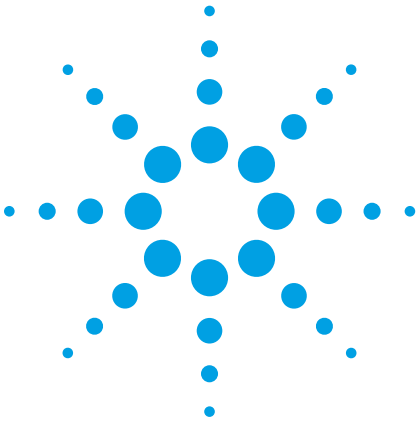


Adding Retention Times to the METLIN Personal Metabolite Database to Improve Compound Identification

Authors

Steven Fischer and Theodore Sana
Agilent Technologies

A key step in metabolomics research is the identification of metabolites. Searching accurate-mass LC/MS data against metabolite databases is a powerful technique for identifying metabolites. Due to the presence of isobaric masses such as compound isomers and diastereomers, database searching by mass alone is generally insufficient for positive identification. Public databases do not contain retention time information and thus there is no way to resolve the issue of isobaric masses. However, with Agilent Technologies' METLIN Personal metabolite database, empirically derived, method-specific retention times (RTs) can be added to compounds in the database, creating an accurate-mass retention time (AMRT) database. Retention times then provide a second match criterion orthogonal to, and independent of, compound mass. Searching an AMRT database provides a narrower, more focused list of possible metabolite identities. This technical note discusses the addition of retention times to the METLIN Personal metabolite database and provides a faster batch approach for creating an AMRT database.

Adding retention times to create an AMRT database

Retention times can be added to the METLIN Personal metabolite database in a manual process of running metabolite standards, selecting the proper database entry for each standard, and updating the retention time in the database. This approach is effective, but can be slow. A faster, easier approach uses the Agilent MassHunter Mass Profiler software to preprocess the data file from the analysis of metabolite standards before database searching. This approach filters out non-standards and simplifies the assignment of standard matches and addition of retention times in the METLIN software. The workflow is outlined in Figure 1.

Analyzing the metabolite standards

The first step in the workflow is to acquire accurate-mass data from metabolite standards. As many standards as is convenient may be analyzed in a single sample; however, each standard must have a unique mass range around it that is free of interfering masses. If two or more metabolite standards have masses within 5 mDa (or 50 ppm, whichever is larger) of each other, they should be analyzed in separate samples.

For the retention times to be useful, the metabolite standards and experiment samples must be analyzed using the exact same chromatographic method. Ideally, a "universal" chromatographic method that produces good results for both positive and negative ion polarities and for both electrospray and APCI should be used. This makes the resulting AMRT database useful for a wider range of compounds. Agilent recommends a water/methanol mobile phase with approximately 5 mM ammonium acetate and 0.1% acetic acid.

Chromatographic methods optimized for a particular ionization polarity, ionization technique, or analyte can be used, but the resulting retention times will only be meaningful for compounds analyzed with that specialized method.

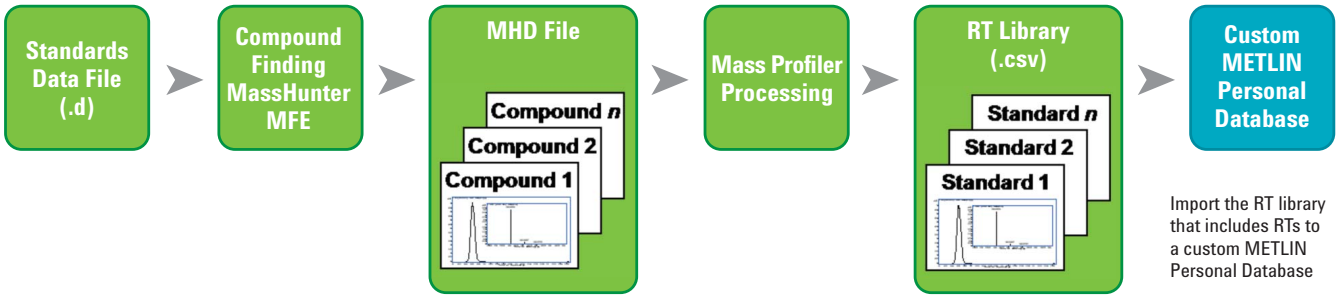


Figure 1. Workflow for batch addition or update of retention times in the METLIN Personal metabolite database.

Once the metabolite standard sample(s) have been analyzed, process the data using the MassHunter Qualitative Analysis software. Use the Molecular Feature Extraction (MFE) algorithm to locate the compounds and save the resulting data file.

Processing the standards data with the MassHunter Mass Profiler software

The data acquired by analyzing the metabolite standards are processed in the Mass Profiler software to generate a Feature Summary table of the standards and retention times. This table is uploaded to the METLIN Personal metabolite database and is used to add or update retention times for the metabolite entries.

1. Create a new project in the Mass Profiler software. Specify one group in the project and add the data file(s) (.mhd) created from the metabolite standards (see Figure 2).
2. Open the data file from the standards analysis and select Method → Edit Method on the menu bar.
3. In the Pre-Analysis Filters tab, choose the Special Masses option Limit to These (Figure 3). Enter or paste the exact neutral masses of the metabolite standards and click OK. Entering the exact neutral masses here allows the Mass Profiler software to filter out all non-standards (for example, chemical noise or impurities) that may have been detected in the metabolite standards sample. This significantly simplifies subsequent METLIN database searching and retention time assignment or updating.
4. Save the method so that it can be applied to the same type of standard mix in the future.
5. Export the feature table as a comma separated value file (.csv).

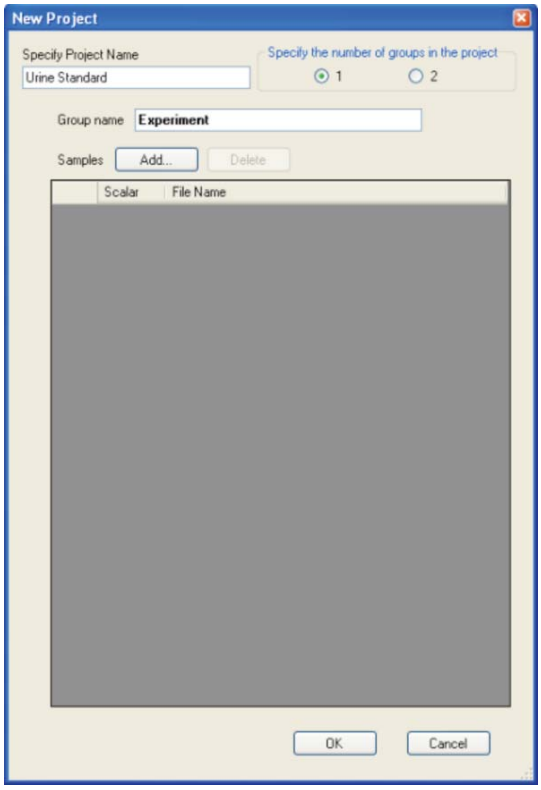


Figure 2. To begin processing the standards data in the Mass Profiler software, create a new project that includes the standards data (.mhd).

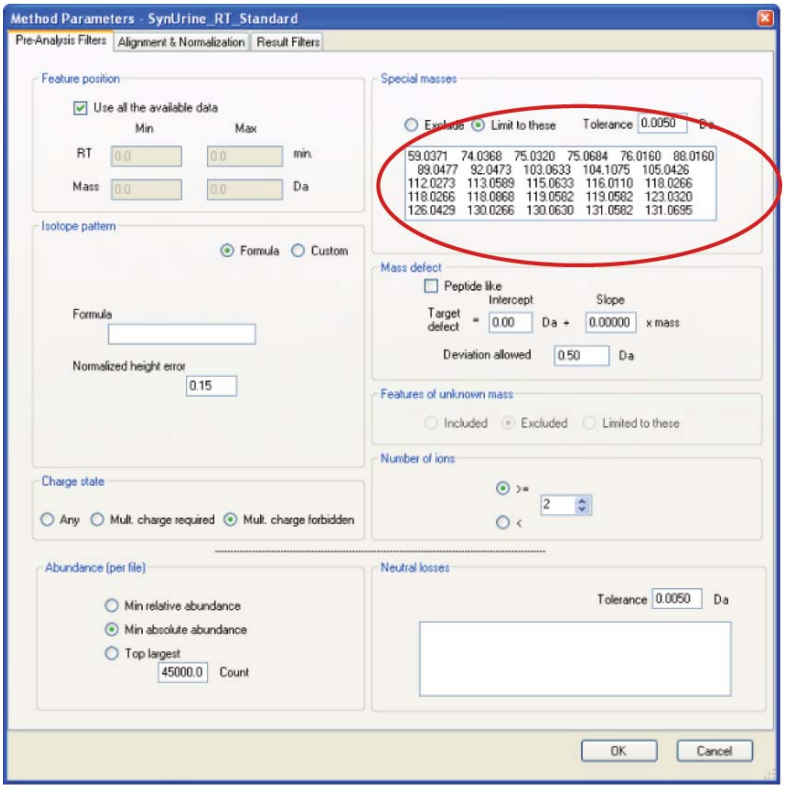


Figure 3. Filter out non-standards data from the analysis by specifying the exact neutral masses of the metabolite standards.

Creating an accurate-mass retention time (AMRT) METLIN database using an imported mass-RT list

After processing the metabolite standards data in the MassHunter Mass Profiler data, use this procedure to create a copy of the METLIN Personal database and add retention times for the metabolite standards to the database.

1. In the METLIN software, select File → New Database. The original METLIN Personal metabolite database is write-protected so a "new" database must be created before retention times can be added.
2. Select an existing database from which to begin (Figure 4). Enter a new name and a description, and click Create.
3. In the Batch Search tab (Figure 5), choose Optional for Retention times .
4. Click File, and in the dialog box that appears, select the Summary Feature table (.csv) for the metabolite standards that was exported from the Mass Profiler software.
5. Click Find Metabolites on the toolbar.

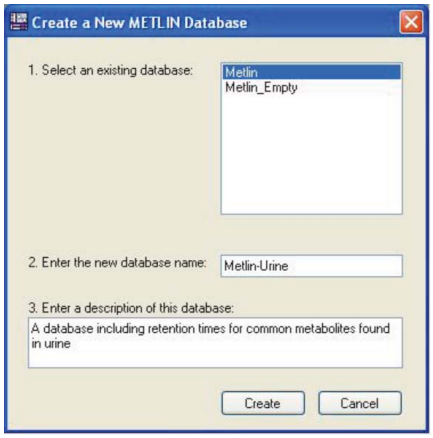


Figure 4. Create a new METLIN Personal metabolite database from an existing METLIN database.

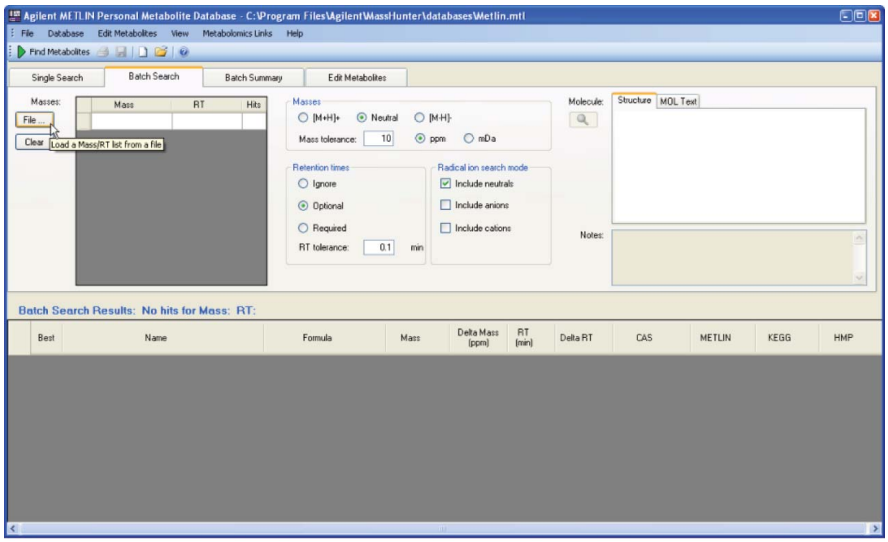


Figure 5. Upload the standards Feature Summary table (.csv) to the new METLIN database.

6. One standard at a time, highlight the standard mass in the Batch Search list and select the correct metabolite identity in the Batch Search Results (Figure 6). In almost all cases, there will be more than one match due to isobaric compounds in the database, but since the identity of each analyzed standard is already known, it should be easy to select the correct metabolite from the search results.

   Note: If a metabolite standard is not present in the database, add the standard in the Edit Metabolites tab. In the Edit Metabolites tab, click the Add New button, enter the appropriate metabolite information, and click Save As New. Return to the Batch Search tab and select the newly added metabolite as the identity of the standard.

7. Once all of the metabolite standards are assigned identities from the database, review the results in the Batch Summary tab (Figure 7). When all of the assignments are correct, you are ready to update the database.

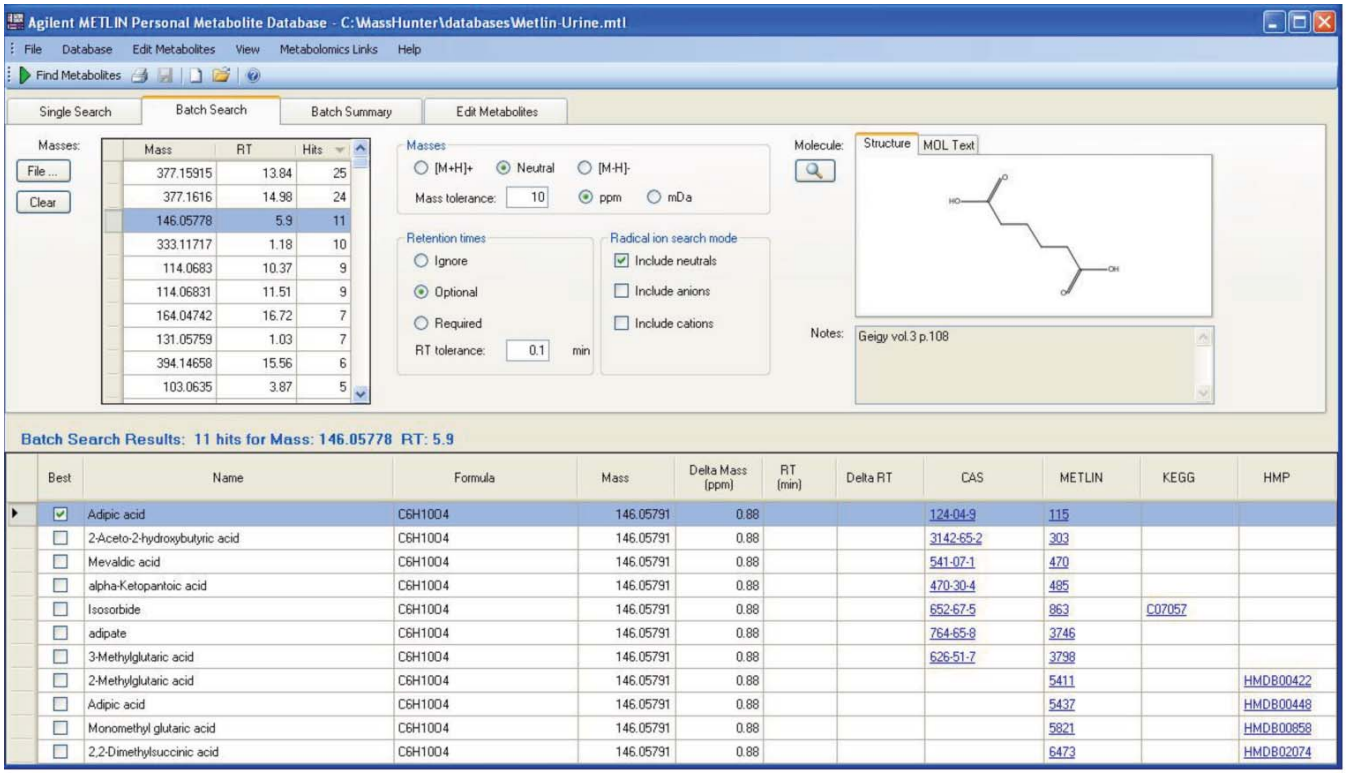


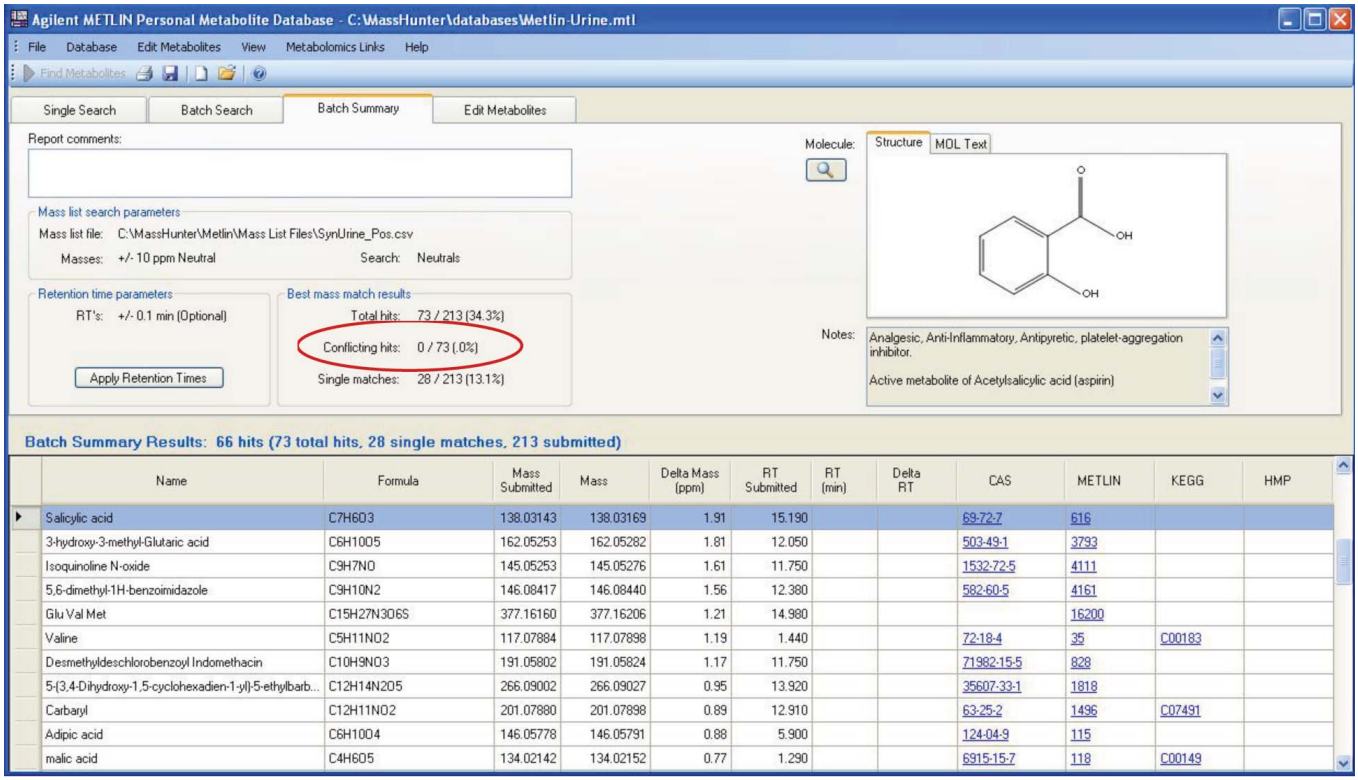
Figure 6. Confirm that each standard mass in the batch search list is correctly identified in the batch search results. If necessary, select the correct identity for a standard in the batch search results.

8. The “Apply Retention Time” tab will become active once the number of conflicting hits is set to zero. Once active, click Apply Retention Times in the Batch Summary tab. The database is updated to include the retention times (Figure 8) and is automatically saved.

Searching a custom METLIN personal database using accurate masses and retention times

After an AMRT METLIN database is created, it is easy to incorporate retention times into subsequent searches. When specifying the search parameters in the Batch Search tab, simply select Optional or Required under Retention times.

Choosing Optional ranks mass matches higher that also match RT; Required returns only mass matches that also have RT matches. As a general rule, choose Optional for general (non-targeted) searches because not all compounds in the database will have retention times. For targeted searches, Required is useful because it eliminates compounds that are not of interest from the results.

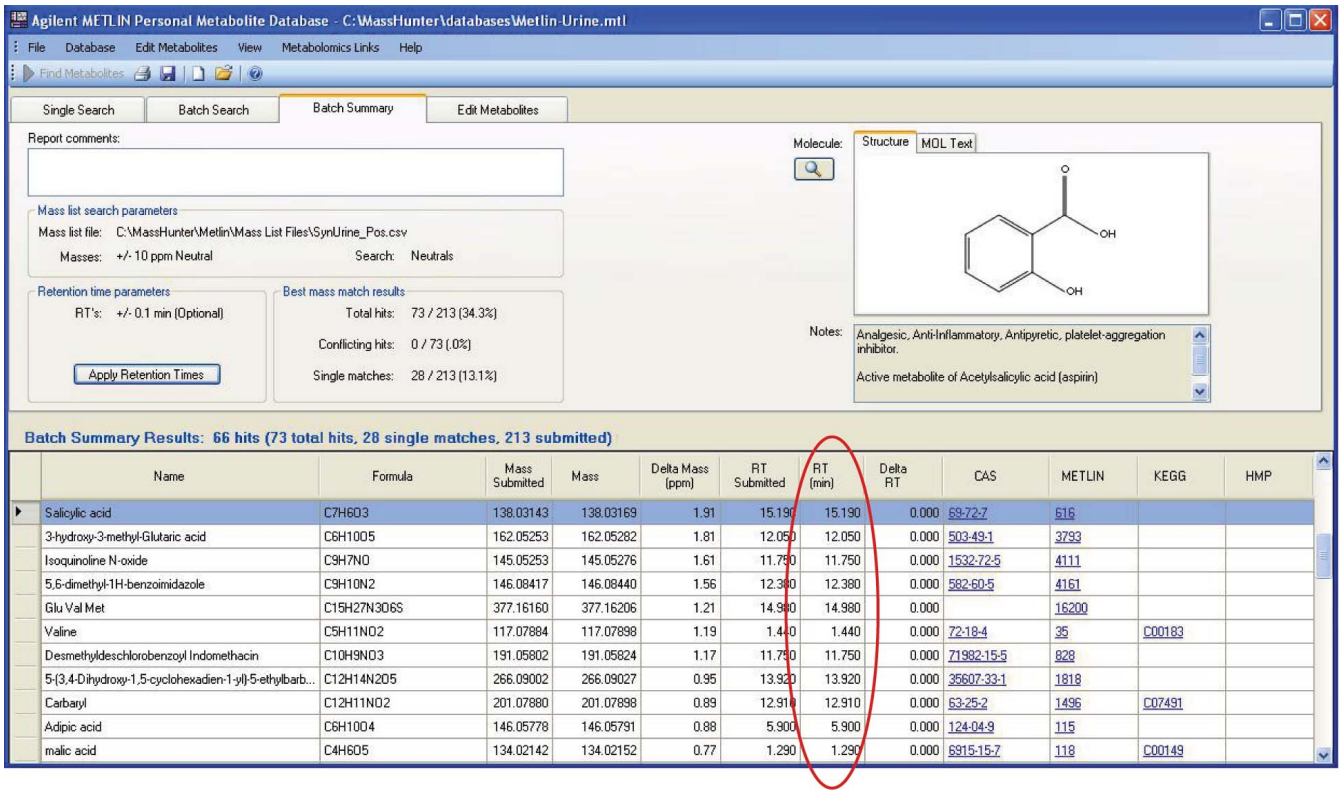


Name	Formula	Mass Submitted	Mass	Delta Mass (ppm)	RT Submitted	RT (min)	Delta RT	CAS	METLIN	KEGG	HMP
Salicylic acid	C7H6O3	138.03143	138.03169	1.91	15.190			69-72-7	616		
3-hydroxy-3-methyl-Glutaric acid	C6H10O5	162.05253	162.05282	1.81	12.050			503-49-1	3793		
Isoquinoline N-oxide	C9H7NO	145.05253	145.05276	1.61	11.750			1532-72-5	4111		
5,6-dimethyl-1H-benzoimidazole	C9H10N2	146.08417	146.08440	1.56	12.380			582-60-5	4161		
Glu Val Met	C15H27N3O6S	377.16160	377.16206	1.21	14.980				16200		
Valine	C5H11NO2	117.07884	117.07898	1.19	1.440			72-18-4	35	C00183	
Desmethyldeschlorobenzoyl Indomethacin	C10H9NO3	191.05802	191.05824	1.17	11.750			71982-15-5	828		
5-(3,4-Dihydroxy-1,5-cyclohexadien-1-yl)-5-ethylbarb...	C12H14N2O5	266.09002	266.09027	0.95	13.920			35607-33-1	1818		
Carbaryl	C12H11NO2	201.07880	201.07898	0.89	12.910			63-25-2	1496	C07491	
Adipic acid	C6H10O4	146.05778	146.05791	0.88	5.900			124-04-9	115		
malic acid	C4H6O5	134.02142	134.02152	0.77	1.290			6915-15-7	118	C00149	

Figure 7. Review the metabolite standard information and confirm that there are no conflicting hits.

Conclusions

Adding retention times to metabolite databases provides a second match criterion orthogonal to, and independent of, compound mass. Searching both accurate mass and retention time will often provide a narrower, more focused list of possible metabolite identities, especially for targeted metabolite searches. Unlike public metabolite databases, retention times can be added to the Agilent METLIN Personal metabolite database. The process of adding retention times to the METLIN database can be significantly simplified and accelerated by preprocessing data from metabolite standards with the Agilent MassHunter Mass Profiler software.



Name	Formula	Mass Submitted	Mass	Delta Mass (ppm)	RT Submitted	RT (min)	Delta RT	CAS	METLIN	KEGG	HMP
Salicylic acid	C7H6O3	138.03143	138.03169	1.91	15.190	15.190	0.000	69-72-7	616		
3-hydroxy-3-methyl-Glutaric acid	C6H10O5	162.05253	162.05282	1.81	12.050	12.050	0.000	503-49-1	3793		
Isoquinoline N-oxide	C9H7NO	145.05253	145.05276	1.61	11.750	11.750	0.000	1532-72-5	4111		
5,6-dimethyl-1H-benzoimidazole	C9H10N2	146.08417	146.08440	1.56	12.380	12.380	0.000	582-60-5	4161		
Glu Val Met	C15H27N3O6S	377.16160	377.16206	1.21	14.980	14.980	0.000		16200		
Valine	C5H11NO2	117.07884	117.07898	1.19	1.440	1.440	0.000	72-18-4	35	C00183	
Desmethyldeschlorobenzoyl Indomethacin	C10H9NO3	191.05802	191.05824	1.17	11.750	11.750	0.000	71982-15-5	828		
5-(3,4-Dihydroxy-1,5-cyclohexadien-1-yl)-5-ethylbarb...	C12H14N2O5	266.09002	266.09027	0.95	13.920	13.920	0.000	35607-33-1	1818		
Carbaryl	C12H11NO2	201.07880	201.07898	0.89	12.910	12.910	0.000	63-25-2	1496	C07491	
Adipic acid	C6H10O4	146.05778	146.05791	0.88	5.900	5.900	0.000	124-04-9	115		
malic acid	C4H6O5	134.02142	134.02152	0.77	1.290	1.290	0.000	6915-15-7	118	C00149	

Figure 8. The AMRT METLIN database is updated to include RT information from the standards Feature Summary table.

About Agilent Technologies

Agilent Technologies is a leading supplier of life science research systems that enable scientists to understand complex biological processes, determine disease mechanisms, and speed drug discovery. Engineered for sensitivity, reproducibility, and workflow productivity, Agilent's life science solutions include instrumentation, microfluidics, software, microarrays, consumables, and services for genomics, proteomics, and metabolomics applications.

Buy online:
www.agilent.com/chem/store

Find an Agilent customer center in your country:
www.agilent.com/chem/contactus

U.S. and Canada
1-800-227-9770
agilent_inquiries@agilent.com

Asia Pacific
adinquiry_aplsca@agilent.com

Europe
info_agilent@agilent.com

Printed in the U.S.A. May 23, 2008

5989-8546EN

Our measure is your success.

Agilent Technologies