

## 危机环境和本土安全事件中利用快速响应 GC/MS 流动实验室提高未知物识别的确认能力

Jianping Chen、顾明、王永东、Don Kuehl

——本文发表于 LC/GC

### 摘要

本文描述一个新方法，利用 GC/MS 分析挥发性有机物并大大提高其对“未知”化合物识别的准确度和可靠性。方法是基于先进的数据处理方法，对 GC/MS 仪器获得的典型名义质量数据进行处理，并结合 GC/MS 谱库大大提高了化合物识别的准确度。该方法基于建好的 EPA 方法用于分析挥发性有机物，验证了总共 90 个环境相关化合物。

### 引言

康涅狄格环境保护局(CT DEP)位于康涅狄格州的 Windsor, 是装备 GC/MS 等分析仪器的唯一流动实验室，负责本土安全和应急环境事件的现场、快速未知化合物的测定。CT DEP 应急响应队员 24/7 待命，通常是救火队员、地方或联邦执法人员遇到不明的可能化学危险物时电话求助，他们的责任是快速识别这些物质，以便事故指挥官和执法人员作出适当的应变措施，在这种情况下，快速准确识别至关重要，因为这有助于迅速决定是对无毒物质进行简单控制处理还是撤离周围居民。

GC/MS 是快速识别挥发和半挥发有机物的选择工具，标准 EPA 方法能识别超过 200 种环境相关的化合物，这些方法具有很好的灵敏度，根据保留时间和 GC/MS 谱图检索可以准确识别这些目标化合物。然而，对于不能用这些方法表征的化合物，仅仅利用 GC/MS 谱库检索会产生很大的不确定性。错误识别对于先遣急救员和执法者，可能导致昂贵的、危险的甚至致命的错误响应，确认识别的不确定性或延迟也能导致公共资源的极大浪费。

为了提高未知物正确识别的可能性，需要一些分析方法进行二次确证分析，如 GC/MS 谱库匹配后接下来利用精确质量 GC/MS 进行分子式确证。对于化学物识别的初期应答者，挑战是非常严峻的，因为分析必须具有高度的可信度，必须快速（通常在 1-3 小时内），而且必须依靠流动实验室中已有的分析仪器。对这些应用，对于 GC/MS 结果的确信和准确性需要额外的确证工具。

最近在质谱数据处理上的进步，对于在单位分辨质谱上实现准确的分子式识别显示了巨大的潜力<sup>[1]</sup>。该应用基于精确质量数和线形校正的新技术，在单位分辨质谱上得到精确质量数和精确的同位素峰形匹配（谱图准确度）用于分子式识别，得到的结果与高分辨测量仪器具有很好的可比性。

本文呈现一个新方法，通过标准 GC/MS 谱库检索，并结合在标准单位质量分辨四极杆 GC/MS 仪器的分子式识别进行结果确认，提高 GC/MS 对未知物识别的速度和准确性。方法是基于 VOC 分析的标准 EPA 方法<sup>[2]</sup>（EPA 8260B），这里定义为 EVOCA（扩展的 VOC 分析）。

### 实验

使用标准 EPA 8260B 方法中 7 个标准替代化合物和内标，建立用于评价 EVOCA 的测试方案，该测试方案使用 73 个与环境相关的目标化合物，在另一个单独运行中有额外 17 个化合物，所以共有 90 个化合物用于评价该方法。

73 个目标化合物测试混合物溶解在 10 ml 水中，每种组分的浓度约为 20 ug/l，通过吹扫捕集系统和 50:1 的分流，每种大约 4 ng 的组分被引入质谱系统。17 种额外化合物制成约 2000 ug/ml 的浓度，进样 1 ul 并通过相同比例的分流，每种约 40 ng 的组分引入质谱系统。

Agilent 6890-5973 GC/MS 系统安装在 CT DEP 流动实验室用于全部分析。分析使用标准

---

#### 北京绿绵科技有限公司

北京市北四环西路 68 号左岸工社 806-807 室  
(100080)  
Tel: 010-8267 6061/2/3/4/5/6/7  
Fax: 010-8267 6068

沈阳市和平区南一马路力创大厦 503 室  
(110001)  
Tel: 024-2387 9100/2387 8588/2387 3099  
Fax: 024-2387 6558

杭州市天目山路 224 号中融城市花园 6 幢 2 单元 702 室 (310012)  
Tel: 0571-8765 8020/8717 8257  
Fax: 0571-8717 8257

EPA 8260B 方法，除了 MS 被设定为 Agilent 的"raw"模式（原始数据模式）采集谱图数据，原始采集模式采集和保存的 MS 数据是连续的，这点对于分子式识别的谱图校正是必要的。利用 Agilent Chemstation 软件对每种目标测试化合物在 NIST 02 谱库中进行谱图检索，并列出现在前面的检索结果。图 1 显示了 73 个目标化合物运行的色谱图和一个感兴趣化合物离子分析结果。额外的 17 种 VOCs 在独立的运行中使用同样方法进行分析，仅仅是样品引入方式不同。

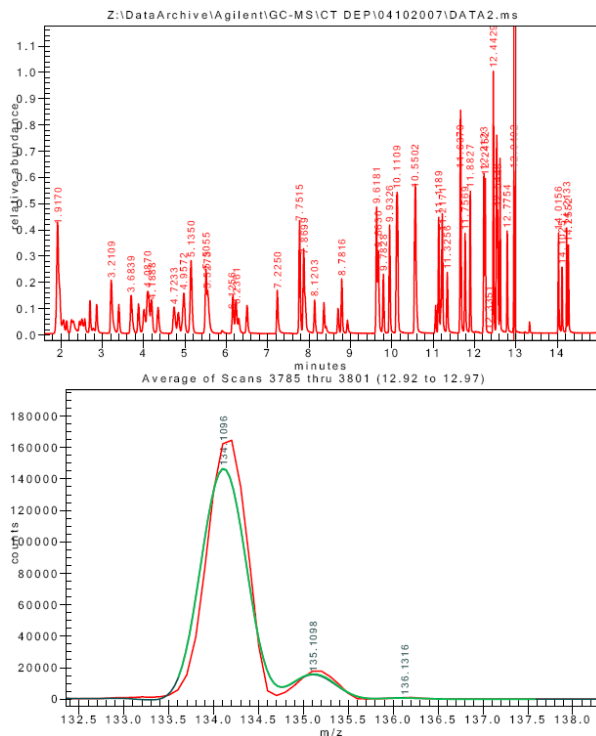


图 1. 顶图显示 73 种目标 EPA 化合物的总离子流图。底图显示 n-丁苯未校正（红色）和校正后及理论（绿和黑色）谱图的匹配结果，99.85%的光谱准确度匹配几乎完全重叠

每次运行后数据被装载进 MassWorks 软件 (Cerno Bioscience, Danbury, CT)，利用表 1 中列出替代化合物和内标的六个选择离子校正质谱数据，得到精确质量数和线形谱图，选择的校正离子质量涵盖了整个感兴趣的质量范围，提供了全面的谱图校正信息。校正然后应用在整个运行测试上，并对 73 种 EPA 目标化合物和 17 个额外 VOCs 中每种化合物谱图进行平均。表 2 显示了 73 种目标物中的部分化合物。根据全部分析结果的精确质量信息，利用每个化合物的单一同位素峰位最大 20 mDa 质量偏差，对感兴趣的离子进行分子式检索。既然，方法代表真正未知物分析，设定未知物的元素组成为 C、H、N、O、F、Br、Cl、S 和 P 进行检索，这些也是影响环境和健康的相关化学品中最常见元素。一旦分子式列表产生，MassWorks 校正线形谱图同位素轮廓搜索 (CLIPS) 的分子式检索功能被应用到选择离子，并根据谱图准确度列出可能分子式列表。谱图准确度是测量校正的线形谱图与每个候选分子式理论计算谱图匹配完美程度的指标：

$$\text{谱图准确度 (Spectral Accuracy)} = (1 - |RMSE|) * 100$$

这里 RMSE 是校正谱图与理论谱图误差的均方根，因此数据 100 是完美符合，图 2 说明了匹配的例子。得到的分子式结果然后利用谱库检索进行确认，以验证混合物测试的可靠性。

## 北京绿绵科技有限公司

北京市北四环西路 68 号左岸工社 806-807 室  
(100080)  
Tel: 010-8267 6061/2/3/4/5/6/7  
Fax: 010-8267 6068

沈阳市和平区南一马路力创大厦 503 室  
(110001)  
Tel: 024-2387 9100/2387 8588/2387 3099  
Fax: 024-2387 6558

杭州市天目山路 224 号中融城市花园 6 幢 2 单元 702 室 (310012)  
Tel: 0571-8765 8020/8717 8257  
Fax: 0571-8717 8257

表1. 用于质谱校正的离子和校正误差

Ion	Formula	Mass(Da)	Error (mDa)	Error (PPM)
Fragment 1,2-Dichloroethane-d4	C <sub>2</sub> Cl <sup>(2)H</sup> <sub>3</sub>	65.0112	0.1	0.8
Fragment Chlorobenzene-d5	C <sub>6</sub> ( <sup>2</sup> H) <sub>5</sub>	82.07	-0.5	-6.7
Fragment 4-Bromofluorobenzene	C <sub>6</sub> H <sub>4</sub> F	95.0297	-0.5	-5.4
Fragment 1,4-Dichlorobenzene-d4	C <sub>6</sub> ( <sup>2</sup> H) <sub>4</sub> Cl <sub>2</sub>	115.0253	-0.7	-5.8
1,4-Dichlorobenzene-d4	C <sub>6</sub> ( <sup>2</sup> H) <sub>4</sub> Cl <sub>2</sub>	149.9941	0.2	1.5
Dibromofluoromethane	C <sub>2</sub> HBr <sub>2</sub> F	189.8429	1	5.3

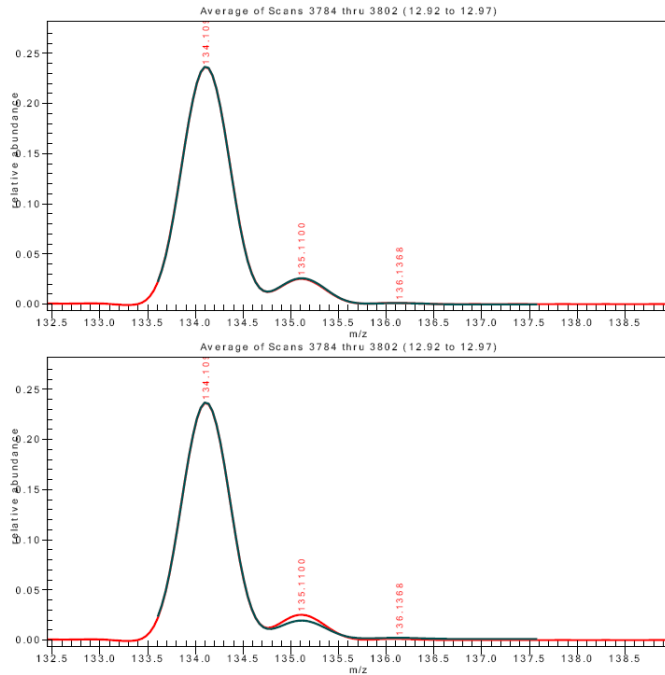


图2. CLIPS检索后n-丁苯谱图匹配结果。顶部显示谱图准确度为99.85%排名第一的校正谱图（红色）与理论谱图（黑色）的叠加图；底部显示谱图准确度为97.53%排名第二（C<sub>7</sub>H<sub>15</sub>OF）匹配图

表2. 94种目标VOCs中的部分化合物

Compound	CAS #	Formula	Mass (Da)
Toluene	108-88-3	C <sub>7</sub> H <sub>8</sub>	92.0626
(trans)-1,3-Dichloropropene	10061-02-6	C <sub>3</sub> H <sub>4</sub> Cl <sub>2</sub>	109.969
1,1,2-Trichloroethane	79-00-5	C <sub>2</sub> H <sub>3</sub> Cl <sub>3</sub>	131.93
Tetrachloroethene	127-18-4	C <sub>2</sub> Cl <sub>4</sub>	163.8754
1,3-Dichloropropane	142-28-9	C <sub>3</sub> H <sub>6</sub> Cl <sub>2</sub>	110.9768
Dibromochloromethane	124-48-1	CHClBr <sub>2</sub>	205.8134
1,2-Dibromoethane	106-93-4	C <sub>2</sub> H <sub>4</sub> Br <sub>2</sub>	185.868
Chlorobenzene	108-90-7	C <sub>6</sub> H <sub>5</sub> Cl	112.008
1,1,1,2-Tetrachloroethane	630-20-6	C <sub>2</sub> H <sub>2</sub> Cl <sub>4</sub>	165.8911
Ethylbenzene	100-41-4	C <sub>8</sub> H <sub>10</sub>	106.0783
m-Xylene + p-Xylene	108-38-3	C <sub>8</sub> H <sub>10</sub>	106.0783
o-Xylene	95-47-6	C <sub>8</sub> H <sub>10</sub>	106.0783
Styrene	100-42-5	C <sub>8</sub> H <sub>8</sub>	104.0626
Bromoform	75-25-2	CHBr <sub>3</sub>	249.7628
iso-Propylbenzene	98-82-8	C <sub>9</sub> H <sub>12</sub>	120.0939

## 结果和讨论

### 北京绿绵科技有限公司

北京市北四环西路 68 号左岸工社 806-807 室  
(100080)  
Tel: 010-8267 6061/2/3/4/5/6/7  
Fax: 010-8267 6068

沈阳市和平区南一马路力创大厦 503 室  
(110001)  
Tel: 024-2387 9100/2387 8588/2387 3099  
Fax: 024-2387 6558

杭州市天目山路 224 号中融城市花园 6 幢 2 单元 702 室 (310012)  
Tel: 0571-8765 8020/ 8717 8257  
Fax: 0571-8717 8257

对于大部分化合物，通过 NIST 谱库匹配，分子离子用于确定元素组成，但在某些情况下分子离子非常弱或没有分子离子峰，在这种情况下，可以用一个或更多的碎片离子用于确证正确的分子式。73 种目标化合物和 17 种额外 VOCs 质量准确度的标准偏差分别为 13 mDa 和 7 mDa，保守设定最大 20 mDa 的质量窗口用于元素组成搜索，由于设定的元素组成搜索范围很宽，虽然通常有 50-100 待选化合物，但对于更高分子量的化合物可能会得到更多的待选化合物。即使在高分辨 GC-TOF 仪器，理想条件下的 5 ppm<sup>[3]</sup>质量精度，利用这些限制元素组成搜索时也会得到约 20 种待选分子式，而对于“真实样品”分析，质量准确度可能低 2-4 倍。这说明通过同位素峰形进一步限制待选分子式数量对于得到更可靠的分析结果是非常重要的。

图 2 说明了使用 CLIPS 的同位素峰形匹配。谱图校正不但校正了线形谱图也校正了质量轴位置，利用列表中分子式计算出的理论谱图与测量谱图进行准确对比，不但计算出谱图准确度值，而且可以看到同位素轮廓比较图，验证其匹配程度。对于 NIST 或其它谱库中不存在的真正未知物，对其进行分析也是可能的，在这种情况下，该方法至少能提供未知物的化学组成信息，结合化学及其它线索帮助识别化合物。

表 3 显示了结果总结。值得注意的是，对于每个测试的化合物，得到的有用分子式信息用于谱库检索结果的确证。最坏的情况，苯乙烯 (C<sub>8</sub>H<sub>8</sub>, MW=104.0626 Da) 的分子离子 (名义质量为 *m/z* 104 Da) 用于化合物识别，谱图准确度排名第五，进一步研究表明，二甲苯与苯乙烯是共流出物，它也在 *m/z* 104 Da 产生 M-H<sub>2</sub> 的干扰碎片离子，苯乙烯分子离子的谱图准确度为 89.3%，远低于观察的大部分离子的 98% 典型值及更好的 99%。低谱图准确度对于不正确化合物识别是很好的指示，而质量准确度做不到这点。干扰现象很容易通过对比测量谱图和理论谱图进行证实 (如图 3)，明显看到不好的匹配结果，可以猜测是干扰造成的。当仔细研究和检查色谱峰后，猜测二甲苯是一个共流出物，为了证实这点，将二甲苯的分子离子和它的 M-H 离子包含进来重新进行 CLIPS 搜索，搜索自动计算这三个离子的最好匹配，得到 99.75% 更好谱图准确度，因此可以确认差的谱图准确度是干扰造成的。

表 3. 全部化合物的测试结果总结。根据谱图准确度排名，90 个化合物中除了 4 个外正确离子都排在前两位

Total # Compounds	Ranking by Spectral Accuracy				SD Mass Accuracy (mDa)
	1	2	3-5	<10	
73 VOC Target	65	3	4	0	13
17 VOC Test	16	2	0	0	7

该方法的优点是完全以现存的分析方法基础，使用替代化合物和内标，不需要额外的谱图校正标样如 PFTBA，因此对现存的分析方法不需改变。这简化了分析过程且改进了谱图校正，仅需要额外的数据分析步骤就得到分子式结果，这意味该方法很容易适合其它实验室。

将 CLIPS 搜索应用到多个离子碎片，很容易在 GC/MS 上进一步提高分子式识别的确证程度 (见表 4)，因为碎片离子的组成元素一定来源于其分子离子。这里，对我们有利的事实是 EI 离子化提供了丰富的化学信息，利用这个相同的方法，我们可以解决 NIST 谱库检索中分子离子峰很弱或没有的情况。在完全未知物时，碎片丢失也能给出结构线索。通过识别不相关分子和碎片离子，这些信息也能用于帮助识别化合物是共流出物。

## 北京绿绵科技有限公司

北京市北四环西路 68 号左岸工社 806-807 室  
(100080)  
Tel: 010-8267 6061/2/3/4/5/6/7  
Fax: 010-8267 6068

沈阳市和平区南一马路力创大厦 503 室  
(110001)  
Tel: 024-2387 9100/2387 8588/2387 3099  
Fax: 024-2387 6558

杭州市天目山路 224 号中融城市花园 6 幢 2 单元 702 室 (310012)  
Tel: 0571-8765 8020/8717 8257  
Fax: 0571-8717 8257

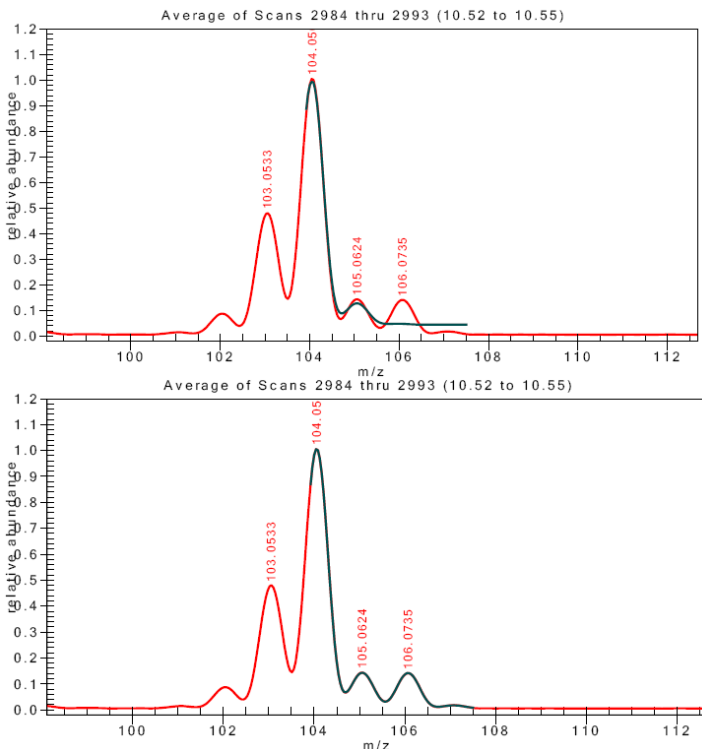


图 3. 顶图显示  $m/z$  104 苯乙烯离子 ( $C_8H_8$ , 红色) 与其理论谱图 (黑色) 叠加图, 匹配结果很差, 暗示有干扰离子存在。底图为包含了二甲苯 ( $C_8H_{10}$ ) 和脱氢的二甲苯的匹配图, 混合物匹配的谱图准确度为 99.75%, 确认存在二甲苯干扰。

表 4. 甲基丙烯酸乙酯两个碎片离子的 CLIPS 检索结果。正确的分子离子排在第二位, 通过分析另一个  $m/z$  69 离子, 正确的分子式能够根据  $m/z$  69 碎片 ( $C_4H_5O$ ) 丢失  $C_2H_5O$  推断出来

Search results for Molecular ion at 114						
Row	Formula	Mono Isotope	Mass Error (mDa)	Mass Error (PPM)	Spectral Accuracy	DBE
1	$C_8H_{11}P$	114.0598	-1.6369	-14.3511	98.0144	2
2	$C_8H_{10}O_2$	114.0681	-9.8796	-86.6186	97.9428	2
3	$C_8H_7N_2F$	114.0593	-1.1264	-9.876	97.9173	3
4	$C_8H_8NO_2$	114.0555	2.6965	23.6414	97.8515	2.5
5	$C_8H_{10}S$	114.0503	7.879	69.0787	97.7656	2
6	$C_4H_8N_3O$	114.0667	-8.5369	-74.8468	97.6723	2.5

Search results for fragment at 69						
Row	Formula	Mono Isotope	Mass Error (mDa)	Mass Error (PPM)	Spectral Accuracy	DBE
1	$C_4H_5O$	69.034	1.3602	19.7032	99.0029	2.5
2	$C_3H_5N_2$	69.0453	-9.8732	-143.016	98.8248	2.5
3	$CHN_4$	69.0201	15.2789	221.3205	98.5772	3.5
4	$C_2H_7F_2$	69.0516	-16.1816	-234.396	98.3387	-1.5
5	$CH_5O_2F$	69.0352	0.2174	3.1486	98.2205	-1.5
6	$CH_7FP$	69.0269	8.4601	122.5468	98.2194	-1.5

## 结论

正像表3说明的, 90个测试化合物中, 81个化合物确证的分子式在CLIPS检索列表中排在第一位, 5个化合物排在列表的第二位, 总共超过95%化合物具有很高的确证效率, 其余的四个化

合物排在3-5的位置，该结果清楚表明，通过这个方法得到的谱图准确度和质量精度对于GC/MS谱库检索提供了高度可靠的确证。这个方法关键优势在于，能够运行在成本较低的流动GC/MS仪器，通常可用于个人负责的初始响应形式，这个额外确认能力在紧急时刻甚至生命危机时刻是非常关键的。

同样方法也很容易扩展到其它应用领域，GC/MS系统对未知半挥发有机化合物（SVOCs）进行快速识别，包括香精香料、药物降解、天然产物识别和污染等等。这种方式结合化学离子化GC/MS对于未知化合物结构解析提供额外的有价值信息。

### 参考文献

- [1] Gu, Ming, et al "Accurate mass filtering of ion chromatograms for metabolite identification using a unit mass resolution liquid chromatography/mass spectrometry system", Rapid Commun. Mass Spectrom. 2006; 20: 764–770.
- [2] EPA METHOD 8260B VOLATILE ORGANIC COMPOUNDS BY GAS CHROMATOGRAPHY/MASS SPECTROMETRY (GC/MS) (see [www.epa.gov](http://www.epa.gov))
- [3] Manufacturers published specification for the Waters GCT Premier. Waters Corporation, Milford, MA.
- [4] Kind, T. et al "Metabolomic database annotations via query of elemental compositions: Mass accuracy is insufficient even at less than 1 ppm", BMC Bioinformatics 2006, 7, 234 ([www.biomedcentral.com/1471-2105/7/234](http://www.biomedcentral.com/1471-2105/7/234)).
- [5] S.E. Stein and D.R. Scott, "Optimization and testing of mass spectral search algorithms for compound identification", J. Am. Soc. Mass Spectrom 5 (1994), pp. 859–866.

---

## 北京绿绵科技有限公司

北京市北四环西路 68 号左岸工社 806-807 室  
(100080)  
Tel: 010- 8267 6061/2/3/4/5/6/7  
Fax: 010- 8267 6068

沈阳市和平区南一马路力创大厦 503 室  
(110001)  
Tel: 024- 2387 9100/2387 8588/2387 3099  
Fax: 024- 2387 6558

杭州市天目山路 224 号中融城市花园 6 幢 2 单元 702 室 (310012)  
Tel: 0571- 8765 8020/ 8717 8257  
Fax: 0571- 8717 8257