

# LC/MS Unknown Identifications Using MSMS Libraries

## Part VIII Identification of Unknowns with “Spectraless” Libraries

12/27/20

James Little

[tvasailor@gmail.com](mailto:tvasailor@gmail.com)

<https://littlesandsailing.wordpress.com/>

Kingsport, TN

- *Retired\* Research Fellow, Eastman Chem. Co.*
- *42 years experience unknown identification*
- *Now Consultant, MS Interpretation Services*
- *Specialties<sup>1</sup> EI GC-MS, LC-MS/MS, Chemical Ionization,<sup>2</sup> Accurate Mass, Derivatization,<sup>3,4</sup> MS library management, SciFinder, ChemSpider, Surfactant ID,<sup>5</sup> NMR, GC-IR, organic synthesis, matrix ionization effects,<sup>6</sup> etc.*



Eastman Chemical Company, Main Site, Kingsport, TN  
50 Manufacturing Sites Worldwide, ~14,500 Employees

\* [https://en.wikipedia.org/wiki/Eastman\\_Chemical\\_Company](https://en.wikipedia.org/wiki/Eastman_Chemical_Company)



>50 Mass Specs Networked  
Worldwide

## Table of Contents

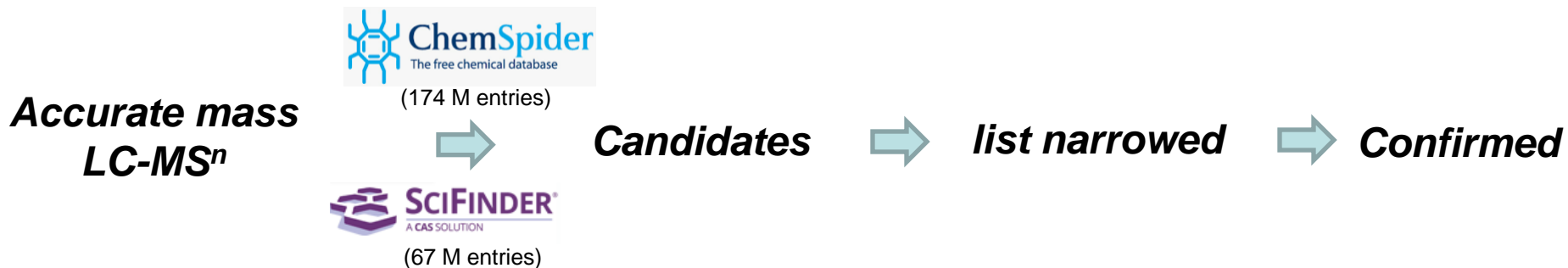
Topic	Slide #
▪ Introduction of Presenter, James Little .....	1
▪ Table of Contents .....	2
▪ Series Topics .....	3
▪ Introduction to Approach .....	4
▪ Origin of Term "Known Unknowns" .....	5
▪ Overall Process Including NIST Library Searches of EI and Tandem Spectra .....	6
▪ ChemSpider Internet Searches .....	7
▪ CAS Registry Using SciFinder and STN Express.....	8
▪ Summary Comparing ChemSpider and CAS Registry Approaches .....	9
▪ Example 1: SciFinder Search of CAS Registry for UV Additive in Polymer .....	10
▪ Example 2: SciFinder Search of CAS Registry for More Obscure Compound .....	13
▪ Example 3: ChemSpider Identification of Additive in Polymer by Monoisotopic Mass ..	14
▪ Live Demo YouTube Presentation .....	16
▪ Presentation References (Internet Links) .....	17
▪ Acknowledgements.....	18

## LC/MS Unknown Identifications Using MSMS Libraries

- Part I: Overview of Software and User Customized Configurations
- Part II: NIST MSMS Search Software and Libraries
- Part III: More Detailed Discussion of MSMS Hybrid Search
- Part IV: Importing MSMS Spectra
- Part V: NIST Structure Searches
- Part VI: MS Interpreter Correlation of Substructure to MSMS Ions
- Part VII: Using and Creating Other MSMS Libraries
- Part VIII: Identification of Unknowns with "Spectraless" Libraries

# Part VIII Identification of Unknowns with “Spectraless” Libraries<sup>8-19</sup>

## Introduction



### Overview:

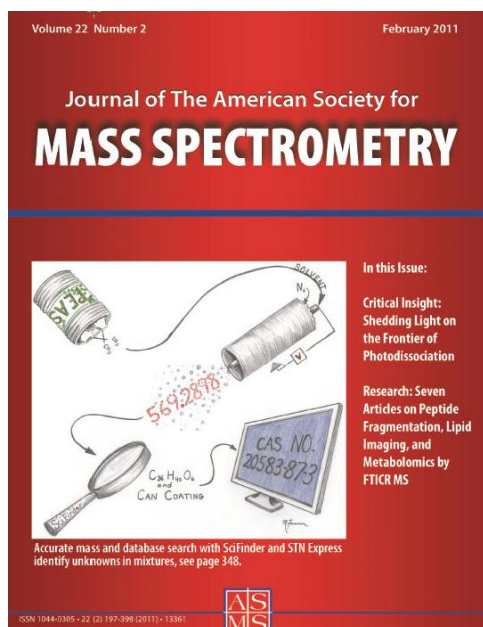
- Need to identify compounds **not found** in Commercial or “Crowd-Sourced” Libraries
- Compounds can be **readily identified** with accurate mass LC-MS<sup>n</sup> data and large databases
- Most useful databases: Scifinder<sup>12-15</sup> (CAS Registry) and ChemSpider<sup>16-18</sup> (Royal Society of Chemistry)
- These are **primarily** “Spectraless” databases
- **Candidate** compounds are **found** by searching either molecular formulae and/or accurate mass (+/- error)
- Candidates **list narrowed** by no. of associated references, key words, substructure, etc.
- **Confirmed** by interpreting MSMS spectrum, no. of exchangeable protons, UV-VIS spectrum, relative retention time, micro-synthesis of mixture, purchasing, etc.

## Origin of Term “Known Unknowns”

- Compounds whose identities are **known** in the world, but **unknown** to us
- Term “known unknowns” **stolen by us** from Rumsfeld<sup>7</sup>

“. . . there are *known knowns*; there are things we know we know. We also know there are *known unknowns*; that is to say we know there are some things we do not know. But there are also *unknown unknowns* -- the ones we don't know we don't know. And if one looks throughout the history of our country and other free countries, it is the latter category that tend to be the difficult ones.”

Donald Rumsfeld, weapons of mass destruction in Iraq News Briefing, Feb. 12, 2002



Reference<sup>14</sup>



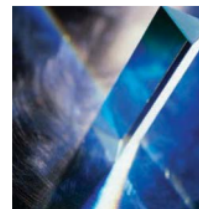
# Part VIII Identification of Unknowns with “Spectraless” Libraries

## Overall Process Including NIST Library Searches of EI and Tandem Spectra

114 LCGC NORTH AMERICA VOLUME 31 NUMBER 2 FEBRUARY 2013

www.chromatograph

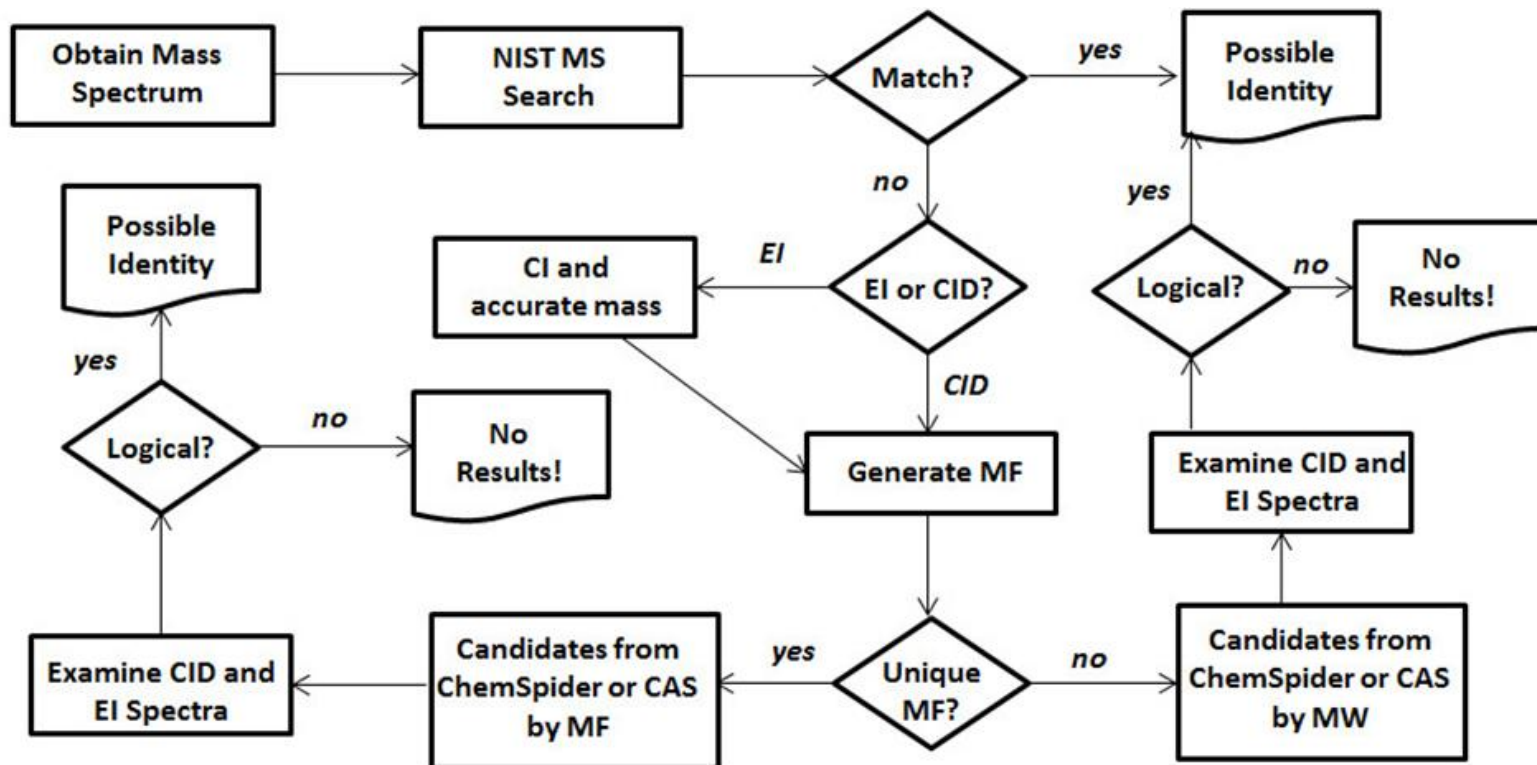
- The SciFinder/ChemSpider is **part** of our overall search
- Total approach shown in flowchart below (1)
- Discussed in LCGC Article<sup>9</sup> (2)



MS – THE PRACTICAL ART

Identifying “Known Unknowns” in Commercial Products by Mass Spectrometry **2**

### 1 Simplified Flowchart for Identifying “Known Unknowns”

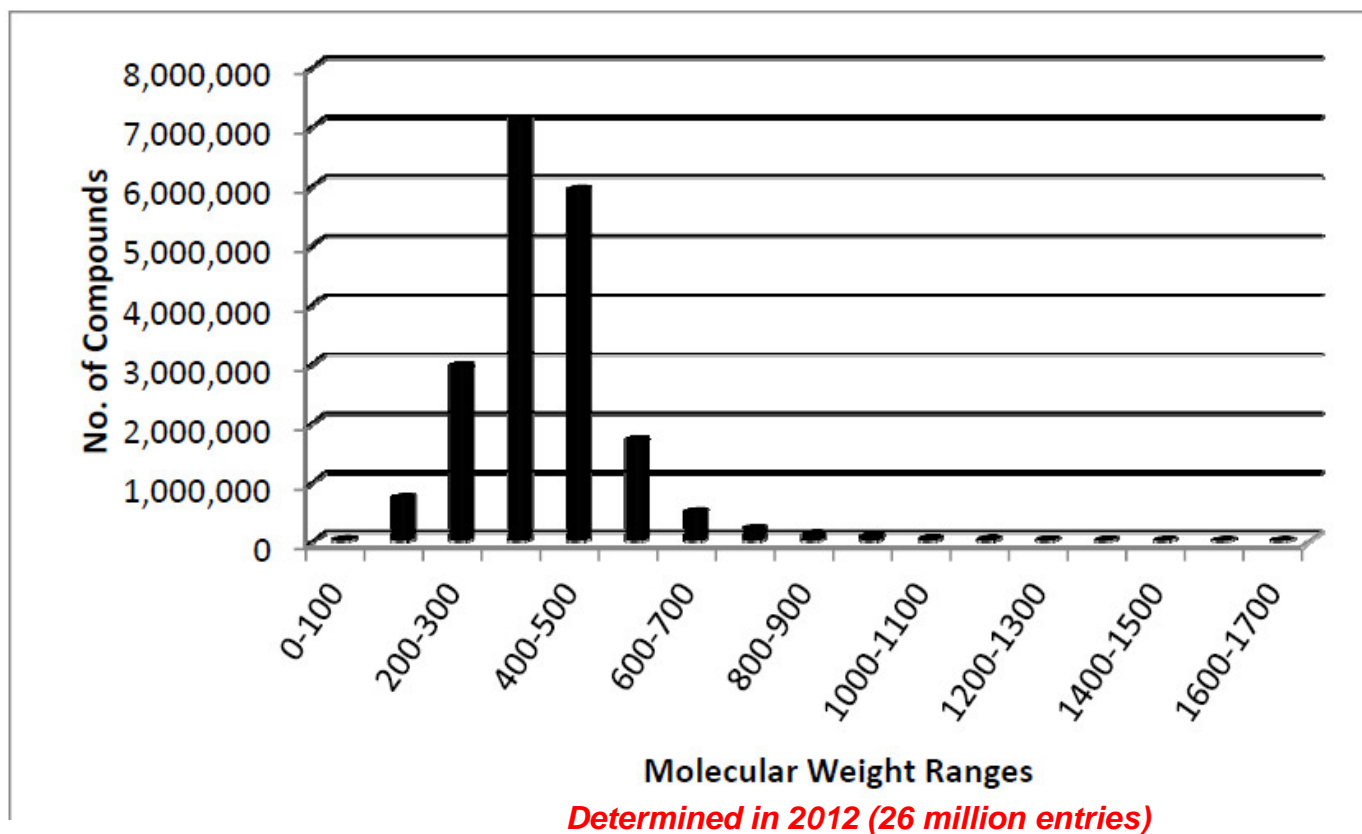


MF = molecular formula; MW = molecular weight; CID = collision-induced dissociation; CI = chemical ionization; EI = electron ionization



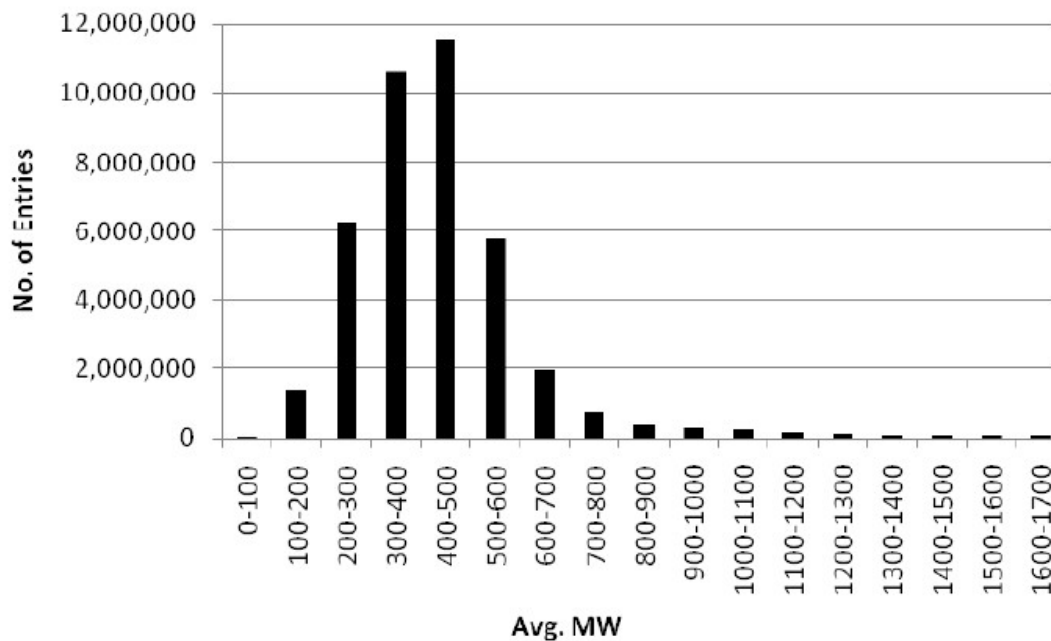
## ChemSpider Internet Searches

- ChemSpider has **67 million** structures from 100's of data sources
- **Free** service with internet interface offered by Royal Society of Chemistry
- **Narrow** candidate list by sorting in descending order by the number of associated references
- **Best method** is to use molecular formula at lower molecular weights
- **Better** to search by *monoisotopic mass at molecular weights > 600* because *difficult to narrow to one specific formula* in that range



## CAS Registry Using SciFinder and STN Express

- CAS Registry with SciFinder has 174 million unique organic and inorganic chemical substances
  - Fee based service from American Chemical Society
  - **Narrow** candidate list by number of references, substructure, key words from sample history, etc.
  - **Best method** is to use molecular formula
  - **Does not** search by monoisotopic mass, only molecular weight using STN Express
  - STN Express is a command base interface, SciFinder much easier to use
  - Error for latter is 4-5 times greater
- 



Average MW distribution of compounds in CAS Registry (Dec. 2009) determined by `s /mw` command in STN Express.



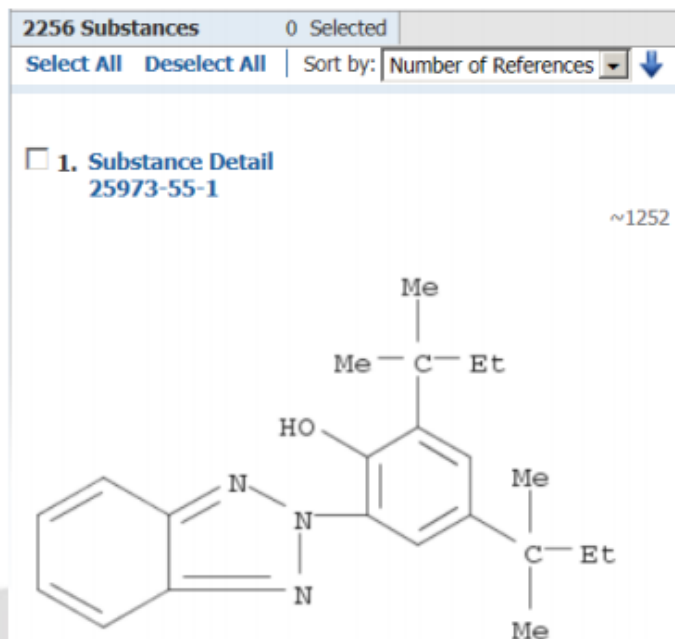
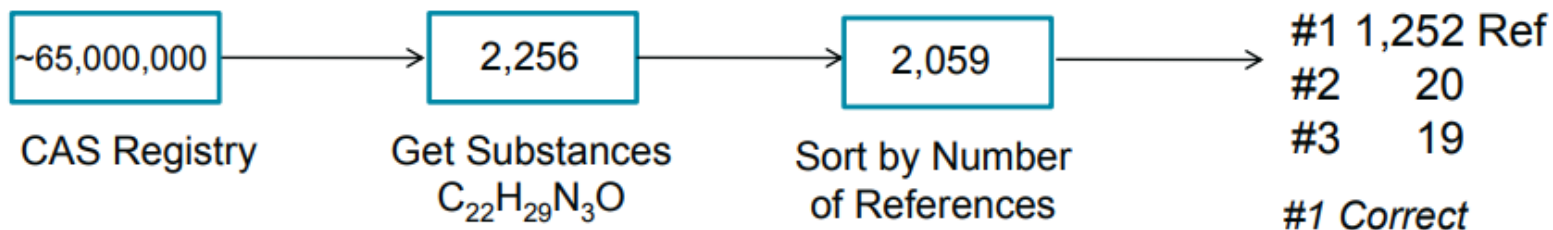
## Summary Comparing ChemSpider and CAS Registry Approaches

Search Approach	Pros	Cons
ChemSpider	<ul style="list-style-type: none"> <li>-free via internet</li> <li>-good user interface</li> <li>-automation by instrument manufacturer using Web API (Application Program Interface)</li> <li>-ability to search monoisotopic mass</li> </ul>	<ul style="list-style-type: none"> <li>-smaller No. of entries (~26 M) and references</li> <li>-can't refine by key word</li> </ul>
CAS Registry with SciFinder or STN Express	<ul style="list-style-type: none"> <li>-larger No. of entries (~65 M) and references</li> <li>-refine by key word</li> <li>-good SciFinder user interface for MF searches on internet</li> </ul>	<ul style="list-style-type: none"> <li>-fee charged</li> <li>-no API available for instrument manufacturer automation</li> <li>-no ability to search by monoisotopic mass, <b>only</b> MW with <b>complicated</b> STN Express interface</li> </ul>

## Example 1: SciFinder Search of CAS Registry

### Search CAS Registry with SciFinder by MF: UV Additive in Polymer

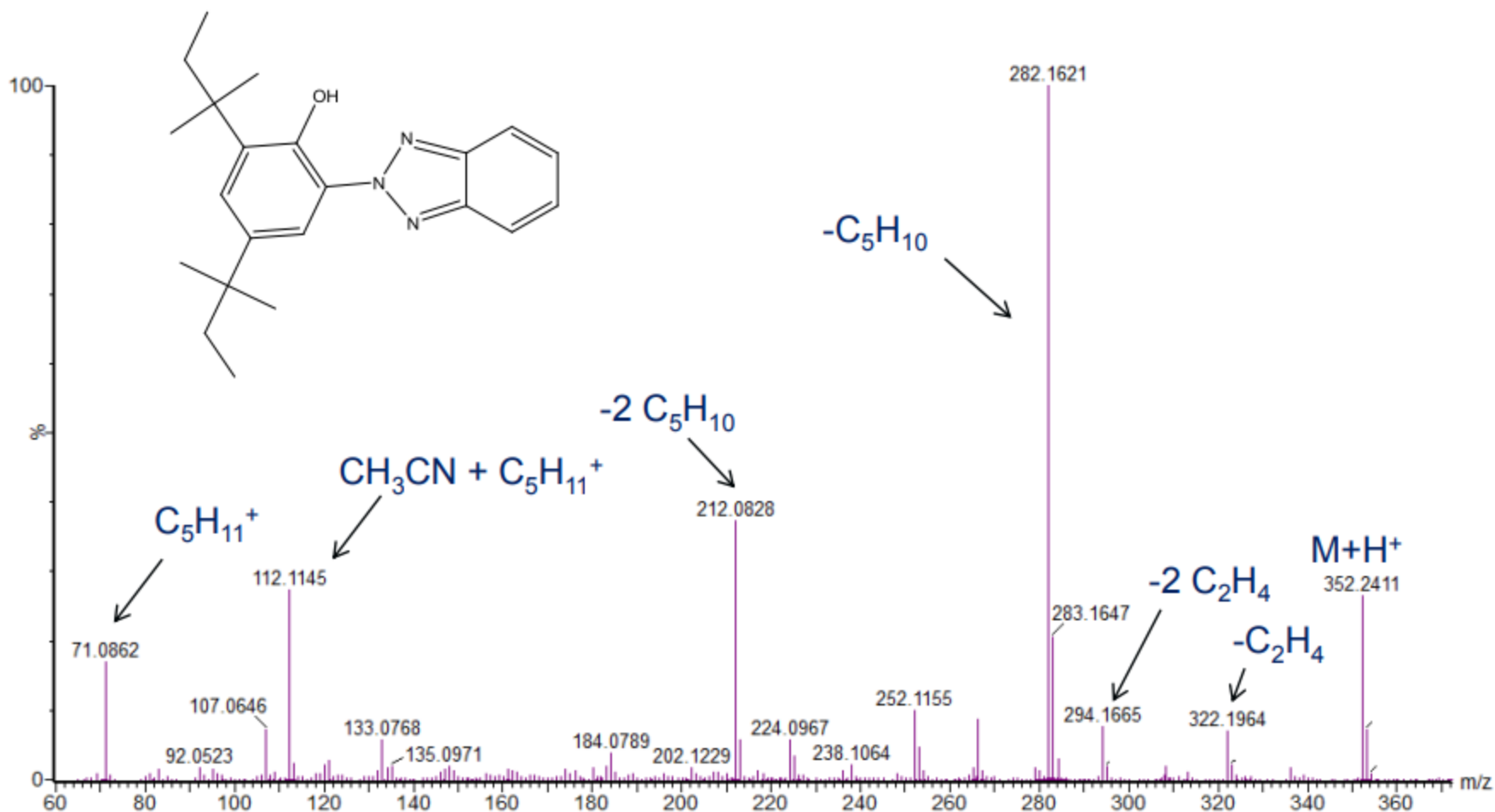
- Request to Identify UV additive in polymer
- Found in LC chromatogram by characteristic UV spectrum
- MF by accurate mass and isotope pattern



# Example 1: SciFinder Search of CAS Registry

## Confirmation of Candidate Structure

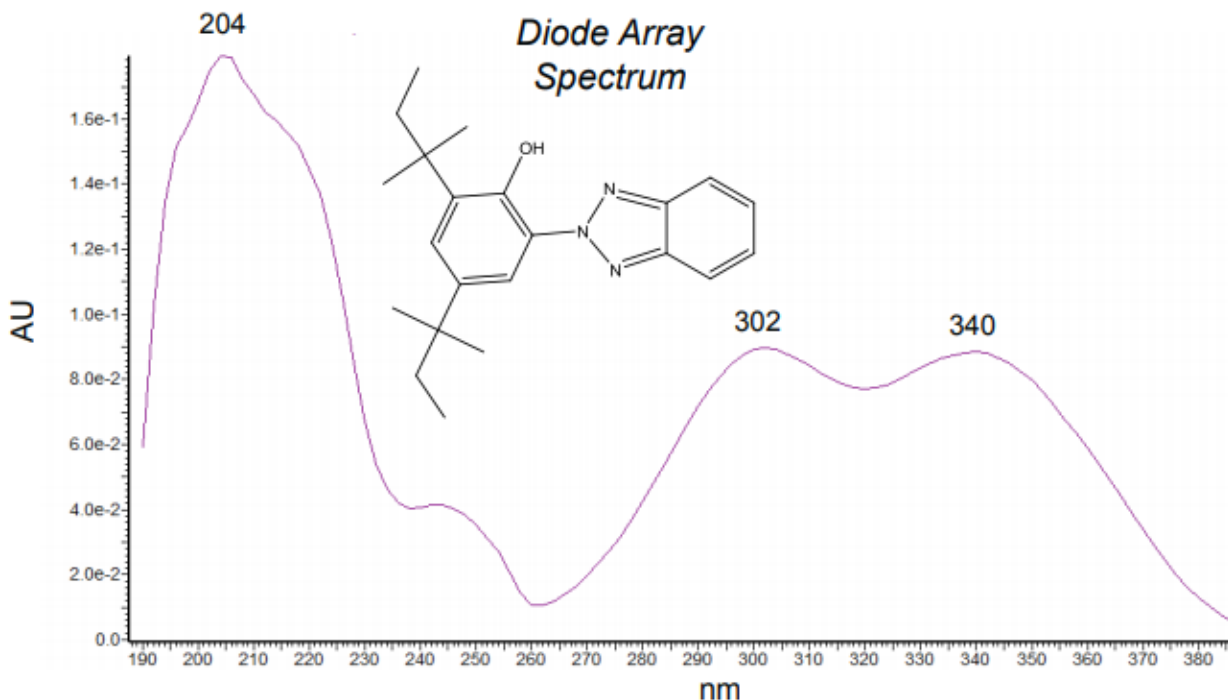
### ➤ Interpretation of MS/MS Spectrum



## Example 1: SciFinder Search of CAS Registry

### Confirmation of Candidate Structure (continued)

- UV spectrum characteristic of UV Absorber

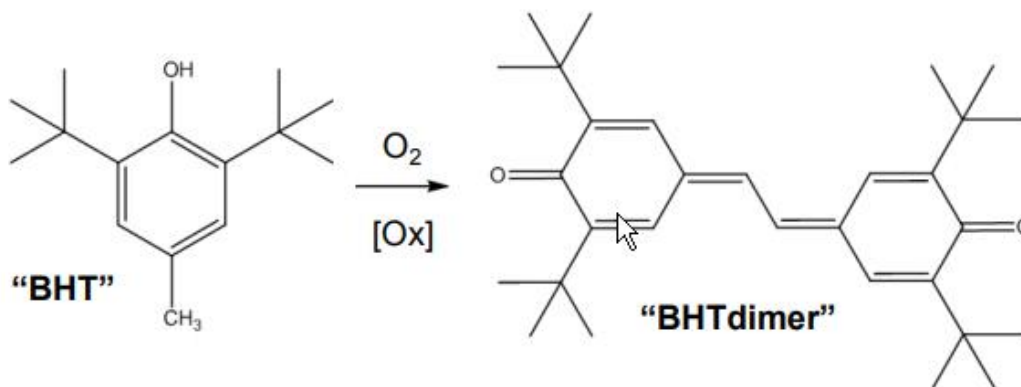
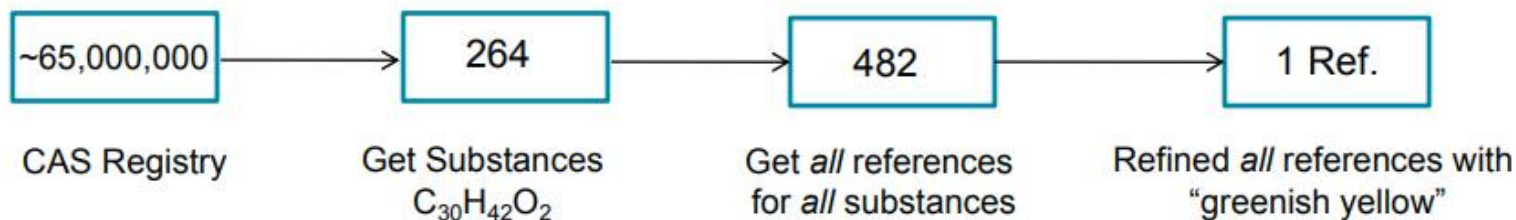


- One exchangeable proton by deuterium exchange *via infusion* in solvent mix with *minimal* D<sub>2</sub>O
- Ultimately confirmed by purchased standard

## Example 2: SciFinder Search of CAS Registry

### MF Search Refined with Key Word for More Obscure Compounds

➤ Identification of *greenish yellow* species in *baby diaper adhesive*



➤ Ironic, BHT included to stabilize, led to color from oxidation

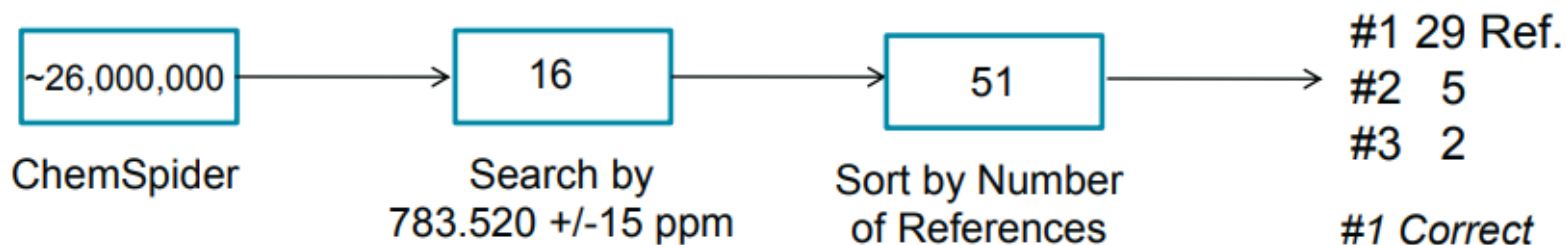


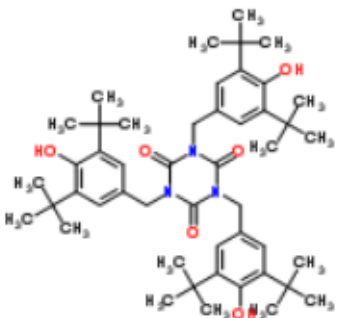
*Problem Solved!*

## Example 3: ChemSpider

### Search by Monoisotopic Mass: Antioxidant in Polymer

- Identification of additive in polymer
- Search ChemSpider by monoisotopic mass

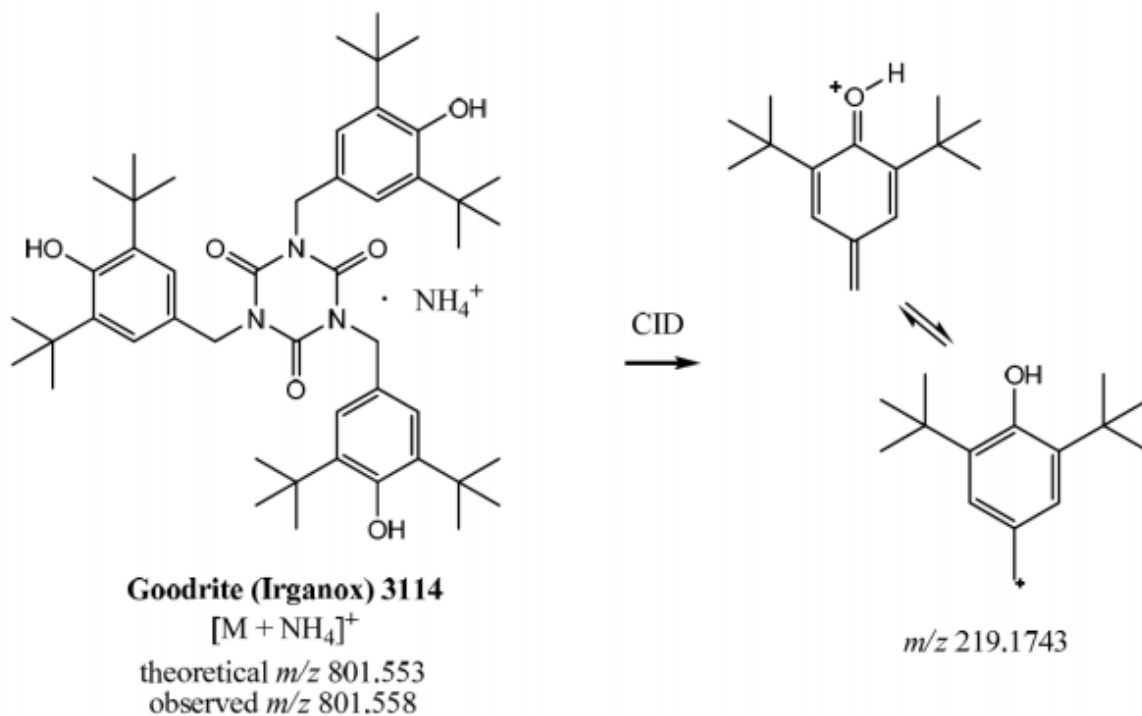


ID	Structure	Molecular Formula	Molecular Weight	# of Data Sources	# of References
<a href="#">84061</a>		$C_{48}H_{69}N_3O_6$	784.078	17	29

## Example 3: ChemSpider

### Confirmation of Antioxidant Candidate Structure

- Initially confirmed by MS/MS fragmentation pattern



- 3 exchangeable protons by infusion in solvent mix with D<sub>2</sub>O
- Ultimately confirmed by purchased standard



***Live Demo*** on YouTube

**LC/MS Unknown Identifications Using MSMS Libraries  
Part VIII Identification of Unknowns with “Spectraless” Libraries**

## Presentation References (*Internet Links*)

1. [James Little Mass Spectral Resource Website](#)
2. [Chemical Ionization for MW Determination](#)
3. [Trimethylsilyl Derivatives for GC-MS](#)
4. [Methyl Ester Derivatives for GC-MS](#)
5. [Surfactant Identification](#)
6. [Lipid Matrix Ionization Effects in LC-MS](#)
7. [Donald Rumsfield "Known Unknown" Quote](#)
8. [Overview on Webpage of "Known Unknown" Identifications](#)
9. [LC/GC Overview Article: Identifying "Known Unknowns" in Commercial Products by Mass Spectrometry](#)
10. [Importance of Mass Spectral Search to Chemical Manufacturers](#)
11. [PittCon Overview Talk on "Known Unknowns"](#)
12. [Accurate Mass Measurements: Identifying Known Unknowns Using Publicly Accessible Databases](#)
13. [Spectraless Databases-CAS SciFinder "Known Unknowns"](#)
14. [CAS/SciFinder Prepress Article "Known Unknowns"](#)
15. [Further Refining CAS SciFinder Search with Substructure, Adam Howard](#)
16. [Accurate Mass Measurements: Identifying "Known Unknowns" Using ChemSpider](#)
17. [Identifying "Known Unknowns" Using ChemSpider and Automated MS/MS Structure Correlation](#)
18. [ChemSpider Prepress Article "Known Unknowns"](#)
19. [YouTube Video: Using MS to Identify Unknowns and Known Unknowns Using Databases and Mass Spectral Software, David Sparkman](#)
20. [NIST Known Unknown Databases in NIST11 Library Release](#)

## Acknowledgements

- **Eastman:** Bill Tindall, Kent Morrill, Curt Cleven, Adam Howard, Jean Coffman, Mike Ramsey, Sen Li
- **ETSU School of Pharmacy:** Stacy Brown
- **CAS:** Anthony Machosky
- **Waters Corporation:** Jim Lekander
- **Agilent Technologies:** Mike Scott
- **Art Work for Journal Cover:** Minta Fannon
- **NIST:** Steve Stein
- **NIST Contractor:** David Sparkman