

DOI: 10.13376/j.cbbs/2014039

文章编号: 1004-0374(2014)03-0261-09



黄超兰 研究员

黄超兰研究员实验室着重在质谱和基于质谱的蛋白质组学的方法开发。主要集中在: (1) 应用原位和离子淌度质谱 (Native and Ion Mobility Mass Spectrometry) 以至上而下的途径对大分子复合物的鉴定, 以及结构和动态研究; (2) ncRNA 和 ncRNA 相互作用蛋白的直接和高通量的检测与定量; (3) 低丰度蛋白和翻译后修饰 (PTM) 的富集。

RNA的质谱分析研究进展

彭 超, 黄超兰*

(中国科学院上海生命科学研究院生物化学与细胞生物学研究所, 国家蛋白质科学中心, 上海 201210)

摘要: 核糖核酸 (RNA) 是一类重要的遗传物质, 在生物体内起着至关重要的作用。传统的 RNA 测序鉴定技术既费时又费力, 而且会丢失大量的 RNA 转录后修饰等方面的信息。随着近年来仪器的更新和技术的飞速发展, 质谱技术在生物大分子研究方向得到了很好的应用, 特别是基于质谱技术的蛋白质组学。相比蛋白质组学的研究来说, 基于质谱技术的 RNA 鉴定分析更为困难, 发展相对缓慢。但是, 硬件方面的革新和软件方面的开发也为其发展带来了很好的契机。重点通过对质谱技术在 RNA 分析鉴定的介绍, 从样品制备、离子化方式、碎裂方式、分析方法、数据后处理软件等几个方面展现了基于质谱技术的 RNA 研究进展。

关键词: 电喷雾电离; 基质辅助激光解吸电离; 静电场轨道阱; 核糖核酸; 数据库检索

中图分类号: Q524; Q-33 **文献标志码:** A

The developments of RNA analysis using MS technology

PENG Chao, Catherine C.L. Wong*

(National Center for Protein Science (Shanghai), Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 201210, China)

Abstract: Ribonucleic acids (RNA) are transcribed from unannotated genomic regions and many of them play an important role in a variety of cellular processes, such as, chromatin remodeling, transcriptional regulation, precursor messenger, etc. The conventional biochemical tool for analysis of RNA species is time consuming and often ineffective for characterizing modified RNA. Due to inefficient ionization of the gas-phase ions, MS has been less frequently used for RNA analysis than it has been for protein research. Recent technical improvements in both instrumentation and software make MS a powerful tool for RNA analysis because it can now be used to sequence, quantify, and chemically analyze oligonucleotides and intact nucleic acid. This review focuses on the principle of the MS-based RNA studies to give a view of RNA analysis using mass spectrometry, such as, sample preparation, ionization, software for data analysis.

Key words: ESI; MALDI; Orbitrap; RNA; mass mapping

收稿日期: 2014-02-12

*通信作者: E-mail: catherine_wong@sibcb.ac.cn

核糖核酸 (ribonucleic acid, RNA) 是以脱氧核糖核酸 (deoxyribonucleic acid, DNA) 为模板转录而成的重要遗传物质, 是遗传信息传递过程中的重要桥梁。RNA 在生物体内的很多细胞生理过程中发挥非常重要的作用, 如染色质重组、细胞内的转录调控、前体 mRNA 的加工、基因沉默、翻译调控等等^[1-2]。而与核糖核蛋白结合的非编码 RNA 在 RNA 的剪切与连接、RNA 特异性位点转录后修饰、DNA 的甲基化以及端粒的合成等生理过程中亦起着至关重要的作用。通过基于质谱技术的蛋白质组学技术, 人们已经证实了一些核糖核蛋白复合物中所包含的相关组成蛋白质^[3-5]。大部分实验也证实了 RNA, 如小核 RNA (snRNA)、核糖体 RNA (rRNA) 等, 在核糖核蛋白的合成以及后续的功能行使中起到了一定的作用^[6-7]。所以, 为了更加清楚地了解 RNA 在生物过程中的作用和调节机制, 除了开展 RNA 相关联蛋白质的研究鉴定以外, 对 RNA 自身的研究及鉴定也是相关研究中非常关键的内容。

近几十年来, 基于质谱技术的蛋白质组学得到了飞速的发展, 也为进一步的生物学各项研究提供了很大的帮助。而基于质谱技术的 RNA 研究却少有报道, 进展缓慢。相对于蛋白质来说, RNA 主要有以下几点不同。(1) 蛋白质由 20 多种氨基酸残基所组成, 而 RNA 是由 4 种核糖核苷酸 (A) 腺嘌呤核苷酸; G, 鸟嘌呤核苷酸; C, 胞嘧啶核苷酸; U, 尿嘧啶核苷酸; 其中, U 取代了 DNA 中的 T, 如图 1) 组成, 序列的重复性高、特异性较差。(2) 从性质上来说, 蛋白质来源的多肽易于形成正离子, 而且离子化的效率比较高, 便于质谱检测; RNA 则由于骨架中存在大量的磷酸基团而具有很强的亲水性, 容易形成负离子, 离子化比较困难、效率相对低下, 不便于分离与检测。传统意义上的 RNA 鉴定方法主要有两种: 一, 利用同位素标记和 RNA 酶解的方法^[8-9]; 二, 利用反转录酶得到相应的 cDNA, 利用 PCR 扩增以后进行 DNA 测序从而获得相应 RNA 的信息^[10]。前者需要大量的样品, 而且样品处理比较费时费力; 后者虽然在 RNA 定性与定量方面均得到了很好的应用, 却会丢失 RNA 中很重要的转录后修饰信息。

但是, 质谱技术尤其是用于研究生物大分子的生物质谱技术的飞速发展, 如电喷雾电离 (electrospray ionization, ESI) 和基质辅助激光解吸电离 (matrix-assisted laser desorption/ionization, MALDI) 离子化技术的发现, 高分辨率、灵敏度、准确度质谱仪的研发,

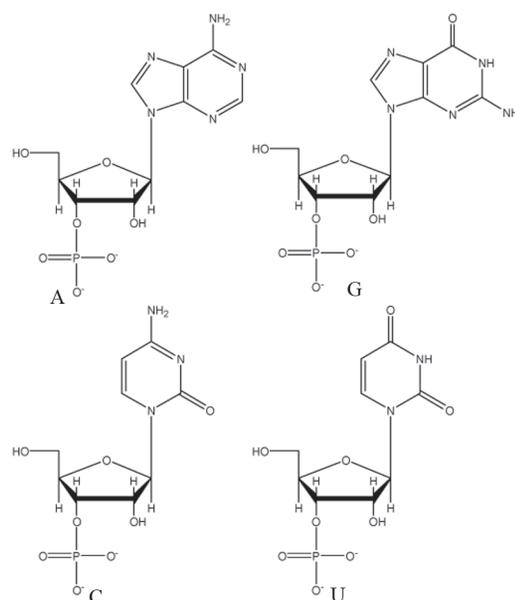


图1 组成RNA的四种不同核糖核苷酸结构

串联质谱技术的开发等等, 使得生物质谱技术在核酸研究领域也取得了一定的进展。

本文将具体从以下几个方面来介绍生物质谱技术在核酸分析领域中的研究进展。

1 质谱仪的基本原理

质谱仪是一类通过测定不同物质的相对分子质量 (M_r) 来对其进行研究的工具。简单地说, 质谱的操作部件主要由硬件的质量电荷比 (m/z) 检测和软件的数据分析两部分组成。硬件部分比较核心的几个组件, 包括了样品离子化、质量分析器和离子检测器等; 软件部分则主要是对质谱控制以及数据的后处理分析。

样品的离子化方法有很多, 在生物质谱中应用最广的主要有两种软电离方式: 电喷雾电离 (ESI) 和基质辅助激光解吸电离 (MALDI)。ESI 和 MALDI 这两种离子化方式所需条件均比较温和, 特别适合生物大分子的质谱分析。它们的发现及其在生物大分子研究方向的应用大大促进了生物质谱技术的发展。

ESI^[11] 的特点是可以产生多电荷离子, 使质荷比降低到多数质量分析仪器都可以检测的范围, 因而大大扩展了仪器对相对分子质量的分析范围; MALDI^[12] 则是通过激光首先使基质带电, 而气化的带电基质和样品之间发生碰撞, 再把激光的能量传递给样品, 从而导致样品的离子化, 适用于混合

物及生物大分子的鉴定分析。而串联质谱的开发,则是通过对质量分析器不同的组合使用,从而达到更为特异性的目的,如生物大分子蛋白质的鉴定,蛋白质后修饰的定性与定量分析等等。

质量分析器是质谱仪的核心,目前主要分为四类:离子阱(Ion trap)、飞行时间(time of flight, TOF)、四极杆(Quadrupole)和傅立叶变换离子回旋共振(Fourier transform ion cyclotron resonance, FT-ICR)。它们在设计 and 构造上各有不同,因而各有优缺点。质量分析器决定整个机器的分辨率、质量准确性、灵敏性和质量检测范围。其中具有满足生物大分子质谱分析的质量分析器有三类:傅里叶变换-离子回旋共振质量分析器、静电场轨道阱(Orbitrap)质量分析器和时间飞行质量分析器。FT-ICR-MS具有超高质谱分辨率、高质量测量准确度、回旋池内现场反应等显著优点,但是,价格昂贵,仪器操作复杂,维护成本高;Orbitrap,是第一个在静电场中进行离子捕获的高性能质量分析器,结合线性离子阱技术可以实现离子分离、裂解以及多级质谱功能,在质量准确度、分辨率、动态范围、灵敏度以及多级质谱能力等方面具有明显优势,在降低运行成本的同时又可以得到高分辨率的数据结果;MALDI-TOF质谱则采用一系列的高新技术来提升其各个方面的性能,如提供二阶无网离子反射器来延长离子在飞行管中的飞行距离,创新地使用LIFTTM技术提升能量来高速完成高质量的MS/MS质谱数据,采用独有的PANTM全景宽域聚焦技术来实现宽质量范围内的高分辨率等等。

近10年来,质谱技术的重要进展还体现在另外两方面:(1)串联质谱的开发,就是对上述质量分析器进行不同的组合,以达到多级质谱分析等特异性目标;(2)质谱仪另一个重要进展不在于技术层面上,而是在仪器化方面,商业化的仪器不断更新,推动了质谱在各个应用领域的快速发展。各厂家为满足客户的需要,尤其是生命科学领域的需要,组合不同的特殊电离技术以及各种质量检测器,生产出超高分辨率、高灵敏度、宽质量范围的质谱仪;把质谱与气相色谱、高效液相色谱系统联用,大大拓宽了质谱的应用范围。正是各个方面的这些改进,大大地推动了质谱技术在核糖核苷酸分析方面的进展。

2 RNA的质谱分析进展

基于质谱技术的蛋白质组学研究,主要是通过

对蛋白质或其酶解产物的串联质谱分析数据与理论数据库进行比对,来对蛋白质进行定量与定性方面的分析。其主要分为以下几个步骤(如图2所示):首先,从生物样品中分离得到总RNA或者纯化得到特定的RNA;紧接着对得到的RNA进行酶解处理,得到合适大小的RNA片段;所得到的样品进行一级或者多级的质谱分析,获得相关的质谱数据;最后,结合各种不同的分析软件或者数据库检索方法,完成对RNA及其后修饰的鉴定。

在质谱仪中RNA的断裂方式(图3)与蛋白质比较类似,主要是发生在骨架结构中的磷酸基团上,从而产生带电的a、b、c、d离子片段和相应的w、x、y、z离子片段。通过对离子片段的比对分析,首先可以确定寡聚核苷酸的序列信息,而结合多个寡聚核苷酸的鉴定结果就可以完成对RNA的定性分析。

2.1 RNA质谱分析中用到的两种电离方式

与基于质谱技术的蛋白质组学研究类似,RNA的离子化仍然主要依赖于两种软电离方式:电喷雾电离ESI和基质辅助激光解吸电离MALDI。20世纪80年代,Fenn等^[11]和Covey等^[13]提出并证实了软电离的离子化方式可以很好地用于质量数超过质谱仪器质荷比上限的物质检测。ESI的电离方式可以使生物大分子带上多电荷,从而大大地降低其质荷比,达到质谱检测的目的。ESI要求物质能够在溶液状态下带上一定的电荷,而RNA骨架上的磷酸基团正好可以使其带上负电,满足了ESI离子

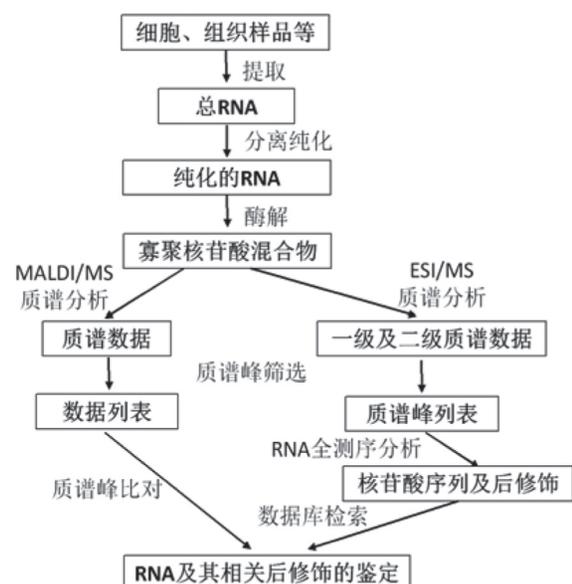


图2 RNA样品质谱分析流程

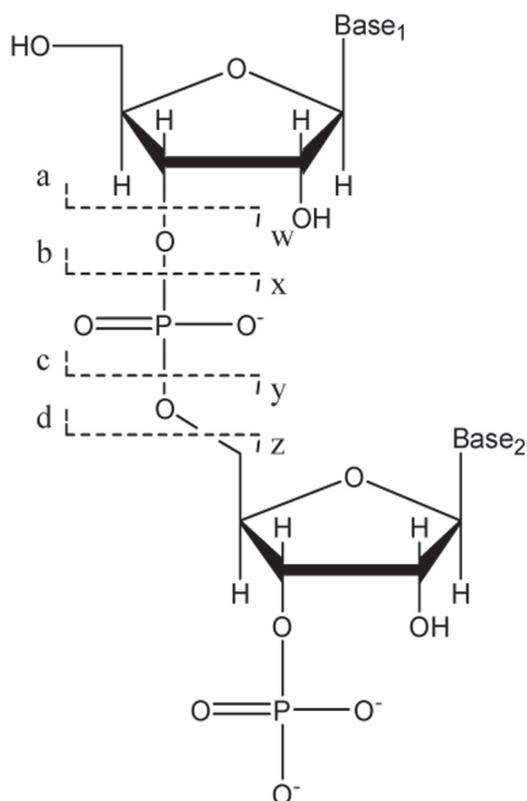


图3 质谱检测中RNA的碎裂方式

化以后质谱负电荷检测模式的要求。随后，在仪器与实验方法方面相关的发展更是开启了高准确率、高灵敏性地分析核糖核苷酸的大门。相比来说，MALDI的电离方式更加容易产生单电荷的离子，对于大分子样品，也会产生多电荷的离子。MALDI结合不同性质的基质，使其可以兼容不同物质的离子化和质谱分析，应用的范围更为广泛。在MALDI的离子化条件下，核糖核苷酸可以实现正离子和负离子两种模式的检测，其中低相对分子质量的样品更加适合负离子模式的检测，而高相对分子质量的样品则更加适合正离子模式的检测。这主要是由于现有的仪器增强了正离子模式下高相对分子质量离子峰的检测效率，而并不是由于高相对分子质量的样品在正离子模式下的离子化效率得到了很大的提高。

2.2 RNA质谱分析中的样品制备方法

样品的前期处理在整个质谱分析过程中起着至关重要的作用。对于MALDI与ESI两种不同的离子化方式来说，样品前期处理的基本要求比较类似，也有一些针对不同离子化方式的独特的样品处理方法。由于非挥发性正离子（如 Na^+ 、 K^+ 等）与RNA

形成的加合峰会大大影响RNA在质谱中的分析，所以，RNA质谱分析的主要目的就是排除这些因素的影响。

在ESI-MS质谱分析中，有很多脱盐或者纯化方法可以用于核酸样品的前处理，主要分为以下几类：(1) 利用离子交换的方法，把RNA上加合的 Na^+ 或者 K^+ 置换为 H^+ 或者 NH_4^+ ；(2) 有机溶剂沉淀的方法，如把溶于醋酸铵的寡聚核苷酸通过预冷的乙醇沉淀法，从而排除离子加合物的干扰；(3) 微量透析法，Muddiman等^[14]于1996年首次报道，他们成功并高效地利用微量透析法完成了PCR聚合产物样品混合物的脱盐，并成功地鉴定到了一个89 nt的寡聚核苷酸；(4) 在蛋白质组学中常用的C18反相分析柱脱盐法在处理某些RNA样品时也有很好的效果，如Little和McLaffery^[15]于1995年结合C18脱盐法很好地检测到了酵母来源的tRNA，其中残留了非常少量的金属离子加合峰。

相对于ESI的样品处理来说，MALDI的离子化手段与盐的兼容性会更好一点，但是，阳离子加合离子的存在还是会非常大地降低检测的灵敏度和分辨率。在ESI中所用到的脱盐法在MALDI中基本都可以取得很好的效果。除此之外，还有一些其他的脱盐方法在MALDI离子化方法中也有很好的效果。(1) 醋酸铵处理过的阳离子交换树脂法：这种树脂可以在点板之前，加入样品、基质甚至样品基质混合物中，直接点到样品板上，但是，它的载量有一点限度而且会影响样品的结晶；(2) 有机溶剂法：有机溶剂特别是有机碱，与基质一同加入样品来减少阳离子加合物的产生。Pieles等^[12]于1995年首先提出和尝试了有机溶剂法，他们在THAP基质中加入DAHC达到了降低阳离子加合物的目的。其他效果较好的有机溶剂还有三乙胺、哌啶、咪唑等。另外，有报道指出聚氨类化合物（如精胺、亚精胺等）也可以很好的加强MALDI-MS对于核酸样品的检测。

2.3 RNA质谱分析方法

RNA的自下而上的质谱分析(bottom-up MS analysis)流程如下(以RNP complex为例)：首先，从分离纯化的RNP complex中，提取出纯化的RNA混合物；利用RNA特异性内切酶对纯化得到的RNA混合体系进行剪切，产生一定长度范围的寡聚核苷酸混合物；进一步进行串联质谱分析得到相应寡聚核苷酸的一级和二级质谱谱图；最后，结合不同分析软件和数据库，完成RNA的测序分析及

其相应的转录后修饰的定性分析。

而 RNA 自上而下的质谱分析,则是结合高分辨质谱仪与多种核糖核苷酸的碎裂方式,如 EDD、CAD、CID 等,对完整的 RNA 大分子直接进行二级质谱的分析,避免了繁琐的酶切和寡聚核糖核苷酸分离的步骤,保留了更多的结构和转录后修饰的信息,但是,对于仪器和样品的要求更高。Brain Chait 指出只要能够从自上而下的质谱分析获得足够的片段信息,就可以了解到 RNA 完整一级结构、相关转录后修饰情况以及后修饰之间的相互联系等等重要信息。2009 年,Scott A. McLucky 实验室就通过离子阱 CID 的碎裂方式完成了对于 tRNA 的自上而下的质谱分析,覆盖率达到了 60%^[16]。2012 年,BreuKer 实验室通过结合 EDD 与 CAD 两种碎裂方式,对 22 nt 和 34 nt 长度的 RNA 分析的覆盖率可以达到 100% 和 97%,对高度后修饰的 tRNA 分析的覆盖率亦可达到 89% 以上^[17]。

2.4 RNA 质谱分析的后期数据处理

现有用于处理 RNA 相关的质谱数据分析软件主要可以分为四类:(1)Peak Picking 软件:主要用于根据核苷酸数据库中已有序列的理论 M_r 的计算结果,以及原始质谱数据的处理,如 Mongo Oligo Mass Calculator、SpliceCmd 等;(2)MS-based mass mapping 软件,如 RNAccess、RMM 等;(3)MS/MS-based de-novo sequencing 软件,如 SOS、COMPAS 等;(4)数据库检索(Database searching)软件,如 Ariadne 等。

2.4.1 Peak picking 软件

Mongo Oligo Mass Calculator 主要可以用于计算所给 RNA 或者 DNA 的理论 M_r 或者在 CID 碎裂方式下的质谱信号碎片峰,并可以根据要求拟合出在一定的条件下, RNA 或者 DNA 的理论质谱信号图谱。在 Mongo Oligo Mass Calculator 中,除了常见的 5 个碱基以外,也可以加入自定义的残基,或者是加入已知的相关后修饰类型,切换正离子或者负离子模式。

此类软件的另外一个功能就是同位素质谱信号峰的辨别与筛选。在质谱数据分析中,只有化合物的单同位素峰是完全由一种同位素组成的,它对应的 M_r 才是化合物的理论 M_r ,对于化合物的定性和定量分析才具有至关重要的作用。所以,对于化合物的单同位素峰的辨别和筛选,是质谱数据分析中关键的一步。

此步骤也可以手动完成,但是,由于数据量的

巨大,相关数据分析软件是非常有必要的。通过分析软件对原始数据的平滑处理、信号峰的识别与辨认,可以很好地得到单同位素的质谱信息以及所带电荷数情况。由日本 Mitsui Knowledge 公司推出的 SpliceCmd 软件在此方面得到了很好的应用。2005 年,Emili 实验室通过比较发现由于在多肽质谱数据分析中也存在类似的过程,而核苷酸与多肽存在着一定的相似性,很多用于多肽分析的软件只要进行特定的设置,在核苷酸分析方面也具有很好的应用^[18]。

而在核苷酸分析过程中有一点必须十分注意,C 与 U 的相对分子质量差别是 0.984 Da,而 C 的两个同位素(¹²C 与 ¹³C)的差别是 1.003 Da,非常相近。这样的差别(0.019 Da)很容易造成不同信号峰的重叠,需要更高分辨率的质谱分析器(分辨率大于 100 000)。

2.4.2 MS-based mass mapping 软件

此类软件只要用于通过比较实际检测到的 RNA 酶解产物碎片峰与理论值来对 RNA 进行鉴定。2006 年,Wilson 实验室首先利用此种方法结合质谱分析实现了特定 16S rRNA 的鉴定。他们首先建立了一个由 1 921 个 16S rRNA 的 RNA 水解酶 T1 或者 RNA 水解酶 A 的碎片数据库,发现 RNA 水解酶 T 所得数据库中大于 9 个碱基的片段在质谱上具有很好的识别度^[19]。通过所得质谱数据与此数据库的比对分析,他们就可以完成给定来源的单个 16S rRNA 或 16S rRNA 混合物的鉴定。随后,Jackson 等^[20]通过加入“谱图可信度”功能,可以进一步了解所得到的可信度方面的信息。2007 年,Hossain 和 Limbach^[21]则是采取了另外一个不同的方法:他们利用不同 RNA 中特异性的核苷酸片段实现了对每个 tRNA 的鉴定。据此结果,他们开发了一个相关的分析软件——RNAccess,可以很好地用于高含量 RNA(如 rRNA、tRNA 等)的质谱数据分析。

2009 年,Taoka 等^[22]通过统计发现,人类基因组所产生大约 2 000 个 21 nt 长的核糖核苷酸可以很好的在分辨率足够好的质谱仪上得到很好的分离,而不会出现同位素峰重叠的情况。以此理论为基础,Matthiesen 和 Kirpekar 等^[23]随后开发了一款适合于所有 RNA 质谱数据分析的软件——RMM(RNA mass mapping)。其原理如图 4 所示:通过实际实验所得到的 RNA 酶解核苷酸片段的质谱数据与理论上拟合的质谱数据进行比对来实现 RNA 的

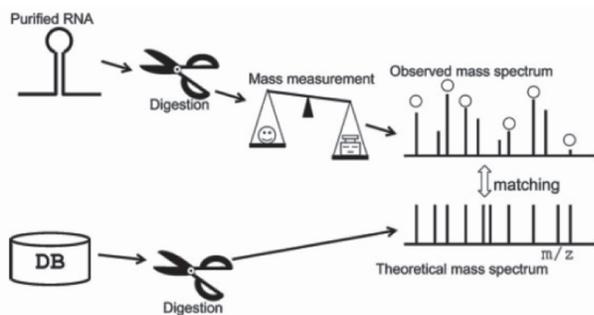


图4 MS-based mass mapping软件数据分析原理

鉴定；在原理上，与蛋白质组学中自上而下的鉴定方法具有一定的相似性。在核苷酸修饰分析方面，RMM具有一定的局限性，会大大增加分析的时间。

2.4.3 MS/MS-based *de-novo* sequencing软件

1992年，McLuckey等^[24]首次报道了带多电荷的核苷酸在ESI-CID质谱模式下的碎裂情况。在前人工作的基础上，Rozenski和McCloskey^[25]于2002年发布了一款依赖于二级质谱数据进行RNA与DNA全测序的数据分析软件——SOS (Simple Oligonucleotide Sequencer)。2006年，Guymon等^[26]更是通过实验证明SOS可以用来分析鉴定*T. thermophiles*来源16S rRNA上的化学后修饰位点。类似的分析软件还有2002年Oberacher等^[27]发布的COMPASS软件，实验证实其可以成功地鉴定相对较大的脱氧核糖核苷酸序列(如长度80 bp左右的核苷酸、50 bp左右核苷酸的突变位点等)。

SOS、COMPASS主要是设计用来分析DNA数据，但是，在RNA数据分析中也得到了很好的应用。进一步的分析发现，RNA与DNA的碎裂方式在低能量的CID-MS分析模式下是截然不同的(如图5所示)：DNA主要是丢掉一个碱基，形成了w和(a-B)离子；而RNA则主要是在核心骨架的磷酸基团上发生P-O的断裂形成c和y离子。

2.4.4 Database Searching软件

在纯度较高或者体系比较简单的情况下，以上的软件得到了很好的应用。但是，对于真正意义上复杂的生物学样品来说，它们都有一定的局限性。相对于蛋白质组学中广泛应用的搜索引擎——Mascot和Sequest来说，RNA的质谱研究还存在着一定的欠缺，主要原因在于以下两方面。(1)只有2%的人类基因组最终翻译成了蛋白质，而RNA则是相对复杂的多，在哺乳动物中有超过40%的基因

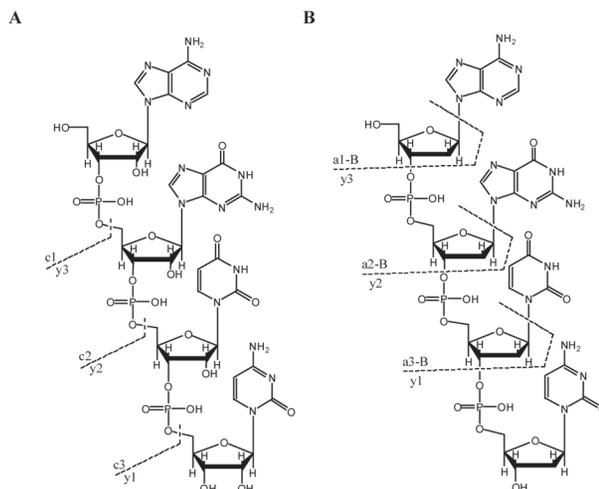


图5 RNA(A)与DNA(B)不同的碎裂方式

被转录成了RNA；而且，蛋白质的编码基因(开放阅读框)均具有很好的规律性，而从基因序列的信息很难得知非编码RNA的区域及其后续转录后的过程(如后修饰等等)。(2)相对于蛋白质组成的20个常规氨基酸残基来说，RNA的组成只有4种且不同RNA所得到的酶解片段特异性太差，所以在复杂的体系中，当没有足够多或者足够长的核苷酸片段信息的情况下，很难通过搜索引擎对相应的RNA进行鉴定。

2009年，Nakayama等^[28]首次发展了一款以RNA二级质谱数据库为基础的搜索引擎——Ariadne。与蛋白质组学中应用的Mascot和Sequest类似，第一步，它通过从质谱数据提取的寡聚核糖核苷酸的二级谱图与数据库中理论谱图比对得到片段的序列信息，并对可能的序列进行排队打分；紧接着，所有可能的寡聚核糖核苷酸再与数据库中的所用RNA进行比对排序打分，从而完成对RNA的鉴定。在随后的实验中，Nakayama等^[28]也成功地利用此软件结合相应的质谱数据从商业化的酵母tRNA提取物中鉴定到了一个“未知的”RNA；同时，在tRNA的样品中还鉴定到了甲基化的核苷酸与二羟基尿苷。

3 结语与展望

基于质谱的代谢组学和蛋白质组学技术已经日趋成熟，结合其他技术手段在生物学研究中起到了越来越重要的作用。由于RNA自身性质的缘故，质谱技术在RNA研究中的应用一直困难重重，且

并没有受到很大的关注。随着近些年来质谱仪器性能上的突破、质谱技术的发展以及在相关数据分析软件方面的开发,质谱方法高效地分析复杂的生物学 RNA 样品的能力得到了很大的增强。另一方面,由于传统 RNA 鉴定和测序方法的弊端,质谱方法在 RNA 研究(如 RNA 后修饰研究、RNA 与蛋白质相互作用等)中的应用将会是一个非常有效的突破口。因此, RNA 分析相关的质谱技术在硬件和软件两个方面进一步的发展将会给 RNA 研究带来新的契机。

[参 考 文 献]

- [1] Sharp PA. The centrality of RNA. *Cell*, 2009, 136(4): 577-80
- [2] Wang X, Arai S, Song X, et al. Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. *Nature*, 2008, 454(7200): 126-30
- [3] Wahl MC, Will CL, Luhrmann R. The spliceosome: design principles of a dynamic RNP machine. *Cell*, 2009, 136(4): 701-18
- [4] Takahashi N, Isobe T. Proteomic biology using LC/MS. Hoboken: John Wiley & Sons, 2007: 1-254
- [5] Takahashi N, Yanagida M, Fujiyama S, et al. Proteomic snapshot analyses of preribosomal ribonucleoprotein complexes formed at various stages of ribosome biogenesis in yeast and mammalian cells. *Mass Spectrom Rev*, 2003, 22(5): 287-317
- [6] Wu L, Belasco JG. Let me count the ways: mechanisms of gene regulation by miRNAs and siRNAs. *Mol Cell*, 2008, 29(1): 1-7
- [7] Valadkhan S. The spliceosome: caught in a web of shifting interactions. *Curr Opin Struct Biol*, 2007, 17(3): 310-5
- [8] Donis-Keller H, Maxam AM, Gilbert W. Mapping adenines, guanines, and pyrimidines in RNA. *Nucleic Acids Res*, 1977, 4(8): 2527-38
- [9] Peattie DA. Direct chemical method for sequencing RNA. *Proc Natl Acad Sci USA*, 1979, 76(4): 1760-4
- [10] Lockhart DJ, Dong H, Byrne MC, et al. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotechnol*, 1996, 14(13): 1675-80
- [11] Fenn JB, Mann M, Meng CK, et al. Electrospray ionization for mass spectrometry of large biomolecules. *Science*, 1989, 246(4926): 64-71
- [12] Pielek U, Zurcher W, Schar M, et al. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry: a powerful tool for the mass and sequence analysis of natural and modified oligonucleotides. *Nucleic Acids Res*, 1993, 21(14): 3191-6
- [13] Covey TR, Bonner RF, Shushan BI, et al. The determination of protein, oligonucleotide and peptide molecular weights by ion-spray mass spectrometry. *Rapid Commun Mass Spectrom*, 1988, 2(11): 249-56
- [14] Muddiman DC, Wunschel DS, Liu C, et al. Characterization of PCR products from bacilli using electrospray ionization FTICR mass spectrometry. *Anal Chem*, 1996, 68(21): 3705-12
- [15] Little DP, McLafferty FW. Sequencing 50-mer DNAs using electrospray tandem mass spectrometry and complementary fragmentation methods. *J Am Chem Soc*, 1995, 117: 6783-4
- [16] Huang TY, Liu J, McLuckey SA. Top-down tandem mass spectrometry of tRNA via ion trap collision-induced dissociation. *J Am Soc Mass Spectrom*, 2010, 21(6): 890-8
- [17] Taucher M, Breuker K. Characterization of modified RNA by top-down mass spectrometry. *Angew Chem Int Ed Engl*, 2012, 51(45): 11289-92
- [18] Listgarten J, Emili A. Statistical and computational methods for comparative proteomic profiling using liquid chromatography-tandem mass spectrometry. *Mol Cell Proteomics*, 2005, 4(4): 419-34
- [19] Zhang Z, Jackson GW, Fox GE, et al. Microbial identification by mass cataloging. *BMC Bioinformatics*, 2006, 7: 117
- [20] Jackson GW, McNichols RJ, Fox GE, et al. Bacterial genotyping by 16S rRNA mass cataloging. *BMC Bioinformatics*, 2006, 7: 321
- [21] Hossain M, Limbach PA. Mass spectrometry-based detection of transfer RNAs by their signature endonuclease digestion products. *RNA*, 2007, 13(2): 295-303
- [22] Taoka M, Yamauchi Y, Nobe Y, et al. An analytical platform for mass spectrometry-based identification and chemical analysis of RNA in ribonucleoprotein complexes. *Nucleic Acids Res*, 2009, 37(21): e140
- [23] Matthiesen R, Kirpekar F. Identification of RNA molecules by specific enzyme digestion and mass spectrometry: software for and implementation of RNA mass mapping. *Nucleic Acids Res*, 2009, 37(6): e48
- [24] McLuckey SA, Van Berkel GJ, Glish GL. Tandem mass spectrometry of small, multiply charged oligonucleotides. *J Am Soc Mass Spectrom*, 1992, 3(1): 60-70
- [25] Rozenski J, McCloskey JA. SOS: a simple interactive program for ab initio oligonucleotide sequencing by mass spectrometry. *J Am Soc Mass Spectrom*, 2002, 13(3): 200-3
- [26] Guymon R, Pomerantz SC, Crain PF, et al. Influence of phylogeny on posttranscriptional modification of rRNA in thermophilic prokaryotes: the complete modification map of 16S rRNA of *Thermus thermophilus*. *Biochemistry*, 2006, 45(15): 4888-99
- [27] Oberacher H, Wellenzohn B, Huber CG. Comparative sequencing of nucleic acids by liquid chromatography-tandem mass spectrometry. *Anal Chem*, 2002, 74(1): 211-8
- [28] Nakayama H, Akiyama M, Taoka M, et al. Ariadne: a database search engine for identification and chemical analysis of RNA using tandem mass spectrometry data. *Nucleic Acids Res*, 2009, 37(6): e47

黄超兰报告讨论

付向东 (加州大学圣地亚哥分校)

Q: 你们这个非常强大的方法, 除了等别人来合作, 你们自己想做什么?

A: 我们的研究主要是利用非变性和离子淌度 MS (native ion mobility (IM) MS) 研究生物大分子复合物, 发展 RNA- 蛋白质相互作用的直接鉴定方法。

Q: 如果做大分子结构, 大分子结构有的地方是动态的, 没办法做出晶体。比如 Primosome 可以做出来, 是因为结构很稳定; 但是 Splicesome 不稳定, 用信息学的方法做不出来, 你们这种质谱方法能不能做出来?

A: Native IM 质谱其中一种应用是观察分子复合物的动态变化。

Q: 对于蛋白质, 你是用离子化方法, 而对于 RNA 的检测你觉得主要障碍是什么?

A: 其中一个技术障碍是要用负的离子化模式 (negative ion mode), 而负的离子化模式灵敏度比较低。

施蕴渝 (中国科学技术大学)

Q: 请问对于大分子复合物, 你可以得到哪些性质?

A: 我们可以得到其拓扑学信息, 例如是何种多聚体, 如何组装, 还有结合位点。

Q: 你需要用到氘交换吗?

A: 是的。氘交换是其中的一种主要手段。

Q: 你们用的是快速混合的方法?

A: 是的。

Q: 你们的 ion mobility 前面安装了这个装置吗?

A: 还没有。

蔡刚 (中国科学技术大学)

Q: 想请教下, 大概能达到什么分辨率? 能看到各个亚基的边界吗?

A: 质谱测的是荷质比的分辨率。现在最大能测到 megadalton(MDa)。

施蕴渝 (中国科学技术大学)

Q: 你们的分析仪和普通的有什么区别?

A: 主要的区别是, 我们的 IM-MS 用了定制的四级杆 (Quadruple), 能扫描离子到 32 kDa, 能更好地检测超大分子。

陈润生 (中国科学院生物物理研究所)

Q: 研究肽段序列的话, 是依赖数据库信息的。目前有肽段质谱图数据库, 也有蛋白质的, 请问 RNA 有吗?

A: 有。我这里有个表格, 白色的代表已经有数据的。关于 RNA 的数据不多只有几个, 而且大多是科研人员自己做出来的。

Q: 从头测序 (*De novo seq*) 有没有可能?

A: 困难比较大。

叶克穷 (北京生命科学研究所)

Q: 我对非变性质谱很感兴趣, 如果在中国做成功是非常好的, 很有应用价值。你现在的仪器可以做到 250 kDa, 是吗?

A: 对。

Q: 有改进的空间吗? 能到 1 MDa 吗?

A: 这个要试。250 kDa 的样品条件比较容易。越大的复合物, 需要调整的参数越多。250 kDa 的图谱还是很漂亮的, 但是 2.3 MDa 分辨率就差了很多。

Q: 你这个是专门做这方面的质谱吗?

A: 是的。

蔡刚 (中国科学技术大学)

Q: 还有一个问题, 你这个蛋白质 / 复合物用量是多少?

A: 比起核磁及电镜, 我们的用量很小, 微摩尔的量级。