PMI SCIENCE
PHILIP MORRIS INTERNATIONAL

# High-Resolution Mass Spectrometry in Combination with *In Silico* Prediction Tools to Improve Accuracy for Compound Identification

---

*Environmental & Food MS Meeting - Manchester, February 7th, 2017*

*Eric Dossin, Elyette Martin, Antonio Castellon, Pierrick Diana, Pavel Pospisil, Mark Bentley, Philippe Guy*

*Philip Morris International R&D*

PMI SCIENCE
PHILIP MORRIS INTERNATIONAL

***Reduced-Risk Products ("RRPs") is the term we use to refer to products that present, are likely to present, or have the potential to present less risk of harm to smokers who switch to these products versus continued smoking.***

***We have a range of RRPs in various stages of development, scientific assessment and commercialization. Because our RRPs do not burn tobacco, they produce far lower quantities of harmful and potentially harmful compounds than found in cigarette smoke.***

# Outline

- Challenge for complex matrix characterization
- GC-HR-MS instrumentation
- Linear retention index (LRI) modeling
- Compound identification in complex matrices
- *In silico* prediction (use of fragmentation software)
- Improvement in compound identification
- Conclusion

PMI SCIENCE
PHILIP MORRIS INTERNATIONAL

# Chemical Characterization of Diverse Samples is Challenging

- More than 6'000 chemicals reported as present in tobacco plant and smoke [1]
- Many possible flavor compounds used in e-liquids or smoking products [2,3]


[1] The Chemical Components of Tobacco and Tobacco Smoke, A. Rodgman, T.A. Perfetti, 2013, *2nd Ed. CRC press*.
[2] Leffingwell, J. C.; Young, H. J.; Bernasek, E. Tobacco flavoring for Smoking Products, *R. J. Reynolds Tobacco Company, Winston-Salem,* **1972**.
[3] EFSA flavoring substances database

PMI SCIENCE
PHILIP MORRIS INTERNATIONAL

# Analytical Technique: GC-High Resolution (GC-HR-MS)

**GC-HR-MS_1**
**(7200A Agilent Q-TOF-MS)**

Volatiles and semi-volatiles
LRI from 500 to 1’900

**GC-HR-MS_2**
**(7200B Agilent Q-TOF-MS)**

Apolar and polar
LRI from 1’000 to 3’000

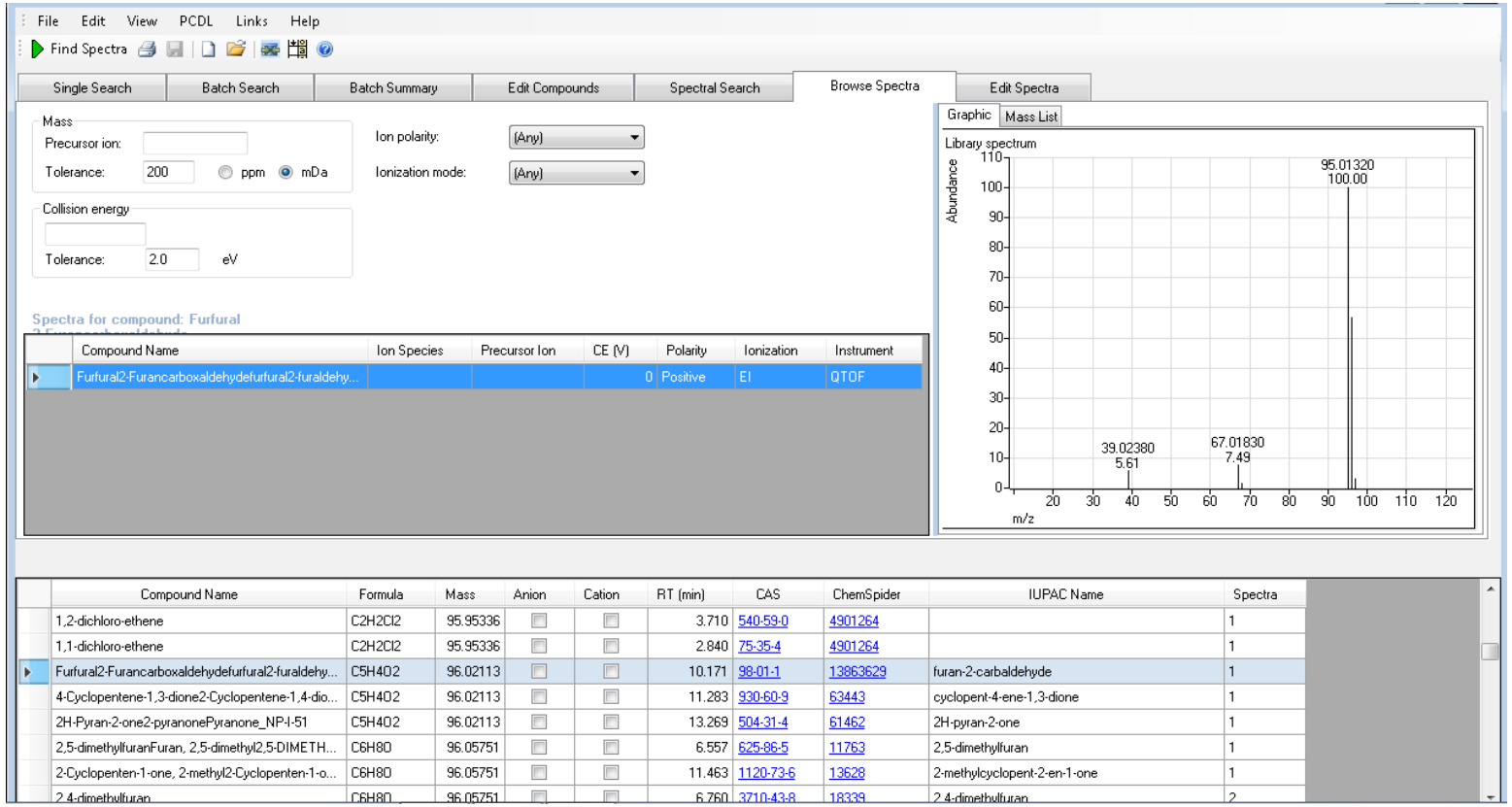
**Goal is to screen the broadest range of smoke constituents**

**Non-Targeted Screening**

PMI SCIENCE
PHILIP MORRIS INTERNATIONAL

# Volatile, Semi-volatile and Polar Compound Library

- 822 reference compounds were analyzed by GC-HR-MS
- Accurate mass spectra associated with their Linear Retention Indices were registered in the Personal Compound Database Library




**An accurate LRI calculation and high quality spectra is key!!!**

PMI SCIENCE
PHILIP MORRIS INTERNATIONAL

# LRI Model Creation

- To predict LRI, two software combination were used:
  1. RapidMiner-Dragon
  2. ACD/Labs ChromGenius

- To build a relevant LRI prediction system 552 molecules were used:
  1. Experimental LRI
  2. Quantitative Structure-Property Relationship (QSPR) and structure similarities

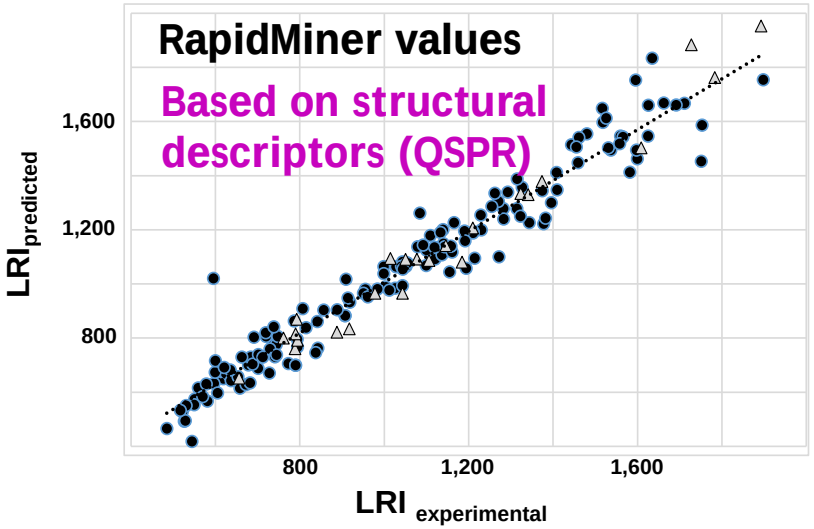
➢ The experimental linear retention indices were randomly split as training (n=401) and test (n=151) sets

➢ Validation set (n=23) confirmed the great performance of both prediction models

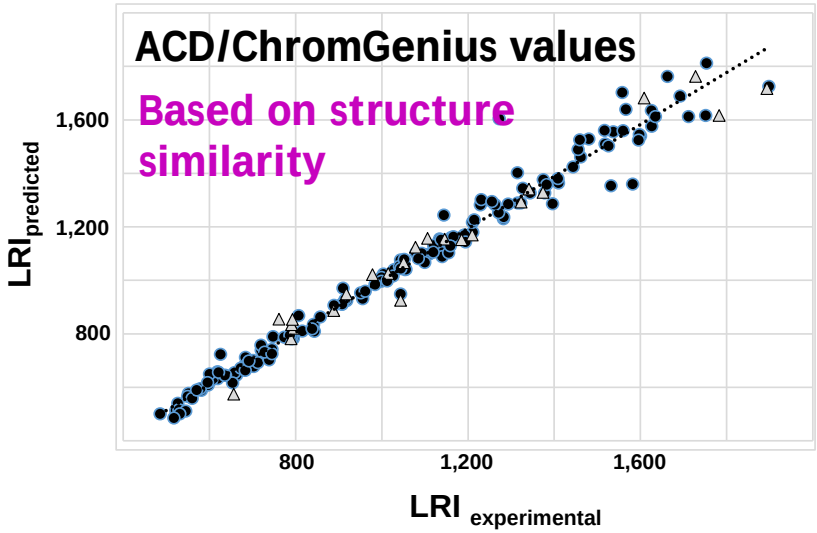
➢ Discrepancy between two LRI predictions can highlight possible errors in the predicted values

PMI SCIENCE
PHILIP MORRIS INTERNATIONAL

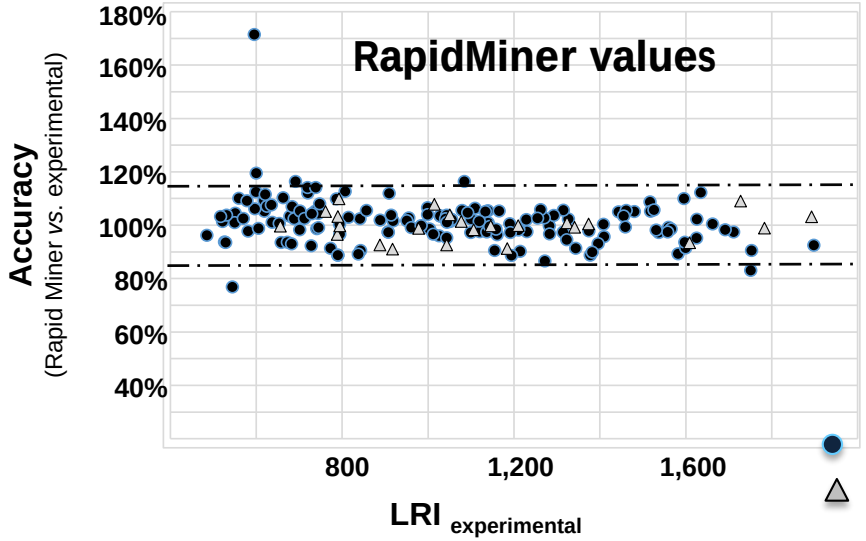
# Assessment of the Prediction Models (Test & Validation Sets)

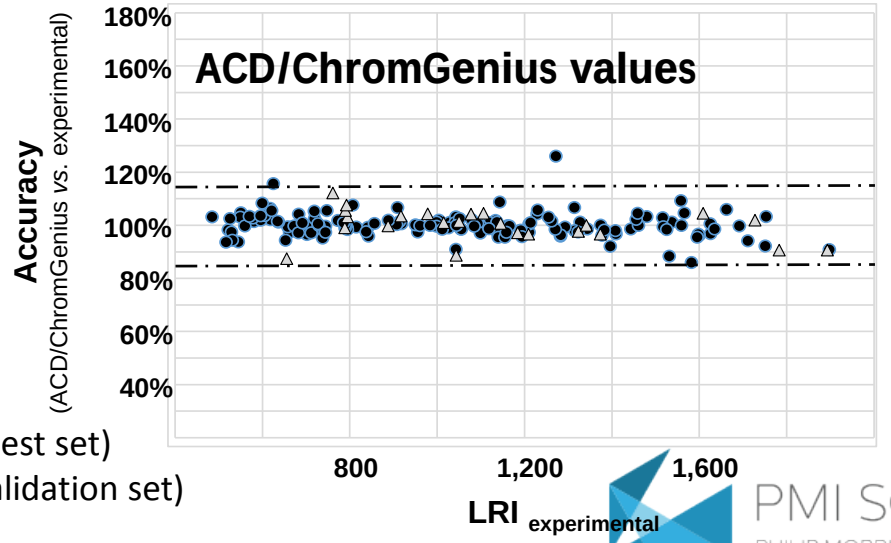



$R^2$: 0.949




$R^2$: 0.976



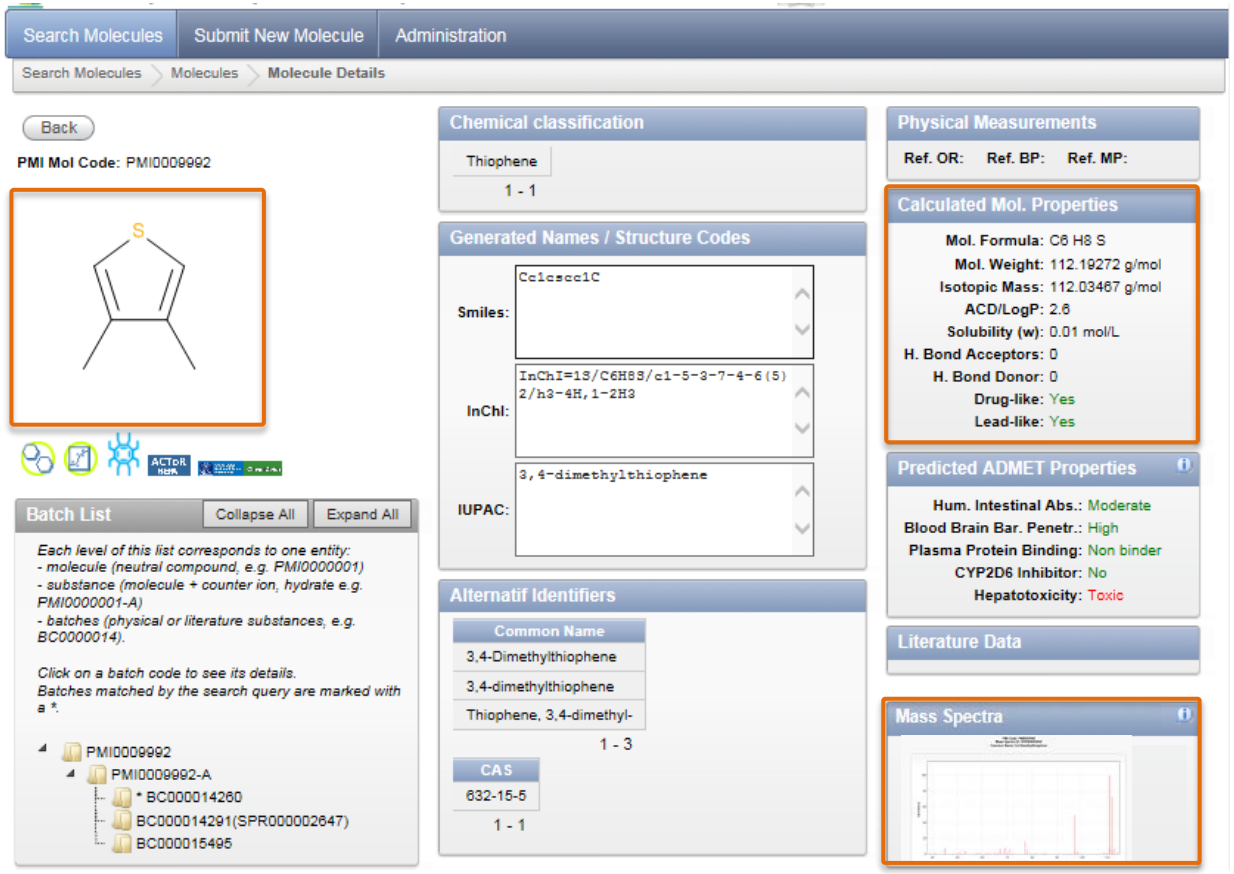

n=151 reference standards (test set)

△ n=23 reference standards (validation set)

E. Dossin *et al., Anal. Chem*. **2016**, *88*, 15, p7539-7547.

PMI SCIENCE
PHILIP MORRIS INTERNATIONAL

# Unique Compound & Spectra Database (UCSD)




UCSD is our in-house database that contains:

- 11’567 molecules
- 1’013 accurate mass spectra and $LRI_{exp}$

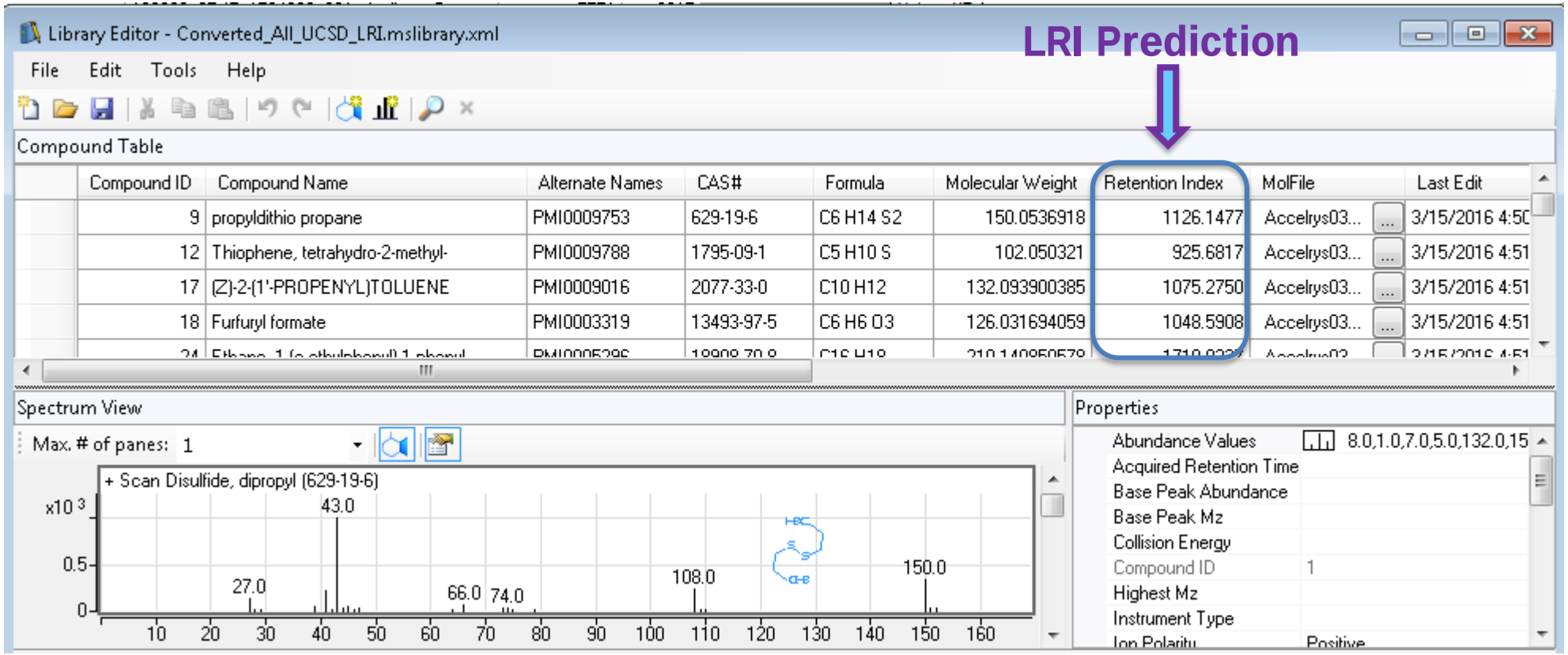
The 2 models were used to predict the LRI values of UCSD database (suspect list)

Our GC-MS conditions are suitable to analyze potentially 6’053 molecules

Of these 6’053 molecules, **3’646** have an available nominal EI Mass Spectra (NIST or Wiley)

PMI SCIENCE
PHILIP MORRIS INTERNATIONAL

# Creation of the Suspected Compound Library



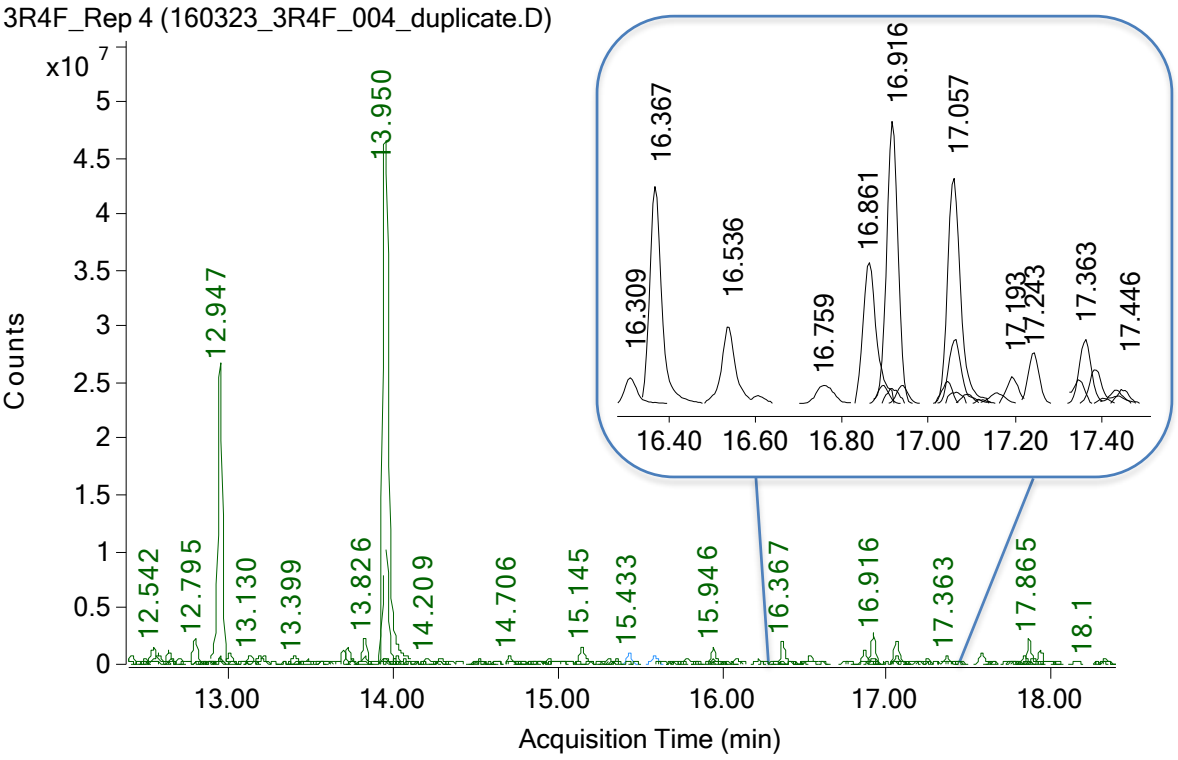

**LRI prediction values were associated with the 3'646 nominal EI mass spectra extracted from commercial libraries**

PMI SCIENCE
PHILIP MORRIS INTERNATIONAL

# Compound Identification in Complex Matrices

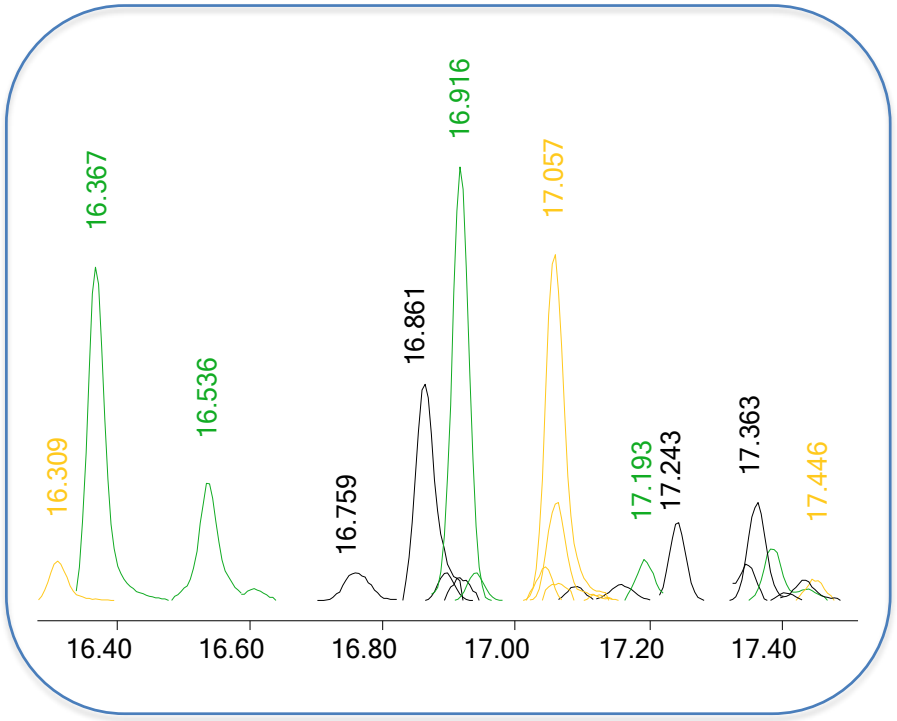
**Data Acquisition (+EI)** → **Deconvolution** → **MassHunter Unknown Analysis Software** → **Identification**

- Stop when identified:
  - ➔ Accurate Mass Library
    **(In house PMI Library≈800 spectra)**
  - ➔ Suspected Compound Library
  - ➔ **NIST 14**
- Export .cef and review
- Purchase of ref std if available




| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 1 | Compound Name | Formula | CAS# | RT | Match Factor | Library File |
| 378 | 2,5-Pyrrolidinedione (CAS) | C4H5NO2 | 123-56-8 | 16.309 | 67.5 | UCSD |
| 379 | 3-Hydroxypyridine | C5H5NO | 109-00-2 | 16.367 | 52.8 | FingerprintingDB624 |
| 380 | 4-Ethylphenol | C8H10O | 123-07-9 | 16.536 | 76.4 | FingerprintingDB624 |
| 381 | 3-Ethyl phenol | C8H10O | 620-17-7 | 16.605 | 66.3 | FingerprintingDB624 |
| 382 | 4-Fluorobenzoic acid, hex-4-yn-3-yl ester | C13H13FO2 | 1000299-15-3 | 16.759 | 61.1 | NIST14 |
| 383 | Butanoic acid, 3-methyl-, 3-methyl-3-butenyl ester | C10H18O2 | 54410-94-5 | 16.862 | 65.5 | NIST14 |
| 384 | Phenol, 3-amino- | C6H7NO | 591-27-5 | 16.894 | 60.7 | NIST14 |
| 385 | Sarcosylsarcosine, N-methoxycarbonyl-, heptyl ester | C15H28N2O5 | 1000314-83-0 | 16.914 | 66.4 | NIST14 |
| 386 | Tridecane | C13H28 | 629-50-5 | 16.916 | 74.0 | FingerprintingDB624 |
| 387 | 3,4-Dimethylphenol | C8H10O | 95-65-8 | 16.934 | 60.3 | FingerprintingDB624 |
| 388 | 4-Methyl-1,2-dihydronaphthalene | C11H12 | 4373-13-1 | 17.044 | 76.8 | UCSD |
| 389 | Pentanedioic acid, 2-oxo-, dimethyl ester (CAS) | C7H10O5 | 13192-04-6 | 17.057 | 49.4 | UCSD |
| 390 | Nonanoic acid (CAS) | C9H18O2 | 112-05-0 | 17.061 | 54.4 | UCSD |
| 391 | Cyclohexasiloxane, dodecamethyl- | C12H36O6Si6 | 540-97-6 | 17.158 | 60.7 | NIST14 |
| 392 | Phenol,2,3,4-trimethyl | C9H12O | 526-85-2 | 17.193 | 46.4 | FingerprintingDB624 |
| 393 | Succinic acid, 2,5-dimethylphenyl ethyl ester | C14H18O4 | 1000329-86-0 | 17.243 | 56.4 | NIST14 |
| 394 | Naphthalene, 1,2-dihydro-3-methyl- (CAS) | C11H12 | 2717-44-4 | 17.347 | 65.5 | UCSD |
| 395 | 1,4:3,6-Dianhydro-.alpha.-d-glucopyranose | C6H8O4 | 1000098-14-8 | 17.363 | 77.8 | NIST14 |
| 396 | 4-isopropylphenol | C9H12O | 99-89-8 | 17.385 | 47.4 | FingerprintingDB624 |
| 397 | Divinylmethyl(chloromethyl)silane | C6H11ClSi | 25202-02-2 | 17.433 | 56.1 | NIST14 |
| 398 | 1H-Pyrrole-2,5-dione, 3-ethyl-4-methyl- (CAS) | C7H9NO2 | 20189-42-8 | 17.446 | 53.3 | UCSD |

# Reviewing Step




| 1 | Compound Name | Formula | CAS# | Component RI | Library RI | Δ RI | Match Factor | Library File |
|---|---|---|---|---|---|---|---|---|
| 378 | 2,5-Pyrrolidinedione (CAS) | C4H5NO2 | 123-56-8 | 1260.70 | 1249.71 | -10.99 | 67.5 | UCSD |
| 379 | 3-Hydroxypyridine | C5H5NO | 109-00-2 | 1264.45 | 1272.07 | 7.62 | 52.8 | FingerprintingDB624 |
| 380 | 4-Ethylphenol | C8H10O | 123-07-9 | 1275.38 | 1279.99 | 4.61 | 76.4 | FingerprintingDB624 |
| 381 | 3-Ethyl phenol | C8H10O | 620-17-7 | 1279.84 | 1282.00 | 2.16 | 66.3 | FingerprintingDB624 |
| 382 | 4-Fluorobenzoic acid, hex-4-yn-3-yl ester | C13H13FO2 | 1000299-15-3 | 1289.82 | | | 61.1 | NIST14 |
| 383 | Butanoic acid, 3-methyl-, 3-methyl-3-butenyl ester | C10H18O2 | 54410-94-5 | 1296.49 | | | 65.5 | NIST14 |
| 384 | Phenol, 3-amino- | C6H7NO | 591-27-5 | 1298.61 | | | 60.7 | NIST14 |
| 385 | Sarcosylsarcosine, N-methoxycarbonyl-, heptyl ester | C15H28N2O5 | 1000314-83-0 | 1299.86 | | | 66.4 | NIST14 |
| 386 | Tridecane | C13H28 | 629-50-5 | 1300.00 | 1300.00 | 0.00 | 74.0 | FingerprintingDB624 |
| 387 | 3,4-Dimethylphenol | C8H10O | 95-65-8 | 1301.35 | 1302.07 | 0.71 | 60.3 | FingerprintingDB624 |
| 388 | 4-Methyl-1,2-dihydronaphthalene | C11H12 | 4373-13-1 | 1309.40 | 1311.61 | 2.21 | 76.8 | UCSD |
| 389 | Pentanedioic acid, 2-oxo-, dimethyl ester (CAS) | C7H10O5 | 13192-04-6 | 1310.40 | 1279.05 | -31.34 | 49.4 | UCSD |
| 390 | Nonanoic acid (CAS) | C9H18O2 | 112-05-0 | 1310.71 | 1323.04 | 12.33 | 54.4 | UCSD |
| 391 | Cyclohexasiloxane, dodecamethyl- | C12H36O6Si6 | 540-97-6 | 1317.81 | | | 60.7 | NIST14 |
| 392 | Phenol,2,3,4-trimethyl | C9H12O | 526-85-2 | 1320.41 | 1325.05 | 4.64 | 46.4 | FingerprintingDB624 |
| 393 | Succinic acid, 2,5-dimethylphenyl ethyl ester | C14H18O4 | 1000329-86-0 | 1324.09 | | | 56.4 | NIST14 |
| 394 | Naphthalene, 1,2-dihydro-3-methyl- (CAS) | C11H12 | 2717-44-4 | 1331.73 | 1322.44 | -9.29 | 65.5 | UCSD |
| 395 | 1,4:3,6-Dianhydro-.alpha.-d-glucopyranose | C6H8O4 | 1000098-14-8 | 1332.95 | | | 77.8 | NIST14 |
| 396 | 4-isopropylphenol | C9H12O | 99-89-8 | 1334.55 | 1340.26 | 5.71 | 47.4 | FingerprintingDB624 |
| 397 | Divinylmethyl(chloromethyl)silane | C6H11ClSi | 25202-02-2 | 1338.12 | | | 56.1 | NIST14 |
| 398 | 1H-Pyrrole-2,5-dione, 3-ethyl-4-methyl- (CAS) | C7H9NO2 | 20189-42-8 | 1339.08 | 1364.74 | 25.66 | 53.3 | UCSD |

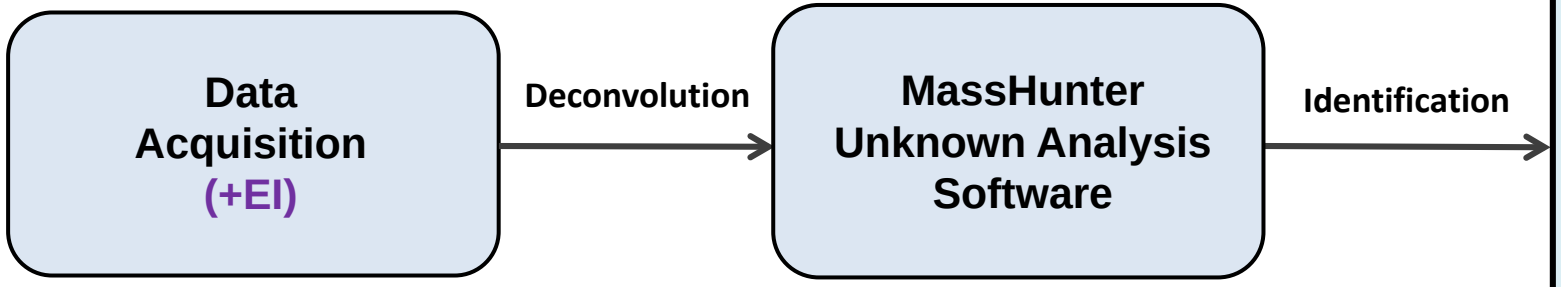
Identified with PMI Accurate Mass Library (n=7 out of 21)

Suspect compound library (n=6 out of 21)

NIST proposal (n=8 out of 21)

PMI SCIENCE
PHILIP MORRIS INTERNATIONAL

# Level of Confidence for Commercial Library proposals




- Stop when identified:
  - ➔ Accurate Mass Library
    (In house PMI Library≈800 spectra)
  - ➔ Suspected Compound Library
  - ➔ NIST 14
- Export .cef and review
- Purchase of ref std if available

| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 1 | Compound Name | Formula | CAS# | Component RI | Library RI | Δ RI | Match Factor | Library File |
| 378 | | | | | | | | |
| 379 | | | | | | | | |
| 380 | | | | | | | | |
| 381 | | | | | | | | |
| 382 | 4-Fluorobenzoic acid, hex-4-yn-3-yl ester | C13H13FO2 | 1000299-15-3 | 1289.82 | | | 61.1 | NIST14 |
| 383 | Butanoic acid, 3-methyl-, 3-methyl-3-butenyl ester | C10H18O2 | 54410-94-5 | 1296.49 | | | 65.5 | NIST14 |
| 384 | Phenol, 3-amino- | C6H7NO | 591-27-5 | 1298.61 | | | 60.7 | NIST14 |
| 385 | Sarcosylsarcosine, N-methoxycarbonyl-, heptyl ester | C15H28N2O5 | 1000314-83-0 | 1299.86 | | | 66.4 | NIST14 |
| 386 | | | | | | | | |
| 387 | | | | | | | | |
| 388 | | | | | | | | |
| 389 | | | | | | | | |
| 390 | | | | | | | | |
| 391 | Cyclohexasiloxane, dodecamethyl- | C12H36O6Si6 | 540-97-6 | 1317.81 | | | 60.7 | NIST14 |
| 392 | | | | | | | | |
| 393 | Succinic acid, 2,5-dimethylphenyl ethyl ester | C14H18O4 | 1000329-86-0 | 1324.09 | | | 56.4 | NIST14 |
| 394 | | | | | | | | |
| 395 | 1,4:3,6-Dianhydro-.alpha.-d-glucopyranose | C6H8O4 | 1000098-14-8 | 1332.95 | | | 77.8 | NIST14 |
| 396 | | | | | | | | |
| 397 | Divinylmethyl(chloromethyl)silane | C6H11ClSi | 25202-02-2 | 1338.12 | | | 56.1 | NIST14 |
| 398 | | | | | | | | |

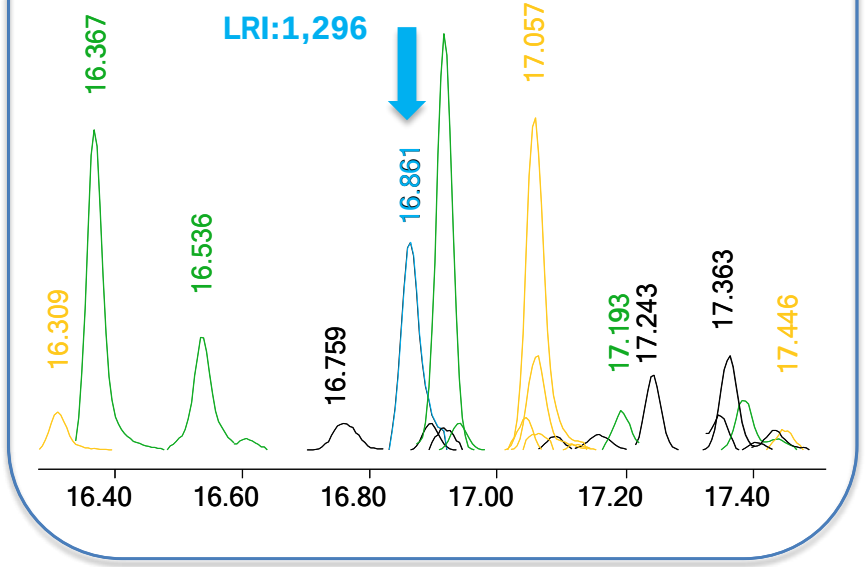
| Pred LRI Rapid Miner | Pred LRI Chrom Genius | |
|---|---|---|
| 1627.9 | 1776.4 | Very different Δ = 412 |
| 1164.0 | 1253.9 | In agreement with exp Δ = 87 |
| 1502.9 | 1477.3 | Very different Δ = 191 |
| 2304.4 | 2279.2 | Very different Δ = 992 |
| 4054.1 | 1635.9 | Very different Δ = 1677 |
| 1914.4 | 1934.4 | Very different Δ = 600 |
| 1354.0 | 1551.6 | In agreement with exp Δ =120 |
| 990.8 | 943.7 | Very different Δ = 371 |

The identification confidence is low to medium with a commercial library search

**Filter out false positive results proposed by existing MS libraries**

# Case Study: Compound with a Poor Match Factor

| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 1 | Compound Name | Formula | CAS# | Component RI | Library RI | Δ RI | Match Factor | Library File |
| 383 | Butanoic acid, 3-methyl-, 3-methyl-3-butenyl ester | C10H18O2 | 54410-94-5 | 1296.49 | | | 65.5 | NIST14 |




| Pred LRI Rapid Miner | Pred LRI Chrom Genius |
|---|---|
| 1164.0 | 1253.9 |

In agreement with exp Δ = 87

**Additional MS experiment are needed and alternative approaches need to be implemented**

**Data Acquisition**
**CI full scan MS**
**Targeted MS/MS**

**LRI prediction**

**MetFrag**
***in silico***
**fragmentation software**

**Data Acquisition**
**EI full scan MS**

**Candidates**

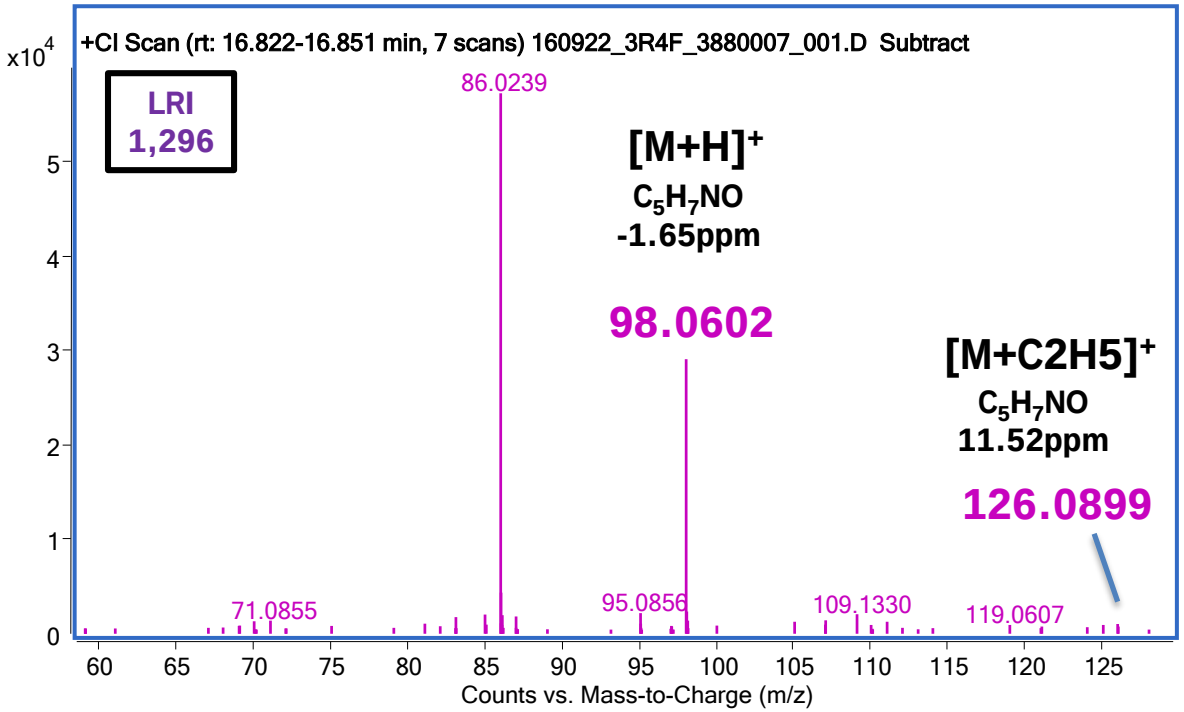
**CFM-ID**
***in silico***
**fragmentation software**

**Final Score**

**Purchase of reference standards if available**

# GC-HR-MS in Chemical Ionization Mode & MS/MS

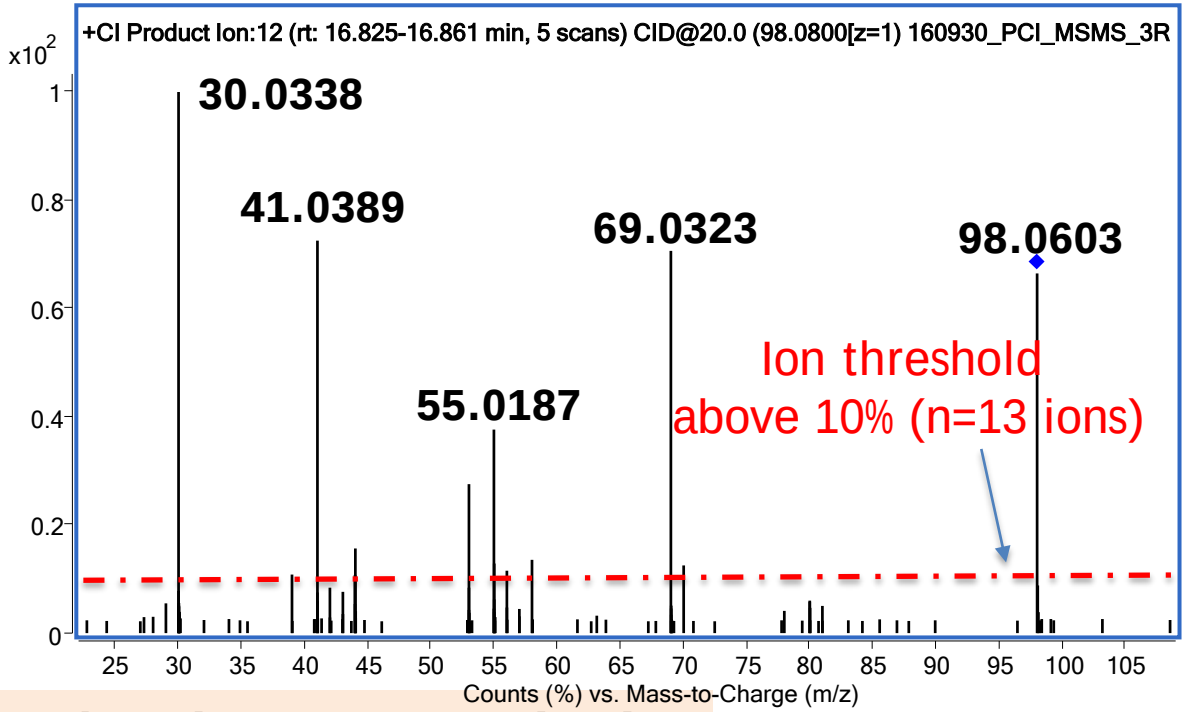
**GC-HR-MS (Full Scan MS)**
**Positive Chemical Ionization (PCI)**




**Determination of elemental formula (adduct ion species)**

**M: $C_5H_7NO$**

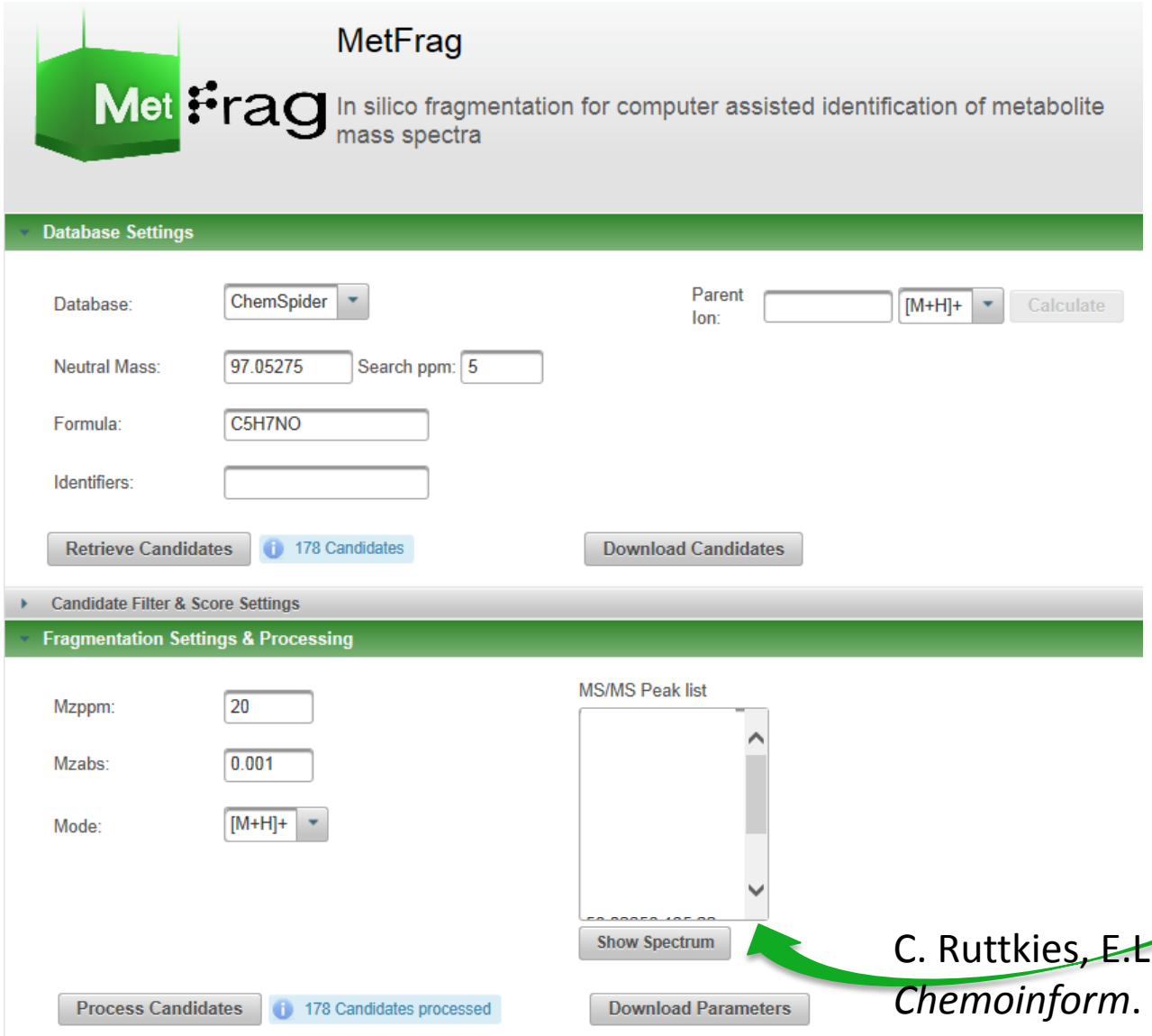
**GC-HR-MS (Full Scan MS/MS)**
**PCI data acquisition CID of 98.08004**

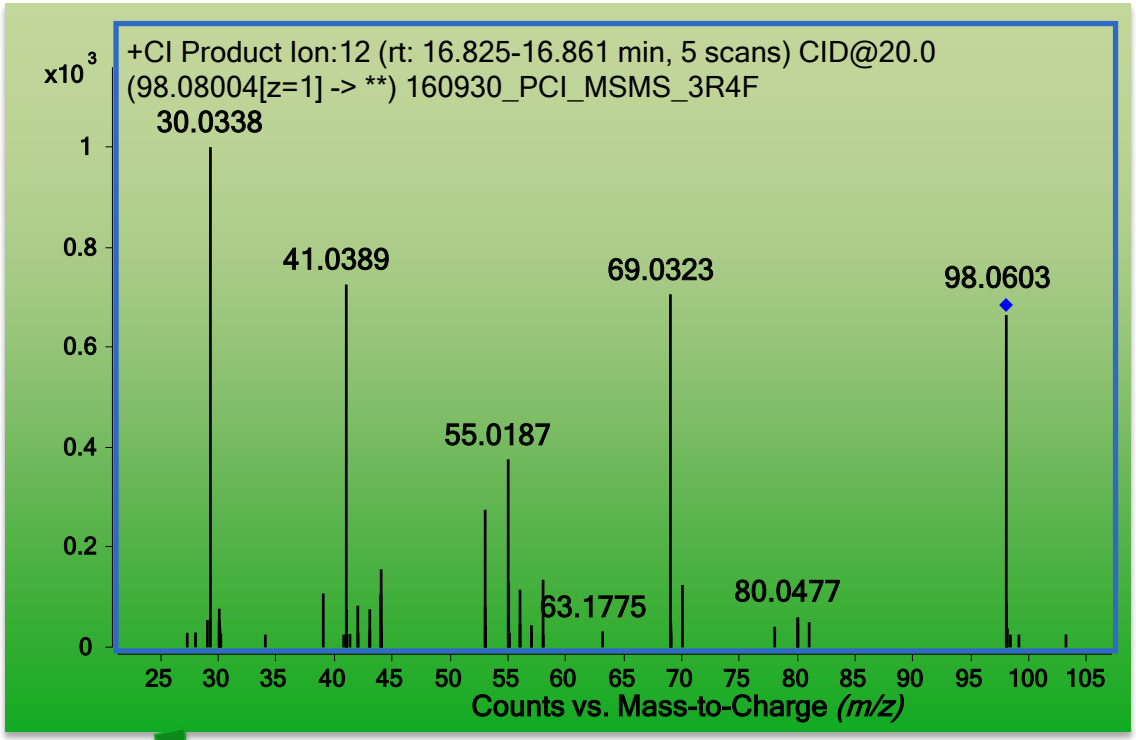



**MS/MS data processed using a larger chemical database with *in silico* predicted fragmentation software**

PMI SCIENCE
PHILIP MORRIS INTERNATIONAL

# Compound Identification with *In Silico* Fragmentation Software




C. Ruttkies, E.L. Schymanski, S. Wolf, J. Hollender, S. Neumann *J. Cheminform*. **2016**, *8*, 3, 16-115.

PMI SCIENCE
PHILIP MORRIS INTERNATIONAL

# Metfrag Results

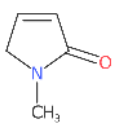
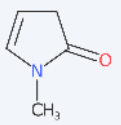
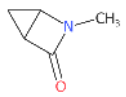
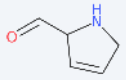
Results

Download Results

Filter Candidates by explained MS/MS Peaks

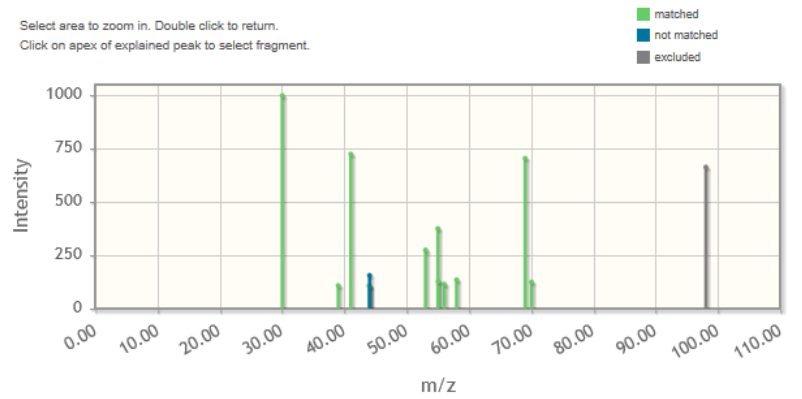
MS/MS Peaks | Filter Candidates

1 2 3 4 5 6 7 8 9 10

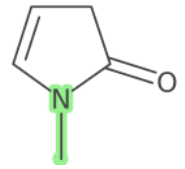
| # | Molecule | Identifier | Mass | Formula | FinalScor | Details |
|---|---|---|---|---|---|---|
| 1 |  3-Pyrrolin-2-one, 1-methyl- | 466735 InChIKeyBlock1 = VHGGRTWHRJRQKU | 97.05277 | $C_5H_7NO$ | 1.0 | Peaks: 12 / 13 Fragments Scores Download |
| 2 |  1-Methyl-1,3-dihydro-2H-pyrrol-2-one | 4483194 InChIKeyBlock1 = VIETUFSZPCIVQL | 97.05277 | $C_5H_7NO$ | 0.9913 | Peaks: 12 / 13 Fragments Scores Download |
| 3 |  2-Methyl-2-azabicyclo[2.1.0]pentan-3-one | 45531574 InChIKeyBlock1 = SYIPZAVVNRAIEV | 97.05277 | $C_5H_7NO$ | 0.9758 | Peaks: 12 / 13 Fragments Scores Download |
| 4 |  | 32989542 InChIKeyBlock1 = | 97.05277 | $C_5H_7NO$ | 0.9589 | Peaks: 10 / 13 Fragments Scores |

1 2 3 4 5 6 7 8 9 10

Fragments View




Fragments

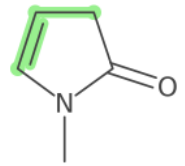


**Fragment 1**

Peak m/z: 30.0338

Fragment Mass: 30.03383 Da

Fragment Formula: $[CH_3N]+H^+$

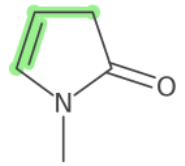


**Fragment 2**

Peak m/z: 39.0234

Fragment Mass: 39.02293 Da

Fragment Formula: $[C_3H_4-H]^+$



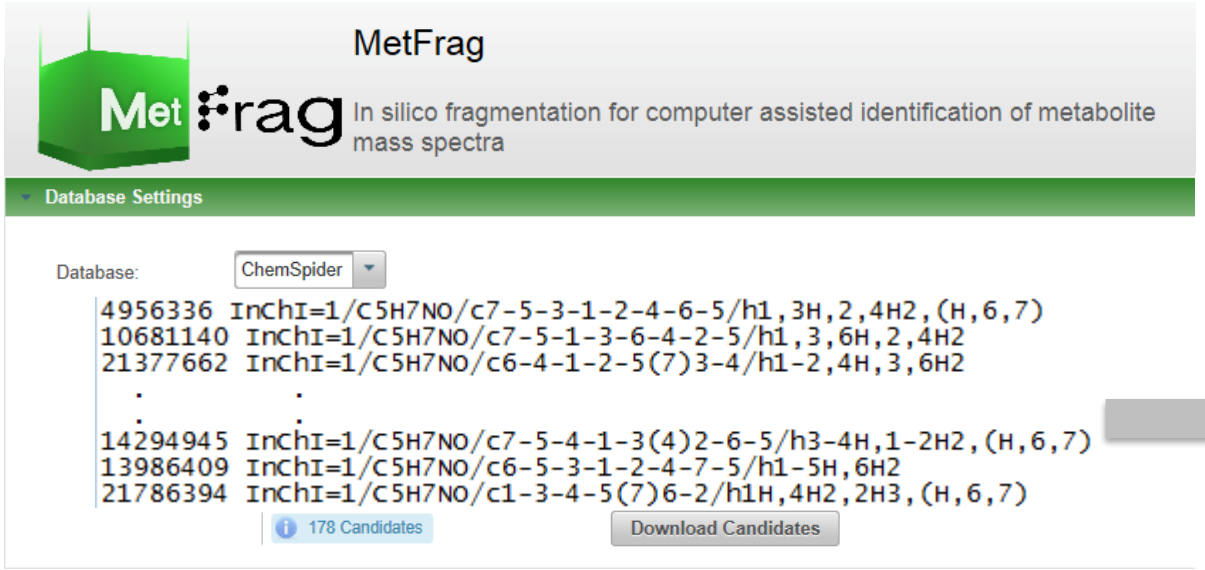
**Fragment 3**

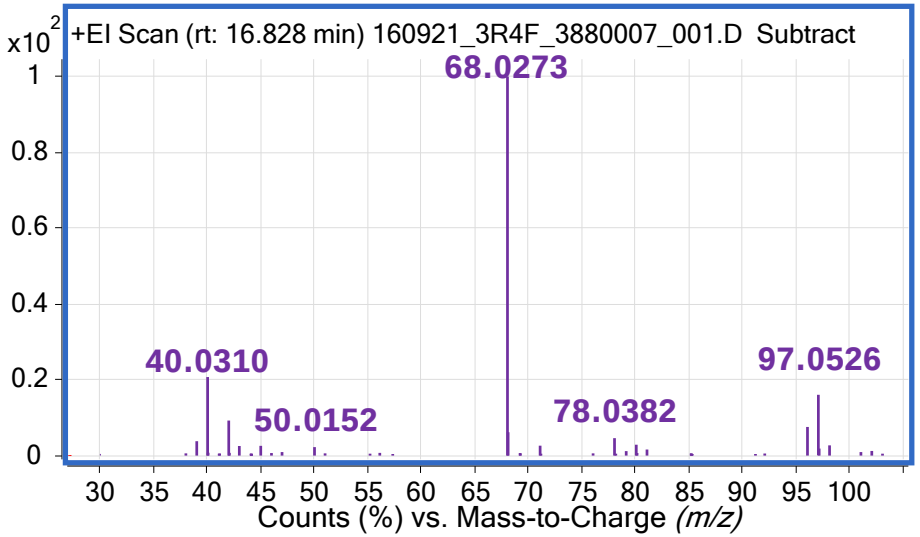
Peak m/z: 41.0389

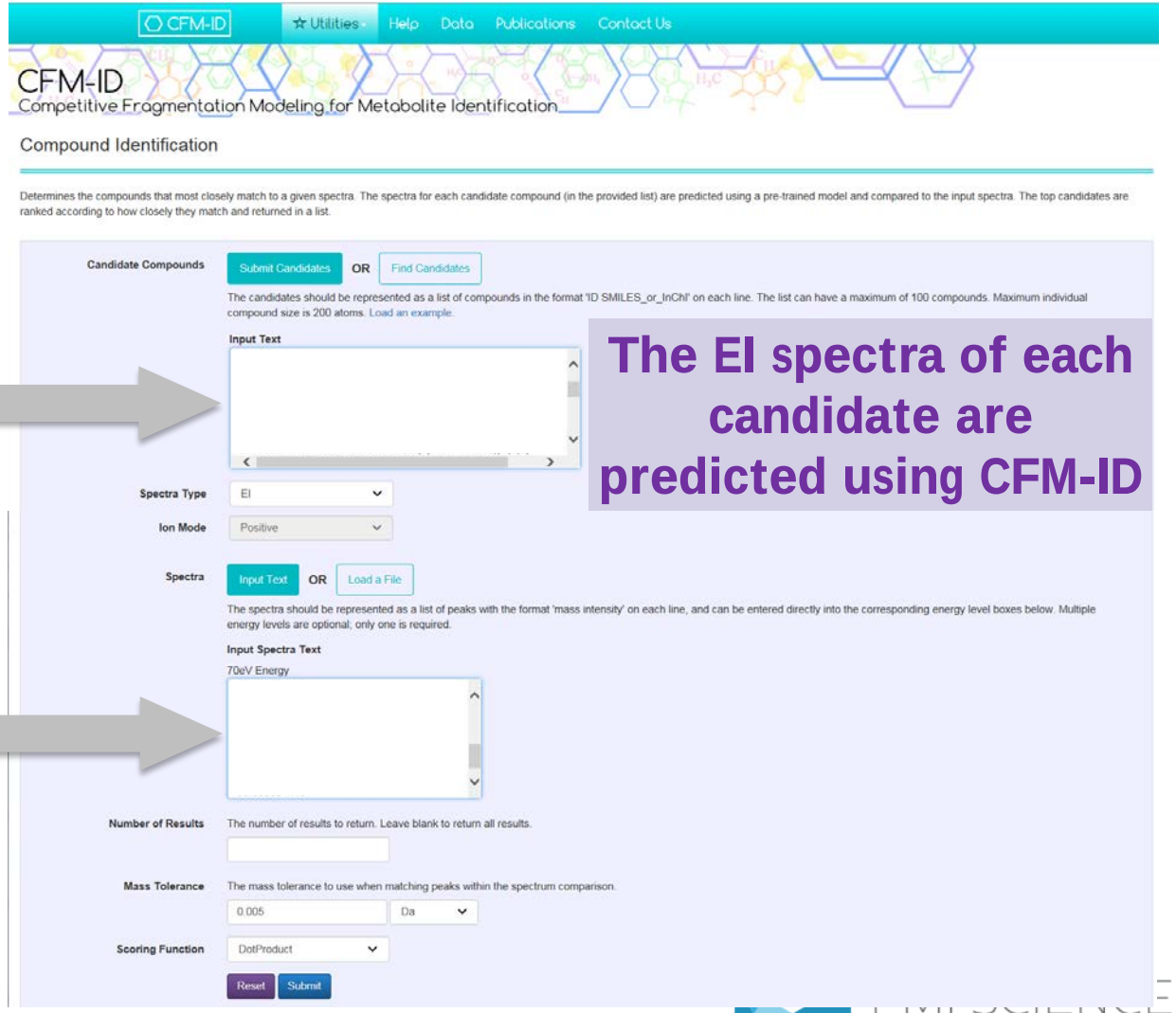
Fragment Mass: 41.03858 Da

Fragment Formula: $[C_3H_4]+H^+$

# Compound Identification with Fragmentation Modeling Software




**The EI spectra of each candidate are predicted using CFM-ID**

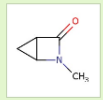
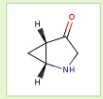
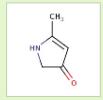
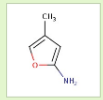
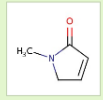
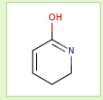
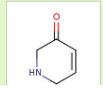
F. Allen, R. Greiner, D. Wishart *Metabolomics*. **2015**, *11*, 1, 98-110.

# CFM-ID Results

Results: Download

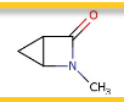
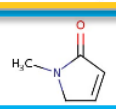
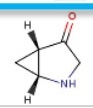
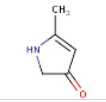
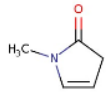
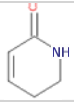
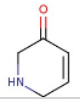
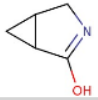
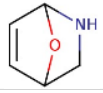
Input spectra are shown below in blue. If a database was queried, candidate spectra are overlayed on top for comparison. The top ranking candidate spectra is shown by default; to compare other database candidates use the "Compare" buttons on the list of ranked candidate compounds that follow the spectra.

Candidate Rankings

| Rank | Score | Structure | ID | Chemical Formula | InChI/SMILES | Compare |
|---|---|---|---|---|---|---|
| 1 | 0.27059567 |  | 45531574 | C5H7NO | InChI=1/C5H7NO/c1-6-4-2-3(4)5(6)7/h3-4H,2H2,1H3 | Current |
| 2 | 0.23287233 |  | 32990381 | C5H7NO | InChI=1/C5H7NO/c7-5-2-6-4-1-3(4)5/h3-4,6H,1-2H2/t3-,4+/m0/s1 | Compare |
| 3 | 0.21118675 |  | 10340193 | C5H7NO | InChI=1/C5H7NO/c1-4-2-5(7)3-6-4/h2,6H,3H2,1H3 | Compare |
| 4 | 0.19719996 |  | 27268587 | C5H7NO | InChI=1/C5H7NO/c1-4-2-5(6)7-3-4/h2-3H,6H2,1H3 | Compare |
| 5 | 0.19212561 |  | 466735 | C5H7NO | InChI=1/C5H7NO/c1-6-4-2-3-5(6)7/h2-3H,4H2,1H3 | Compare |
| 6 | 0.19070942 |  | 4956336 | C5H7NO | InChI=1/C5H7NO/c7-5-3-1-2-4-6-5/h1,3H,2,4H2,(H,6,7) | Compare |
| 7 | 0.1824996 |  | 14395305 | C5H7NO | InChI=1/C5H7NO/c7-5-2-1-3-6-4-5/h1-2,6H,3-4H2 | Compare |

PMI SCIENCE
PHILIP MORRIS INTERNATIONAL

# Final Score (MetFrag / CFM-ID)

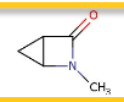
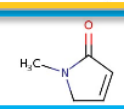
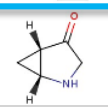
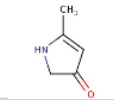
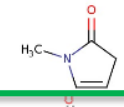
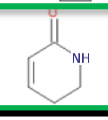
| InChI | CompoundName | Molecular Formula | Identifier | Structure | Rank MetFrag PCI | Score MetFrag PCI | Rank CFM-ID | Score CFM-ID |
|---|---|---|---|---|---|---|---|---|
| c1-6-4-2-3(4)5(6)7/h3-4H,2H2,1H3 | 2-Methyl-2-azabicyclo[2.1.0]pentan-3-one | C5H7NO | 45531574 |  | 3 | 0.9758 | 1 | 0.271 |
| c1-6-4-2-3-5(6)7/h2-3H,4H2,1H3 | 3-Pyrrolin-2-one, 1-methyl- | C5H7NO | 466735 |  | 1 | 1.0 | 5 | 0.192 |
| c7-5-2-6-4-1-3(4)5/h3-4,6H,1-2H2/t3-,4+/m0/s1 | (1R,5S)-2-Azabicyclo[3.1.0]hexan-4-one | C5H7NO | 32990381 |  | 16 | 0.9089 | 2 | 0.233 |
| c1-4-2-5(7)3-6-4/h2,6H,3H2,1H3 | 5-Methyl-1,2-dihydro-3H-pyrrol-3-one | C5H7NO | 10340193 |  | 15 | 0.915 | 3 | 0.211 |
| c1-6-4-2-3-5(6)7/h2,4H,3H2,1H3 | 1-Methyl-1,3-dihydro-2H-pyrrol-2-one | C5H7NO | 4483194 |  | 2 | 0.9913 | 11 | 0.167 |
| c7-5-3-1-2-4-6-5/h1,3H,2,4H2,(H,6,7) | 5,6-Dihydro-2(1H)-pyridinone | C5H7NO | 4956336 |  | 5 | 0.9354 | 6 | 0.191 |
| c7-5-2-1-3-6-4-5/h1-2,6H,3-4H2 | 1,6-Dihydro-3(2H)-pyridinone | C5H7NO | 14395305 |  | 8 | 0.9294 | 7 | 0.182 |
| c7-5-4-1-3(4)2-6-5/h3-4H,1-2H2,(H,6,7) | 3-Azabicyclo[3.1.0]hexan-2-one | C5H7NO | 14294945 |  | 12 | 0.9173 | 8 | 0.182 |
| c1-2-5-6-3-4(1)7-5/h1-2,4-6H,3H2 | 7-Oxa-2-azabicyclo[2.2.1]hept-5-ene | C5H7NO | 29542634 |  | 22 | 0.8925 | 9 | 0.181 |

$$\text{Combined Score} = (\text{MetFrag} \times 0.66 + \text{CFM} - \text{ID}_{Norm} \times 0.33)$$

PMI SCIENCE
PHILIP MORRIS INTERNATIONAL

# Final Score (MetFrag / CFM-ID / LRI Predictions)

**Unknown Peak**
**$LRI_{exp}$:1,296**

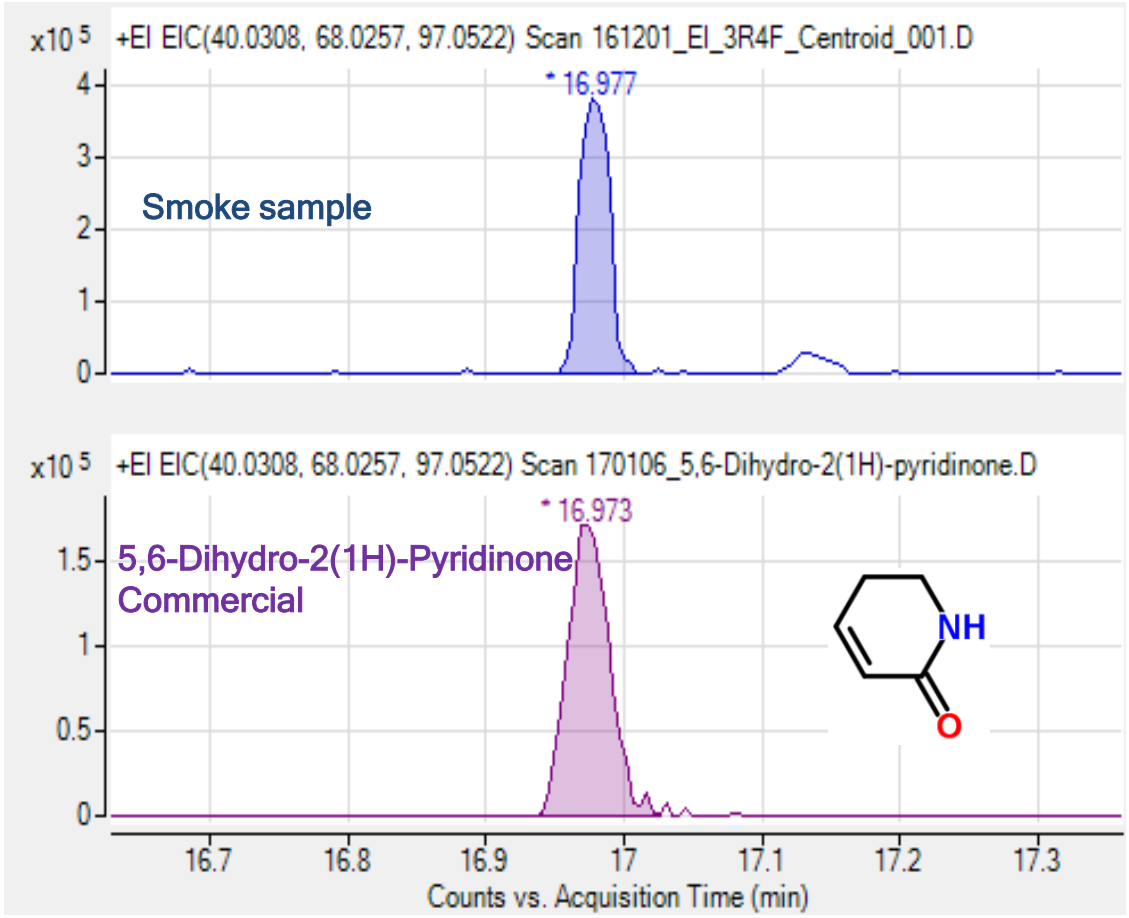
| InChI | CompoundName | Molecular Formula | Identifier | Structure | Rank MetFrag PCI | Score MetFrag PCI | Rank CFM-ID | Score CFM-ID | Score MetFrag & CFM-ID | LRI RM | LRI CG | Final Score |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| c1-6-4-2-3(4)5(6)7/h3-4H,2H2,1H3 | 2-Methyl-2-azabicyclo[2.1.0]pentan-3-one | C5H7NO | 45531574 |  | 3 | 0.9758 | 1 | 0.271 | 0.974 | 962.51 | 1036.66 | 0.65 |
| c1-6-4-2-3-5(6)7/h2-3H,4H2,1H3 | 3-Pyrrolin-2-one, 1-methyl- | C5H7NO | 466735 |  | 1 | 1.0 | 5 | 0.192 | 0.894 | 995.41 | 1168.08 | 0.66 |
| c7-5-2-6-4-1-3(4)5/h3-4,6H,1-2H2/t3-,4+/m0/s1 | (1R,5S)-2-Azabicyclo[3.1.0]hexan-4-one | C5H7NO | 32990381 |  | 16 | 0.9089 | 2 | 0.233 | 0.884 | 1126.04 | 1015.88 | 0.64 |
| c1-4-2-5(7)3-6-4/h2,6H,3H2,1H3 | 5-Methyl-1,2-dihydro-3H-pyrrol-3-one | C5H7NO | 10340193 |  | 15 | 0.915 | 3 | 0.211 | 0.861 | 1057.08 | 918.43 | 0.56 |
| c1-6-4-2-3-5(6)7/h2,4H,3H2,1H3 | 1-Methyl-1,3-dihydro-2H-pyrrol-2-one | C5H7NO | 4483194 |  | 2 | 0.9913 | 11 | 0.167 | 0.858 | 994.94 | 1083.00 | 0.60 |
| c7-5-3-1-2-4-6-5/h1,3H,2,4H2,(H,6,7) | 5,6-Dihydro-2(1H)-pyridinone | C5H7NO | 4956336 |  | 5 | 0.9354 | 6 | 0.191 | 0.850 | 1110.21 | 1304.80 | 0.74 |
| c7-5-2-1-3-6-4-5/h1-2,6H,3-4H2 | | | | | | | | | 0.836 | 1170.32 | 1049.83 | 0.64 |
| c7-5-4-1-3(4)2-6-5/h3-4H,1-2H2,(H,6,7) | 3- | | | | | | | | 0.827 | 1041.30 | 1303.53 | 0.69 |
| c1-2-5-6-3-4(1)7-5/h1-2,4-6H,3H2 | 7-O | | | | | | | | 0.810 | 1098.64 | 893.42 | 0.54 |

1) We use Inchi information to draw the compound and LRI values were predicted for all 178 proposals
2) Final SCORE was calculated using:
   - ✓ MetFrag Score
   - ✓ CFM-ID Score
   - ✓ $LRI_{exp.}$ against $LRI_{RM}$
   - ✓ $LRI_{exp.}$ against $LRI_{CG}$ …

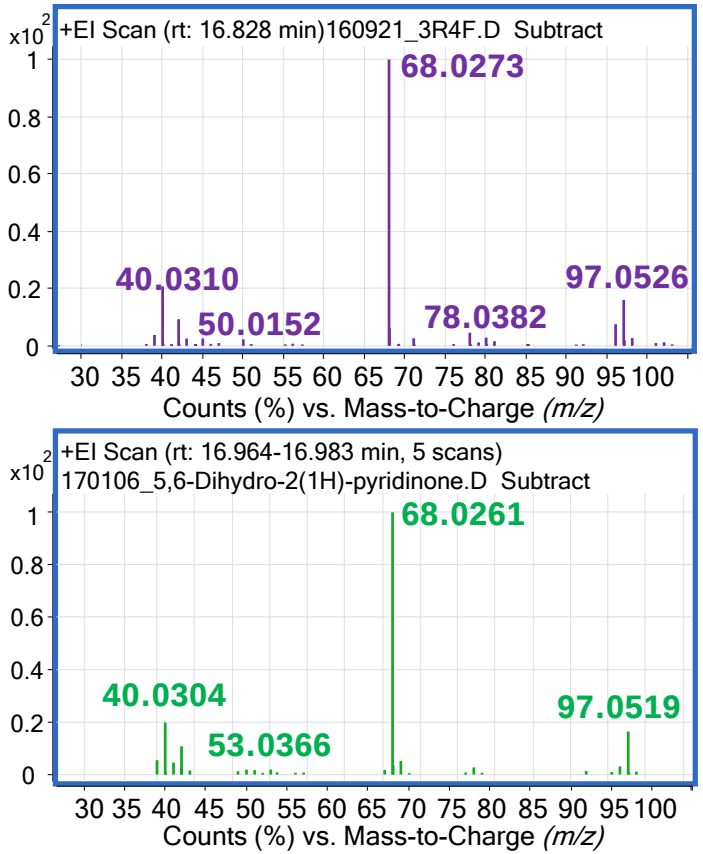
$$Final_{Score} = \frac{(\text{MetFrag} \times 0.66 + \text{CFM} - \text{ID}_{Norm} \times 0.33) \times LRI_{exp} \times LRI_{exp}}{\left[LRI_{exp} + \sqrt{(LRI_{CG} - LRI_{exp})^2}\right] \times \left[LRI_{exp} + \sqrt{(LRI_{RM} - LRI_{exp})^2}\right]}$$

PMI SCIENCE
PHILIP MORRIS INTERNATIONAL

# Final Confirmation




Same RT

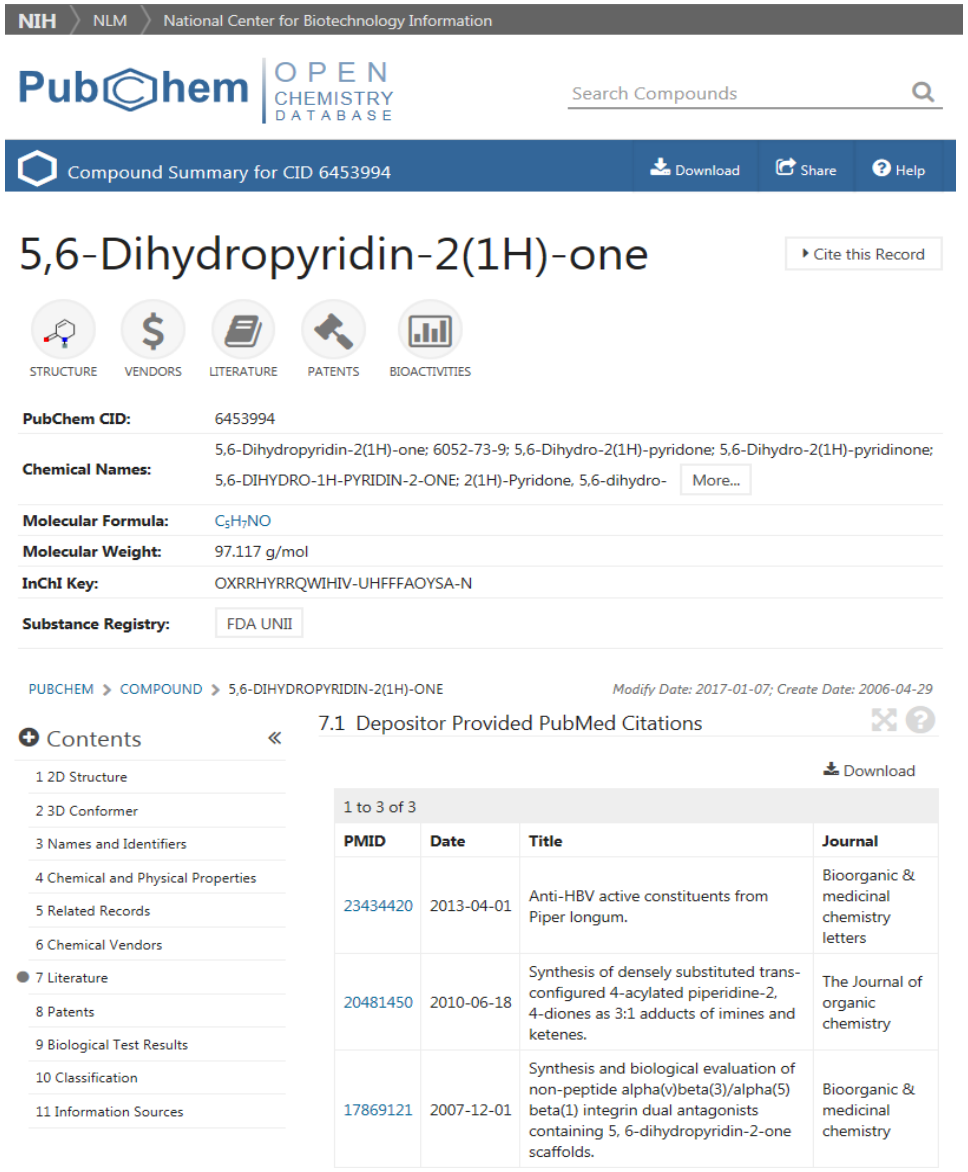



Similar EI Mass Spectra

**5,6-Dihydro-2(1H)-Pyridinone** ✓

PMI SCIENCE
PHILIP MORRIS INTERNATIONAL

# Bibliography


NIH › NLM › National Center for Biotechnology Information

PubChem OPEN CHEMISTRY DATABASE

Search Compounds

Compound Summary for CID 6453994 — Download — Share — Help

5,6-Dihydropyridin-2(1H)-one

Cite this Record

STRUCTURE VENDORS LITERATURE PATENTS BIOACTIVITIES

| | |
|---|---|
| **PubChem CID:** | 6453994 |
| **Chemical Names:** | 5,6-Dihydropyridin-2(1H)-one; 6052-73-9; 5,6-Dihydro-2(1H)-pyridone; 5,6-Dihydro-2(1H)-pyridinone; 5,6-DIHYDRO-1H-PYRIDIN-2-ONE; 2(1H)-Pyridone, 5,6-dihydro- More... |
| **Molecular Formula:** | $C_5H_7NO$ |
| **Molecular Weight:** | 97.117 g/mol |
| **InChI Key:** | OXRRHYRRQWIHIV-UHFFFAOYSA-N |
| **Substance Registry:** | FDA UNII |

PUBCHEM › COMPOUND › 5,6-DIHYDROPYRIDIN-2(1H)-ONE *Modify Date: 2017-01-07; Create Date: 2006-04-29*

Contents

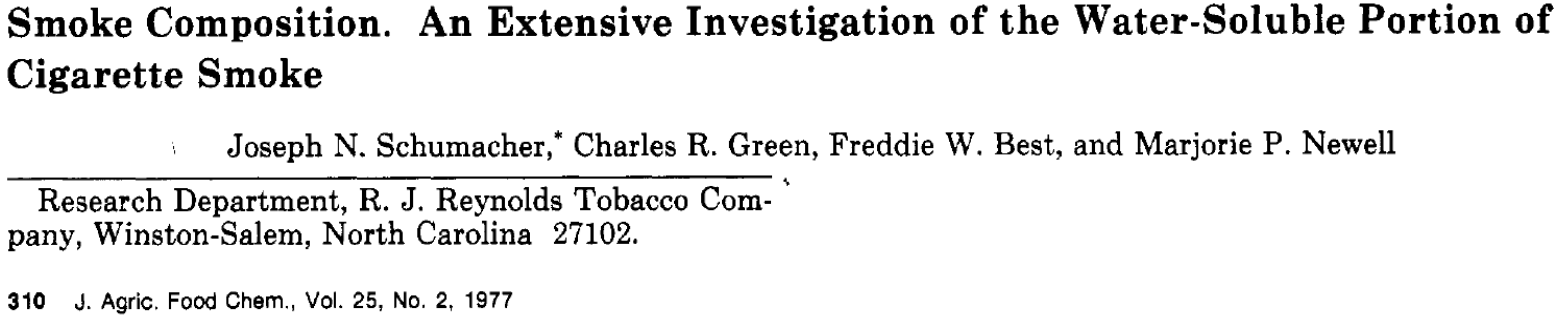
1 2D Structure
2 3D Conformer
3 Names and Identifiers
4 Chemical and Physical Properties
5 Related Records
6 Chemical Vendors
7 Literature
8 Patents
9 Biological Test Results
10 Classification
11 Information Sources

7.1 Depositor Provided PubMed Citations

Download

1 to 3 of 3

| PMID | Date | Title | Journal |
|---|---|---|---|
| 23434420 | 2013-04-01 | Anti-HBV active constituents from Piper longum. | Bioorganic & medicinal chemistry letters |
| 20481450 | 2010-06-18 | Synthesis of densely substituted trans-configured 4-acylated piperidine-2, 4-diones as 3:1 adducts of imines and ketenes. | The Journal of organic chemistry |
| 17869121 | 2007-12-01 | Synthesis and biological evaluation of non-peptide alpha(v)beta(3)/alpha(5) beta(1) integrin dual antagonists containing 5, 6-dihydropyridin-2-one scaffolds. | Bioorganic & medicinal chemistry |


**Smoke Composition. An Extensive Investigation of the Water-Soluble Portion of Cigarette Smoke**

Joseph N. Schumacher,* Charles R. Green, Freddie W. Best, and Marjorie P. Newell

Research Department, R. J. Reynolds Tobacco Company, Winston-Salem, North Carolina 27102.

310 J. Agric. Food Chem., Vol. 25, No. 2, 1977

- **This compound was reported first from cigarette smoke by Schumacher et *al*. in 1977**
- **NOT registered in NIST 14 NOR in Wiley 11**
- **Strengthen the need to implement *in silico* predictions when dealing with untargeted analysis.**

PMI SCIENCE
PHILIP MORRIS INTERNATIONAL

# Conclusion

- **Existing MS libraries are not exhaustive; the implementation of chemoinformatic tools is needed to postulate compounds that are not registered in any library**
- **The combined use of MetFrag and CFM-ID software has been demonstrated to be a good complementary tool to propose reliable compound hits**
- **Addition of LRI predictions demonstrated the ability to correctly rank putative hits proposed by *in silico* fragmentation software**

PMI SCIENCE
PHILIP MORRIS INTERNATIONAL

PMI SCIENCE
PHILIP MORRIS INTERNATIONAL

# Thank You

*eric.dossin@pmi.com*

Eric Dossin, Elyette Martin, Antonio Castellon, Pierrick Diana, Pavel Pospisil, Mark Bentley, Philippe Guy

Philip Morris International R&D, Philip Morris Products S.A.,
(part of Philip Morris International group of companies)