WILEY

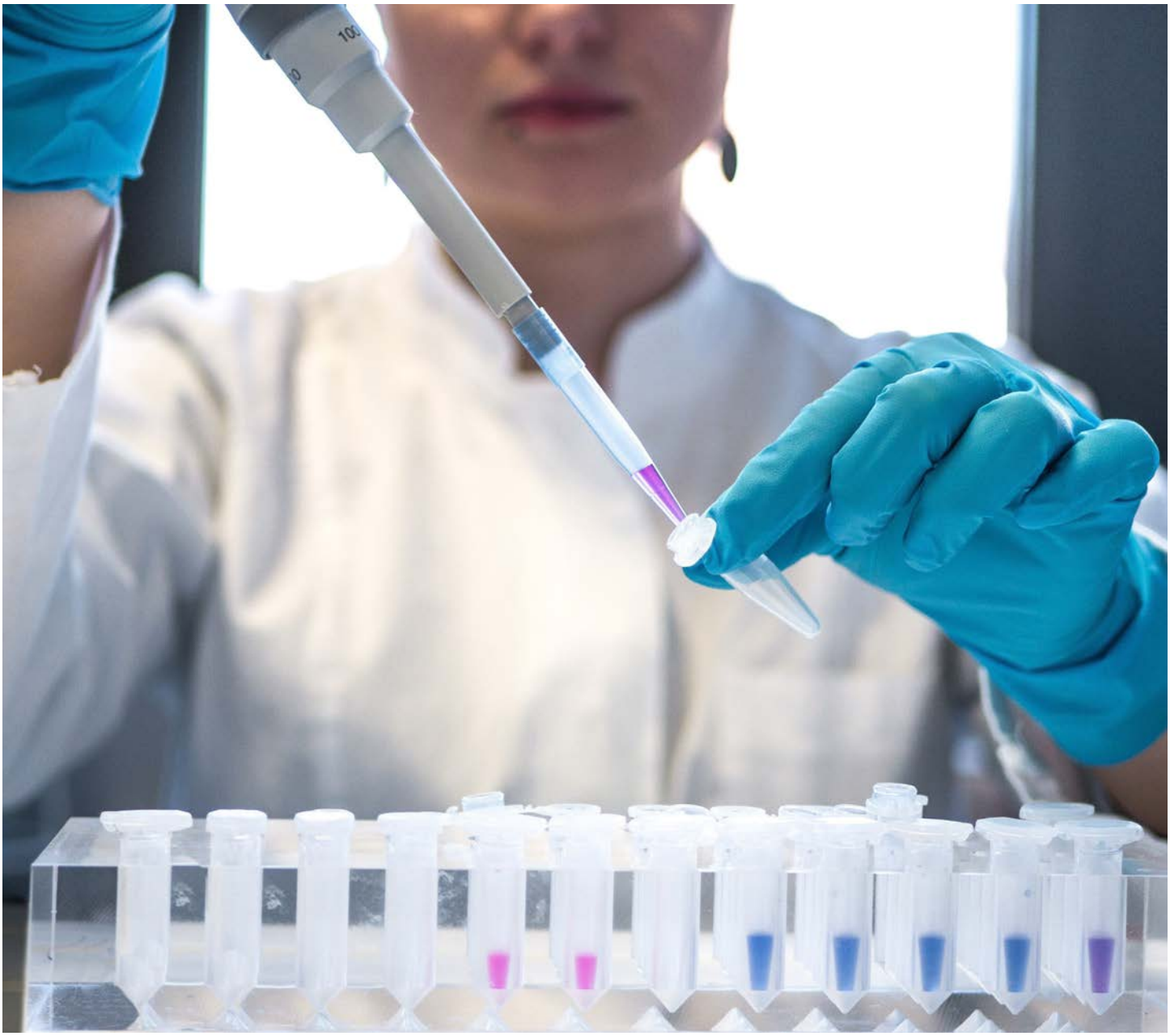
ENABLING DISCOVERY | POWERING EDUCATION | SHAPING WORKFORCES

Automated Non-Targeted GC-MS Analysis by KnowItAll MS Expert

Table of Contents

Introduction

GC-MS data are information rich. Their analysis can be time consuming, especially when examining complex analytes. A new software application in KnowItAll[1] release 2023, **MS Expert**, automatically processes, deconvolutes, and analyzes GC-MS data. Combined with KnowItAll's fast database search, **MS Expert** matches knowns and suggests unknowns. Although **MS Expert** generally does a good job, sometimes it is necessary to send coeluted components to the sophisticated mixture analysis application for further examination, KnowItAll facilitates this seamlessly. Novel compounds' structural characteristics can be further deduced by applying the patent-pending MS **Adaptive Search**[2] that uses fragmentation and structural information to propose probable structural details.



Algorithm

MS Expert processes the total ion chromatogram by following individual m/z values across multiple spectra and extracts a pure spectrum from the data for each individual component, while separating components with overlapping m/z value peaks. If high-resolution data is available, **MS Expert** can determine the most accurate m/z values by taking into account the instrument's resolving power.

During the deconvolution step, individual m/z values are pursued across multiple raw spectra, and a component spectrum is extracted from the data for each individual. The details of the algorithm are summarized by the papers[3-6] in the **Reference** section of this paper.

Additional steps automatically detect components with low intensity **R**econstructed **T**otal **I**on **C**urrent (RTIC) chromatographic peaks as long as they can be distinguished from neighboring components.

Method

The software used for automatic GC-MS analysis is **MS Expert** available in the KnowItAll 2023 release. Sample files used in this article are experimental GC-MS data sets obtained from external sources. The sample files are distributed as part of the KnowItAll 2023 installation in the "C:\Users\Public\Documents\Wiley\KnowItAll\Samples\MS Expert" folder. The entire **KnowItAll Mass Spectral Database Collection**[7], plus optional databases, over 1.2 million spectra, was used as the reference data.

Unit Resolution GC-MS Data

Figure 1 shows the deconvoluted GC-MS data of unit resolution where m/z values are integers, along with the database search results for each component. When RTIC chromatogram ranges are manually selected in the **Chromatogram** pane (A), the calculations are re-run automatically to perform a new analysis. In the **Components Table** (E), the number of matched reference compounds per component can be adjusted.

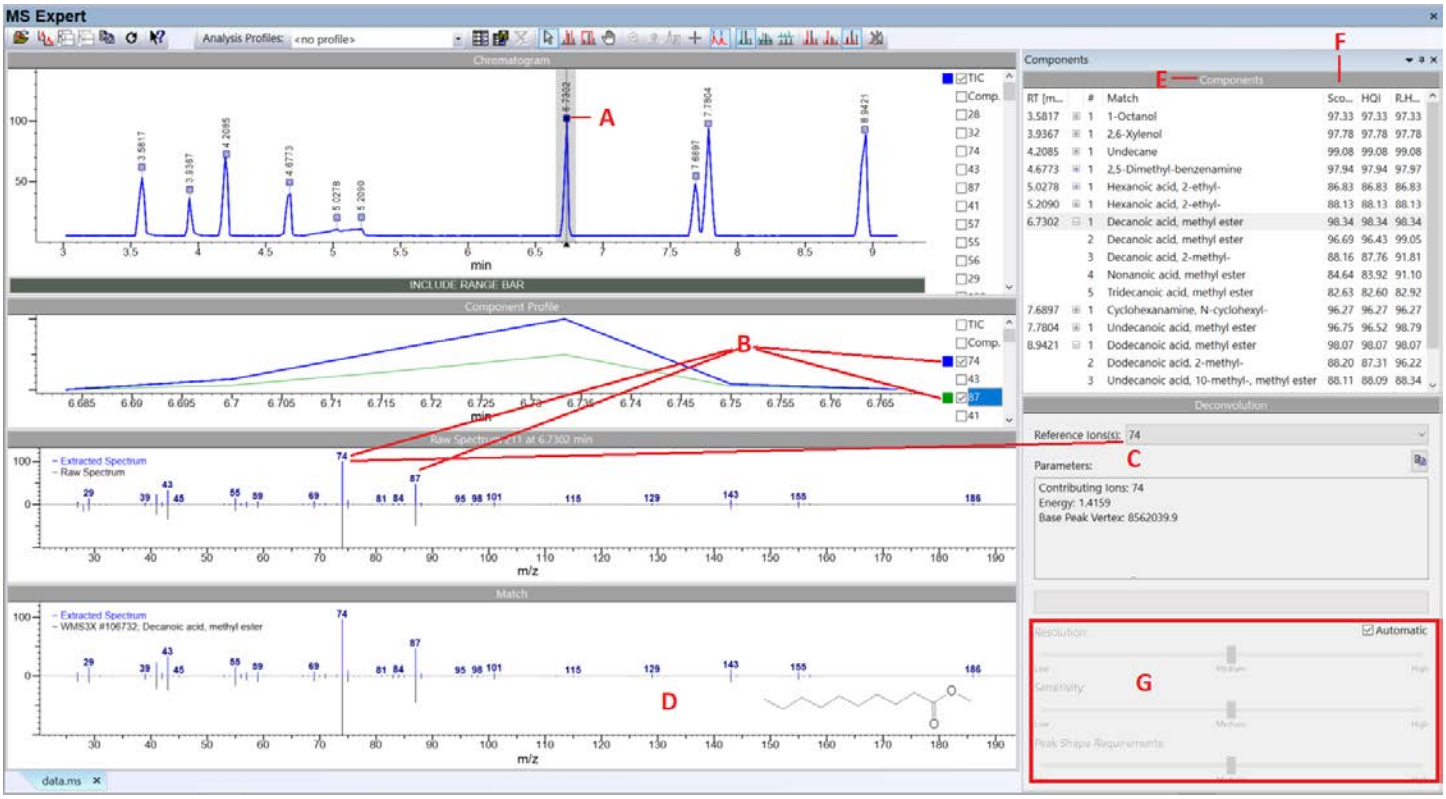


Figure 1. A – Deconvoluted RTIC peaks whose extracted MS spectrum found matches from reference MS databases; B – Selected ions in a component; C – Reference ion(s) used to model a component; D – Extracted MS spectrum (top) vs reference MS spectrum (bottom); E – Components table; F – Combined spectrum search and reverse MS search Hit Quality Index (HQI) in %; G – Analytical parameters (adjustable).

Tables 1 – 5 were used to display test results consistently.

- The **RT (MIN)** column lists the retention time in minutes for each deconvoluted component.
- The **HIT** column lists the position of a MS database record in a hit list.
- The **SCORE** column lists the match quality percentage between extracted MS spectrum vs. that of the database record, where 100 would be a perfect match and 0 is no match at all.
- The name of the database record is listed in the **MS EXPERT COMPONENT** column.

  - If a database record name and manually assigned component name are synonymous to each other, "(*synonym*)" was suffixed to the database record name.
  - For erroneous assignments made by **MS Expert**, "(*error*)" was appended to the database record names.
- The **MANUAL ASSIGNMENT** column lists the actual compound assigned by the scientist who had contributed the data.
- If the same component is associated with multiple **RT (MIN)** values, "(*duplicate*)" was appended to the name(s) following the first one.
- Some chromatogram peaks are from the atmosphere or solvent, we labeled them with "(*ignored*)" appendix.

The result of the first sample of automatic GC-MS deconvolution and component identification is shown in **Table 1**. As seen in columns 4 and 5, automatically assigned components and manually assigned components agree with each other. The **HIT**s were all the first choice among matches between the extracted MS and the database record. The **SCORE**s of the matches were also well within the 90 percentiles.

Table 1. Sample 1

RT (MIN)	HIT	SCORE	MS EXPERT COMPONENT	MANUAL ASSIGNMENT
3.5883	1	98.41	1-Octanol	1-Octanol
3.9367	1	99.1	Phenol, 2,6-dimethyl-	Phenol, 2,6-dimethyl-
4.2085	1	99.08	Undecane	Undecane
4.6773	1	97.94	2,5-Dimethyl-benzenamine (*synonym*)	1-AMINO-2,6-DIMETHYLBENZENE
5.0278	1	93.19	Hexanoic acid, 2-ethyl-	Hexanoic acid, 2-ethyl-
5.209	1	92.5	Hexanoic acid, 2-ethyl- (*duplicate*)	Hexanoic acid, 2-ethyl-
6.7302	1	98.34	Decanoic acid, methyl ester	Decanoic acid, methyl ester
7.6897	1	98.07	Cyclohexanamine, N-cyclohexyl-	Cyclohexanamine, N-cyclohexyl-
7.7804	1	98.31	Undecanoic acid, methyl ester	Undecanoic acid, methyl ester
8.9421	1	98.37	Dodecanoic acid, methyl ester	Dodecanoic acid, methyl ester

Sample 2 is a mixture of solvents. The matches were excellent except that in the last component row, **MS Expert** suggested an extra component, silane, tetrafluoro-, with a low **SCORE** value of 55.01%.

Table 2. Sample 2

RT (MIN)	HIT	SCORE	MS EXPERT COMPONENT	MANUAL ASSIGNMENT
4.7821	1	97.48	Methanol	Methanol
5.794	1	98.23	Ethanol	Ethanol
6.3659	1	94.08	Acetone	Acetone
6.5407	1	97.65	Isopropyl Alcohol	Isopropyl Alcohol
6.7635	1	99.78	Acetonitrile	Acetonitrile
7.0053	1	98.24	Methane, dichloro-	Methane, dichloro-
8.0065	1	94.45	1-Propanol	1-Propanol
8.7426	1	93.31	Ethyl Acetate	Ethyl Acetate
9.0787	1	98.82	Tetrahydrofuran	Tetrahydrofuran
9.8496	1	92.18	Acetic acid, 1-methylethyl ester	Acetic acid, 1-methylethyl ester
10.4866	1	89.14	1-Butanol	1-Butanol
12.2444	1	98.7	Pyridine	Pyridine
12.3922	1	99.76	Toluene	Toluene
13.7601	1	78.35	2-Furanol, tetrahydro-	2-Furanol, tetrahydro-
13.8912	1	89.92	Dimethylformamide	Dimethylformamide
17.8148	1	88.12	2(3H)-Furanone, dihydro-	2(3H)-Furanone, dihydro-
17.8244	1	55.01	Silane, tetrafluoro-	

Sample 3 is a complex mixture. Forty-eight components had been manually assigned while **MS Expert** found forty-seven. It made one erroneous assignment at **RT (MIN)** 25.2945."

Table 3. Sample 3

RT (MIN)	HIT	SCORE	MS EXPERT COMPONENT	MANUAL ASSIGNMENT
0.7241	1	99.22	Argon (*ignored*)	Argon (*ignored*)
0.9833	1	99.95	Carbamic acid, monoammonium salt (*error, ignored*)	Carbon dioxide (*ignored*)

Results

RT (MIN)	HIT	SCORE	MS EXPERT COMPONENT	MANUAL ASSIGNMENT
2.0297	1	98.75	Methane, chloro-	Methane, chloro-
3.1198	1	97.68	Methane, oxybis-	Methane, oxybis-
3.2225	1	99.61	Methyl bromide	Methyl bromide
4.2028	1	97.51	Ethene, chloro-	Ethene, chloro-
5.3466	1	99.82	Ethyl chloride (*duplicate*)	Ethyl chloride (*duplicate*)
7.6648	1	99.79	Methylene chloride	Methylene chloride
8.5025	1	95.43	Acetone	Acetone
9.1442	1	98.82	Dithioxomethane	Dithioxomethane
9.7463	1	99.85	Trichloromonofluoromethane	Trichloromonofluoromethane
10.4328	1	99.27	Ethene, 1,1-dichloro-	Ethene, 1,1-dichloro- and Methyl acetate
10.5386				2-Propanone, 1-hydroxy- ?
10.7473	1	96	Methylal (*Synonym*)	METHANE, DIMETHOXY-
11.0188	1	99.38	Methane, bromochloro-	Methane, bromochloro-
11.9073	1	98.99	Ethane, 1,1-dichloro-	Ethane, 1,1-dichloro-
12.6332	1	99.64	Ethylene, 1,2-dichloro-, (E)-	Ethylene, 1,2-dichloro-, (E)-
12.1746				Argon? (*duplicate*)
13.3568	1	99.45	Methane, trichloro-	Methane, trichloro-
14.0243	1	93.36	1,2-Dichloroethane-d4	1,2-Dichloroethane-d4 and 1,2-Dichloroethane
14.1112	1	95.04	2-Butanone	2-Butanone
14.1366	1	85.38	Ethane, 1,2-dichloro-	Ethane, 1,2-dichloro- (*duplicate*)
14.3376				2-Butanone (*duplicate*)
15.516	1	99.39	Ethane, 1,1,1-trichloro-	Ethane, 1,1,1-trichloro-
15.7739	1	97.88	1,4-Dioxane	1,4-Dioxane

Results

RT (MIN)	HIT	SCORE	MS EXPERT COMPONENT	MANUAL ASSIGNMENT
15.926	1	98.29	Carbon Tetrachloride	Carbon Tetrachloride
15.9646				Ethanamine, N-ethyl-N-[(ethylthio) methyl]-
16.0176	1	97.90	2,3-Butanedione (*synonym*)	Acetic acid ethyl ester
16.5463	1	98.92	Methane, bromodichloro-	Methane, bromodichloro-
17.4897	1	93.27	2-Butanone, 3-methyl-	2-Butanone, 3-methyl-
17.9705	1	90.66	Propane, 1,2-dichloro-	Propane, 1,2-dichloro-
18.2397	1	98.08	1-Propene, 1,3-dichloro-, (Z)-	1-Propene, 1,3-dichloro-, (Z)-
18.7663	1	99.66	Trichloroethylene	Trichloroethylene
19.2624	1	99.44	Benzene	Benzene
19.6127	1	56.74	Methane, dibromochloro-	Methane, dibromochloro-
19.6533	1	95.7	Ethane, 1,1,2-trichloro-	Ethane, 1,1,2-trichloro-
20.7145	1	97.23	Ethene, (2-chloroethoxy)-	Ethene, (2-chloroethoxy)-
21.0187	1	92.22	Butanoic acid, methyl ester	
21.9853	1	98.28	Benzene, 1,4-difluoro-	Benzene, 1,4-difluoro-
22.5206	1	93.94	Methane, tribromo-	Methane, tribromo-
22.8777	1	96.97	Methyl Isobutyl Ketone	Methyl Isobutyl Ketone
23.7254	1	88.39	Butanoic acid, 2-methyl-, methyl ester	Butanoic acid, 2-methyl-, methyl ester
24.5232	1	95.25	2-Hexanone	2-Hexanone
24.8421	1	96.98	Tetrachloroethylene	Tetrachloroethylene
24.9644	1	78.29	Ethane, 1,1,2,2-tetrachloro-	Ethane, 1,1,2,2-tetrachloro-
25.2945	1	63.22	1-Butylvinyl methyl ether (*error*)	Oxirane, 2-methyl-2-(2-methylpropyl)-
25.9844	1	95.96	Toluene-D8	Toluene-D8
26.1962	1	97.79	Toluene	Toluene

Results

RT (MIN)	HIT	SCORE	MS EXPERT COMPONENT	MANUAL ASSIGNMENT
26.4586				1,4-Dichlorobenzene-D4
27.4897	1	89.22	Chlorobenzene-D5	Chlorobenzene-D5
27.6442	1	95.57	Benzene, chloro-	Benzene, chloro-
30.3625	1	99.07	Ethylbenzene	Ethylbenzene
34.0456	1	95.08	p-Bromofluorobenzene	p-Bromofluorobenzene
36.0513	1	98.94	Styrene	Styrene
37.7879				m-Xylene (*duplicate*)
37.8439	1	98.34	o-Xylene (*positional isomer*)	m-Xylene

MS Expert did not find a component at **RT (MIN)** 10.5386. The "2-propanone, 1-hydroxy- ?" assigned by a scientist suggests the uncertainty of that assignment. Similarly, **MS Expert** did not find ethanamine, N-ethyl-N-[(ethylthio)methyl]- at **RT (MIN)** 15.9646 either. It did not find **RT(MIN)** 26.4586, 1,4-dichlorobenzene-D4.

According to the scientist, the component at **RT (MIN)** 10.4328 could be a mixture of ethene, 1,1-dichloro- and methyl acetate. **MS Expert** assigned it to one component, ethene, 1,1-dichloro-. **RT (MIN)** 14.0243 should be a mixture of 1,2-dichloroethane-d4 and 1,2-dichloroethane, while **MS Expert** only found one component, 1,2-dichloroethane-d4.

For the last component, **RT (MIN)** 37.8439, **MS Expert** picked the positional isomer of the structure identified by the scientist, which is acceptable for MS.

Encouragingly, **MS Expert** picked up an extra component, butanoic acid, methyl ester, with a **SCORE** 92 at **RT (MIN)** 21.0187.

Coeluted Components (Mixture) Analysis

It is interesting that **MS Expert** gave component at **RT (MIN)** 19.6127 a very low match score, 56.74%, when compared to the reference spectrum of Methane, dibromochloro-. From the bottom spectrum pane (**Figure 2** red-board box), one can see that something else existed in the extracted spectrum (top) in addition to Methane, dibromochloro- (bottom). Obviously, we had a mixture.

Results

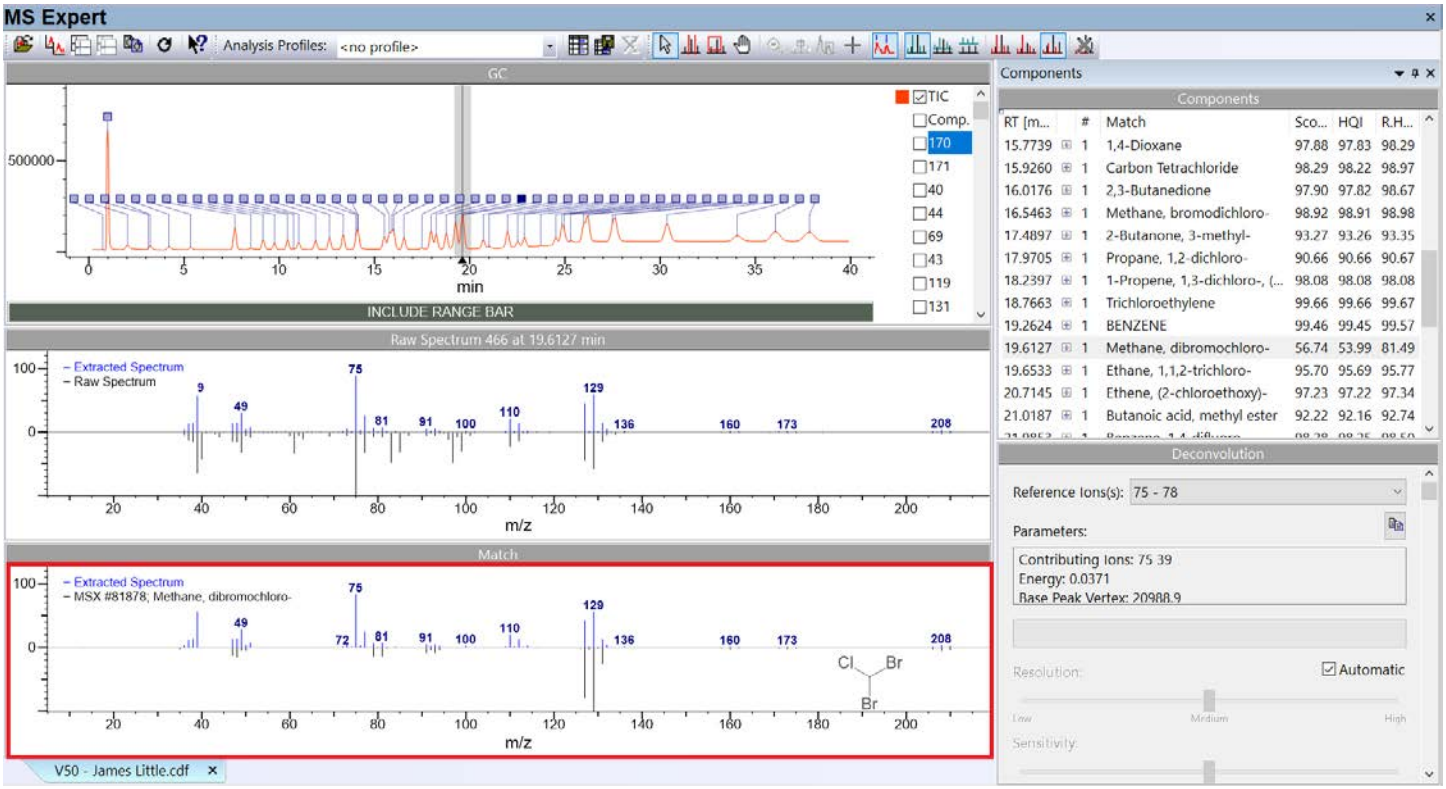


Figure 2. Extracted spectrum is a mixture

MS Expert allows one to transfer this extracted spectrum to another application, **SearchIt**, for quick automatic mixture analysis. The first hit of all possible component combinations is shown in **Figure 3**. The first row is the **Composite Spectrum**, a mixture of methane, dibromochloro- and 1-propene, 1,3-dichloro-. The **Composite Spectrum** had a high HQI value, 94.09%, when compared against the extracted spectrum. Rows 2 and 3 were each component's contributions. Row 4 was the difference between extracted spectrum and **Composite Spectrum**.

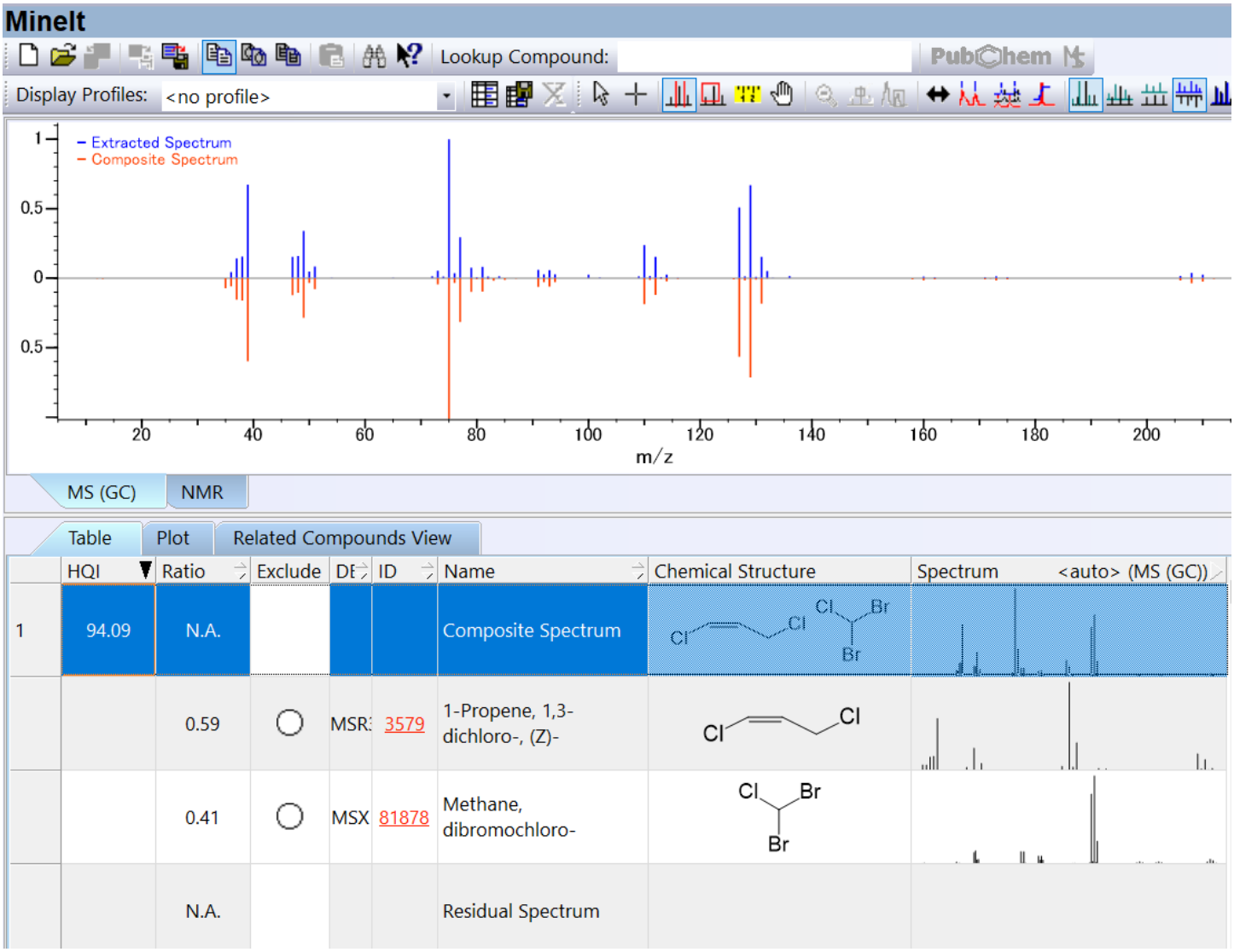


Figure 3. Mixture analysis of extracted spectrum

It is tricky to assess the accuracy of **MS Expert** for this sample. Out of forty-eight user assigned components, **MS Expert** missed three completely. Two partially assigned components were counted as one missed. There was one erroneous assignment. Therefore, overall accuracy is 43/48, i.e. 89.58%. **MS Expert** did pick up an extra component, which made the overall accuracy "89.58% +".

As shown by the three examples above, **KnowItAll MS Expert** has excellent accuracy when analyzing unit resolution GC-MS data.

Accurate m/z Value GC-MS Data

For accurate m/z value data, without knowing the instrument resolving power, there is no safe algorithm to automatically calculate this value. Increasing the m/z value accuracy too much risks splitting an individual m/z value into individual mass spectral peaks which should be a single peak. Decreasing the m/z value accuracy too much may cause individual mass spectral peaks to merge, resulting in incorrect accurate m/z values. Based on our research, **MS Expert** generates a reasonable value by default. Empirically, this works in most cases. However, in most cases, users will know the instrument resolution, which is the recommended approach for getting the best results (**Figure 4**)

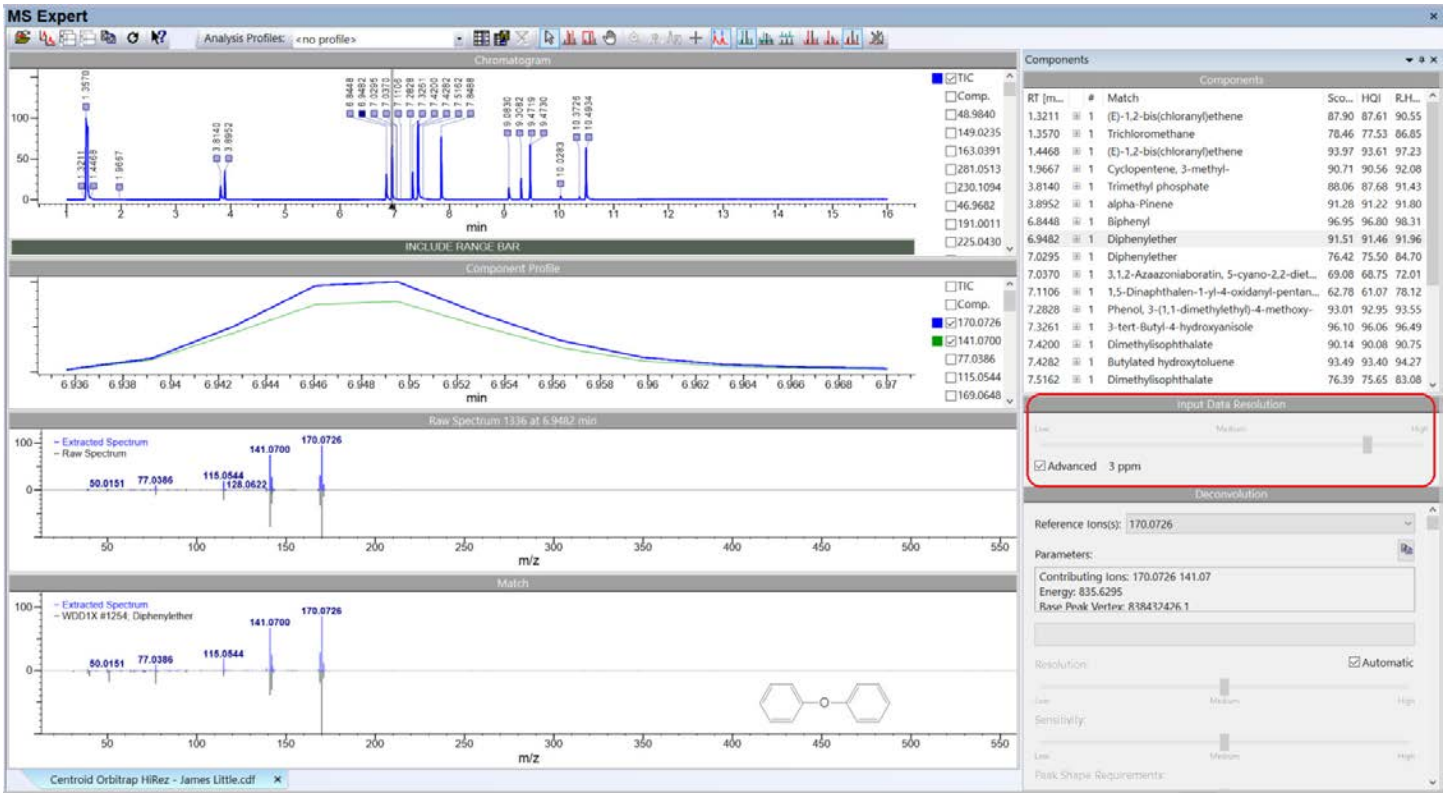


Figure 4. Accurate m/z GC-MS data analysis where one can use the highlighted **Input Data Resolution** to set the resolution for the data file.

MS Expert shows its unique strength in finding components in accurate mass GC-MS data. Many extra components were discovered in **Sample 4 (Table 4)**. Interestingly, it considered cyclopentene, 3-methyl-, an isomer of cyclohexene, to be a good match for extracted MS at **RT (MIN)** 1.9667. It also considered m-terphenyl, a positional isomer, to be a top match for the last component, p-terphenyl with a **SCORE** of 97.02. The database MS spectrum of p-terphenyl had a **SCORE** of 96.94, which put this correct result second in the hit list.

Table 4. Sample 4

RT (MIN)	HIT	SCORE	MS EXPERT COMPONENT	MANUAL ASSIGNMENT
1.3211	1	81.53	(E)-1,2-bis(chloranyl)ethene	
1.357	1	76.33	Trichloromethane	
1.365	1	53.96	Methanethiol	
1.3852	1	87.65	Methane, dichloro-	Methane, dichloro-
1.4468	1	95.41	(E)-1,2-bis(chloranyl)ethene (*duplicate*)	
1.9667	1	90.92	Cyclopentene, 3-methyl- (*isomer*)	Cyclohexene
3.814	1	88.2	Trimethyl phosphate	Trimethyl phosphate
3.8952	1	91.36	alpha-Pinene	alpha-Pinene
6.8449	1	96.72	Biphenyl	Biphenyl
6.9482	1	91.56	Diphenylether	Diphenylether
7.0295	1	87.03	Diphenylether (*duplicate*)	Diphenylether (*duplicate*)
7.037	1	69.83	3,1,2-AZAAZONIABORATIN, 5-CYANO-2,2-DIETHYL-2,3-DIHYDRO-1,3-DIDEUTERO-4,6-BIS(TRIDEUTEROMETHYL)-	
7.111	1	72.44	Biphenyl (*duplicate*)	
7.2828	1	93.03	Phenol, 3-(1,1-dimethylethyl)-4-methoxy-	Phenol, 3-(1,1-dimethylethyl)-4-methoxy-
7.3261	1	96.11	3-tert-Butyl-4-hydroxyanisole	3-tert-Butyl-4-hydroxyanisole
7.4193	1	90.17	Dimethylisophthalate	Dimethylisophthalate
7.4282	1	93.17	Butylated hydroxytoluene	Butylated hydroxytoluene
7.5162	1	76.54	Dimethylisophthalate (*duplicate*)	
7.8488	1	94.57	Diethylphthalate	Diethylphthalate
9.083	1	96.12	Caffeine	Caffeine
9.3082	1	93.31	Hexadecanoic acid, methyl ester	Hexadecanoic acid, methyl ester

RT (MIN)	HIT	SCORE	MS EXPERT COMPONENT	MANUAL ASSIGNMENT
9.4719	1	95.43	1,4-Dibutyl benzene-1,4-dicarboxylate	
9.473	1	93.82	Dibutyl phthalate	Dibutyl phthalate
10.0283	1	95.72	Drometrizole	Drometrizole
10.3739	1	93.88	Phenol, 4,4'-(1-methylethylidene)bis-	Phenol, 4,4'-(1-methylethylidene)bis-
10.4934	2	96.94	p-Terphenyl	p-Terphenyl

Adjust to Instrument Resolution

In the above example, the original instrument resolution was not included in the data file. **MS Expert** assumed a reasonably good value. However, when we looked at the component at **RT (MIN)** 6.9482 (Diphenylether), the accurate mass 170.0717 was slightly low. By adjusting the **Input Data Resolution** manually to the instrument value of 3 ppm, this component's m/z value became 170.0726, which is correct.

Table 5 contains many empty cells in the **MANUAL ASSIGNMENT** column, indicating that **MS Expert** found many extra components in **Sample 5**. The high **SCORE** values for most of the matched database records is an indication of the accuracy of automatic analysis. It is worth noting that at **RT (MIN)** 19.2284, the second choice in the database match list agreed with the manual assignment of 1-hexadecanol, with a **SCORE** of 94.98. The first choice would be cyclotetradecane whose **SCORE** was 95.29, just slightly greater than the second choice. The consensus amongst the MS community is that one should always examine the hit list of extracted MS and database matches rather than accepting the first match. Importantly, most manual assignments agreed with the **MS Expert** automatic assignments.

Table 5. Sample 5

RT (MIN)	HIT	SCORE	MS EXPERT COMPONENT	MANUAL ASSIGNMENT
1.8389	1	80.68	1,1-Diethoxyethane	Ethane, 1,1-diethoxy
2.073	1	97.65	Ethanol	Ethanol
2.1313	1	93.97	1,3-Propanediol, 2-nitro-2-[(nitrooxy) methyl]-, dinitrate (ester)	
2.8304	1	87.6	Butanoic acid, ethyl ester	
2.8837	1	98.24	Ethanol (*duplicate*)	
2.9346	1	97.96	Ethanol P248 (*duplicate*)	

Results

RT (MIN)	HIT	SCORE	MS EXPERT COMPONENT	MANUAL ASSIGNMENT
3.4974	1	88.44	1-Propanol, 2-methyl-	
3.6973	1	96.15	1-Butanol, 3-methyl-, acetate	
4.9095	1	94.12	1-Butanol, 3-methyl-	1-Butanol, 3-methyl-
5.2095	1	98.41	Hexanoic acid, ethyl ester	Hexanoic acid, ethyl ester
5.3268	1	89.43	Hexanoic acid, ethyl ester (*duplicate*)	
5.8228	1	75.14	Silane, triethylfluoro-	
6.6175	1	88.23	Heptanoic acid, ethyl ester	
7.419	1	83.36	2-Dodecanone	
8.0825	1	98.46	Octanoic acid, ethyl ester	Octanoic acid, ethyl ester
8.3181	1	96.12	Acetic acid	
8.4844	1	81.66	(4R)-4-Hydroxymethyl-4-methylcyclopent-2-en-1-one	
8.7777	1	48.44	Benzoic acid, 4-methyl-, dimethyl(pentafluorophenyl)silyl ester	
9.2638	1	73.22	Benzaldehyde	
10.6947	1	98.36	Decanoic acid, ethyl ester	Decanoic acid, ethyl ester
10.8586	1	91.69	Octanoic acid, 3-methylbutyl ester	
11.0116	1	71.58	Benzoic acid, ethyl ester	
11.0505	1	85.21	4-Ethoxy-4-oxobutanoic acid	
11.5905	1	71.21	Decanoic acid, propyl ester	
11.9513	1	84.11	n-Capric acid isobutyl ester	
11.9904	1	68.59	1-Hexadecanol	
12.6939	1	94.87	Acetic acid, 2-phenylethyl ester	
13.0114	1	97.59	Dodecanoic acid, ethyl ester	Dodecanoic acid, ethyl ester

Results

RT (MIN)	HIT	SCORE	MS EXPERT COMPONENT	MANUAL ASSIGNMENT
13.1823	1	92.76	Pentadecanoic acid, 3-methylbutyl ester	
13.7568	1	74.13	4-Anilino-4-keto-2-phenyl-butyric acid	
13.7749	1	96.65	Phenylethyl Alcohol	Phenylethyl Alcohol
14.3433	1	89.83	1-Decanol	
15.0314	1	93.98	Undecanone	
15.2301	1	69.42	1,6,10-Dodecatrien-3-ol, 3,7,11-trimethyl-, [S-(Z)]-	
15.3582	1	96.09	Tetradecanoic acid, ethyl ester	Tetradecanoic acid, ethyl ester
15.4909	1	98.05	Octanoic acid	Octanoic acid
15.5692	1	76.69	Hexane, 2,3-dimethyl-	
15.649	1	66.12	3-Octanol, 3,6-dimethyl-	
16.7976	1	97.62	Tetradecanol <n-> (*synonym*)	1-Dodecanol
17.8032	1	94.59	Palmitate <ethyl-> (*synonym*)	Hexadecanoic acid, ethyl ester
17.8821	1	76.35	Cyclopentanol, nitrate	
18.0304	1	98.54	n-Decanoic acid	n-Decanoic acid
18.13	1	91.19	Ethyl 9-hexadecenoate	Ethyl 9-hexadecenoate
19.2284	2	94.98	1-Hexadecanol	1-Hexadecanol
20.4895	1	98.07	Dodecanoic acid (*duplicate*)	Dodecanoic acid (*duplicate*)
22.813	1	92.89	Myristic acid	Myristic acid
25.0184				Hexadecanoic acid <n->
25.0732	1	87.48	4-Pentenenitrile, 3-hydroxy-	4-Pentenenitrile, 3-hydroxy-
25.4155	1	74.93	10-Undecenoic acid	
34.2401	1	96.18	Cyclotetrasiloxane, octamethyl-	

RT (MIN)	HIT	SCORE	MS EXPERT COMPONENT	MANUAL ASSIGNMENT
34.2659	1	95.74	Cyclotetrasiloxane, octamethyl- (*duplicate*)	
34.3045	1	95.95	Cyclotetrasiloxane, octamethyl- (*duplicate*)	

As shown in **Table 5**, by using default analysis parameters, **MS Expert** missed to assign one component initially, hexadecanoic acid <n-> at **RT (MIN)** 25.0184. Zooming in region 24 – 26 min (**Figure 5**), three tiny peaks are seen in large scale (**Figure 6**).

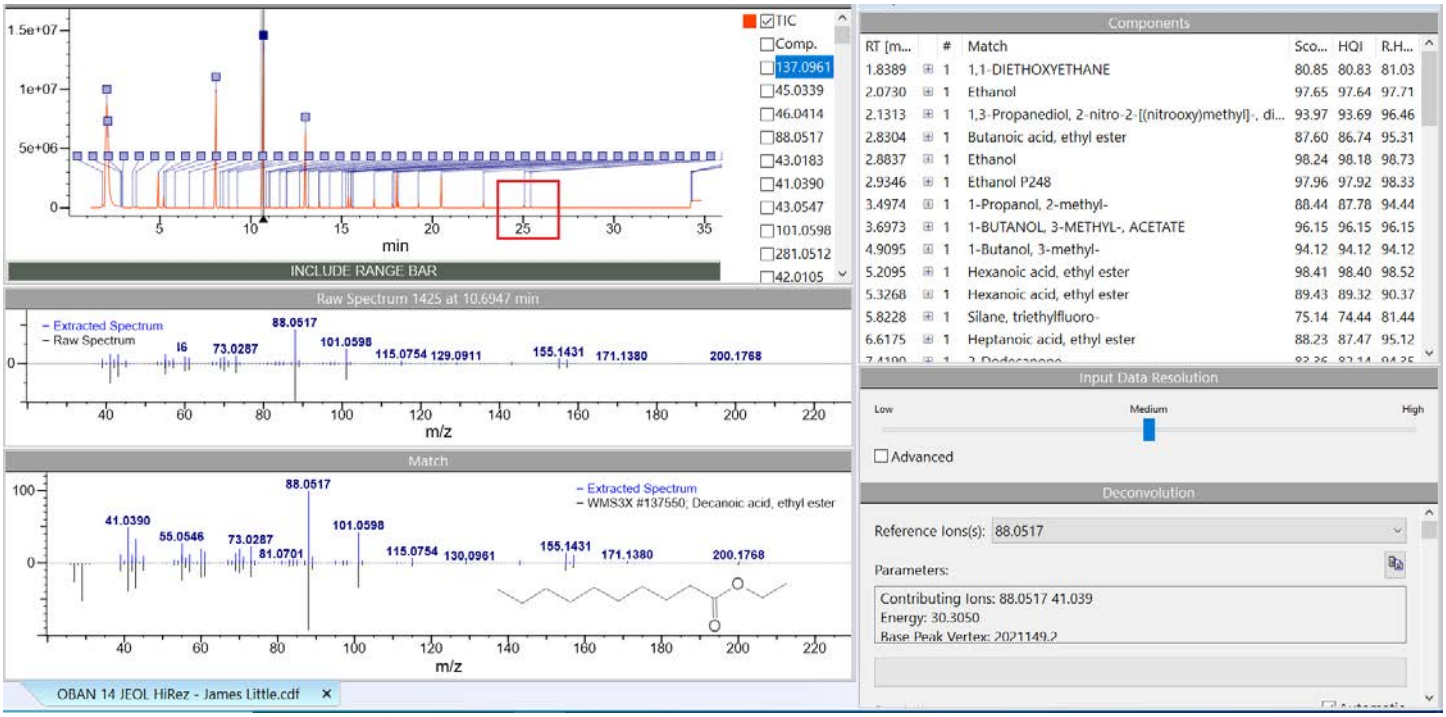


Figure 5. Interesting small peaks around **RT (MIN)** 25 min region.

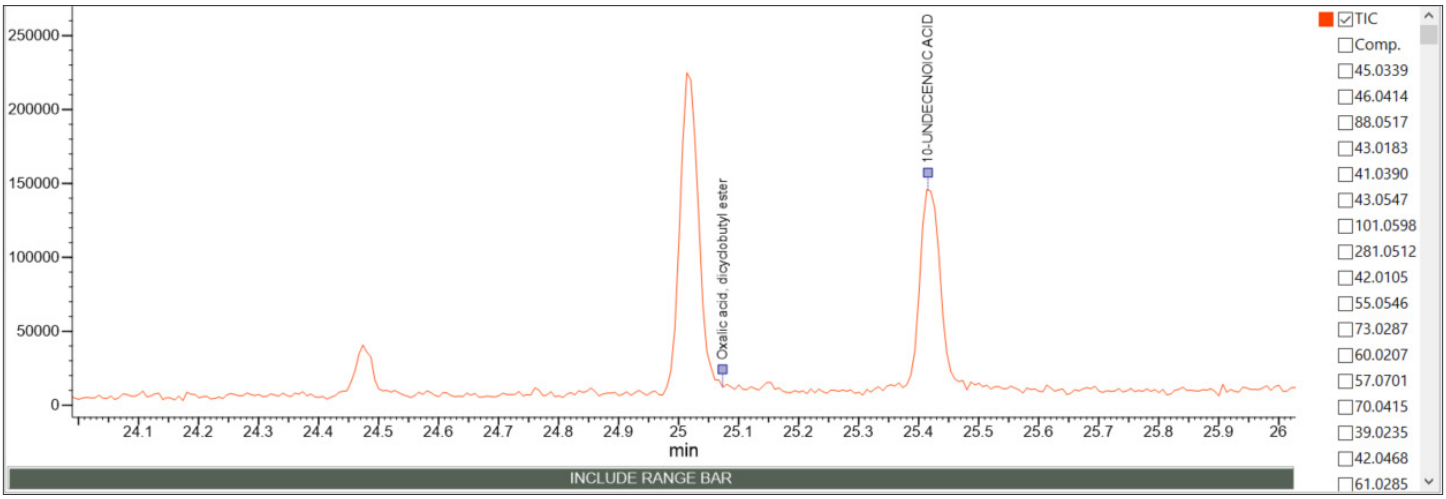


Figure 6. Small peaks.

There is an issue that a component was assigned to the baseline.

Adjust Algorithm Sensitivity

By manually adjusting the sensitivity used by **MS Expert** deconvolution algorithm, the appropriate component, hexadecanoic acid, missed previously, was assigned to the peak, shown in **Figure 7**.

Figure 7. Sensitivity is one of the adjustable parameters

Transfer Unknown to Adaptive Search

For the small peak around 24.245 -24.50 (min) in **Figure 7**, restricting analysis region to this peak yielded no hits. By transferring the **Raw Spectrum** from this region to **SearchIt** and performing an Adaptive Search, a good similar compound can be found.

Accuracy Assessment

A summary of the accuracy of **MS Expert** is in **Table 6**.

Table 6. MS Expert Performance

SAMPLE	MASS	#AUTO ASSIGNED	#MANUAL ASSIGNED	#ASSIGN ERROR	#MISSED	#PARTIALLY ASSIGNED	#2ND HIT BY AUTO ASSIGN	#EXTRA	ACCURACY
1	unit	9	9	0	0	0	0	0	100
2	unit	17	16	0	0	0	0	1	100 +
3	unit	47	48	1	3	2	0	1	89.58 +
4	accurate	22	17	0	0	0	0	5	100 +
5	accurate	45	19	0	0	0	1	26	100 +

Although a limited sample, this suggests that MS Expert performs as well as an industry expert for analysis of unit mass spectra and may significantly outperform experts in analysis of complex high-resolution spectra.

Speed of Analysis

MS Expert processes data as follows:

1. Deconvolutes data into components
2. Simultaneously, for each component, performs the following:
    a. A normal MS database search of the entire Wiley MS spectral collection7
    b. A reverse search over the same reference dataset
    c. Combines both the normal and reverse search **HQIs** values into a **SCORE** value and sorts the hit lists based on the **SCORE** values

The processing of the five presented samples each completed within seconds. Scientists who contributed the jum nyoriginal datasets estimated that it would take hours to days to analyze these kinds of samples manually. Therefore, **MS Expert** is a great efficiency booster.

Even the largest reference data pool will not contain all compounds. **MS Expert** allows users to add their own databases to the reference data pool. Components are easily reported and exported by this system.

Summary

In conclusion, with proven accuracy and high execution speeds, **KnowItAll MS Expert** can reliably improve the efficiency of time-consuming GC-MS data analysis for both unit mass and high-resolution GC-MS TICs. Although a limited set of samples, our findings suggest that **MS Expert** performs as well as an industry expert for analysis of unit mass spectra and may significantly outperform experts in analysis of complex high-resolution spectra. Application of this workflow will be especially useful for non-targeted screening of complex biological and environmental analytes.

References

1. *KnowItAll Analytical Edition Software - Wiley Science Solutions.* https://sciencesolutions.wiley.com/knowitall-spectroscopy-software/ (accessed on 2022-09-01).

2. Nedwed, K.; D'Souza, M.; Abshear, T. for Adaptive Search Mass Spectrometer Spectral Analysis. Patent Pending

3. Stein, S .E. An Integrated Method for Spectrum Extraction and Compound Identification from Gas Chromatography/Mass Spectrometry Data. *J. Am. Soc. Mass Spectrom*. **1999**, 10 (8), 770-781.

4. Dromey, R. G.; Stefik, M. J.; Rindfleisch, T. C.; Duffield, A. M. Extraction of Mass Spectra Free of Background and Neighboring Component Contributions from Gas Chromatography/Mass Spectrometry Data. Anal. Chem. **1976**, *48* (9), 1368-1375.

5. Biller, J. E.; Biemann, K. Reconstructed Mass Spectra, A Novel Approach For The Utilization Of Gas Chromatography-Mass Spectrometer Data. *Anal. Lett.* **1974**, 7, 515-528.

6. Colby, B. N. Spectral Deconvolution for Overlapping GC/MS Components. *J. Am. Soc. Mass Spectrom.* **1992**, 3, 558-562.

7. *KnowItAll Mass Spectral Database Collection* - Wiley Science Solutions. https://sciencesolutions.wiley.com/solutions/technique/gc-ms/knowitall-ms-collection/ (accessed on 2022-09-01).

8. O. D. Sparkman, J. Little. Identification of Fentanyl-Related Species by Wiley Adaptive MS and IR Searches. *James Little*. YouTube, August 16, 2022. https://www.youtube.com/watch?v=mdCsxj4HC7A (accessed on 2022-09-01).