Applications of mass spectrometry to the characterization of oligonucleotides and nucleic acids

Pamela F Crain* and James A McCloskey†

Mass spectrometry-based techniques continue to undergo active development for applications to nucleic acids, fueled by methods based on electrospray and matrix-assisted laser desorption ionization. In the past two years, notable advances have occurred in multiple interrelated areas, including sequencing techniques for oligonucleotides, approaches to mixture analysis, microscale sample handling and targeted DNA assays, and improvements in instrumentation for greater sensitivity and mass resolution.

Addresses
University of Utah, Department of Medicinal Chemistry, 30 South 2000 East, Salt Lake City, UT 84112-5820, USA
*e-mail: pamela.crain@rna.pharm.utah.edu
†e-mail: james.mccloskey@n.cc.utah.edu

Current Opinion in Biotechnology 1998, **9**:25–34

http://biomednet.com/elecref/0958166900900025

Abbreviations

ESI	electrospray ionization
MALDI	matrix-assisted laser desorption ionization
MS	mass spectrometry
TOF	time-of-flight

Introduction

Mass spectrometry (MS) deals with the production and gas phase measurement of molecular mass or of the masses of gas phase dissociation products, either of which are then related to structure of the intact molecule. The revolution in applications of MS to biopolymers, such as nucleic acids and proteins, is now in its second decade, having been fueled by the parallel discoveries of electrospray ionization (ESI) [1] and matrix-assisted laser desorption ionization (MALDI) [2]. These techniques, now widely available in a range of commercial instruments, have dramatically moved the boundaries of routine applicability of MS, in terms of the molecular size and polarity of the molecules that can be studied, to include oligonucleotides, small nucleic acids, and proteins [3].

As the most polar of the biopolymers, it may be argued that nucleic acids have been the principal beneficiaries of these new ionization methods [4], which permit efficient transfer of solution-phase polynucleotide ions to the gas phase, and allow accurate mass measurements at the picomole level and below with errors typically in the range 0.01–0.1% (e.g. ±1–10 Da at 10 kDa). One consequence of both the availability and relative ease with which such data can be acquired has been a decided shift in the location of ESI- and MALDI-based instruments to biological laboratories rather than being exclusively in specialized facilities. This has been accompanied by wider applicability (sensitivity and mass range) to biological problems and has led to increased development of novel integrated protocols (such as enzymology plus MS), directed toward the solution of specific types of problems, as has been notable in the protein field [5].

This review focuses primarily on the recent literature, 1996–1997, and on developments associated specifically with the use of ESI and MALDI for studies of nucleic acids. Summaries of the earlier literature are available through reviews covering multiple topics [4,6••], one book devoted to ESI [7], as well as reviews concentrating on sequencing [8,9••], MALDI [10] and ESI [11] of nucleic acids, and very readable summaries of the ESI and MALDI techniques [12,13]. It is unfortunate that the recent literature in this field tends to be strongly segregated by ionization method, often without recognition that either method may suffice for application to a given problem. Also, as a consequence of a rapidly growing field, some papers represent only modest incremental steps forward, some of which are largely redundant, and others quickly superseded. We have, therefore, attempted to emphasize the most recent reports rather than those in which initial studies on a given topic were reported. The categories selected are intended to give the reader the flavor of (arguably) the most interesting or important aspects of recent developments.

Sequence determination

Because the structural elements of sequence in DNA and RNA are represented by differences in mass, mass spectrometry is intrinsically attractive as a sequencing method. Instrumentation cost and limited practical applicability to chain lengths beyond ~30–50 are detractions compared with gel-based or other conventional methods; however, a major advantage lies in applications involving modified subunits where conventional methods may be uninformative. Such circumstances may arise if modification is unexpected or structurally unknown, or in the case of mobility-based methods (chromatographic or electrophoretic) if the influence of the modification on mobility is not known. Also, while prospects for large scale genomic sequencing by MS have faded [14•], the potential for high-throughput targeted sequencing in conjunction with genotyping or similar assays has aroused recent interest [15•] (also see the following section).

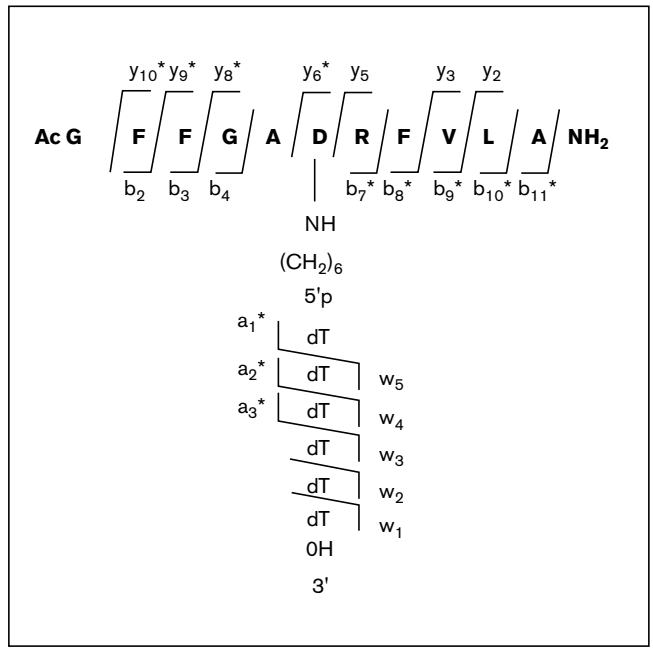
From the standpoint of methodology, MS-based approaches [9••] can be broadly divided into two categories as sequence information can be derived either directly

from gas phase fragmentation, or by mass measurement of chain cleavage products formed in the condensed phase (by chemical or enzymatic digestion). In either case, the advent of ESI and MALDI, as well as of expanded and refined designs of mass analyzers (e.g. [16•]), has led to a number of proposed sequencing strategies [8] that were not possible ten years ago. However, two important distinctions must be kept in mind: first, many reported protocols represent novel ideas but have not passed the test of routine utilization in a problem-solving setting; second, there is inherent ambiguity in the phrase 'verification of sequence', which often simply means that a model of known sequence was studied as a putative test of the method, without knowledge of how well the method would work if an incorrect or truly unknown sequence were encountered.

Direct sequencing by gas phase fragmentation (for example see [17,18•]) was shown to be simple from an experimental standpoint but is limited to chain lengths below ~30 nucleotides as a consequence of the amount and distribution of energy available to the dissociating ion. In the case of MALDI, the identity of the matrix (used both for absorption of laser energy and its transfer to the analyte) plays a crucial role and can be chosen to influence the extent of fragmentation [19,20]. The relative complexity of mass spectra produced by dissociation of ESI-generated polyanionic nucleotides can be used to advantage for *de novo* sequencing [18•], and can be extended in certain cases to unresolved mixtures of oligonucleotides [21]. Presence of a charged backbone was shown not to be a prerequisite to fragmentation reactions in the mass spectrometer, although proposed cleavage mechanisms require a charge; thus, direct sequence information can be obtained from methylphosphonate oligonucleotides in which the backbone is not charged [22]. Contrary to conventional logic (charge state in solution dictates preferred charge state in the gas phase), the formation and dissociation of polyprotonated (rather than deprotonated) oligonucleotides produces informative mass spectra, reflecting the identities and sequence locations of the heterocyclic bases, which differ in basicity and, hence, extent of protonation during electrospray [23]. Positive ions may also be effective for simultaneous sequencing of peptides and oligonucleotides in covalently linked conjugates. The potential of this interesting approach is illustrated by the data in Figure 1 [24•].

On the other hand, indirect sequencing methods [9••], although requiring pre-MS sample treatments of varying complexity, are capable of reaching greater chain lengths. Capitalizing on the advantages of the delayed ion extraction technique (see the instrumental techniques section) MALDI sequencing of Sanger dideoxy [25] and cycle sequencing reactions [26] has been applied to analysis of 40 and 50-base templates. Hahner *et al.* [27] have recently proposed the use of various RNases (T1, U2, PhyM, A, CL3, cusativin) and partial alkaline hydrolysis as a means of sequencing RNA using MALDI. The principle is straightforward, but the results are confounded by lack of enzyme cleavage specificity in some cases and mass errors averaging ± 0.4–0.9 Da (of consequence in distinguishing U versus C, 1 Da). Applications to determination of modified nucleotides is suggested but the data presented do not encourage applications to non-model (unknown) RNA sequences at this time.

Figure 1



Positive ion fragmentation patterns observed in tandem mass analysis of peptide–linker–dT_6 (amidated at the carboxyl-terminus), which permits sequencing of both the peptide and oligonucleotide moieties. Nucleotide fragment ions are designated as a_n or w_n and peptide fragment ions are designated as b_n or y_n. The $a_n{}^*$ ions are oligonucleotide-fragment ions with the peptide attached through the C_6-linker. The w_n oligonucleotide-fragment ions each contain the oligonucleotide 3′-terminus. The three b_n peptide-fragment ions and the four $y_n{}^*$ ions, which are y_n peptide-fragment ions with the linker–dT_6 attached, confirm the amino-terminal sequence Ac–Gly–Phe–Phe–Gly–Ala–, while the three y_n fragment ions and the five $b_n{}^*$ ions, which are b_n peptide-fragment ions with the linker–dT_6 attached, confirm the carboxy-terminal sequence –Arg–Phe–Val–Leu–Ala–NH_2. The attachment of the aminolinker–dT_6 to the Asp residue is readily deduced from the difference in mass between $y_6{}^*$ and y_5. Reproduced with permission from [24•].

Genotyping and other targeted DNA assays

Genotyping and similar assays are broadly considered as measurement of the molecular masses of DNA oligonucleotides whose sequence(s) are defined by the enzyme substrates and primers used for their synthesis. Both ESI [28–30] and MALDI [31–33,34•,35•,36,37] have been used for these measurements. The merits of one ionization method over the other are those inherent to the type of mass spectrometric measurement performed [6••], with greater mass accuracy accruing to ESI-based methods, and greater simplicity and potentially higher

throughput with MALDI-based methods. In particular, the technologies that will enhance application of MALDI include chip-based strategies (for example, see [38,39]) and miniaturized sample application devices [40•], which can lead to significant improvement in detection limits (as shown in Figure 2). It should be noted from recent work [41], however, that chip-based approaches can be adapted to ESI, although they are at a much earlier stage of development.

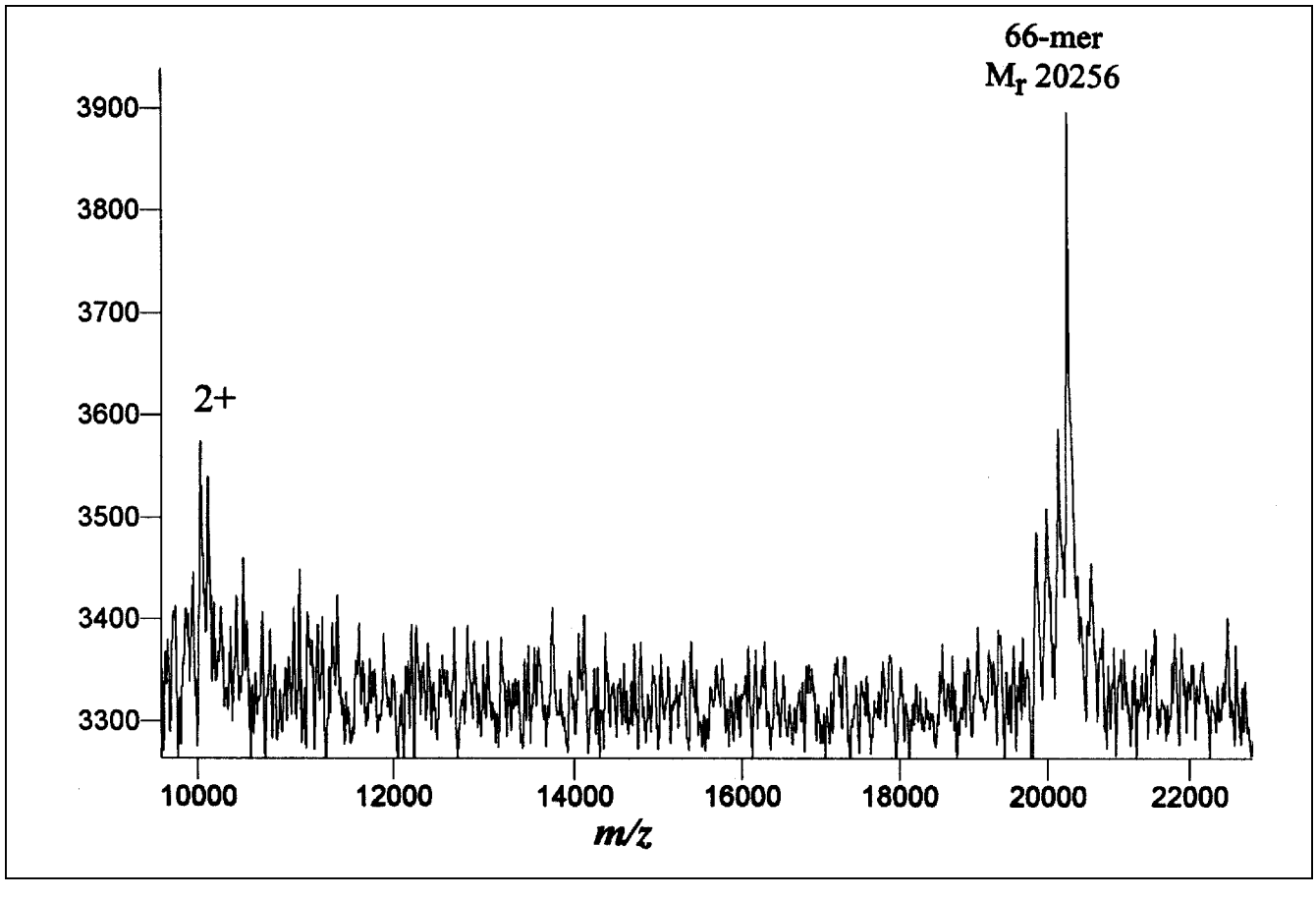
Owing to the poorer mass resolution and decreased ion abundance observed for large DNA oligonucleotides with MALDI time-of-light (TOF) [6••], generation of relatively short (versus full length) PCR products can improve detection of the products. Thus, MALDI can be used for analysis of the products of ligase chain reactions [32], and of PROBE [35•,42] and PinPoint [34•] assays, via measurement of molecular masses of products. Multiplexing is possible if individual PCR probes are tagged with oligo-dT segments of unique masses [43].

At present, these developments to a large extent consist of highly innovative [44••] (see this issue, Graber *et al.* pp 14–18) building blocks designed to ultimately produce practical devices and protocols that can compete with existing methods of targeted DNA assay. Individually they may well find application in numerous less routine research-oriented applications. In any event, time and further development in this fascinating area will need to pass before judgement, measured by acceptance and usage, can be rendered.

Measurements of non-covalent associations

Increasing attention has been given in the past several years to the potential use of ESI mass spectrometry to study the stoichiometry, and in some cases kinetics (e.g. [45]), of non-covalent associations [11,46•]. Due to the gentle nature of the ESI process, it was determined in earlier studies (see [4]) that some associations formed in solution, such as complementary oligonucleotide duplexes, can be transferred with fidelity to the gas phase. Mass measurement of the complexes formed, which are highly accurate, then provides compelling evidence for the identities and stoichiometries of the constituent subunits. The key question is whether such results do, in every case, reflect true solution structures. The number of apparently

Figure 2



MALDI-TOF mass spectrum from 12 femtomoles of a mixed base synthetic DNA 66-mer, M_r 20,243 (theoretical), dispensed in a 6 nL volume with a piezoelectric pipet into the well of a silicon chip containing 6 nL of matrix solvent. Reproduced with permission from [40•].

successful measurements of this type is quite impressive [46•], but cautionary reports [47] are reminders that such experiments are highly dependent on experimental conditions [11]; negative results are usually not reported. Thus, it is essential that in previously unstudied systems control experiments be included, such as gel shift assays, spectroscopic data or ion abundance changes as a function of solution concentrations [46•,47]. Furthermore, it should be recognized that experimentally measured mass often infers, but does not directly determine, structure. Thus, in the case of gas phase oligonucleotide duplexes there is no direct evidence, to date, that the geometries closely resemble solution phase structures even though they may be (loosely) described as Watson-Crick base pairs.

In notable recent work, Smith and co-workers [48] demonstrate the use of the ESI method for stoichiometric measurement of protein–oligodeoxynucleotide complexes made with gene V protein (M_r 9,688) from bacteriophage f1. The first such protein–oligonucleotide measurements had been reported several months earlier using bovine serum albumin (M_r 66,497) and phosphorothioate 20-mer [45]. Selective interactions between RNA and protein can also be measured, as shown by a careful study of gas phase complexes of HIV-1 tat peptide and tat proteins with TAR RNA [49]. Competitive binding experiments between tat peptide and TAR RNA and also several mutants in which a structurally important CUC bulge was modified or deleted, demonstrated greater affinity for the bulge-containing RNA, consistent with previous solution phase studies. Analogous applications used to probe the effects of conformational variations on drug–DNA binding also warrant continued exploration. A successful example, which had been well studied in solution using fluorescence measurements, was the binding of post-activated neocarzinostatin to a bulged DNA hairpin [50]. The experiments outlined above entail the measurement of mass and abundances of gas phase complexes. By contrast, two novel approaches of note have been reported in which the products of collision-induced dissociation of gas phase complexes were used as a site-selective probe of structure, one involving metal ion binding [51•], and the other stability consequences of base pair mismatches in duplexes [52•]. Further experiments along these lines are likely in the future as knowledge of the properties and differences between solution-phase and solventless gas-phase non-covalent interactions increases.

Analysis of oligonucleotide mixtures by liquid chromatography/mass spectrometry

The direct combination of chromatography (gas or liquid) with mass spectrometry has historically proven to be a method of immense analytical power. Because ESI is effected from a flowing liquid, HPLC is intrinsically compatible with the electrospray ion source, and potential benefits of a routine LC/MS procedure are considerable, with applications to areas such as oligonucleotide chemistry and synthesis, characterization of metabolic products, and molecular biology. It is somewhat surprising, therefore, that until recently the development of LC/ESI-mass spectrometry in this area was relatively limited (reviewed in [11]); however, this is largely attributed to the incompatible demands of the two methods. Anion exchange and reversed-phase HPLC operate on ion pair principles, utilizing salt-containing mobile phases for efficient separation, which are deleterious to ESI, and reduce sensitivity and mass spectral peak quality due to gas phase cation adduction by the nucleotide backbone. A significant solution to this problem has been the recent development by Apffel *et al.* [53••], of a reversed-phase HPLC system employing 1,1,1,3,3,3-hexafluoro-2-propanol as a mobile phase additive, adjusted to pH 7.0 with triethylamine. The advantages demonstrated for this new system [53••,54•] include high ESI sensitivity, suppressed cation adduction levels, and excellent chromatographic peak dispersion both between chain lengths and for mixed base compositions at the same chain length. Applicability was demonstrated to phosphorothioate and phosphodiester oligodeoxynucleotides, up to the 20-mer level, and with mass measurement errors of around ±0.2% using a quadrupole ion trap mass spectrometer [54•]. The solvent system is equally applicable to RNA (PF Crain, unpublished results), further suggesting utility in a range of oligonucleotide analysis problems.

Sample preparation for ESI and MALDI mass spectrometry

It is well accepted that both ESI and MALDI require relatively salt-free samples for best results, with MALDI being somewhat more tolerant to their presence [6••,11]. Sample preparation and clean-up is critical to success of either method, regardless of which nucleic acid is being analyzed. Sample pre-treatments utilizing streptavidin–biotin affinity methods are proving useful, especially for MALDI-based applications (for example, see [27,44••]). To a certain degree, however, the deleterious influence of residual salt can be negated by sample manipulation at the point of sample introduction into the mass spectrometer.

In an important study, it was shown that suppression of cationization can be effected by addition of volatile organic bases such as piperidine, imidazole or tetraethylammonium, alone or in combination with the solvent used for sample infusion with ESI-MS [55]. Mg^{2+} presents a special problem in that nucleic acids have high affinity for this cation, which may not be removed by simple desalting procedures. 1,2-Diaminocyclohexane-*N*, *N*, *N′*, *N′*-tetraacetic acid has a greater affinity for Mg^{2+} than EDTA does, is compatible with ESI and its use in combination with tetraethylammonium in the infusion solvent permitted M_r measurements of a set of natural tRNAs and 5S rRNA to within 0.008% or better [56] (1,2-diaminocyclohexone-N,N,N′,N′-tetraacetic acid is also effective as an additive to the matrix used for MALDI-MS of tRNA [57]).

Sample introduction through a liquid chromatograph provides desalting *in situ* during LC/MS of oligonucleotides, independent of whether or not full chromatographic fractionation of the components is effected [53••,58•]. A microdialysis probe has been described that produces an impressive degree of desalting; useful oligonucleotide mass spectra could be obtained from solutions in which the NaCl concentration was 2 M [59]. A relatively high flow rate requirement (2 µl/min) limits on-line applicability for analysis of trace amounts of material.

The effects on MALDI analysis of impurities that may be encountered in the products of enzymatic DNA sequencing reactions have been evaluated and classified into two categories, type I or type II [60]. The former category includes salts and buffer constituents that form adducts in a concentration-dependent manner, while the latter adversely affect signal stability and sample-matrix co-crystallization. As a supplement to the usual methods for nucleic acid desalting for MALDI-MS [6••], organic base can be added to the matrix to suppress residual cations [61]. The homogeneity of the sample plus matrix on the sample target also influences ion intensity and spectral reproducibility. Use of isopropanol/water in place of the commonly used acetonitrile/water for preparation of the matrix solution was found to improve the MALDI behavior of oligonucleotides [62]. DNA desorbed from a target coated with nitrocellulose shows lower levels of salt adduction than samples desorbed from an untreated metal target [33].

A novel alternative matrix system has been described that utilizes frozen thin films of substituted phenols [63]. High quality spectra are produced as a consequence of several factors, including homogeneity of the matrix relative to the usual crystal matrices, and lessened extent of fragmentation.

Modification of PCR products has been shown to enhance their detectability by MALDI-TOF. Base- or sugar-modified deoxycytidine derivatives were incorporated into model oligonucleotides [64] and shown to suppress fragmentation, thus extending the accessible mass range as well as sensitivity of detection. Quaternary ammonium derivatives were added to the 5′ end of oligonucleotides, which were then alkylated on the backbone [65]; this combination of procedures improved detectability more than 100-fold.

Instrumental techniques

During the period before about 1990, centre stage in biological MS was occupied by the development of ionization methods, such as fast atom bombardment and thermospray, which held promise for extending the polarity or mass range of compounds amendable to mass spectrometry. At that time interest in the development of new mass analyzers was modest, but changed rapidly with the advent of ESI [1] and MALDI [2], and the necessity (or opportunity) to mass-measure and dissociate multiply-charged and/or large ions. As a consequence, considerable contemporary interest has been devoted to improved instrumental techniques for analysis of ESI- and MALDI-generated ions.

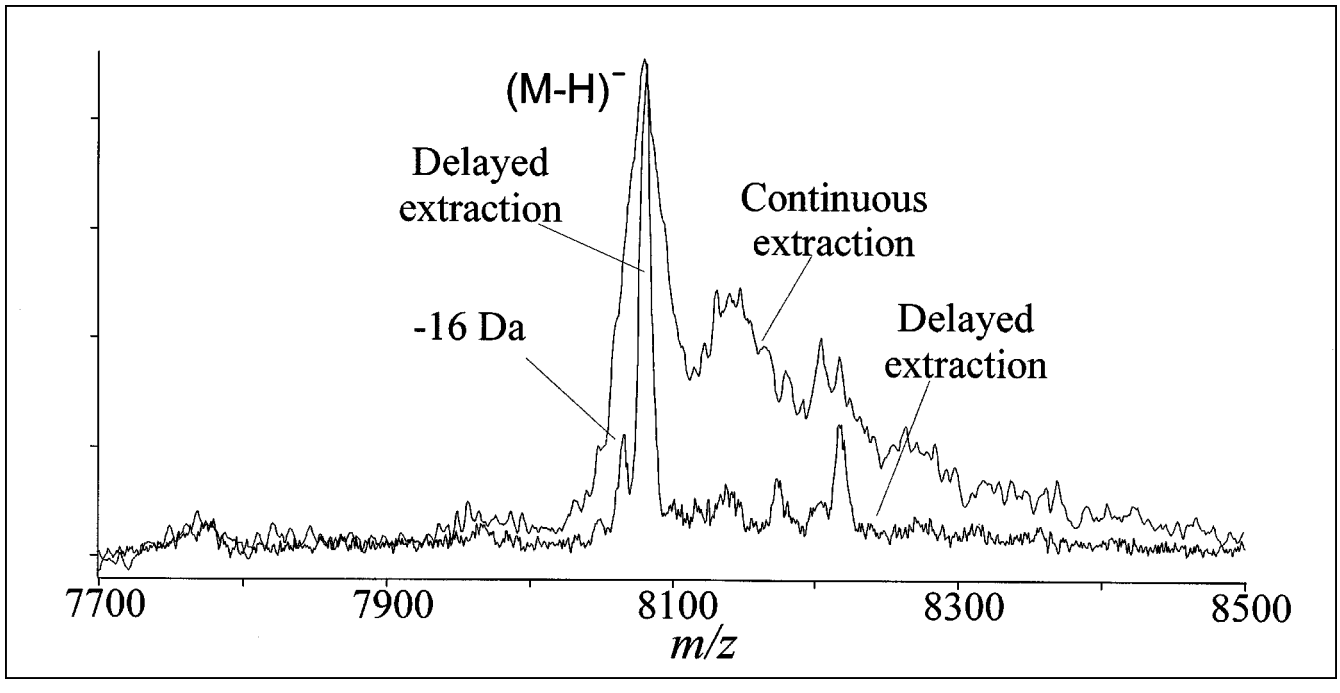
A major advance in the performance of MALDI-TOF instruments has been the result of development and commercialization of the delayed ion extraction technique [66,67]. Following the laser pulse leading to the ionization/desorption event, a delay period of typically 100–300 nsec is employed before the ion extraction pulse, allowing dissipation of higher energy ions, and resulting in narrowing of the energy distribution of the extracted ions. This results in significantly increased MS resolution as shown in Figure 3 [68••], and also higher mass measurement accuracy. This modification is now becoming available on most commercial MALDI-TOF systems, and is reflected in improved data published from such instruments acquired after about mid-1996.

The return in recent years of the TOF mass analyzer as a widely used instrument is attributable to its twofold complementarity to MALDI: both operate in a pulsed mode; and the TOF analyzer has unlimited mass range, of importance for measurement of large m/z value ions because (unlike ions from electrospray) z is typically 1. A number of hybrid mass analyzers incorporating the TOF analyzer have been described [16•], but one of the most impressive in terms of performance is the quadrupole/orthogonal-acceleration TOF tandem mass spectrometer [69•] which takes advantage of the high efficiency of the TOF analyzer as MS-2 (the second mass analyzer, which follows the collision cell).

Of practical importance has been greater understanding and availability of the nanospray [70•] or microspray (see [4]) ESI device in which nL/min rather than µL/min range flow rates are employed. The needle tip is positioned very close to the entry orifice to the mass spectrometer, resulting in transmission of a much larger fraction of ions into the vacuum system compared with conventional ESI sources. This leads to femtomole level [70•] (and below [71]) sensitivity levels; however, sample solution consumed during initial sample preparation and transfer must also be taken into account when interpreting such reports in terms of routine applications.

The difficult experimental problem of detecting and mass-measuring very large ions (>1 MDa) of DNA has been approached using two different techniques, but both employ ESI. Using Fourier transform ion cyclotron resonance mass spectrometry, individual plasmid DNA ions of mass 1.95 MDa were detected and measured with an accuracy of ± 0.2% [72]. One of the problems surrounding such an extraordinary measurement, however, is heterogeneity in the DNA, which is not easily recognized because ion populations are not sampled. A

Figure 3



Comparison of continuous and delayed ion extraction MALDI mass spectra of a phosphorothioate oligonucleotide 26-mer (M_r 8,081 [theoretical]), showing increased resolution and decreased peak tailing on the high mass side. The minus 16 Da peak is assigned as a phosphorothioate oxidation impurity (C=SC=O). Adapted with permission from [68••].

novel electrostatic ion trap in which a population of pBR322 DNA was measured as 2.88 MDa (Na form) was recently reported and the instrument described in detail [73]. In both cases, the techniques used are less accurate and less practical (as was recognized) than M_r measurements by conventional means; however, they represent significant stepping stones in types of gas phase measurements thought impossible a few years ago.

Applications to problem solving

In areas of rapidly developing technology, a careful distinction should be drawn between demonstrations of applicability to model systems (solution to the problem known in advance or obvious from earlier work), and actual applications to problem solving (answer not known in advance). The number of methodological proposals, often having minor or questionable advantages, usually greatly outnumber true applications that allow a more realistic assessment of the advantages and problems likely to be encountered. Compared with our earlier review in 1995 [4], a sharply increased number of applications of MALDI and ESI methods to nucleic acids have recently appeared, to some of which the above caveat can be applied. Some of those which constitute good examples of the range of potential applications one may expect to see over the near future include the characterization of incorrect or unexpected products of oligonucleotide solid phase synthesis [74], characterization of products of oligonucleotide platination reactions [75•,76] (or other modified oligonucleotide synthesis products), studies of structure and distribution (by tissue type) of *in vivo* metabolic products of antisense oligonucleotides [54•,58•], use of the ESI method for study of the stoichiometry or kinetics of non-covalent associations [45,49], determination of post-transcriptional modifications in RNA [77•], and characterization of the structures of RNA (or DNA) protein crosslinks [78•,79].

In particular, two of these approaches will be of unusual value in future studies in their respective fields. The first is in the extension of LC/MS analysis methods [53••,54•,58•] (see earlier section of this review) to a wide range of problems involving mixtures of oligonucleotides. Because of the relative high accuracy of mass measurement afforded by ESI mass spectrometry, a number of problems can be solved simply by measurement of individual molecular masses (typically ± 0.2–0.3 Da for a 15-mer) in multicomponent mixtures in a single LC/MS run without conventional isolation of individual components. Such measurements can be made using relatively simple LC/MS instruments, such as the quadrupole ion trap [54•]. The second area is in the structural characterization of protein–nucleic acid cross-linked products [4,24•,78•], an area fraught with experimental difficulty using more conventional methods. Here also, accurate measurement of mass is a highly effective parameter, because the amino acid and nucleotide compositions and sequences of linked fragments are constrained by the starting (unlinked) structures. Interestingly, in some cases the

exact structure, and therefore the mass, of the cross-linked moiety may not be known. This then constitutes an initial hindrance to the interpretation of mass measurement data, although in principle the measurements may be sufficiently accurate to distinguish candidate possibilities. An interesting example (of perhaps more to come) was the conclusion that the ε-amino group of lysine in methionyl-tRNA formyltransferase was bound to AMP (following oxidation) from the tRNA 3′-terminus, by a morpholino, but not hydroxymorpholino or Schiffs base, linkage [79].

Conclusions

The number of publications dealing with mass spectrometry of nucleic acids has risen sharply over the past 23 years. In terms of methodology reported, we view the development and availability of the MALDI delayed ion extraction technique to be the single most important advance with regard to immediate impact. On the other hand LC/MS using ESI is, in general, under utilized in respect to the number of potential applications involving analysis of oligonucleotide mixtures. Sequencing techniques, of which a number have been proposed, have (judging by the published literature) been slow to find their way into routine applications. This will probably change and acceptance will increase with greater awareness of the utility of the technique. The use of MS in targeted DNA assays (genotyping, DNA diagnostics) is one of the most fascinating areas of current interest, but the jury may be out for several more years before the impact of these novel technologies still under development can be fairly assessed. To a large extent, while MS methodologies as such will continue to develop, a relatively greater fraction of effort in the future will be devoted to sample preparation and manipulation, a trend which is already in place.

Acknowledgements

The authors are grateful for contributions from various colleagues who provided copies of manuscripts prior to publication, materials for Figure 1–3, and for support from the National Institute of General Medical Sciences (GM29812).

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM: **Electrospray ionization for mass spectrometry of large biomolecules.** *Science* 1989, **246**:64-71.

2. Karas M, Hillenkamp F: **Laser desorption ionization of proteins with molecular masses exceeding 10,000 Daltons.** *Anal Chem* 1988, **60**:2299-2301.

3. Burlingame AL, Boyd RK, Gaskell SJ: **Mass spectrometry.** *Anal Chem* 1996, **68**:599R-651R.

4. Limbach PA, Crain PF, McCloskey JA: **Characterization of oligonucleotides and nucleic acids by mass spectrometry.** *Curr Opin Biotechnol* 1995, **6**:96-102.

5. Roepstorff P: **Mass spectrometry in protein studies from genome to function.** *Curr Opin Biotechnol* 1997, **8**:6-13.

6. •• Nordhoff E, Kirpekar F, Roepstorff P: **Mass spectrometry of nucleic acids.** *Mass Spectrom Rev* 1996, **15**:67-138.
This extensive review, with 190 citations, covers the literature until July 1996. Both ESI and MALDI are included, with a section devoted to comparison of some aspects of the two ionization methods, and good summaries of gas phase chemistry including proposed fragmentation mechanisms.

7. Cole RB (Ed): *Electrospray Ionization Mass Spectrometry. Fundamentals, Instrumentation, and Applications.* New York: John Wiley & Sons; 1997.

8. Murray KK: **DNA sequencing by mass spectrometry.** *J Mass Spectrom* 1996, **31**:1203-1215.

9. •• Limbach PA: **Indirect mass spectrometric methods for characterizing and sequencing oligonucleotides.** *Mass Spectrom Rev* 1996, **15**:297-336.
A clear presentation of all contemporary MS-based methods proposed for oligonucleotide sequencing, which concentrates on indirect methods involving mass measurement of chain cleavage products formed in solution.

10. Fitzgerald MC, Smith LM: **Mass spectrometry of nucleic acids: the promise of matrix-assisted laser desorption-ionization (MALDI) mass spectrometry.** *Annu Rev Biophys Biomol Struct* 1995, **24**:117-140.

11. Crain PF: **Nucleic acids and their constituents.** In *Electrospray Ionization Mass Spectrometry. Fundamentals, Instrumentation and Applications.* Edited by RB Cole. New York: John Wiley & Sons; 1997:421-427.

12. Stults JT: **Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS).** *Curr Opin Struct Biol* 1995, **5**:691-698.

13. Hop CECA, Bakhtiar R: **An introduction to electrospray ionization and matrix-assisted laser desorption/ionization mass spectrometry: essential tools in a modern biotechnology environment.** *Biospectroscopy* 1997, **3**:259-280.

14. • Henry C: **Can MS really compete in the DNA world?** *Anal Chem* 1997, **69**:243A-246A.
A lively summary of viewpoints regarding expectations and reality in the use of MS for DNA sequencing, including large scale efforts (unlikely) and DNA diagnostics (more promising).

15. • Smith LM: **Sequence from spectrometry: a realistic prospect?** *Nat Biotechnol* 1996, **14**:1084-1085.
A brief and readable assessment of the issues facing development of practical MALDI-based sequencing schemes.

16. • Gilhaus M, Mlynski V, Selby D: **Perfect timing: time-of-flight mass spectrometry.** *Rapid Commun Mass Spectrom* 1997, **11**:951-962.
This excellent paper reviews the operating principles and characteristics of the TOF mass analyzer and describes several novel multianalyzer (hybrid) configurations.

17. Nordhoff E, Karas M, Cramer R, Hahner S, Hillenkamp F, Kirpekar F, Lezius A, Muth J, Meier C, Engels JW: **Direct mass spectrometric sequencing of low-picomole amounts of oligodeoxynucleotides with up to 21 bases by matrix-assisted laser desorption/ionization mass spectrometry.** *J Mass Spectrom* 1995, **30**:99-112.

18. • Ni J, Pomerantz SC, Rozenski J, Zhang Y, McCloskey JA: **Interpretation of oligonucleotide mass spectra for determination of sequence using electrospray ionization and tandem mass spectrometry.** *Anal Chem* 1996, **68**:1989-1999.
Procedures are described for manual or rapid computer-based interpretation of oligonucleotide mass spectra from collision-induced dissociation. Application to completely unknown sequences and to oligonucleotides modified in the base or sugar are emphasized.

19. Zhu L, Parr GR, Fitzgerald MC, Nelson CM, Smith LM: **Oligodeoxynucleotide fragmentation in MALDI/TOF mass spectrometry using 355-nm radiation.** *J Am Chem Soc* 1995, **117**:6048-6056.

20. Zhu YF, Taranenko NI, Allman SL, Taranenko NV, Martin SA, Haff LA, Chen CH: **Oligonucleotide sequencing by fragmentation in matrix-assisted laser desorption/ionization time-of-flight mass spectrometry.** *Rapid Commun Mass Spectrom* 1997, **11**:897-903.

21. Pomerantz SC, McCloskey JA, Tarasow T, Eaton BE: **Deconvolution of combinatorial oligonucleotide libraries by electrospray ionization tandem mass spectrometry.** *J Am Chem Soc* 1997, **119**:3861-3867.

22. Bartlett MG, McCloskey JA, Manalili S, Griffey RH: **The effect of backbone charge on the collision-induced dissociation of oligonucleotides.** *J Mass Spectrom* 1996, **31**:1277-1283.

23. Ni J, Mathews MAA, McCloskey JA: **Collision-induced dissociation of polyprotonated oligonucleotides produced by electrospray ionization.** *Rapid Commun Mass Spectrom* 1997, **11**:535-540.

24. • Jensen ON, Kulkarni S, Aldrich JV, Barofsky DF: **Characterization of peptide oligonucleotide heteroconjugates by mass spectrometry.** *Nucleic Acids Res* 1996, **24**:3866-3872.
This paper explores analytical strategies using MALDI and ESI, for detailed sequence analysis of both the peptide and oligonucleotide moieties in a model conjugate (see Figure 1). Although the selection of experimental parameters (e.g. positive versus negative ions) may vary depending on structural features, the data presented illustrate the potential of mass spectrometry in this area.

25. Roskey MT, Juhasz P, Smirnov IP, Takach EJ, Martin SA, Haff LA: **DNA sequencing by delayed extraction-matrix-assisted laser desorption/ionization time of flight mass spectrometry.** *Proc Natl Acad Sci USA* 1996, **93**:4724-4729.

26. Taranenko NI, Chung CN, Zhu YF, Allman SL, Golovlev VV, Isola NR, Martin SA, Haff LA, Chen CH: **Matrix-assisted laser desorption/ionization for sequencing single-stranded and double-stranded DNA.** *Rapid Commun Mass Spectrom* 1997, **11**:386-392.

27. Hahner S, Lüdemann H-C, Kirpekar F, Nordhoff E, Roepstorff P, Galla H-J, Hillenkamp F: **Matrix-assisted laser desorption/ionization mass spectrometry (MALDI) of endonuclease digests of RNA.** *Nucleic Acids Res* 1997, **25**:1957-1964.

28. Naito Y, Ishikawa K, Koga Y, Tsuneyoshi T, Terunuma H, Arakawa R: **Molecular mass measurement of polymerase chain reaction products amplified from human blood DNA by electrospray ionization mass spectrometry.** *Rapid Commun Mass Spectrom* 1995, **9**:1484-1486.

29. Wunschel DS, Fox KF, Fox A, Bruce JE, Muddiman DC, Smith RD: **Analysis of double-stranded polymerase chain reaction products from the *Bacillus cereus* group by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry.** *Rapid Commun Mass Spectrom* 1996, **10**:29-35.

30. Naito Y, Ishikawa K, Koga Y, Tsuneyoshi T, Terunuma H, Arakawa R: **Genetic diagnosis by polymerase chain reaction and electrospray ionization mass spectrometry: detection of five-base pair deletion from blood DNA of a familial adenomatous polyposis patient.** *J Am Soc Mass Spectrom* 1997, **8**:737-742.

31. Hurst GB, Doktycz MJ, Vass AA, Buchanan MV: **Detection of bacterial DNA polymerase chain reaction products by matrix-assisted laser desorption/ionization mass spectrometry.** *Rapid Commun Mass Spectrom* 1996, **10**:377-382.

32. Jurinke C, van den Boom D, Jacob A, Tang K, Wörl R, Köster H: **Analysis of ligase chain reaction products via matrix-assisted laser desportion/ionization time-of-flight-mass spectrometry.** *Anal Biochem* 1996, **237**:174-181.

33. Srinivasan JR, Liu Y-H, Venta PJ, Siemieniak D, Killeen AA, Zhu Y, Lubman DM: **Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry as a rapid screening method to detect mutations causing Tay-Sachs disease.** *Rapid Commun Mass Spectrom* 1997, **11**:1144-1150.

34. • Haff LA, Smirnov IP: **Single-nucleotide polymorphism identification assays using a thermostable DNA polymerase and delayed extraction MALDI-TOF mass spectrometry.** *Genome Res* 1997, **7**:378-388.
The PinPoint assay described here probes the identity of a single nucleotide at the location of interest. PCR is initiated with a primer terminating one nucleotide upstream of the target site and four ddNTPs. The primer is elongated by only one residue, corresponding to the base complementary to the site of interest, and gives a short, readily analyzed DNA fragment.

35. • Braun A, Little DP, Reuter D, Müller-Mysok B, Köster H: **Improved analysis of microsatellites using mass spectrometry.** *Genomics* 1997, **46**:in press.
The generation of informative short PCR products using a modification of the primer oligo base extension protocol is described. By adjusting the dNTP/ddNTP ratio, unexpected second site mutations in could be detected in microsatellite DNA.

36. Ross PL, Belgrader P: **Analysis of short tandem repeat polymorphisms in human DNA by matrix-assisted laser desorption/ionization mass spectrometry.** *Anal Chem* 1997, **69**:3966-3972.

37. Ross PL, Lee K, Belgrader P: **Discrimination of single-nucleotide polymorphisms in human DNA using peptide nucleic acid probes detected by MALDI-TOF mass spectrometry.** *Anal Chem* 1997, **69**:41974202.

38. Tang K, Fu D, Kötter S, Cotter RJ, Cantor CR, Köster H: **Matrix-assisted laser desorption/ionization mass spectrometry of immobilized duplex DNA probes.** *Nucleic Acids Res* 1995, **23**:3126-3131.

39. O'Donnell MJ, Tang K, Köster H, Smith CL, Cantor CR: **High density, covalent attachment of DNA to silicon wafers for analysis by MALDI-TOF mass specctrometry.** *Anal Chem* 1997, **69**:2438-2443.

40. • Little DP, Cornish TJ, O'Donnell MJ, Braun A, Cotter RJ, Köster H: **MALDI on a chip: analysis of arrays of low-femtomole to subfemtomole quantities of synthetic oligonucleotides and DNA diagnostic products dispensed by a piezoelectric pipet.** *Anal Chem* 1997, **69**:4540-4546.
This paper addresses an important component of protocols under development for high-throughput automated analyses. A piezoelectric pipet is used to dispense low nanoliter volumes of matrix and DNA into individual wells, so that the sample area is approximately the same size as the laser irradiation area. The result is high sensitivity (Figure 2), greater spot-to-spot uniformity, and avoidance of the need to search the target surface for optimal sensitivity regions.

41. Xue Q, Foret F, Dunayevskiy YM, Zavracky PM, McGruer NE, Karger BL: **Multichannel microchip electrospray mass spectrometry.** *Anal Chem* 1997, **69**:426-430.

42. Braun A, Little DP, Köster H: **Detecting CFTR gene mutations by using primer oligo base extension and mass spectrometry.** *Clin Chem* 1997, **43**:1151-1158.

43. Haff LA, Smirnov IP: **Multiplex genotyping of PCR products with MassTag-labeled primers.** *Nucleic Acids Res* 1997, **25**:3749-3750.

44. •• Köster H, Tang K, Fu D-J, Braun A, van den Boom D, Smith CL, Cotter RJ, Cantor CR: **A strategy for rapid and efficient DNA sequencing by mass spectrometry.** *Nat Biotechnol* 1996, **14**:1123-1128.
This optimistic paper summarizes recent experiments and future plans for the application of MALDI to high-throughput sequencing. Two approaches are described, utilizing immobilized synthetic template or immobilized duplex probes, as elements of a future automated method for mass measurement of sequencing ladders.

45. Greig MJ, Gaus H, Cummins LL, Sasmor H, Griffey R: **Measurement of macromolecular binding using electrospray mass spectrometry. Determination of dissociation constants for oligonucleotide:serum albumin complexes.** *J Am Chem Soc* 1995, **117**:10765-10766.

46. • Loo JA: **Studying noncovalent protein complexes by electrospray ionization mass spectrometry.** *Mass Spectrom Rev* 1997, **16**:123.
This extensive review includes protein–nucleic acid complexes and a good discussion of experimental variables, as well as comparison of solution phase versus gas phase binding characteristics, and the complementarity of the ESI technique to various biophysical methods.

47. Ding J, Anderegg RJ: **Specific and nonspecific dimer formation in the electrospray ionization mass spectrometry of oligonucleotides.** *J Am Soc Mass Spectrom* 1995, **6**:159-164.

48. Cheng X, Harms AC, Goudreau PN, Terwilliger TC, Smith RD: **Direct measurement of oligonucleotide binding stoichiometry of gene V protein by mass spectrometry.** *Proc Natl Acad Sci USA* 1996, **93**:7022-7027.

49. Sannes-Lowery KA, Hu P, Mack DP, Mei H-Y, Loo JA: **HIV-1 Tat peptide binding to TAR RNA by electrospray ionization mass spectrometry.** *Anal Chem* 1997, **69**:in press.

50. Gao Q, Cheng X, Smith RD, Yang CF, Goldberg IH: **Binding specificity of post-activated neocarzinostatin chromophore**

drug–bulged DNA complex studied using electrospray ionization mass spectrometry. *J Mass Spectrom* 1996, **31**:31-36.

51. • Wu Q, Cheng X, Hofstadler SA, Smith RD: **Specific metal-oligonucleotide binding studied by high resolution tandem mass spectrometry.** *J Mass Spectrom* 1996, **31**:669-675.
In this thought provoking paper, low energy collision-induced dissociation of Na^+, Mg^{2+} and UO^{2+} complexes with oligodeoxynucleotides is used to derive information on binding sites and relative binding specificities of the metal ions. Uncertainties in the validity of this approach *vis a vis* solution structure clearly remain but this work points the way to additional studies.

52. • Griffey RH, Greig MJ: **Detection of base pair mismatches in duplex DNA and RNA oligonucleotides using electrospray mass spectrometry.** *Proc Int Soc Optical Eng (SPIE)* 1997, **2985**:82-86.
Innovative experiments show that gas-phase oligonucleotide duplexes that contain a single base pair mismatch undergo selective backbone cleavage at that site upon collisional dissociation. The authors suggest that such measurements could ultimately be used to selectively identify genetic mutations, but other applications, such as probes of secondary structure, might also be possible.

53. •• Apffel A, Chakel JA, Fischer S, Lichtenwalter K, Hancock WS: **Analysis of oligonucleotides by HPLC-electrospray ionization mass spectrometry.** *Anal Chem* 1997, **69**:1320-1325.
The authors attribute the effectiveness of this novel solvent system to the volatility of the hexafluoro-2-propanol buffer constituent, which is preferentially depleted at the droplet surface during desolvation. The pH then increases to ~10 so that the oligonucleotide-triethylamine ion pairs dissociate, releasing the oligonucleotide into the gas phase.

54. • Griffey RH, Greig MJ, Gaus HJ, Liu K, Monteith D, Winniman M, Cummins LL: **Characterization of oligonucleotide metabolism *in vivo* via liquid chromatography/electrospray tandem mass spectrometry with a quadrupole ion trap mass spectrometer.** *J Mass Spectrom* 1997, **32**:305-313.
Complete identification of phosphorothioate oligonucleotide metabolites from pig kidney was achieved using LC/ESI tandem mass spectrometry in a quadrupole ion trap mass spectrometer, using an HFIP mobile phase (see Apffel *et al.* 1997 [53••]).

55. Greig M, Griffey RH: **Utility of organic bases for improved electrospray mass spectrometry of oligonucleotides.** *Rapid Commun Mass Spectrom* 1995, **9**:97-102.

56. Limbach PA, Crain PF, McCloskey JA: **Molecular mass measurement of intact ribonucleic acids using electrospray ionization quadrupole mass spectrometry.** *J Am Soc Mass Spectrom* 1995, **6**:27-39.

57. Gruć-Sovulj I, Lüdemann H-C, Hillenkamp F, Weygand-Durašević I, Kućan Ž, Peter-Katalinić J: **Matrix-assisted laser desorption/ionisation mass spectrometry of transfer ribonucleic acids isolated from yeast.** *Nucleic Acids Res* 1997, **25**:1859-1861.

58. • Gaus HJ, Owens SR, Winniman M, Cooper S, Cummins LL: **On-line HPLC electrospray mass spectrometry of phosphorothioate oligonucleotide metabolites.** *Anal Chem* 1997, **69**:313-319.
This paper, although written prior to report of the HFIP solvent system described by Apffel *et al.* 1997 [53••], provides an excellent example of the uses of mass spectrometry in the identification of oligonucleotide metabolites, 2-deoxyphosphorothioates in this case. The distribution of metabolites in rat plasma, liver and kidney were determined, and were consistent with exonuclease degradation, (primarily) at the 3′-termini.

59. Liu C, Wu Q, Harms AC, Smith RD: **On-line microdialysis sample cleanup for electrospray ionization mass spectrometry of nucleic acid samples.** *Anal Chem* 1996, **68**:3295-3299.

60. Shaler TA, Wickham JN, Sannes KA, Wu KJ, Becker CH: **Effect of impurities on the matrix-assisted laser desorption mass spectra of single-stranded oligodeoxynucleotides.** *Anal Chem* 1996, **68**:576-579.

61. Simmons TA, Limbach PA: **The use of a co-matrix for improved analysis of oligonucleotides by matrix-assisted laser desorption/ionization time of-flight mass spectrometry.** *Rapid Commun Mass Spectrom* 1997, **11**:567-572.

62. Börnsen KO, Gass MAS, Bruin GJM, von Adrichem JHM, Biro MC, Kresbach GM, Ehrat M: **Influence of solvents and detergents on matrix-assisted laser desorption/ionization mass spectrometry measurements of proteins and oligonucleotides.** *Rapid Commun Mass Spectrom* 1997, **11**:603-609.

63. Hunter JM, Lin H, Becker CH: **Cryogenic frozen solution matrixes for analysis of DNA by time-of-flight mass spectrometry.** *Anal Chem* 1997, **69**:3608-3612.

64. Tang W, Zhu L, Smith LM: **Controlling DNA fragmentation in MALDI-MS by chemical modification.** *Anal Chem* 1997, **69**:302-312.

65. Gut IG, Jeffrey WA, Pappin DJC, Beck S: **Analysis of DNA by charge tagging and matrix-assisted laser desorption/ionization mass spectrometry.** *Rapid Commun Mass Spectrom* 1997, **11**:43-50.

66. Vestal ML, Juhasz P, Martin SA: **Delayed extraction matrix-assisted laser desorption time-of-flight mass spectrometry.** *Rapid Commun Mass Spectrom* 1995, **9**:1044-1050.

67. Christian NP, Colby SM, Giver L, Houston CT, Arnold RJ, Ellington AD, Reilly JP: **High resolution matrix-assisted laser desorption/ionization time-of-flight analysis of single-stranded DNA of 27 to 68 nucleotides in length.** *Rapid Commun Mass Spectrom* 1995, **9**:1061-1066.

68. •• Juhasz P, Roskey MT, Smirnov IP, Haff LA, Vestal ML, Martin SA: **Applications of delayed extraction matrix-assisted laser desorption ionization time-of-flight mass spectrometry to oligonucleotide analysis.** *Anal Chem* 1996, **68**:941-946.
The authors report notable improvements in instrument performance in the linear, and in particular the reflector mode, of operation when using delayed extraction techniques, demonstrated up to 50 nucleotides in length.

69. • Morris HR, Paxton T, Dell A, Langhorne J, Berg M, Bordoli RS, Hoyes J, Bateman RH: **High sensitivity collisionally activated decomposition tandem mass spectrometry on a novel quadrupole/orthogonal-acceleration time-of-flight mass spectrometer.** *Rapid Commun Mass Spectrom* 1996, **10**:889-896.
The instrument (Q-TOF) described consists of a quadrupole mass filter, hexapole collision cell and time-of-flight mass analyzer with reflectron. Using ESI, sequencing of peptides in low femtomole to attomole amounts was demonstrated. These initial sensitivity results suggest a favorable future for this instrument design in a range of applications.

70. • Wilm M, Mann M: **Analytical properties of the nanoelectrospray ion source.** *Anal Chem* 1996, **68**:1-8.
This paper discusses the theory and practical operating aspects of the nanospray ion source. This ion source utilizes very small diameter electrospray capillary needles (1–2 µm inner diameter), which are loaded with less than 1 µl of sample solution and operate without necessity of a solvent pump. In general, the operation of the device becomes less routine, compared with the somewhat larger scale microspray ion source (see Limbach *et al.* 1995 [4]), as the flow rate and orifice diameter are decreased.

71. Valaskovic GA, Kelleher NL, Little DP, Aaserud DJ, McLafferty FW: **Attomole-sensitivity electrospray source for large-molecule mass spectometry.** *Anal Chem* 1995, **87**:3802-3805.

72. Cheng X, Camp IDG II, Wu Q, Bakhtiar R, Springer DL, Morris BJ, Bruce JE, Anderson GA, Edmonds CG, Smith RD: **Molecular weight determination of plasmid DNA using electrospray ionization mass spectrometry.** *Nucleic Acids Res* 1996, **24**:2183-2189.

73. Benner WH: **A gated electrostatic ion trap to repetitiously measure the charge and *m/z* of large electrospray ions.** *Anal Chem* 1997, **69**:4162-4168.

74. Fearon KL, Stults JT, Bergot BJ, Christensen LM, Raible AM: **Investigation of the 'n - 1' impurity in phosphorothioate oligodeoxynucleotides synthesized by the solid-phase beta-cyanoethyl phosphoramidite method using stepwise sulfurization.** *Nucleic Acids Res* 1995, **23**:2754-2761.

75. • Lowe G, McCloskey JA, Ni J, Vilavian T: **Mass spectrometric investigation of the reaction between 4,4-vinylenedipyridine bis[2,2:6,2-terpyridine platinum (II)] and the self-complementary oligonucleotide d(CpGpTpApCpG).** *Bioorg Med Chem* 1996, **4**:1007-1013.
The authors report formation of one terpyridine Pt(II) ligand on the 3′-terminal guanine and demonstrate the use of several ESI techniques: sequencing by collision-induced dissociation; determination of binding stoichiometry from molecular mass; and LC/MS analysis of platinated dG released by enzymatic hydrolysis. Unusual ion charge patterns in fragment ions from the oligonucleotide adduct were observed, as a consequence of the presence of Pt^{2+}, which neutralizes the anionic phosphodiester at the site of adduction.

76. Gonnet F, Kocher F, Blais JC, Bolbach G, Tabet JC: **Kinetic analysis of the reaction between d(TTGGCCAA) and**

$[Pt(NH_3)_3(H_2O)]^{2+}$ by enzymatic degradation of the products and ESI and MALDI mass spectrometries. *J Mass Spectrom* 1996, **31**:802-809.

77. • Kowalak JA, Bruenger E, Hashizume T, Peltier JM, Ofengand J, McCloskey JA: **Structural characterization of U*-1915 in domain IV from *E. coli* 23S ribosomal RNA as 3-methylpseudouridine.** *Nucleic Acids Res* 1996, **24**:688-693.

Example of a protocol for location and characterization of modified residues in large RNAs, based on recognition of modifications in RNase T1 hydrolysis products, from mass increments (e.g. 14 Da for methyl) not allowed by the corresponding gene sequence.

78. • Urlaub H, Thiede B, Müller E-C, Brimacombe R, Wittmann-Liebold B: **Identification and sequence analysis of contact sites between ribosomal proteins and rRNA in *Escherichia coli* 30 S subunits by a new approach using matrix-assisted laser desorption/ionization-mass spectrometry combined with N-terminal microsequencing.** *J Biol Chem* 1997, **272**:14547-14555.

This excellent paper demonstrates the considerable power of mass spectrometry for the determination of nucleic acid–protein cross-link sites. Peptide-oligonucleotide complexes formed by protease-RNase treatments were isolated by HPLC. The RNA 8-mer linked product was sequenced by measurement of a mass ladder generated by partial alkaline hydrolysis.

79. Gite S, RajBhandary UL: **Lysine 207 as the site of cross-linking between the 3′-end of *Escherichia coli* initiator tRNA and methionyl-tRNA formyltransferase.** *J Biol Chem* 1997, **272**:5305-5312.