.

AMDIS allows you to run a procedure to update the retention time assigned to an external marker's RI value. The RI for a compound stored in the library is calculated from the retention times of the FAME markers. You can run a sample containing the FAME markers using different methods to obtain the correct FAME RT values. You then replace the AMDIS Fiehn.cal RT values with the values obtained during the FAMEs run on this new method. The RIs in the Fiehn.cal file are not changed during this update process, permitting the library RI relationship to remain.

Before running this method, you must make it an RTL method using the GC/MS ChemStation and a sample containing the locking compound. You then run a sample containing the FAME markers with this locked method. This data file is then used to update the RT times as shown in [“Update the AMDIS .cal file”](#) on page 40.

This different method could not be used for searches on the GC/MS ChemStation where RT is necessary for metabolite confirmation. The RT contained in the Fiehn.L library or screener file is only valid for the unmodified Fiehn.M method.

Existing data files containing FAME markers

If you have data files obtained with an RTL method, that is different from the Fiehn.M method but using the same FAME markers, you can use AMDIS to identify the metabolites using the concepts presented in this guide.

You first use the data file containing the FAME markers to update the RTs in the Fiehn.cal file. See [“Update the AMDIS .cal file”](#) on page 40. Then run this data file on AMDIS using an **Analysis Type** of **Use RI Calibration Data**.

Using Agilent MassHunter Qualitative Analysis

You can use the Agilent MassHunter Qualitative Analysis software to search either spectra or compounds. If you want to export the results to the Agilent Mass Profiler Professional program, you first search compounds that you created and then you export a CEF file which contains the information on each of the compounds.

- “Install the Fiehn library to use with MassHunter Qualitative Analysis” on page 63
- “Search Unit Mass Library on spectra” on page 64
- “Search Unit Mass Library with Retention Time Matching on compounds” on page 68
- “Export results to Mass Profiler Professional using CEF file” on page 73

Install the Fiehn library to use with MassHunter Qualitative Analysis

When you install the Fiehn library and you have MassHunter Qualitative Analysis installed, the library is automatically placed in the *\Program Files (x86)\Agilent\Fiehn Metabolomics RTL Database (G1676AA)* folder. The library needs to be copied to the *\MassHunter\Library* folder.

The data files need to be translated and copied to the *\MassHunter\Data* folder.

- 1 Copy the library from *\Program Files (x86)\Agilent\Fiehn Metabolomics RTL Database (G1676AA)\GCMS Libraries\Fiehn.L* to the *\MassHunter\Library* folder. The Qualitative Analysis program cannot use the library if it is in the *\Program Files (x86)* folder.
- 2 Translate the GC/MSD data files which are in the *\Program Files (x86)\Agilent\Data* folder and put the files in the *\MassHunter\Data* folder. You need to install the translator program first which is available on the GC/MS Supplemental Software disk.
 - a Double-click the *GC MSD Translator* icon.

- b** Select the **Folder containing Data Files to Import**. Click **Browse**. In the Browse For Folder dialog box, navigate to the *C:\Program Files (x86)\Agilent\Agilent Fiehn GCMS Metabolomics RTL Database (G1676AA)\Data* folder and click **OK**.
- c** Select the **Output Folder**. Click **Browse**. In the Browse For Folder dialog box, navigate to the *\MassHunter\Data* folder and click **OK**.
- d** Clear the **In-place Translation** check box.
- e** Click **Start Translation**.

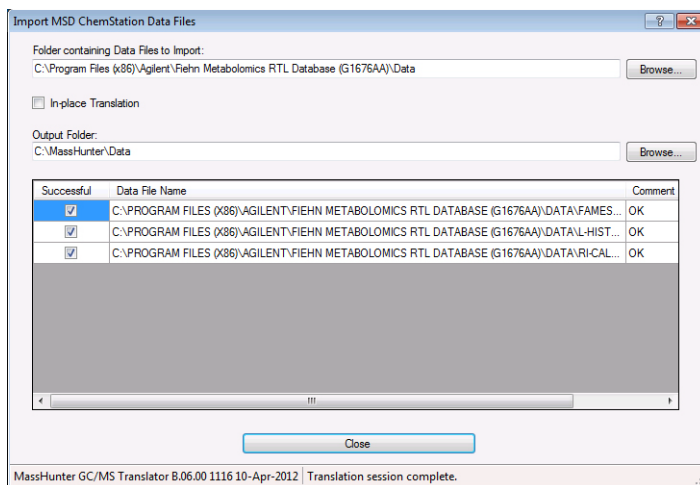


Figure 10 Import MSD ChemStation Data Files

- f** Verify that all of the data files were translated. If the check box in the Successful column is marked, then the translation was successful.

Search Unit Mass Library on spectra

You can use the Search Unit Mass Library algorithm on either a compound or a spectrum in the Agilent MassHunter Qualitative Analysis program. This example shows searching a spectrum. See “[Search Unit Mass Library with Retention Time Matching on compounds](#)” on page 68 for an example showing how to search a compound.

The Agilent Fiehn Library compound naming convention includes a bracketed RT (retention time) value as part of the library name. See “Compound Naming in the Agilent Fiehn Library” on page 48 for more details.

This example uses the Fiehn.L library stored in the **Library** folder.

- 1 Open the FAMES-2COL-SPLIT01.d data file. Click **File > Open Data File**. Select this data file and click **Open**.
- 2 Configure the user interface to include GC algorithms and to include advanced parameters. Click **Configuration > User Interface Configuration**. Mark the **GC** check box and the **Unit mass (Q, QQQ)** check box. Mark the **Show advanced parameters** check box, and click **OK**. You can press **F1** to get help on any dialog box or window.
- 3 Integrate and extract peak spectra from the data file.

Click **Chromatograms > Integrate and Extract Peak Spectra**. You can also right-click the chromatogram and click **Integrate and Extract Peak Spectra**.

You can limit the number of peaks integrated when you mark the **Limit (by height) to the largest** check box on the **Integrate (MS) > Peak Filters tab** in the Method Editor window.

The averaged spectra for each of the integrated peaks is shown in the MS Spectrum Results window. In this example the peak with the 12.28 minute retention time was chosen.

- 4 In the Method Explorer window, click **Identify Compounds** and then click **Search Unit Mass Library**.
- 5 In the Method Editor window, click **Add Library** and add the *Fiehn.l* library to the **Library selection**.
- 6 In the Method Editor window on the Scoring tab, clear the **Use RT match** check box.

5 Analyzing your Data

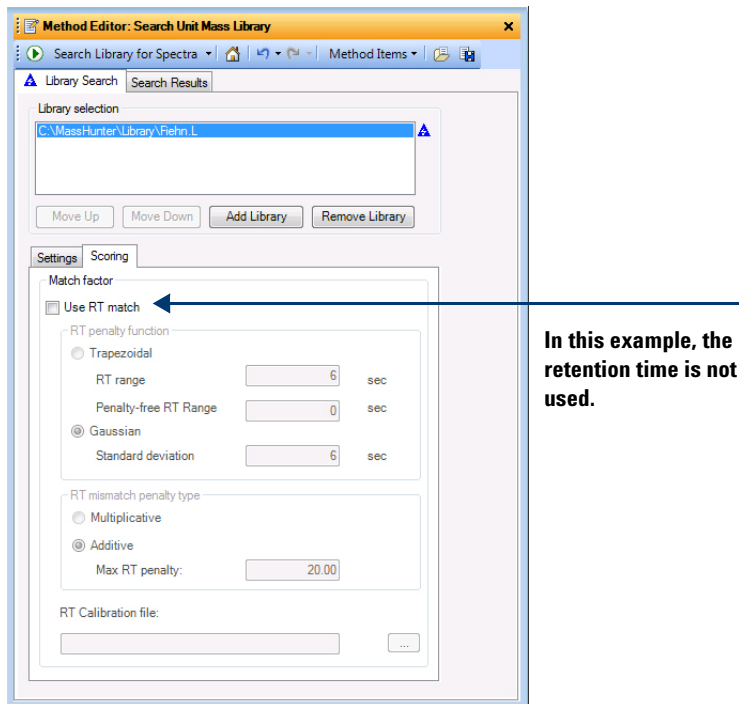


Figure 11 Library Search tab in the Method Editor window

- 7 In the Data Navigator window, click **User Spectra** to highlight all of the spectra.
- 8 Click **Spectra > Search Library for Spectra**. You can also right-click the spectrum and click **Search Library for Spectra**. If you highlight more than one spectra, all of the highlighted spectra are searched.
- 9 In the Data Navigator window, click **Scan (12.261-12.298 min)**.

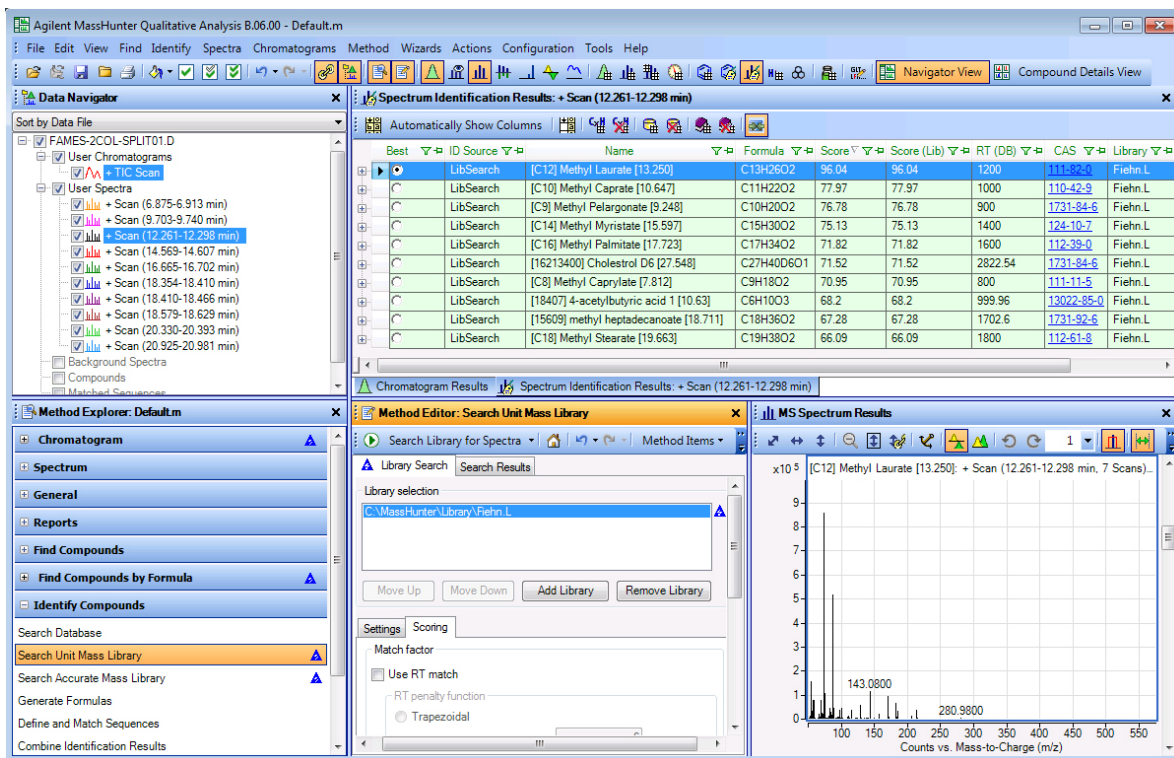


Figure 12 Qualitative Analysis after Search Unit Mass Library

The match with the highest **Score** is automatically chosen as the **Best** match. In this example, **[C12] Methyl Laureate [13.250]** has a **Score** of 96.04 and is selected as the **Best** match. You can select a different **Best** match by clicking on the row. The number in brackets at the end of the name is the retention time of the entry.

The **RT (DB)** column shows the retention index for this match. The retention index does not have a direct correlation with the retention time. Instead, the retention index is used to look up the retention time in the RT Calibration file. The RT Calibration file is used in the “**Search Unit Mass Library with Retention Time Matching on compounds**” on page 68.

10 (optional) Print a report. Click **File > Print > Analysis Report**. Mark the **Print Preview** check box and click **OK**.


You can change the results that are included in this report in the method. In the Method Explorer window, select **Reports** and then select **Analysis Report**. In the Method Editor window, mark the results that you want to include in the report.

- 11 (optional) Close the data file. Click **File > Close Data File**. Do not save the results (Click **No** in the **Do you want to save the results** message box.)

Search Unit Mass Library with Retention Time Matching on compounds

You can use the Search Unit Mass Library algorithm on either a compound or a spectrum in the Agilent MassHunter Qualitative Analysis program. This example shows searching a compound. See “[Search Unit Mass Library on spectra](#)” on page 64 for an example showing how to search a spectrum.

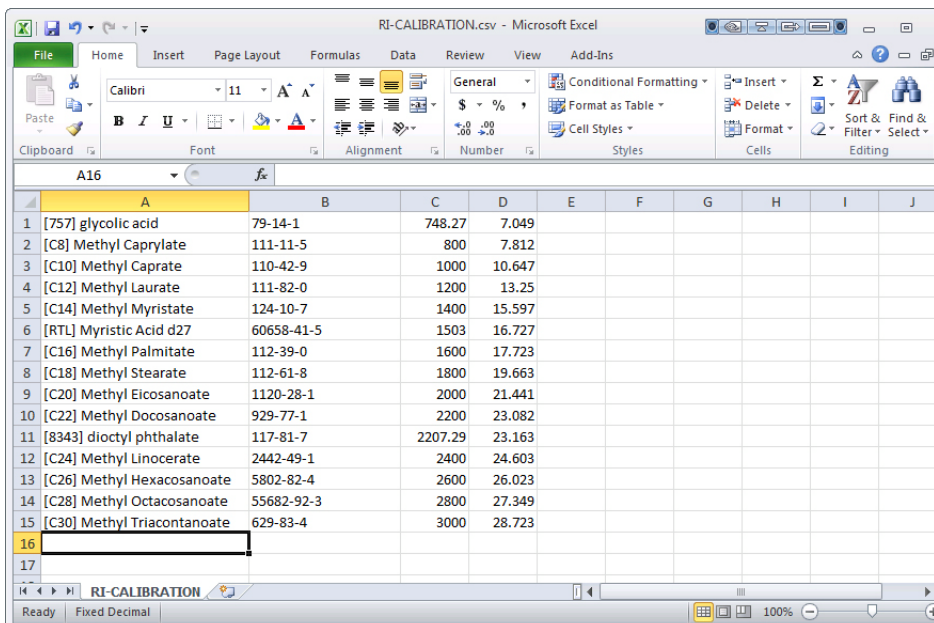
- 1 Open the RI-CALIBRATION.d data file.
- 2 Configure the user interface to include GC algorithms. Click **Configuration > User Interface Configuration**. Mark the **GC** check box and the **Unit mass (Q, QQQ)** check box, and click **OK**. You can press **F1** to get help on any dialog box or window.
- 3 In the Method Explorer window, click **Find Compounds** and then click **Find by Chromatogram Deconvolution**.
- 4 In the Method Editor window, review the parameters for this algorithm. For GC/MS data, a value of 100 is a good starting point for the **RT window size factor**.
- 5 Click **Find > Find Compounds by Chromatogram Deconvolution**. With default parameters, 43 compounds are found.
- 6 In the Data Navigator window, click **Compounds** to highlight all compounds for this data file.
- 7 In the Method Explorer window, click **Identify Compounds** and then click **Search Unit Mass Library**.
- 8 In the Method Editor window, click **Add Library** and add the Fiehn.l library to the **Library selection**.
- 9 On the Scoring tab, mark the **Use RT match** check box.

10 Click the  button to select the RT Calibration file.

11 Select the RI-CALIBRATION.csv file.

An RT Calibration file is a CSV file with four values per line:

- Name
- CAS number
- Retention index
- Retention time



	A	B	C	D	E	F	G	H	I	J
1	[757] glycolic acid	79-14-1	748.27	7.049						
2	[C8] Methyl Caprylate	111-11-5	800	7.812						
3	[C10] Methyl Caprate	110-42-9	1000	10.647						
4	[C12] Methyl Laurate	111-82-0	1200	13.25						
5	[C14] Methyl Myristate	124-10-7	1400	15.597						
6	[RTL] Myristic Acid d27	60658-41-5	1503	16.727						
7	[C16] Methyl Palmitate	112-39-0	1600	17.723						
8	[C18] Methyl Stearate	112-61-8	1800	19.663						
9	[C20] Methyl Eicosanoate	1120-28-1	2000	21.441						
10	[C22] Methyl Docosanoate	929-77-1	2200	23.082						
11	[8343] dioctyl phthalate	117-81-7	2207.29	23.163						
12	[C24] Methyl Linocerate	2442-49-1	2400	24.603						
13	[C26] Methyl Hexacosanoate	5802-82-4	2600	26.023						
14	[C28] Methyl Octacosanoate	55682-92-3	2800	27.349						
15	[C30] Methyl Triacontanoate	629-83-4	3000	28.723						
16										
17										

Figure 13 Example RT Calibration File

5 Analyzing your Data

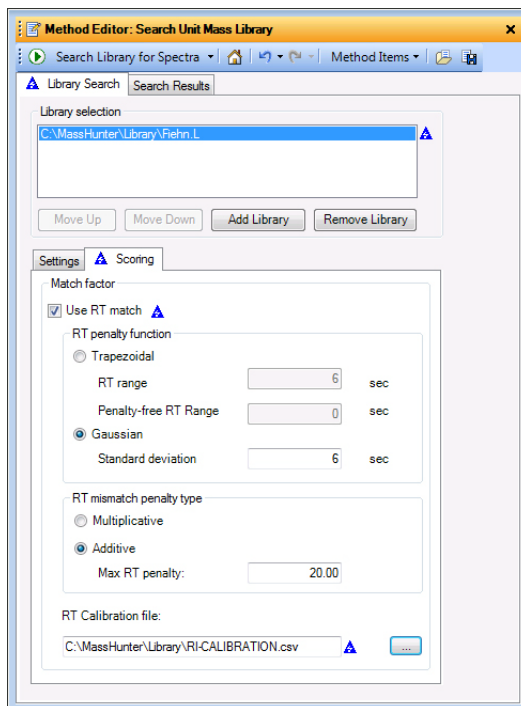


Figure 14 Library Search > Scoring tab in the Method Editor window

12 Click **Identify** > **Search Library for Compounds**.

13 Close the Method Editor window. In the Data Navigator window, highlight **Compound 16**.

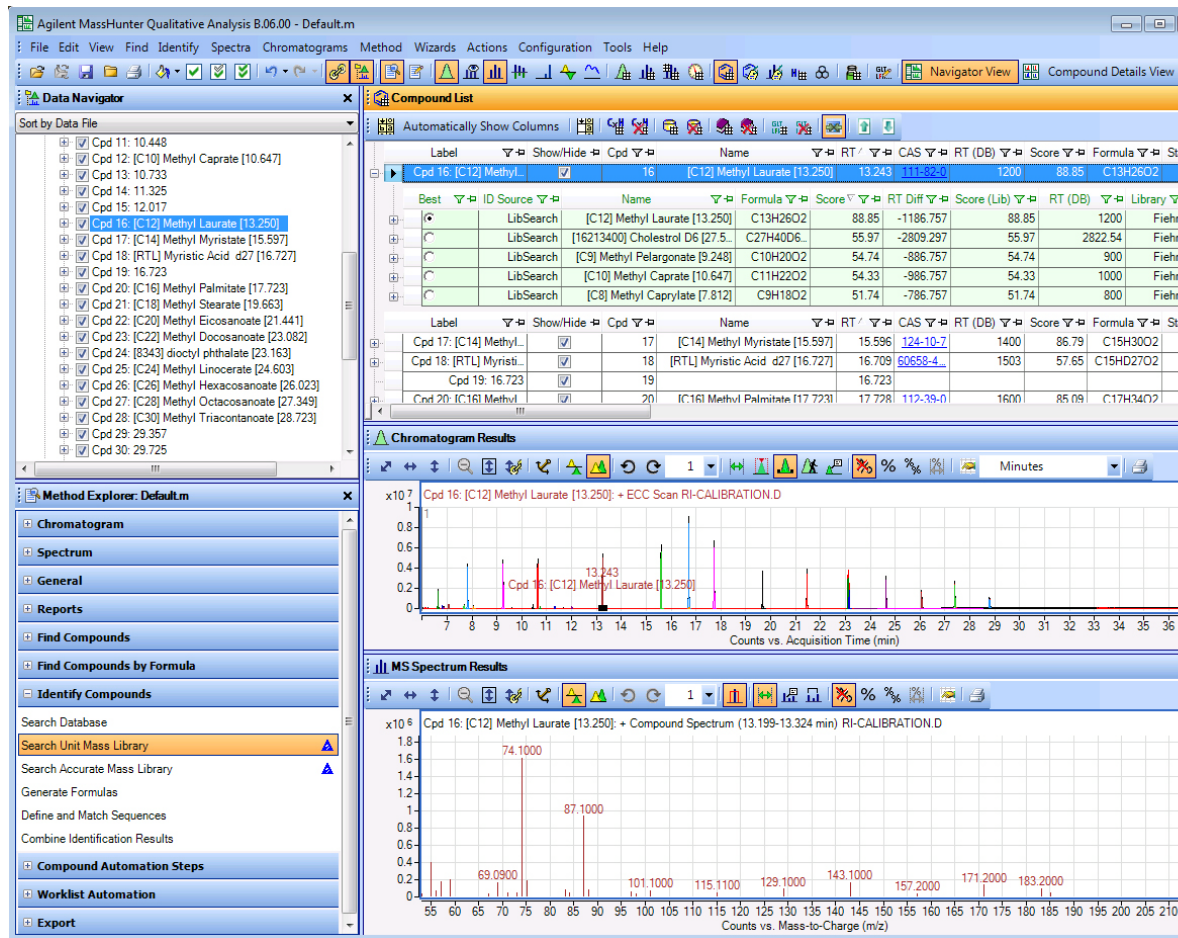


Figure 15 Qualitative Analysis after Search Unit Mass Library on compounds

The compounds are shown in the Compound List window. You can change which columns are visible in the table when you right-click in the title of the table and click **Add/Remove Columns**.

Label	Show/Hide	Cpd	Name	RT	CAS	RT (DB)	Score	Formula	Start
Cpd 16: [C12] Methyl...	<input checked="" type="checkbox"/>	16	[C12] Methyl Laurate [13.250]	13.243	111-82-0	1200	88.85	C13H26O2	13.243
Best									
Best	ID Source	Name	Formula	Score	RT Diff	Score (Lib)	RT (DB)	Library	
<input checked="" type="checkbox"/>	LibSearch	[C12] Methyl Laurate [13.250]	C13H26O2	88.85	-1186.757	88.85	1200	Fiehn.L	
<input type="checkbox"/>	LibSearch	[16213400] Cholesterol D6 [27.5...]	C27H40D6...	55.97	-2809.297	55.97	2822.54	Fiehn.L	
<input type="checkbox"/>	LibSearch	[C9] Methyl Pelargonate [9.248]	C10H20O2	54.74	-886.757	54.74	900	Fiehn.L	
<input type="checkbox"/>	LibSearch	[C10] Methyl Caprate [10.647]	C11H22O2	54.33	-986.757	54.33	1000	Fiehn.L	
<input type="checkbox"/>	LibSearch	[C8] Methyl Caprylate [7.812]	C9H18O2	51.74	-786.757	51.74	800	Fiehn.L	
Label	Show/Hide	Cpd	Name	RT	CAS	RT (DB)	Score	Formula	Start
Cpd 17: [C14] Methyl...	<input checked="" type="checkbox"/>	17	[C14] Methyl Myristate [15.597]	15.596	124-10-7	1400	86.79	C15H30O2	15.596
Cpd 18: [RTL] Myristi...	<input checked="" type="checkbox"/>	18	[RTL] Myristic Acid d27 [16.727]	16.709	60658-4	1503	57.65	C15HD27O2	16.709
Cpd 19: 16.723	<input checked="" type="checkbox"/>	19		16.723					16.723
Cpd 20: [C16] Methyl...	<input checked="" type="checkbox"/>	20	[C16] Methyl Palmitate [17.723]	17.728	112-39-0	1600	85.09	C17H34O2	17.728

Figure 16 Compound List window in the Qualitative Analysis program

In the first level of the table, you see results for each compound. In the second level of the table, you see results for each identification technique that you tried. The second level shows each of the possible matches for the selected compound.

You can compare the retention time value (**RT**) with the number in brackets at the end of the **Name** of the match. For example for compound 16, the retention time (**RT**) of the compound is 13.243, and the retention time shown in the **Name** is 13.250.

The **CAS** number for the **Best** match is included in the first level of the table.

The **RT (DB)** column shows the retention index of the match in the library. The retention index is used as an index into the RT calibration file.

Ignore the **RT (Diff)** column when you run the Search Unit Mass Library algorithm. It shows the difference between the retention time of the compound and the **RT (DB)** value for the match. The **RT (DB)** value is actually a retention index, so it is meaningless to subtract the actual retention time and the retention index.

- 14** (optional) Print a compound report. Click **File > Print > Compound Report**. Click the **All Results** button if you want the report to include all compounds. Mark the **Print Preview** check box and click **OK**.

You can change the results that are included in this report in the method. In the Method Explorer window, select **Reports** and then select **Analysis Report**. In the Method Editor window, mark the results that you want to include in the report.

You can also customize the report template. See the **Report Designer** training disk for more information.

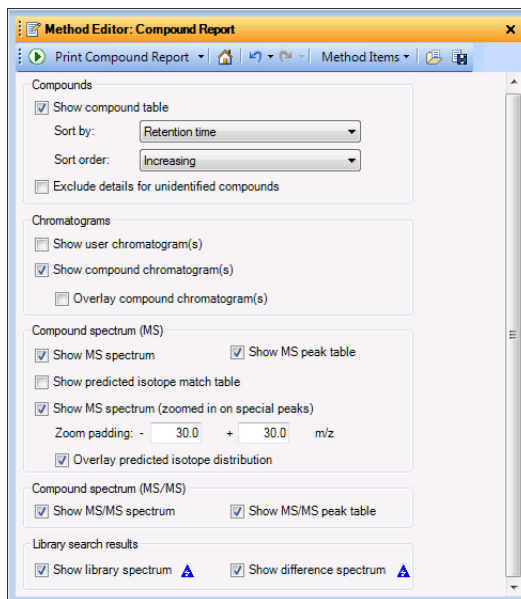


Figure 17 Report > Compound Report section in the Method Editor window

Export results to Mass Profiler Professional using CEF file

After you have created compounds and searched those compounds with the Search Unit Mass Library algorithm, you can export those results to a compound exchange format (CEF) file. The Agilent Mass Profiler Professional program can use this CEF file in experiments in a project.

- 1 Create compounds. See “[Search Unit Mass Library with Retention Time Matching on compounds](#)” on page 68.
- 2 Click **File > Export > as CEF**. The Export CEF Options dialog box opens.

- 3 Click **All results**.
- 4 Review the other parameters and click **OK**.

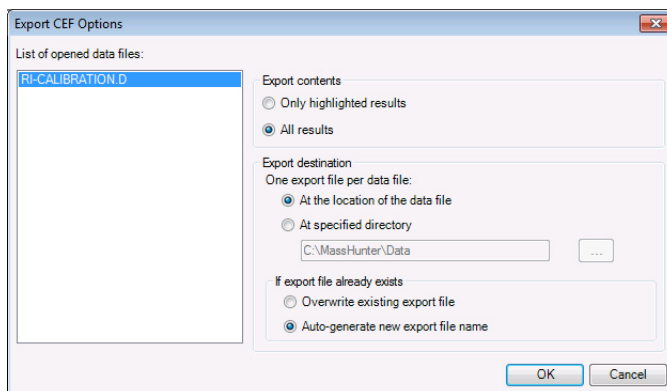


Figure 18 Export CEF Options dialog box

Using Agilent MassHunter Mass Profiler Professional

Agilent MassHunter Mass Profiler Professional (MPP) software is a powerful chemometrics platform designed to exploit the high information content of mass spectra (MS) data and can be used in any MS-based differential analysis to determine relationships among two or more sample groups and variables. MPP provides advanced statistical analysis and visualization tools for GC/MS, LC/MS, CE/MS, ICP-MS, and NMR data analysis. MPP also integrates smoothly with Agilent MassHunter Workstation, Spectrum Mill, ChemStation, and AMDIS software and is the only platform that provides integrated identification/ annotation of compounds and integrated pathway analysis for metabolomic and proteomic studies.

This section covers the following features of analyzing your data using MPP:

- “Overview of a metabolomics experiment using MPP” on page 75
- “Elements to consider in planning your experiment” on page 77
- “Create a new project and experiment” on page 79
- “Import and organize your data” on page 84
- “Perform your initial analysis” on page 99
- “Save your project” on page 113
- “Functional areas of MPP” on page 114

Overview of a metabolomics experiment using MPP

MPP is used to import, organize, and analyze the data you acquired. A metabolic profiling experiment may include the following steps: (1) prepare for your experiment, (2) acquire your data, (3) find the spectral features, (4) import and organize your data, (5) create your initial analysis, (6) identify the features, (7) save your project, and (8) perform advanced analysis operations. [Figure 19](#) on page 76 shows the Agilent tools used in metabolic profiling. MPP is used for steps 4 through 8.

MPP helps you analyze your metabolomics data through the use of sequential dialog boxes and wizards. The sequential tasks performed using MPP are shown in Figure 20 on page 76.

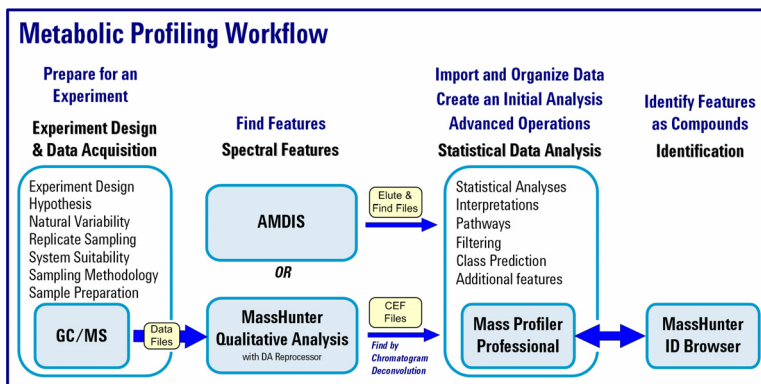


Figure 19 The steps involved in a metabolic profiling experiment using Mass Profiler Professional

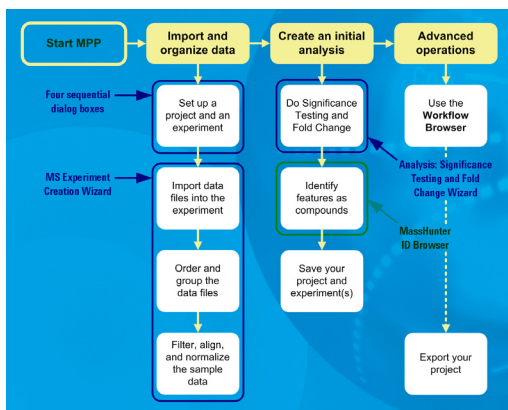


Figure 20 Overview of the wizards that help you use MPP

Elements to consider in planning your experiment

This section helps you understand an experiment and steps you can take to improve your results before analyzing your data with MPP.

An experiment consists of the analysis of a set of replicate samples collected over a range of well defined parameters, treatments, and/or exposures known as independent variables, including parameter controls representing minimal or normal perturbations (control samples). The results from the changes observed in the samples are designed to provide an answer to your hypothesis. The hypothesis may be proved or disproved by analyzing the correlation of the independent variables on the resulting expression of a large number of dependent variables - the features (compounds) that are measured in your samples. The results must be significant beyond natural variability.

After you obtain your samples, acquire your data, and find the features in your sample data, MPP takes you through data extraction, processing, and statistical analysis so that you can prove or disprove your hypothesis.

The hypothesis

The hypothesis is the question that is answered by your analysis. For example, the question may be a statement that proposes a possible correlation, or cause and effect, between a set of independent variables and the resulting features in your data.

Natural variability

It is important to understand how any one sample in your data represents the population as a whole. Because of natural variability and the uncertainties associated with both the measurement and the population, no assurance exists that any single sample from a population represents the mean of the population. Thus, increasing the sample size greatly improves the accuracy of the sample set in describing the characteristics of the population.

Replicate sampling

Sampling the entire population is not typically feasible because of constraints imposed by time, resources, and finances. On the other hand, fewer samples increase the probability of making a false positive or false negative correlation.

System suitability

System suitability involves collecting data to provide you with a means to evaluate and compensate for drift and instrumental variations to assure quality results. Techniques employed by Agilent and AMDIS software include (1) retention time alignment, (2) intensity normalization, (3) chromatographic deconvolution, and (4) baselining to produce the highest quality results. The best results are achieved by maintaining your instrument and using good chromatography.

Sampling methodology

Improved data quality comes from matching the sampling methodology to the experimental design so that replicate data is collected to span the parameter values for each parameter. A larger number of samples appropriate to the population under study results in a better answer to your hypothesis. An understanding of the methodologies used in sampling and using more than one method of sample collection have a positive impact on the significance of your results.

Create a new project and experiment

When you launch MPP you are guided through four sequential dialog boxes (Figure 21) to create a new project and experiment to receive your data:

Step 1. Startup: Select the option to create a new project.

Step 2. Create New Project: Enter descriptive information about your project.

Step 3. Experiment Selection: Select the option to create a new experiment as part of your project.

Step 4. New Experiment: Set up the information to store with your experiment and to guide the analysis process.

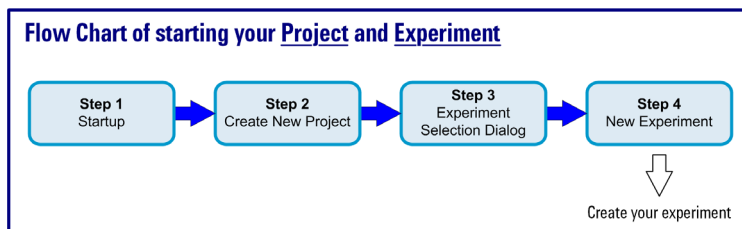


Figure 21 Steps to start a new project and an experiment

Step 1: Startup

- 1 Click **Create new project** in the **Startup** dialog box after launching MPP.

If MPP is already open click **Project > Close Project** to close and automatically save any open analysis. Then click **Project > New Project** to begin creating a new project at “Step 2: Create New Project”.

- 2 Click **OK**.

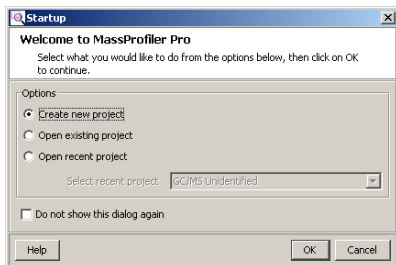


Figure 22 Startup dialog box

Step 2: Create New Project

- 1 Type a descriptive, unique name for your project in **Name** in the **Create New Project** dialog box.
- 2 Type information about your project in **Notes**.
- 3 Click **OK**.

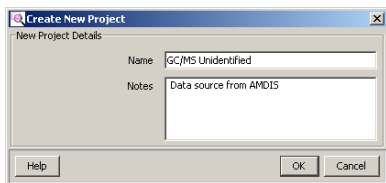


Figure 23 Create New Project dialog box

Step 3: Experiment Selection

- 1 Click **Create new experiment** in the **Experiment Selection Dialog** dialog box.
- 2 Click **OK**.

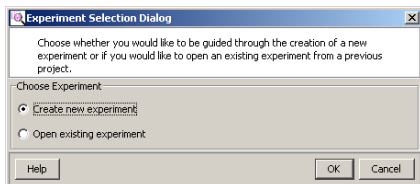


Figure 24 Experiment Selection Dialog dialog box

Step 4: New Experiment

A summary table of selections and entries for the **New Experiment** dialog box is shown in [Table 2](#) on page 82.

- 1 Type a descriptive name for your experiment in **Experiment name** in the **New Experiment** dialog box.
- 2 Select **Mass Profiler Professional** for **Analysis Type**.
- 3 Select the **Experiment type** that matches your source data.

If your data has been process using AMDIS:

- Select **Identified** to use AMDIS find data files.
- Select **Unidentified** to use AMDIS elute data files.
- Select **Combined (Identified + Unidentified)** to use both AMDIS data file types.

If your data has been process using Qualitative Analysis:

- Select **Identified** when the compounds have been identified by compound, formula, and/or CAS number.
- Select **Unidentified** when the compounds have only been identified by Find by Chromatogram Deconvolution resulting in mass and retention time.
- Select **Combined (Identified + Unidentified)** when you are unsure if the data has been identified in full or in part, or when MassHunter Qualitative Analysis has been previously used to identify some of the compounds.

A summary table of data sources and file types based on the selection for **Experiment type** is shown in [Table 3](#) on page 83.

- 4 Select **Analysis: Significance Testing and Fold Change** for **Workflow type**.

Regardless of your personal expertise, it is recommended to select the **Analysis: Significance Testing and Fold Change** for the **Workflow type** to provide you with quality control for your analysis that improves your results.

- 5 Type information about your experiment in **Notes**.
- 6 Click **OK**.

5 Analyzing your Data

Figure 25 New Experiment dialog box

Dialog Box Field	Your Choices	Comments
Experiment name	<none>	Edit field to describe this experiment
Analysis type	Mass Profiler Professional <other choices depending on Order IDs>	"Mass Profiler Professional" must be selected
Experiment type	Unidentified Identified Combined (Identified + Unidentified)	<see the next table for more information>
Workflow type	Analysis: Significance Testing and Fold Change Class Prediction: Build and Test Model Data Import Wizard	"Analysis: Significance Testing and Fold Change" is recommended to provide quality control and improve your results
Experiment notes	<none>	Edit field to enter other experimental notes

Table 2 Summary of selections and entries in the New Experiment dialog box

Experiment Type	Data Source	File Types	Comments
Identified	MH Quant		Compounds identified by MassHunter Quantitative Analysis
	Chemstation	.fin	Compounds identified by Chemstation Quantification or Screener processes
	MH Qual	.cef	Qualitative Analysis, Find by Formula (FbF)
	MH Qual (GC Scan)	.cef	Identify by Unit Mass Library
	ICP-MS	.csv	Identified by ICP-MS software
	AMDIS	.fin	Compound identified by an AMDIS target library
	Generic	.xls .xlsx .csv .txt	Entries identified by Compound (column C), Formula (column D), CASID (column E)
Unidentified	MH Qual	.cef	Find By Molecular Feature Extractor (MFE)
	MH Qual (GC Scan)	.cef	Find by Chromatographic Deconvolution
	ICP-MS	.csv	Identified by ICP-MS software
	AMDIS	.elu	Components identified by AMDIS that are not identified by an AMDIS target library
	Generic	.xls .xlsx .csv .txt	Entries NOT identified by Compound (column C), Formula (column D), CASID (column E)
	Combined	MH Qual	.cef
MH Qual (GC Scan)		.cef	Find by Chromatographic Deconvolution and Library Search
ICP-MS		.csv	Identified by ICP-MS software
AMDIS		.fin .elu	Targets and components discovered by AMDIS
Generic		.xls .xlsx .csv .txt	Combination of entries identified by and not identified by Compound (column C), Formula (column D), CASID (column E)

Table 3 Summary of data sources and file types based on the selection for Experiment Type in the New Experiment dialog box

If you selected **Analysis: Significance Testing and Fold Change** or **Data Import Wizard** for the **Workflow type** in the **New Experiment** dialog box, the MS Experiment Creation Wizard is automatically started after this step.

Import and organize your data

After you set up your project and create an experiment, the MS Experiment Creation Wizard (Figure 26) immediately guides you through the nine (9) of eleven (11) steps to organize your experiment, import your data, define your experiment variables, and prepare your data for analysis; data preparation includes grouping, filtering, alignment, normalization, and baselining.

Step 1. Select Data Source: Select the data source that generated the molecular features you are using for your experiment.

Step 2. Select Data to Import: Select the molecular feature sample files.

Step 3 and Step 4 are skipped by the MS Experiment Creation Wizard for GC/MS data.

Step 5. Sample Reordering: Organize your samples by selecting individual samples and reordering the selection to group the samples based on the independent variables.

Step 6. Experiment Grouping: Define the sample grouping with respect to your independent variables, including the replicate structure of your experiment.

Step 7. Filtering: Filter the molecular features by abundance, mass range, number of ions per feature, and charge state.

Step 8. Alignment: Align the features across the samples based on tolerances established by retention time and mass. This step is omitted when the experiment type is “identified” because identified compounds are treated as aligned by identification.

Step 9. Sample Summary: Display a mass versus retention time plot, spreadsheet, and compound frequency for the distribution of aligned and unaligned entities in the samples. Compound Frequency charts provide a quick view into the effectiveness of the alignment of unidentified experiment types. The back and next buttons in the wizard let you easily review the effects of different alignment and filter options.

Step 10. Normalization Criteria: Scale the signal intensity of sample features to a value calculated by the specified algorithm or an external scalar.

Step 11. Baseline Options: Compare the signal intensity of each sample to a representative value calculated across all of the samples or the control samples.

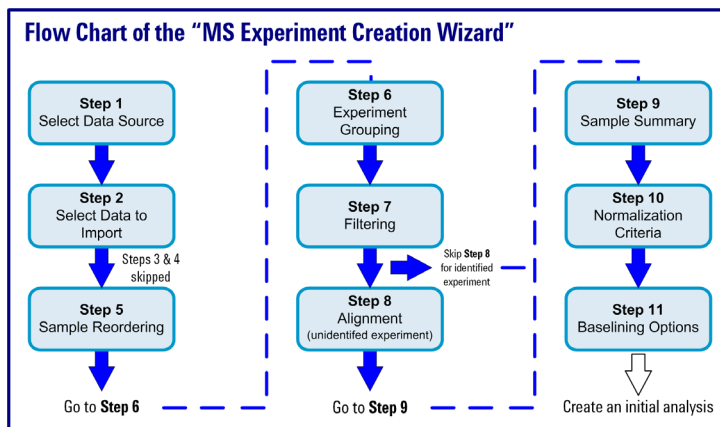


Figure 26 Steps performed by the MS Experiment Creation Wizard

Step 1 of 11: Select Data Source

- 1 Select the source of your data files. The available data source selection depends on the **Experiment type** you selected in the **New Experiment** dialog box (Figure 25 on page 82).

For metabolomics analyses using GC/MS data, you typically select **MassHunter Qual**, **Chemstation**, or **AMDIS**.

- 2 Select the **Organism** represented by your experiment.

Selecting an **Organism** is most important when you use the Pathway Analysis features of MPP.

- 3 Click **Next**.

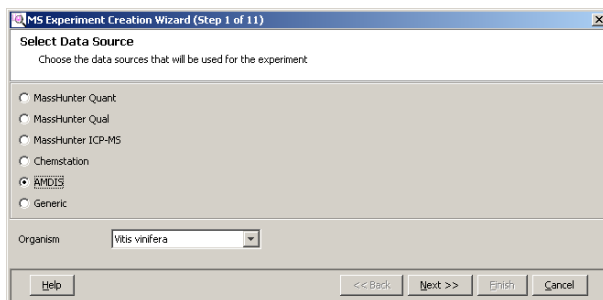


Figure 27 Select Data Source page (Step 1 of 11) of the MS Experiment Creation Wizard

Step 2 of 11: Select Data Source

- 1 Click **Select Data Files** to display the file selection dialog box (Figure 28).
- 2 Select files you want to import into the experiment in the **Open** dialog box.

Data processed using the Agilent Fiehn Library results in two output files that can be used for statistical analysis: elute data file (.elu) and find compound data file (.fin).

- The elute file contains a detailed ion and abundance data listing for all of the peaks (features) found and is imported with **Identified** and **Combined (Identified + Unidentified)** experiment types.

Note: The .fin files contain some AMDIS parameters that are important to properly importing the unidentified data contained on the .elu files. When you import .elu data files for an **Unidentified** experiment, the .fin files must also be located in the same folder in order for the sample files to import properly.

- The find compound file includes all of the information contained in the elute file and additionally contains the Agilent Fiehn Library search assignment for each peak (feature) and is imported with **Unidentified** and **Combined (Identified + Unidentified)** experiment types.

Data files from single quadrupole MSDs (mass selective detectors) may also be processed using Agilent MassHunter Qualitative Analysis. Qualitative Analysis generates a single output compound exchange format (CEF) file (.cef).

- The CEF file is imported with **Identified**, **Unidentified**, and **Combined (Identified + Unidentified)** experiment types.
- 3 Click **Open** to load the selected files.
 - 4 Review the selected files and use **Select Data Files** or **Select Samples** to adjust the selected data files based on your experiment and/or advanced search parameters.
 - 5 Click **Next**.

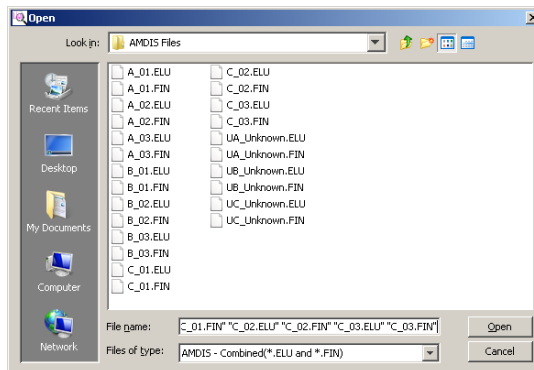


Figure 28 Data file selection dialog box

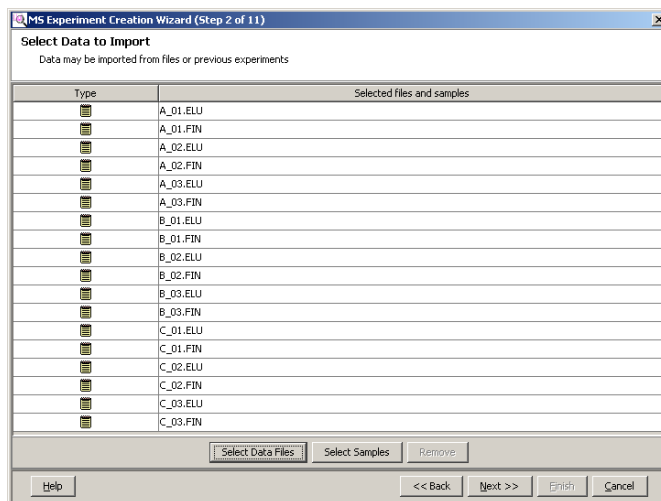


Figure 29 Select Data to Import page (Step 2 of 11) of the MS Experiment Creation Wizard

Step 5 of 11: Sample Reordering

- 1 Click one or more samples that you want to reorder. A selected sample is highlighted using a background color for the sample row.
- 2 Reorder the samples as you like using the reorder buttons up, down, and restore as often as necessary.
- 3 Mark the sample names using the **Select** column that you want to import into your experiment.
- 4 Click **Next**.

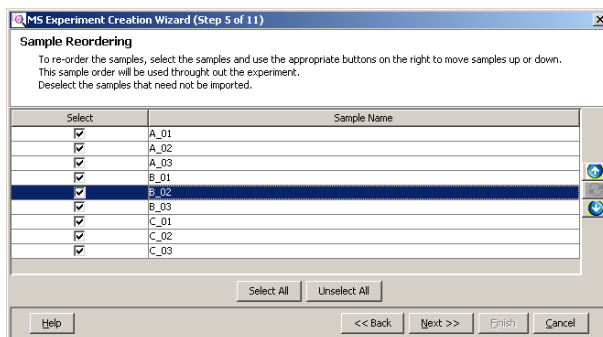



Figure 30 Sample Reordering page (Step 5 of 11) of the MS Experiment Creation Wizard

Step 6 of 11: Experiment Grouping

1 Click Add Parameter.

You can click the **Load experiment parameters from file** button  to apply a previously saved experiment grouping.

If you use a previously saved experiment grouping you can skip the remaining steps and click **Next**.

2 Type a name for your **Parameter name** in the **Add/Edit Experiment Parameter** dialog box.

An independent variable is an essential element, constituent, attribute, or quality in a data set that is deliberately controlled in your experiment. An independent variable is referred to as a parameter and is assigned a parameter name.

The attribute values within an independent variable are referred to as parameter values. Samples with the same parameter value are treated as replicates.

3 Select the **Parameter type** for your grouping.

- Select **Non-Numeric** if the grouping is not a quantitative value.
- Select **Numeric** if the grouping value is quantitative or a value that reflects a degree of proportionality among the samples with respect to an independent variable. A numeric parameter type allows some data plots in MPP to be scaled by the parameter values.

- 4 Click your replicate **Samples** that share the same first parameter value in your data. A selected sample is highlighted using a background color for the sample row.
- 5 Click **Assign Value**.

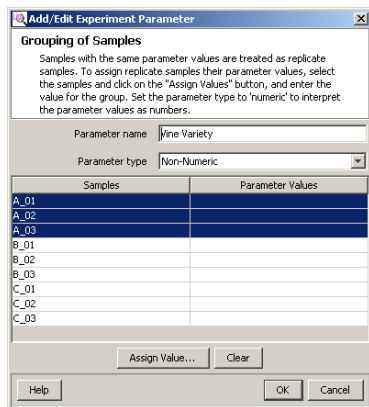


Figure 31 Add/Edit Experiment Parameter dialog box

- 6 Type the value for your first grouping in the **Assign Value** dialog box.
- 7 Click **OK**.

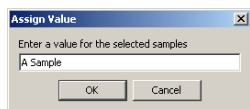


Figure 32 Assign Value dialog box

- 8 Click your replicate **Samples** that share the same second parameter value in your data.
- 9 Click **Assign Value**.
- 10 Type the value for your second grouping in the **Assign Value** dialog box.
- 11 Click **OK**.
- 12 Review your entries and grouping assignment accuracy in the **Add/Edit Experiment Parameter** dialog box. Repeat the **Sample** selection and

Assign Value steps until all of your samples are assigned a parameter value.

13 Click **OK** when the grouping for this **Parameter name** is complete.

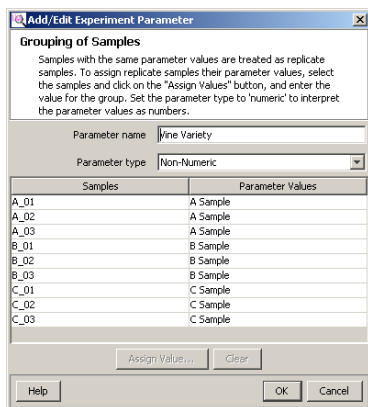



Figure 33 Add/Edit Experiment Parameter dialog box

14 Repeat **Add Parameter** if your data has more than one independent variable.

- Click **Add Parameter**.
- Repeat the **Sample** selection and **Assign Value** steps until all of your samples are assigned a parameter value for the new **Parameter name**.

15 Click the **Save experiment parameters to file** button  to save your experiment parameters and parameter values to a file. Saving your experiment parameters to a file saves time when you create multiple experiments.

16 Click **Next** when you have assigned a parameter name, type, and value to all of your data.

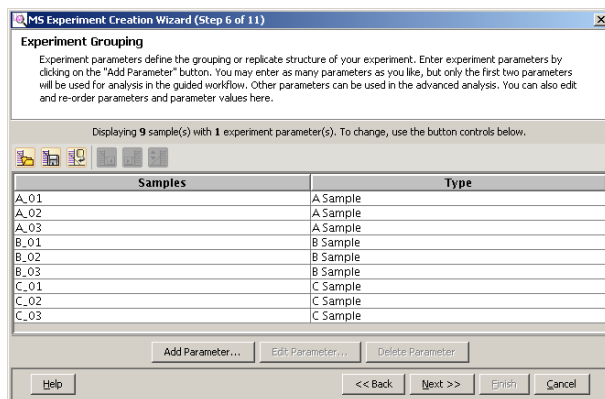


Figure 34 Experiment Grouping page (Step 6 of 11) of the MS Experiment Creation Wizard

Step 7 of 11: Filtering

1 Mark an **Abundance filtering** option.

- **Minimum absolute abundance** limits the data that is imported to absolute abundances greater than the specified value. The abundance counts can be between 1 and 1,000,000,000.
- **Limit to the largest** limits the number of compounds that imported for each sample file to the number specified. The compounds are selected starting with the highest abundance value in descending order. The number of compounds can be between 1 and 999.
- **Minimum Relative Abundance** filters out the compounds in a sample file that have a relative abundance less than the value specified with respect to the most abundant compounds in the file. The relative abundance is calculated by dividing abundance value for each compound in a sample by the abundance value of compound with the largest abundance in the same sample. The percentage can be between 1 and 100%.

For GC/MS datasets the **Minimum Relative Abundance** filtering is performed on identified and unidentified compounds separately.

2 Mark **Use all available data** or enter the retention time range for the compounds to import for **Retention time filtering**.

- 3 Mark **Use all available data** or enter the mass range for the compounds to import for **Mass filtering**.
- 4 Enter the minimum number of model ions required in order for a compound to be imported in **Number of ions required** \geq . For AMDIS data the default is five (5).
- 5 Select the applicable **Charge states** parameter allowed for your experiment. This parameter is not available for AMDIS data sets.
- 6 Enter the **Minimum Quality Score** for the **Compound Quality Score Filter**. Compounds with a quality score less than the value specified are filtered out. The quality score value can be between 1 and 100.

This parameter is only available for AMDIS experiments.

$$\text{Quality Score} = a * \text{MO} + b * \text{SNR} + c * \text{Log}_{10}(\text{abundance})$$

MO is the number of model ions and SNR is the signal to noise ratio. The coefficients a, b, and c are 1 by default. Click **Tools > Options** to launch the **Configuration Dialog** dialog box. Then click **MS > AMDIS Compound Quality Score Parameters** to adjust the weighting coefficients used by the quality score calculation.

- 7 Click **Next**.

The screenshot shows the 'MS Experiment Creation Wizard (Step 7 of 11)' window. The title bar includes a search icon and the text 'MS Experiment Creation Wizard (Step 7 of 11)'. The main content area is titled 'Filtering' and contains the following sections:

- Abundance filtering:**
 - Minimum absolute abundance: 10000 counts
 - Limit to the largest: [] compounds
 - Minimum relative abundance: [] %
- Retention time filtering:**
 - Use all available data
 - Min RT (0.78): 0.78
 - Max RT (60.0): 60.0
- Mass filtering:**
 - Use all available data
 - Min Mass (28.9): 28.9
 - Max Mass (263.1): 263.1
- Number of ions:**
 - Number of ions required \geq : 3
- Charge states:**
 - All charge states permitted
 - Multiple charge states required
 - Multiple charge states forbidden
- Compound Quality Score Filter:**
 - Minimum Quality Score: 25

At the bottom of the window, there are buttons for 'Help', '<< Back', 'Next >>', 'Finish', and 'Cancel'.

Figure 35 Filtering page (Step 7 of 11) of the MS Experiment Creation Wizard

Step 8 of 11: Alignment Parameters

This step is applicable for **Unidentified** experiment types. **Identified** compounds are treated as aligned by identification and this step is omitted.

The formula for performing the alignment is described in section “3.2.7 Alignment Parameters (Step 8 of 11)” in the *Mass Profiler Professional User Manual*. Click **Tools > Options** to launch the **Configuration Dialog** dialog box. Then click **MS > GCMS Peak Alignment Parameters** to adjust the calculation coefficients.

- 1 Type a value for **Retention Time Tolerance (min)**. The default value is 0.05 min.
- 2 Type a value for **Match Factor**. The default value is 0.3, and the value can be between 0 and 1. 0 denotes completely different spectra, while a value of 1 denotes identical spectra.
- 3 Type a value for **Delta MZ (Low Resolution)**. The default value is 0.2 m/z for low resolution spectra. For high resolution spectra the default value is 0.005 m/z .
- 4 Click **Next**.

The screenshot shows a dialog box titled "MS Experiment Creation Wizard (Step 8 of 11)" with a sub-header "Alignment Parameters". Below the sub-header is a descriptive paragraph: "Unidentified compounds from different samples are aligned or grouped together if their retention times are within the specified tolerance window and the mass spectral similarity as determined by a simple dot product calculation above the specified level." There are three input fields: "Retention Time Tolerance (min)" with a value of 0.08, "Match Factor" with a value of 0.3, and "Delta MZ (Low Resolution)" with a value of 0.2. At the bottom of the dialog box, there are five buttons: "Help", "<< Back", "Next >>", "Finish", and "Cancel".

Figure 36 Alignment Parameters page (Step 8 of 11) of the MS Experiment Creation Wizard

Step 9 of 11: Sample Summary

- 1 Clear the **Export For Recursion** check box.
- 2 Review the table and chart in the **Compound Frequency** tab.
- 3 Assess the quality of your sample alignment by reviewing the frequency of compounds that appear in only one sample.

A “one-hit wonder” is a compound that appears in only one sample and is absent from the replicate samples. Therefore, a “one-hit wonder” compound does not provide any utility for statistical analysis and you want to filter such compounds from your analysis.

If most of the “one-hit wonders,” the number of compounds in the compound frequency table with unit frequency (Frequency = 1), have a low relative abundance, less than 10% and ideally less than 1%, your sample data alignment is likely good. If the “one-hit wonders” have a high relative abundance (i.e., in the 30-100% column) then you may need to improve your sample data alignment. See [Figure 37](#) on page 95.

4 Click Next.

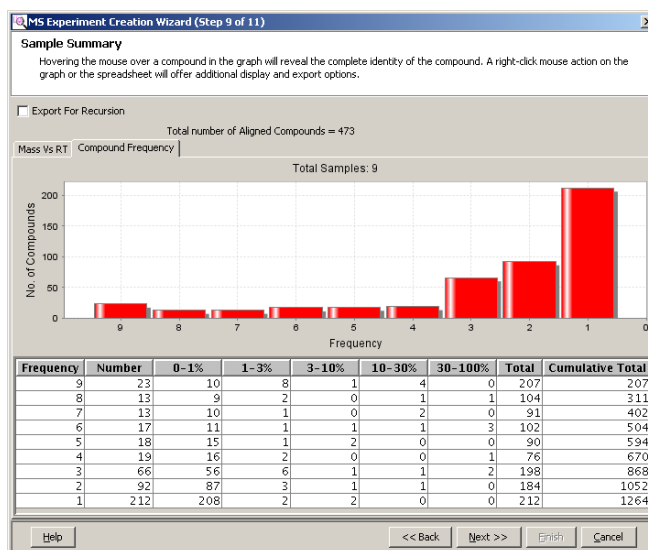


Figure 37 Compound Frequency view in the Sample Summary page (Step 9 of 11) of the MS Experiment Creation Wizard

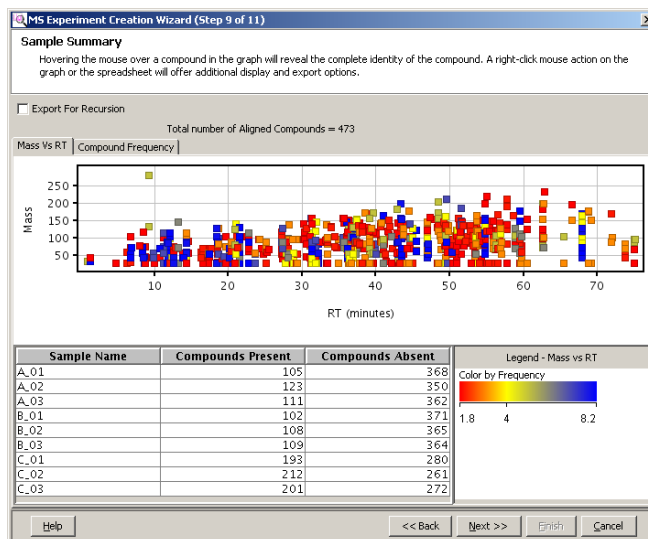


Figure 38 Mass vs. RT view in the Sample Summary page (Step 9 of 11) of the MS Experiment Creation Wizard

Step 10 of 11: Normalization Criteria

1 Select the **Normalization Algorithm** in the **Normalization** tab.

- **Internal Standard** may be applied only to a compound whose frequency equals the number of samples in the experiment are available for selection. An internal standard is a compound added to each sample before your GC/MS acquisition and used to reduce the abundance variability caused by sample preparation and instrument response.

Select the compound you want to use for the internal standard from the table that appears after selecting **Internal Standard**.

- **Percentile Shift** is an experiment-wide normalization that adjusts the abundance of all of the compounds in all of your samples by the abundance value that represents the entered percentile. The abundance for the percentile entered is subtracted from the abundance of each entity. The **Percentile Shift** value can be between 0 and 100. A value of 50 represents the median abundance.

- **Quantile** normalization is a method of normalization that makes the distribution of abundance values of all samples in an experiment identical. After normalization, all statistical parameters for each sample (i.e., mean, median, and percentile) are identical. Quantile normalization reduces variance between samples.
 - **None** does not perform any normalization. The abundance values are transformed to a \log_2 scale.
- 2 Mark **Use External Scalar** in the **External Scalar** tab. If you use an external scalar, enter the value to scale each sample file in the **Scale to Value** column.

The **Normalization** options can be applied alone or in conjunction with **External Scalar**. When both are selected, then External Scalar is applied first followed by the chosen Normalization option.

- 3 Click **Next**.

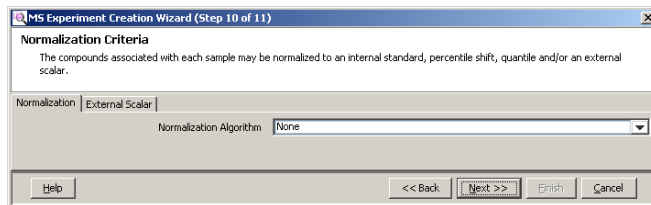


Figure 39 Normalization Criteria page (Step 10 of 11) of the MS Experiment Creation Wizard

Step 11 of 11: Baseline Options

- 1 Select the baselining **Option**.
 - Select **None** if only a few features in the samples exist.
 - Select **Z-Transform** if the data sets are very dense, i.e., with data where very few instances of compounds are absent from any sample, such as a quantitation data set from recursion.
 - Select **Baseline to ___ of all samples** to reduce the effect of very abundant and very weak compound features on later statistical analyses. The abundance for each compound is normalized to its selected statistical abundance across all of the samples.

- Select **Baseline to ___ of control samples** to weight the compound features to a known value that is considered to be normal in the population while reducing the effect of large and small compound features. The abundance for each compound is normalized to its selected statistical abundance across just the samples selected as the control samples.

When you use this baselining option you must select the samples that are the control samples from the table.

2 Click **Finish**.

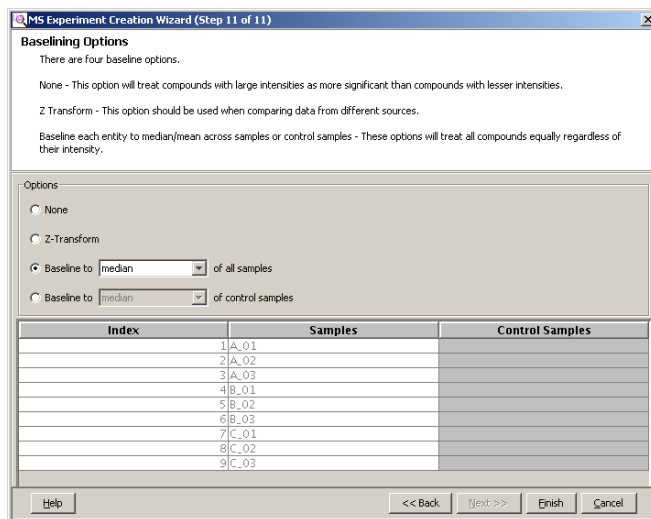


Figure 40 Baselining Options page (Step 11 of 11) of the MS Experiment Creation Wizard

If you chose **Data Import Wizard** for **Workflow type** in the **New Experiment** dialog box you are done creating your experiment and can use the operations available in the Workflow Browser

If you chose **Analysis: Significance Testing and Fold Change** for **Workflow type** in the **New Experiment** dialog box you are immediately directed to “Perform your initial analysis” on page 99.

Perform your initial analysis

The **Analysis: Significance Testing and Fold Change Wizard** guides you through the eight (8) steps to enter parameters and values that improve the quality of your results and produce an initial differential expression for your analysis. The steps performed during the **Significance Testing and Fold Change Wizard** are illustrated in [Figure 41](#).

Step 1. Summary Report: Displays a summary view of your experiment based on the parameters you provided in the Import Data wizard. A profile plot with the samples on the x-axis and normalized abundance values on the y-axis is displayed. If the number of samples is more than 30, the data is represented by a spreadsheet view instead of a profile plot.

Step 2. Experiment Grouping: Independent variables and the attribute values of the independent variables must be specified to define grouping of the samples. An independent variable is referred to as a parameter name. The attribute values within an independent variable are referred to as parameter values. Samples with the same parameter values within a parameter name are treated as replicates.

Step 3. Filter Flags: The compounds created during the experiment creation are now referred to as entities. The entities are filtered (removed) from further analysis based on their presence across samples and parameter values (now referred to as a condition).

It is recommended to process your data through at least through this step. The Filter Flags step is used to require that an entity must be present in at least two samples in your data. This requirement removes entities referred to as “one-hit wonders.” A “one-hit wonder” is an entity that appears in only one sample, is absent from the replicate samples, and does not provide any utility for statistical analysis.

Step 4. Filter by Frequency: Entities are further filtered based on their frequency of presence in specified samples and conditions. This filter removes irreproducible entities.

Step 5. Quality Control on Samples: The samples are presented by grouping and the current Principal Component Analysis (PCA). PCA calculates all the possible principal components and visually represents them in a 3D scatter plot. The scores shown by the axes scales are used

to check data quality. The scatter plot shows one point per sample colored-coded by the experiment grouping. Replicates within a group should cluster together and be separated from samples in other groups

Step 6. Significance Analysis: The entities are filtered based on their p-values calculated from a statistical analysis. The statistical analysis performed depends on the samples and experiment grouping.

Step 7. Fold Change: Compounds are further filtered based on their abundance ratios or differences between a treatment and a control that are greater than a specified cut-off or threshold value.

Step 8. ID Browser Identification: You have the option to import your final entity list directly into ID Browser for identification.

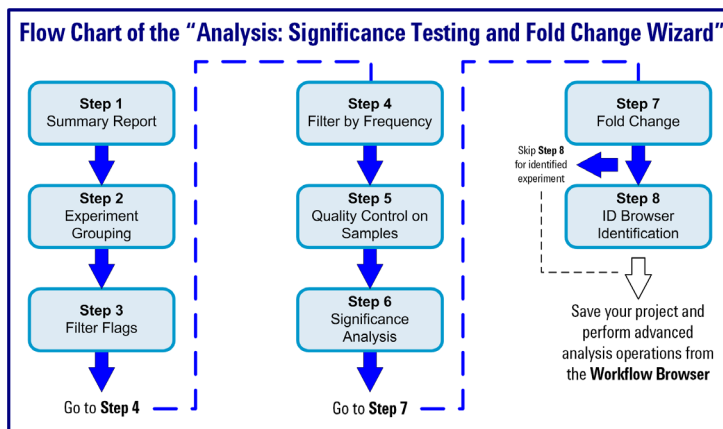


Figure 41 Steps performed by the Analysis: Significance Testing and Fold Change Wizard

Step 1 of 8: Summary Report

1 Review the Summary Report.

The Summary Report is displayed as a spreadsheet view when you have more than 30 samples

2 Click and right-click features on the plot, or spreadsheet, to review the data, change the plot view, export selected data, or export the plot to a file.

3 Click **Next**.

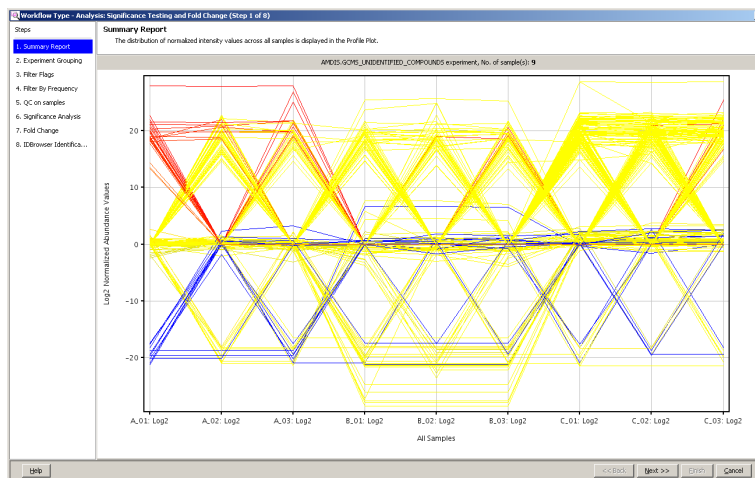


Figure 42 Summary Report page of the Analysis: Significance Testing and Fold Change (Step 1 of 8) workflow

Step 2 of 8: Experiment Grouping

- 1 Review your **Experiment Grouping**.
- 2 Click **Add Parameter** to define or adjust your experiment grouping.
- 3 Follow the steps in “Step 6 of 11: Experiment Grouping” on page 89.

Note: In order to proceed to the next step at least one parameter with two parameter values must be assigned.

- 4 Click **Next** when you have completed your experiment grouping.

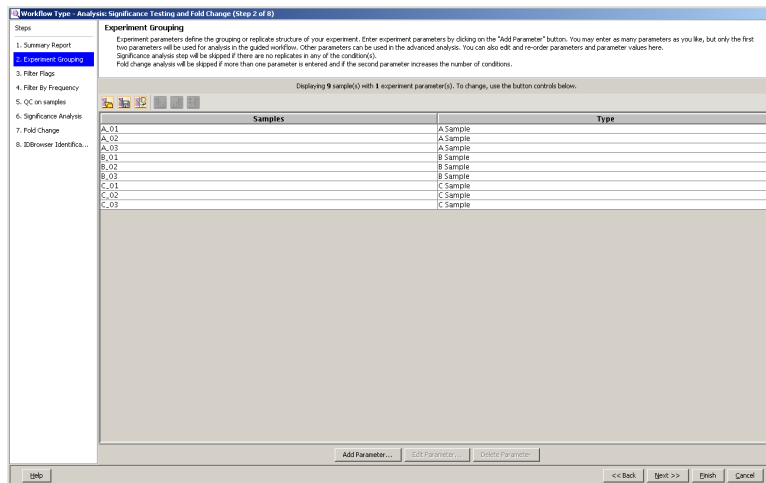


Figure 43 Experiment Grouping page of the Analysis: Significance Testing and Fold Change (Step 2 of 8) workflow

Step 3 of 8: Filter Flags

- 1 Review your data, change the plot view, export selected data, or export the plot to a file.
- 2 Click **Re-run Filter** to enter parameters in the **Filter Parameters** dialog box (Figure 45 on page 103).

A flag is used to denote the quality of an entity within a sample. A flag indicates if the entity was detected in each sample as follows:

- Present means the entity was detected.
- Absent means the entity was not detected.
- Marginal means the signal for the entity was saturated.

- 3 Mark the **Present** and **Marginal** check boxes.
- 4 Clear the **Absent** check box.
- 5 Click at least ___ out of X samples have acceptable values and type 2 in the entry box.

By setting this parameter to a value of two (2) or more, “one-hit wonders” are filtered.

- 6 Click **OK**.

This filter removes irreproducible entities from further consideration as you continue your analysis.

- 7 Review the profile plot and repeat the **Re-run Filter** until you obtain the best results for your experiment. Compare the entity summary above each profile plot in [Figure 44](#) and [Figure 46](#) on page 104; the displayed number of entities declined from 473 to 261 when “one-hit wonders” were removed.

- 8 Click **Next**.

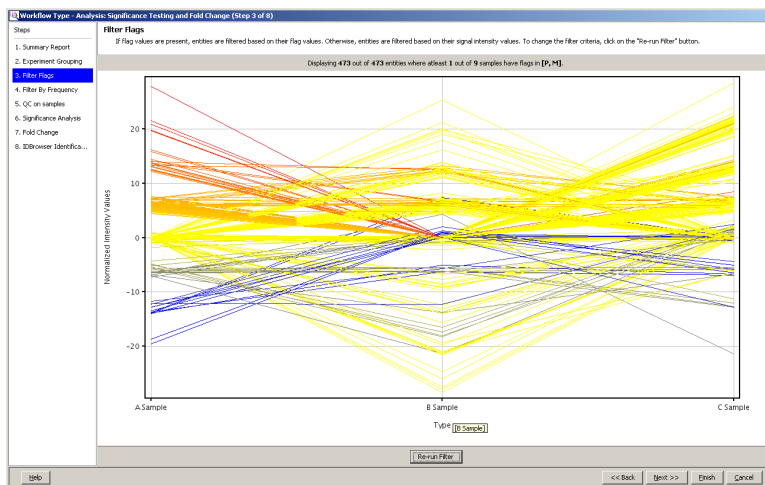


Figure 44 Filter Flags page of the Analysis: Significance Testing and Fold Change (Step 3 of 8) workflow before filtering “one-hit wonders”

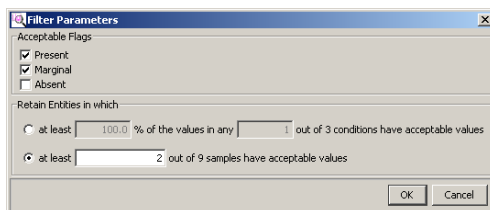


Figure 45 Filter Parameters dialog box

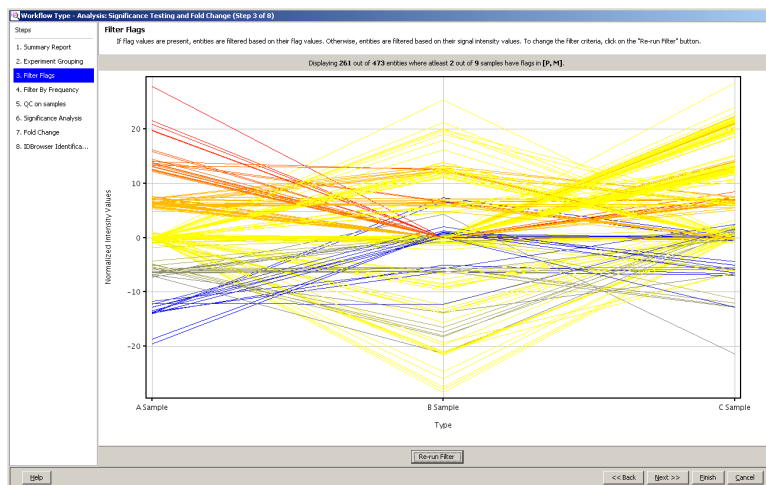


Figure 46 Filter Flags page of the Analysis: Significance Testing and Fold Change (Step 3 of 8) workflow after filtering “one-hit wonders”

Step 4 of 8: Filter by Frequency

- 1 Review your data, change the plot view, export selected data, or export the plot to a file.
- 2 Click **Re-run Filter** to enter parameters in the **Filter Parameters** dialog box (Figure 48 on page 105).

Filter Frequency allows you to filter the input data based upon the frequency with which any compound appears in each sample in the experiment. The filter is specified by typing the minimum percentage and selecting the applicable condition.

- 3 Type **100** in the **Retain entities that appear in at least** box.
- 4 Click of **samples in at least one condition**.
- 5 Click **OK**.
- 6 Review the profile plot and repeat the **Re-run Filter** until you obtain the best results for your experiment. Compare the entity summary above each profile plot in Figure 46 on page 104 and Figure 47; the displayed number of entities declined from 261 to 131.
- 7 Click **Next**.

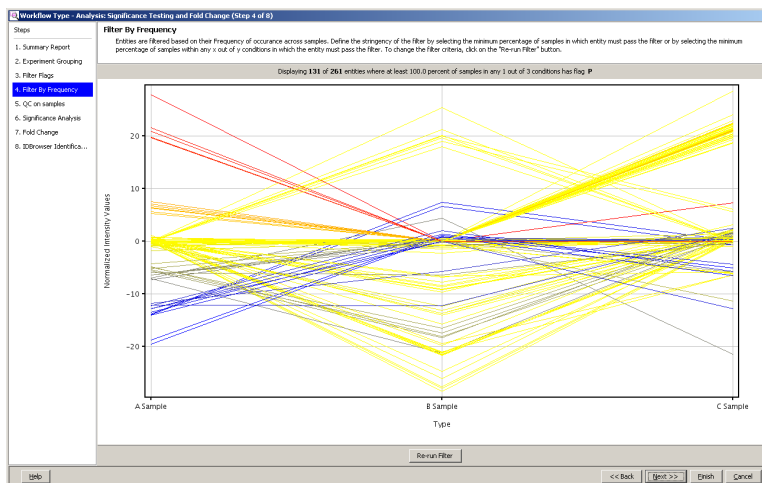


Figure 47 Filter by Frequency page of the Analysis: Significance Testing and Fold Change (Step 4 of 8) workflow

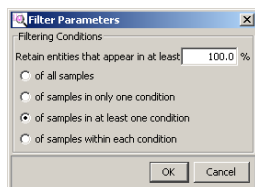


Figure 48 Filter Parameters dialog box

Step 5 of 8: QC on samples

- 1 Review your data, change the plot view, export selected data, or export the plot to a file.

QC on samples provides you with the first view of the data using a Principal Component Analysis (PCA). PCA allows you to assess the data by viewing a 3D scatter plot of the calculated principal components.

You want your samples to form discrete groups in the 3D PCA Scores view based on their parameter assignments.

- 2 Click **Back** to make adjustments to prior steps in the workflow to improve the PCA results.

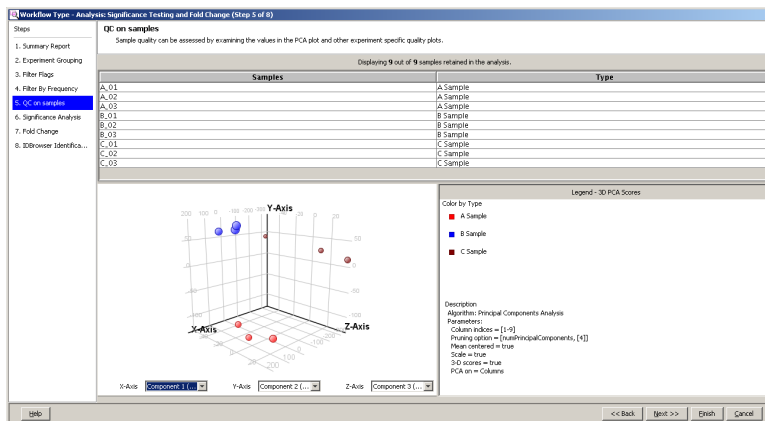
3 Click **Next**.

Figure 49 QC on samples page of the Analysis: Significance Testing and Fold Change (Step 5 of 8) workflow

Step 6 of 8: Significance Analysis

1 Review the summary plot.

The statistical analysis is either a T-test or an Analysis of Variance (ANOVA) based on your samples and experiment grouping.

2 Move the **p-value cut-off** slider or type a value to change the **p-value cut-off** value. A larger p-value passes a larger number of entities.

The last row of data in the **Result Summary** spreadsheet shows the number of entities that would be expected to meet the significance analysis by random chance based on the p-value specified in each column heading. If the number of entities **Expected by chance** is much smaller than those based on the **Corrected p-value**, your entities show significance among the parameter values.

3 Click **Next**.

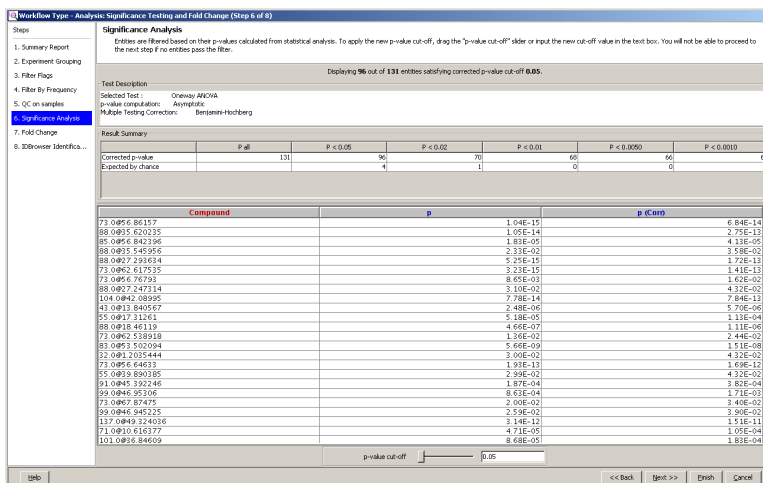


Figure 50 Significance Analysis page of the Analysis: Significance Testing and Fold Change (Step 6 of 8) workflow

Step 7 of 8: Fold Change

1 Review the summary plot.

Fold change is a signed value that describes how much an entity changes from its initial to its final value. For example, when an entity changes from a value of 60 to a value of 15, the fold change is -4. The quantity experienced a four-fold decrease. Fold change is the ratio of the final value to the initial value.

Fold change analysis is used to identify entities with abundance ratios, or, for example, differences between a treatment and a control, that are in excess of specified cut-off or threshold value. Fold change is calculated between the conditions where Condition 1 and another condition, Condition 2, are treated as a single group.

2 Move the **Fold change cut-off** slider or type a value to change the **Fold change cut-off**. The default value is 2.0. A larger cut-off value passes a smaller number of entities through to the final results.

3 Select a value for the **Minimum number of pairs** of conditions that must have entities with a fold change greater than the cut-off. The default value is 1.

- 4 Select the experiment grouping parameter value that is the **Control Group** for the fold change calculation.

Fold Change value per entity = |(Condition1 entity abundance) / (Condition 2 entity abundance)| where Condition 2 is the control group.

- 5 Click **Next**. An entity list is created in the Experiment Navigator for the entities that satisfy the cut-off parameters.

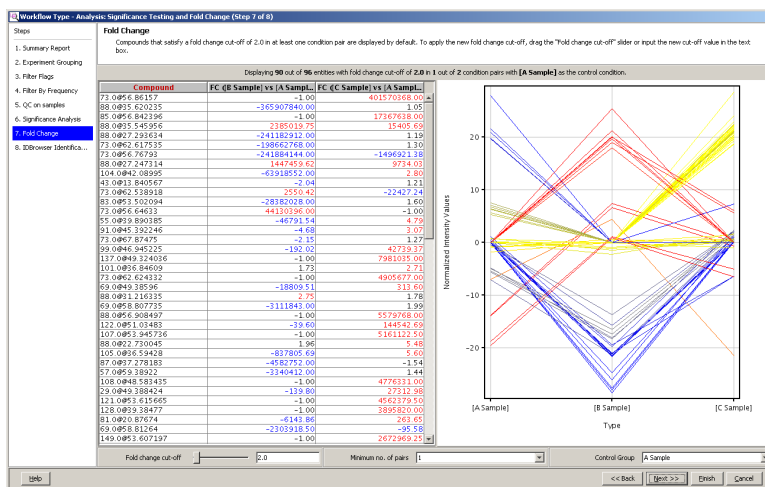


Figure 51 Fold Change page of the Analysis: Significance Testing and Fold Change (Step 7 of 8) workflow

Step 8 of 8: ID Browser Identification

- 1 Click **ID Browser Identification** to export your entity list to Agilent MassHunter ID Browser. ID Browser is started and automatically prompts you to set up your identification method parameters.

Processing your entities with ID Browser performs the following automatically: save the selected entity list into a CEF file, open Agilent MassHunter ID Browser, and import the saved CEF file for identification.

Once identification is completed, ID Browser returns an identified CEF file. This CEF file is imported into the MPP experiment and annotations are automatically updated.

- 2 Select the compounds to identify and mark the identification method for your experiment in the **Compound Identification Wizard** dialog box (Figure 53 on page 110).
- 3 Click **Next**.
- 4 Enter the path and file name for the Fiehn library in the **Spectral library path** in the **Settings** tab (Figure 54 on page 110).
- 5 Adjust the remaining parameters in the **Settings**, **Scoring**, and **Search Results** tabs to complete setting up your identification method.
- 6 Click **Finish** when you have the method set up for your experiment. ID Browser automatically begins identifying your entities and shows a progress bar.

Compound	FC (E Sample) vs.	Log FC (E Sample)	FC (abx (E Sample))	Regulation (E Sa.)	FC (C Sample) vs.	Log FC (C Sample)	FC (abx (C Sample))	Regulation (C Sa.)
88.0825.620235	-345.907840.00	-28.45	1.00	down	401370168.00	23.78	16.00	up
85.0825.642396	-1.00	0.00	1.00	down	17167618.00	24.05	16.00	up
88.0825.545095	2385017.75	7.13	16.00	up	15405.00	11.93	16.00	up
88.0827.231634	-241182311.00	-27.85	16.00	down	1.19	0.26	1.19	up
87.0826.217535	-130662768.00	-27.57	16.00	down	1.10	0.38	1.10	up
73.0825.676792	-241884144.00	-27.85	16.00	down	-1496921.18	-20.53	16.00	down
88.0827.247514	-1447459.62	-20.47	16.00	up	9774.03	13.25	16.00	up
104.0824.088995	-43238555.00	-25.92	16.00	down	2.80	1.48	2.80	up
43.0813.840567	-2.04	-1.03	2.04	down	1.21	0.27	1.21	up
73.0826.238938	2550423	11.32	16.00	up	-22487.24	-14.45	16.00	down
83.0823.502094	-28382028.00	-24.76	16.00	down	1.60	0.68	1.60	up
72.0825.64623	41130396.00	25.40	16.00	up	-1.00	0.00	1.00	down
25.0829.890385	-487915.4	-15.51	16.00	down	4.79	2.25	4.79	up
91.0825.392246	-4.88	-2.23	4.88	down	3.07	1.62	3.07	up
73.0827.67495	-2.15	-1.11	2.15	down	1.27	0.34	1.27	up
99.0846.845225	-132.02	-7.59	16.00	down	42739.37	15.38	16.00	up
117.0849.224056	-1.00	0.00	1.00	up	7981035.00	22.93	16.00	up
101.0825.846508	-1.73	0.78	1.73	up	2.11	1.44	2.71	up
73.0826.624332	-1.00	0.00	1.00	down	4905677.00	22.23	16.00	up
63.0849.385396	-188095.3	-14.70	16.00	down	133.60	0.29	16.00	up
88.0821.216135	2.75	1.46	2.75	up	1.78	0.83	1.78	up
63.0825.607235	-2311841.00	-21.57	16.00	down	1.99	0.99	1.99	up
88.0825.608497	-1.00	0.00	1.00	down	5579768.00	25.43	16.00	up
112.0825.03483	-39.60	-5.11	16.00	down	144942.69	17.14	16.00	up
107.0825.145726	-1.00	0.00	1.00	down	536132.50	22.20	16.00	up
88.0822.730045	-1.96	0.97	1.96	up	5.48	1.45	5.48	up
105.0825.54628	-837807.69	-19.46	16.00	down	5.60	1.48	5.60	up
87.0827.278183	-4582752.00	-22.13	16.00	down	-1.54	-0.63	1.54	down
87.0825.38922	-2340412.00	-21.67	16.00	down	1.44	0.53	1.44	up
108.0848.293435	-1.00	0.00	1.00	down	4776311.00	22.19	16.00	up
29.0849.388424	-139.80	-7.13	16.00	down	27332.98	14.74	16.00	up
113.0825.615665	-1.00	0.00	1.00	down	4924789.50	21.12	16.00	up
118.0829.38477	-1.00	0.00	1.00	down	3895502.00	21.89	16.00	up
81.0820.87074	-6145.86	-12.78	16.00	down	253.65	1.04	16.00	up
63.0825.81264	-2302342.50	-21.14	16.00	down	-95.18	-0.58	16.00	down
149.0823.607197	-1.00	0.00	1.00	down	2672969.25	21.35	16.00	up
113.0846.9129	-1.00	0.00	1.00	down	2644952.50	21.89	16.00	up
63.0846.823502	-1.00	0.00	1.00	down	2928281.25	21.48	16.00	up
151.0827.11029	1.00	0.00	1.00	down	2463513.00	21.22	16.00	up

Figure 52 ID Browser page of the Analysis: Significance Testing and Fold Change (Step 8 of 8) workflow before identification

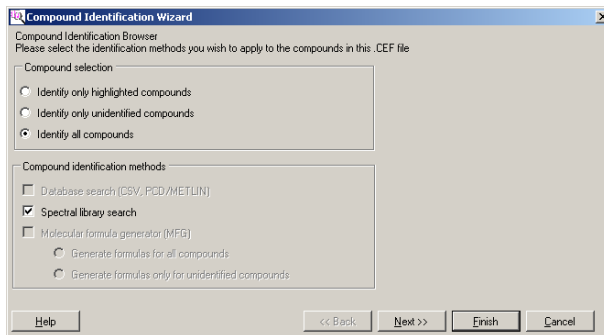


Figure 53 Compound Identification Wizard dialog box, compound

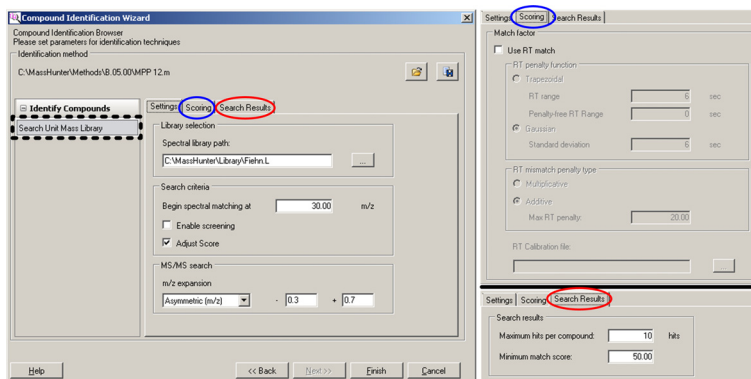


Figure 54 Compound Identification Wizard dialog box, method

- 7 Review and make adjustments to the entity identifications as necessary using the ID Browser interface (Figure 55 on page 111).
- 8 Click **Save and Return** to export your entity list back to your experiment in MPP. You are automatically returned to the MPP user interface.
- 9 Review your identified entity list in the ID Browser Identification results. The molecular formula now replaces the mass and retention time for identified entities in the compound column (Figure 56 on page 112).
- 10 Click **Finish** when you have completed the ID Browser Identification.

The Analysis: Significance Testing and Fold Change workflow is now complete and you are immediately returned to the main MPP interface similar to that shown in Figure 57 on page 112.

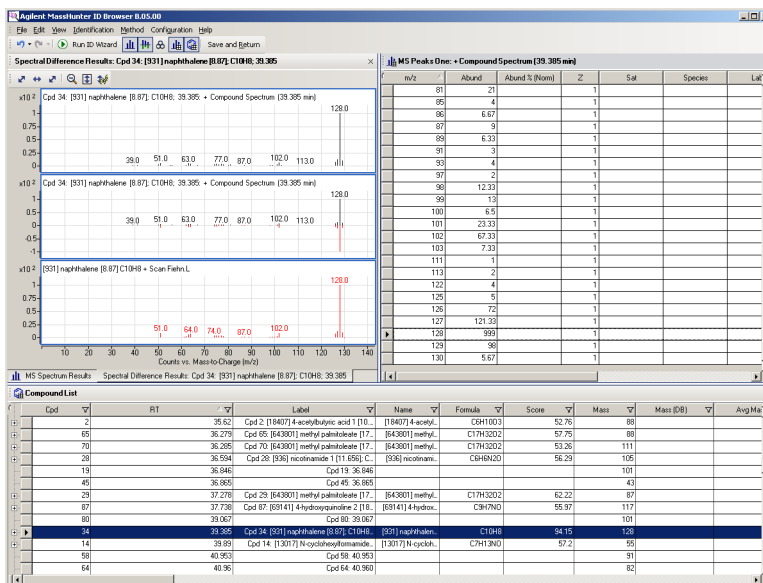


Figure 55 ID Browser user interface for reviewing the results

Workflow Type: Analysis: Significance Testing and Fold Change (Step 8 of 8)

Steps

- Summary Report
- Experiment Grouping
- Filter by Frequency
- QC on samples
- Significance Analysis
- Fold Change
- ID Browser**

ID Browser Identification

To identify the Entries that passed the fold change cut-off with ID browser click on the "ID browser Identification" button.

Identify Entries with ID browser	ID browser Identification
Compound	Regulation (C.S.)
(C) MetMyl Peargonate [2.248]	down
(184071)-acetylcholinic acid 1 [110.63]	down
(C) MetMyl Peargonate [2.248] 5.6.64...	down
(184071)-acetylcholinic acid 1 [110.63]	down
(C) MetMyl Capnitate [7.812]	down
(184071)-acetylcholinic acid 1 [110.63]	down
(C) MetMyl Peargonate [2.248] 5.6.74...	down
(C) MetMyl Capnitate [7.812] 27.2473...	down
104-0846-08995	down
43-0893-840567	down
(184071)-acetylcholinic acid 1 [110.63]	down
(130271)-cyclohexanofornamide [9.814]	down
(C) MetMyl Peargonate [2.248] 5.6.64...	down
(130271)-cyclohexanofornamide [9.814]	down
(18170) benzyl thioacetate [10.969]	down
(114) Methyl Myristate [5.597]	down
(12652)-squalene 1 [8.723]	down
137-0849-224036	down
103-0845-34609	up
(184071)-acetylcholinic acid 1 [110.63]	down
(130271) squalene [2.251]	down
(C) MetMyl Peargonate [2.248] 31.23...	down
(130271) squalene [2.251] 158.807...	down
(141031) methyl pantoic acid [17.466]	down
122-0851-01483	down
107-0845-346736	down
88-0822-730045	down
(25) nicotinamide 1 [11.956]	down
(141031) methyl pantoic acid [17.466]	down
(25) 191 tetraacetate [29.185]	down
108-0846-52415	down
(130271) squalene [2.251] 49.388...	down
(1742) N,N-dimethyl-L-phenylethyl...	down
(25) 191 tetraacetate [29.185]	down
81-0820-87674	down
(130271) squalene [2.251] 56.81264...	down
(1742) N,N-dimethyl-L-phenylethyl...	down
121-0846-3116	down
65-0846-303502	down

Figure 56 ID Browser page of the Analysis: Significance Testing and Fold Change (Step 8 of 8) workflow after identification

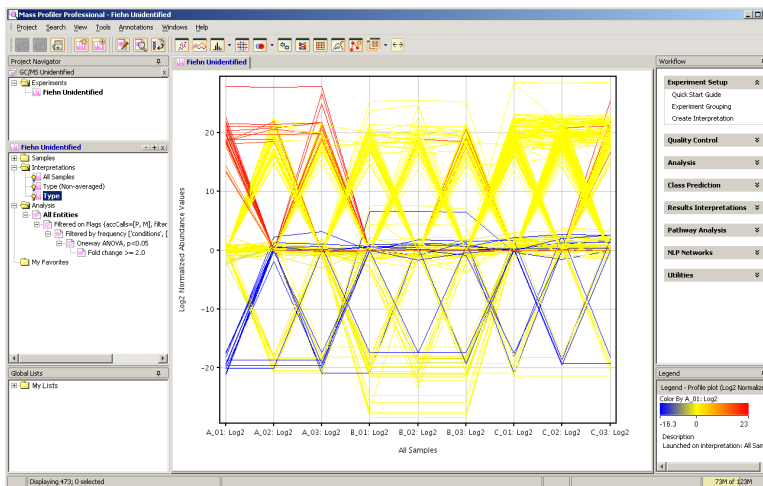


Figure 57 Example view of MPP after creating a new project and experiment.

Save your project

Save your current analysis as a TAR file for archiving, restoration of any future analysis to the current results, sharing the data with a collaborator, or sharing the data with Agilent customer support.

- 1 Click **Project > Export Project**.
- 2 Mark the check box next to the experiments you wish to save.
- 3 Click **OK**.

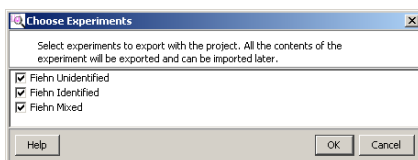


Figure 58 Choose Experiments dialog box

- 4 Select or create the file folder.
- 5 Type the **File name**.
- 6 Click **Save**.

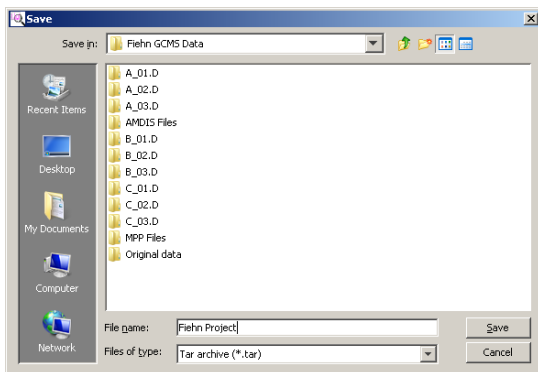


Figure 59 Save dialog box

- 7 Click **OK** in the **Information** dialog box that confirms your project was saved.

Functional areas of MPP

The main functional areas of the Mass Profiler Professional screen are illustrated in Figure 60.

The main Mass Profiler Professional window consists of four parts:

Menu Bar - access to actions that are used for managing your projects, experiments, pathways, and display pane views

Toolbar - access to buttons for commonly used tasks grouped by project, experiment, entity, statistical plot, and sidebar tasks

Display Pane - organized into functional areas that help you navigate through your project, experiments, analyses, and available operations

Status Bar - information related to the current view, cursor position, entity, and system memory

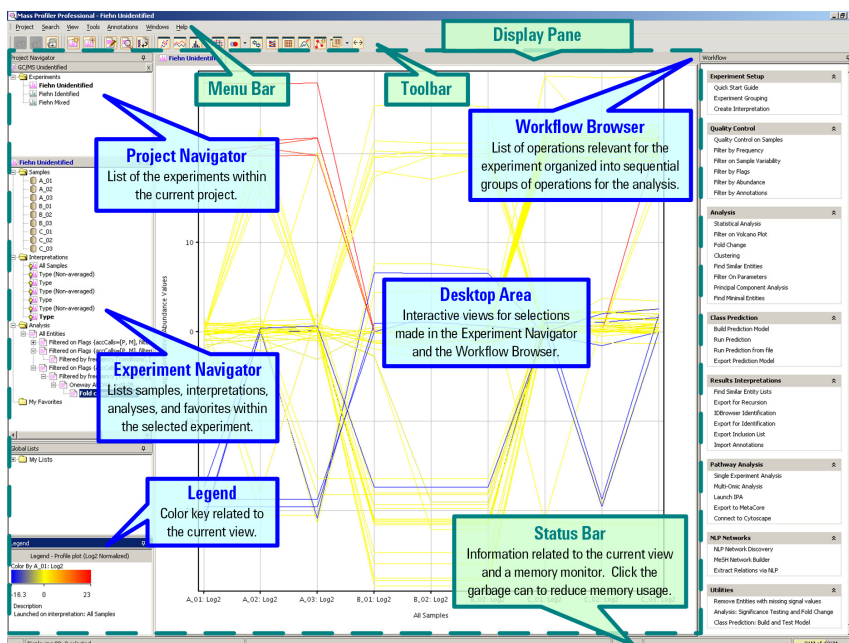
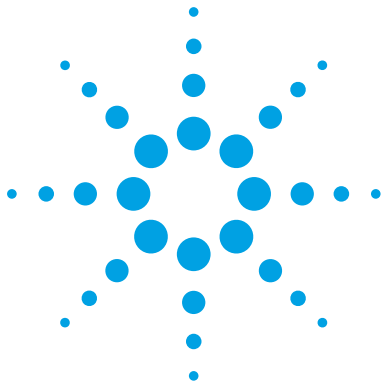


Figure 60 The main functional areas of Mass Profiler Professional



6 AMDIS Getting Started

Running AMDIS	116
To simultaneously view AMDIS and its Help window	116
To put AMDIS into an initialized state	116
Confirm Window and Results Window	118
Confirm Window Appearance	119
Results Window Appearance	126
Analysis Types	128
Tutorials Overview	130
Analysis Type - Simple Analysis Type	132
Analysis Type - Use Retention Time	141
Analysis Type - RI Calibration/Performance	148
Analysis Type - Use Retention Index Data	156
Analysis Type - Use Internal Standards for RI	163
Analysis Type - Use RI Calibr. Data + Internal Std.	170
Analysis Type - Performance Check for RI	176
Recap	183

This chapter is written in *tutorial style*. Follow the examples to learn about the capabilities of AMDIS. The whole process should take about 30 minutes.

For the examples in this chapter, AMDIS was installed in the root directory of the C:\ drive. The actual location of AMDIS is chosen at installation time. The default location is **C:\Program Files\NISTMS\AMDIS_32**.



Running AMDIS

To run the AMDIS program, select the **AMDIS** icon on your desktop or click **Start > All Programs > AMDIS32 > AMDIS_32**. If the NIST05 Mass Spectral Database and Search Algorithm is installed, select **Start > All Programs > NISTMS > AMDIS** instead.

To simultaneously view AMDIS and its Help window

- 1 Click **Help > Contents** to open the Help window.
- 2 Ensure that AMDIS and its Help window are not minimized.
- 3 Ensure that all other windows are minimized.
- 4 Right-click in any *unused* area in the Windows task bar to bring up its menu and click **Tile Windows Vertically**. The task bar is usually at the bottom of the screen and has the start button at the left.

To exit from this mode:

- 5 Right-click in any *unused* area in the Windows task bar and click **Undo Tile**.

To put AMDIS into an initialized state

To ensure that the examples in this chapter can be easily followed, AMDIS must be in an initialized state (state following installation).

- 1 Close the AMDIS program.
- 2 Delete the **onsite.ini** file located in the AMDIS32 folder.
- 3 Run AMDIS and select **Agilent Files** in the Set Default Instrument dialog (Figure 61).

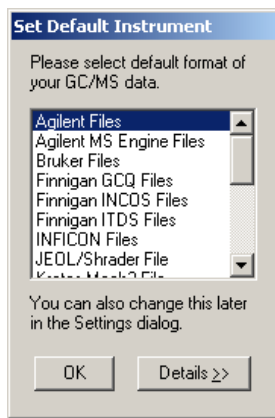


Figure 61 Set Default Instrument

At this point, all of the AMDIS parameters are set at default values. If result files already exist for the tutorial data files, the results you see can differ slightly from the examples. To start in a clean state, do the next steps.

- 4 When you are told to open a file, click **File > Open** and select any other file. Click **Open**. The Delete Result Files dialog appears (Figure 62).

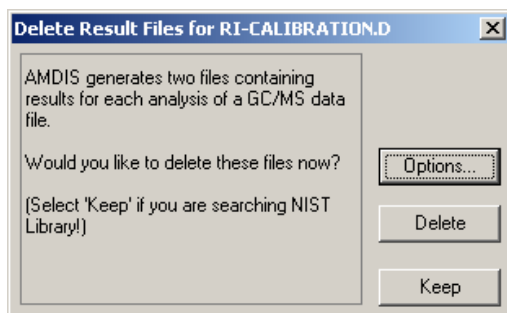


Figure 62 Delete Results Files dialog box

- 5 Click **Delete** and the newly selected file opens.
- 6 Click **File > Open** and select the file you originally opened according to the tutorial's instructions. Click **Open**.

You are now in the correct state for the tutorial.

Confirm Window and Results Window

AMDIS offers two user interfaces (windows) for you to work with.

The Confirm window is a full-functionality window, offering a menu bar and one or more child windows (for example, [Figure 65](#) on page 121). Each child window has a toolbar, several graphical areas, and a text area. The menu bar provides access to the full functionality of AMDIS, and the graphical areas permit interactive manipulations and detailed analysis of the results of AMDIS processing.

The Confirm window is fully described in AMDIS on-line Help and is the window of choice for most users.

The Results window offers a more limited set of functionality than the Confirm window and in particular does not display graphics, nor permit library manipulations. It can be useful when it is sufficient to simply review the results of AMDIS processing (for example, [Figure 67](#) on page 126).

The Results window is fully described in AMDIS online Help.

To move between windows

- In the Confirm window, click **File > Go to Results** to go to the Results window.
- In the Results window, click **Confirm** to go to the Confirm window.

Starting AMDIS from the GC/MS ChemStation or the Windows Start menu always starts in the Confirm window.

Confirm Window Appearance

The Confirm window has two basic appearances according to whether or not it is showing the results of an analysis. When results are being shown, the contents of the Confirm window are also regulated using the Mode command. See the AMDIS online Help for more information.

When a new data file is initially opened and does not have previously kept results, the Confirm window has two areas consisting of a Chromatogram Display and a Mass Spectral Display as shown in Figure 63.

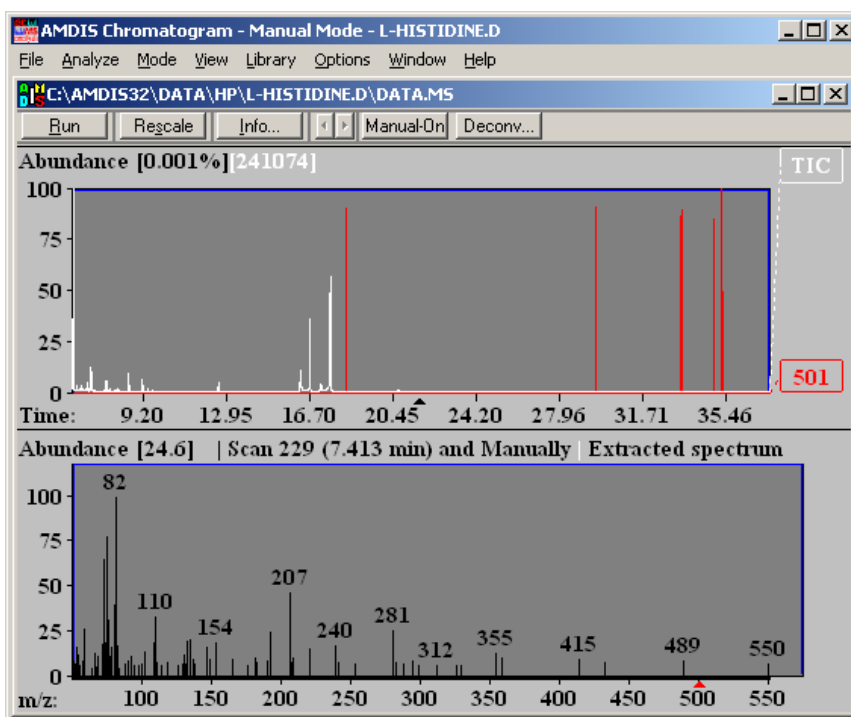


Figure 63 Confirm window showing unprocessed data

When a data file has been analyzed, or a file is opened with previously kept results, the Confirm window has four basic areas:

- Chromatogram Display
- Profile Display

- Information Lists
- Mass Spectral Display

If you wish to adjust the scaling of the Chromatogram Display or Mass Spectral Display areas, you can do this interactively.

After bringing AMDIS to a known state, the profile display shows the TIC and the model ion for the selected target or component. To show additional ions, select the **Select M/Z** item from the **Options** menu and increase the number of major m/z from 1 to 3.

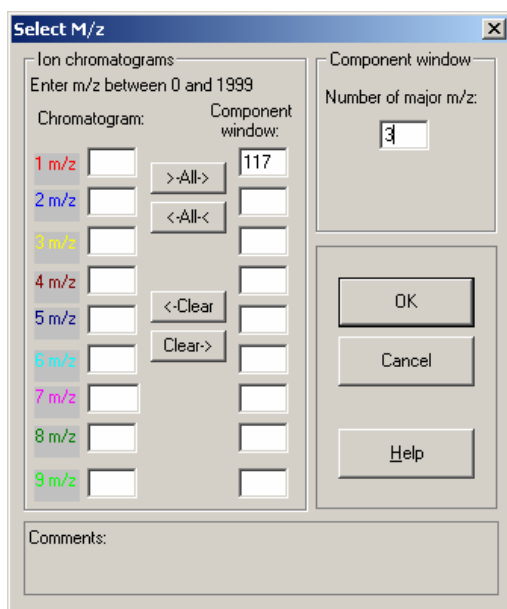


Figure 64 Showing additional ions

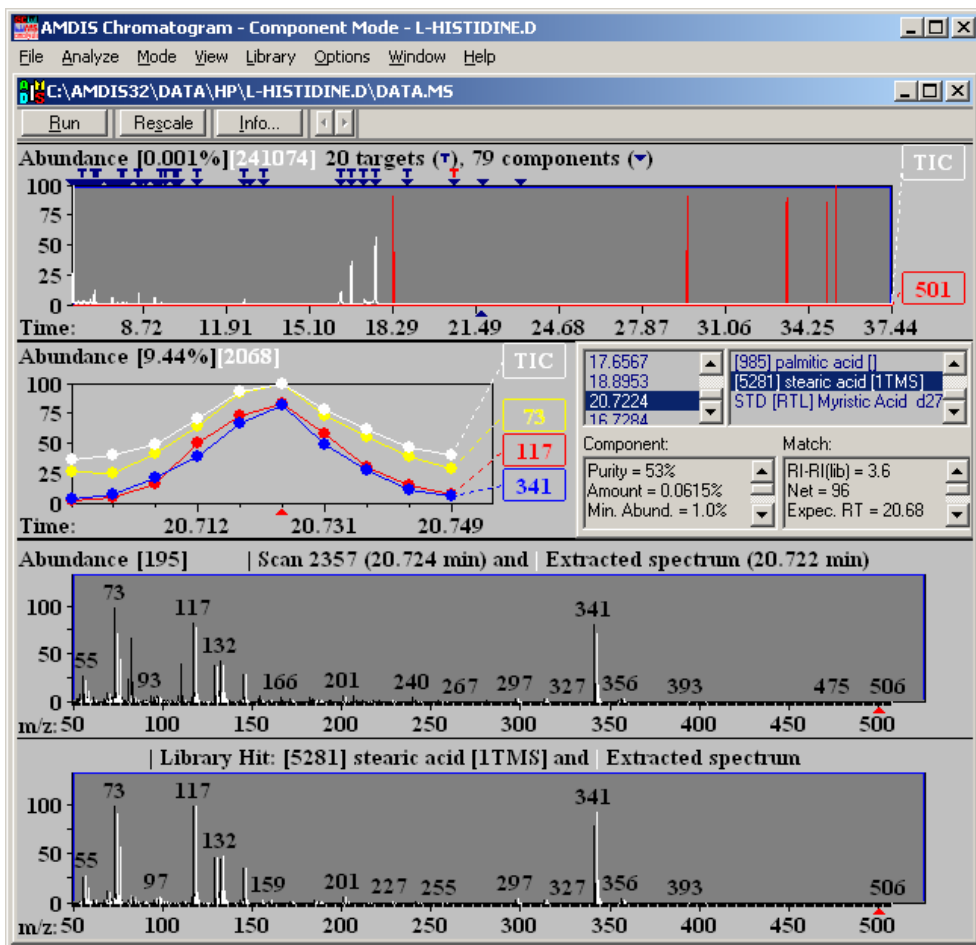


Figure 65 Confirm window showing the graphical results of an analysis for a component

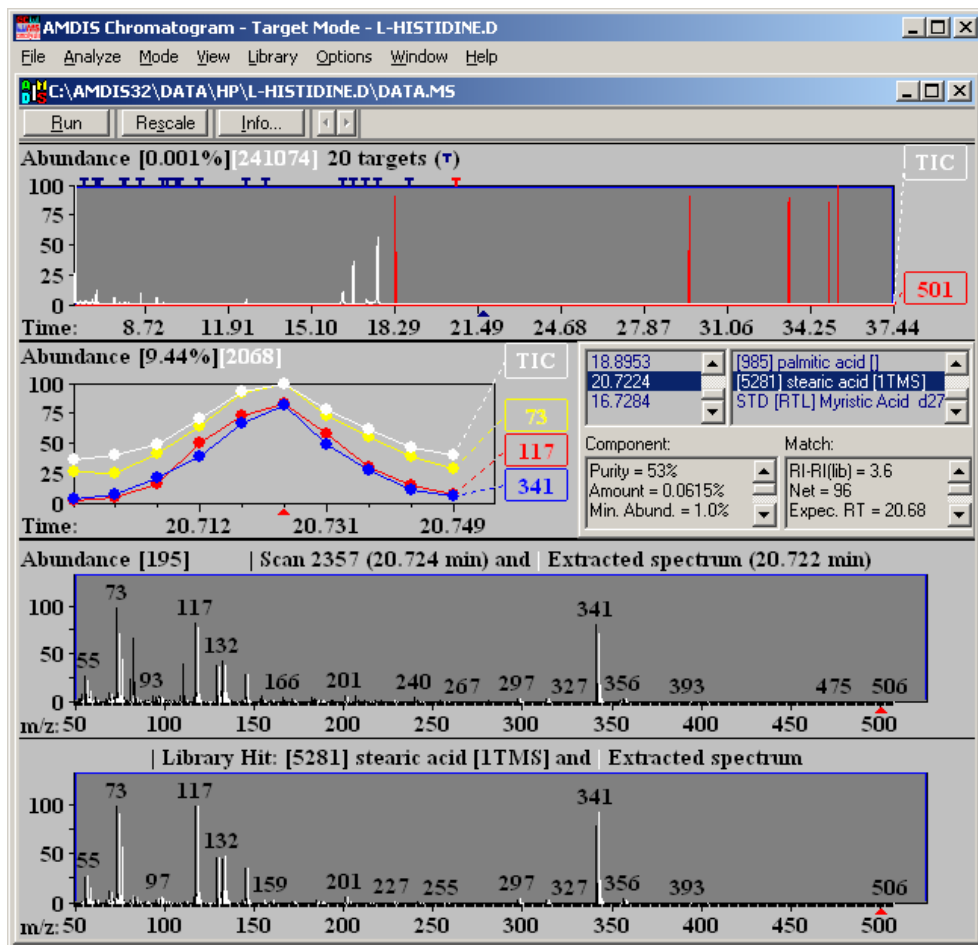


Figure 66 Confirm window showing the graphical results of an analysis for a target

Chromatogram Display The Chromatogram Display area generally shows the TIC chromatogram, but may also show one or more mass chromatograms.

The X-axis shows retention time in decimal minutes and the Y-axis is a relative abundance percentage scale, with the absolute abundance value representing 100% given at the top left of the display (for example, 241074 in Figure 65 and Figure 66).

To the right of the abundance value for analyzed data (and as appropriate) are the numbers of found targets (**T** symbol) and components (**▼** symbol), where the count of components includes the targets.

Along the top of the chromatogram graphical display area are rows of **T** and **▼** symbols, which are dark blue and positioned at the retention times at which these compounds were determined by the deconvolution process.

The currently selected target or component is signified by its symbol being red. The rest of the Confirm window's data relate to this compound. You can click any of the symbols to make it the current selection.

For full details of the Chromatogram Display area, see AMDIS online Help.

Profile Display

The Profile Display area generally shows the TIC chromatogram and major characteristic ion(s) of the deconvoluted compound over the retention time window in which it was determined. The y-axis is a relative abundance percentage scale with the absolute abundance value representing 100% given at the top left of the display (for example, 2068 in [Figure 66](#)).

For full details of the Profile Display area see AMDIS online Help.

Information Lists

The Information Lists area changes somewhat according to whether the currently selected compound is being displayed as a component (**▼**) or as a target (**T**).

For a component there are two side-by-side sections. For a target there are four sections arranged more or less as quadrants.

For full details of the Information Lists area, see AMDIS online Help.

Target

The upper-left quadrant is a list of retention times for all targets.

The upper-right quadrant is a list of *hits* (identifications) from a target library.

The lower two quadrants provide various details about the target as determined during the deconvolution process.

Component

The left-hand section is a list of retention times for all components.

The right-hand section provides various details about the component as determined during the deconvolution process.

You can click an RT value and all areas of the whole display changes to reflect the chosen component.

You can click a value in either upper section and all areas of the display change to reflect the chosen retention time or hit.

Mass Spectral Display

The Mass Spectral Display area changes according to whether the currently selected compound is being displayed as a target (T) or as a component (▼).

The X-axis shows m/z values. Mass spectral peaks are displayed rounded to integer m/z .

For a component there is one Mass Spectral Display showing the raw (original) spectrum overlaid with the extracted spectrum resulting from the deconvolution process. Ions from the raw spectrum are shown in black and those from the extracted spectrum in white. Note that for clarity they are slightly offset from each other along the x-axis.

The Y-axis is a relative abundance percentage scale with the absolute abundance value representing 100% given at the top left of the display, (for example, 195 in [Figure 65](#)).

For a target there are two Mass Spectral Displays. The upper display shows the same information as just described for a component. The lower display is similar, but shows a library hit (black ions) overlaid with the extracted spectrum (white ions). Again they are slightly offset for visual clarity. Since library spectra are only stored with relative ion abundances, no absolute abundance value representing 100% is given at the top left of this display.

For full details of the Mass Spectral Display area see AMDIS online Help.

Changing the display range

Often the Chromatogram Display is very crowded with target and component symbols overlapping. To see the detail of the chromatogram(s) and to easily choose (by clicking it) a specific target or component symbol, it is necessary to first zoom (magnify) the display. Similarly, you can zoom the Mass Spectral Display to see greater detail.

By using the *click-drag-release* technique, the display can be zoomed individually in either the X-direction or the Y-direction or in both directions simultaneously.

Click-drag-release is as follows. First click and hold the left mouse button at one end of the desired linear range, or one corner of the desired rectangular area. Now drag the mouse pointer to the opposite end or corner of the desired selection. Finally, release the mouse button.

To zoom in the X-direction, do a click-drag-release operation just below the X-axis. When you release the mouse button, the selected area is redrawn to fill the whole width of the display area.

To zoom in the Y-direction, do a click-drag-release operation just to the left of the Y-axis. When you release the mouse button, the selected area is redrawn to fill the whole height of the display area.

To zoom a rectangular area of the display, do a click-drag-release operation within the display area to *rubber band* (select) the area to be magnified. When you release the mouse button, the selected area is redrawn to fill the whole display area.

To cancel the effects of a zoom, right-click anywhere in the relevant display and click **Unzoom** or, if available, **Unzoom All** on the offered menu.

Results Window Appearance

The Confirm window shows results for targets and components and the Results window only shows information for targets.

The Results window has three basic areas and is fully described in AMDIS online Help.

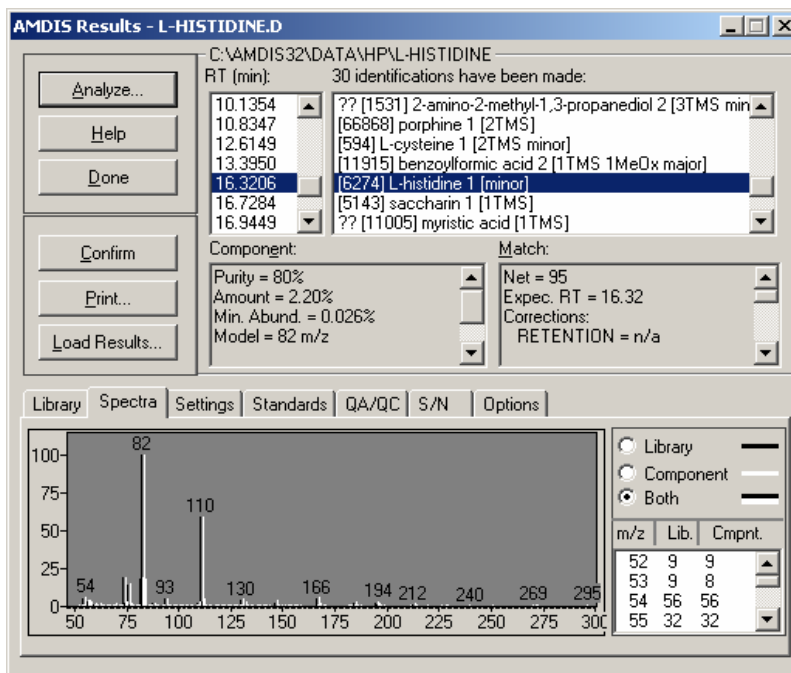


Figure 67 Results window shows only the non graphical results of an analysis

The three areas of the Results window are:

- Control Panel
- Information Lists
- Information Tabs

The Control Panel is a series of command buttons that control the operations available in the Results window.

The Information Lists area has four sections arranged more or less as quadrants. The upper-left quadrant is a list of retention times for all targets. The upper-right quadrant is a list of *hits* (identifications) from a target library. The lower two quadrants provide various details about the target as determined during the deconvolution process. When you click a value in either upper quadrant, the lower two quadrants and certain of the tabs in the Information Tabs area changes to reflect the chosen retention time or hit.

The Information Tabs area has a number of selectable tabs. The Library and Spectra tabs change to reflect the selection in the Information Lists. Other tabs show data that are independent of the particular targets and so do not change as different selections are made.

Analysis Types

AMDIS supports several different *analysis types* for determining whether or not deconvoluted spectra from samples match any of the target compounds (targets) stored in a referenced *target library*. The analysis type you choose depends to a large extent on the type of data that you are working with. Of particular significance is whether or not you are using retention indices and, if you are, whether you are using internal standard compounds within a sample for their computation, or whether you are using an external calibration file holding retention index values.

Table 4 Analysis type

Analysis type	Description
Simple	Only mass spectra are used in determining target matches; so no use is made of retention time or retention indices.
Use Retention Time	Following a target match determined by mass spectral comparison, the compound's retention time is compared against the library value to confirm the match.
RI Calibration/Performance	A clean mixture of known composition is acquired and the data used to create an RI Calibration Data file for use by other appropriate analysis types.
Use Retention Index Data	Following a target match determined by mass spectral comparison, the compound's RI is computed using an external RI Calibration Data file so that the quality of the match can be further assessed.
Use Internal Standards for RI	Known compounds (the <i>internal standards</i>) are co-injected with the sample. Mass spectral matching determines which targets and internal standards are present. The retention times of the found internal standards are used to compute retention indices for the targets so that the quality of their match can be further assessed.

Table 4 Analysis type (continued)

Analysis type	Description
Use RI Calibr. Data + Internal Std.	Internal standards are co-injected with the sample, but target matching is determined as for the Use Retention Index Data type and does not take any account of the internal standard retention times. The internal standards are simply used for reporting purposes and permit you to check that the acquisition system is functioning correctly.
Performance Check for RI	A clean mixture of known composition is acquired and the data for <i>performance check</i> compounds are recorded and reported so that you can determine how well the acquisition system is performing.

Tutorials Overview

AMDIS can run seven different types of analysis and includes a tutorial to demonstrate of each of these analyses. Each tutorial describes any prerequisites, the operational flow, and some post-analysis options relevant to the particular analysis type. Each tutorial uses demo data files and mainly uses default settings of AMDIS parameters. Throughout the tutorials, related topic are occasionally briefly introduced.

Although each tutorial section is self-contained, it is recommended that you initially go through them in order because they do to a certain extent lead from one to the other. If later you want to simply refresh your memory on one section, go to it directly.

A summary of the seven different types of analysis covered in the tutorials is listed in [Table 5](#), along with the demo data files required to run the analysis yourself.

Table 5 AMDIS Analysis Type

AMDIS analysis type	Fiehn use description	Required AMDIS files
Simple	This analysis mode is used to process data files that have not been acquired using the Fiehn.m RTL method or data files acquired without any RI calibration. Results from this analysis mode are harder to interpret because there are more false positives that need to be evaluated. This mode can be used to locate compounds in the data file that are spectrally and structurally similar to compounds in the library.	Fiehn.msl
Use Retention Time	This analysis mode is similar to the Simple analysis mode, but it penalizes potential target hits if the target retention time differs from the retention time in the library.	Fiehn.msl
RI Calibration/Performance	This analysis mode should be used to periodically update the RTs of the RI standards. The Agilent Fiehn Metabolomics Library is calibrated with fatty acid methyl esters (FAMES) between C8 and C28.	Fiehn.msl Fiehn.cls New.cal (created)
Use Retention Index Data	This analysis mode is the preferred mode for analyzing unknown samples that are acquired with Fiehn.m. It is important to use Fiehn.m as an RTL method where d27 myristic acid elutes at 16.72 minutes.	Fiehn.msl Fiehn.cal

Table 5 AMDIS Analysis Type (continued)

AMDIS analysis type	Fiehn use description	Required AMDIS files
Use Internal Standards for RI	This analysis mode is used when one or more of the compounds in the Fiehn.isl files are part of the acquired data file. This mode can produce superior RI results because the RI standards are acquired at the same time as the target compounds. The disadvantage of this mode is that some target compounds coelute with the internal standards.	Fiehn.msl Fiehn.isl
Use RI Calibr. Data + Internal Std.	This analysis mode uses the CAL file for the RI calculations and uses the internal standards as a functional check of the acquisition system. This mode expects the RI standards to be injected with the sample.	Fiehn.msl Fiehn.isl Fiehn.cal
Performance Check for RI	This analysis mode is used to evaluate the performance of the acquisition system. This mode expects the RI standards to be injected.	Fiehn.csl Fiehn.cal

Analysis Type - Simple Analysis Type

This example illustrates the most basic operation of AMDIS, which is to locate and extract significant spectra from a GC/MS data file and then compare them against an AMDIS Target Compounds Library.

It also describes the interactive ability to search a component's spectrum against the main NIST mass spectral database and any other NIST-format libraries (see “Searching the NIST database” on page 140).

A Target Compounds Library has an .msl file extension. A Target Compounds Library must exist for a simple analysis to run.

Before following these instructions, ensure that AMDIS is in a known state by following the instructions in “To put AMDIS into an initialized state” on page 116. An *empty* Confirm window should now be displayed.

- 1 Click **File > Open** and in the Select Data File dialog, navigate to the **AMDIS32\DATA\HP** folder, select the file **L-HISTIDINE.d** and click **Open**. The window shows the file's TIC in the upper half (Chromatogram Display). The spectrum at the center of the displayed TIC range is shown in the lower half (Mass Spectral Display). See [Figure 63](#) “Confirm window showing unprocessed data” on page 119.
- 2 Click **Analyze > Analyze GC/MS Data** to invoke the Analyze GC/MS Data dialog.

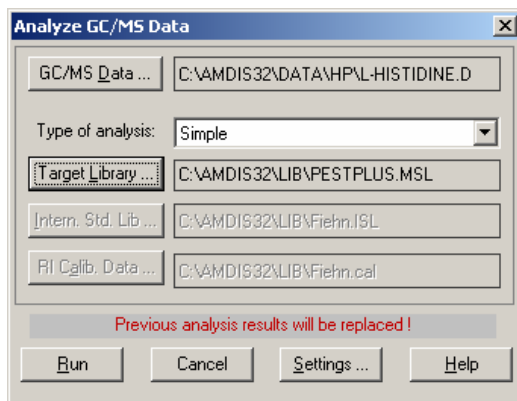


Figure 68 Analyze GC/MS Data dialog (Simple)

The text box to the right of the **GC/MS Data** button shows the current data file selection and Type of analysis shows Simple.

- 3 Click **Target Library...** to invoke the Analysis Settings dialog, showing the **Libr.** tab. Click **Select New...** to specify Fiehn.msl, the searchable target compounds library.

Fiehn.msl and other supplied libraries are normally installed in AMDIS32\LIB\.

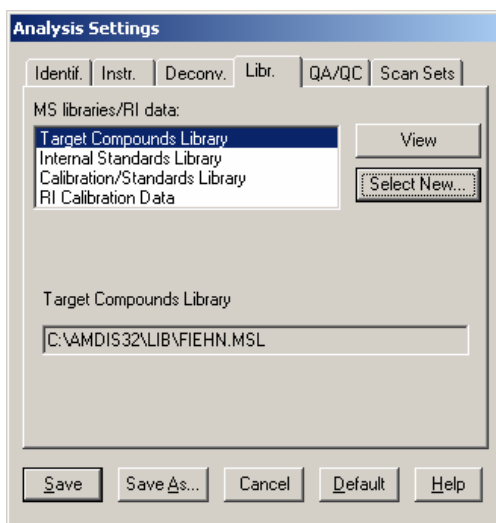


Figure 69 Analysis Settings dialog (Libraries tab)

- 4 Click **Open** to return to the Analysis Settings dialog with the full path of the Fiehn.msl file shown after Target Compounds Library in the lower half of the dialog.
- 5 Click **Save** to return to the Analyze GC/MS Data dialog.

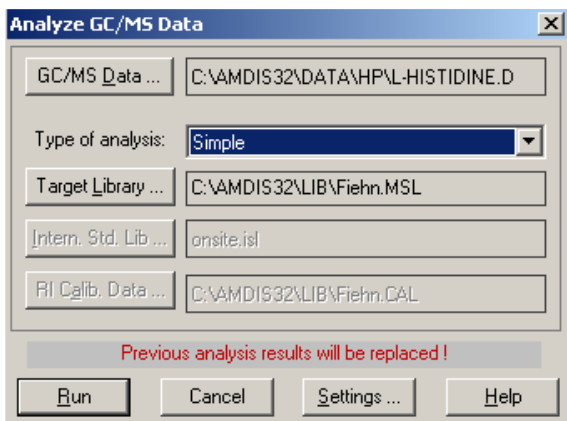


Figure 70 Analyze GC/MS Data dialog readied (Simple)

The **Intern. Std. Lib ...** and **RI Calib. Data ...** buttons in the dialog are disabled. A Simple analysis does not involve the use of internal standards, nor does it take any account of retention index information. You are now ready to perform the Simple analysis. Click **Run**.

- 6 Select **Run**. The progress bars indicate the processing operations being performed. Within a few seconds the Confirm window is redisplayed, showing the results of the deconvolution process.

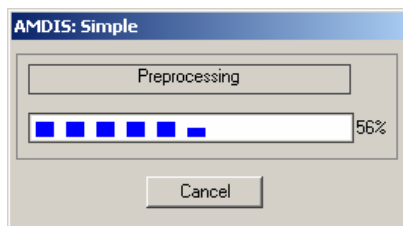


Figure 71



Figure 72 The results of a Simple analysis

The file's TIC is presented in the Chromatogram Display area and some, but not all, of the strong peaks have a T mark above them. These are the components that have been identified as target compounds in the Agilent Fiehn Target Compounds Library (27 of its 122 entries have been identified).

There are also a large number of other components, as shown by the ▼ symbols without a T above them. These are compounds that have been extracted from the data file, but have no match in the library. Many of these unidentified components are in the first few minutes of the data file.

- 7 Expand the display to see components in more detail. As an example, position the mouse pointer at the 25% position of the TIC graphical display area at about 0.00 min. While holding down the left mouse button, drag the mouse to the bottom of the TIC window at about 8.45 min; a rubber band box is displayed as you drag the mouse. When you let go of the left mouse button the display expands to make it easier to determine which components remain unidentified.

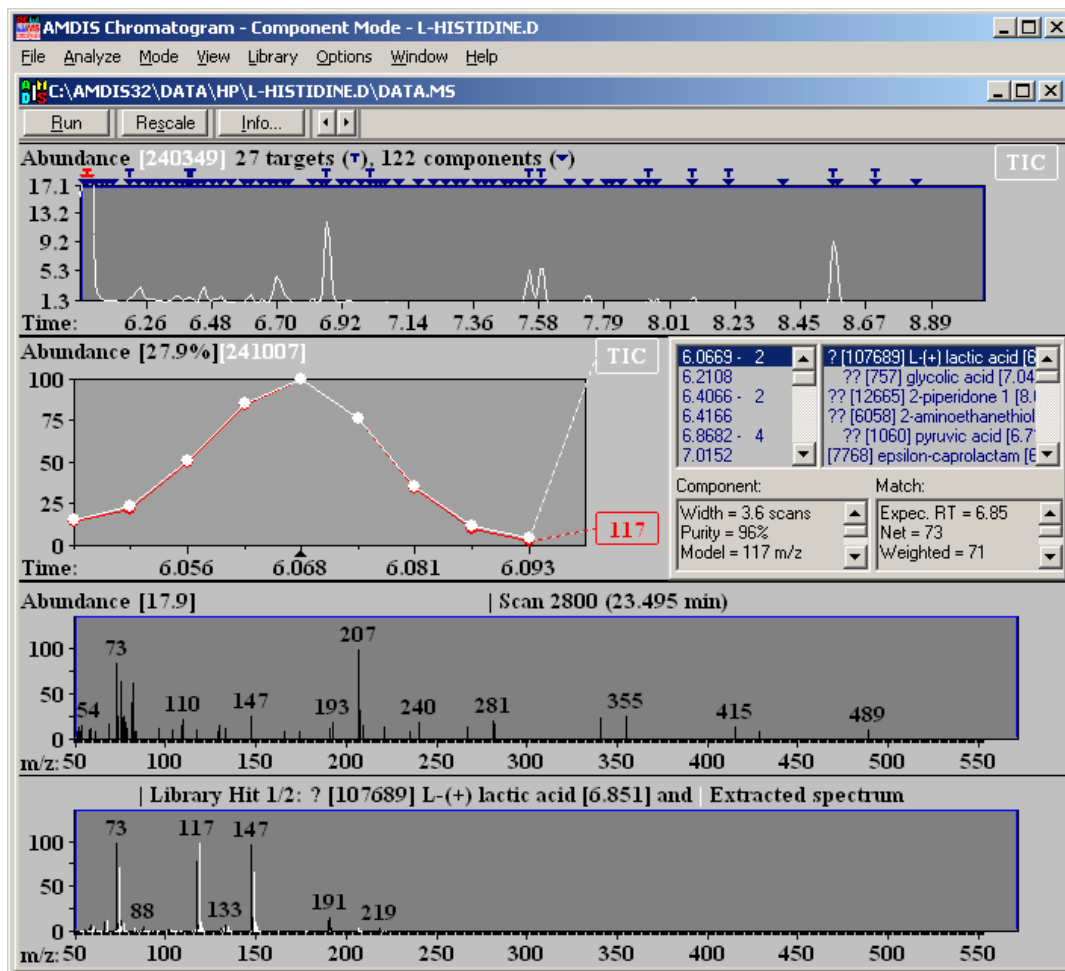


Figure 73 Expanded chromatogram

If you make a mistake with the rubber-banding, you can always right-click anywhere in the Chromatogram Display area and then click **Unzoom** to correct the most recent rubber-banding, or if appropriate, **Unzoom All** to immediately restore the display to its full range.

Individual components can be selected by either clicking the blue ▼ above them, or by clicking their RT value in the Information Lists area. The currently selected component is identified by its ▼ changing from blue to red.

The same selection mechanism and color changes apply when viewing targets (T). Also, as a component or target is selected, all of the other information in the Confirm window changes to reflect the new selection.

The deconvolution process does have a few parameters that can be changed by the user and so let's see where these are.

- 8 On the Analyze menu click **Settings**, and in the Analysis Settings dialog, click **Deconv.** to select its tab.

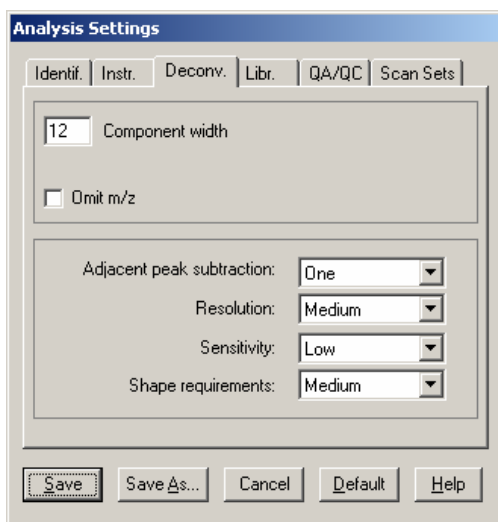


Figure 74 Deconvolution tab of Analysis Settings dialog

The Analysis Settings dialog is described in detail in the AMDIS online Help file, but for now simply change **Sensitivity** to **Low** and click **Save**. A message appears noting that the parameters have changed and asking if you want to reanalyze. Click **Yes**.

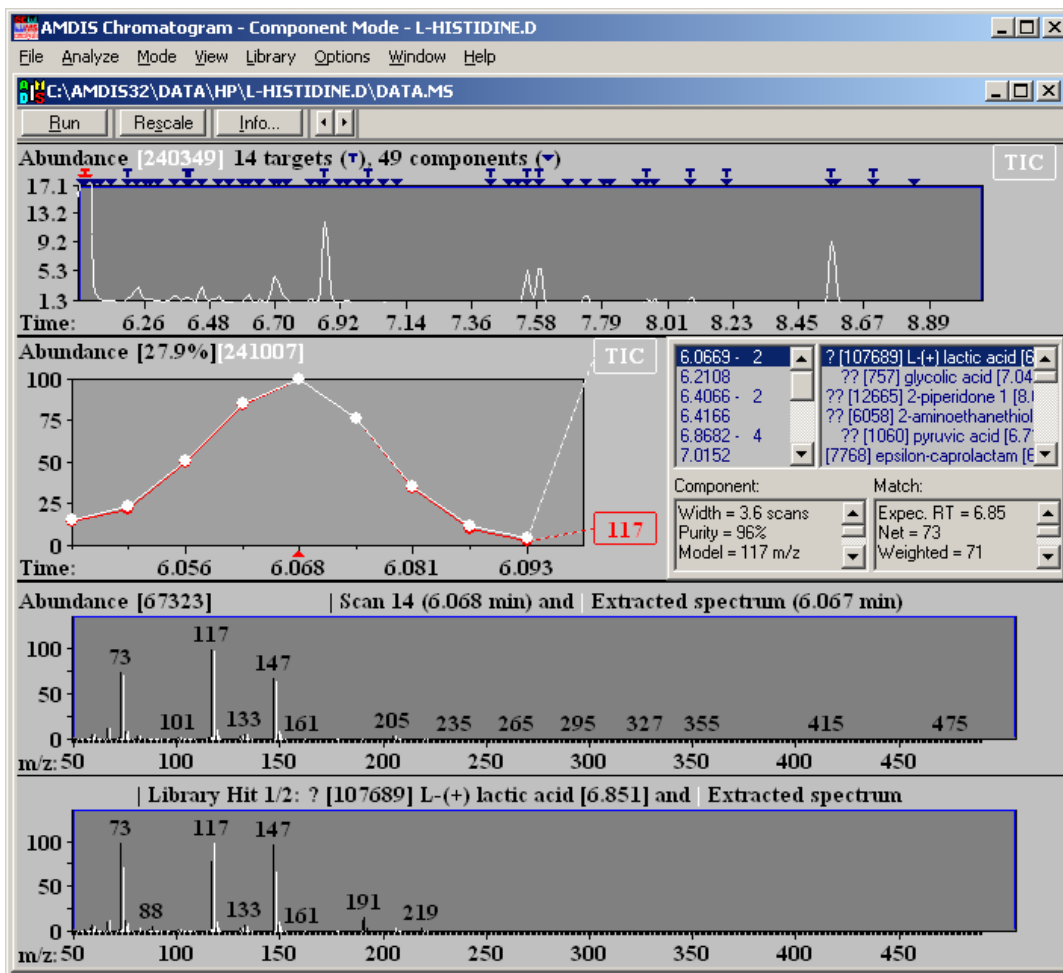


Figure 75 After adjusting the deconvolution parameters

The resulting display differs in two respects to the previous display.

First, the header lists 14 targets and 49 components (compared previously to 27 and 122, respectively). This is because when a deconvolution analysis is run, it applies to the current display range only. In the first instance the analysis was run when the full chromatogram was being displayed.

The exact numbers reported for this range depend on the actual retention time range you previously selected by rubber-banding.

Many parameters on the various tabs in the Analysis Settings dialog control deconvolution, detection, and reporting in AMDIS. Fortunately, in most cases the default values either work *as is*, or only require minor changes.

Searching the NIST database

From AMDIS you can search the full NIST mass spectral database with a component. This capability requires the G1033A NIST Library.

- 1 Click the ▼ above the large peak at 17.6567 min (or select it from the retention time list).
- 2 On the Analyze menu, click **Go to NIST MS Program**.

The NIST MS Search program is invoked and a dialog shows the progress of the search of the selected component's spectrum (this can take up to a minute or two when first invoked). The component is identified as Oleic Acid, a component of the analyzed oil.

As noted earlier, the Agilent Fiehn Library reports compounds as metabolites, "L-Histidine" whereas the NIST Search program correctly identifies the metabolite that has been analyzed as a trimethylsilyl (TMS) derivative. The exact library result here and elsewhere depend on the installed libraries and the search parameters currently set in the NIST MS Search program. The result here occurs when using default search parameters and searching the Main and Replib libraries.

- 3 From the MS Search program use the ALT+TAB key, or the Windows task bar, to get back to AMDIS.

Analysis Type - Use Retention Time

If you wish to use retention times as an additional aid in identifying targets, AMDIS can use those stored in a Target Compounds Library.

It is a prerequisite of this type that a Target Compounds Library exists. Before following the instructions below, ensure that AMDIS is in a known state by following the instructions in “To put AMDIS into an initialized state” on page 116. There should now be an *empty* Confirm window displayed.

- 1 Click **File > Open** and in the Select Data File dialog, navigate to the **AMDIS32\DATA\HP** folder, select the file **L-HISTIDINE.d** and click **Open**. As normal, the window shows the file’s TIC in the upper half (Chromatogram Display) and the spectrum at the center of the displayed TIC range is shown in the lower half (Mass Spectral Display); see Figure 63 “Confirm window showing unprocessed data” on page 119.
- 2 Click **Analyze > Analyze GC/MS Data** to invoke the Analyze GC/MS Data dialog and set Type of analysis to **Use Retention Time**.

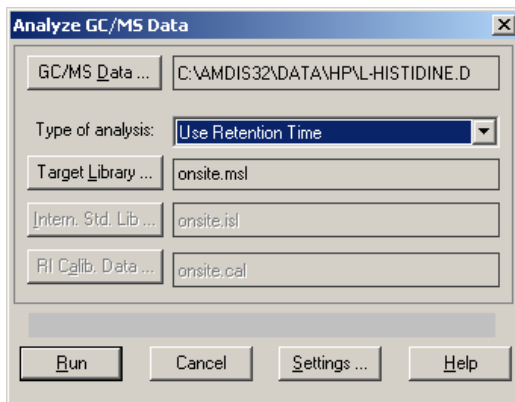


Figure 76 Analyze GC/MS Data dialog (Use Retention Time)

- 3 Click **Target Library** to invoke the Analysis Settings dialog and on the **Libr.** tab (with Target Compounds Library highlighted) click Select **New**. In the Target Compounds Library dialog, navigate to the **AMDIS32\LIB** folder, select the file **Fiehn.msl** and click **Open** to return to the Analysis Settings dialog.

4 Click **Save** to return to the Analyze GC/MS Data dialog.

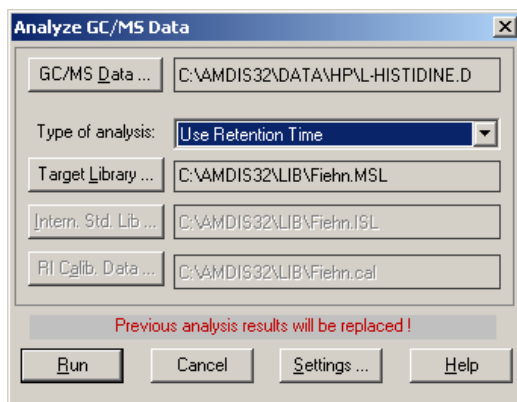


Figure 77 Analyze GC/MS Data dialog readied (Use Retention Time)

The **Intern. Std. Lib...** and **RI Calib. Data...** buttons in the dialog are disabled. A Use Retention Time analysis does not involve the use of internal standards, nor does it take any account of retention index information, but there are a few analysis parameters that are specific to the Use Retention Time analysis.

5 Click **Settings** and in the Analysis Settings dialog click **Identif.**

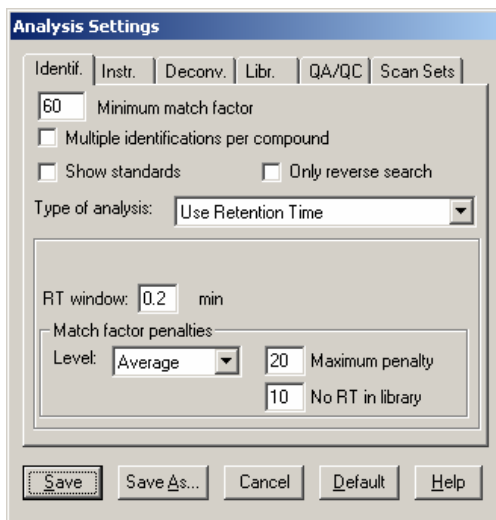


Figure 78 Analysis Settings dialog (Use Retention Time)

RT and its associated Match factor penalties are used by this analysis type only. These and the other parameters on this tab are fully described in online Help.

Select the **Deconv.** tab and specify **Very High Sensitivity**.

You are now ready to perform the analysis

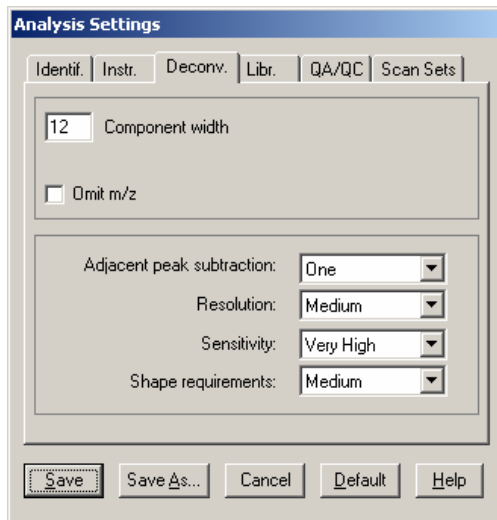


Figure 79 Analysis Settings dialog (Deconv. sensitivity)

- 6** Click **Save** to exit the Analysis Settings dialog and return to the Analyze GC/MS Data dialog.
- 7** Click **Run** to start the analysis and within a few seconds the results are displayed.

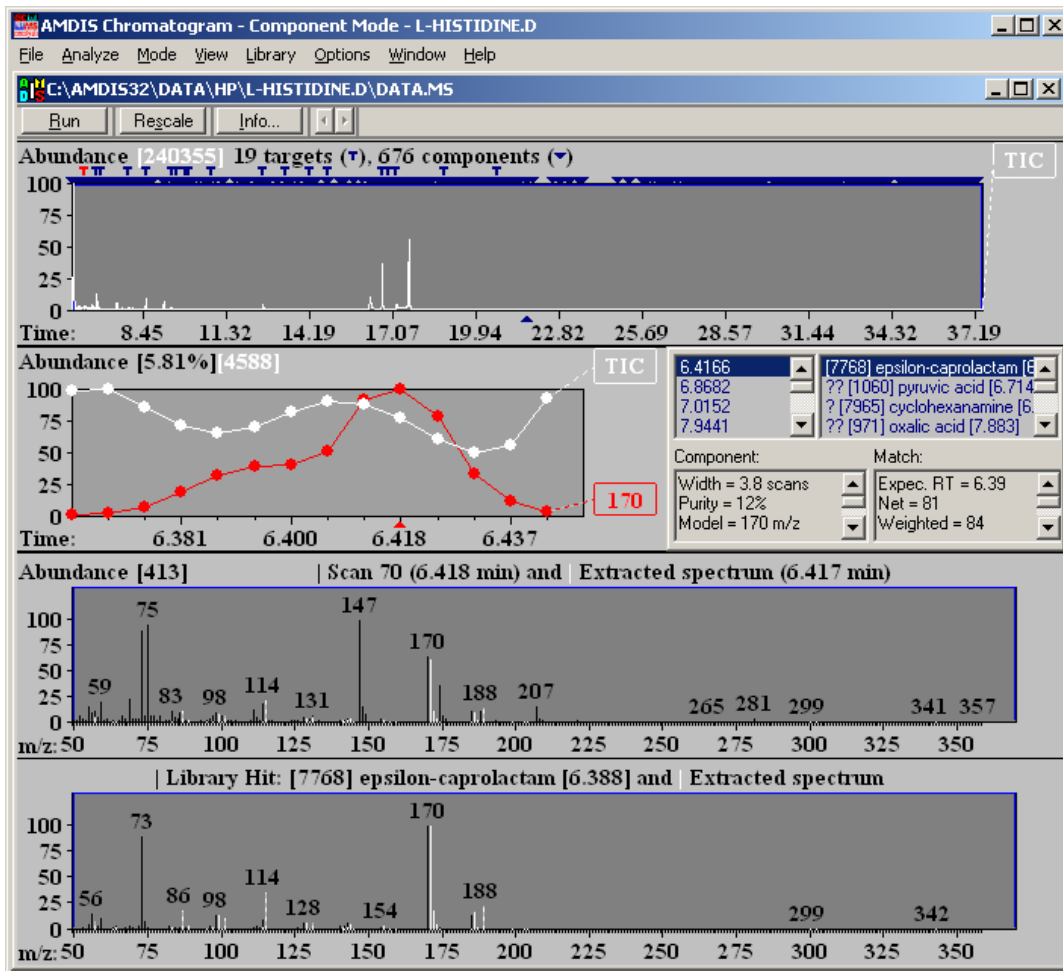


Figure 80 Use Retention Time analysis results

- To more easily examine the AMDIS results, on the **Mode** menu click **Target Only**. Now, only targets are displayed. You can switch back to viewing both by clicking **Mode > Component**.

While viewing targets, right-click anywhere in the upper right-hand quadrant of the Information Lists area and view a 2-entry menu.



Figure 81 Controlling the displayed hits

Click **Best Hits Only** to display only the single best hit for each target.

Now click the target symbols above the TIC chromatogram display, or click retention times in the Information Lists area, to see just the *best hit* name of each target displayed; with the currently selected one highlighted. Further, the displayed Component and Match information relates to the selected target.

- 9 To examine the details of the deconvolution and library matching, select the **UNDOCK** menu item of the **View** menu. Once the Information list is detached from the AMDIS window, the size of the window can be enlarged to show the detailed match parameters. In particular, select the component at retention time 9.4429 - 2-amino-1-phenylethanol.

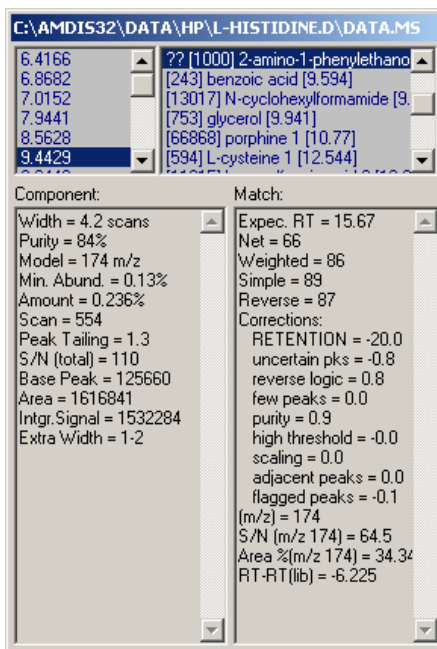


Figure 82 RETENTION correction to Net value

For most targets the retention time correction to the Net match factor is 0.0 (RETENTION is 0.0). However, 2-amino-1-phenylethanol has a value of -20.

A value of 0.0 means that the target has been found within the RT window you specified on the **Identif.** tab of the Analysis Settings dialog. When the library's expected retention time for 2-amino-1-phenylethanol is compared to the found retention time for it, the difference is greater than the **Identif.** tab's RT window and hence a penalty is computed (-20.0 in this case).

The RETENTION penalty is applied to the computed Net (match) value for the target and reduce the latter's magnitude; hence making the component a less favorable hit for the target. In this way, as a target is found further away from its expected time, a larger penalty is applied (up to the Maximum penalty value on the **Identif.** tab of the Analysis Settings dialog).

Analysis Type - RI Calibration/Performance

If you wish to use retention indices as an additional aid in identifying targets, or for checking the performance of the acquisition system, AMDIS can use an RI Calibration Data file that holds a correspondence between retention time and retention index. The compounds used when creating an RI Calibration Data file are referred to as the *retention index standards*. If compounds are used for performance checking purposes, they are referred to as *calibration check compounds*.

An RI Calibration Data file has a .cal file extension. An RI Calibration Data file is used in three of the analysis types that are examined in the coming sections.

This tutorial section illustrates how an RI Calibration Data file is created by analyzing a clean mixture of known composition (often a *C-series* or *homologous series* sample).

It is a prerequisite of this analysis type that a Calibration & Standards Library exists. A Calibration & Standards Library has a .csl file extension.

Before following the instructions below, ensure that AMDIS is in a known state by following the instructions in [Figure “To put AMDIS into an initialized state”](#) on page 116. There should now be an *empty* Confirm window displayed.

- 1 Click **File > Open** and in the Select Data File dialog, navigate to the **AMDIS32\DATA\HP** folder; select the file **RI-CALIBRATION.d** and click **Open**.

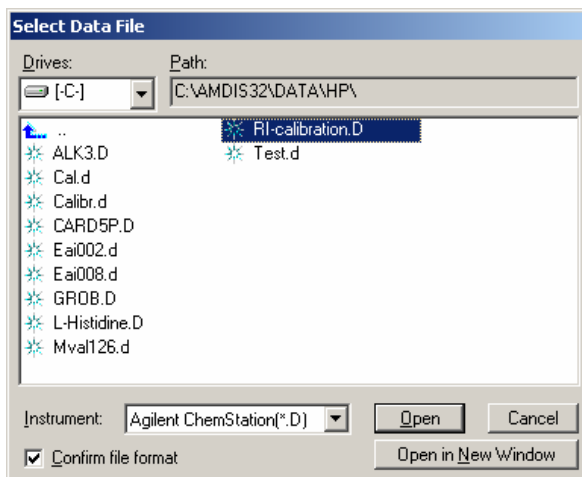


Figure 83 AMDIS32\DATA\HP folder (RI-CALIBRATION.d)

- 2 Click **Analyze > Analyze GC/MS Data** to invoke the Analyze GC/MS Data dialog and set **Type of analysis** to **RI Calibration/Performance**.

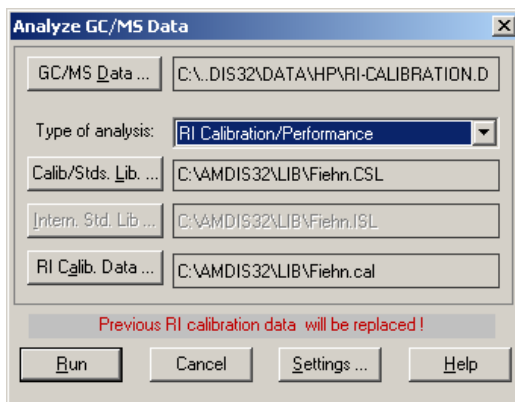


Figure 84 Analyze GC/MS Data dialog on entry (RI Calibration/Performance)

The button that was labeled Target Library for a Simple and Use Retention Time analysis is now labeled **Calib/Std. Lib..**

- 3 Click **Calib\Stds. Lib.** to invoke the Analysis Settings dialog and on the **Libr.** tab (with Calibration/Standards Library highlighted) click **Select New**. In the Calibration/Standards Library dialog, navigate to the **AMDIS32\LIB** folder, select the file **Fiehn.csl** and click **Open** to return to the Analysis Settings dialog with the full path of the Fiehn.csl file shown as the Calibration/Standards Library in the lower half of the dialog. Click **Save** to return to the Analyze GC/MS Data dialog.
- 4 Click **RI Calib. Data** to invoke the Analysis Settings dialog and on the **Libr.** tab (with RI Calibration Data highlighted) click **Select New**. Type **New.cal** as the File name. Click **Open** to return to the Analysis Settings dialog with the full path of the **New.cal** file shown as the RI Calibration Data in the lower half of the dialog. Click **Save** to return to the Analyze GC/MS Data dialog.

A new file, **New.cal**, is created to avoid overwriting the **Fiehn.cal** file.

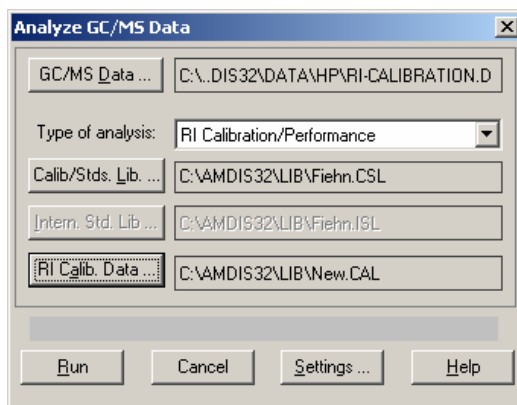


Figure 85 Analyze GC/MS Data dialog readied (RI Calibration/Performance)

- 5 The **Intern. Std. Lib** in the dialog is disabled. An RI Calibration/Performance analysis does not involve the use of internal standards.

You are now ready to perform the analysis. Click **Run**.

A confirmation dialog appears for you to approve the file selections.

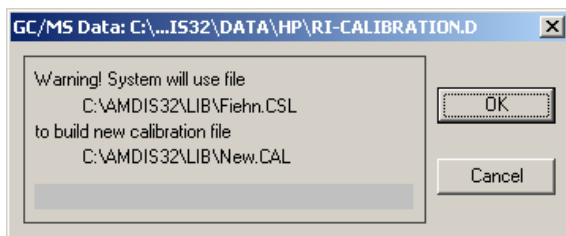


Figure 86 Confirmation dialog (RI Calibration/Performance analysis)

- 6 Click **OK** to proceed. Within a few seconds the Confirm window is redisplayed showing the results of the analysis.

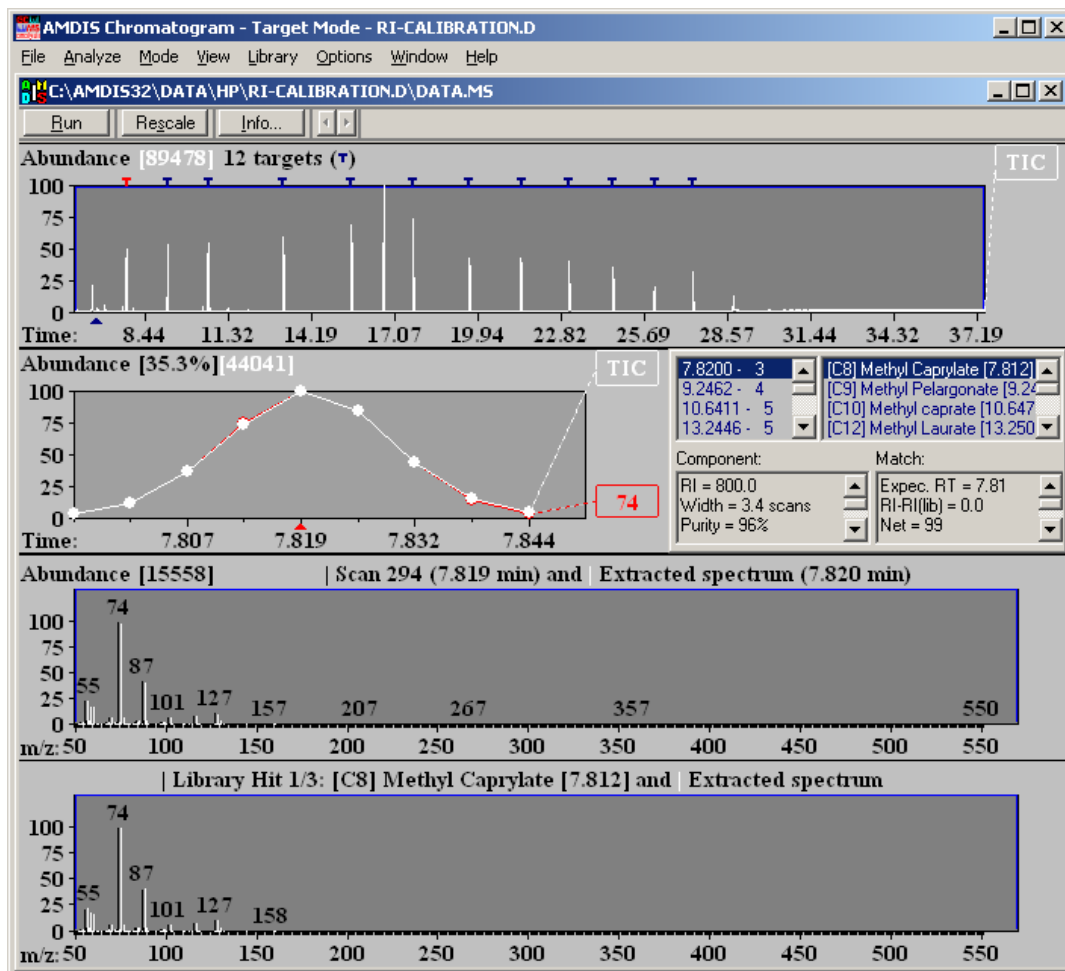
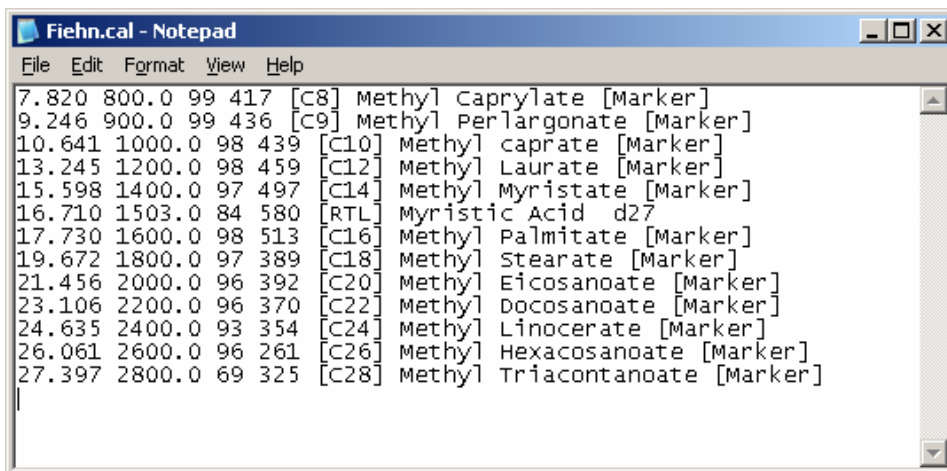


Figure 87 RI Calibration/Performance analysis results

All of the TIC peaks have been detected as Targets.

- The RI Calibration Data file (.cal) that you created is not directly viewable within AMDIS, but is a simple text file and can be easily viewed, for example, using Notepad.



```

File Edit Format View Help
7.820 800.0 99 417 [C8] Methyl Caprylate [Marker]
9.246 900.0 99 436 [C9] Methyl Perlargonate [Marker]
10.641 1000.0 98 439 [C10] Methyl caprate [Marker]
13.245 1200.0 98 459 [C12] Methyl Laurate [Marker]
15.598 1400.0 97 497 [C14] Methyl Myristate [Marker]
16.710 1503.0 84 580 [RTL] Myristic Acid d27
17.730 1600.0 98 513 [C16] Methyl Palmitate [Marker]
19.672 1800.0 97 389 [C18] Methyl Stearate [Marker]
21.456 2000.0 96 392 [C20] Methyl Eicosanoate [Marker]
23.106 2200.0 96 370 [C22] Methyl Docosanoate [Marker]
24.635 2400.0 93 354 [C24] Methyl Linocerate [Marker]
26.061 2600.0 96 261 [C26] Methyl Hexacosanoate [Marker]
27.397 2800.0 69 325 [C28] Methyl Triacontanoate [Marker]

```

Figure 88 RI Calibration Data file listing

The pertinent information that it contains per entry is a retention time and the corresponding retention index (for example, 10.641 and 1000.0, respectively, in [Figure 88](#)). The retention time is that of a deconvoluted peak and the retention index is taken from the corresponding entry in the Calibration & Standards Library (.csl).

8 Examine the Calibration & Standards Library with which you started.

On the Library menu click **Build One Library**. In the Build One Library editor that appears, click **Files** and in the next dialog click **Load Library**. Change **Files of type** to Calibr. & Stds Library(.csl) and navigate to the **AMDIS32/LIB** folder. Finally click **Fiehn.csl** and **Open** to load the library into the Build One Library editor.

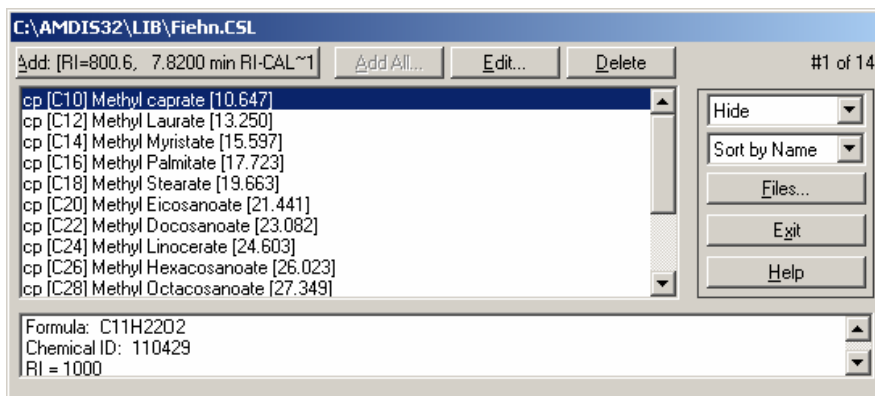


Figure 89 Calibration & Standards Library dialog

Here you can see some of the target compounds that you just successfully located in the **RI-CALIBRATION.d** data file.

The *c* to the left of the compound name shows that the compound is a *calibration check* compound and is used as a calibrant in this analysis type. A *p* signifies that a compound is used as a *performance check* compound.

If you click **Edit** you get the Spectrum Editor (Calibration/Standards Library) dialog for the currently highlighted compound and you can see the sort of information stored for an entry. Click **Cancel** to exit from the dialog and leave the values unchanged.

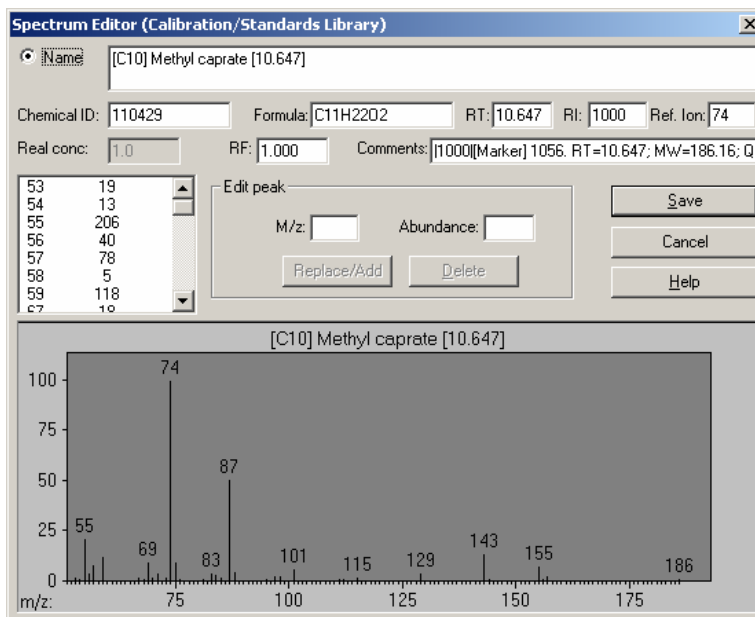


Figure 90 Spectrum editor

Analysis Type - Use Retention Index Data

If you have retention index (RI) information stored in a Target Compounds Library, an RI Calibration Data file can be used as an additional quality check on the results determined by the standard AMDIS spectral comparison process.

A Target Compounds Library has a .msl file extension and an RI Calibration Data file has a .cal file extension.

The RI Calibration Data file is created as you just saw using the RI Calibration/Performance analysis type see [“Analysis Type - RI Calibration/Performance”](#) on page 148.

In the Use Retention Index Data analysis type the RI Calibration Data file is acting as an *external calibration* file. In practice you only need to create a new one of these when the chromatographic conditions you are using have changed significantly such that the correlation between retention time and retention index in the calibration file is no longer accurate enough.

The Use Retention Index Data analysis type first identifies target compounds using spectral comparisons only (exactly as in the Simple analysis type). Next it uses the actual retention times of the identified targets to compute their retention index values; by linear interpolation techniques using the calibration data held in the RI Calibration Data file. Only then does it determine what to do with these retention indices.

If you have elected to use RI information (see below), for each identified target the software compares its calibration-computed RI value with the value held in the Target Compounds Library and adjusts the target's match value according to how good the fit is.

If you are not using RI information, the target match value is not adjusted and the RI information is simply used in reporting the deviation between computed and library RI values.

It is a prerequisite of this analysis type that a Target Compounds Library and an RI Calibration Data file exist. Further, the calibration information can only be of any use for compounds in the library that have RI values stored for them.

Before following the instructions below, ensure that AMDIS is in a known state by following the instructions in Figure “To put AMDIS into an initialized state” on page 116. There should now be an *empty* Confirm window displayed.

- 1 Click **File > Open** and in the Select Data File dialog, navigate to the **AMDIS32\DATA\HP** folder, select the file **L-HISTIDINE.d** and click **Open**.

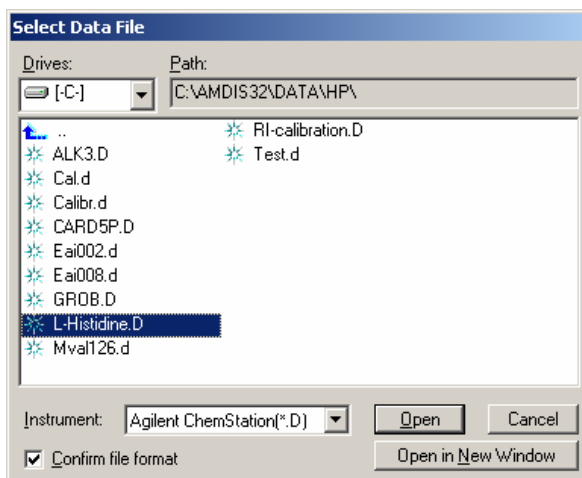


Figure 91 Select data file dialog (L-HISTIDINE.d)

- 2 Click **Analyze > Analyze GC/MS Data** and in the Analyze GC/MS Data dialog set Type of analysis to **Use Retention Index Data**. Use the **Target Library** and **RI Calib. Data** to select the **Fiehn.msl** and **Fiehn.cal** files, respectively. The dialog should now resemble the following:

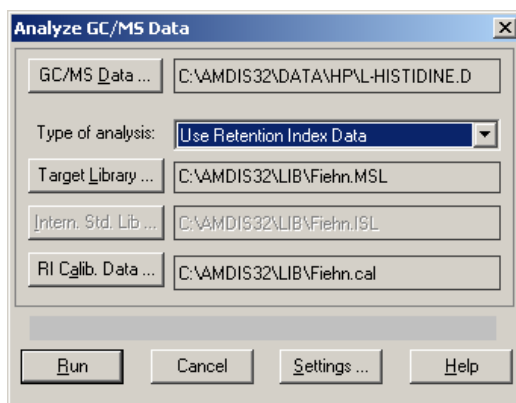


Figure 92 Analyze GC/MS Data dialog (Use Retention Index Data)

The **Intern. Std. Lib** in the dialog is disabled. This type of analysis does not involve the use of internal standards.

- 3 Although retention index values are automatically computed in this analysis type, you still need to tell AMDIS exactly how to use them. Click **Settings** and in the Analysis Settings dialog click **Identif.**

Now set **Minimum match factor** to **60**. For later, note that the Match factor penalties group has its Maximum penalty value set at 20 and its Level set to Infinite.

You are now ready to perform the analysis.

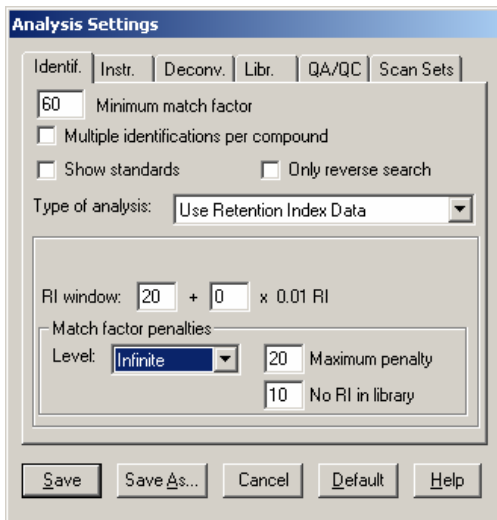


Figure 93 Settings dialog (Use Retention Index Data)

- 4 Click **Save** to exit the Analysis Settings dialog and return to the Analyze GC/MS Data dialog.

If the selected GC/MS data file has previously been analyzed, the message **Previous analysis results will be replaced!** appears above the bottom row of buttons; otherwise this area is blank.

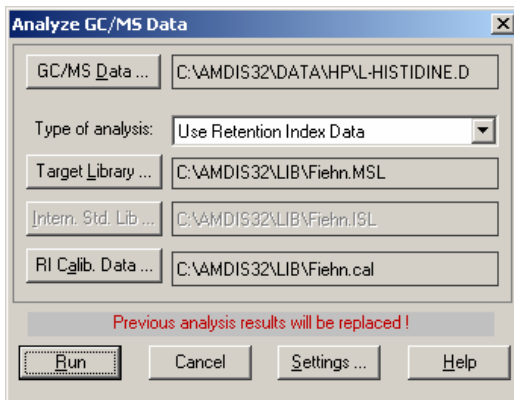


Figure 94 Analyze GC/MS Data dialog showing warning (Use Retention Index Data)

When first running this example, the analysis results file for L-HISTIDINE.d should not exist. In any case, previous analysis results for a sample are *always overwritten* when a new analysis is performed and so this warning is completely normal as you reanalyze samples with different parameters. Click **Run** to start the analysis. Within a few seconds the results are displayed in the format you are now accustomed to.

- Learn a little more about the **Information Lists** area of the display. You previously saw how to control the display of all hits or just the best hits. Now, you learn how to undock the Information Lists area so that you can conveniently view all of a hit's results information at once without the need to scroll.

First select the target at retention time 17.6567 minutes.

Now right-click in the upper left-hand quadrant of the Information Lists area to launch a 2-entry menu.

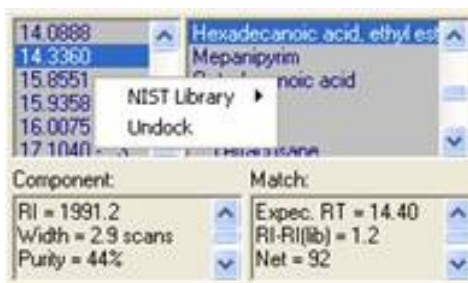


Figure 95 Controlling docking and undocking

Click **Undock**. The Information Lists area now becomes a window.

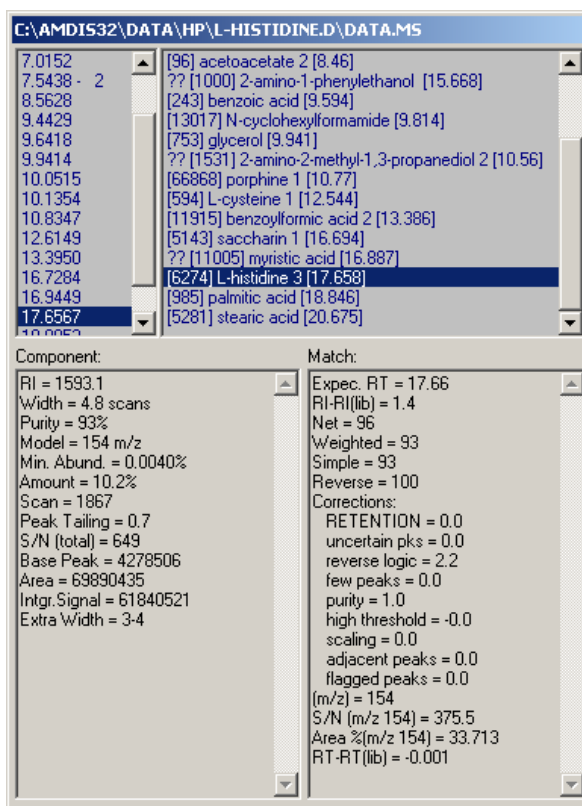


Figure 96 Information Lists area in **Undock** mode

You can now use Windows techniques to drag and expand the window. Experiment moving and stretching it. Note especially that, if you stretch the window at the top edge, it increases the relative size of the upper quadrants, whereas if you stretch it at the bottom edge, it increases the relative size of the lower quadrants.

If you right click in the lower left quadrant, the Component list of information, you can view the lower level parameters. Within this dialog, you can determine which parameters to be displayed and the order in which to display the parameters.

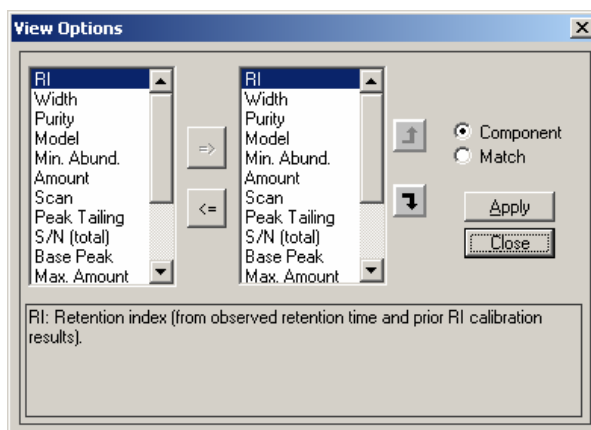


Figure 97

The Information Lists area is fully explained in the online Help, but for now we shall just highlight a few values relating to RI.

In the Match area L-Histidine has an RI-RI(lib) value of 1.4, indicating how close the RI match is. This value is the difference in RI between the value computed for this component from the RI Calibration Data file (.cal) and the value held in the Target Compounds Library (.msl). The calculated value is shown in the Component area and is 1593.1 in this case; with 1591.7 being the library value.

If the RI-RI(lib) label and value are absent the library entry does not have an RI value specified.

The Net value (96 in this case) is an overall indicator of the matching process and takes all Corrections: into account. Finally note that in the Corrections: list, the RETENTION value is 0.0. Note that the Net match factor is increased over the weighted value of 93 due to the high purity of the component and the reverse search logic.

Analysis Type - Use Internal Standards for RI

If you have retention index (RI) information stored in a Target Compounds Library, an Internal Standards Library (.isl) can be used to aid the target identification process. To take advantage of this it is also necessary that one or more of the internal standard compounds represented by the library are co-injected with the sample.

A Target Compounds Library has a .msl file extension and an Internal Standards Library has a .isl file extension.

The Use Internal Standards for RI analysis type first identifies target and internal standard compounds using spectral comparisons only (exactly as in the Simple analysis type). Next, using the known RI values and actual retention times of the internal standards that have been found, it computes RI values for all of the identified library targets using linear interpolation techniques.

Only now does it determine what to do with these retention indices. If you have elected to use RI information (see below), for each identified target it the computed RI value is compared with the value held in the Target Compounds Library and the target's match value is adjusted according to how good the fit is. Otherwise the target match value is not adjusted and the RI information is simply used in reporting the deviation between computed and library RI values.

At the completion of every analysis, a QA/QC report is created that provides information about all of the internal standards that have been found, as well as those that were not found.

It is a prerequisite of this analysis type that a Target Compounds Library and an Internal Standards Library exist. Further, the internal standards' information can only be of any use for compounds in the library that have RI values stored for them.

Before following the instructions below, ensure that AMDIS is in a known state by following the instructions in [“To put AMDIS into an initialized state”](#) on page 116. There should now be an *empty* Confirm window displayed.

- 1 Click **File > Open** and in the Select Data File dialog, navigate to the AMDIS32\DATA\HP folder, select the file **RI-Calibration.d** and click **Open**.
- 2 Click **Analyze > Analyze GC/MS Data** and in the Analyze GC/MS Data dialog set Type of analysis to **Use Internal Standards** for RI. Use **Target Library ...** and **Intern. Std. Lib ...** to select the **Fiehn.msl** and **Fiehn.isl** files, respectively. The dialog should now resemble:

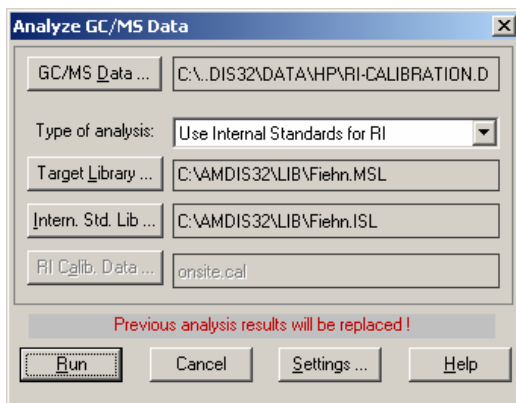


Figure 98 Analyze GC/MS Data dialog (Use Internal Standards for RI)

The **RI Calib. Data...** in the dialog is disabled. This type of analysis does not involve the use of an RI Calibration Data file.

- 3 Although retention index values are automatically computed in this analysis type, you still need to tell AMDIS exactly how to use them.

Click **Settings** and in the Analysis Settings dialog click **Identif...** Now set Minimum match factor to 65 and check **Show Standards**.

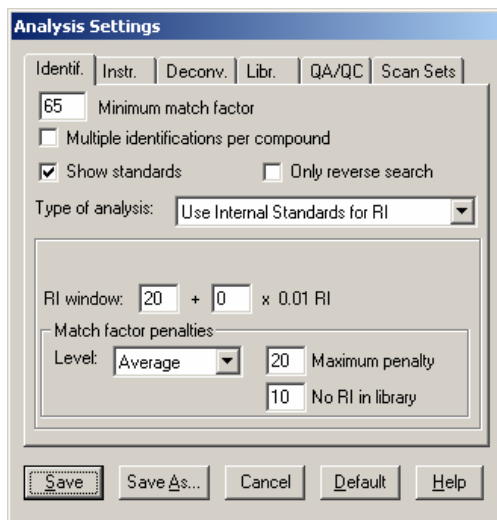


Figure 99

- 4 We are now ready to perform the analysis. Click **Save** to exit the Analysis Settings dialog and return to the Analyze GC/MS Data dialog.

Click **Run** to start the analysis. Within a few seconds the results are displayed in the usual format.

Set the Information Lists area to view only the best hits and then examine RI-RI(lib) values for the found targets. If the RI-RI(lib) label and value are absent the library entry does not have an RI value specified.

Methyl Palmitate (found at 17.7296 minutes) has a value of 0 and is a perfect match.

C:\AMDIS32\DATA\HP\RI-CALIBRATION.D\DATA.MS	
6.5314 - 2	?? [8871] 2-hydroxypyridine [6.519]
6.8141 - 4	? [757] glycolic acid [7.049]
6.8717 - 2	[107689] L-(+) lactic acid [6.851]
7.8200 - 4	[8091] octanoic acid methyl ester [7.812]
8.1653 - 6	??? [3469] gentisic acid [16.123]
9.2462 - 5	[C9] Methyl Pelargonate [9.248]
9.4038 - 6	?? [1000] 2-amino-1-phenylethanol [15.668]
9.5952 - 3	[243] benzoic acid [9.594]
9.9781 - 2	[753] glycerol [9.941]
10.5729 - 2	??? [239] Beta- alanine 2 [14.555]
10.6411 - 6	[C10] Methyl caprate [10.647]
10.7731	[66868] porphine 1 [10.77]
13.2446 - 4	[C12] Methyl Laurate [13.250]
15.5978	[C14] Methyl Myristate [15.597]
17.7296	[C16] Methyl Palmitate [17.723]
18.9013	[985] palmitic acid [18.846]
19.6725	[C18] Methyl Stearate [19.663]
20.7122	[5281] stearic acid [20.675]
21.4565	[C20] Methyl Eicosanoate [21.441]
23.1061	[C22] Methyl Docosanoate [23.082]
23.1524	[8343] dioctyl phthalate [23.163]
24.6349	[C24] Methyl Linocerate [24.603]
26.0611	[C26] Methyl Hexacosanoate [26.023]
27.3935	?? [C28] Methyl Octacosanoate [27.349]

Component:	Match:
RI = 2800.0	Exp. RT = 27.35
Width = 3.6 scans	RI-RI(lib) = 0.
Purity = 100%	Net = 69
Model = 99 m/z	Weighted = 59
Min. Abund. = 0.040%	Simple = 92
Amount = 2.13%	Reverse = 92
Scan = 3423	Corrections:
Peak Tailing = 1.7	RETENTION = 0.0
S/N (total) = 329	uncertain pks = -0.4
Base Peak = 425014	reverse logic = 10.1
Area = 10144531	few peaks = 0.0
Intgr. Signal = 9264819	purity = 1.0
Extra Width = 2-4	high threshold = -0.0
	scaling = 0.0
	adjacent peaks = 0.0
	flagged peaks = -0.0
	(m/z) = 74
	S/N (m/z 74) = 131.8
	Area %(m/z 74) = 16.085

Figure 100

- 5 Examine the QA/QC report that was mentioned earlier and review the AMDIS Results dialog.

Click **Info** (found near the top left of the Confirm window) to display the AMDIS Results dialog.

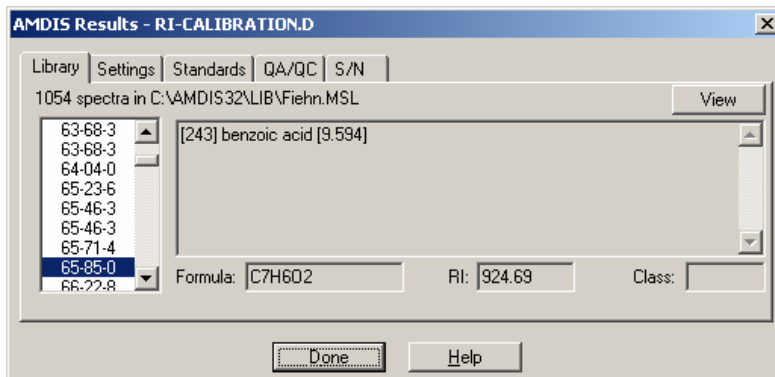


Figure 101 AMDIS Results dialog (Library tab)

When it appears, the dialog presents the **Library** tab and shows information from the Target Compounds Library for the currently selected target (Benzoic Acid in our example). Note the textual header giving the name of the Target Compounds Library and saying how many spectra there are in it.

Click **Standards** to display a tab listing pertinent non-spectral information about all of the internal standards in the Internal Standards Library.

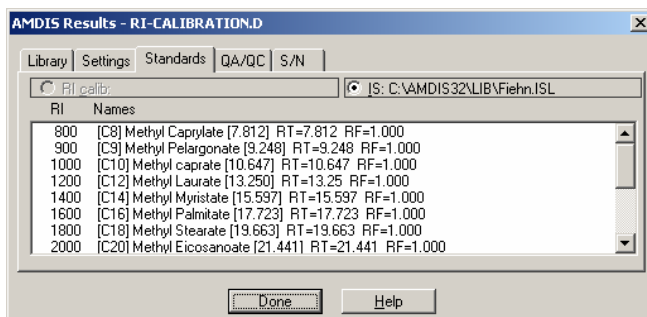


Figure 102 AMDIS Results dialog (Standards tab)

The **RI calib:** button is disabled. The *RI Calibration* information is only available when an RI Calibration Data file (.cal) has been used in an analysis (briefly examined in the next tutorial section).

Now click **QA/QC** to view the *QA/QC Results* report and scroll down to the bottom of the report. This report indicates the retention index standards that have been found and those that were missed.

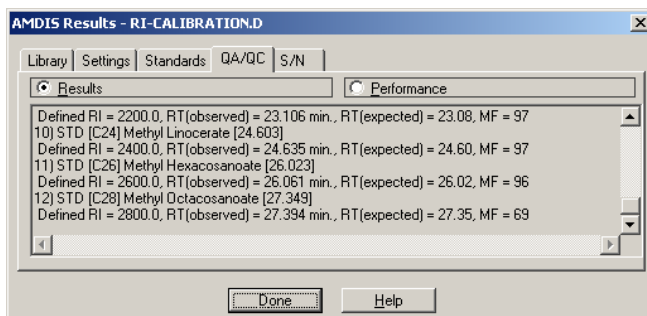


Figure 103 AMDIS Results - QA/QC report

Information about the internal standards that have been found is presented. If any of the internal standards have not been found, a separate list of those would also be presented. The **Performance** button is described later see “[Analysis Type - Performance Check for RI](#)” on page 176.

The **Settings** tab presents important parameter values used for the most recent analysis and the S/N tab provides various information relating to the signal-to-noise analysis performed on the data file.

It is important to understand that much of the information presented in the AMDIS Results dialog relates to the analysis results currently being examined. As such, these values are not necessarily those that would be used for a new analysis, as analysis settings may have been modified in the Analysis Settings dialog. The exception is that the library information is taken from the current state of the library; which in general may have been edited since the analysis.

Click **Done** to close the dialog.

6 Take a quick look at the Internal Standards Library.

On the Library menu click **Build One Library**. In the Build One Library editor that appears, click **Files** and in the next dialog click **Load Library**. Change Files of type to Internal Std. Library(.isl) and navigate to the **AMDIS32/LIB** folder. Finally click **FIHN.isl** and **Open** to load the library into the Build One Library editor.

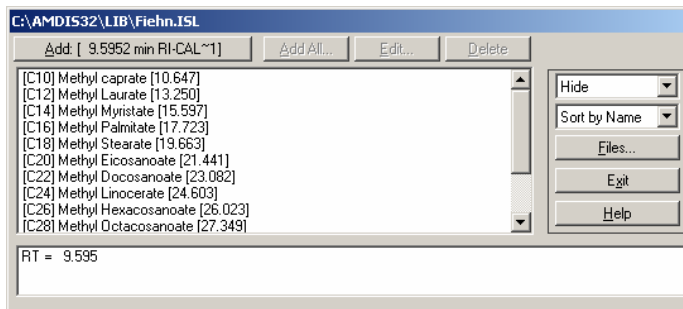


Figure 104 Internal Standards Library dialog

Here you can see the list of internal standard compounds that you just located in the RI-Calibration.d data file.

If you click **Edit** you get the Spectrum Editor (Internal Standards Library) dialog for the currently highlighted compound and you can see the sort of information stored for an entry. Click **Cancel** to exit from the dialog and leave the values unchanged.

Click **Exit** to return to the Confirm window.

Analysis Type - Use RI Calibr. Data + Internal Std.

The Use RI Calibr. Data + Internal Std. analysis type is equivalent in operation to the Use Retention Index Data analysis type that you have already learned about. Therefore, to derive maximum benefit, you need a Target Compounds Library with retention index (RI) information stored in it and an RI Calibration Data file.

An Internal Standards Library is also required, but unlike in the Use Internal Standards for RI analysis type you just examined, the internal standards are used only as a functional check of the acquisition system and are not used in predicting target RI values; since this is done using the RI Calibration Data file.

A Target Compounds Library has a .msl file extension, an RI Calibration Data file has a .cal file extension and an Internal Standards Library has a .isl file extension.

As in the Use Internal Standards for RI analysis type, internal standards are co-injected with the sample.

The Use RI Calibr. Data + Internal Std. analysis type first identifies target and internal standard compounds using spectral comparisons only and then computes retention index values for the identified targets using the calibration data held in the RI Calibration Data file.

If you have elected to use RI information (using the Analysis Settings dialog), for each identified target the calibration-computed RI value is compared with the value held in the Target Compounds Library and adjusted to the target's match value according to how good the fit is. If you are not using RI information, the target match value is not adjusted and the RI information is simply used in reporting the deviation between computed and library RI values.

The deviation between the expected and observed retention times of each internal standard is given in the QA/QC report (see below for details).

It is a prerequisite of this analysis type that a Target Compounds Library, an RI Calibration Data file and an Internal Standards Library exist. Further, the calibration information can only be of any use for compounds in the target library that have RI values stored for them. For the purpose of this example, suitable files are provided.

Before following the instructions below, ensure that AMDIS is in a known state by following the instructions in “To put AMDIS into an initialized state” on page 116. There should now be an *empty* Confirm window displayed.

- 1 Click **File > Open** and in the Select Data File dialog, navigate to the **AMDIS32\DATA\HP** folder, select the file **RI-Calibration.d** and click **Open**.
- 2 Click **Analyze > Analyze GC/MS Data** and in the Analyze GC/MS Data dialog set Type of analysis to **Use RI Calibr. Data + Internal Standard..** Use **Target Library...**, **Intern. Std. Lib...** and **RI Calib. Data...** to select the **Fiehn.msl**, **FIEHN.isl** and **Fiehn.cal** files, respectively. The dialog should now resemble:

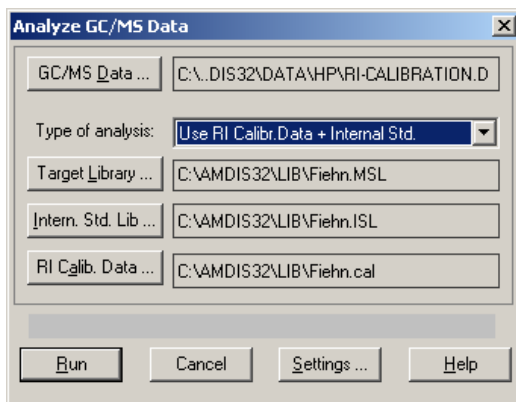


Figure 105 Analyze GC/MS Data dialog (Use RI Calibr. Data + Internal Std.)

- 3 Although retention index values are automatically computed in this analysis type, you still need to tell AMDIS exactly how to use them.
Click **Settings** and in the Analysis Settings dialog click **Identif..** Now set Minimum match factor to 65.
We are now ready to perform the analysis.
- 4 Click **Save** to exit the Analysis Settings dialog and return to the Analyze GC/MS Data dialog.
Click **Run** to start the analysis. Within a few seconds the results are displayed in the usual format.
Set the Information Lists area to view only the best hits of the targets and then examine RI-RI(lib) values for the found targets.

If the RI-RI(lib) label and value are absent the library entry does not have an RI value specified.

If you check the value for the Methyl Palmitate internal standard you can see that it is no longer exactly 0; and so is unlike the previous Use Internal Standards for RI analysis method where it was exactly 0. This time the internal standards are being treated more or less the same as normal targets and in this case are subject to the normal RI computation (from the RI Calibration Data file) and RI adjustments.

A discussion on this in more detail follows, but first click **Info** to display the AMDIS Results dialog and then click **Standards** to display RI calib data. This is a list of all compounds in the RI Calibration Data file and shows their stored RT, RI and Names information.

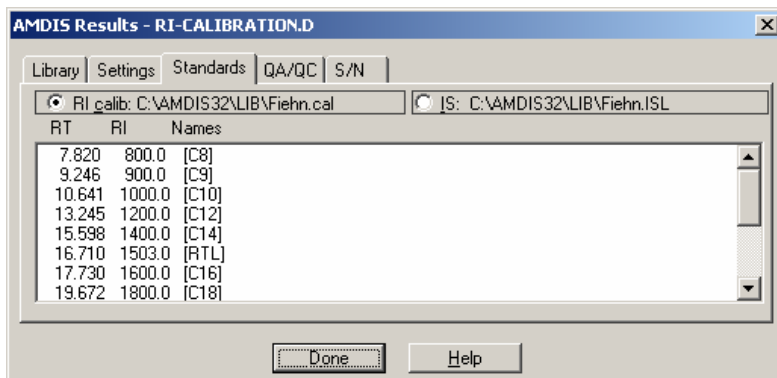


Figure 106 AMDIS Results dialog (Standards tab, RI calib)

It is this RT and RI calibration information that is used to predict the RI of a target from its found RT (see below).

If you click **IS**, the same type of information as described previously for the **Use Internal Standards for RI** analysis type is displayed.

Similarly, click **QA/QC** and scroll down to the bottom of the Results section. Both of the internal standards have been found. The difference between RT(observed) and RT(expected), together with the match value (MF), an indication of the overall acquisition system is performing.

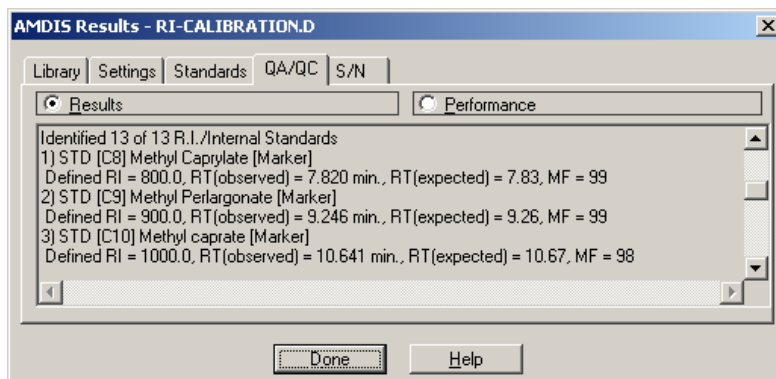


Figure 107 QA/QC results showing the internal standards

It is important to understand that much of the information presented in the AMDIS Results dialog relates to the analysis results currently being examined. As such, these values are not necessarily those that would be used for a new analysis, as analysis settings may have been modified in the Analysis Settings dialog. The exception is that the library information is taken from the current state of the library; which in general may have been edited since the analysis.

- 5 Examine how the RI Calibration Data file's RT and RI information are used.

Click **Done** to exit from the AMDIS Results dialog and then select **Palmitic acid** in the undocked Information Lists area. the results look similar to:

C:\AMDIS32\DATA\HP\RI-CALIBRATION.D\DATA.MS	
6.5314	?? [8871] 2-hydroxypyridine [6.519]
6.8141	[757] glycolic acid [7.049]
6.8717	[107689] L-(+) lactic acid [6.851]
7.8200	[8091] octanoic acid methyl ester [7.812]
9.2462 - 2	[C9] Methyl Pelargonate [9.248]
9.5952	[243] benzoic acid [9.594]
9.9781	[753] glycerol [9.941]
10.6411 - 2	[C10] Methyl caprate [10.647]
10.7731	[66868] porphine 1 [10.77]
13.2446	[C12] Methyl Laurate [13.250]
15.5978	[C14] Methyl Myristate [15.597]
17.7296	[C16] Methyl Palmitate [17.723]
18.9013	[985] palmitic acid [18.846]
19.6725	[C18] Methyl Stearate [19.663]
20.7122	[5281] stearic acid [20.675]
21.4565	[C20] Methyl Eicosanoate [21.441]
23.1061	[C22] Methyl Docosanoate [23.082]
23.1524	[8343] dioctyl phthalate [23.163]
24.6349	[C24] Methyl Linocerate [24.603]
26.0611	[C26] Methyl Hexacosanoate [26.023]
27.3935	?? [C28] Methyl Octacosanoate [27.349]

Component:	Match:
RI = 1721.5	Expec. RT = 18.85
Width = 3.9 scans	RI-RI(lib) = 3.1
Purity = 72%	Net = 94
Model = 117 m/z	Weighted = 95
Min. Abund. = 4.1%	Simple = 81
Amount = 0.0202%	Reverse = 94
Scan = 2066	Corrections:
Peak Tailing = 1.0	RETENTION = 0.0
S/N (total) = 31	uncertain pks = 0.0
Base Peak = 4148	reverse logic = 0.0
Area = 96059	few peaks = 0.0
Intr. Signal = 88176	purity = 0.8
Extra Width = 1.3	high threshold = -0.8
	scaling = 0.0
	adjacent peaks = 0.0
	flagged peaks = -0.9
	(m/z) = 313
	S/N (m/z 313) = 11.1
	Area %(m/z 313) = 12.371
	RT-RT(lib) = 0.055

Figure 108 Results for Palmitic acid

The RT of palmitic acid is 18.9013 and so the RI Calibration Data file is examined for the closest RT value below this and the closest above (in this case 17.730 and 19.672) (see [Figure 106](#) “AMDIS Results dialog (Standards tab, RI calib)” on page 172.) These two values and their corresponding RI values (1600 and 1800) are used to predict the RI of palmitic acid by simple interpolation (yielding RI = 1721.48). The difference

between the computed RI value for Palmitic acid and its value in the Target Compounds Library (1718.38) is presented as RI-RI(lib) in the Match results area (1721.48-1718.38=3.1).

Analysis Type - Performance Check for RI

The Performance Check for RI analysis type is used as a check to see how well the acquisition system is performing. You do this by first acquiring a clean mixture of known composition containing one or more *performance check* compounds. You then analyze the results using a Calibration & Standards Library and an RI Calibration Data file; both of which was explained in “Analysis Type - RI Calibration/Performance” on page 148.

A Calibration & Standards Library has a .csl file extension and an RI Calibration Data file has a .cal file extension.

The Calibration & Standards Library is a list of compounds that can be used in the creation of a RI Calibration Data file (as described in “Analysis Type - RI Calibration/Performance” on page 148) and/or for performance checking purposes.

First of all the Performance Check for RI analysis type attempts to identify all compounds in the Calibration & Standards Library using spectral comparisons only and then computes retention index values for the identified targets using the calibration data held in the RI Calibration Data file.

If you have elected to use RI information (using the Analysis Settings dialog), for each identified target is compared to its calibration-computed RI value with the value held in the RI Calibration Data file and adjustment is made to the target's match value according to how good the fit is. If you are not using RI information, the target match value is not adjusted and the RI information is simply used in reporting the deviation between computed and library RI values.

Finally, any *performance check* compounds that have been identified have their data added to the Performance Log file. This file can be inspected at any time and by examining the details of the information, it is possible to detect any gradual (or sudden) deterioration in the acquisition system that might indicate the need for remedial action, or in some cases simply the creation of a new RI Calibration Data file.

It is a prerequisite of this analysis type that a Calibration & Standards Library and an RI Calibration Data file exist. Further, the Calibration & Standards Library should have at least one compound identified as a *performance check* compound. For the purpose of this example, suitable files are provided.

Before following the instructions below, ensure that AMDIS is in a known state by following the instructions in “[To put AMDIS into an initialized state](#)” on page 116. There should now be an *empty* Confirm window displayed.

Since this tutorial illustrates the Performance Log, it would also be beneficial to start with that in a known state. The Performance Log is actually a collection of 3 files that are located in the folder in which AMDIS is installed. Their names are `perform.idx`, `perform.tDB` and `perform.txt`. If these files already exist and you know you want to preserve them, use standard Windows techniques to rename each of them before running this tutorial and then afterwards delete the newly-created files and rename your saved files back to their original names.

- 1 Click **File > Open** and in the Select Data File dialog, navigate to the **AMDIS32\DATA\HP** folder, select the file **RI-Calibration.d** and click **Open**.
- 2 Click **Analyze > Analyze GC/MS Data** and in the Analyze GC/MS Data dialog set Type of analysis to **Performance Check for RI**. Use the **Calib/Std. Lib. ...** and **RI Calib. Data ...** buttons to select the **Fiehn.csl** and **Fiehn.cal** files, respectively. The dialog should now resemble:

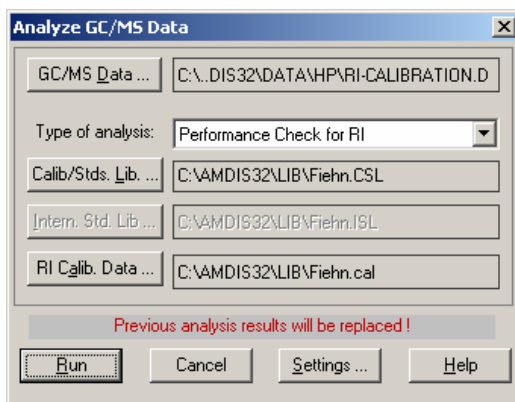


Figure 109 Analyze GC/MS Data dialog (Performance Check for RI)

The **Intern. Std. Lib. ...** button in the dialog is disabled. This type of analysis does not involve the use of internal standards.

- 3 Click **Settings** and in the Analysis Settings dialog click **Identif.**. Now set Minimum match factor to 65.

You are now ready to perform the analysis.

- 4 Click **Save** to exit the Analysis Settings dialog and return to the Analyze GC/MS Data dialog.

Click **Run** to start the analysis. Within a few seconds the results are displayed in the usual format.

If you undock and expand the Information Lists area and set it to show only best-hit data, the target list appears as in [Figure 110](#) on page 178.

C:\AMDIS32\DATA\HP\RI-CALIBRATION.D\DATA.MS	
7.8200 - 3	[C8] Methyl Caprylate [7.812]
9.2462 - 4	[C9] Methyl Pelargonate [9.248]
10.6411 - 4	[C10] Methyl caprate [10.647]
13.2446 - 3	[C12] Methyl Laurate [13.250]
15.5978 - 3	[C14] Methyl Myristate [15.597]
17.7296	[C16] Methyl Palmitate [17.723]
19.6725	[C18] Methyl Stearate [19.663]
21.4565	[C20] Methyl Eicosanoate [21.441]
23.1061	[C22] Methyl Docosanoate [23.082]
24.6349	[C24] Methyl Linocerate [24.603]
26.0611	[C26] Methyl Hexacosanoate [26.023]
27.3935	? [C28] Methyl Octacosanoate [27.349]

Component:	Match:
RI = 800.0	Expec. RT = 7.81
Width = 3.4 scans	RI-RI(lib) = -0.0
Purity = 96%	Net = 99
Model = 74 m/z	Weighted = 97
Min. Abund. = 0.011%	Simple = 98
Amount = 3.23%	Reverse = 100
Scan = 294	Corrections:
Peak Tailing = 1.0	RETENTION = n/a
S/N (total) = 417	uncertain pks = 0.0
Base Peak = 1515921	reverse logic = 1.0
Area = 15337399	few peaks = 0.0
Intr. Signal = 14588658	purity = 1.0
Extra Width = 1-1	high threshold = -0.0
	scaling = 0.0
	adjacent peaks = 0.0
	flagged peaks = 0.0
	(m/z) = 74
	S/N (m/z 74) = 248.8
	Area %(m/z 74) = 35.590
	RT-RT(lib) = 0.008

Figure 110 Information Lists area after Performance Analysis

5 Examine the Calibration & Standards Library that you used.

On the Library menu click **Build One Library**. In the Build One Library editor that appears, click **Files** and in the next dialog click **Load Library**. Change Files of type to Calibr. & Stds Library(.csl) and navigate to the **AMDIS32/LIB** folder. Finally click **Fiehn.csl** and **Open** to load the library into the Build One Library editor.

Select **Sort by RI** and examine the full list of compounds by scrolling.

Retention Time	Compound Name
7.8200 - 3	[C8] Methyl Caprylate [7.812]
9.2462 - 4	[C9] Methyl Pelargonate [9.248]
10.6411 - 4	[C10] Methyl caprate [10.647]
13.2446 - 3	[C12] Methyl Laurate [13.250]
15.5978 - 3	[C14] Methyl Myristate [15.597]
17.7296	[C16] Methyl Palmitate [17.723]
19.6725	[C18] Methyl Stearate [19.663]
21.4565	[C20] Methyl Eicosanoate [21.441]
23.1061	[C22] Methyl Docosanoate [23.082]
24.6349	[C24] Methyl Linocerate [24.603]
26.0611	[C26] Methyl Hexacosanoate [26.023]
27.3935	? [C28] Methyl Octacosanoate [27.349]

Component:	Match:
RI = 800.0	Expec. RT = 7.81
Width = 3.4 scans	RI-R(lib) = -0.0
Purity = 96%	Net = 99
Model = 74 m/z	Weighted = 97
Min. Abund. = 0.011%	Simple = 98
Amount = 3.23%	Reverse = 100
Scan = 294	Corrections:
Peak Tailing = 1.0	RETENTION = n/a
S/N (total) = 417	uncertain pks = 0.0
Base Peak = 1515921	reverse logic = 1.0
Area = 15337399	few peaks = 0.0
Intr. Signal = 14588658	purity = 1.0
Extra Width = 1-1	high threshold = -0.0
	scaling = 0.0
	adjacent peaks = 0.0
	flagged peaks = 0.0
	(m/z) = 74
	S/N (m/z 74) = 248.8
	Area %(m/z 74) = 35.590
	RT-RT(lib) = 0.008

Figure 111 Calibration & Standards Library (Performance Check)

The compound entries that are flagged with cp or #p are the *performance check* compounds and in this case you can see that there are a total of 14. Click **Exit** to leave the dialog.

- 6 Click **Info** to display the AMDIS Results dialog and then click QA/QC, followed by Performance to display data for the *performance check* compounds that have been found.

AMDIS Results - RI-CALIBRATION.D								
Library Settings Standards QA/QC S/N								
Results Performance								
RT	RI	Height	Width	Tailing	Match	Date	File	
[C10]	Methyl caprate	[10.647]	10.641	1000.0	439	3.6	1.00	98 13:06 Apr 24 2008
[C12]	Methyl Laurate	[13.250]	13.245	1200.0	459	3.8	1.00	98 13:06 Apr 24 2008
[C14]	Methyl Myristate	[15.597]	15.598	1400.0	497	3.8	0.90	98 13:06 Apr 24 2008
[C16]	Methyl Palmitate	[17.723]	17.730	1600.0	513	3.8	0.90	97 13:06 Apr 24 2008

Figure 112 AMDIS Results dialog (Performance Check)

This example has only one set of performance data recorded. As this analysis is repeated over time, the data build up until many more lines of data are acquired for each compound. By observing the reported RT, RI, Height, Width, Tailing and Match data, you can spot a gradual change over time (or a sudden larger change) of one or more of these values.

Click **Done** to exit.

- 7 Print or save the results.

On the **File** menu click **Print Spectra** to invoke the Print Spectra dialog.

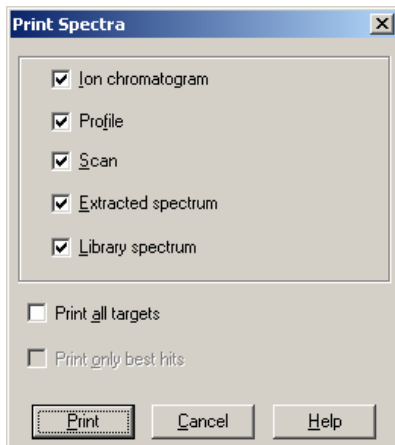


Figure 113 Print Spectra dialog

The settings allow you some control over the graphical items that are printed and are fully described by clicking on **Help**. Experiment if you wish by using **Print**, but finally return to the Confirm window.

- 8 On the **File** menu click **Print Text Report** to invoke the Print Text Report dialog.

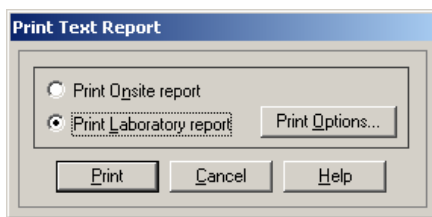


Figure 114 Print Text Report dialog

If you click **Print Options** the ensuing dialog's settings allow you some control over the textual items that can be printed and are fully described in online Help. Experiment if you wish by using **Print**, but finally return to the Confirm window.

- 9 It is possible to save most of the results of an analysis in textual format for subsequent examination, or simply for historical recording, or any other reason. On the **File** menu click **Generate Report** to invoke the Generate Report dialog.

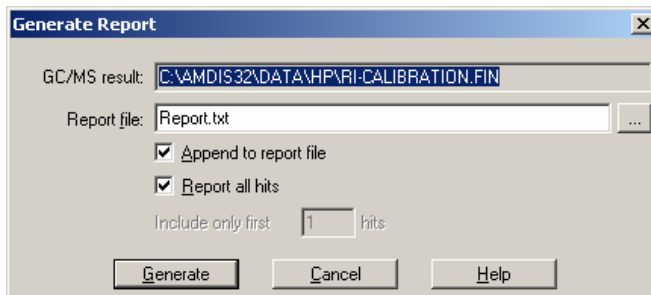


Figure 115 Generate Report dialog

Click **C**ancel to return to the Confirm window.

Recap

In this section you were introduced to the most commonly used main window of AMDIS (the Confirm window) and the other major window was briefly described (the Results window). You have learned the basics of manipulating the display of the Confirm window and have worked through each analysis type in turn. In doing this you have used all of the file types that AMDIS supports and you have been introduced to many of the most significant commands and some parameters. You have been shown how to view and perform basic interpretation of results and it was mentioned how to print and create reports.

All of this should give you the confidence to begin *serious* work with AMDIS.

The rest of this User Guide is written in the style of a reference document giving full details of all features with little or no reference made to supplied example files. However, it does cross-reference tutorial sections where relevant to enable you to refresh your memory of a technique that you perhaps use only occasionally.

Recommended Settings

The settings that are shown in the following figures are the recommended settings that are used for most analyses. These recommendations are best used as a starting point to become familiar with AMDIS. The AMDIS online help provides detailed information for each setting. Individual data files, sample matrices, and experimental objectives may require other settings.

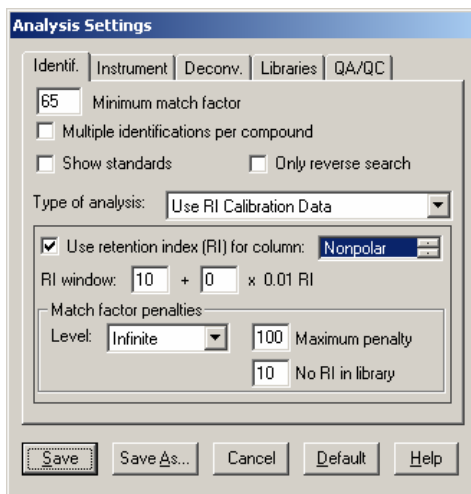


Figure 116

It is important to note that the recommended settings are extremely prejudiced against prospective targets that do not meet the retention indice requirements (± 10 RI units).

Select the **Instrument** tab and adjust values as shown below.

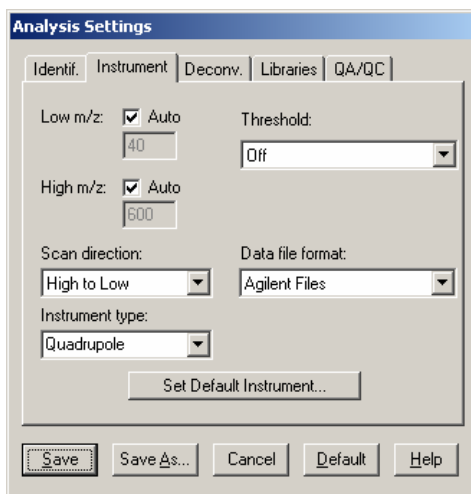


Figure 117

Select the **Deconv.** Tab and adjust values as shown below.

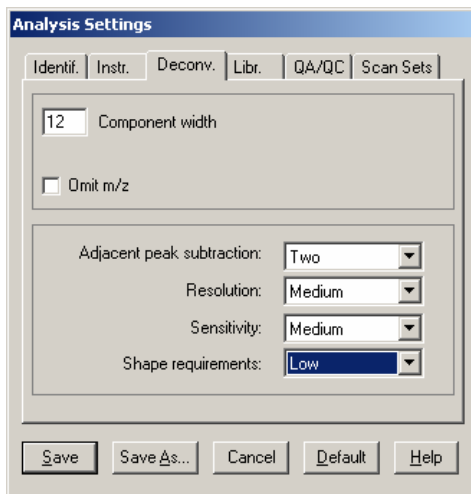


Figure 118 Deconv. Tab in the Analysis Settings dialog box

Select the **Libraries** tab and if the Target Compounds Library is not Fiehn.msl, select **Select New...** to browse for it.

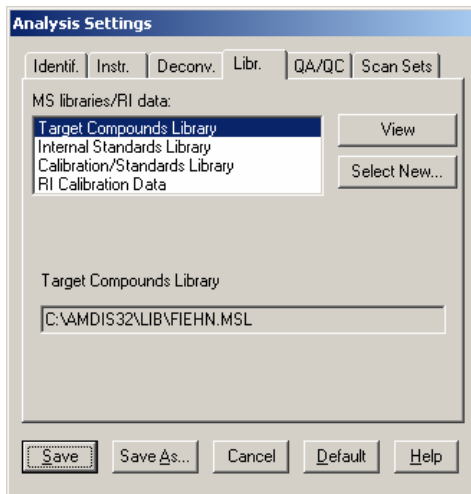


Figure 119

If the Internal Standards Library is not Fiehn.isl, select **Select New...** to browse for it.

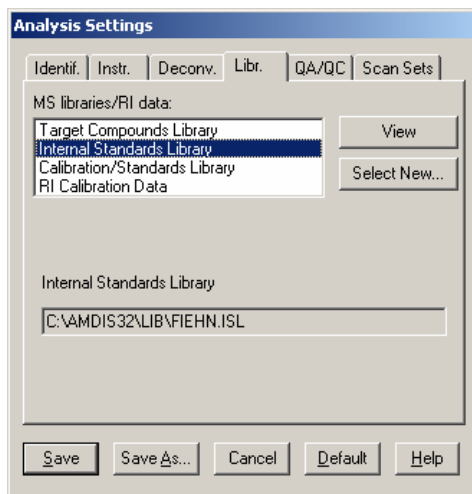


Figure 120

If the Calibration/Standards Library is not Fiehn.csl, select **Select New...** to browse for it.

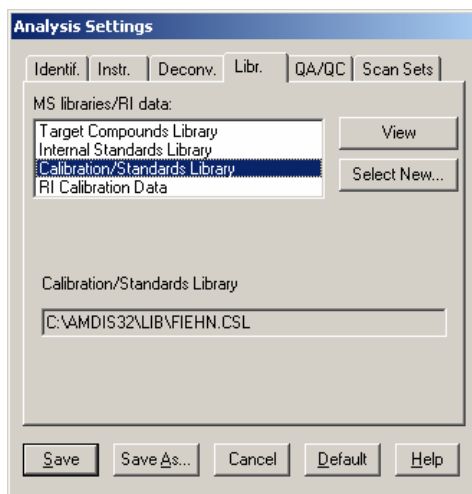


Figure 121

If the RI Calibration Data is not Fiehn.cal, select **Select New...** to browse for it.

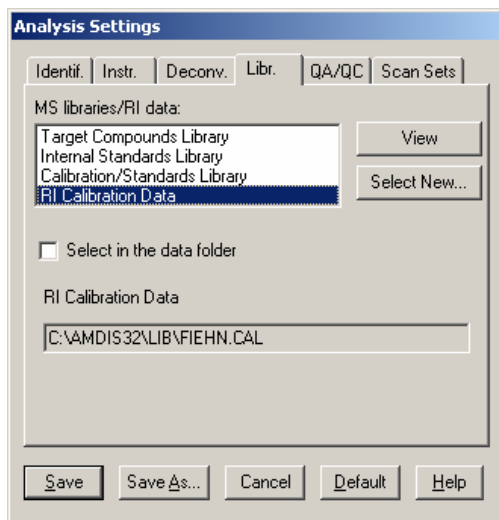


Figure 122

Select the **QA/QC** tab and adjust values as shown below.

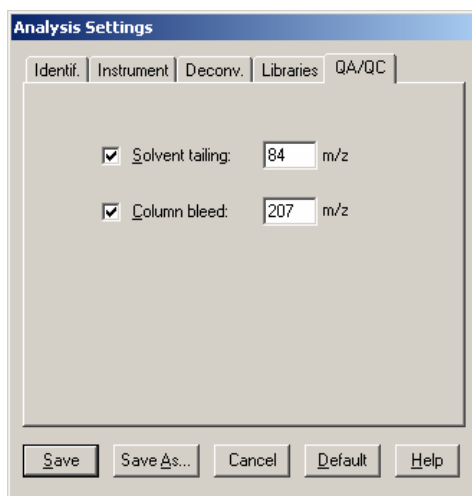


Figure 123

The components of the Agilent Fiehn Library are the recommended libraries for all of the types of analysis.

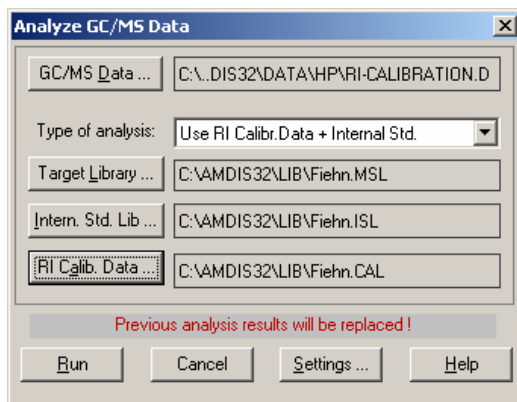


Figure 124

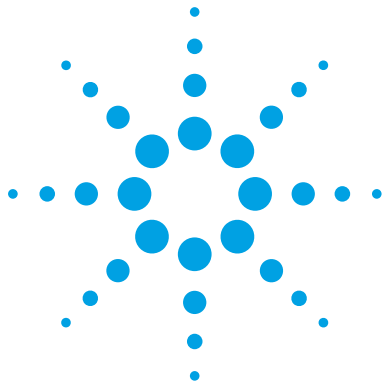
Finally select **SAVE** to save the analysis settings. The settings are saved in the AMDIS32 folder. The name of the file is onsite.ini. Before exiting the dialog, it is possible to save the current analysis settings for future reuse. Click **Save As** to invoke the Save Setting As dialog, enter a File name and Save in location of your choice and click **Save** to create your named INI file and to exit from that dialog back to the Analysis Settings dialog.

If you wish to subsequently reuse the settings in a saved INI file, proceed as follows:

- 1 **Exit** from AMDIS
- 2 Delete the file onsite.ini in the AMDIS32 folder.
- 3 Locate the saved INI file, whose values you wish to restore for use in AMDIS.
- 4 Make a copy of it in the AMDIS32 folder and then rename it as onsite.ini.
- 5 Restart AMDIS. The restored values are now in use.

It is important to understand that the Analysis Settings dialog reflects the values you last set in it and not always therefore the values used to analyze the data file, whose results you are currently examining. If you want to see the pertinent values used for the current file, use the **View > Auxiliary Information** command and select its dialog's **Settings** tab.

A single metabolite may be derivatized into several compounds that elute at different retention times. Each of the derivatized compounds share the same CAS number. For any given search of a data file, AMDIS only reports a compound or CAS number once. AMDIS reports the derivative that possesses the highest match factor (MF). It does not report the other derivatives. The detection of one of these peaks is necessary and sufficient to claim the identification of this compound in a particular sample. The use of the screener version of the library in the ChemStation is more likely to identify and report all of the derivatized compounds.



7 Using CAS and PubChem Numbers

KEGG LIGAND Database	193
MetaCyc Database	196
ChemSpider Database	199
SciFinder Substance Identifier Database	202
PubChem Compound Database	203

This section describes how to use CAS entries and PubChem numbers to retrieve supplemental chemical and biochemical information on metabolites.

The Agilent Fiehn Library contains CAS and PubChem numbers used to identify metabolites in biological samples. You can use these numbers to derive biologically relevant information from quantitative and qualitative comparisons.

It is critical that library searches provide data that can be used to perform biological interpretations. Such interpretations can be performed by placing metabolites into chemical and biochemical context, notably by mapping to biochemical pathways, but also by comparing physicochemical properties or scientific literature.

CAS entries do not necessarily support unique records for each chemical structure, because CAS lists multiple identifiers for different versions of chemicals (e.g. sodium or potassium salts of anions of organic acids). The Fiehn library uses best possible numbers to refer to CAS entries, and in addition, Internet resources often recognize different CAS numbers as variants of a unique metabolite. Contact Agilent if you find inadequate annotation of metabolites with either PubChem or CAS entries.




7 Using CAS and PubChem Numbers

The examples that follow use oxaloacetate with a CAS # 328-42-7, and a PubChem # 970. These examples demonstrate how these compound identifiers are used to query biochemical and chemical databases. Resulting information ranges from lists of synonym names to structural, chemical and biochemical data and comprises both calculated properties as well as links to other external databases.

KEGG LIGAND Database

- 1 To access this database enter <http://www.genome.jp/ligand/> into your internet browser.
- 2 Scroll down to **Search Compound** and select **DBlinks** from the dropdown list. Enter **328-42-7** in the adjacent text box for the CAS number and click **Go**.



KEGG LIGAND Database
Molecular building blocks of life in the chemical space

KEGG2 PATHWAY BRITE MODULE LIGAND COMPOUND GLYCAN REACTION RMODUL

Enter C/G/R/RP/RC numbers (Example) C00389 C05903 C06562 C08650 C09727 C09762

Filter Pathway mapping Brite mapping Get title Get entry Clear

Chemical Substances and Reactions

KEGG LIGAND contains our knowledge on the universe of chemical substances and reactions that are relevant to life. It is a composite database consisting of COMPOUND, GLYCAN, REACTION, RPAIR, RCLASS, and ENZYME databases, whose entries are identified by C, G, R, RP, RC, and EC numbers, respectively. ENZYME is derived from the IUBMB/IUPAC Enzyme Nomenclature, but the others are internally developed and maintained.

Database	Identifier	Content	Specialized entry point
LIGAND	COMPOUND	C number	Chemical compound structures
	GLYCAN	G number	Glycan structures
	REACTION	R number	Biochemical reactions
	RPAIR	RP number	Reactant pair alignments
	RCLASS	RC number	Reaction class
	ENZYME	EC number	Enzyme nomenclature

Specialized entry points: [KEGG COMPOUND](#), [KEGG GLYCAN](#), [KEGG REACTION](#)

Search for

bfind mode bget mode

LIGAND Relational Database

The primary database of KEGG LIGAND is a relational database with the [KegDraw](#) interface, which is used to generate the secondary (flat file) database for [DBGET](#). A read-only copy of the LIGAND relational database is made publicly accessible.

Search COMPOUND

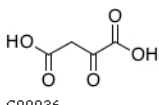
7 Using CAS and PubChem Numbers

This displays a single result in the Compound Data Search Result screen.

Compound Data Search Result Top

Number of entries in a page

Items : 1 - 1 of 1

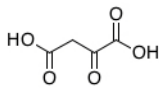
No	Entry	Structure	Name	Formula	DBlinks
1	C00036	 C00036	Oxaloacetate Oxalacetic acid Oxaloacetic acid 2-Oxobutanedioic acid Oxosuccinic acid keto-Oxaloacetate	C ₄ H ₄ O ₅	CAS: 328-42-7

Items : 1 - 1 of 1

[[LIGAND](#) | [KEGG](#)]

- 3 Click on **C000036** in the **Entry** column to display detailed information for the compound. See [Figure 125](#) on page 195

KEGG **COMPOUND: C00036** [Help](#)

Entry	C00036	Compound
Name	Oxaloacetate; Oxalacetic acid; Oxaloacetic acid; 2-Oxobutanedioic acid; Oxosuccinic acid; keto-Oxaloacetate	
Formula	C4H4O5	
Exact mass	132.0059	
Mol weight	132.0716	
Structure	 <p>C00036</p> <p>Mol file KCF file DB search Jmol KegDraw</p>	
Reaction	R00217 R00338 R00339 R00340 R00341 R00342 R00343 R00344 R00345 R00346 R00347 R00348 R00350 R00351 R00352 R00353 R00354 R00355 R00357 R00359 R00360 R00361 R00362 R00363 R00373 R00400 R00431 R00477 R00493 R00695 R00726 R00930 R00931 R01144 R01257 R01447 R01713 R01731 R03735 R05053 R05758 R07164 R07165 R09683	
Pathway	ko00010 Glycolysis / Gluconeogenesis ko00020 Citrate cycle (TCA cycle) ko00250 Alanine, aspartate and glutamate metabolism ko00362 Benzoate degradation ko00620 Pyruvate metabolism ko00630 Glyoxylate and dicarboxylate metabolism ko00680 Methane metabolism ko00710 Carbon fixation in photosynthetic organisms ko00720 Carbon fixation pathways in prokaryotes map01060 Biosynthesis of plant secondary metabolites map01061 Biosynthesis of phenylpropanoids map01062 Biosynthesis of terpenoids and steroids map01063 Biosynthesis of alkaloids derived from shikimate pathway map01064 Biosynthesis of alkaloids derived from ornithine, lysine and nicotinic acid map01065 Biosynthesis of alkaloids derived from histidine and purine map01066 Biosynthesis of alkaloids derived from tyrosine and	

All links

- Ontology (1)
- KEGG BRITE (1)
- Pathway (6068)
- KEGG PATHWAY (63)
- KEGG MODULE (6005)
- Chemical substance (7)
- PubChem (1)
- CHEBI (1)
- 3DMET (1)
- HMDB (1)
- KNAPSAcK (1)
- NIKKAJI (1)
- PDB-CCD (1)
- Chemical reaction (125)
- KEGG ENZYME (46)
- KEGG REACTION (44)
- KEGG RPAIR (35)
- Gene (18797)
- KEGG GENES (18797)
- All databases (24998)

[Download RDF](#)

Figure 125 Detailed information and links displayed in the KEGG Ligand Database

MetaCyc Database

- 1 To access this database enter <http://metacyc.org/> into your internet browser.
- 2 Enter the CAS number **328-42-7** into the **Quick Search** field and click **Quick Search**.

The screenshot shows the MetaCyc website interface. At the top left is the MetaCyc logo with the tagline "A member of the BioCyc database collection". To the right of the logo is a yellow box containing "Pathway Tools" and "Metabolism and". Further right is a search bar with the text "328-42-7" entered, and buttons for "Quick Search" and "Gene Search". Below the search bar is a green navigation bar with links for "Home", "Search", "Tools", and "Help".

The main content area is titled "METACYC OVERVIEW". It contains the following text:

MetaCyc is a database of nonredundant, experimentally elucidated metabolic pathways. MetaCyc contains more than 1928 pathways from more than 2263 different organisms [\[more\]](#), and is curated from the scientific experimental literature. [\[more\]](#)

MetaCyc contains pathways involved in both primary [\[def\]](#) and secondary [\[def\]](#) metabolism, as well as associated compounds, enzymes, and genes. [\[more\]](#)

Motivations

The goal of MetaCyc is to catalog the universe of metabolism by storing a representative sample of each experimentally elucidated pathway. [\[MetaCyc mission\]](#)

MetaCyc is used in a variety of scientific applications, such as providing a reference data set for computationally predicting the metabolic pathways of organisms from their sequenced genomes, supporting metabolic engineering, helping to compare biochemical networks, and serving as an encyclopedia of metabolism. [\[scientific applications\]](#)

Recent Publication

- [The MetaCyc Database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases](#), *Nucleic Acids Research* 40:D742-D753 2012.

On the left side of the page, there is a sidebar with the following sections:

- About MetaCyc**
 - Guide To MetaCyc
 - Querying Examples
 - Update History
 - Advisory Board
- Services**
 - Join MetaCyc Mailing List
 - Faster MetaCyc Search
 - Web Services

At the bottom of the sidebar, there are logos for SRI International, Carnegie, and BTI.

METACYC
A member of the BioCyc database collection

Pathway Tools Metabolism and Phenotypes Workshop
Mar 4th-6th 2013

LOGIN | Why Login? | Create New Account

Quick Search | Gene Search

Searching MetaCyc [change organism database](#)

Home | Search | Tools | Help | Compound

[Add to group](#)

MetaCyc Compound: oxaloacetate

Synonyms: oxaloacetate, keto-oxaloacetate, oxaloacetic acid, oxalacetic acid, oxalacetate

Chemical Formula: $C_4H_2O_5$

Molecular Weight: 130.06 Daltons

Monoisotopic Molecular Weight: 132.0058732389 Daltons

SMILES: C(C([O-])=O)C(=O)C([O-])=O

InChi: InChi=1S/C4H4O5/c5-2(4(8)9)1-3(6)7/h1H2,(H,6,7)(H,8,9)/p-2

Unification Links: CAS:328-42-7, [ChEBI:16452](#), [KEGG:C00036](#), [PubChem:164550](#)

Gibbs Energy of Formation (kcal/mol, estimated): -193.1 [\[Forsythe97\]](#)

Reactions known to consume the compound:

C4 photosynthetic carbon assimilation cycle, NADP ME biosynthesis

Figure 126 Results in the MetaCyc website

- 3 Scroll down to display detailed information on the compound. See [Figure 127](#) on page 198.

7 Using CAS and PubChem Numbers

Activator (Mechanism unknown) of: [isocitrate dehydrogenase phosphatase \[Nimmo84\]](#)

Enzymes inhibited by oxaloacetate, sorted by the type of inhibition, are:

Inhibitor (Competitive) of: [isocitrate dehydrogenase \[Comment 1\]](#) , [4-oxalocitramalate aldolase \[Hara03\]](#) , [acetylpyruvate hydrolase \[Davey75\]](#) , [4-oxalocitramalate aldolase \[Maruyama01\]](#) , [glutamate synthase \(NADH-dependent\) \[Boland77\]](#) , [isocitrate dehydrogenase \[Kanao02\]](#) , [4-hydroxy-2-ketoglutarate aldolase \[Anderson85\]](#) , [succinate dehydrogenase \[Burke82\]](#)

Inhibitor (Noncompetitive) of: [4-carboxy-4-hydroxy-2-oxoadipate aldolase \[Tack72\]](#) , [glutamate dehydrogenase \(NAD-dependent\) \[Bonete96, Comment 2\]](#)

Inhibitor (Allosteric) of: [malate dehydrogenase \[Sanwal69, Sanwal69a, Bologna07\]](#) , [citrate synthase](#)

Inhibitor (Mechanism unknown) of: [phosphoenolpyruvate synthetase \[Chulavatna73\]](#) , [malate dehydrogenase, NAD-requiring \[Yamaguchi73, Comment 3\]](#) , [fumarate reductase \[Iverson99\]](#) , [2-oxoglutarate decarboxylase \[Frank07\]](#) , [citrate lyase \[Helmward89\]](#) , [glutamine synthetase adenyltransferase \[Ebner70, Comment 4\]](#) , [isocitrate dehydrogenase kinase \[Nimmo84\]](#) , [4-oxalocrotonate tautomerase \[Orii06\]](#) , [glutamate dehydrogenase \[Kimura77\]](#) , [succinate dehydrogenase \[Dey97\]](#) , [ATP-citrate lyase \[Kanao01\]](#)

References

[Anderson85](#): Anderson M, Scholtz JM, Schuster SM (1985). "Rat liver 4-hydroxy-2-ketoglutarate aldolase: purification and kinetic characterization." *Arch Biochem Biophys* 236(1):82-97. PMID: 3966804

[Boland77](#): Boland MJ, Benny AG (1977). "Enzymes of nitrogen metabolism in legume nodules. Purification and properties of NADH-dependent glutamate synthase from lupin nodules." *Eur J Biochem* 79(2):355-62. PMID: 21790

[Bologna07](#): Bologna FP, Andreo CS, Drincovich MF (2007). "Escherichia coli malic enzymes: two isoforms with substantial differences in kinetic properties, metabolic regulation, and structure." *J Bacteriol* 189(16):5937-46. PMID: 17557829

[Bonete96](#): Bonete MJ, Perez-Pomares F, Ferrer J, Camacho ML (1996). "NAD-glutamate dehydrogenase from Halobacterium halobium: inhibition and activation by TCA intermediates and amino acids." *Biochim Biophys Acta* 1996;1289(1):14-24. PMID: 8605224

[Burke82](#): Burke JJ, Siedow JN, Moreland DE (1982). "Succinate Dehydrogenase : A Partial Purification from Mung Bean Hypocotyls and Soybean Cotyledons." *Plant Physiol* 70(6):1577-1581. PMID: 16662722

[Chulavatna73](#): Chulavatnatol M, Atkinson DE (1973). "Phosphoenolpyruvate synthetase from Escherichia coli. Effects of adenylate energy

Figure 127 MetaCyc also lists references to activation or inhibition of enzymes.

ChemSpider Database

- 1 To access this database enter <http://www.chemspider.com/> into your internet browser.
- 2 Click on **Search** under the ChemSpider logo to display the search box.
- 3 Enter the CAS number **328-42-7** into the search field.

The screenshot shows the ChemSpider website interface. At the top left is the ChemSpider logo with the tagline "The free chemical database". To the right is the RSC logo. A navigation bar contains links for "About", "More Searches", "Web APIs", and "Help". A search box contains the text "eg. Pyridine". Below this, there are tabs for "Simple search", "Structure search", and "Advanced search". The "Simple search" tab is active, and a search input field contains the CAS number "328-42-7". Below the input field is a table with the following columns and data:

Systematic names	Synonyms	Trade names	Registry numbers	SMILES	InChI
1,2-dihydroxybenzene	AIBN	Aspirin	7732-18-5	O=C(OCC) C	InChI=1/CH4h1H4 C

Below the table is a "Search" button. At the bottom of the page, there are four sections: "What is ChemSpider?", "Search by chemical names", "Search by chemical structure", and "Find important data". Each section contains a list of features or services. The "What is ChemSpider?" section describes the database's capabilities. The "Search by chemical names" section lists search criteria like systematic names, synonyms, trade names, and database identifiers. The "Search by chemical structure" section lists options like structure-based queries, drawing structures, and using structure files. The "Find important data" section lists literature references, physical properties, interactive spectra, and chemical suppliers. A "Generate Leads" button is visible in the bottom right corner.

Figure 128 CAS Search in ChemSpider.

- 4 Click the **Search** button below this field to display compounds found in the search. See [Figure 129](#) on page 200.

ChemSpider
The free chemical database

RSC | Syntheti

About | More Searches | Web APIs | Help | *eg. Pyridine*

Search term: 328-42-7 (Found by approved synonym) ?

Oxaloacetic acid

ChemSpider ID: **945**
 Molecular Formula: C₄H₄O₅
 Average mass: 132.072006 Da
 Monoisotopic mass: 132.005997 Da

▼ Systematic name
2-Oxosuccinic acid

► SMILES and InChIs
 ► Cite this record

2D 3D Save Zoom

► Names and Identifiers

► ChemSpider Searches

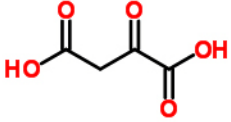
▼ Properties

Experimental data | Predicted - ACD/Labs | Predicted - EPISuite | Predicted - ChemAxon | Generate Leads

Figure 129 ChemSpider finds Oxaloacetic acid.

- 5 Click **Properties** or one of the other heading to see additional information. See [Figure 130](#) on page 201.

Oxaloacetic acid



ChemSpider ID: **945**
 Molecular Formula: $C_4H_4O_5$
 Average mass: 132.072006 Da
 Monoisotopic mass: 132.005997 Da

▼ Systematic name
 2-Oxosuccinic acid

► SMILES and InChIs
 ► Cite this record

2D 3D Save Zoom

► Names and Identifiers

► ChemSpider Searches

▼ Properties ? Print

Experimental data | Predicted - ACD/Labs | Predicted - EPI Suite | Predicted - ChemAxon

Predicted data is generated using the ACD/Labs' ACD/PhysChem Suite, for more information see their [website](#).

ACD/LogP:	-1.821	# of Rule of 5 Violations:	0
ACD/LogD (pH 5.5):	-6.53	ACD/LogD (pH 7.4):	-6.57
ACD/BCF (pH 5.5):	1.00	ACD/BCF (pH 7.4):	1.00
ACD/KOC (pH 5.5):	1.00	ACD/KOC (pH 7.4):	1.00
#H bond acceptors:	5	#H bond donors:	2
#Freely Rotating Bonds:	3	Polar Surface Area:	91.67 Å ²
Index of Refraction:	1.498	Molar Refractivity:	23.739 cm ³
Molar Volume:	80.934 cm ³	Polarizability:	9.411 10 ⁻²⁴ cm ³
Surface Tension:	77.2590026855469 dyne/cm	Density:	1.632 g/cm ³
Flash Point:	174.794 °C	Enthalpy of Vaporization:	64.358 kJ/mol
Boiling Point:	341.931 °C at 760 mmHg	Vapour Pressure:	0 mmHg at 25°C

Figure 130 ChemSpider finds Oxaloacetic acid.

SciFinder Substance Identifier Database

SciFinder provides a proprietary substance identifier database. The example below shows that a query using the same CAS number searched in our other examples returns approximately 5,054 references.

Note: You must have an account with Chemical Abstracts Service (CAS) in order to access the SciFinder Substance Identifier database. More information is available at CAS (<https://www.cas.org/products/scifinder>).

The figure consists of two screenshots of the SciFinder web interface. The top screenshot shows the search page for the substance identifier '328-42-7'. The search box contains the number '328-42-7' and a 'Search' button is highlighted with a blue arrow. The bottom screenshot shows the results page for '328-42-7'. The page displays the chemical structure of Butanedioic acid, 2-oxo- (C₄H₄O₅) and a list of references. A blue arrow points to the '~5054' reference count. The 'References' section is highlighted with a blue box, and a list of filters is visible on the right side of the page.

Figure 131 SciFinder returns a comprehensive list of literature reference information.

PubChem Compound Database

- 1 To access this database enter
<http://www.ncbi.nlm.nih.gov/sites/entrez?db=pccompound>
 into your Internet browser.
- 2 Enter the PubChem ID number **970** into the search **for** field and click **Go**.

One compound was found for this search.

- 3 Click the **3D Conformer** tab.

NCBI
PubChem
 Compound

PubChem Compound Search

Limits Advanced search Help

Oxaloacetic Acid - Compound Summary (CID 970)

Also known as: oxalacetic acid, Ketosuccinic acid, 2-Oxobutanedioic acid, Oxosuccinic acid, 2-Oxosuccinic acid, 2-Ketosuccinic acid, Oxobutanedioic acid

Molecular Formula: C₄H₄O₅ Molecular Weight: 132.07156 InChIKey: KHPXUQMNIQBQEV-UHFFFAOYSA-N

A dicarboxylic acid ketone that is an important metabolic intermediate of the CITRIC ACID CYCLE. It can be converted to ASPARTIC ACID by ASPARTATE TRANSAMINASE. From: MeSH

Table of Contents [Show subcontent titles](#)

- Identification
- Related Records
- Literature
- Patents
- Biomolecular Interactions and Pathways
- Biological Test Results
- Classification
- Chemical and Physical Properties

[Expand all sub-sections](#)

2D Structure 3D Conformer

Properties

Compound ID: 970
 Molecular Weight: 132.07156 [g/mol]
 Molecular Formula: C₄H₄O₅
 XLogP3-AA: -0.6
 H-Bond Donor: 2
 H-Bond Acceptor: 5

BioActivity Data Links

This Compound
[with Similar Compounds](#)
[with Similar Conformers](#)

Related Compounds

[Same Connectivity \(5\)](#)
[Same Any Tautomers \(8\)](#)
[Similar Compounds \(25\)](#)
[Similar Conformers \(1001\) View](#)

Related Substances

[All \(174\)](#)
[Same Structure \(76\)](#)
[Mixture \(98\)](#)

Identification

Figure 132 PubChem information on Oxaloacetic acid.

7 Using CAS and PubChem Numbers

- 4 In the Table of Contents, click the **Identification** line, and the page scrolls to the Identification section.

PubChem lists commercial availability of compounds under 'vendors' in the column on the right. It also shows chemical properties and out links to metabolic pathways and protein data.

The screenshot displays the 'Identification' section of a PubChem record. It is divided into several panels:

- Depositor-Supplied Synonyms:** A list of synonyms including oxalacetic acid, oxaloacetic acid, Ketosuccinic acid, 2-Oxobutanedioic acid, Oxosuccinic acid, 2-Oxosuccinic acid, 2-Ketosuccinic acid, Oxobutanedioic acid, oxaloacetate, and 328-42-7. A link to 'see all 69' is provided.
- Compound Information:** Shows CID 970, Create Date: 2004-09-16, and Descriptors including IUPAC Name (2-oxobutanedioic acid), InChI, InChIKey, and Canonical SMILES.
- Related Records:** A section for 'Related Compounds with Annotation'.
- Other Links:** A list of external links such as Protein Structure (20), PubMed (1011), Gene (70), Taxonomy (1), NLM Toxicology Link, and Chemical Structure Search.
- Chemical Vendors:** A list of vendors and their associated IDs, including ABI Chem, Achemica, AKos Consulting & Solutions, Ambinter, Angene Chemical, Anward, Biosynth, ChEMbase.cn, Chembo, and ChemFinder Chemical Directory.

- 5 Click the **Classification** line in the Table of Contents to access links to Substance Categorization Classification. Then, scroll to the Metabolic Pathways section. See Figure 133.

Comparative Toxicogenomics Database
SID 53787483 - External ID: D020351 ... see all 2

Database Vendor: 1 record
chemicalize.org by ChemAxon
SID 137002063 - External ID: 1588

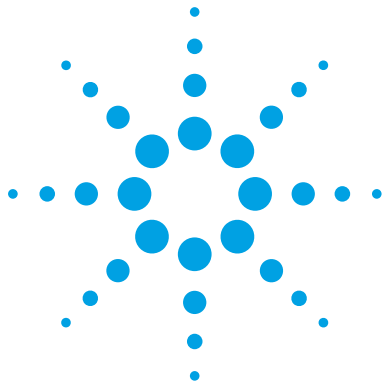
Journal Publishers: 3 records
Nature Chemical Biology
SID 56437667 - External ID: nchembio.145-comp30 ... see all 2
Thomson Pharma
SID 15297251 - External ID: 00057572

Metabolic Pathways: 5 records
BIND
SID 841755 - External ID: 1554
Comparative Toxicogenomics Database
SID 53787483 - External ID: D020351 ... see all 2
Golm Metabolome Database (GMD), Max Planck Institute of Molecular Plant Physiology
SID 92298111 - External ID: F15D7D2F-5E37-4B87-A8D2-824D5097A4DE
KEGG
SID 3338 - External ID: C00036

NIH Molecular Libraries: 1 record
NCGC
SID 124955195 - External ID: NCGC00248960-01

Patents: 3 records
IBM
SID 128383593 - External ID: A176CBC7DE331FD9CDA592160699DF73
SCRIPDB
SID 142338819 - External ID: US06197558-20010306-C00002_3
Thomson Pharma
SID 15297251 - External ID: 00057572

Figure 133 Five links to Metabolic pathways are given by a PubChem Compound search.



8 Reference Information

Definitions	208
Derivatization Instructions	220
References	221



Definitions

This section contains a list of terms and their definitions as used in this workflow. Review of the terms and definitions presented in this section helps you understand the Agilent software wizards and the metabolomics workflow.

- Alignment** Adjustment of the chromatographic retention time of eluting components to improve the correlation among data sets, based on the elution of specific component(s) that are (1) naturally present in each sample or (2) deliberately added to the sample through spiking the sample with a known compound or set of compounds that does not interfere with the sample.
- AMDIS** Acronym for automated mass spectral deconvolution and identification system developed by NIST (<http://www.amdis.net>).
- Amino acid** Biologically significant molecules that contain a core carbon positioned between a carboxyl and amine group in addition to an organic substituent. Dual carboxyl and amine functionalities facilitate the formation of peptides and proteins.
- ANOVA** Abbreviation for analysis of variance which is a statistical method that simultaneously compares the means between two or more attributes or parameters of a data set. ANOVA is used to determine if a statistical difference exists between the means of two or more data sets and thereby prove or disprove the hypothesis. See also t-Test.
- Attribute** Another term for an independent variable. Referred to as a parameter and is assigned a parameter name during the various steps of the metabolomic data analysis.
- Attribute value** Another term for one of several values within an attribute for which exist correlating samples. Referred to as a condition or a parameter value and given an assigned value during the various steps of the metabolomic data analysis.
- Baselining** A technique used to view and compare data that involves converting the original data values to values that are expressed as changes relative to a calculated statistical value derived from the data. The calculated statistical value is referred to as the baseline.
- Bayesian** A term used to refer to statistical techniques named after the Reverend Thomas Bayes (ca. 1702 - 1761).

Bayesian inference	The use of statistical reasoning, instead of direct facts, to calculate the probability that a hypothesis may be true. Also known as Bayesian statistics.
Bioinformatics	The use of computers, statistics, and informational techniques to increase the understanding of biological processes.
Biomarker	An organic molecule whose presence and concentration in a biological sample indicates a normal or altered function of higher level biological activity.
Carbohydrate	An organic molecule consisting entirely of carbon, hydrogen, and oxygen that is important to living organisms.
CEF file	A binary file format called a compound exchange file (CEF) that is used to exchange data between Agilent software. In the metabolomics workflow CEF files are used to share molecular features between MassHunter Qualitative Analysis and Mass Profiler Professional.
Cell	The fundamental unit of an organism consisting of several sets of biochemical functions within an enclosing membrane. Animals and plants are made of one or more cells that combine to form tissues and perform living functions.
Census	Collection of a sample from every member of a population.
Cheminformatics	The use of computers and informational techniques (such as analysis, classification, manipulation, storage, and retrieval) to analyze and solve problems in the field of chemistry.
Chemometrics	A science employing mathematical and analytical processes to extract information from chemical data sets. The processes involve interactive applications of techniques employed in disciplines such as multivariate statistics, applied mathematics, and computer science to obtain meaningful information from complex data sets. Chemometrics is typically used to obtain meaningful information from data derived from chemistry, biochemistry and chemical engineering. Agilent Mass Profiler Professional is designed to employ chemometrics processes to GC/MS and LC/MS data sets to obtain useful information.
Child	A subset of information that is created by an algorithm from an original set of information. An entity list created using Mass Profiler Professional is a child. An original entity list is referred to as the parent of one or more child entity lists.

Co-elution	When compounds elute from a chromatographic column at nominally the same time making the assignment of the observed ions to each compound difficult.
Complex	Class of compounds consisting of more than one protein physically which physically bind each other and are biologically active and stable in their combined form.
Composite spectrum	A compound spectrum generated to represent the molecular feature that includes more than one ion, isotope, or adduct (not just M + H) and is used by Mass Profiler Professional for recursive analysis and ID Browser.
Compound	A metabolite that may be individually referred to as a compound, molecular feature, element, or entity during the various steps of the metabolomic data analysis.
Condition	Another term for one of several values within a parameter for which exist correlating samples. Condition may also be referred to as a parameter value during the various steps of the metabolomic data analysis. See also attribute value.
Data	Information in a form suitable for storing and processing by a computer that represent the qualitative or quantitative attributes of a subject. Examples include GC/MS and LC/MS data consisting fundamentally of time, ion m/z, and ion abundance from a chemical sample.
Data processing	Conversion of data into meaningful information. Computers are employed to enable rapid recording and handling of large amounts of data, i.e. Agilent MassHunter Workstation and Agilent Mass Profiler Professional.
Data reduction	See reduction.
Deconvolution	The technique of reconstructing individual mass and mass spectral data from co-eluting compounds.
Dependent variable	An element in a data set that can only be observed as a result of the influence from the variation of an independent variable. For example, a pharmaceutical compound structure and quantity may be controlled as two independent variables while the metabolite profile presents a host of small-molecule products that make up the dependent variables of a study.
Determinate	Having exact and definite limits on an analytical result that provide a conclusive degree of correlation of the subject to the specimen.

Element	A metabolite that may be individually referred to as a compound, molecular feature, element, or entity during the various steps of the metabolomic data analysis.
Endogenous	Pertaining to cause, development, or origination from within an organism.
Entity	A metabolite that may be individually referred to as a compound, molecular feature, element, or entity during the various steps of the metabolomic data analysis.
Entity List	The compounds that meet the requirements specified by each experiment performed on your data. Entity lists are viewed in the Experiment Navigator.
Enzyme	Proteins acting as biocatalysts in a metabolomic reaction. These entities are particularly important in depicting a biochemical network.
Experiment	Data acquired in an attempt to understand causality where tests or analyses are defined and performed on an organism to discover something that is not yet known, to demonstrate as proof of something that is known, or to find out whether something is effective.
Externality	A quality, attribute, or state that originates and/or is established independently from the specimen under evaluation.
Extraction	The process of retrieving a deliberate subset of data from a larger data set whereby the subset of the data preserves the meaningful information as opposed to the redundant and less meaningful information. Also known as data extraction.
Family	A group of proteins related by structure, function, or another biological parameter.
Feature	Independent, distinct characteristic of a phenomena and data under observation. Features are an important part of the identification of patterns - pattern recognition - within data whether processed by a human or by artificial intelligence, such as Agilent MassHunter Workstation and Agilent Mass Profiler Professional. In metabolomics analysis a feature is a metabolite and may be individually referred to as a compound, molecular feature, element, or entity during the various steps of the metabolomic data analysis.
Feature extraction	The reduction of data size and complexity through the removal of redundant and non-specific data by using the important variables (features) associated with the data. Careful feature extraction yields a

smaller data set that is more easily processed without any compromise in the information quality. This is part of the principal component analysis process employed by Agilent Mass Profiler Professional.

Feature selection	The identification of important, or non-important, variables and the variable relationships in a data set using both analytical and a priori knowledge about the data. This is part of the principal component analysis process employed by Agilent Mass Profiler Professional.
Filter	The process of establishing criteria by which entities are removed (filtered) from further analysis during the metabolomics workflow.
Filter by flag	A flag is a term used to denote a quality of an entity within a sample. A flag indicates if the entity was detected in each sample as follows: Present means the entity was detected, Absent means the entity was not detected, and Marginal means the signal for the entity was saturated
Function	A classification of compounds based on their biological purpose or activity.
Hypothesis	A proposition made to explain certain facts and tentatively accepted to provide a basis for further investigation. A proposed explanation for observable phenomena may or may not be supported by the analytical data. Statistical data analysis is performed to quantify the probability that the hypothesis is true. Also known as the scientific hypothesis.
Hypothetical	A statement based on, involving, or having the nature of a hypothesis for the purposes of serving as an example and not necessarily based on an actuality.
ID Browser	Agilent software that automatically annotates the entity list with the compound names and adds them to any of the various visualization and pathway analysis tools.
Identified compound	Chromatographic components that have an assigned, exact identity, such as compound name and molecular formula, based on prior assessment or comparison with a database. See also Unidentified Compound.
Independent variable	An essential element, constituent, attribute, or quality in a data set that is deliberately controlled in an experiment. For example, a pharmaceutical compound structure and quantity may be controlled as two independent variables while the metabolite profile presents a host of independent small molecule products that make up the dependent variables of a study. An

independent variable may be referred to as a parameter and is assigned a parameter name during the various steps of the metabolomic data analysis.

Inorganic compound	Non carbon and non biological origin compounds such as minerals and salts.
Interpretation	Expression of your data in entity lists after grouping your samples, applying filters, and performing statistical correlation methods. When you open an experiment, the “All Samples” interpretation is active. You can click on another interpretation to activate it.
Lipidomics	Identification and quantification of cellular lipids from an organism in a specified biological situation. The study of lipids is a subset of metabolomics.
Mass variation	Using the mass to charge (m/z) resolution to improve compound identification. Compounds with nearly identical and identical chromatographic behavior are deconvoluted by adjusting the m/z range for extracting ion chromatograms.
Mean	The numerical result of dividing the sum of the data values by the number of individual data observations.
Metabolism	The chemical reactions and physical processes whereby living organisms convert ingested compounds into other compounds, structures, energy and waste.
Metabolite	Small organic molecules that are intermediate compounds and products produced as part of metabolism. Metabolites are important modulators, substrates, byproducts, and building blocks of many different biological processes. In order to distinguish metabolites from larger biological molecules, known as macromolecules such as proteins, DNA and others, metabolites are typically under 1000 Da. A metabolite may be individually referred to as a compound, molecular feature, element, or entity during the various steps of the metabolomic data analysis.
Metabolome	The complete set of small-molecule metabolites that may be found within a biological sample. Small molecules are typically in the range of 50 to 600 Da.

- Metabolomics** The process of identification and quantification of all metabolites of an organism in a specified biological situation. The study of the metabolites of an organism presents a chemical “fingerprint” of the organism under the specific situation. See metabonomics for the study of the change in the metabolites in response to externalities.
- Metabonomics** The metabolic response to externalities such as drugs, environmental factors, and disease. The study of metabonomics by the medical community may lead to more efficient drug discovery and to individualized patient treatment. Meaningful information learned from the metabolite response can be used for clinical diagnostics or for understanding the onset and progression of human diseases. See metabolomic for the identification and quantitation of metabolites.
- NLP** Natural Language Processing (NLP) algorithm that extracts information from published literature.
- Normalization** A technique used to adjust the ion intensity of mass spectral data from an absolute value based on the signal measured at the detector to a relative intensity of 0 to 100 percent based on the signal of either (1) the ion of the greatest intensity or (2) a specific ion in the mass spectrum.
- Null hypothesis** The default position taken by the hypothesis that no effect or correlation of the independent variables exists with respect to the measurements taken from the samples.
- Observation** Data acquired in an attempt to understand causality where no ability exists to (1) control how subjects are sampled and/or (2) control the exposure each sample group receives.
- One-hit wonder** An entity that appears in only one sample, is absent from the replicate samples, and does not provide any utility for statistical analysis. Entities that are one-hit wonders may be filtered using Filter by Flags.
- Organic compound** Carbon-based compounds, often with biological origin.
- Organism** A group of biochemical systems that function together as a whole thereby creating an individual living entity such as an animal, plant, or microorganism. Individual living entities may be multicellular or unicellular. See also specimen.
- p-value** The probability of obtaining a statistical result that is comparable to or greater in magnitude than the result that was actually observed, assuming that the null hypothesis is true. The null hypothesis is stated that no

correlation exists between the independent variables and the measurements taken from the samples. Rejection of the null hypothesis is typically made when the p-value is less than 0.05 or 0.01. A p-value of 0.05 or 0.01 may be restated as a 5% or 1% chance of rejecting the null hypothesis when it is true. When the null hypothesis is rejected, the result is said to be statistically significant meaning that a correlation exists between the independent variables and the measurements as specified in the hypothesis.

- Parameter** Another term for an independent variable. Referred to as a parameter or parameter name and is assigned a parameter name during the various steps of the metabolomic data analysis. See also condition and attribute.
- Parameter value** Another term for one of several values within a parameter for which exist correlating samples. Parameter value may also be referred to as a condition during the various steps of the metabolomic data analysis. See also attribute value.
- Parent** The original set of information that is processed by an algorithm to create one or more subsets of information. A subset entity list is referred to as the child of a parent entity list.
- Peptide** Linear chain of amino acids that is shorter than a protein. The length of a peptide is sufficiently short that it is easily made synthetically from the constituent amino acids.
- Peptide bond** The covalent bond formed by the reaction of a carboxyl group with an amine group between two molecules, e.g. between amino acids.
- Permutation** Any of the total number of subsets that may be formed by the combination of individual parameters among the independent variables. For example the number of permutations of A and B in variable ? in combination with X, Y, and Z in variable ? equals six ($6 = 2 \times 3$) and may be represented as AX, AY, AZ, BX, BY, and BZ. Note that the combinations of parameters within a variable are not relevant such as AB, XY, XZ, and YZ.
- Polarity** The condition of an effect as being positive or negative, additive or subtractive, with respect to some point of reference, such as with respect to the concentration of a metabolite.
- Polymer** A molecule formed by the covalent bonding of a repeating molecular group to form a larger molecule.

- Pooled sample** When the amount of available biological material is very small samples may be combined into a single sample (pooled) and then split into different aliquots for multiple analyses. By pooling the sample, sufficient material exists to obtain replicate analyses of each sample where formerly there was insufficient material to obtain replicate analytical results. The trade-off loss of information about the biological variation that was formerly present in each unique sample is offset by a gain in statistical significance of the results.
- Principal component** Transformed data into axes, or principal components, so that the patterns between the axes most closely describe the relationships between the data. The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible. The principal components often may be viewed, and interpreted, most readily in graphical axes with additional dimensions represented by color and/or shape representing the key elements (independent variables) of the hypothesis. This is part of the principal component analysis process employed by Agilent Mass Profiler Professional.
- Principal component analysis** The mathematical process by which data containing a number of potentially correlated variables is transformed into a data set in relation to a smaller number of variables called principal components which account for the most variability in the data. The result of the data transformation leads to the identification of the best explanation of the variance in the data, e.g. identification of the meaningful information. Also known as PCA.
- Process** Classification of biological processes.
- Protein** Linear chain of amino acids whose amino acid order and three-dimensional structure are essential to living organisms. Also known as a polypeptide.
- Proteomics** The study of the structure and function of proteins occurring in living organisms. Proteins are assemblies of amino acids (polypeptides) based on information encoded in the genes of an organism and are the main components of the physiological metabolic pathways of the organism.
- Quality** A feature, attribute, and/or characteristic element whose presence, absence, or inability to be properly ascertained due to instrumental factors, is factored into whether a sample is or is not representative of the larger specimen.

Recursive Reapplying the same algorithm to a subset of a previous result in order to generate an improved result.

Recursive finding A three-step process in the metabolomics workflow that improves the accuracy of finding statistically significant features in sample data files. Step 1: Find untargeted compounds by molecular feature in MassHunter Qualitative Analysis. Step 2: Filter the molecular features in Mass Profiler Professional. Step 3: Find targeted compounds by formula in MassHunter Qualitative Analysis. Importing the most significant features identified using Mass Profiler Professional back into MassHunter Qualitative Analysis as targeted features improves the accuracy in finding these features from the original sample data files.

Reduction The process whereby the number of variables in a data set is decreased to improve computation time and information quality. For example, an extracted ion chromatogram obtained from GC/MS and LC/MS data files. Reduction provides smaller, viewable and interpretable data sets by employing feature selection and feature extraction. Also known as dimension reduction and data reduction. This is part of the principal component analysis process employed by Agilent Mass Profiler Professional.

Regression analysis Mathematical techniques for analyzing data to identify the relationship between dependent and independent variables present in the data. Information is gained from the estimation, regression, or the sign and proportionality of the effects of the independent variables on the dependent variables. This is part of the principal component analysis process employed by Agilent Mass Profiler Professional. Also known as regression.

Replicate Collecting multiple identical samples from a population so that when the samples are evaluated a value is obtained that more closely approximates the true value.

Sample A part, piece, or item that is taken from a specimen and understood as being representative of the larger specimen (e.g., blood sample, cell culture, body fluid, aliquot) or population. An analysis may be derived from samples taken at a particular geographical location, taken at a specific period of time during an experiment, or taken before or after a specific treatment. A small number of specimens used to represent a whole class or group.

- Sample class prediction** A workflow used to build a model and classify samples from mass spectrometry data. Class prediction is a supervised learning method and involves three steps: validation, training, and prediction. The algorithm learns from samples (training set) with known functional class and builds a prediction model to classify new samples (test set) of unknown class.
- Specimen** An individual organism, e.g., a person, animal, plant, or other organism, of a class or group that is used as a representative of a whole class or group.
- Spike** The specific and quantitative addition of one or more compounds to a sample.
- Standard** A chemical or mixture of chemicals selected for use as a basis of comparing the quality of analytical results or for use to measure and compensate the precise offset or drift incurred over a set of analyses.
- Standard deviation** A measure of variability among a set of data that is equal to the square root of the arithmetic average of the squares of the deviations from the mean. A low standard deviation value indicates that the individual data tend to be very close to the mean, whereas a high standard deviation indicates that the data is spread out over a larger range of values from the mean.
- State** A set of circumstances or attributes characterizing a biological organism at a given time. A few sample attributes may include temperature, time, pH, nutrition, geography, stress, disease, and controlled exposure.
- Statistics** The mathematical process employed in manipulating numerical data from scientific experiments to derive meaningful information. This is part of the principal component analysis process employed by Agilent Mass Profiler Professional.
- Subject** A chemical or biological sample taken from a specimen, or a whole specimen, that undergoes a treatment, experiment, or an analysis for the purposes of further understanding.
- Survey** Collection of samples from less than the entire population in order to estimate the population attributes.
- t-Test** A statistical test to determine whether the mean of the data differs significantly from that expected if the samples followed a normal distribution in the population. The test may also be used to assess statistical significance between the means of two normally distributed data sets. See also ANOVA.

- Unidentified compound** Chromatographic components that are only uniquely denoted by their mass and retention times and which have not been assigned an exact identity, such as compound name and molecular formula. Unidentified compounds are typically produced by feature finding and deconvolution algorithms. See also Identified Compound.
- Variable** An element in a data set that assumes changing values, e.g. values that are not constant over the entire data set. The two types of variables are independent and dependent.
- Volume** The area of the extracted compound chromatogram (ECC). The ECC is formed from the sum of the individual ion abundances within the compound spectrum at each retention time in the specified time window. The compound volume generated by MFE is used by Mass Profiler Professional to make quantitative comparisons.
- Wizard** A sequence of dialog boxes presented by Mass Profiler Professional that guides you through well-defined steps to enter information, organize data, and perform analyses.

Derivatization Instructions

Derivatization of d_{27} -myristic acid is required before using this material for GC/MS retention time locking.

- In a V-bottom GC vial (e.g. Agilent p/n 5184-3554, with screw cap (Agilent p/n 5182-0723), assemble the derivatization reaction detailed below in a fume hood.
 - Use clean graduated glass syringes (for example, Agilent p/n 5182-9615) to transfer each component to the vial.
 - During each addition step, do not allow the syringe to make contact with any material already in the vial.
- 1** Transfer 10 μ l of pyridine to the GC vial, and cap the vial to prevent evaporation.
 - 2** Add 10 μ l of the FAMES/ d_{27} -Myristic Acid Mixture to the vial and cap the vial.
 - 3** Add 80 μ l of MSTFA/1% TMCS to the vial and cap the vial.
 - 4** Incubate the vial at 37°C for 30 minutes.

Upon completion of these steps, the FAMES/ d_{27} -Myristic Acid Mixture is ready for use in GC/MS. Derivatized samples should be analyzed within 24 hours.

NOTE

One single-use vial each of pyridine and MSTFA/1% TMCS is supplied with the kit and may be used in a single derivatization procedure. Thereafter, use the GC/MS-grade pyridine and MSTFA/1% TMCS reagents in routine use in your laboratory

References

This section consists of citations to Agilent manuals, primers, application notes, presentations, product brochures, technical overviews, training videos, and software that help you use Agilent products and perform your metabolomics analyses.

Manuals

Agilent MassHunter Workstation Software Qualitative Analysis Familiarization Guide for GC/MS

Agilent MassHunter Workstation Software Quantitative Analysis - Familiarization Guide

Agilent G3835AA MassHunter Mass Profiler Professional - Quick Start Guide (Agilent publication, [G3835-90009](#), Revision A, November 2012)

Agilent G3835AA MassHunter Mass Profiler Professional - Familiarization Guide (Agilent publication, [G3835-90010](#), Revision A, November 2012)

Agilent G3835AA MassHunter Mass Profiler Professional - Application Guide (Agilent publication, [G3835-90011](#), Revision A, November 2012)

Agilent Metabolomics Workflow - Discovery Workflow Guide (Agilent publication [5990-7067EN](#), Revision B, October 2012)

Agilent Metabolomics Workflow - Discovery Workflow Overview (Agilent publication [5990-7069EN](#), Revision B, October 2012)

Agilent Mass Profiler Professional - (Agilent publication, January 2012)

Primers

Proteomics: Biomarker Discovery and Validation (Agilent publication [5990-5357EN](#), February 11, 2010)

Metabolomics: Approaches Using Mass Spectrometry (Agilent publication [5990-4314EN](#), October 27, 2009)

Application Notes

Multi-omic Analysis with Agilent's GeneSpring 11.5 Analysis Platform (Agilent publication [5990-7505EN](#), March 25, 2011)

An LC/MS Metabolomics Discovery Workflow for Malaria-Infected Red Blood Cells Using Mass Profiler Professional Software and LC-Triple Quadrupole MRM Confirmation (Agilent publication [5990-6790EN](#), November 19, 2010)

Profiling Approach for Biomarker Discovery using an Agilent HPLC-Chip Coupled with an Accurate-Mass Q-TOF LC/MS (Agilent publication [5990-4404EN](#), October 20, 2009)

Metabolite Identification in Blood Plasma Using GC/MS and the Agilent Fiehn GC/MS Metabolomics RTL Library (Agilent publication [5990-3638EN](#), April 1, 2009)

Metabolomic Profiling of Bacterial Leaf Blight in Rice (Agilent publication [5989-6234EN](#), February 14, 2007)

Presentations

Advances in Instrumentation and Software for Metabolomics Research (Agilent publication n/a, September 18, 2012)

Multi-omics Analysis Software for Targeted Identification of Key Biological Pathways (Agilent publication n/a, May 3, 2012)

Metabolomics LCMS Approach to: Identifying Red Wines according to their variety and Investigating Malaria infected red blood cells (Agilent publication n/a, November 3, 2010)

Small Molecule Metabolomics (Agilent publication n/a, November 3, 2010)

Presentation: Metabolome Analysis from Sample Prep through Data Analysis (Agilent publication n/a, November 3, 2010)

Product Brochures

Emerging Insights: Agilent Solutions for Metabolomics (Agilent publication [5990-6048EN](#), April 30, 2012)

Agilent Mass Profiler Professional Software - Discover the Difference in your Data (Agilent publication [5990-4164EN](#), April 27, 2012)

Pathways to Insight - Integrated Biology at Agilent (Agilent publication [5991-0222EN](#), March 30, 2012)

Confidently Better Bioinformatics Solutions (Agilent publication [5990-9905EN](#), February 2, 2012)

Integrated Biology from Agilent: The Future is Emerging (Agilent publication [5990-6047EN](#), September 1, 2010)

Agilent Fiehn GC/MS Metabolomics RTL Library (Agilent publication [5989-8310EN](#), December 5, 2008)

Agilent METLIN Personal Metabolite Database (Agilent publication [5989-7712EN](#), December 31, 2007)

Agilent Metabolomics Laboratory: The breadth of tools you need for successful metabolomics research (Agilent publication 5989-5472EN, January 31, 2007)

**BioCyc Pathway/
Genome
Databases**

Includes BioCyc Pathway/Genome databases from the Bioinformatics Research Group at SRI International®, used under license.



<http://www.biocyc.org/>

Citation based on use of BioCyc

Users who publish research results in scientific journals based on use of data from the EcoCyc Pathway/Genome database should cite:

Keseler et al, Nucleic Acids Research 39:D583-90 2011.

Users who publish research results in scientific journals based on use of data from most other BioCyc Pathway/Genome databases should cite:

Caspi et al, Nucleic Acids Research 40:D742-53 2012.

In some cases, BioCyc Pathway/Genome databases are described by other specific publications that can be found by selecting the database and then going to the Summary Statistics pages under the Tools menu. The resulting page sometimes contains a citation for that database.

www.agilent.com

© Agilent Technologies, Inc. 2013

First Edition, July 2013



G1676-90001



Agilent Technologies