

ACD/LABS [ADVANCED CHEMISTRY DEVELOPMENT, INC.]

## A Streamlined Workflow for the Rapid Identification of Unknowns

Anne Marie Smith

Brent G. Pautler

Richard Lee

Joe DiMartino

### Introduction

When performing any type of screening assay, one of the most significant analysis challenges lies in the identification of unknown components. While high resolution and high mass accuracy enables greater precision when predicting elemental composition, these estimates lack structural information. When faced with a true unknown, even the fragment ion analysis has limitations and can only highlight particular functional groups at best.

In this work, we present a workflow that utilizes the available analytical information, such as predicted elemental formula, fragment ions, and the inclusion/exclusion of known fragments to refine a mass-based local library search of PubChem. Combining these parameters provides an accumulative filter that reduces search results to a manageable number, or in the best-case scenario, a single structure.

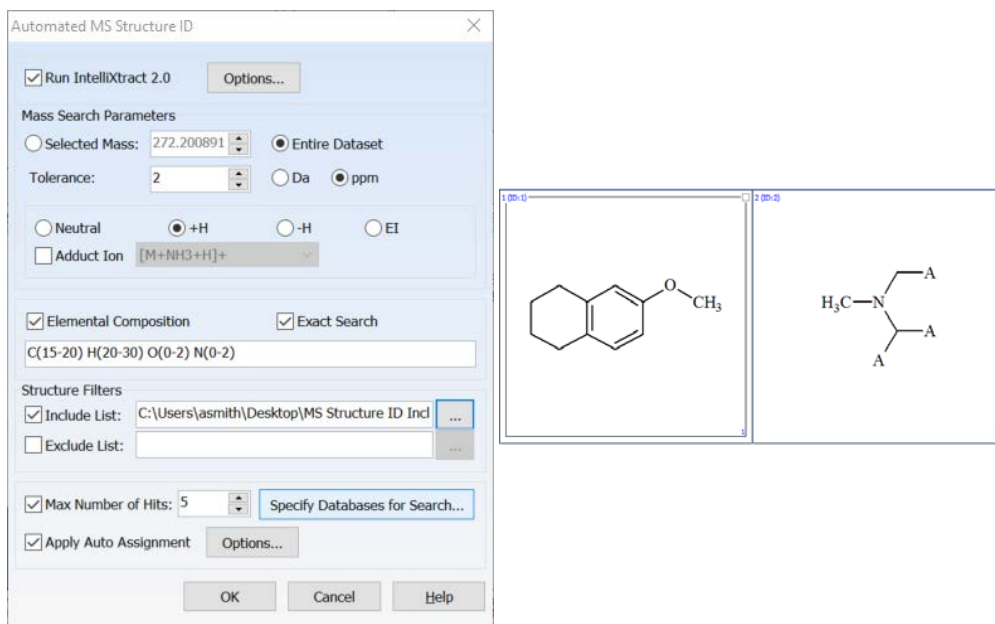
The workflow highlighted in this paper uses an unknown pharmaceutical sample as a test case, but the software may be applied to samples from any number of sources such as food & beverages, formulated products, metabolomics, and environmental analyses.

We demonstrate how MS Structure ID Suite can be used to help you identify known unknowns in your samples.

## Identification of Unknown Components

An example pharmaceutical sample was analyzed by HPLC-FTMS (Thermo Fisher Scientific Orbitrap equipped with an electrospray ionization source). The main peak of interest occurs between 5.05 and 5.15 min with a monoisotopic mass of 272.2011 m/z. In this sample, additional components are expected to be related impurities. The 272.2011 m/z component (RT ~5.1 min) is used to determine a possible elemental composition range for a structure search. The isotopic pattern suggests that only C, H, N, and/or O atoms are present in the molecular formula. The software also suggests C<sub>18</sub>H<sub>25</sub>NO as the best fit based on the accurate mass.

The software will componentize your entire dataset and automatically search by mass in PubChem with a specified mass tolerance (high res or low res data) to provide a list of candidate structures. (PubChem is a commercial library that contains >100 million compounds, available through ACD/Labs). The final review list can be simplified as part of the process through elemental composition and structure filters (Figure 1).



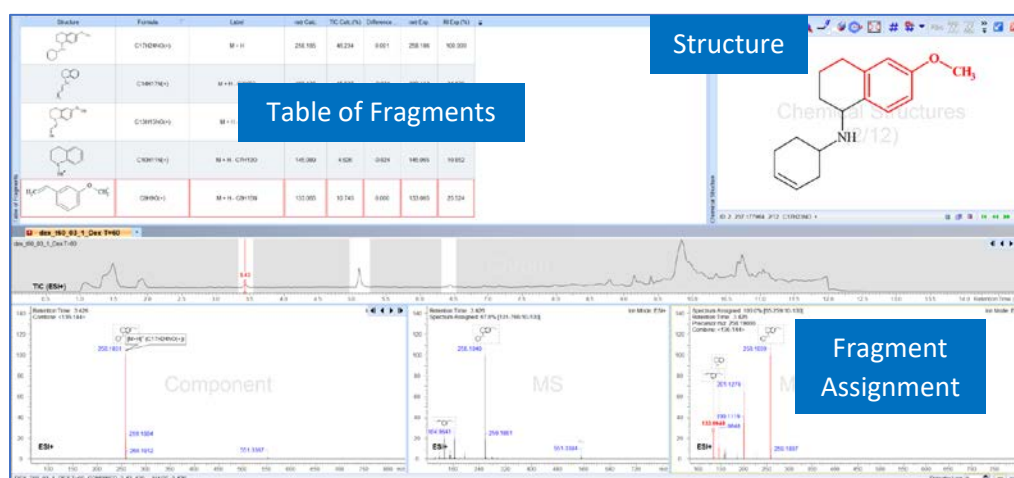
**Figure 1:** Left: Settings for a mass search allow you to review a manageable list of structure results through settings for elemental composition and structure filters. Right: Example structure filters for a simplified candidate list for each component

In this example, hits are found for all selected areas of the chromatogram (highlighted white in Figure 2) within the elemental composition range and a 2 ppm tolerance in PubChem.



**Figure 2:** Structure results for the given compound in MS Structure ID Suite. Users can select the accurate structure and update the Table of Components.

As a result of the stringent criteria, the 5 results gave the same molecular formula and thus the same Formula Fit Quality (FFQ—a measure of how well the calculated and experimental isotope clusters of the expected molecular ion match). Auto Assignment allowed us to look at the d(MS) score to see how the spectral fragments match the structure. In two cases we achieve a score of 1.000. These are the top hits. By transferring a hit to the Table of Components and running Auto Assignment once more, separate from the structure search, structural candidates can be reviewed more closely.



**Figure 3:** Results from Auto Assignment of a structure to the spectrum. A minimum abundance of 10% was required for assignment leading to a 100% assignment score for the MS2 spectra. Analyze the Table of Fragments or the assignment on the spectra to determine if the structure is correct.

## Conclusions

This example demonstrates how the MS Structure ID Suite is able to quickly search a wide range of potential structures, curate a reasonable and relevant list of candidates, and identify the most likely structure for a chromatographic peak, all in an extremely efficient workflow.

To learn more visit [www.acdlabs.com/MSStructureIDSuite](http://www.acdlabs.com/MSStructureIDSuite)