

作图信息可以从 UniGene 和人类基因转录本作图数据库 (Human Transcript Map database) 中获得 (<http://www.ncbi.nlm.nih.gov/genemap98>)。另外, 还有其他一些基因作图联合会的网址可以利用 (<http://www.ncbi.nlm.nih.gov/SCIENCE96/Restools.html>)。特别是白头研究所基因组研究中心 (<http://www.genome.wi.mit.edu/>) 和斯坦福大学人类基因组中心 (<http://www.shgc.stanford.edu/>), 他们将 EST 作图信息整合成非常精确的物理和遗传图谱。

随着模式生物基因组测序工作的完成, 可以预见未来几年 EST 在基因识别及阐明基因功能的工作中将发挥越来越大的作用。而且也会有更多的功能更强大的生物信息学工具问世, 使该领域的研究工作更方便迅捷。

参 考 文 献

- 1 陈润生. 生物信息学. 生物物理学报, 1999, **15** (1): 5~12
Chen R S. Acta Biophysica Sinica, 1999, **15** (1): 5~12
- 2 Adams M D, Kelley J M, Gocayne J D, et al. Complementary DNA sequencing: expressed sequence tags and Human Genome Project. Science, 1991, **252** (5013): 1651~1656

- 3 Boguski M S, Schuler G D. EST establishing a human transcript map. Nat Genet, 1995, **10** (4): 369~371
- 4 Huang X. An improved sequence assembly program. Genomics, 1996, **33** (1): 21~31
- 5 Brookes A J. The essence of SNPs. Gene, 1999, **234** (2): 177~186
- 6 Wang D G, Fan J B, Siao C J, et al. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science, 1998, **280** (5366): 1077~1082
- 7 Picoult-Newberg L, Ldeker T E, Pohl M G, et al. Mining SNPs from EST databases. Genome Research, 1999, **9** (2): 167~174

Get the Best Use of dbEST. WANG Hu Chun, CHEN Qing-Xuan (*Institute of Developmental Biology, The Chinese Academy of Sciences, Beijing 100080, China*).

Abstract dbEST is an important division of Genebank. EST clones represent useful molecular tools for new gene identification, gene expression and recombinant protein expression studies. A description of dbEST and the way to get the best of it were given.

key words expressed sequence tag (EST), computer, cloning

质谱在寡核苷酸药物质量控制中的应用*

杨秉呼 王升启¹⁾

(军事医学科学院放射医学研究所, 北京 100850)

摘要 介绍了质谱在寡核苷酸质量控制方面的应用。合成寡核苷酸及类似物作为反义治疗剂在病毒感染和一些癌症治疗方面有良好的前景。寡核苷酸作为药物, 其结构特性必须进行确证。寡核苷酸浓度和纯度的分析可使用色谱或电泳技术, 对寡核苷酸的碱基组成、序列、同一性, 修饰基团, 色谱或电泳分析方法无能为力。质谱的高鉴别能力使其能有效、灵敏、快速和精确地确定寡核苷酸的这些特性。

关键词 寡核苷酸, 质谱, 质量控制

学科分类号 Q503

常用的寡核苷酸基本上全是由固相 DNA 自动合成仪合成的, 合成仪上合成效率的监测不能作为可靠的寡核苷酸切割和脱保护最终质量指标。确定合成寡核苷酸质量的唯一可靠方法是检测合成的每一个核苷酸。常用的寡核苷酸分析方法有聚丙烯酰胺凝胶电泳 (PAGE)^[1]、毛细管凝胶电泳 (CGE)^[1~3]、高效液相色谱 (HPLC)^[3,4]、高效薄层色谱 (HPTLC)^[5,6]。这些分析方法可以确定寡核苷酸的组分和纯度, 对于长度、同一性、序列、

修饰基团等结构特性的检测无能为力。质谱 (MS) 测定质量的高鉴别能力使它能够快速提供寡核苷酸的这些结构信息^[7~20]。

* 国家“863”(102-08-04-01)及军队“九五”重点课题(96Z007)资助项目。

¹⁾通讯联系人。

Tel: (010) 66932211, E-mail: sqwang@nic.bmi.ac.cn

收稿日期: 1999-07-14, 修回日期: 1999-11-01

1 寡核苷酸的含量及杂质检测

寡核苷酸的含量测定一般使用色谱法或电泳方法。PAGE 法分辨率高，可对多个样品实现高分辨率分离，但电泳后显现需使用放射性标记或染色技术，全长及失败序列的定量仅是半定量的且使用凝胶扫描装置。高效液相色谱提供了很好的分离时间重现性和定量准确性，但对一些寡核苷酸不能给出单碱基分离。CGE 能提供单碱基分离，但最大迁移时间和重现性仅在每次分析使用新鲜电解质时获得，变性缓冲液的使用使其耐久性降低。HPTLC 法重现性差和分析精度低。这些分析方法不能确定寡核苷酸的结构信息。质谱测定质量的高鉴别能力使其既能确定寡核苷酸的结构又能确定其中的极少量杂质。质谱应用于核酸研究，长期为它们的高极性和多阴离子属性所限制。最初使用快速原子轰击(FAB)，仅限于小的寡聚体(< 5 ku)。电喷雾电离(ESI) 和基质辅助激光解吸附电离(MALDI)能从高极性分子产生完整的气体离子，使分析高分子(> 10 ku) 成为可能。ESI 和 MALDI 测量寡核苷酸质量的主要困难是阳离子加合物的存在，它们的存在引起多重峰，导致谱解释复杂化。用铵离子代替钠离子可获得直到 25 ku 高质量质谱，并且可达化学合成寡核苷酸的上限——132 个碱基的寡核苷酸。ESI-MS 在寡核苷酸和各种修饰寡聚体及较小污染物检测和结构特征测定方面，显示了高度的准确性，但 ESI-MS 不适合用最小的劳动获得高通量^[7]，由于这个原因电喷雾质谱没有成为针对典型寡核苷酸合成设备的日常分析工具。MALDI 飞行时间(TOF) 质谱的使用，可对寡核苷酸样品进行快速质量分析。各种滞后离子萃取技术的 MALDI-TOF-MS 和自动样品制备及数据收集技术的结合使用提供了合成寡核苷酸高通量分析方法。滞后离子萃取 MALDI-TOF-MS 可在样品质量理论值 0.13% ± 0.10% 范围内确定每一寡核苷酸质量^[8]。质量的精确性允许以前所未有的可靠程度确定合成寡核苷酸的同一性和监测存在于合成样品中的低水平杂质。使用这种自动化组合，从 12 到 70 碱基的 100 个样品的制备、点样和分析可在少于 90 min 内完成。

1.1 MALDI-TOF-MS 所用样品和基质的制备

在浓氨水中被切割和脱保护的合成寡核苷酸需过柱纯化，使形成加合物的阳离子最小化。纯化后的样品在 UV 260 nm 定量。所用基质制备一般使

用 3-羟基吡啶羧酸或吡啶羧酸以一定比例溶于乙腈水溶液中，为了降低阳离子加合物形成，需在基质中加入阳离子交换树脂或无机铵盐。为了加快样品制备可使用后者，尽管这不能完全消除阳离子加合物形成，但对产生所需分辨率是足够的。基质溶液贮存不要超过一周。

1.2 仪器的标准化和优化

质谱仪最好以阳离子模式收集数据，和阴离子模式比，两种模式所获结果之间没有明显差异，但阳离子模式还有其他优点。激光功率刚好高于使样品离子化所需功率，功率太高将导致明显的碱基丢失。在相当的寡核苷酸长度范围内，照射样品所需功率变化非常小。对 17 至 30 寡聚体，单一功率给出一致的好结果，更长或更短的寡聚体分别增加或降低功率。使用可重现激光功率有助于自动化控制下的数据收集。记录谱可使用手动或自动模式。手动模式收集数据总体给出更强的信号。对长寡核苷酸(> 35 聚体) 手动模式更好，但所需时间延长。对于日常工作，使用外标给出的值可精确到平均值的 0.15%，总体是满意的，因 QC 主要着重确定失败序列。使用内标可获得更精确结果，准确度可到 0.03%。在反射模式中，在满意的情况下，可增加到 0.015%。反射模式倾向于生成 DNA 碎片，只有当精确度需要 1 u 时才用，对 RNA，谱质量通常是好的。

1.3 寡核苷酸的分析范围和杂质的监测

当合成寡聚体长度增加时，相对于长度相差一到两个核苷酸的寡核苷酸杂质，全长 DNA 组分所占百分率降低，并且加合物形成更严重，各种杂质的存在既降低了分辨率又降低了信号质量。当长度超过 35 聚体，样品需 HPLC 纯化^[9]。长度在 60 碱基以内时，分辨率对实现单一核苷酸分离是足够充分的。当长度超过 60 碱基，分辨率降低。实验室合成的寡核苷酸大部分少于 30 碱基，绝大多数少于 50 碱基，因此该质量控制方法有能力处理几乎所有实验室产品，对于较少的更长的寡核苷酸样品，由手动质谱分析或仍由毛细管电泳分析。

MALDI-TOF-MS 能够证实修饰体或衍生化的结合。实验室中，一般所用的修饰有末端磷酸化、末端氨基修饰、生物素和荧光素标记。由于最小的质量修饰(磷酸化质量差为 80 u) 约占平均长度寡核苷酸质量的 1%，对于含小修饰的大寡聚体，适当的更严格控制质量差接收范围可确定磷酸基的存在。即使对绝大多数 80 聚体，磷酸基仍占理论质

量的 0.3%，还是可监测的。

由于一个核苷酸的质量在任一寡聚体全部质量中所占比重超过了 1%，而 MALDI-TOF-MS 法可在样品理论值 0.13% ± 0.10% 范围内确定样品质量，因此该法可检测一个核苷酸的丢失和确定其位置。质谱在确定失败序列方面同 CGE 同样灵敏。但对失败序列的定量测定仅是半定量的，对于通常的长度在 15~30 碱基的寡核苷酸能很好地确定 $n-1$ 失败序列的含量。

2 寡核苷酸的序列分析

寡核苷酸的经典测序方法为 Maxam-Gilbert 化学降解法。该法操作繁琐、耗时、而且需使用放射性同位素。对化学修饰的寡核苷酸不能反映修饰基团的信息。ESI 和 MALDI 质谱技术的出现提供了强有力的寡核苷酸及类似物结构和序列分析方法。

2.1 质谱测序的原理

质谱对寡核苷酸测序目前使用最多的是磷酸二酯酶 (PDEs) 梯带测序法。该法测序的原理是将被测寡核苷酸样品先用外切酶从 3' 或 5' 端进行部分降解，在不同时间内分别取样进行质谱分析，取得寡核苷酸部分降解的分子离子峰信号，通过对相邻两个碎片分子质量进行比较，可以计算出被切割的核苷酸单体分子质量，将其与四个脱氧核苷酸的标准分子质量 (PdA= 313.2, PdC= 289.2, PdG= 329.2, PdT= 304.2) 进行对照，就可以按顺序读出寡核苷酸的完整序列。使用该法，Pieles 等^[10]最先获得了 13 聚体寡核苷酸的序列。使用滞后萃取技术，Smirnov 等^[11]测定了 33 聚体的序列。Wu 等^[12]使用 ESI 质谱技术测定了 22 聚体的寡核苷酸的序列。Faulstich 等^[13]使用 MALDI-TOF-MS 技术第一次测定了 RNA 的序列。

除了梯带测序法，质谱测序的研究还着重于以下两个方面：a. 质谱分析 Sanger 测序的反应产物。利用质谱的高通量、灵敏和操作简便代替凝胶电泳。Roskey 等^[14]使用滞后萃取 MALDI 飞行时间质谱技术测定了引物后的 37 个碱基的序列。Mouradian 等^[15]分析了产生于载体 M13-mp19 的 DNA 片段的序列。b. 基于气相裂解的测序分析。如源后衰变或碰撞诱导解离等。Zhu 等^[16]使用紫外 MALDI 和两组分基质测定了 35 聚体的序列。

2.2 寡核苷酸的降解

质谱对寡核苷酸测序时，被测的寡核苷酸样品先用 PDEs 从 3' 或 5' 端进行部分降解，在不同时间

点分别取样进行质谱分析，能得到寡核苷酸部分降解的分子离子峰信号，从而获得寡核苷酸的序列。无论从 5' → 3' 或从 3' → 5' 降解，最后两个碱基的序列不能确定。要获取寡核苷酸全长序列，必须分别使用 3' 和 5' 外切酶。使用两种酶测序也有其他优点。当 3' 外切酶降解处理到 5' 端时，降解逐渐变慢，噪音增强，要获得 5' 端序列信息需要更多的时间。对于 5' 外切酶降解也是如此。使用两种酶可克服使用一种酶的缺陷。3' 外切酶一般使用蛇毒磷酸二酯酶 (SVP)，5' 外切酶使用牛脾磷酸二酯酶 (CSP)。SVP 降解与序列无关，产生均匀分布的峰。而 CSP 降解与序列有关，一些区域的峰信号弱，难以识别^[11]。因此总是通过 SVP 降解获取绝大多数序列信息，由 CSP 降解获取较少的 5' 端序列信息。SVP 的活性较高，初期切割速度较快，在开始阶段取样间隔要短，否则得不到完整的碎片分子。而 CSP 活性相对较低，切割速度较慢，取样间隔要长一些^[17]。

硫代寡核苷酸由于分子骨架上每个磷酸基上的一个氧为硫取代，增加了对核酸酶降解的阻力，这类化合物适合于反义治疗，但不能直接通过外切酶降解由质谱测序。质谱测序时，该类化合物必须通过氧化使之转化成磷酸二酯寡核苷酸。Schuette 等^[18]用 THF-水-甲基咪唑和碘氧化硫代寡核苷酸，使其转化成磷酸二酯寡核苷酸，再使用 3' 和 5' PDEs 降解后，由质谱测序。

寡核苷酸 5' 或 3' 端标记时，只能使用两种外切酶中的一种降解寡核苷酸。降解在距标记端只有一个碱基时停止。连接标记基团的最后一个碱基应由碰撞诱导解离串联质谱 (CID-MS/MS) 测定^[12]。

2.3 质谱测序的电离方式

质谱测序目前使用 ESI 和 MALDI 两种电离方式。应用最多的是 MALDI。使用该技术的原因是 MALDI 主要产生单电荷离子，形成的质谱图简单。而 ESI 产生多电荷离子，形成的质谱图复杂^[11, 12, 19]。对于较大的寡核苷酸，特别是长度大于 20 碱基时，常规 MALDI 因分辨率低、灵敏度差限制了其应用。为了使 MALDI 电离方式用于较大寡核苷酸的序列分析，可使用以下方法：a. 滞后萃取。可极大提高分辨能力和灵敏度。对于 30 聚体，使用滞后萃取峰宽只有 10~15 u，不用滞后萃取时，峰宽为 55~75 u^[11]。b. 使用 MALDI 兼容缓冲或基质自身作为酶降解的缓冲溶液。使酶降

解条件更严格和合理。c. 用3-羟基吡啶羧酸代替2, 4, 6-三羟基苯乙酮或使用混合基质^[11, 16, 17]。

ESI在测序中应用，存在着如下的问题^[12]：a. 由于形成多电荷离子，电喷雾质谱必须转化成重组分子质量(RMW)谱，以获得可读的序列梯带。然而由于RMW谱质量差，超过700u的分子质量重组未见报道。b. 酶和样品未彻底纯化时，导致金属离子加合物形成，在质谱中产生额外的峰。c. 降解过程中，寡核苷酸的浓度逐渐降低，确定分子质量需更高的灵敏度。这些问题妨碍了ESI在寡核苷酸测序中的应用。通过纯化所有试剂，选择合适的溶剂，控制降解条件及分子质量谱的重组，使用电喷雾或离子喷雾质谱也可获得大于20碱基的寡核苷酸序列。

3 问题及展望

质谱测序能很好地适合于长度小于30个碱基的寡核苷酸，对于长度大于30个碱基的寡核苷酸，由于费用和方法的局限限制了该法的使用。质谱对DNA大规模测序是不可能的^[20]。质谱分析自动化使大量合成寡核苷酸的快速分析成为现实，用普通的核酸分析方法不能解决的化学修饰及标记存在，质谱给出了很好的结果。质谱仪测定质量的精确性及能够识别来源于合成仪或纯化不充分的低浓度杂质，使它能够被用作识别和纠正仪器操作。这些优势使质谱能够对合成寡核苷酸进行质量监测和质量控制，但质谱的定量仅是半定量的。

质谱在寡核苷酸或核苷酸中的应用前景是将色谱或电泳技术的强分离能力同质谱的高鉴别能力结合在一起，其中最有价值的是扩展液相色谱/质谱联用技术。对多组分寡核苷酸无需分离，一次运行就可精确地测定每一组分的质量，从而确定每一寡核苷酸的特性。

参考文献

- DeDionisio L A, Raible A M, Nelson J S. Analysis of an oligonucleotides N^{3'} → P^{5'} phosphoramidate/phosphorothioate chimera with capillary gel electrophoresis. *Electrophoresis*, 1998, **19** (16 ~ 17): 2935~ 2938
- Lowery J D, Ugozzoli L, Wallace R B. Application of capillary electrophoresis to the measurement of oligonucleotide concentration and purity over a wide dynamic range. *Anal Biochem*, 1997, **254** (2): 236~ 239
- Srivatsa G S, Klopchin P, Batt M, et al. Selectivity of anion exchange chromatography and capillary gel electrophoresis for the analysis of phosphorothioate oligonucleotides. *J Pharm Biomed Anal*, 1997, **16** (4): 619~ 630
- Gaus H J, Owens S R, Winnman M O, et al. On-line HPLC electrospray mass spectrometry of phosphorothioate oligonucleotide metabolites. *Anal Chem*, 1997, **69** (3): 313~ 319
- 王升启, 马立人, 张京生, 等. 薄层扫描法测定寡核苷酸含量. 生物化学与生物物理进展, 1993, **20** (2): 134~ 136
Wang S Q, Ma L R, Zhang J S, et al. Prog Biochem Biophys, 1993, **20** (2): 134~ 136
- 王升启, 马立人. 高效薄层色谱法评价合成寡核苷酸的质量. 色谱, 1993, **11** (3): 146~ 148
Wang S Q, Ma L R. Chromatography, 1993, **11** (3): 146~ 148
- Ball R W, Packman L C. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry as a rapid quality control method in oligonucleotide synthesis. *Anal Biochem*, 1997, **246** (2): 185~ 194
- Van Ausdall D A, Marshall W S. Automated high-throughput mass spectrometric analysis of synthetic oligonucleotides. *Anal Biochem*, 1998, **256** (2): 220~ 228
- Zhu Y, He L, Srinivasan J R, et al. Improved resolution in the detection of oligonucleotides up to 60-mers in matrix-assisted laser desorption/ionization time-of-flight mass spectrometry using pulsed-delayed extraction with a simple high voltage transistor switch. *Rapid Commun Mass Spectrom*, 1997, **11** (9): 987~ 992
- Pielies U, Zurcher W, Schär 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 Acid Res*, 1993, **21** (14): 3191~ 3196
- Smirnov I P, Roskey M T, Juhasz P, et al. Sequencing oligonucleotides by exonuclease digestion and delayed extraction matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Anal Biochem*, 1996, **238** (1): 19~ 25
- Wu H, Chan C, Aboleneen H. Sequencing regular and labeled oligonucleotides using enzymatic digestion and ionspray mass spectrometry. *Anal Biochem*, 1998, **263** (2): 129~ 138
- Faulstich K, Worner K, Brill H, et al. A sequencing method for RNA oligonucleotides based on mass spectrometry. *Anal Chem*, 1997, **69** (21): 4349~ 4353
- Roskey M T, Juhasz P, Smirnov I P, et al. DNA sequencing by delayed extraction matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Proc Natl Acad Sci USA*, 1996, **93** (10): 4724~ 4729
- Mouradian S, Rank D R, Smith M L. Analyzing sequencing reactions from bacteriophage M13 by matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun Mass Spectrom*, 1996, **10** (12): 1475~ 1478
- Zhu Y F, Tarangenko N I, Allman S L, et al. Oligonucleotide sequencing by fragmentation in matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom*, 1997, **11** (8): 897~ 903
- 阎庆金, 王升启, 杨松成, 等. MALDI-TOF-MS对寡核苷酸的序列分析. 生物化学与生物物理学报, 1997, **29** (5): 475~ 480
Yan Q J, Wang S Q, Yang S C, et al. Acta Biochem Biophys Sinica, 1997, **29** (5): 475~ 480
- Schuette J M, Pielies U, Maleknia S D, et al. Sequence analysis of phosphorothioate oligonucleotides via matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *J Pharm Biomed Anal*, 1995, **13** (10): 1195~ 1203
- Scalf M, Westphall M S, Krause J, et al. Controlling charge states of large ions. *Science*, 1999, **283** (5399): 194~ 197
- Crain P F, McCloskey J A. Applications of mass spectrometry to the characterization of oligonucleotides and nucleic acids. *Curr Opin Biotechnol*, 1998, **9** (1): 25~ 34

(下转第451页, Continued on page 451)

1966

- 6 Tillyer C R. Error estimation in the quantification of alkaline phosphatase isoenzymes by selective inhibition methods. *Clin Chem*, 1988, **34** (12): 2490~2493
- 7 Bouman A A, de Ridder C M, Nijhof J H, et al. Immuno-adsorption assay for bone alkaline phosphatase compared with wheat germ agglutinin precipitation assay in serum from (pre) pubertal girls. *Clin Chem*, 1996, **42** (12): 1970~1974
- 8 Rosalki S B, Foo A Y, Burlina A, et al. Multicenter evaluation of iso-ALP test kit for measurement of bone alkaline phosphatase activity in serum and plasma. *Clin Chem*, 1993, **39** (4): 648~652
- 9 Rosalki S B, Foo A Y. Two new methods for separating and quantifying bone and liver alkaline phosphatase isoenzymes in plasma. *Clin Chem*, 1984, **30** (7): 1182~1186

Determination of Alkaline Phosphatase Activity by the Agglutinin Precipitation Assay and Its Clinical Application. PENG Zhi Ying (*First Hospital Laboratory Medicine of West China University of Medical Science, Chengdu 610041, China*); DONG

Yun-Hua (*Analyze and Determine Center of West China University of Medical Science, Chengdu 610041, China*); WEI Xiao-Ming (*Department of Laboratory Medicine, West China University of Medical Science, Chengdu 610041, China*); ZHENG Yong (*College of Pharmacology, West China University of Medical Science, Chengdu 610041, China*).

Abstract The agglutinin precipitation assay was used to determine the bone alkaline phosphatase. The advantage of this assay is simple manipulation, and good reproducibility. Coefficient variation within and between runs were 6.12%, 8.5% and 6.4%, 9.5% respectively. It is concluded that serum B-ALP levels is an useful parameter for clinical observation of the bone metabolism.

Key words bone alkaline phosphatase, agglutinin, bone fracture

(上接第 447 页, Continued from page 447)

Application of Mass Spectrometry in Quality Control of Oligonucleotide Drugs. YANG Bing-Hu, WANG Sheng-Qi (*Institute of Radiation Medicine, Academy of Military Medical Sciences, Beijing 100850, China*).

Abstract Synthetic oligonucleotides and their analogs have shown promising potential therapeutic agents in the treatment of viral infections and certain cancers. As drugs, the full characterization of oligonucleotides requires complete verification. Chromatographic or electrophoretic techniques can be used in analysis of concentration and purity of

synthetic oligonucleotides, but the analytic methods have been limited in analysis of base composition, sequence, identity and modification of oligonucleotides. Ability of mass spectrometry to discriminate mass make MS effective, sensitive, rapid and accurate to determine these characterizations of oligonucleotides. application of mass spectrometry in quality control of synthetic oligonucleotides was introduced.

Key words oligonucleotides, mass spectrometry, quality control