KnowItAll Software Training

Simple Spectral Search/Identification with KnowItAll ID Expert

WILEY

Simple Spectral Search/Identification

How to Use KnowItAll ID Expert to Perform a Simple Spectral Search/Identification

Purpose

These exercises demonstrate how to use KnowItAll ID Expert to identify IR, Raman, and other spectra.

Objectives

These exercises will teach you:

- How to perform single component searches, multiple component searches, designer drug classification & functional group analyses simultaneously, and view results on a single screen for complete view of all possibilities for the unknown spectrum
- How to deconvolve industrial material to organic chemicals, inorganic chemicals, and component chemicals
- How KnowItAll ID Expert's patented optimized corrections technology can help you find the optimal search results
- How to generate a PDF report with a single click

Background

The KnowItAll ID Expert spectral identification software combined with the KnowItAll Spectral Libraries provide fast answers for scientists identifying unknown spectra.

It's easy to use. Simply open an unknown spectrum and KnowItAll ID Expert automatically performs single component searches, multiple component searches, possible designer drug classification and functional group analyses simultaneously, and summarizes the results on a single screen to give a complete view of all possibilities for the unknown. It can also perform analysis using only pure organic and inorganic compound spectra, thus breaking down industrial material into basic building blocks. If there are problems with the query spectrum, ID Expert has the spectral intelligence to identify issues and fix them using patented Optimized Corrections. Once the user has identified the unknown spectrum, a PDF report can be generated with a single click.

Training Files Used in This Lesson

C:\Users\Public\Documents\Wiley\KnowItAll\Samples

- \ID Expert\IR\2 - ATR-IR of Unknown Sample 8675309.irf
- \ID Expert\IR\4 - ATR-IR of Unknown Sample 1282013.irf
- \Deformulation\Deformulation Example.irf
- \Optimized Corrections\Raman Spectrum of Mint Candy.wdf

KnowItAll Applications Used

- KnowItAll ID Expert

KnowItAll IR and Raman Search Algorithms

A background in the algorithms used by KnowItAll will be beneficial. For IR and Raman spectral comparison, KnowItAll uses the following algorithms:

Correlation

This is the default algorithm for searching in KnowItAll and it conforms to the industry standard for correlation algorithms. The Correlation algorithm is similar to the Euclidean Distance algorithm. The difference between the two is in the way the spectra are treated before the comparison. Each spectrum is mean-centered prior to performing the dot product normalization. This approach can improve search results for noisy spectra and spectra that have baseline issues, particularly with a baseline offset that is the result of a negative spike or chemical noise. It is slightly more time-consuming than the Euclidean Distance algorithm. The search speed is slower because each spectrum in the database must be mean centered and then normalized prior to the comparison. The search results that are obtained with the Correlation algorithm are spectrally similar to the unknown, even if the unknown compound is not in the database. The Correlation algorithm is heavily weighted by peak areas. Broad features are weighted much more strongly than sharp features. The algorithm is most tolerant to peak shifts and non-linearities in relative band intensities.

Correlation (Classic)

The Correlation algorithm that was found in all versions of KnowItAll prior to KnowItAll 2020 is similar to the Euclidean Distance algorithm. However, it did not conform to the industry standard for correlation algorithms. Beginning with KnowItAll 2020, the Correlation algorithm does conform to the industry standard and it is the default algorithm used for searching in KnowItAll. To provide backward compatibility for customers who want to reproduce prior search results, the previous correlation algorithm is now provided as Correlation (Classic).

Euclidean Distance

The Euclidean Distance algorithm measures the point-to-point differences between a pair of spectra. The results that are obtained with the Euclidean Distance algorithm are spectrally similar to the unknown, even if the unknown compound is not in the database. This algorithm, however, can yield degraded search results when the unknown spectrum has a sloping or offset baseline. The Euclidean Distance algorithm is heavily weighted by peak areas. Broad features are weighed much more strongly than sharp features. This algorithm is most tolerant to peak shifts and non-linearities in relative band intensities.

First Derivative Euclidean Distance

Use this algorithm to reduce the effects of baseline slope or offset in the unknown. Although search speed is slightly slower than with the Euclidean Distance algorithm, the First Derivative Euclidean Distance sometimes gives improved search results, especially when the unknown spectrum is a mixture of two or more compounds. The First Derivative Euclidean Distance algorithm is heavily weighted by changes in slope. Sharp features are weighted much more strongly than broad features. The algorithm is also very sensitive to peak shifts. Small shifts can make the algorithm miss a similar result.

Second Derivative Euclidean Distance Use the Second Derivative Euclidean Distance algorithm to compare the second derivative of a reference spectrum to that of the query spectrum.

Optimized Corrections: A Breakthrough Technology for Spectral Searching

Spectral searching is one of the most important tools scientists use to classify or identify materials, yet it continues to be plagued by errors and imperfections. During a spectral search, a sample spectrum is compared to a database of reference spectra. To ensure that an optimal match is found in the database, spectra can be adjusted to compensate for differences between spectra caused by variability in instruments, accessories, environmental conditions, and other factors.

According to ASTM's guide on spectral searching[1], various algorithms and manual methods exist to adjust spectra to get reasonable match scores when two compared spectra of the same compound differ for various reasons. While these methods may work in selected cases, subtle discrepancies such as a shift of the X-axis are very hard to identify and correct manually. The inflexible mathematical algorithms typically employed do not compensate for these types of errors in spectra that are flawed.

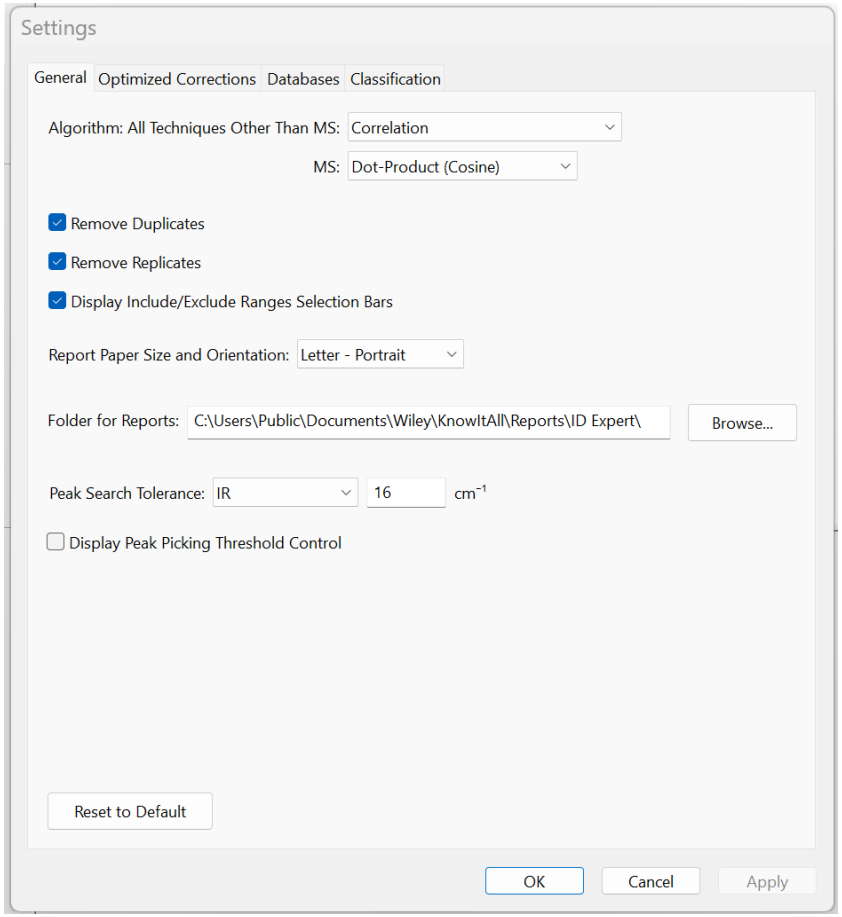
Manual corrections can be made by expert spectroscopists, but those less experienced in spectroscopy are often unaware of how to perform the necessary corrections on their sample spectrum to achieve the best search result. To address this growing concern, Wiley has introduced Optimized Corrections, a breakthrough patented technology that performs a computationally complex set of multiple corrections on query and reference spectra in a search to find the optimal match between the query and each individual reference spectrum. This training guide will demonstrate how the Optimized Corrections technology yields better match between query and reference spectra than can be attained using rigid search algorithms alone or with manual methods to optimize spectra for searching.

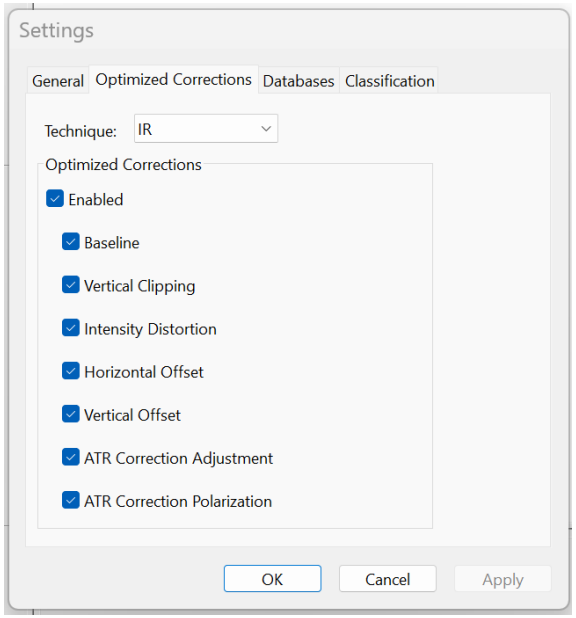
Optimized Corrections consider the full spectrum during a selected range(s) search.

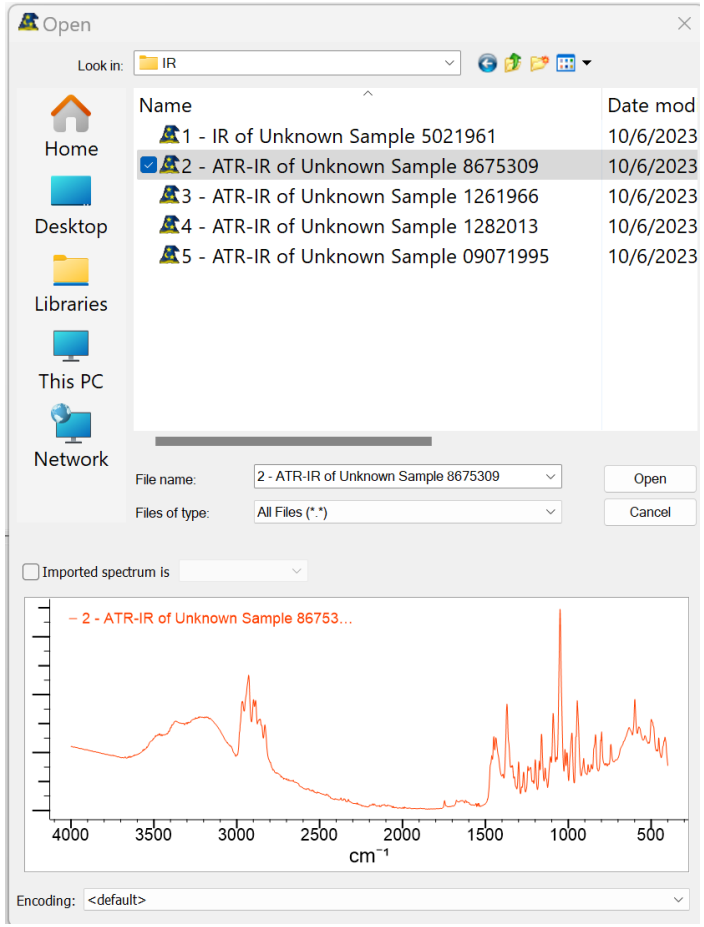
[1] E2310-04 - Standard Guide for Use of Spectral Searching by Curve Matching Algorithms with Data Recorded Using Mid-Infrared Spectroscopy, 2009. ASTM International Web Site. http://www.astm.org/Standards/E2310.htm (accessed March 4, 2015).

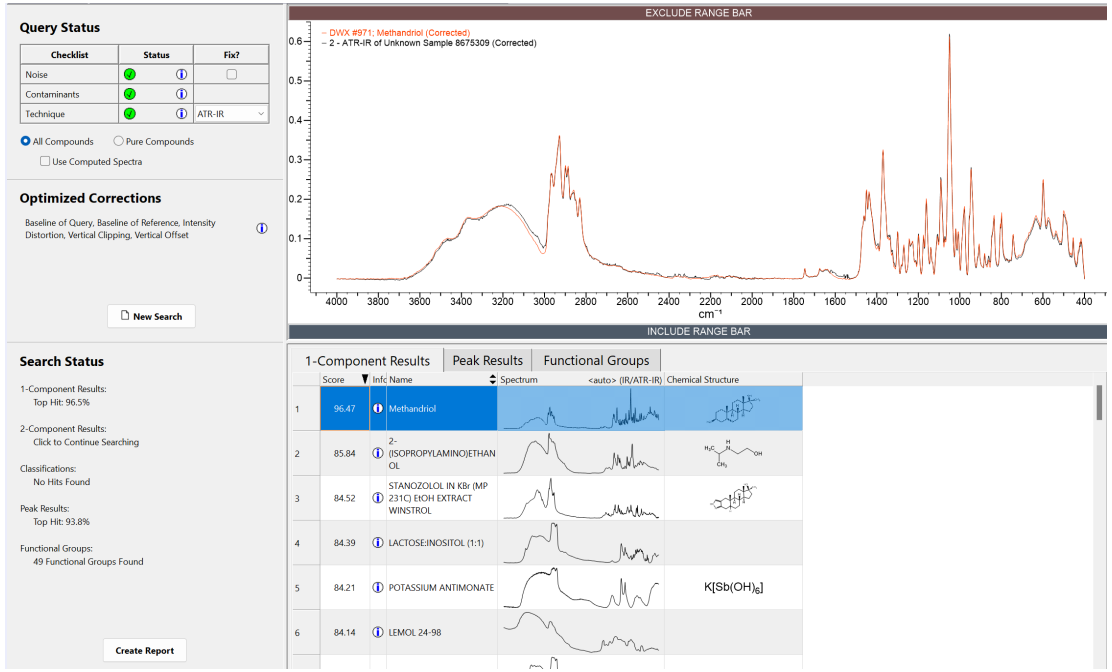
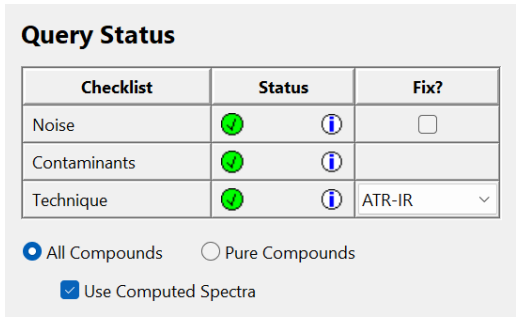
Example 1: 2 - ATR-IR of Unknown Sample 8675309.irf

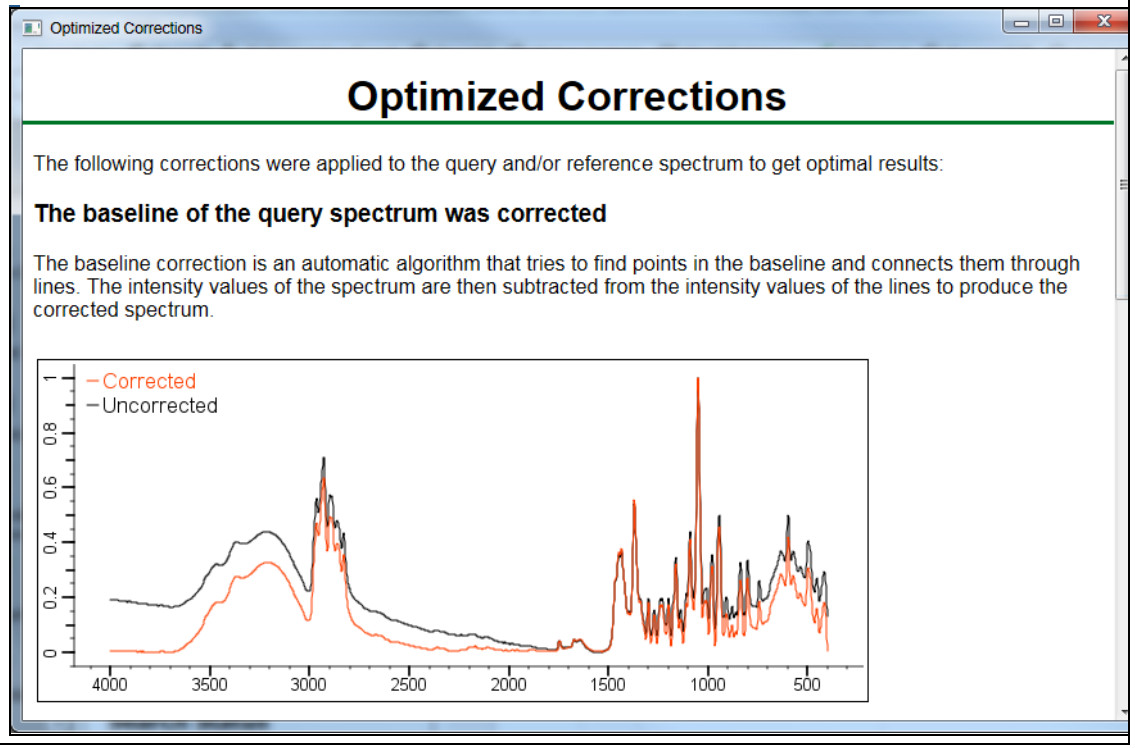
	Action	Result
1	Navigate to the **Data** toolbox and open the **ID Expert** application by clicking the **ID Expert** icon. Alternatively, if the desktop (standalone) application is installed, **ID Expert** can be opened directly by double-clicking on the desktop icon.	The application opens and a Windows **Open** dialog box displays.

	Action	Result
2	Close the **Open** dialog box, then choose **File > Settings**.	The **Settings** form opens. 

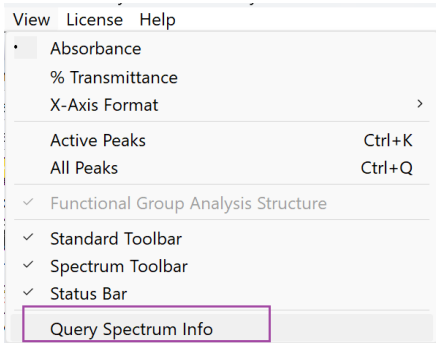
	Action	Result
3	Select the **Optimized Corrections** tab.	 By default, **Optimized Corrections** is enabled for IR, Near IR, Raman, and NMR spectra. With the **Optimized Corrections** checkbox enabled, you can specify which corrections you wish to apply for each spectral technique.
4	Enable all **Optimized Corrections** for IR. Click **OK**.	The **Settings** form closes.

	Action	Result
5	Click **New Search** in the middle left of the window. Navigate to **C:\Users\Public\Documents\Wiley\KnowItAll\Samples\ID Expert\IR** folder Select **2 - ATR-IR of Unknown Sample 8675309.irf**. 'Imported spectrum is' should be checked. Choose **ATR-IR** from the dropdown menu. Click **Open**.	

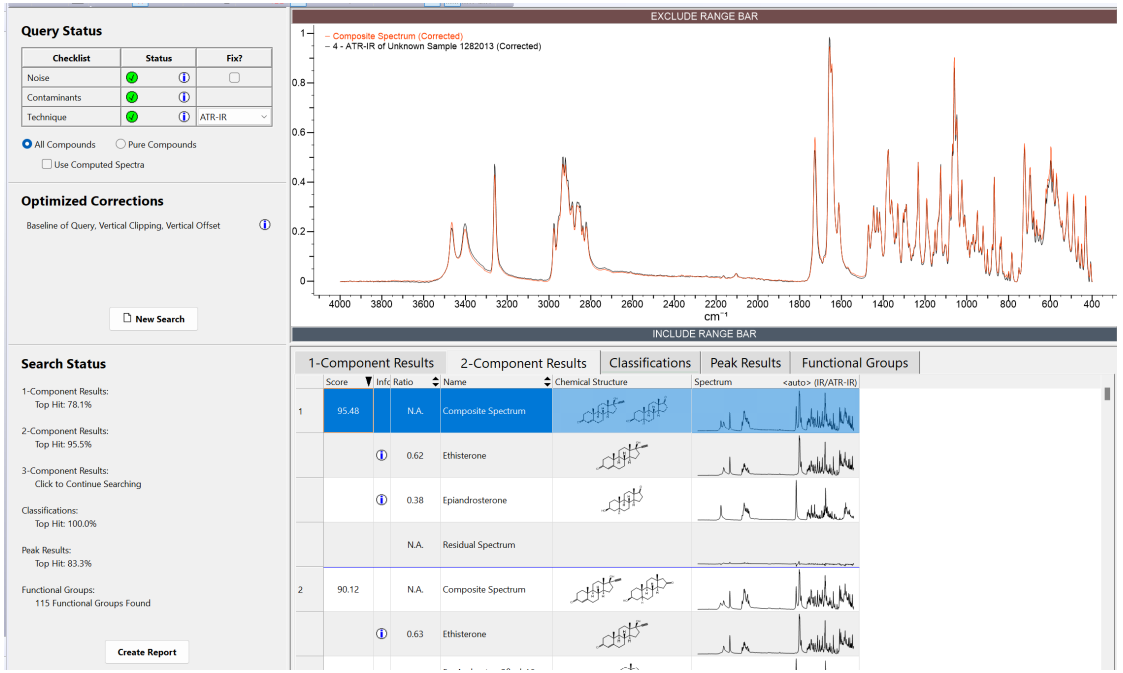
	Action	Result
6	Wait for Search Status bars to show that the search is complete.	The search proceeds automatically.  The **Score** is the Hit Quality Index (**HQI**) in 100% unit – a comparison of reference spectrum to query spectrum. ***Note***: Hit list might be different due to continuous addition of reference data.
7	If **SmartSpectra IR data** is included in one's license, it can be selected by clicking **Use Computed Spectra** so that licensed SmartSpectra data is also searched. Deselect **Use Computed Spectra** for this example.	

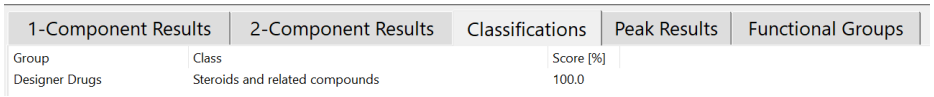
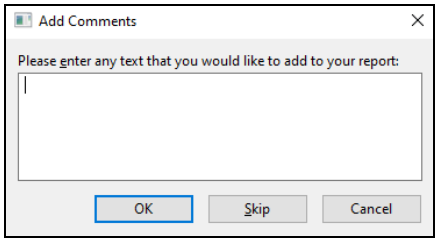
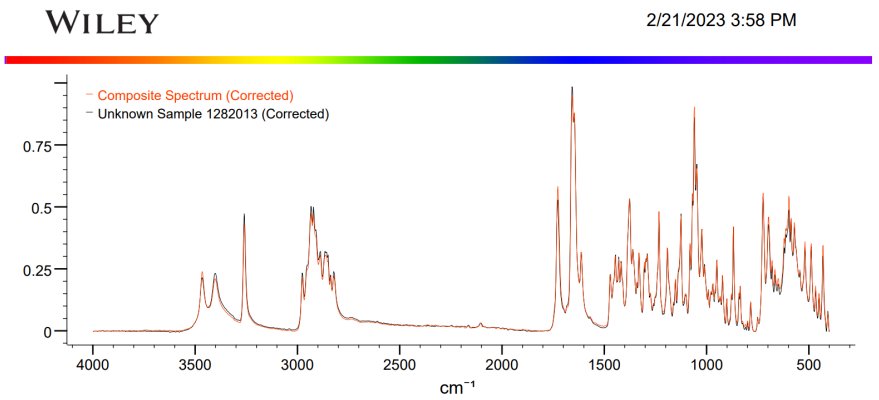
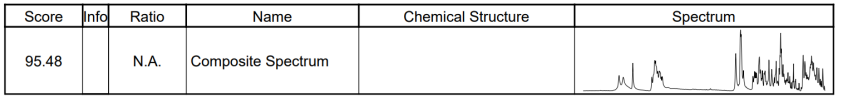
	Action	Result
8	Click the information icon ⓘ in the main window's **Optimized Corrections** section. Scroll down to view information about the optimized corrections that were applied. Close the window when you have finished reviewing the information.	The **Optimized Corrections** window opens. 

Note: one can view the metadata of query spectrum by View > Query Spectrum Info



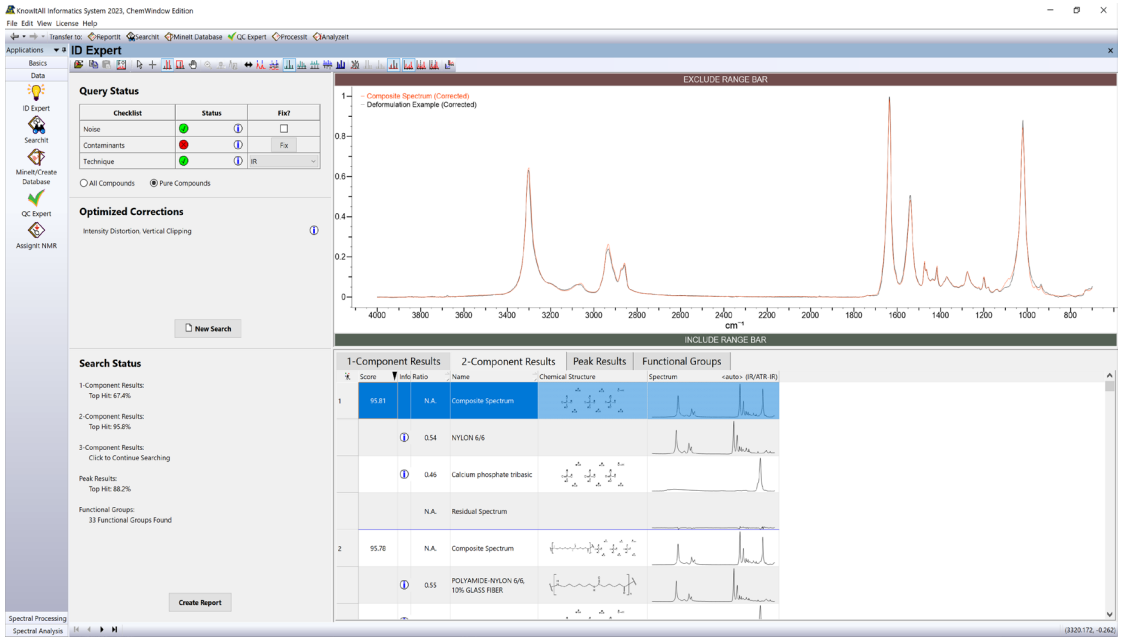
Example 2: 4 - ATR-IR of Unknown Sample 1282013.irf

	Action	Result
1	Click **New Search**.	A Windows **Open** dialog box displays.
2	Navigate to **C:\Users\Public\Documents\Wiley\KnowItAll\Samples\ID Expert\IR**. Open the spectral file **4-ATR-IR of Unknown Sample 1282013.irf**.	The search proceeds automatically. There is no excellent 1-Component match, and KnowItAll ID Expert automatically examines multiple component matches.
3	Allow the search to complete.	The **2-Component Results** tab under **Search Status** flashes indicating that a good match has been found.
4	Click the **2-Component Results** tab.	The Composite Spectrum, individual component spectra, and Residual Spectrum are displayed. 

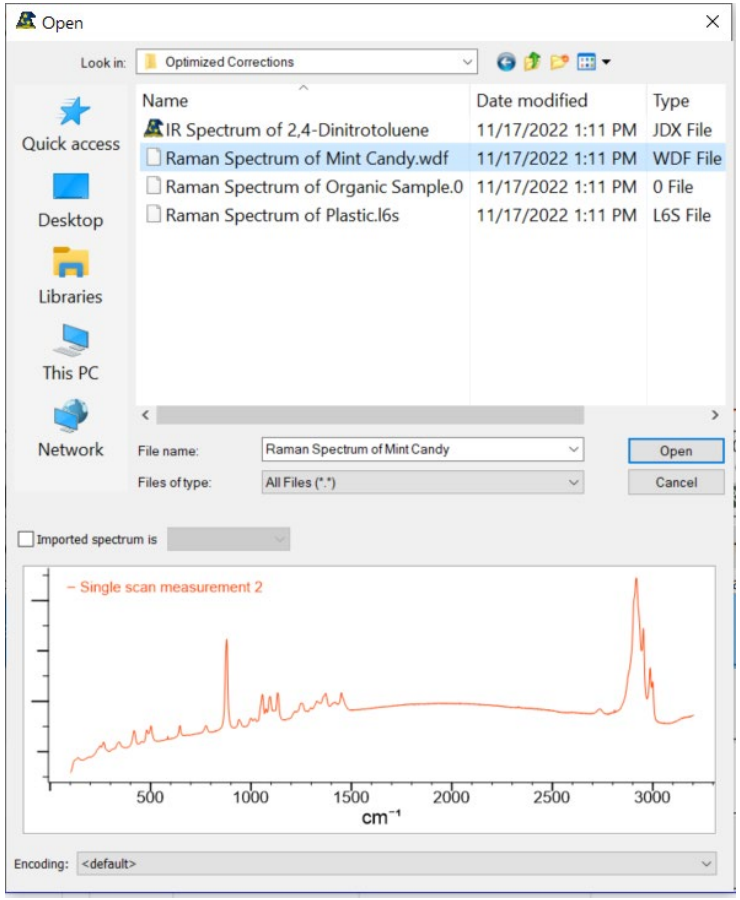
	Action	Result
5	Click the **Classification** tab.	 We can see that this sample is correctly classified by the predictive models in KnowItAll.
6	Click **Create Report** in **Search Status** section.	The **Add Comments** dialog box opens. 
7	Click **OK**.	The report is created and automatically opens.  Manual Corrections: None Ranges: Full Search Algorithm: Correlation Query Path: C:\Users\Public\Documents\Wiley\KnowItAll\Samples\KnowItAll_2023_Example_Training_Files\02 - Simple Spectral Search\ID Expert - IR\4 - ATR-IR of Unknown Sample 1282013.irf 

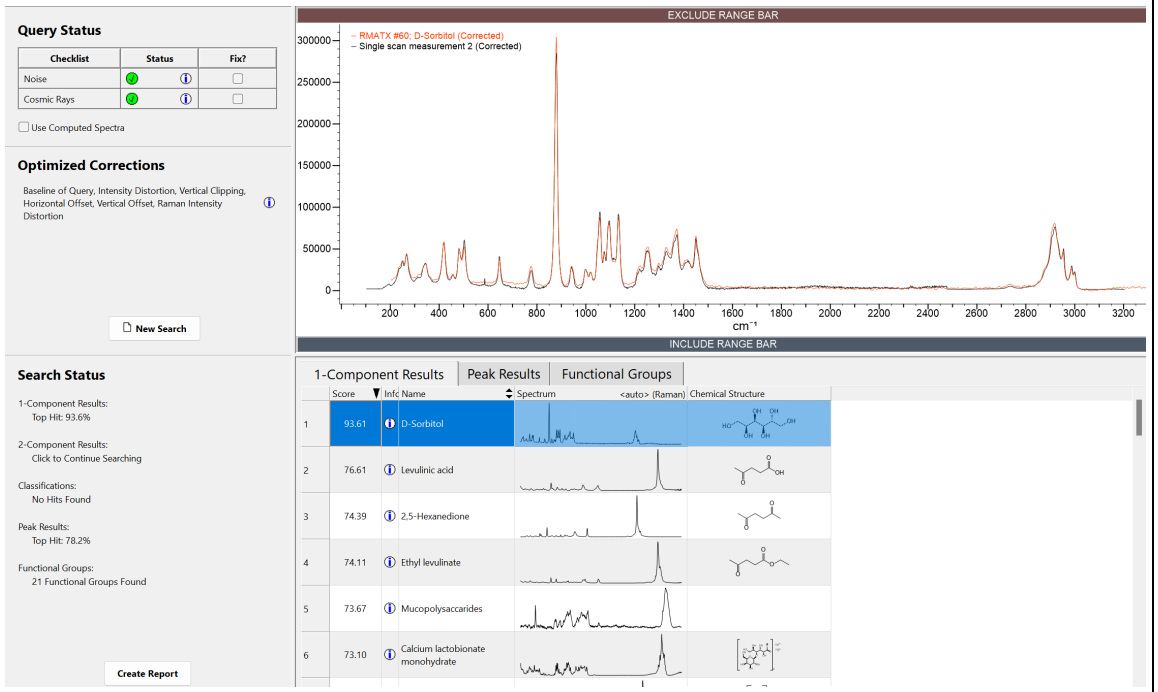
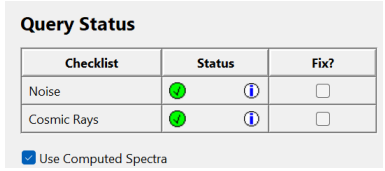
Example 3: Deformulation Example.irf

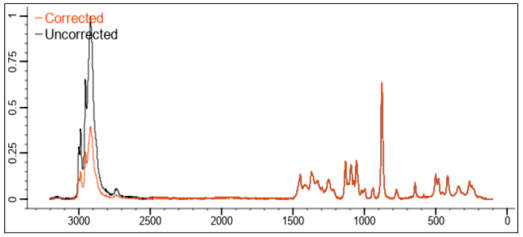
	Action	Result
1	Click the **All Compounds** radio button in the **Query Status** section. (Depending on the previous action taken in ID Expert, the start page may differ. If you do not see the **All Compounds** radio button, close and reopen KnowItAll, then navigate to ID Expert and close the **Open** dialog box. You will then be able to select **All Compounds**.) Click **New Search**.	A Windows **Open** dialog box displays.
2	Navigate to **C:\Users\Public\Documents\Wiley\KnowItAll\Samples\Deformulation**. Select **Deformulation Example.irf**. Click **Open**.	The search proceeds automatically. A perfect match is found: AKOLOUN S223-HM8. However, it is unclear what this material is made of.
3	Click the radio button **Pure Compounds**.	ID Expert performs another search, but only on pure organic and inorganic chemical spectra. The 2-Component Results tab under Search Status flashes indicating that a good match has been found.

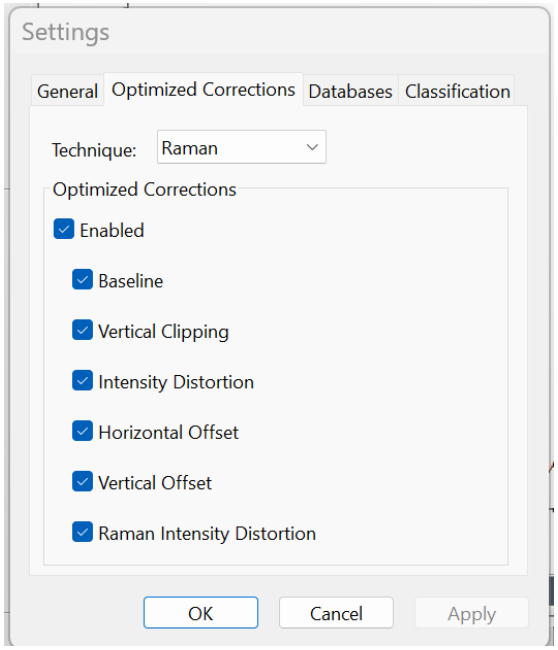
	Action	Result
4	Click the **2-Component Results** tab.	The Composite Spectrum, individual chemical component spectra, and Residual Spectrum are displayed. The sample is possibly made up of NYLON and Calcium salt. 

Example 4: Raman of Mint Candy.wdf

	Action	Result
1	Click **New Search**.	A Windows **Open** dialog box displays.
2	Navigate to **C:\Users\Public\Documents\Wiley\KnowItAll\Samples\Optimized Corrections** folder. Select the spectral file **Raman Spectrum of Mint Candy.wdf**. Click **Open**.	From the preview, this is not a "good" spectrum. 

	Action	Result
3	Allow the search to complete, then scroll through the **Search Status** area.	We have a very good match. 
4	If SmartSpectra Raman data is included in one's license, one can check "**Use Computed Spectra**" so that Raman SmartSpectra participates in search and mixture analysis. Deselect **Use Computed Spectra** for this example.	 Note: From KnowItAll 2025 onwards Cosmic Rays detection is performed and corrected upon the user's request.

	Action	Result
5	Click the ⓘ for **Optimized Corrections**. Scroll through the information, the middle of this page shown on right. When finished, close out of the pop-up window.	**The Raman-specific intensity distortions were corrected by applying an adjustment factor of 60.4% to all regions above 2474.9 cm^{-1}** According to the ASTM E2911-13 Standard Guide for Relative Intensity Correction of Raman Spectrometers, "Generally, Raman spectra measured using grating-based dispersive or Fourier transform Raman spectrometers have not been corrected for the instrumental response (spectral responsivity of the detection system). Raman spectra obtained with different instruments may show significant variations in the measured relative peak intensities of a sample compound. This is mainly as a result of differences in their wavelength-dependent optical transmission and detector efficiencies. These variations can be particularly large when widely different laser excitation wavelengths are used, but can occur when the same laser excitation is used and spectra of the same compound are compared between instruments." To compensate for these intensity variations when comparing Raman spectra that may not have been corrected for instrumental response, the following equation introduces an adjustment factor R and is applied to all spectral regions above the wavenumber value mentioned above: $I_c = I \cdot (100\% - R)$ where I_c is the corrected intensity, and I is the original intensity.  We see that all kinds of spectrum corrections have been done for the first hit. One unique to KnowItAll is the "Raman-specific intensity distortion."

	Action	Result
6	Go to **File > Settings**, **Optimize Correction** tab.	These are the corrections done for Raman sample and reference spectra.  Note on **peak clipping**: KnowItAll makes a conscious effort to reduce the intensity of the strongest peaks.
7	As an exercise, you can check off **Enabled** and repeat the search to see a very different result.	

Note: In the KnowItAll 2026 release, ID Expert can be run against licensed databases rather than all reference databases. Databases to be searched can be specified by:

- Going to **File > Settings.**
- Under the **Databases** tab, check **Search Only License Databases** and click on **Apply.**