The Detection and Stability of DNA Duplexes Probed by MALDI Mass Spectrometry

by **Rajagopalan Sudha**[1]) and **Renato Zenobi***

Department of Chemistry, Swiss Federal Institute of Technology (ETH) Zürich, HCI, CH-8093 Zürich
(fax: 41-1-632 12 92; e-mail: zenobi@org.chem.eth.ch)

Dedicated to Professor Dr. *Dieter Seebach* on the occasion of his 65th birthday

Matrix-assisted laser desorption/ionization (MALDI) with nonacidic matrices is shown to generate intact gas-phase ions of double-stranded DNA. Control experiments show that specific DNA duplexes are detected by this method, while nonspecific adducts form only when high monomer concentrations are present in the MALDI sample. The relative intensity of the duplex-ion signal is found to reflect the solution-phase stability of the double-stranded DNA.

1. Introduction. – Noncovalent interactions such as H-bonding, ion-dipole interactions, hydrophobic interactions, and others play important roles in the biological function of biomacromolecules: they form the basis for molecular recognition and encode information. The advancement in soft ionization methods, in particular matrix-assisted laser desorption (MALDI) and electrospray ionization (ESI), has enabled the mass-spectrometric study of such noncovalent complexes [1][2]. These include protein – protein, enzyme – substrate, metal ion – protein, protein – oligonucleotide, oligonucleotide – drug complexes, and related systems [2 – 7].

Oligonucleotide duplexes can be used as model systems to study noncovalent interactions in the gas phase by mass-spectrometric techniques [8]. Duplex formation in oligonucleotides is governed by *Watson-Crick* base pairing and hydrophobic base-stacking interactions. Studies under solvent-free conditions facilitate the understanding of the role of solvents in stabilizing these noncovalent interactions. There are several reports in the literature on the detection of oligonucleotide duplexes in the gas phase by MS techniques. ESI Mass spectrometry has been used preferentially [9 – 17], *e.g.*, for estimating dissociation constants of noncovalent complexes, as the solution-state structural properties are believed to be largely preserved under the mild ESI conditions. The data have been shown to compare well with the melting curves obtained with solution-phase techniques [18][19]. In contrast, application of MALDI for the studies of oligonucleotide duplexes has been fairly limited [20 – 22]. The difficulties encountered in the analysis of oligonucleotide duplexes by MALDI are: *i*) weak noncovalent interactions are less likely to be preserved during the crystallization, desorption, and ionization processes, and *ii*) nonspecific cluster formation and fragmentation can occur under MALDI conditions. For example, it has been observed

[1]) Present address: Molecular Biophysics Unit, Indian Institute of Science, 560012 Bangalore, India

that oligonucleotides with random sequences fragment much more easily and are, thus, more difficult to detect by MALDI compared to thymidine-rich sequences, presumably due to differences in the gas-phase basicity of the four nucleobases [23–25]. On the other hand, the advantages of MALDI over ESI are its greater tolerance for salts and buffers and higher overall sensitivity.

The aim of this study was to investigate oligonucleotide duplexes by MALDI and to attempt to correlate the abundance of the mass spectroscopic duplex signal with the duplex stability in solution. The stability of the DNA duplexes depends on the number of complementary base pairs and on the nature of the bases, which, in turn, determines the number of H-bonds. *Kirpekar et al.* [21] have reported that a minimum of twelve base pairs is necessary for the detection of DNA duplexes under MALDI conditions. Because of the intrinsic stability of T-rich oligonucleotides, we started by investigating DNA nucleotides with A and T complementary base pairing. However, the solution-phase melting temperatures for AT-rich oligonucleotides are smaller than those of GC-rich oligonucleotides of equal length, because a GC pair forms three H-bonds, whereas an AT pair forms only two H-bonds. Hence, GC-rich oligonucleotides with varying GC content were also included in order to study the effect of H-bonding on the stability of the duplex. Generally, peak intensities in MS data decrease with higher molecular weight. Hence, sequences with varying GC content but equal length (16-mers) were studied. Sample concentration and sample-to-matrix molar ratio were adjusted so that nonspecific adduct formation was minimized. The issue of specificity in duplex formation is also addressed. To differentiate between the nonspecific dimers and the specifically bound duplexes, oligonucleotides were chosen to be non-selfcomplementary sequences. The mass difference between the complementary sequences was more than 100 Da, allowing the unambiguous distinction of the duplex from the nonspecific dimers of the single strands.

2. Results and Discussion. – MALDI Matrices such as 3-hydroxypicolinic acid (HPA), 3-amino-4-methyl-5-nitropyridine (AMNP), 6-aza-2-thiothymine (ATT), and 2,4,6-trihydroxyacetophenone (THAP) were investigated for the detection of oligonucleotides in the negative-ion-detection mode. Unusual matrices, including HPA [26], THAP [27], ATT [20][28], other nonacidic matrices [29][30], and matrices for IR-MALDI [31] have been advocaded for MALDI-MS of DNA by various research groups working in this field. Due to the presence of the negatively charged phosphate backbone, adduct formation with alkali metal ions (Na^+ and K^+) is often observed in MALDI spectra of DNA, resulting in peak broadening and loss of sensitivity. Matrix additives such as ammonium citrate can be used for sample preparation to suppress this undesirable cationization. Ammonium ions exchange with the Na^+ and K^+ ions, which, in turn, are complexed with citrate. Upon crystallization of the exchanged sample with the matrix, ammonia is liberated, leading to the formation of protonated oligonucleotides. Other matrix additives such as fucose, spermine, or spermidine [32–34] have also been reported for the analysis of oligonucleotides. Among all the matrix systems tested, ATT with ammonium citrate as a comatrix was found to give the best results. Optimization of the matrix-to-analyte ratio and sample preparation was important for the detection of the oligonucleotides and their duplexes; a matrix-to-analyte molar ratio of 1500 : 1 was found to be optimal. Repeated warming and cooling of the sample

resulted in the loss of signal intensity in the mass spectra; fresh aliquots of samples were thus always used. The oligodeoxynucleotide samples studied and the abbreviations used for them here are listed in *Table 1*. Fragmentation and formation of oligonucleotide sequence ions was observed at elevated laser power; even at a laser pulse energy only slightly above the threshold for ion formation (*ca.* 20 μJ/pulse) it could not be completely avoided.

Table 1. *Oligonucleotides Studied, Nonspecific and Specific Dimer Complexes, and Their Molecular Weights* (M.W.)

Code	Sequence	M.W. [Da] of monomer	M.W. [Da] of homo-dimers	M.W. [Da] of specific dimer
16G10C6	d(GGGCCGCGGCGCCGGG)	4965	9930	9770
16C10G6	d(CCCGGCGCCGCGGCCC)	4805	9610	
16G9C3	d(GGGCGGATTAGCCGGG)	5003	10006	9766
16C9G3	d(CCCGCCTAATCGGCCC)	4763	9526	
16G6C2	d(GGGCATTATAATCGGG)	4961	9922	9762
16C6G2	d(CCCGTAATATTAGCCC)	4801	9602	
T15	d(T)15	4501	9002	9137
A15	d(A)15	4636	9272	

Fig. 1 shows mass spectra of 1:1 mixtures of three complementary DNA strands with content of CG base pairs decreasing from 100% (*a*) to 75% (*b*) to 0% (*c*). First, alkalimetal adducts are almost completely absent in the data, in the monomer mass range of *ca.* 4800 Da as well as in the duplex mass range of *ca.* 9600 Da. This shows that the ammonium citrate comatrix performs well to eliminate Na^+ and K^+ adducts from the spectra. The mass spectra for the oligonucleotides 16G10C6 and 16G9C3, annealed with the corresponding complementary sequences at pH 8, show the presence of pure duplex only in the dimer region. For the 16C10G6 – 16G10C6 pair, the peak observed at 9767 Da (calculated average M.W. 9769 Da for the deprotonated duplex) is strong, reaching a peak intensity of 70% for the most intense single-strand peak, the signal of the 16G10C6 monomer at 4964 Da (*Fig. 1*,*a*). A total of 48 H-bonds can form between the two complementary monomer strands. The other peaks in the duplex range are not due to nonspecific duplex formation, but rather to fragmentation. The peaks observed at $m/z = 9656$, 9618, 9478, and 9437 (marked with 'F') are interpreted as fragments resulting from loss of cytosine (calculated $\Delta m = 110$ Da), guanine (150 Da), cytidine (289 Da), and guanosine (329 Da), respectively. It is quite unlikely that these are duplexes formed in the gas phase by recombination of a single-stranded DNA with a fragment of the complementary strand, because the signal intensities of the corresponding ions in the monomer range are not intense enough for this pathway to be viable. The binding energy for duplex DNA is not readily available, and depends on a number of factors, including oligonucleotide length, GC content, buffer content of the solution, *etc*. However, ΔH can be estimated [35] to be *ca.* 46 kJ/mol per GC base pair and 33kJ/mol per AT base pair, and is, to first approximation, additive as the length of the oligonucleotide duplex increases. Thus, the overall binding enthalpy for the 16-mer duplexes studied here can easily reach or even exceed 600 kJ/mol, which is larger than the binding enthalpy of an average covalent single bond. Fragmentation of covalently

bound units from the duplex rather than dissociation/unzipping of the duplex in the mass spectrometer is therefore expected to occur.

A weaker duplex intensity is observed for the 16C9G3 – 15G9C3 pair (*Fig. 1,b*), which is somewhat surprising, because there are still 44 H-bonds between the monomers in this case. The duplex intensity is only *ca.* 10% of the 16G9C3 monomer signal intensity at 5002 Da. However, the melting temperature for this duplex is also significantly lower, at 54°, than that of the 16C10G6 – 16G10C6 pair (64°). The specific 16C9G3 – 15G9C3 duplex ion is observed at 9765 Da. The only other notable signal in the duplex mass range is at 9615 Da, corresponding to loss of guanine (calculated $\Delta m = 150$ Da), not to a nonspecific dimer. In the case of the A15 – T15 pair, no signal is observed in the dimer range at all (*Fig. 1,c*). There would be 30 H bonds present in such a duplex, and the melting temperature is estimated to be only 20°.

The spectra show signals in addition to those of the monomers. Most of these can be assigned to fragments, for example, the signals at 4324 Da and 4197 Da in *Fig. 1,c* that are assigned to loss of one nucleotide unit each from the A15 and T15 strands, respectively. Sequence ions are observed down to much lower m/z, where they begin to overlap with matrix cluster signals. The small peaks above the mass of the highest-molecular-weight monomers are interpreted as impurities due to sequence errors, probably introduced in the synthetic procedure. Such sequence errors are easily detected by mass spectrometry. For example, the signal at 5144 Da in *Fig. 1,b*, might be a sequence variant of 16G9C3 or a fragment thereof.

Specificity of Duplex Formation. It is an important issue in mass spectrometry of noncovalent complexes to assure that the peaks observed correspond to specifically bound subunits rather than nonspecific aggregates or clusters. Here, we discuss some controls that can be employed to address this issue for the case of oligonucleotides. First, some of the mass spectra were recorded up to 15 kDa; no trimer peaks that would represent nonspecific adducts were observed in the data. Second, conditions can be chosen such that nonspecific adducts do show up in the duplex mass range. This is demonstrated in *Fig. 2* for the 16C6G2 – 16G6C2 pair (50% GC content). Whereas no duplex signals were observed under the standard experimental conditions used here (*Fig. 2,a*), higher monomer concentrations (5×10^{-4} M) and higher laser power resulted in the formation of nonspecific dimer adducts. The duplex region of the mass spectra showed the presence of three peaks corresponding to the two homoduplex signals in addition to the heteroduplex, indicating that the dimers observed are nonspecific adducts, probably formed in the gas phase under these mass-spectral conditions. In addition, some fragmentation is again seen in the dimer mass range as well as in the monomer mass range. In this case, the peaks below m/z 9602, the 16C6G2 homodimer, are interpreted as clusters formed in the gas phase between monomer strands and monomeric fragments. Under the conditions of high monomer concentration, clustering involving monomers fragments is probable, due to an increased density of compounds in the plume. The broad hump that includes a number of sharp peaks between m/z 5900 – 7400 is also interpreted as fragmentation from the dimer range. Some of the mass differences between pronounced peaks in this mass range do correspond to nucleosides or nucleotides, but a detailed analysis and interpretation was not attempted.

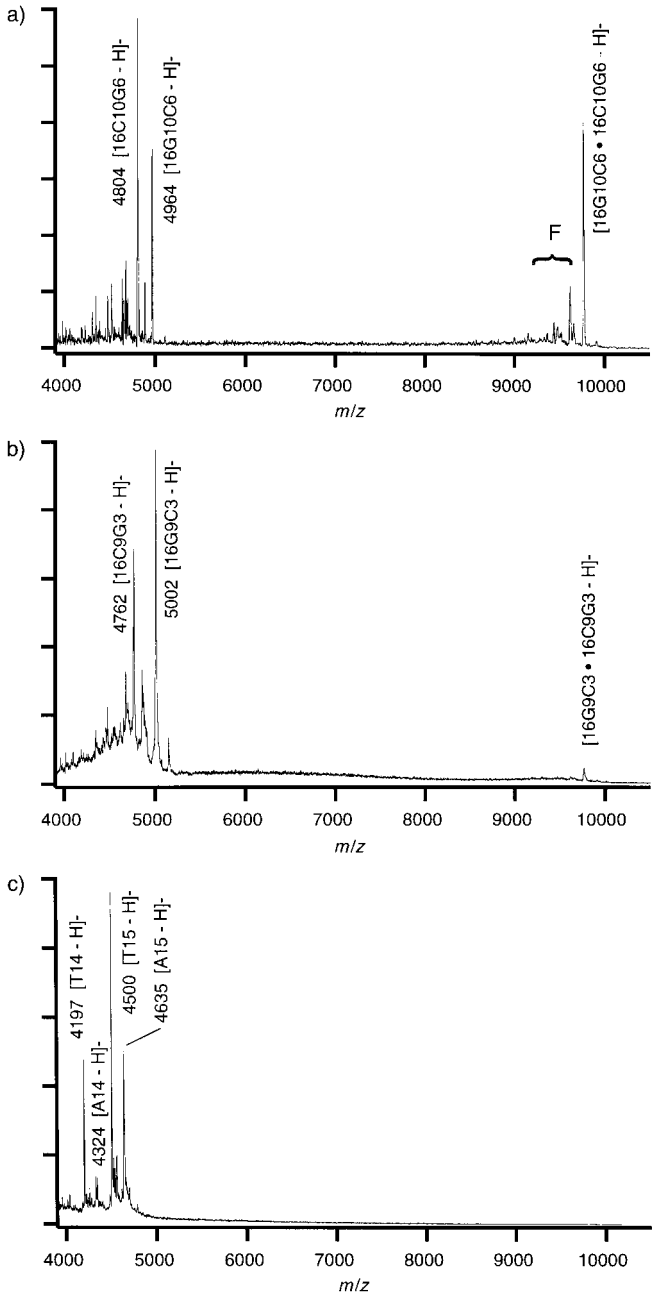

Fig. 1. *Negative-ion-mode MALDI mass spectra of equimolar mixtures of* a) *16C10G6 and 16G10C6 hexadecamers* (100% GC content), b) *16C9G3–16G9C3 hexadecamers* (75% GC content), *and* c) *A15–T15 pentadecamers* (0% CG content). ATT was used as the matrix, and a 1500 : 1 (matrix/monomer) molar ratio was used for sample preparation.

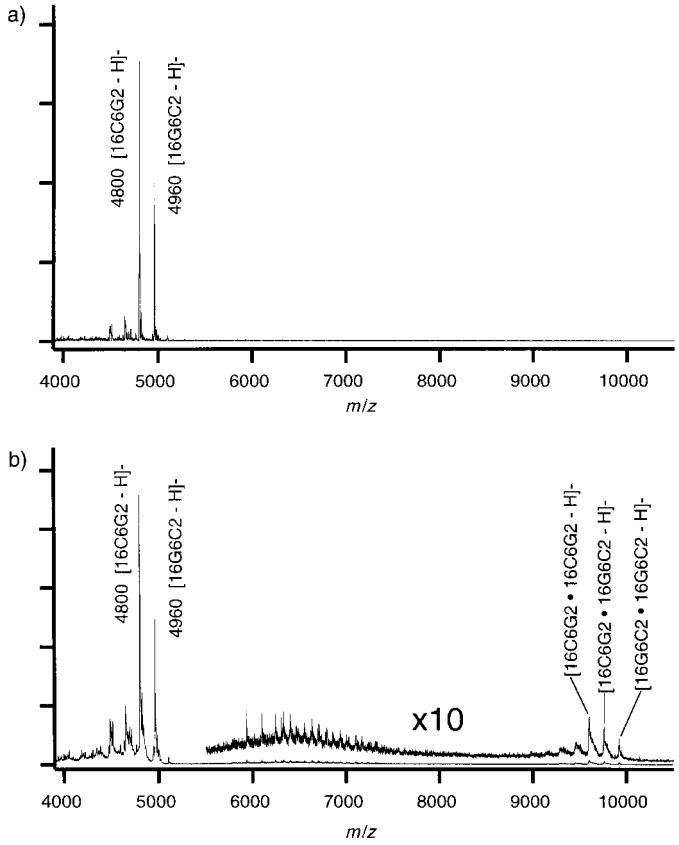

Fig. 2. *Negative-ion-mode MALDI mass spectra of an equimolar mixture of 16G6C2 and 16C6G2 hexadecamers* (50% GC content). *a*) Spectrum taken under the same experimental conditions as in *Fig. 1*; *b*) spectrum recorded at elevated monomer concentration to induce the formation of nonspecific clusters. The laser power employed was identical for *a*) and *b*).

The data in *Fig. 1* were all recorded under conditions where this clustering was absent. No peak corresponding to the nonspecific dimer of the single strands are observed. Theoretical melting temperatures calculated for the various sequences studied are given in *Table 2*. The stability of the duplex in the solution is greater for GC rich oligonucleotides than for AT rich oligonucleotides. In the mass spectra, the relative peak intensity of the duplex with respect to the single-strand signals was found to be greater for oligonucleotides with higher GC content. Interestingly, there is a clear correlation between calculated melting temperature of the duplex and the relative intensity of the most intense duplex peak. For melting temperatures less than 43°, corresponding to an estimated noncovalent binding enthalpy of 535 kJ/mol (16-mer with a CG content of 50%), the duplex cannot be observed any more by MALDI mass spectrometry. It is currently unknown whether dissociation occurs in the MALDI matrix or during the desorption/ionization event.

Table 2. *Relative Peak Intensities of Duplexes Observed and the Calculated Melting Temperature*

Oligonucleotide sequence[a])	GC content [%]	Calculated melting temperature[b]) of the duplex with complementary sequence [°]	Relative peak intensity of the duplex *vs.* single strand
d(GGGCCGCGGCGCCGGG)	100	64	70%
d(GGGCGGATTAGCCGGG)	75	54	10%
d(GGGCATTATAATCGGG)	50	43	No duplex observed
d(A)15	0	20	No duplex observed

[a]) Only the sequence of one strand (5′-3′) is indicated. The other strand contains the complementary base sequence. [b]) The 'oligo properties calculator' was used [35].

3. Conclusions. – Intact specific oligonucleotide duplexes were observed in the gas phase by MALDI mass spectrometry, with ATT as the matrix and ammonium citrate as an additive. The solution-phase binding stability (melting temperature) of the duplex is reflected in the ratio of intensities of duplex to monomer peaks in the mass spectra, indicating that the noncovalent interaction strength in solution is reflected in the MALDI mass spectra.

Experimental Part

Oligonucleotides used in this study were purchased from *Microsynth* (CH-Balgach) in HPLC-purified form and used as received. 3-Hydroxypicolinic acid, 3-amino-4-methyl-5-nitropyridine, 6-aza-2-thiothymine, 2,4,6-trihydroxyacetophenone, and ammonium citrate were from *Sigma* (CH-Buchs) and used without further purification. Sample preparation for MALDI mass spectrometry was done as follows: 1 µl of an aq. soln. of DNA nucleotides (1×10^{-5} M) was mixed with 3 µl of matrix (5×10^{-3} M) dissolved in a 1:1 mixture of MeCN/aq. ammonium citrate soln. (2×10^{-2} M). The mixture was placed on the sample holder, dried under high vacuum ('dried-droplet' sample-preparation method) and analyzed with a commercial MALDI-TOF mass spectrometer (*Axima CFR*, *Kratos/Shimadzu*, Manchester, UK) operated in linear negative-ion detection mode. The output of a N_2 laser (337 nm wavelength, *ca.* 3-ns pulse width) was focused to a spot of *ca.* 150 µm diameter for laser desorption/ionization. A laser-pulse energy of 20 µJ/shot was used for all spectra.

For the detection of oligonucleotide duplexes (specific dimers), 5×10^{-4} M soln. of complementary oligonucleotides were mixed in equal volumes, and the pH of the mixture was adjusted to 8 with NH_4OAc (1×10^{-1} M). The mixture was heated to 15° above the theoretically calculated melting temp. for the duplex and cooled slowly to r.t. The soln. was diluted to 1×10^{-5} M and prepared for MALDI analysis by the 'dried droplet' sample preparation method with ATT as the matrix.

For the observation of nonspecific dimers, annealing of the complementary sequences was done at pH 8 as mentioned above. However, a much higher concentration (5×10^{-4} M) of the samples was employed, and the soln. was analyzed without further dilution. Usually, nonspecific dimers started to appear at a monomer concentration in aq. soln. greater than 1×10^{-4} M at slightly elevated laser pulse energies.

R. S. thanks the *ICSC World Laboratory* (Lausanne, Switzerland) for a *Wilhelm-Simon* fellowship. The authors would like to thank *Sebastian D. Friess* for helpful comments and for careful proofreading of the manuscript.

REFERENCES

[1] R. D. Smith, K. J. Light-Wahl, B. E. Winger, J. A. Loo, *Org. Mass Spectrom.* **1992**, *27*, 811.
[2] J. A. Loo, *Mass Spectrom. Rev.* **1997**, *16*, 1.
[3] R. D. Smith, Z. Zhang, *Mass Spectrom. Rev.* **1994**, *13*, 411.
[4] M. Przybylski, M. O. Glocker, *Angew. Chem., Int. Ed.* **1996**, *35*, 806.
[5] R. Winston, M. C. Fitzgerald, *Mass Spectrom. Rev.* **1997**, *16*, 165.
[6] B. N. Pramanik, P. L. Bartner, U. A. Mirza, Y.-H. Liu, A. K. Ganguly, *J. Mass Spectrom.* **1998**, *33*, 911.
[7] J. M. Daniel, S. D. Friess, S. Rajagopalan, S. Wendt, R. Zenobi, *Int. J. Mass Spectrom.* **2002**, *216*, 1.
[8] S. A. Hofstadler, R. H. Griffey, *Chem. Rev.* **2001**, *101*, 377.
[9] D. R. Goodlett, D. G. Camp II, C. C. Hardin, M. Corregan, R. D. Smith, *Biol. Mass Spectrom.* **1993**, *22*, 181.
[10] B. Ganem, Y.-T. Li, J. D. Henion, *Tetrahedron Lett.* **1993**, *34*, 1445.
[11] E. Bayer, T. Bauer, K. Schmeer, K. Bleicher, M. Maier, H.-J. Gaus, *Anal. Chem.* **1994**, *66*, 3858.
[12] M. J. Doktycz, S. Habibi-Goudarzi, S. A. McLuckey, *Anal. Chem.* **1994**, *66*, 3416.
[13] J. Ding, R. J. Anderegg, *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 159.
[14] D. C. Gale, R. D. Smith, *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1154.
[15] D. J. Aaserud, N. L. Kelleher, D. P. Little, F. W. McLafferty, *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 1266.
[16] M. J. Greig, H.-J. Gaus, R. H. Griffey, *Rapid Commun. Mass Spectrom.* **1996**, *10*, 47.
[17] K. X. Wan, M. L. Gross, T. Shibue, *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 450.
[18] V. Gabelica, E. De Pauw, *J. Mass Spectrom.* **2001**, *36*, 397.
[19] P. D. Schnier, J. S. Klassen, E. F. Strittmatter, E. R. Williams, *J. Am. Chem. Soc.* **1998**, *120*, 9605.
[20] P. Lecchi, L. K. Pannell, *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 972.
[21] F. Kirpekar, S. Berkenkamp, F. Hillenkamp, *Anal. Chem.* **1999**, *71*, 2334.
[22] A. Meyer, N. Spinelli, J.-C. Brès, C. Dell'Aquila, F. Morvan, I. Lefevre, B. Rayner, J.-L. Imbach, J.-J. Vasseur, *Nucleosides Nucleotides Nucleic Acids* **2001**, *20*, 963.
[23] L. Zhu, G. R. Parr, M. C. Fitzgerald, C. M. Nelson, L. M. Smith, *J. Am. Chem. Soc.* **1995**, *117*, 6048.
[24] J. Gross, A. Leisner, F. Hillenkamp, S. Hahner, M. Karas, J. Schäfer, F. Lützenkirchen, E. Nordhoff, *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 866.
[25] J. Gross, F. Hillenkamp, K. X. Wan, M. L. Gross, *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 180.
[26] K. J. Wu, T. A. Shaler, C. H. Becker, *Anal. Chem.* **1994**, *66*, 1637.
[27] Y. F. Zhu, C. N. Chung, N. I. Taranenko, S. L. Allmann, S. A. Martin, L. Haff, C. H. Chen, *Rapid Commun. Mass Spectrom.* **1996**, *10*, 383.
[28] F. Kirpekar, E. Nordhoff, K. Kristiansen, P. Roepstorff, S. Hahner, F. Hillenkamp, *Rapid Commun. Mass Spectrom.* **1995**, *9*, 525.
[29] S. Jespersen, W. M. A. Niessen, U. R. Tjaden, J. van der Greef, *J. Mass Spectrom.* **1998**, *33*, 1088.
[30] M. C. Fitzgerald, G. R. Parr, L. M. Smith, *Anal. Chem.* **1993**, *65*, 3204.
[31] E. Nordhoff, F. Kirpekar, M. Karas, R. Cramer, S. Hahner, F. Hillenkamp, K. Kristiansen, P. Roepstorff, A. Lezius, *Nucleic Acid Res.* **1994**, *22*, 2460.
[32] J. M. Asara, J. Allison, *Anal. Chem.* **1999**, *71*, 2866.
[33] S.-W. Cheng, D. T.-W. Chan, *Rapid Commun. Mass Spectrom.* **1996**, *10*, 907.
[34] A. M. Distler, J. Allison, *Anal. Chem.* **2001**, *73*, 5000.
[35] http//www.basic.nwu.edu/biotools/oligocalc.html, Northwestern University, Copyright 1997–2000.

Received May 30, 2002