

碳氢化合物质谱的计算机解释

朱大模 王洛秋 余建文 许崇德 刘人余
(中国科学院大连化学物理研究所)

[摘要]本文用谱库检索和智能解释程序相串联鉴定碳氢化合物。谱库检索应用新的正、反检结合型原理建成。智能解释程序,应用从大约 2000 个烃类化合物质谱提取的结构-谱图联系规律,解释质谱推导分子结构。

一、前 言

质谱法是现代有机结构分析的重要仪器分析方法。色谱/质谱联用是分析微量复杂有机混合物的最优方法。目前,质谱、色谱/质谱联用仪在我国已普遍用于石油地质、开采、炼制及石油化工样品的分析中。近年来我国从国外引进的质谱或色谱/质谱联用仪都带有一个质谱数据库检索系统,内含近两千个碳氢化合物的质谱,提供用质谱检索法鉴定这些化合物。实践证明从国外引进的这些检索程序已不能满足我国石油及石油化工工业发展提出的分析课题的需要,迫切要求解释质谱鉴定烃类化合物分子结构的方法。本文报导的碳氢化合物质谱的计算机解释系统是一个质谱检索程序和一个质谱解释程序的串联形式。利用检索程序可以快速、准确地鉴定数据库中有标准质谱数据的化合物。解释程序用来解释数据库中无标准谱图数据的试样质谱推导其分子结构。

二、方 法

图 1 表示出质谱解释系统的工作过程。研究试样的低分辨电子轰击质谱数据进入计算机后,质谱逆检索程序把数据库中的每一标准质谱调来与之比较,若有良好匹配存在,程序将最佳匹配的几个标准数据化合物的名称、分子式、分子量按相似度顺序输出当作试样物质的鉴定结果,供分析者选用。如果没有良好匹配(或分析者对检索结果不满意),程序自动地解释样品谱、推导分子结构。质谱解释以推导化合物的分子量开始。数据解释功能依据分子结构与质谱之间的联系规律解释质谱,详尽地提供试样分子的各个片断结构。分子结构式组建功能按一切可能方式利用这些片段结构组合成符合分子量要求的各

1988 年 1 月 20 日收

个可能的分子异构体。质谱预测功能利用分子结构与质谱之间的联系规律,对每一推得的分子结构式提供一个预测质谱。把各个预测质谱同样品的实测质谱比较评价各个结构式的优劣,并按顺序排列输出。整个系统在 IBM/PC-XT(或 AT)上运行。

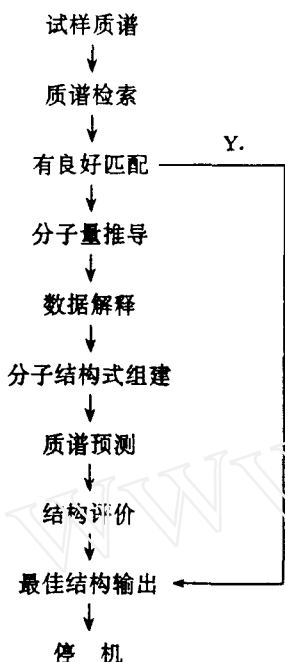


图1 质谱解释系统的功能逻辑

正检索同时承认已知的“标准质谱”和样品质谱都是纯化合物的完整质谱数据。这一假定在某些情况是不被良好满足的。例如,从 GC/MS 联用得到的质谱数据中经常包含有因色谱固定液流失的质谱峰,因相邻组分未能完全分开而带来的干扰物质峰。所以,存在着一定程序的“试样不纯”。这是问题的一方面。另一方面,分析者总是把自己得到的试样质谱数据原原本本地送入计算机做质谱检索,同数据库中的“标准数据”比较。同什么样的“标准质谱”数据比较? L. Domokos 等^[1]的研究,指出 EPA/NIH 和 CORNELL 的质谱数据库中的质谱存在着严重的质量问题。我们发现 EPA/NIH MSDB (1978)中 $M=112-185$ 的质谱中有 4.5% 存在着错误。这两家及 NBS 的数据库质谱是当今世界上的主要商品形式,它们都源于 Wiley 公司。这就是目前国际上质谱检索发生错误的原因。

逆检索只承认“标准质谱”数据是标准的,至于样品质谱,它可是纯化合物的,也可是混合物的。从

原则上讲,反检索有直接从混合物质谱鉴定混合物组成成分的能力。但是,对反检索的这一优点不能强调过分。因为,任何大分子的复杂质谱,都可以分解为较小分子的简单质谱加“杂质”的质谱之和。这就是反检索发生错误的原因。McLafferty^[2] 学派对其 PBM 的补充证明了这点。本工作中使用正反检结合型原理的检索程序。正检索原理用来建立预检索程序,用反检索原理建立终止程序。预检索在宽松的条件下工作,它不怕杂质的干扰。反检索是在预检索结果中进行,也就克服了自身单独使用时的缺陷。所以,以正反检结合型原理为基础的程序,特别适合于色谱联用分析。除基本原理外,检索程序的另一优点,是其标准质谱数据可靠性强。我们修正和删除了美国 NBS 质谱数据库(42000 谱图)中的近一千张错误质谱数据,补充了近期文献质谱四千余化合物。总计四万五千余化合物。

关于解释程序,方法的基本假定是:任何复杂化合物的分子结构都可用一定种类的基本结构单元来组合产生。这些基本结构单元都有自己确定的质谱特征,或表现为一组特征的正离子,或表现为一组中性丢失,或同时表现在这两方面。这些结构单元的质谱特征,即结构——谱图规律,可由相关化合物的质谱数据获得。用这样的结构——谱图联系规律为基础建立的数据解释功能和质谱预测功能程序就具有解释质谱、预测质谱的能力。由此可见,该系统的结构鉴定能力决定于所用的结构——谱图联系规律的结构类型完善和质量可靠的程度。目前系统利用了从大约 2000 个烃类化合物的质谱中抽出的结构——谱图

规律,它的结构鉴定能力大于 100 万个烃类化合物。

分子结构式组建程序组建的各个异构体的分子量必须等于推得的分子量。因此,分子量限定了某些结构特征在组建分子结构式时是否需要被重复利用。建立分子量推导程序的基础是合理的中性丢失。所以,当试样质谱中有分子离子时,它可由合理的中性丢失得到确证;在没有分子离子时,利用合理的中性丢失规律可以推导分子量。

三、应 用

质谱检索,方法简单,对数据库中有数据完整的正确标准谱图的化合物,都能得到正确鉴定,无讨论的必要。检索方法不能鉴定数据库中无标准数据的化合物,是其严重缺陷。解释程序则有鉴定数据库中无标准数据化合物乃至新化合物的能力。用本程序解释图 2 的质谱数据。在第一次解释中操作者告诉计算机,强度 $\geq 3\%$ 的离子是可靠的,解释结果列于表 1。表 2 的结果对应于 $\geq 1\%$ 强度值的离子是可信的情况。表中推出结构的可靠性由不相似度 DISS 指示;最小的不相似度对应于最高的可靠性。化学合成证明 2,2,5,6-四甲基庚烷是正确结构。在 $I \geq 1\%$ 为可信的情况,由于有更多的信息可应用,推得结果更准确,这时正确结构由 $I \geq 3\%$ 时的第三位变为第一位。 $I \geq 3\%$ 的离子是可信的,意味着计算机把 $I < 3\%$ 的离子当作杂质看待,即图 2 中 m/z 99 以上的峰都被略去了。表 1 的结果说明用该程序解释被“污染”的正确质谱的能力是良好的。

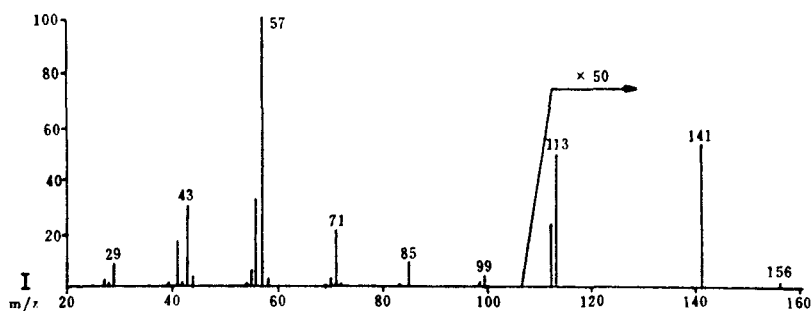


图 2

表 1 计算机推出结构 ($I=3.0\%$)

DISS	MW	结 构 式
0.048	114	$ \begin{array}{c} \text{C} \quad \quad \text{C} \\ \quad \quad \\ \text{C}-\text{C}-\text{C}-\text{C}-\text{C} \\ \\ \text{C} \end{array} $
0.052	128	$ \begin{array}{c} \text{C} \quad \quad \quad \text{C} \\ \quad \quad \quad \\ \text{C}-\text{C}-\text{C}-\text{C}-\text{C}-\text{C} \\ \\ \text{C} \end{array} $

DISS	MW	结构式
0.056	156	<pre> C C C C---C---C---C---C---C C </pre>
0.065	156	<pre> C C C---C---C---C---C---C---C C </pre>
0.069	156	<pre> C C C C---C---C---C---C---C C </pre>

表 2 计算机推出结构 (I=1.0%)

DISS	MW	结构式
0.042	156	<pre> C C C C---C---C---C---C---C C </pre>
0.052	170	<pre> C C C C C---C---C---C---C---C C </pre>
0.052	142	<pre> C C C---C---C---C---C---C C </pre>
0.055	156	<pre> C C C---C---C---C---C---C---C C </pre>
0.060	156	<pre> C C C---C---C---C---C---C C C </pre>

参 考 文 献

- [1] L. Domokos, D. Henneberg and B. Weimann, *Anal Chim Acta*, 150 (1983) 37
- [2] 1985 Pittsburgh Conference and Exposition on Analytical Chemistry and Applied Spectroscopy, Abstracts, NO. 499

A Computer Program for Mass Spectral Interpretation of Hydrocarbons

Zhu Damo, Wang Luoqiu, She Jianwen

Xu Chongde, Liu Renyu

(Dalian Institute of Chemical Physics, Academia Sinica)

Received 20, Jan. 1988

Abstract

A series program which consists of a library search and an intelligence interpretation program was developed for hydrocarbon identification. The library search program uses a new combined forward and reverse search principle. The intelligence program infers molecular structure of hydrocarbon using spectrum-structure correlation rules inferred from about 2000 reference spectra of hydrocarbons.