Oligodeoxynucleotide Fragmentation in MALDI/TOF Mass Spectrometry Using 355-nm Radiation

Lin Zhu, Gary R. Parr, Michael C. Fitzgerald, Christine M. Nelson, and Lloyd M. Smith*

Contribution from the Department of Chemistry, University of Wisconsin, Madison, Wisconsin 53706

Received June 23, 1994[⊗]

Abstract: The fragmentation of small oligodeoxynucleotides using matrix-assisted laser desorption/ionization (MALDI) mass spectrometry with 355-nm radiation from the matrix 2,5-dihydroxybenzoic acid is studied. Negative ion mass spectra of the homopolymer oligodeoxynucleotides $d(A)_n$, $d(C)_n$, and $d(G)_n$ (n = 4, 6, 8, and 10) show substantial cleavage at *N*-glycosidic and phosphodiester bonds, in contrast to $d(T)_n$ which shows little fragmentation. The number and intensity of the fragment peaks increases significantly with the length of the parent molecules. A series of asymmetric oligodeoxynucleotides were synthesized to study this fragmentation in greater detail. In dT_4GT_{10}, $dT_4G_4T_7$, and $dT_7G_4T_4$, the primary fragmentation pathway is loss of a base followed by backbone cleavage at the 3′ C−O bond of the corresponding deoxyribose. Similar fragmentation patterns were observed with $dT_4N_4T_7$ samples where N = C or A. A statistical cleavage model described the observed patterns of fragmentation well; the model yields fractions of backbone cleavage at A, C, and G of 0.13, 0.26, and 0.27, respectively. In contrast to the fragmentation observed using 2,5-dihydroxybenzoic acid as a matrix, little or no fragmentation is observed with the matrix 3-hydroxypicolinic acid. Fragmentation of oligodeoxynucleotides in UV/MALDI is thus both nucleobase and matrix dependent.

Introduction

Matrix assisted laser desorption/ionization (MALDI) mass spectrometry has evolved into a rapid, accurate, and sensitive method for the mass analysis of high molecular weight synthetic and biologically important polymers. The first MALDI results reported by Karas and Hillenkamp[1] and by Tanaka *et al.*[2] demonstrated that intact protein molecules could be desorbed, ionized, and subsequently mass analyzed by incorporating analyte molecules in an appropriate matrix and irradiating the sample with a high-intensity, pulsed laser beam. Karas and Hillenkamp used a nicotinic acid matrix and 266 nm radiation output from a frequency-quadrupled Nd:YAG laser, while Tanaka *et al.* used cobalt powder dispersed in glycerol and 337 nm radiation output from a nitrogen laser. These and subsequent experiments have shown that the vaporization and ionization of intact, nonvolatile, high molecular weight analytes is possible with the appropriate combination of matrix and wavelength. Since MALDI emerged in 1988, a number of new matrix materials of substantial utility have been reported. Proteins, peptides, glycopeptides, carbohydrates, oligosaccharides, polar and nonpolar synthetic polymers, and oligonucleotides have proven amenable to MALDI provided an appropriate matrix is used. An excellent summary of these matrices is given in ref 3.

Different classes of compounds have enjoyed varying degrees of success with the MALDI technique. Clearly, one of the most successful applications of MALDI has been the analysis of proteins. The discovery of matrix compounds such as 2,5-dihydroxybenzoic acid (2,5-DHBA),[4] 2-(4-hydroxyphenylazo)-benzoic acid,[3] and several cinnamic acid derivatives,[5] including ferulic, caffeic, and sinapinic acids, has made possible the analysis of a wide variety of high molecular weight proteins (<500 000 Da). Significantly, no limitations to the MALDI analysis of proteins due to primary, secondary, or tertiary structures are apparent. However, in sharp contrast to the good results obtained with proteins, MALDI analysis of mixed base oligonucleotides has been limited to relatively small molecules (<30 000 Da) and strongly dependent upon base composition and choice of matrix. Spengler *et al.*[6] first reported positive-ion mass spectra of 3 to 6 base oligonucleotides using nicotinic acid as the matrix. Results from Parr *et al.*[7] have shown that 2,5-DHBA is useful for the analysis of homopolymers of thymidine or uridine and short, single-stranded oligonucleotides (<10 nucleotides) of mixed base composition. Homopolymers of thymidine as large as 100 nucleotides in length have also been analyzed with 2,5-DHBA,[8] but other homopolymer oligodeoxynucleotides and mixed sequences longer than 10 nucleotides do not generally yield high-quality results using the same matrix. Although several studies[9−11] have focused on the search for new matrices for the analysis of nucleic acids which might eliminate these constraints of size and base composition, only a limited number of matrices have been discovered that

[⊗] Abstract published in *Advance ACS Abstracts*, May 15, 1995.

(1) Karas, M.; Bachmann, D.; Bahr, U.; Hillenkamp, F. *Int. J. Mass Spectrom. Ion. Proc.* **1987**, *78*, 53−68.

(2) Tanaka, K.; Waki, H.; Ido, Y.; Akita, S. *Rapid Commun. Mass. Spectrom.* **1989**, *3*, 436−439.

(3) Juhasz, P.; Costello, C. E.; Biemann, K. *J. Am. Soc. Mass Spectrom.* **1993**, *4*, 399−409.

(4) Strupat, K.; Karas, M.; Hillenkamp, F. *Int. J. Mass Spectrom. Ion Proc.* **1991**, *111*, 89−102.

(5) Beavis, R. C.; Chait, B. T. *Rapid Commun. Mass Spectrom.* **1989**, *3*, 432−435.

(6) Spengler, B.; Pan, Y.; Cotter, R. J.; Kan, L.-S. *Rapid Commun. Mass Spectrom.* **1990**, *4*, 99−102.

(7) Parr, G. R.; Fitzgerald, M. C.; Smith, L. M. *Rapid Commun. Mass Spectrom.* **1992**, *6*, 369−372.

(8) Tang, K.; Allman, S. L.; Jones, R. B.; Chen, C. H.; Araghi, S. *Rapid Commun. Mass Spectrom.* **1993**, *7*, 63−66.

(9) Fitzgerald, M. C.; Parr, G. R.; Smith, L. M. *Anal. Chem.* **1993**, *65*, 3204−3211.

(10) Tang, K.; Allman, S. L.; Jones, R. B.; Chen, C. H. *J. Org. Mass Spectrom.* **1992**, *27*, 389−392.

(11) Schneider, K.; Chait, B. T. *Org. Mass Spectrom.* **1993**, 1353−1361.

0002-7863/95/1517-6048$09.00/0

are applicable to the analysis of oligonucleotides. Currently, the most successful matrix for the UV–MALDI analysis of oligodeoxynucleotides is 3-hydroxypicolinic acid (3-HPA).[12] However, the largest mixed-base, single-stranded oligodeoxynucleotide analyzed with 3-HPA is still only 89 bases in length.[13] Recent results describing the IR–MALDI analysis of nucleic acids have shown that oligodeoxynucleotides up to 26 nucleotides in length and oligoribonucleotides up to 142 nucleotides can be analyzed using a succinic acid matrix.[14] These results also suggested that the relative instability of the *N*-glycosidic bond in cytosine, adenine, and guanine oligodeoxynucleotides during the desorption process is a major factor limiting the IR-MALDI analysis of larger oligodeoxynucleotides.

The current limitations of size and base composition in the MALDI analysis of oligodeoxynucleotides have generated significant interest in characterizing the unique problems associated with the application of the technique to oligodeoxynucleotide analysis. As the MALDI process itself is not generally very well understood, it has been difficult to pinpoint the specific problems related to oligodeoxynucleotide analysis. However, several important factors are likely to influence MALDI results including the following: the desorption and/or ionization efficiency of analyte molecules, the physical distribution of analyte molecules in matrix crystals, and the extent of analyte fragmentation during the MALDI process. The successful MALDI results previously reported for polythymidilic acids suggest that, at least in some cases, the desorption and/or ionization process for oligodeoxynucleotides is effective. Investigations of the physical distribution of analytes in crystals of different matrices have provided evidence for analyte incorporation into matrix crystals.[11] Details of mass spectral results obtained from different oligonucleotides using a variety of matrices[7,11] have indicated that fragmentation may limit the MALDI analysis of some oligonucleotide samples. In addition to the IR–MALDI results reported by Nordhoff and coworkers,[14] Schneider and Chait have shown that a polyguanosine heptamer is more susceptible to fragmentation than a polythymidine heptamer using UV–MALDI with 2,5-DHBA.[11]

The objective of this work was to study base-specific fragmentation during the desorption/ionization process and analyze its significance in the UV–MALDI analysis of oligodeoxynucleotides. Here we report results obtained using the matrix 2,5-DHBA to analyze a series of homopolymer oligodeoxynucleotides $d(A)_n$, $d(C)_n$, $d(G)_n$, and $d(T)_n$ where $n = 2$, 4, 6, 8, and 10. The $d(G)_n$ and $d(C)_n$ were most susceptible to fragmentation, the $d(A)_n$ was somewhat less susceptible, while the $d(T)_n$ samples showed little or no fragmentation. Results with two short oligodeoxynucleotides of sequence dT_4GT_{10} and $dT_4G_4T_7$ also showed significant fragmentation at guanosines. The primary fragmentation products for these samples resulted from cleavages of the phosphate backbone at the 3′ C–O bond of guanosine. Similar fragmentation patterns were also observed with $dT_4N_4T_7$ samples where N = C or A. A statistical cleavage model was developed to describe the fragmentation patterns observed for these samples. Significantly, results with a second matrix, 3-HPA, showed little or no fragmentation with the same samples, showing that fragmentation is not only nucleobase dependent but also matrix dependent.

(12) Wu, K. J.; Steding, A.; Becker, C. H. *Rapid Commun. Mass Spectrom.* **1993**, *7*, 142–146.
(13) Wu, K. J.; Shaler, T. A.; Becker, C. H. *Anal. Chem.* **1994**, *66*, 1637–1645.
(14) Nordhoff, E.; Cramer, R.; Karas, M.; Hillenkamp, F.; Kirpekar, F.; Kristiansen, K.; Roepstorff, P. *Nucl. Acids Res.* **1993**, *21*, 3347–3357.

Experimental Section

Oligodeoxynucleotides $d(A)_n$, $d(C)_n$, $d(G)_n$, and $d(T)_n$ where $n = 2$, 4, 6, 8, and 10 were custom synthesized by the University of Wisconsin Biotechnology Center on an Applied Biosystems 394 DNA synthesizer. All other oligonucleotides including the following 15mers—dT_4GT_{10}, $dT_4G_2T_9$, $dT_4G_4T_7$, $dT_4G_8T_3$, dT_4AT_{10}, $dT_4A_4T_7$, dT_4CT_{10}, $dT_4C_4T_7$, and $dT_7G_4T_4$—were synthesized in our laboratory on an Applied Biosystems 391 DNA synthesizer. Oligodeoxynucleotide sequences are written in the 5′ to 3′ direction. All oligodeoxynucleotide samples were prepared with dimethoxytrityl (DMT) groups at their 5′ termini and further purified by passage over a C18 SEP-PAK cartridge (Waters, Milford, MA 01757) as described previously.[9] Concentrated stock solutions of each purified oligonucleotide were stored in Milli-Q water (18 Mohm) at −20 °C.

Saturated solutions of 2,5-DHBA and 3-HPA (Aldrich, Milwaukee, WI) were prepared daily. The 2,5-DHBA and 3-HPA matrix solutions were prepared respectively in 9:1 and 1:1 mixtures of water and acetonitrile.

The saturated matrix solutions and oligodeoxynucleotide solutions (1–10 μM) were treated overnight with a cation exchange resin (AG 50W-X8, 200–400 mesh, ammonium form; Bio-Rad Laboratories, Richmond, CA) to remove alkali cations.[15] Typically, 10–20 mg of exchange resin ws used to treat 20 μL of solution. MALDI samples were prepared by combining appropriate amounts of ion exchanged matrix and analyte solutions and depositing 2 μL on a solid sample probe tip followed by solvent evaporation at room temperature. Typically, 1–10 pmol of analyte was mixed with a 1000-fold molar excess of matrix on a 3.1 mm^2 probe tip.

All mass spectra were recorded using a Vestec Model VT 2000 laser desorption, linear time-of-flight mass spectrometer which has been described previously.[9] Briefly, samples were ablated/ionized using the focused output of a 355 nm frequency tripled Lumonics Model HY 400 Nd:YAG laser (Lumonics, Kanata, ON, Canada). The laser pulse width was 10 ns and the pulse frequency was 10 Hz. Ions were accelerated through a dual-stage source to a total potential of 30 keV and detected by a 20-stage focused-mesh electron multiplier (Becton-Dickinson Model MM1-1SG) at the end of the drift tube. All spectra shown were obtained in negative-ion mode and summed over 50 laser pulses. Time-to-mass conversion of parent ions was accomplished by calibration using the polydeoxythymidylic acids $d(T)_4$, $d(T)_{10}$, or $d(T)_{17}$ as internal standards, depending on the mass of the parent ion. After the parent ion mass was confirmed, the *m/z* assignments of fragment ions were determined by calibration using the parent ion and a matrix ion as internal standards.

Experiments were performed to examine the possibility of oligodeoxynucleotide degradation during sample preparation or storage. In particular, the possibility of base hydrolysis due to the acidic conditions of the matrix solutions (pH 1.7 for 2,5-DHBA and 3.8 for 3-HPA, determined using colorpHast indicator strips, EM Science, Cherry Hill, NJ) was examined. In these experiments a 50-pmol/μL sample of dT_4GT_{10} was analyzed by HPLC after treatment with a variety of conditions as outlined in Table 1. In each case, 1 μL of the oligonucleotide solution was combined with 5 μL of the matrix solution and the mixture was incubated at room temperature for the indicated period prior to HPLC analysis performed either directly or after drying and redissolution in 50 μL of water. Chromatography was performed under the following conditions: system, Shimadzu SIL-6A (Shimadzu Corporation, Kyoto Japan); column, C_{18} reversed-phase (5 μm × 25 cm, 86-203-C5, Rainin Instrument Co., Inc. Woburn, MA); solution A, 0.01 M triethylammonium acetate (TEAA) in 4% acetonitrile; solution B, 0.01 M TEAA in 40% acetonitrile; gradient, 0–5 min, constant 10% B; 5–50 min, linear ramp from 10% to 70% B; 50–52 min, linear ramp from 70% to 100% B; 52–57 min, constant 100% B; 57–59 min, linear ramp from 100% to 10%B; 59–64 min, constant 10% B; flow rate, 1 mL/min.

No detectable degradation was observed in samples 1–6 of Table 1; however, significant degradation was apparent after 2 days with 2,5-DHBA (sample 7). Accordingly, for all UV–MALDI experiments

(15) Nordhoff, E.; Ingendoh, A.; Cramer, R.; Overberg, A.; Stahl, B.; Karas, M.; Hillenkamp, F. *Rapid Commun. Mass Spectrom.* **1992**, *6*, 771–776.

Table 1. Sample Preparation for HPLC Experiments

		matrix vol, μL					injection	
sample no.	sample vol, μL	H_2O	0.01 M HCl	2,5-DHBA	3-HPA	incubation period	direct	dried and redissolved
1	1	5				30 min	√	
2	1		5			30 min	√	
3	1		5			30 min		√
4	1			5		30 min	√	
5	1			5		30 min		√
6	1				5	30 min		√
7	1			5		2 days	√	

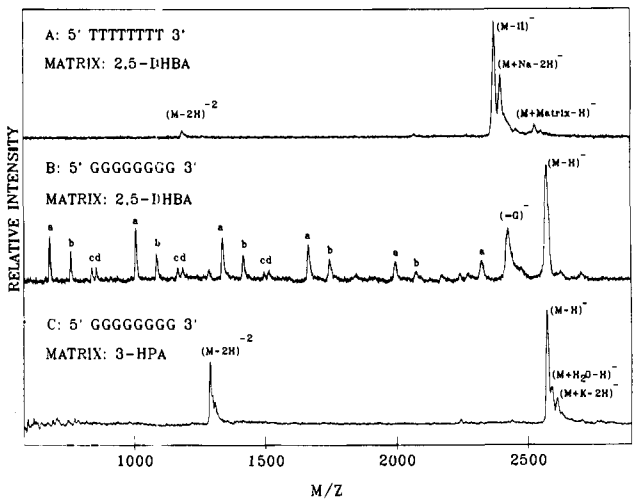


Figure 1. UV−MALDI mass spectra of $d(T)_8$ and $d(G)_8$ using 2,5-DHBA (panels A and B) and 3-HPA (panel C) as the matrix.

reported here samples were prepared fresh and analyzed immediately. The absence of sample degradation prior to UV−MALDI analysis is also confirmed from the results obtained with the matrix 3-HPA, which consistently produced spectra without degradation products (see Results).

Results and Discussion

The UV−MALDI behavior of a series of short homopolymer oligodeoxynucleotides was investigated initially. UV−MALDI analysis was performed on $d(A)_n$, $d(C)_n$, $d(G)_n$, and $d(T)_n$ (where $n = 2, 4, 6, 8$, and 10) using the matrices 2,5-DHBA and 3-HPA. Using 2,5-DHBA, little or no fragmentation of the dimer and tetramer occurred, whereas substantial fragmentation was observed from the longer $d(A)_n$, $d(C)_n$, and $d(G)_n$. In contrast, little or no fragmentation was observed in any of the $d(T)_n$ samples.

Figures 1A and 1B show spectra obtained from $d(T)_8$ and $d(G)_8$ using 2,5-DHBA. The mass spectrum of $d(T)_8$ exhibits a major peak corresponding to the singly charged deprotonated parent ion and a much smaller peak corresponding to the doubly charged species. A singly charged sodium salt peak $[M + Na - 2H]^-$ is also observed as well as a peak corresponding to the matrix adduct $[M + Matrix - H]^-$. The spectrum obtained from $d(G)_8$ under identical conditions is in marked contrast to the simple $d(T)_8$ spectrum. A complex pattern of peaks is observed containing four series of fragmentation peaks marked as **a**, **b**, **c**, and **d**. The members of each series are separated by the mass of a single deoxyguanosine nucleotide unit (329.2 daltons). Tentative assignments could be made to these peaks on the basis of their *m/z* ratios (see Table 2) with the aid of the fragmentation model shown in Figure 2. Three distinct cleavage processes were considered in this model: *N*-glycosidic bond cleavage (base loss); cleavage at the 3′ C−O bond of the phosphate backbone (site I); and cleavage at the 5′ C−O bond of the phosphate backbone (site II). In Figure 2 the products of *N*-glycosidic bond cleavage or 3′ C−O bond cleavage are shown on the left and the products of 5′ C−O bond cleavage on the right, and the products obtained in the case of two cleavage events are shown in the middle. Throughout this paper, the nomenclature for product ions was adopted from McLuckey[16] and McLafferty.[17] Each cleavage at site I generates two fragments: a (the 5′ end product) and w (the 3′ end product). Each cleavage at site II generates another two fragments: d (the 5′ end product) and z (the 3′ end product).

Table 2 shows the calculated and observed *m/z* ratios for these fragments. The four series of peaks **a**, **b**, **c**, and **d** evident in Figure 1B correspond to several different types of fragmentation products. Series **a** corresponds to w_n or d_m (containing respectively of a 3′ or 5′ phosphate group), generated by single cleavages at the phosphate backbone. Series **b** corresponds to two different possible fragmentation products, both the result of double fragmentation: these products are a_m-G or z_n-G, and $w_\alpha+P$, or $d_\beta+P$. Series **c**, w_n-G or d_m-G, also corresponds to the products of double cleavage. Series **d**, w_n-G-H_2O or d_m-G-H_2O, corresponds to the *m/z* expected from loss of a water molecule from series **c**; interestingly, peaks corresponding to a_m or z_n (second column of Table 2) are not observed. The reason for this observation will be explained below.

Although the model in Figure 2 provides much information about the nature of the fragmentation products of Figure 1B, it also leaves many issues unresolved. The fact that the products of cleavage at the 3′ C−O bond and the 5′ C−O bond have the same *m/z* ratios means that it is not possible to determine which bond is the site of cleavage. This is reflected in the structure of Table 2 which shows two possible products of the same *m/z* in each of the six columns. Also, the *m/z* difference between a_m-G and $w_\alpha+P$ (Table 2, **b** series) is only two mass units, which is unresolvable with the mass resolution and accuracy of these UV−MALDI experiments. This makes definitive product assignments impossible. In order to resolve these and other issues, a series of asymmetric oligodeoxynucleotides, dT_4GT_{10}, $dT_4G_2T_9$, $dT_4G_4T_7$, and $dT_4G_8T_3$, were synthesized and their fragmentation patterns analyzed (see below).

In sharp contrast to the UV−MALDI results obtained with 2,5-DHBA, no detectable fragmentation was observed for homopolymers analyzed with 3-HPA. This is illustrated in Figures 1B and 1C showing the results obtained in the analysis of $d(G)_8$ using these two different matrices. 3-HPA yields a relatively clean mass spectrum in which the singly and doubly deprotonated ions of $d(G)_8$ constitute the major peaks in the spectrum. Thus the fragmentation of oligodeoxynucleotides during MALDI analysis is not only sequence dependent but also matrix dependent. The greatly reduced fragmentation obtained from 3-HPA is presumably an important factor in its utility for the UV−MALDI analysis of oligonucleotides.

(16) McLuckey, S. A.; Berkel, G. J. V.; Glish, G. L. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 60−70.
(17) Little, D. P.; Chorush, R. A.; Speir, J. P.; Senko, M. W.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1994**, *116*, 4893−4897.

Table 2. Fragmentation Products of $d(G)_8$[a]

	single cleavage products	double cleavage products			
		a series	**b** series	**c** series	**d** series
nucleotide no.	a_m or z_n	w_n or d_m	a_m-G or z_n-G	$w_\alpha+P$ or $d_\beta+P$	w_n-G or d_m-G	w_n-G-H_2O or d_m-G-H_2O
7	(2224.7)	(2320.7)	(2074.6)	(2072.4)		
	not obsd	2322.7		2074.8		
6	(1895.4)	(1991.4)	(1745.3)	(1743.3)		
	not obsd	1994.4		1744.5		
5	(1566.2)	(1662.2)	(1416.1)	(1414.0)	(1512.1)	(1494.1)
	not obsd	1663.0		1415.5	1512.0	1494.0
4	(1237.0)	(1333.0)	(1086.9)	(1084.7)	(1182.9)	(1164.9)
	not obsd	1333.0		1084.5	1183.0	1164.0
3	(907.7)	(1003.7)	(757.6)	(755.5)	(853.6)	(835.6)
	not obsd	1004.8		754.3	853.8	835.7
2	(578.5)	(674.5)				
	not obsd	675.0				

[a] The numbers in parentheses are the calculated fragment *m/z* ratios; the numbers below are the experimentally observed fragment *m/z* ratios. The absence of an experimental value indicates that the fragment was not observed in the mass spectrum. The calculated values shown here correspond to cleavages at the indicated positions in Figure 2, less 1 amu to provide a negative charge.

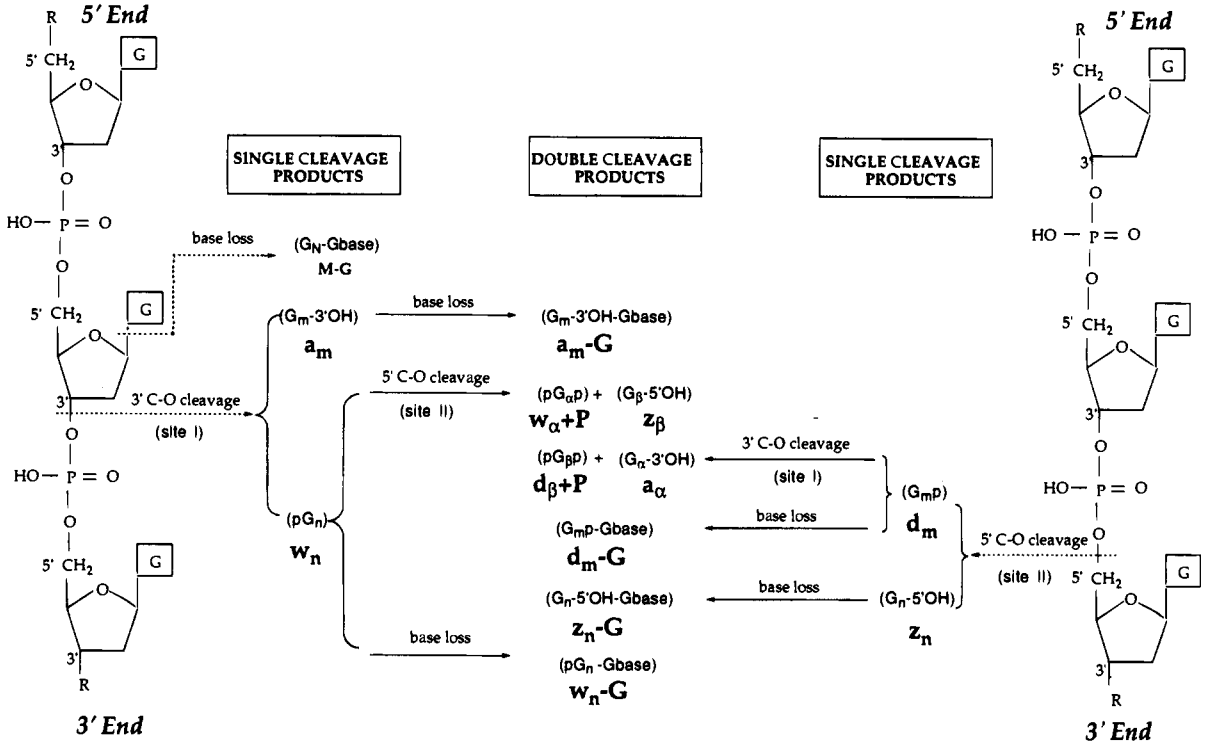


Figure 2. Fragmentation pathways for $d(G)_N$. Cleavage reactions considered are base loss, cleavage of the phosphate backbone at the 3′ C−O bond (site I), or cleavage of the phosphate backbone at the 5′ C−O (site II); the products of both single and double cleavage are shown. The letters *m*, *n*, α, and *β* represent integer numbers smaller than *N*; $N = m + n$; *n*, $m = \alpha + \beta$.

Figure 3 shows mass spectra obtained for the simplest of the asymmetric oligonucleotides, dT_4GT_{10}, using the matrices 2,5-DHBA and 3-HPA. As expected from the homopolymer studies described above, no substantial fragmentation was observed using 3-HPA. However, in 2,5-DHBA, significant peaks corresponding to loss of the guanine base as well as to cleavage of the phosphodiester backbone at the G site are evident. No cleavage was observed between thymidines, consistent with the homopolymer results. The HPLC analysis of dT_4GT_{10} described in the experimental section shows that the observed fragmentation products do not result from degradation during sample preparation but rather occur in the MALDI process.

Cleavage of the phosphodiester bond at the G site in this oligomer could occur in principle at either site I or site II as discussed above, resulting in different fragmentation products (see diagram at the top of Figure 3). Cleavage at site I would yield fragments corresponding to a_5 and w_{10}, whereas cleavage at site II would yield d_4 and z_{11}. The fact that the w_{10}, peak maximum is observed at *m/z* 3058.9 (calculated *m/z* 3059.3) but peaks corresponding to d_4 and z_{11} are not observed shows that cleavage is specific for site I. Interestingly, the complementary fragmentation product a_5 is not observed, just as the corresponding fragments of G_8 are not (Table 2, column 2). The implications of these observations for the fragmentation mechanism are discussed below.

Similar fragmentation patterns were also observed with $dT_4G_2T_9$, $dT_4G_4T_7$, and $dT_4G_8T_3$. A general trend in the mass spectra of these 15-mers was that as the number of deoxyguanosines in the oligomers increased, the intensities of the parent ion peaks decreased and the number of fragmentation peaks increased. This effect is illustrated in the spectrum of $dT_4G_4T_7$ shown in Figure 4. Predicted *m/z* ratios were calculated for fragments generated by cleavage at site I or site II of G nucleosides and are listed along with observed values in Table 3. Again, the major fragments observed were those corresponding to the 3′ end fragments from cleavage at one of the four

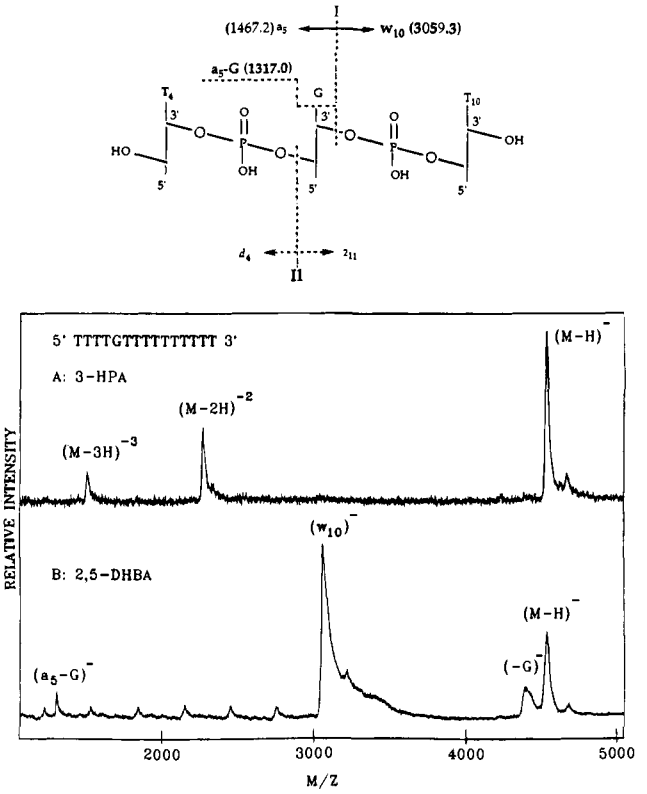


Figure 3. UV−MALDI mass spectra of dT_4GT_{10} using 3-HPA and 2,5-DHBA as matrices.

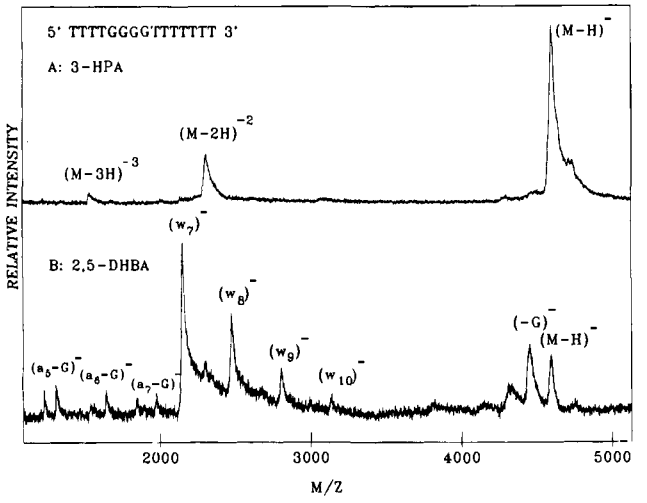
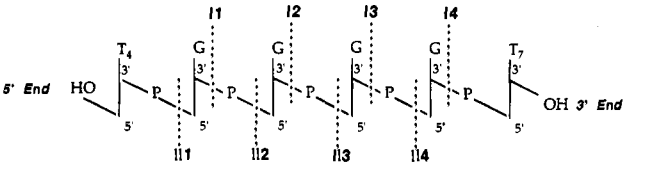


Figure 4. UV−MALDI mass spectra of $dT_4G_4T_7$ using 2,5-DHBA and 3-HPA as matrices.

possible site I's. A number of weaker peaks marked by (a_n−G) were also present; although they were not all assigned conclusively, several could be identified on the basis of their *m/z* ratio as 5′ end fragments lacking the G base. Thus, the overall behavior is similar to that for dT_4GT_{10} above.

The greater intensity of the 3′ w_{10} fragment raised the question of whether this was because the w_{10} fragment was longer or, alternatively, because it was at the 3′ end. This issue was examined by mass analysis of the oligomer $dT_7G_4T_4$, in which the longer dT_7 stretches are at the 5′ terminus rather than the 3′ terminus. A comparison of Figure 4B and Figure 5 shows a similar pattern of peaks with, however, two significant differences. First, the lower *m/z* peaks corresponding to the 5′ end fragments (a_{5-7}−G) for Figure 4B and to the 3′ end fragments (w_{4-6}) for Figure 5 are much more intense for the latter. This is consistent with the mechanism discussed below in which a

Table 3. Fragmentation Products of $dT_4G_4T_7$[a]



	3′ C−O cleavage (site I)			5′ C−O cleavage (site II)	
cleavage site	5′ end fragments		3′ end fragments	5′ end fragments	3′ end fragments
1	a_5−G	a_5	w_{10}	d_4	z_{11}
	(1316.1)	(1466.2)	(3133.3)	(1232.9)	(3366.6)
	1316.4		3139.2		
2	a_6−G	a_6	w_9	d_5	z_{10}
	(1645.3)	(1795.4)	(2804.1)	(1562.2)	(3037.3)
	1645.4		2808.4		
3	a_7−G	a_7	w_8	d_6	z_9
	(1974.5)	(2124.6)	(2474.8)	(1891.4)	(2708.1)
	1974.2		2477.0		
4	a_8−G	a_8	w_7	d_7	z_8
	(2303.7)	(2453.8)	(2145.6)	(2220.6)	(2378.9)
			2148.2		

[a] The numbers in parentheses are the calculated fragment *m/z* ratios; the numbers below are the experimentally observed fragment *m/z* ratios. The absence of an experimental value indicates that the fragment was not observed in the mass spectrum. The calculated values shown here correspond to cleavages in the above diagram, less 1 amu to provide a negative charge.

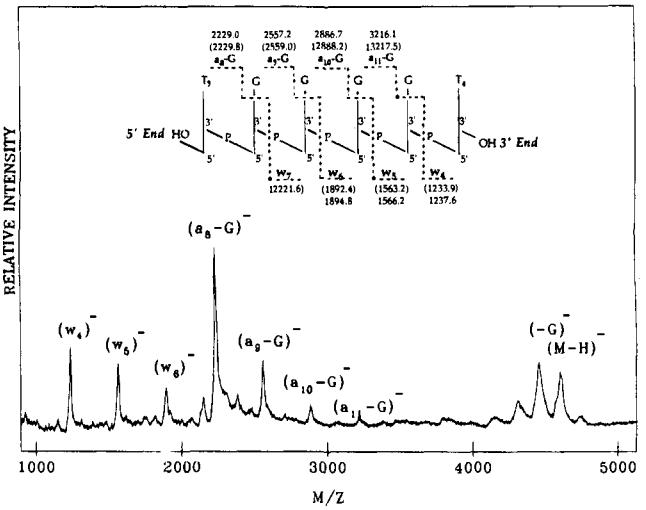


Figure 5. UV−MALDI mass spectrum of $dT_7G_4T_4$ using 2,5-DHBA as the matrix.

negative charge is generated on the 3′ end fragmentation products. Second, the higher *m/z* peaks corresponding to the 3′ end fragments (w_{7-10}) for Figure 4B and to the 5′ end fragments (a_{8-11}−G) for Figure 5 are comparable in size relative to the parent peaks. This suggests an additional effect is operative, increasing the intensity of these longer 5′ end fragments. The simplest explanation for this is that the probability of a fragment possessing a negative charge is proportional to its length, as longer molecules possess more possible sites for charge formation.

The fragmentation behavior of dT_4AT_{10} and dT_4CT_{10} was also investigated. UV−MALDI spectra of these samples using 2,5-DHBA (not shown) were similar to those obtained from dT_4GT_{10}, showing major peaks corresponding to w_{10} and $[M-A]^-$ or $[M-C]^-$ in addition to the parent ion peak. Figure

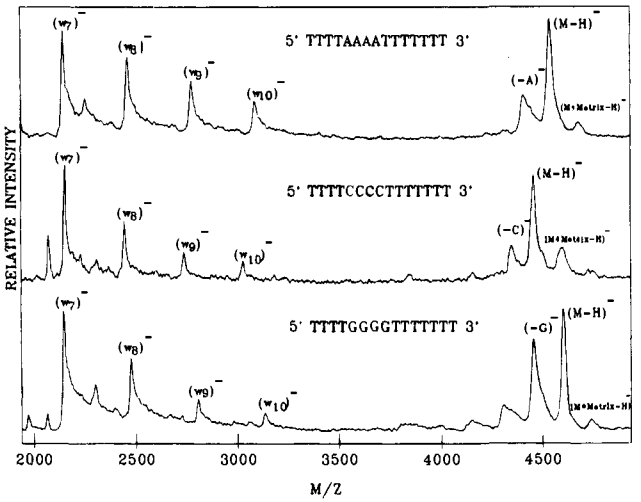


Figure 6. UV−MALDI mass spectra of $dT_4N_4T_7$ (N = G, C, and A) using 2,5-DHBA as the matrix.

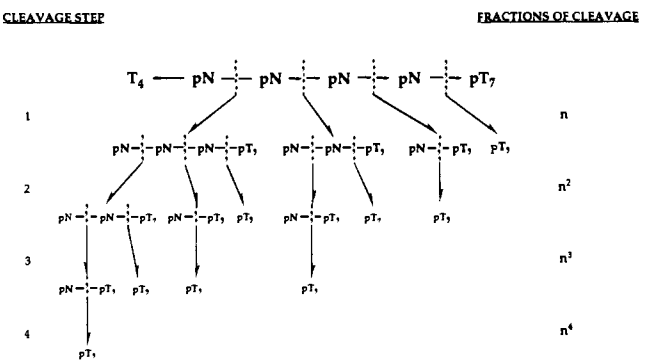


Figure 7. Fragmentation model for $dT_4N_4T_7$.

6 displays the mass spectra obtained from $dT_4A_4T_7$ and $dT_4C_4T_7$ which show the same fragmentation pattern as obtained from $dT_4G_4T_7$.

The magnitude of fragment and parent ion peaks varies in different regions of the sample and with laser power. However, the *relative* heights of the four 3′ fragment peaks shown in Figure 6 were found to be very reproducible, in order $pT_7(w_7)$ > $pNT_7(w_8)$ > $pN_2T_7(w_9)$ > $pN_3T_7(w_{10})$. This reproducibility stimulated our interest in attempting to develop a simple statistical model to explain the fragmentation pattern of these $dT_4N_4T_7$ oligonucleotides.

Statistical Cleavage Model. The hypothesized fragmentation scheme is diagrammed in Figure 7. In this model, it is assumed that the fragment ions observed can be produced in up to four serial cleavages occurring in the source region of the spectrometer. The fraction of 3′ C−O bonds cleaved at a given position is defined as n, a parameter to be determined from the data. n is assumed to be identical for each cleavage site irrespective of whether it is an initial or secondary cleavage event. The model only considers site I cleavage at the internal A, C, or G nucleotides.

In the model, the final amount of a given 3′ end fragment observed is the sum of that generated at each cleavage stage, less the amount of the fragment consumed in subsequent stages. The four fragments differ in the number of different cleavage paths leading to their production. The fragment pN_3T_7 can only be generated in a single fragmentation reaction, the fragment pN_2T_7 in two different stages (from the parent as well as from pN_3T_7), the fragment pNT_7 in three stages, and so on. This may be quantified as follows: let P = the initial amount of the parent ion $dT_4N_4T_7$ formed, and n = the fraction of sites cleaved at a given position. Then summing the production and consumption of the different fragments leads to the following expressions:

$$pN{-}pN{-}pN{-}pT_7 = Pn(1 - 3n) \quad (1)$$

$$pN{-}pN{-}pT_7 = Pn(1 - n - 2n^2) \quad (2)$$

$$pN{-}pT_7 = Pn(1 + n - n^2 - n^3) \quad (3)$$

$$pT_7 = Pn(1 + 3n + 3n^2 + n^3) \quad (4)$$

As mentioned above, the relative intensity of the parent ion peak varies considerably from spectrum to spectrum. The data were therefore normalized to the pT_7 fragment peak height, yielding the following equations.

$$pN{-}pN{-}pN{-}pT_7/pT_7 = (1 - 3n)/(1 + 3n + 3n^2 + n^3) \quad (5)$$

$$pN{-}pN{-}pT_7/pT_7 = (1 - n - 2n^2)/(1 + 3n + 3n^2 + n^3) \quad (6)$$

$$pN{-}pT_7/pT_7 = (1 + n - n^2 - n^3)/(1 + 3n + 3n^2 + n^3) \quad (7)$$

The only unknown in these equations is n. To evaluate n, it is desirable to have statistically reliable values for the measured ratios of the fragment peaks. Accordingly, multiple determinations (between 39 and 46) of the peak ratios were made for each of the three oligomers studied (Table 4). The spectral peaks are notably asymmetric and broad, with tails on the higher mass side. The peak maxima, however, correspond closely to calculated m/z values. These features indicate that both prompt fragmentation and delayed fragmentation in the source region[18] contribute significantly to the observed peak shapes. In view of this, peak intensities were quantified by measurement of peak heights rather than peak areas: this tends to emphasize the prompt fragmentation component of the peak rather than the more delayed fragmentation. Values for n were determined independently for each of the three fragment ratios by solving eqs 5−7. Thus, solving eq 5 yields the n value corresponding to the experimentally determined ratio of pN_3T_7/pT_7; solving eq 6 yields the n value corresponding to pN_2T_7/pT_7; and solving eq 7 yields the n value corresponding to pNT_7/pT_7. Each of eqs 5−7 is cubic, yielding one real root and two complex roots; the solution was taken to be the real root. These n values were averaged separately for each of the three ratios, with the results shown in column 4 of Table 4. The n values are in reasonably good agreement, supporting the validity of the fragmentation model employed. On this basis all the n values were pooled and are shown along with the pooled standard deviations in columns 6 and 7 of Table 5. The n values determined in this manner are 0.13 for dA, 0.26 for dC, and 0.27 for dG. Thus dC and dG have essentially the same frequency of fragmentation, whereas dA is roughly 2-fold more stable.

The assumption that n is the same for each cleavage step was examined in some experiments by analysis of the fragmentation patterns obtained from $dT_4G_2T_7$ and $dT_4G_3T_7$ (data not shown). In these simpler cases it is possible to independently determine the values of n_2 and n_3 (cleavage frequencies for the second and third cleavage steps, respectively). The values of n_2 and n_3 obtained in this way were equal and approximately the same as the value for n obtained as described

(18) Hercules, D. M.; Day, R. J.; Balasanmugam, K.; Dang, T. A.; Li, C. P. *Anal. Chem.* **1982**, *54*, 280A−305A.

Table 4. Fragmentation Results from $dT_4N_4T_7$ Samples

fragment ratios	no. of samples	averaged ratios	avg fraction of cleavage sites (n)	stand dev	final averaged n values	final stand dev
pAT_7/pT_7	42	0.771	0.132	0.0789		
pA_2T_7/pT_7	42	0.564	0.137	0.0446	0.133	0.0567
pA_3T_7/pT_7	42	0.434	0.129	0.0394		
pCT_7/pT_7	39	0.573	0.270	0.0484		
pC_2T_7/pT_7	39	0.285	0.269	0.0370	0.260	0.0400
pC_3T_7/pT_7	39	0.146	0.240	0.0242		
pGT_7/pT_7	46	0.592	0.255	0.0480		
pG_2T_7/pT_7	46	0.240	0.297	0.0312	0.271	0.0401
pG_3T_7/pT_7	46	0.108	0.261	0.0234		

Table 5. Calculation Results of Energy Terms

	energy terms	
dN	$[\Delta H_f(BH) - \Delta H_f(SBH^+)]^a$ (kcal/mol)	E_{taut} (kcal/mol)
A	3.0	6.7,[b,c] 19.6[d]
C	1.8	0.5[b]
G	2.8	1.5[b]

[a] $[\Delta H_f(BH) - \Delta H_f(SBH^+)] = \Delta H_f(BH) - \Delta H_f(SB) - PA(SB) - \Delta H_f(H^+)$, heats of formation for both bases and nucleosides were determined using geometry optimized AM1 calculations, input geometries are based on X-ray data.[32-34] $\Delta H_f(H^+) = 367.2$ kcal/mol.[35] Proton affinities are given in ref 28. [b] Values shown are taken from ref 31 and are identical with those calculated in the present work. [c] N7 protonation. [d] N1 protonation.

above. Therefore the simpler approach of using only a single n value was chosen for this work. The similar values obtained for n in each stage of the cleavage process are reasonable given the prompt nature of the fragmentation under study; such prompt fragmentation necessarily occurs immediately during or subsequent to the laser pulse and as such all of the serial cleavage events are occurring at similar times and in similar regions of the source.

It should also be noted that although this model and the resultant fragmentation frequencies describe the observed fragmentation pattern well for these $dT_4N_4T_7$ oligonucleotides, when the resultant n values are used to predict fragmentation patterns for other mixed sequence oligomers, good agreement with prediction is not obtained (data not shown). This suggests that local sequence context affects cleavage; further work will be required to clarify this issue.

Fragmentation Pathway. The fragmentation patterns observed in these UV–MALDI experiments using the matrix 2,5-DHBA provide important clues as to the fragmentation pathway. Major spectral peaks generally include the following: (a) the parent ion; (b) the parent ion from which a base (A, C, or G) has been lost; (c) the intact 3′ end product of phosphodiester bond cleavage at site I; and (d) the 5′ end product of phosphodiester bond cleavage at site I missing a base (A, C, or G). These data suggest that the oligodeoxynucleotide fragmentation pathway involves two cleavage events: base loss and phosphate backbone cleavage. Significantly, the spectra include neither peaks corresponding to the 3′ end fragmentation products missing a base nor the 5′ end fragmentation products containing all the bases. If base loss occurred after fragmentation, one would presumably see peaks corresponding to both the 5′ and 3′ end products from which a base had been lost. The absence of the 3′ end fragmentation product lacking a base thus suggests that base loss precedes subsequent cleavage. Conversely, if base loss was not required for fragmentation to occur, one presumably would see 5′ end fragmentation products containing all the bases. The absence of such products also argues that base loss must precede the fragmentation. This hypothesis is consistent with the ubiquitous presence in the spectra of peaks corresponding to the parent ion lacking a base. A likely initiation step for base loss is protonation of the base during laser ablation. Base protonation in aqueous solutions is known to withdraw electron density from the *N*-glycosidic bond catalyzing base loss.[19] Such base loss facilitates cleavage of the phosphodiester backbone[20,21] at the 3′ C–O bond (site I), generating a 5′ end fragment lacking a base and a 3′ end fragment containing a base. Chemical ionization mass spectral studies of a number of nucleosides and nucleoside analogs by McCloskey *et al.* also demonstrated *N*-glycosidic bond cleavage subsequent to base protonation.[22]

The hypothesis that base protonation initiates fragmentation also provides a possible explanation for why different matrices show varying amounts of fragmentation. The distribution of excess protons between matrix and analyte will depend on the relative proton affinities of the two species.[23,24] The proton affinities of 2,5-DHBA and 3-HPA have not been experimentally measured. However, they can be estimated from the calculated heats of formation of the geometry optimized neutral and protonated molecules using Austin Model 1 (AM1) calculations[25] followed by single-point Hartree–Fock calculations with 3-21G basis sets.[26] The proton affinities of 2,5-DHBA and 3-HPA calculated in this manner are 213 and 230 kcal/mol, respectively. Comparing these with the known nucleoside proton affinities[22,27,28] (the most recent values (in kcal/mol) from ref 28 are dG = 234.4, dA = 233.6, dC = 233.2, and dT = 224.9), 2,5-DHBA's proton affinity is substantially lower, whereas 3HPA's is similar. The lower proton affinity of 2,5-DHBA makes it a more effective proton donor to the nucleobases, facilitating nucleobase protonation and subsequent backbone fragmentation. This provides an interesting possible explanation for the observed differences in fragmentation for different matrices. Experimental measurements of the proton affinities of a series of matrices varying in their propensity to induce fragmentation could help to confirm this hypothesis.

The proposed fragmentation pathway is diagrammed in Figure 8 and is similar to the fragmentation pathway proposed by McLuckey and co-workers.[16,29] Using electrospray and collisional activation conditions, they found that multiply charged single-stranded oligonucleotide anions fragment by loss of a base, followed by cleavage at the 3′ C–O bond of the sugar from which the base is lost. They proposed that both steps proceed via 1,2-elimination to yield a stable substituted furan on the 5′ end products. The formation of a stable substituted furan compound is consistent with the fragment m/z ratios observed here (see 5′ end products a_n–G in Figures 4 and 5). However, the mass accuracy of these spectra is not sufficient

(19) Kochetkow, H. K.; Budovskii, E. I. *Organic Chemistry of Nucleic Acids*; Plenum Publishing Company Ltd.: London–New York, 1972.
(20) Brown, D. M.; Todd, A. R. *Nucleic Acids*; Academic Press: New York–London, 1955; Vol. 1.
(21) Michelson, A. M. *The Chemistry of Nucleosides and Nucleotides*; Academic Press: London–New York, 1963.
(22) Wilson, M. S.; McCloskey, J. A. *J. Am. Chem. Soc.* **1975**, *97*, 3436–3444.
(23) Cooks, R. G.; Kruger, J. L. *J. Am. Chem. Soc.* **1977**, *99*, 1279–1230.
(24) McLuckey, S. A.; Cameron, D.; Cooks, R. G. *J. Am. Chem. Soc.* **1981**, *103*, 1313–1317.
(25) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902–2903.
(26) Frisch, M.; Foresman, J.; Frisch, A. *Gaussian 92* User's *Guide*; Gaussian, Inc.: Pittsburgh, PA, 1992.
(27) Greco, F.; Liguori, A.; Sindona, G.; Uccella, N. *J. Am. Chem. Soc.* **1990**, *112*, 9092–9096.
(28) Liguori, A.; Napoli, A.; Sindona, G. *Rapid Commun. Mass Spectrom.* **1994**, *8*, 89–93.
(29) McLuckey, S. A.; Habibi-Goudarzi, H. *J. Am. Chem. Soc.* **1993**, *115*, 12085–12095.

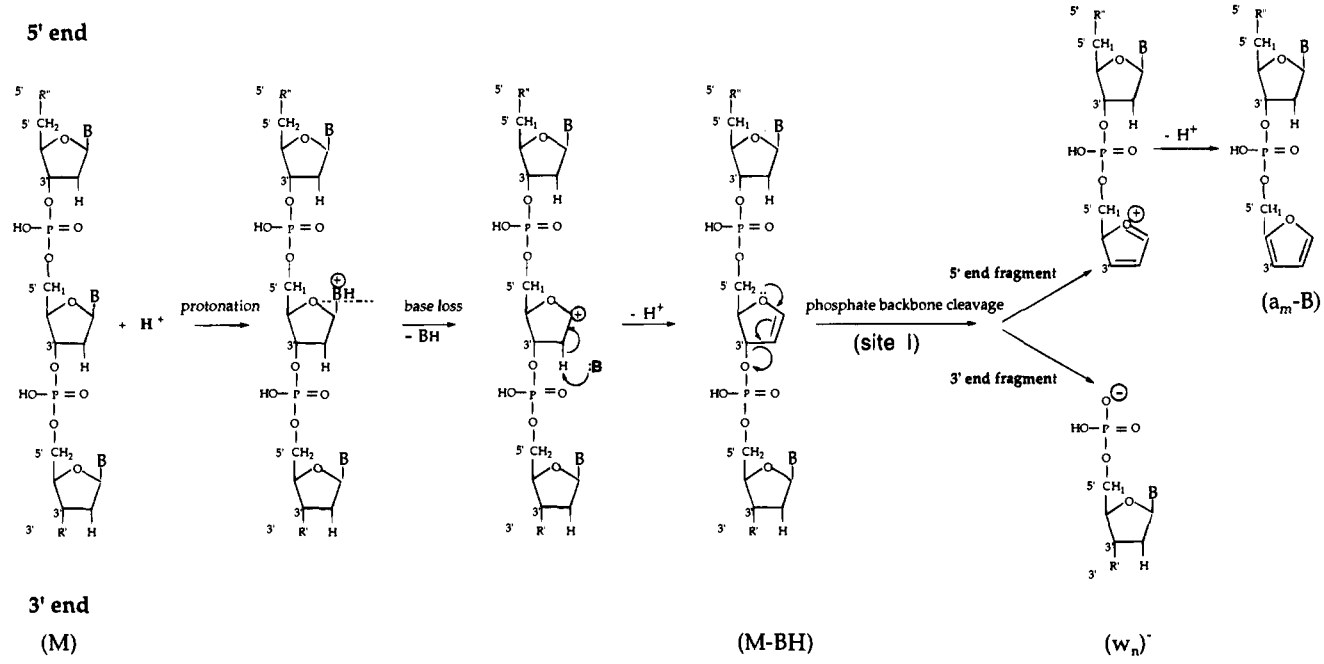


Figure 8. Proposed oligodeoxynucleotide fragmentation pathway.

to prove this conclusively. The putative furan products have masses 2 u lower than the calculated values for the 5′ end products in Tables 2 and 3 and Figure 5. Also in agreement with our results, Hettich and Buchanan determined in collision-induced dissociation experiments employing UV–MALDI and Fourier transform mass spectrometry (266-nm nicotinic acid matrix) that the primary fragmentation pathway involved phosphodiester bond cleavage at site I.[30]

Species observed in the mass spectra are singly-charged negative ions corresponding to the parent $(M-H)^-$, the parent lacking a base $(M-BH)^-$, the 5′ end product $(a_m-B)^-$, and the 3′ end product w_n^-. The location of the negative charge on these ions is only shown in Figure 8 for w_n^-; although the nature and location of the negative charge on the other species is not known conclusively, one likely possibility is that it is produced by loss of a proton on one of the backbone phosphate groups.[14,15,30]

Armed with this improved understanding of fragmentation, the four series of fragmentation peaks observed in the analysis of the G_8 homopolymer (Figure 2) may be assigned with greater confidence. The **a** series corresponds to the intact 3′ end products (w_n in Table 2) and the **b** series to the 5′ end products lacking a base (a_m-G in Table 2). The **b** series are presumed not to correspond to $w_\alpha+P$ or $d_\beta+P$ because the production of these fragments requires 5′ C–O bond (site II) cleavage, an event which was not observed in the $dT_4N_4T_7$ samples. The **c** series corresponds to the fragments w_n-G, which are formed from the **a** series after base loss, and the **d** series are obtained from the **c** series upon loss of a water molecule.

It is of interest to determine if the cleavage mechanism proposed above can provide insight into the different cleavage frequencies observed for the four bases A, C, G, and T (0.13, 0.26, 0.27, and ~0, respectively). The apparent requirement for base loss to occur prior to phosphate backbone cleavage suggests that the differential cleavage may be determined by the base loss step. As shown in Figure 8, base loss proceeds in two sequential steps: base protonation, followed by *N*-glycosidic bond cleavage releasing the neutral base and leaving a positive charge on the ribose ring. As discussed above, the

(30) Hettich, R.; Buchanan, M. *J. Am. Soc. Mass Spectrom.* **1991**, *2*, 402–412.

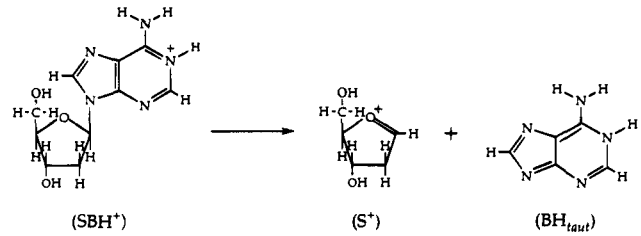


Figure 9. Depurination reaction of protonated nucleoside dA.

first step of base protonation depends upon the proton affinities of the nucleobases. The proton affinities given above for adenine, cytosine, and guanine are very similar, whereas the proton affinity of thymine is significantly lower. The lower proton affinity of dT relative to the other three bases provides an adequate explanation for its much greater resistance to fragmentation; it does not protonate and hence cleavage does not occur. However, the similar proton affinities of the three other bases does not correlate with their relative fragmentation frequencies (G ≈ C > A). A possible explanation may come, however, from consideration of the tautomeric forms of the base. An example of this reaction is shown in Figure 9 for the protonated nucleoside dA. The enthalpy of this reaction (ΔH_{reac}) is related to the heats of formation (ΔH_f) of reactants and products as given in eq 8.

$$\Delta H_{reac} = \Delta H_f(S^+) + \Delta H_f(BH_{taut}) - \Delta H_f(SBH^+) \quad (8)$$

The energy of tautomerization, E_{taut}, is the difference in energy between a tautomeric form of the base and the lowest energy conformation.

$$E_{taut} = \Delta H_f(BH_{taut}) - \Delta H_f(BH) \quad (9)$$

Combining eqs 8 and 9 yields

$$\Delta H_{reac} = \Delta H_f(S^+) + E_{taut} + [\Delta H_f(BH) - \Delta H_f(SBH^+)] \quad (10)$$

For all bases, $\Delta H_f(S^+)$ will be the same. Thus, the difference in the enthalpy of reaction for G, C, and A will depend only on the energies of tautomerization (E_{taut}) and the differences in

enthalpies of formation between the bases and the protonated nucleosides ($[\Delta H_f(BH) - \Delta H_f(SBH^+)]$). Using the proton affinities measured by Liguori *et al.*[28] and geometry optimized AM1 calculations to determine the heat of formation of the bases and nucleosides,[25] we calculated $[\Delta H_f(BH) - \Delta H_f(SBH^+)]$ and found them to vary by no more than 1−2 kcal/mol for G, C, and A (Table 5, column 2).

Rogers *et al.*[31] have determined the values of E_{taut} for G, C, and A (Table 5, column 3). They found that whereas both guanine and cytosine have tautomeric forms corresponding to the favored protonation sites which are within 2 kcal/mol of the lowest energy conformation, adenine's lowest energy tautomer is 6.7 kcal/mol higher in energy than adenine's ground state. However, this conformation does not correspond to the preferred N1 protonation site in adenine. Using AM1 calculations, we determined the heat of tautomerization for that conformer to be 19.6 kcal/mol. Thus, the much greater value of E_{taut} for adenine, reflecting the higher energy of adenine as a leaving group, is probably responsible for the lower propensity of fragmentation at dA sites than at dC or dG sites.

Conclusions

Fragmentation of oligodeoxynucleotides in UV−MALDI using 355-nm radiation is shown to be both nucleobase dependent and matrix dependent. The nature of the products indicates that base protonation is the initial step in fragmentation, followed by cleavage of the *N*-glycosidic bond resulting in base loss, which in turn leads to phosphodiester backbone cleavage at the 3′ C−O bond of adenosine, cytidine, and guanosine nucleotides. A simple statistical model describes the observed fragmentation patterns well, permitting the frequencies of bond cleavage to be determined by using 2,5-DHBA; these frequencies were 0.13, 0.26, and 0.27 for adenosine, cytidine, and guanosine, respectively. No cleavage was observed at thymidine. The lack of cleavage at thymidine is attributed to its relatively low proton affinity; the lower propensity for fragmentation at adenine is attributed to the relatively high energy of the adenine tautomer generated by base loss.

Acknowledgment. The authors are grateful to Professors Jim Skinner and Frank Weinhold for helpful advice. This work was supported by Department of Energy Human Genome Program Grant No. DE-FG02-91 ER61130 and NSF Grant No. CHE-9007850.

JA942002N

(31) Rodgers, M. T.; Campbell, S.; Marzluff, E. M.; Beauchamp, J. L. *Int. J. Mass. Spectrom. Ion Processes* (in press).

(32) Klooster, W. T.; Ruble, J. R.; Craven, B. M. *Acta Crystallogr.* **1991**, *B47*, 376−383.

(33) Post, M. L.; Birnbaum, G. I.; Huber, C. P.; Shugar, D. *Biochim. Biophys. Acta* **1977**, *479*, 133−142.

(34) Thewalt, U.; Bugg, C. E.; Marsh, R. E. *Acta Crystallogr.* **1970**, *B26*, 1089−1101.

(35) Rosenstock, H. M.; Draxl, K.; Steiner, B. W.; Herron, J. T. *J. Phys. Chem. Ref. Data* **1977**, Suppl. 1.