A New Method for Sequencing Fully Protected Oligonucleotides Using ^{252}Cf-Plasma Desorption Mass Spectrometry. 2. Negative Ions of Subunits in the Stepwise Synthesis of a Heptaribonucleotide

Catherine J. McNeal,*,† Kelvin K. Ogilvie,‡ Nicole Y. Theriault,‡ and Mona J. Nemer‡

Contribution from the Departments of Chemistry, Texas A&M University, College Station, Texas 77843, and McGill University, Montreal, Quebec, Canada H3A 2K6.
Received May 18, 1981

Abstract: A new approach to the sequence analysis of oligonucleotides containing unnatural (e.g., phosphotriester, phosphite, etc.) internucleotidic links has been developed based upon californium-252 plasma desorption mass spectrometry. Oligonucleotides which are not amenable to analysis by chemical or enzymatic degredation exhibit a highly specific negative ion fragmentation pattern containing diagnostic sequence ions which resemble the summed products of two enzymatic digests, employing 5′- and 3′-exonucleases. The method is rapid, reliable, and virtually nondestructive.

In the first paper of this series, we introduced the concept that a complete analysis of simple synthetic gene fragments in the chemically-protected stage could be obtained from the ^{252}Cf-plasma desorption positive and negative ion mass spectra. From an analysis of the negative ion spectra of a series of dinucleoside monophosphates we identified the dominant fragmentation pathways and showed how the base sequence could be verified.

In this paper, we extend the study to the level of a heptanucleoside hexaphosphate and demonstrate that despite the complexities of these large structures the systematic fragmentation patterns observed for the dinucleoside monophosphate are also applicable to the higher-order oligonucleotides.

Experimental Details

The methods of measurement by ^{252}Cf-PDMS, sample preparation, and synthetic procedures are described in the first paper of this series.

Results and Discussion

Reliability of Mass Determination. The molecular weights of the molecules studied in this work extend beyond 4000 u. The fragment ions containing the sequence information are in the m/z 500–4000 range. In order to obtain reliable sequence information it is essential that accurate mass values are obtained in the analysis. The problem is compounded by the low mass resolution of the time-of-flight (TOF) mass spectrometer ($\sim$450 $M/\Delta M$ in this study). A reliable mass calibration above m/z 1000 is a problem with all mass spectrometers because of the lack of suitable mass standards. For the ^{252}Cf-PDMS method, the TOF is measured digitally and with high precision. The centroid of a mass peak can be measured in time to 1 part in 10^4 and the relationship between TOF and $m^{1/2}/z$ is linear. If the TOF of two ions in the mass spectrum (e.g., H^-, C_2H^-) are measured precisely and simultaneously the entire mass range of interest is recorded, a mass calibration curve unique to that measurement can be derived which is valid over the entire mass range. The m/z of a particular ion is calculated from the *centroid* of the time distribution for that ion in the TOF spectrum. If the peak is cleanly resolved from adjacent masses, the centroid gives a mass which is very close to the calculated principal ion mass (to within 1 part in 2500). Between m/z 450 and 1000, the isotopic satellites are not cleanly resolved and the time centroid is shifted to a higher value resulting in a measured mass that is between the principal ion mass and the isotopically averaged mass. Above m/z 1000, the isotopic satellites are convoluted into a broad time distribution whose centroid corresponds to the isotopically averaged m/z value.

The low resolution in ^{252}Cf-PDMS is primarily due to the initial kinetic energy distribution of the ions as they leave the surface of the sample. There are other factors that can also contribute to the broadening. Metastable ions that decay during the acceleration stage or in the field-free drift region of the spectrometer can also contribute to the broadening. An additional contribution is from delayed ion emission which gives rise to a tail on the high mass side of a peak. This effect is particularly strong for the alkali metal ions but also is a significant component of the PO_3^- ion peak. As a consequence of all of these effects, the spectra presented in this paper contain a collection of peaks with varying shapes which depend upon the dynamic temporal properties of these ions. Fortunately, these effects do not noticeably perturb the location of the centroids, so that reliable masses can be obtained, even for the metastable ions. Listed in Table I are the experimental masses for the ions shown in Figure 1 together with a comparison of the values of the principal ion and isotopically averaged ion masses. The transition from principal to isotopically averaged mass is clearly discernable, and the overall good agreement between measured and calculated masses attests to the validity of our method for mass determination.

Table I. Comparison of Experimental and Calculated Masses for the Proposed Fragment Ions in Figure 1

m/z(exptl)[a]	average m/z(calcd)	principal ion[b] m/z(calcd)	proposed structure[b]
574.41	573.74	574.22	(a − R′ + H)$^-$
634.50	635.24	634.20	(a − 2Cl)$^-$
670.32	670.69	669.17	(a − Cl)$^-$
706.07	706.14	704.14	a$^-$
883.57			
909.12	908.86	907.22	(b − Cl)$^-$
944.41	944.31	942.19	b$^-$
1080.10			
1229.23	1229.70	1227.29	(c − R′ + H)$^-$
1251.32	1251.68	1249.27	(c − R′ + Na)$^-$
1290.34	1289.18	1286.26	(c − Cl − HCl)$^-$
1326.09	1325.64	1322.23	(c − Cl)$^-$
1360.60	1361.09	1357.20	c$^-$
1564.28	1563.80	1560.29	(d − Cl)$^-$
1598.67	1599.25	1595.26	d$^-$
1621.22	1621.23	1617.24	(d + Na − H)$^-$
1910.74	1909.17	1906.63	(M − R′ − HCl)$^-$
1945.76	1945.63	1942.60	(M − R′)$^-$
2007.68	2006.11	2002.58	(M − Cl − HCl)$^-$
2043.16	2042.57	2038.56	(M − Cl)$^-$
	2078.02	2072.52	M$^-$
2077.73	2077.01	2073.53	(M − H)$^-$
2113.09	2112.46	2108.50	(M + Cl)$^-$

[a] The error associated with each mass is, on the average, one part in 2500 for masses above 1000 m/z and below 400 m/z. [b] Structures of fragments a, b, c, and d are indicated in Figure 2.

Subunits in the Stepwise Synthesis of a Heptaribonucleotide. Spectra of a series of fully protected ribooligonucleotides extending to a heptanucleotide have been obtained. These cognates were obtained at each stage in the stepwise synthesis of the 3′-terminal

† Texas A&M.
‡ McGill University.

0002-7863/82/1504-0976$01.25/0

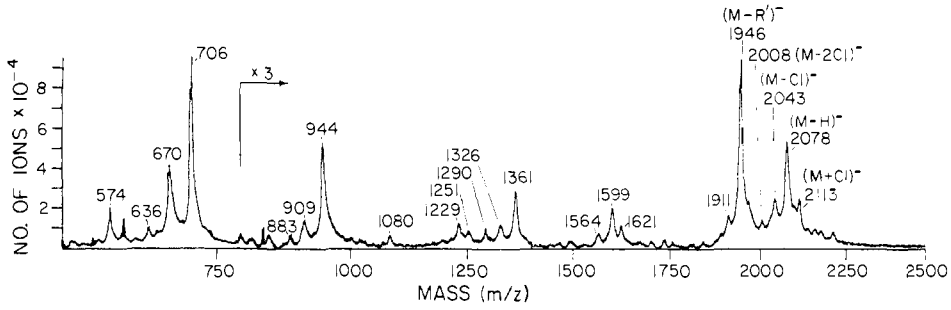

Figure 1. ^{252}Cf-PD negative ion spectrum of MTr(bz)C_p^{Si}(TCE)bzC_p^{Si}(TCE)A_{Si}^{Si} above m/z 500. Experimental masses appear above the peaks. Data were accumulated for 3.6×10^4 s. The ordinate represents the number of ions detected in each 4 ns wide time bin.

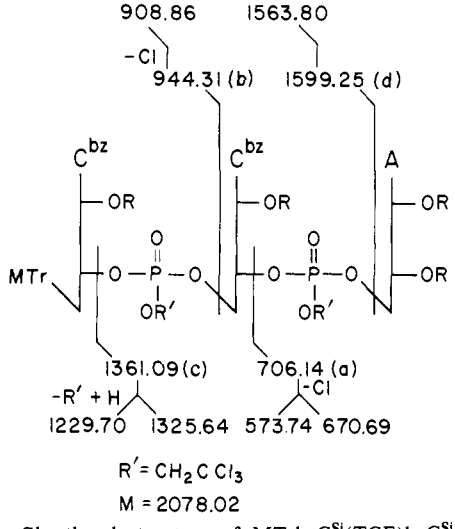

Figure 2. Shorthand structure of MTrbzC_p^{Si}(TCE)bzC_p^{Si}(TCE)A_{Si}^{Si}. Calculated average masses of fragments a, b, c, and d are indicated by each cleavage site. Calculated average masses of fragments formed by Cl loss from the sequence ions appear above (in band d) or below and to the right side (c and a) of these values. Calculated average masses of fragments formed by the substitution of hydrogen for R′ in fragments c and a and the molecular weight are also indicated. MTr = monomethoxytrityl, C^{bz} = benzoylcytidine, A = adenine, R = TBDMS.

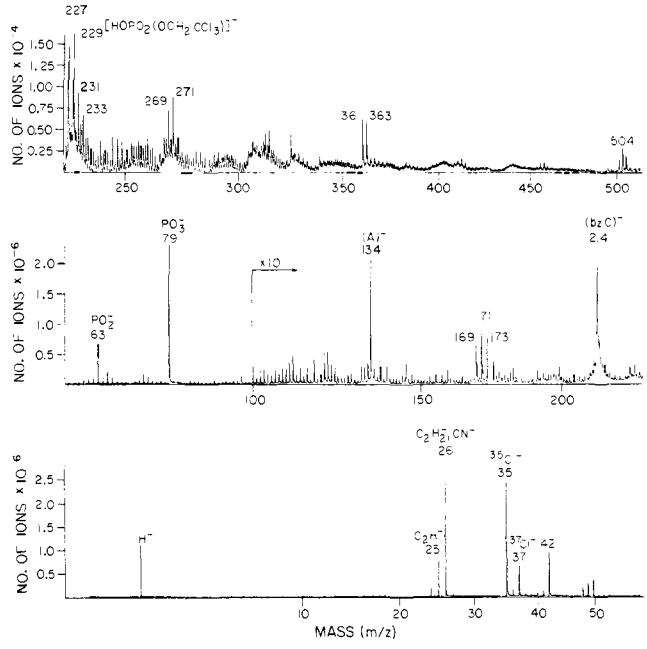

Figure 3. ^{252}Cf-PD negative ion spectrum of MTr(bz)C_p^{Si}(TCE)bzC_p^{Si}(TCE)A_{Si}^{Si} below m/z 500. Experimental masses are indicated above the peaks. Data were accumulated for 3.6×10^4 s and are plotted in 1 ns wide time bins.

heptanucleotide sequence of fMet-tRNA (formylmethionyl transfer-RNA).[1] Each oligonucleotide bears a 5′-monomethoxytrityl group and a TBDMS group at each 2′ position and the terminal 3′ position. Free amino groups on guanine and cytidine residues are benzoylated. The β,β,β-trichloroethyl (TCE) moiety is employed as the phosphate blocking group. The negative ion spectrum of the trimer in the region above m/z 500 is shown in Figure 1. Sites of bond cleavages leading to the production of the primary sequence ions are depicted in Figure 2. This oligonucleotide differs from the dinucleoside monophosphate in Figure 1A of the preceding paper in this series only by the presence of an additional benzoylcytidine nucleotide residue at 5′-terminus. The spectra of these two oligonucleotides therefore contain the same 5′- and 3′-terminal sequence ions ([MTr(bz)C_p^{Si}(TCE)]⁻ and [(TCE)$_p A_{Si}^{Si}$]⁻, at m/z 944 and 706, respectively). Similar peak patterns are also apparent in the region of the molecular ion in both spectra. This includes the ions (M − R′)⁻, (M − Cl)⁻, (M − H)⁻, and (M + Cl)⁻. The trinucleoside diphosphate spectrum contains, in addition, two more clusters of peaks. The mass of the most prominent ion in each cluster corresponds to the mass of each of the two new dinucleotide sequence ions introduced by the addition of one more residue (peaks at m/z 1361 and 1599). Loss of HCl or chlorine from each new sequence ion is also detected as evidenced by the peaks at m/z 1326 and 1564. Replacement of the trichloroethyl group by hydrogen on the 3′-dinucleotide sequence fragment is observed (m/z 1229) yielding an ion similar in structure to the mononucleotide fragment at m/z 575 in this spectrum and in Figures 1B and 1C of the previous paper.

A variety of peaks of low intensity are also apparent in Figure 1. In the parent ion region a peak is detected 35 u below the (M − Cl)⁻ mass. We assume that this is the result of multiple chlorine loss and the peak is therefore assigned the structure (M − 2Cl)⁻. There is also a weak indication of the same kind of ions in spectrum 1B in part 1 of this series. As indicated in Table I the identity of the negative molecular ion cannot be unambiguously assigned as either (M − H)⁻ or M⁻·. Inspection of the protected mononucleotide spectra in the region of the molecular ion suggests the peak is actually a composite of both ions. The cluster of weak peaks above m/z 2113 are molecular ion adducts formed by the attachment of anions such as PO_2^-, PO_3^-, and HPO_4^-.

The peak at m/z 1911 is the result of chlorine loss from (M − R′)⁻. The high mass shoulder on the (M − R′)⁻ peak is attributed to sodium replacement of one of the TCE groups. This is commonly observed in samples rich in sodium. The peak at m/z 1251 is a dinucleoside monophosphate fragment that also contains one sodium atom instead of the TCE moiety. Sodium addition is also believed to be reesponsible for the peak at m/z 1621. In this case, however, sodium has replaced an acidic proton on the molecule. The three additional primary sequence ions also show some indication of this as suggested by the high mass shoulders.

The clusters of weak peaks between m/z 1650–1800, 1400–1550, and 750–900 are largely unidentified. The somewhat prominent peak at m/z 1080 is also unidentified and appears in all negative ion spectra of this series of oligonucleotides. It is,

(1) Ogilvie, K. K.; Theriault, N. Y. *Can. J. Chem.* **1979**, *23*, *3140*–3144.

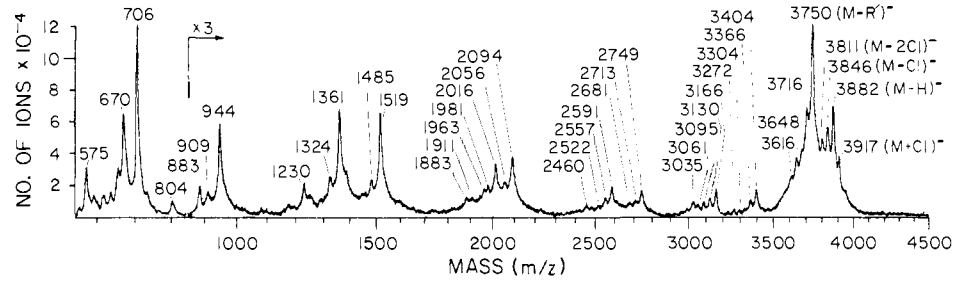

Figure 4. ^{252}Cf-PD negative ion spectrum of the hexamer (structure given in Figure 6). Experimental masses are indicated above the peaks. Data were accumulated for 5×10^4 s and are plotted in 4 ns wide time bins.

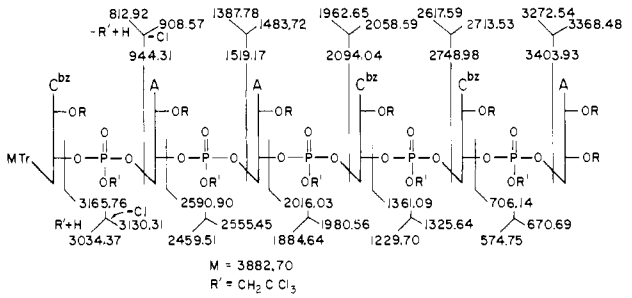

Figure 5. Shorthand structure of the hexamer. Abbreviations are indicated previously. Calculated average masses of each sequence ion are indicated by the cleavage sites. Calculated masses of fragments formed by Cl loss from each sequence ion appear to the right of these values; masses of fragments formed by substitution of hydrogen for one R′ group appears to the left.

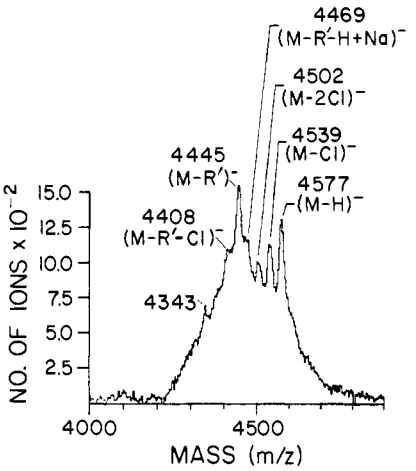

Figure 6. ^{252}Cf-PD negative ion spectrum of the heptamer in the region of the molecular ion. Experimental masses and structural assignments are indicated by each peak. The oligonucleotide has a calculated average mass of 4577.67 u. Mass values of several of the less prominent peaks are skewed in the direction of the interfering adjacent peak. Data were accumulated for 5×10^4 s and are plotted in 8 ns wide time bins.

however, the most intense in the trimer spectrum. The peak at m/z 636 is due to multiple chlorine loss from sequence ion a (see Figure 2).

The mass spectrum of the trimer in the region below m/z 500 is shown in Figure 3. Except for the presence or absence of the different nucleotide base fragments, spectra of all of the protected ribooligonucleotides containing the same phosphate and 5′-, 3′-, and 2′-protecting groups are very similar. Clusters of chlorine-containing peaks are distinguished by the characteristic isotopic peak pattern. The phosphate ester fragment formed by protonation and rupture of both the C3′–OP and C5′-OP bonds is evident at m/z 227. The $^{35}Cl/^{37}Cl$ isotopic peak pattern confirms the presence of three chlorine atoms. Fragment ions are formed by the rupture of the glycosidic bond producing diagnostic base ions at m/z 134 (adenine) and 214 (benzoylcytinine). This species of ion is observed for every base. Relative peak intensities of these ions appear to be related to the extent the negative charge can be delocalized over the molecular surface.[2] Additional identifiable fragments include low molecular weight hydrocarbons (e.g., C_2H^-) and the PO_2^- and PO_3^- ions. Ions in this mass range can be used to identify the presence of the bases and protecting groups but otherwise provide little structural information.

Spectra of the tetramer ($MTrA_p^{Si}(TCE)bzC_p^{Si}(TCE)bzC_p^{Si}(TCE)A_{Si}^{Si}$) and pentamer ($MTrA_p^{Si}(TCE)A_p^{Si}(TCE)bzC_p^{Si}(TCE)bzC_p^{Si}(TCE)A_{Si}^{Si}$) contained two and four more clusters of sequence ions, respectively. Both spectra also contained the characteristic negative molecular ion fingerprint pattern.

The negative ion spectrum of the hexamer is shown in Figure 4. The base sequence and structures of the primary sequence ions are depicted in Figure 5. In this fragmentation scheme masses of ions formed by the loss of HCl or chlorine appear to the right of the masses of the major sequence ions; masses of ions formed by exchange of the β,β,β-trichloroethyl group for hydrogen appear to the left. The ten primary 5′- and 3′-sequence ion peaks are readily distinguishable as the principal peak in each of the ten fragment ion clusters. Experimental masses of these ions are in good agreement with the calculated average masses indicated in Figure 4. Experimental masses of ions formed by loss of HCl or chlorine do not agree well with calculated average masses because they cannot be well resolved from adjacent peaks. Calculated masses deviate less from the experimentally observed values if it is assumed that a chlorine atom is lost rather than HCl. However, loss of a stable neutral molecule such as HCl should be the energetically preferred fragmentation pathway. If it is assumed that HCl is lost all experimental masses are then consistently higher than the calculated masses. It is very likely, however, that the mass values should be skewed in this direction because of the presence of the more intense parent sequence ion ~35 u above each peak of this type. The molecular ion and the penta- and tetranucleotide fragments are observed to also lose a second chlorine or HCl group as evidence by the peaks of low intensity at m/z 2552, 2681, 3095, 3330 (not indicated), and 3811. Multiple chlorine loss is not observed in the spectra of dinucleoside monophosphates which contain only one β,β,β-trichloroethyl group and to only a limited extent in the spectra of trinucleoside diphosphates (Figure 2, Table I).

Replacement of the TCE group by hydrogen is observed only for fragment ions bearing a 3′-terminal phosphodiester group and the parent molecule. The peaks at m/z 1963 and 3272 have the same nominal mass as the 5′-trinucleotide and 5′-pentanucleotide fragments in which the TCE group has been replaced by hydrogen but since no other 5′-sequence ions of this variety are observed it seems doubtful that the ions are formed by this process. Loss of multiple TCE groups does not occur for any sequence ion in any sample studied thus far.

Minor peaks between m/z 575 and 670 appear in many spectra but have not been identifed. Additional unidentified peaks in this spectrum include the ions at m/z 804, 1911, 3061, 3616, and 3648. The peak at m/z 3716 is due to loss of HCl or chlorine and one TCE group from the parent molecule.

(2) McNeal, C. J. Ph.D. Dissertation, Texas A&M University, College Station, TX, 1980.

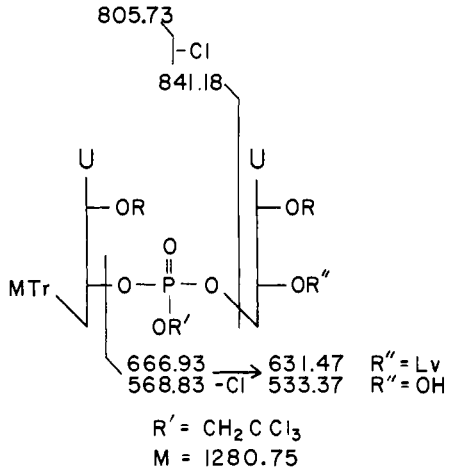

Figure 7. Shorthand structure of $MTrU_p^{Si}(TCE)U_{Lv}^{Si}$. Calculated average masses of the primary molecular and fragment ions are indicated for this form and the delevulinated form (bearing a 3′-hydroxyl group.

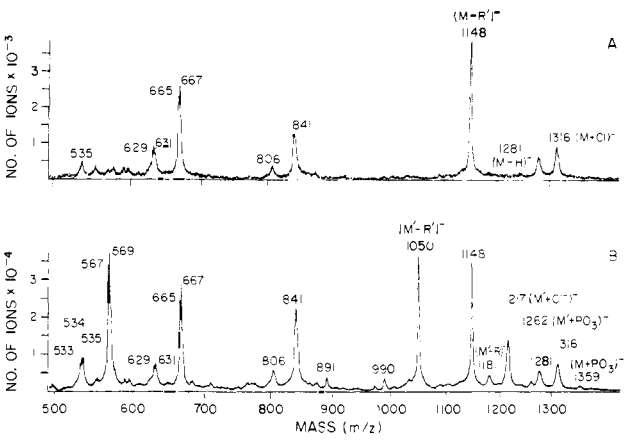

Figure 8. ^{252}Cf-PD negative ion spectrum of (a) $MTrU_p^{Si}(TCE)U_{Lv}^{Si}$ and (b) the same dinucleoside monophosphate containing the delevulinated impurity. Spectrum a was accumulated for 1.8×10^3 s; spectrum b was accumulated for 2.0×10^4 s. both are plotted in 2 ns wide time bins. Experimental masses appear above the peaks.

The negative ion spectrum of the heptamer of this [MTr-(bz)G_p^{Si}(TCE)BzC_p^{Si}(TCE)A_p^{Si}(TCE)A_p^{Si}(TCE)bzC_p^{Si}(TCE)bz-C_p^{Si}(TCE)A_{Si}^{Si}] in the region of the molecular ion is shown in Figure 6. A minor ion peak at m/z 4462, $(M - R' - H + Na)^+$, is present in addition to the usual peaks, $(M - R')^-$, $(M - Cl)^-$, and $(M - H)^-$ or $M^{-\cdot}$. This species is often observed in the ^{252}Cf-PDMS spectra of protected oligonucleotides containing the guanine moiety due to the acidity of the N1 proton.[3] As a result of this all 5′-sequence ions also form sodium ion adducts giving rise to peaks 22 u above the mass of each sequence ion. 3′-sequence ions which necessarily do not contain the benzoylguanine residue do not experience sodium exchange. We have found this reaction to be a convenient indicator for the presence of this base.

All 12 sequence ions were identifiable in the spectrum of the heptamer. Due to the presence of 18 chlorine atoms and the diminishing intensity of the higher-order fragments, masses of the terminal hexanucleotide fragments were more difficult to determine. However, since the sequence information is duplicated by virtue of forming both 5′-terminus and 3′-terminus fragments the base identity of the terminal 3′ nucleotide was easily confirmed by the very prominent ion at m/z 706, $[(TCE)pA_{Si}^{Si}]^-$.

Each oligonucleotide in this series also formed negative dimer ions by losing a TCE group, i.e., $(2\,M - TCE)^-$. These peaks were very broad due to the preponderance of chlorine atoms and the existence of multiple fragment ions.

Characterization of Nucleotide Impurities. Analysis of the uridine dinucleoside to be described in this section was expected to be routine due to the simplicity of the structure. This compound was being investigated because a new hydroxyl protecting group, the levulinyl group, had been introduced at the 2′-end of molecule.[4] We had already established that variations in the protecting groups had little effect on the essential contours of the spectrum. The results obtained were therefore totally unexpected but ultimately illustrated several of the unique features of sequence analysis by ^{252}Cf-PDMS.

The structure and masses of the expected sequence ions for a protected dinucleoside monophosphate containing two uridine bases, a 5′-monomethoxytrityl group, a 3′-levulinyl group, and 2′-TBDMS groups are depicted in Figure 7. Two samples were obtained; for the purpose of this discussion, the results of the second sample will be described first. The negative ion spectrum of the second sample is shown in Figure 8A. The diagnostic 5′- and 3′-mononucleotide fragments are present in the spectrum at m/z 841 and 667. Peaks due to loss of chlorine from each of these fragments and loss of R′ from the 3′-mononucleotide ion (m/z 535) are also evident. This peak pattern and the parent ion signature, e.g., $(M - R')^-$, $(M - H)^-$, and $(M + Cl)^-$, clearly verifies the structure of this compound. The positive ion spectrum also confirmed the singular composition of this sample.

The results of the first analysis of this dinucleoside are shown in Figure 8B. In addition to the two sequence ions, a prominent peak appears at m/z 569. New peaks are also detected at m/z 1050, 1181, and 1217. In addition, the positive ion spectrum contain new peaks 98 u lower than the molecular ion peaks. Such different fragmentation patterns had not been previously observed for two identical molecules. Futhermore, there are not any ionization parameters, such as temperature, that can be altered to produce different fragmentation pathways. Therefore, it was clearly apparent from the data that first sample (Figure 8B) contained an impurity. Examination of the pattern of new peaks suggested that this impurity was a nucleotide.

The appearance of three new peaks each approximately 98–100 u below the $(M - R')$, $(M - H)^-$, and $(M + Cl)^-$ peaks indicated that the second component had a molecular weight of 1182 ± 1 u. There was no evidence of a peak 98 u below the mass of the 5′-mononucleotide fragment (m/z 841) in Figure 8B but a new peak does arise at m/z 569, 98 u below the mass of the 3′-mononucleotide fragment (m/z 667). Because this peak appears 98 u below the mass of only the 3′ sequence it establishes that any differences in the structures of these two components occur only on the 3′-terminal nucleotide unit. Any modification of both nucleosides would result in the mass difference being split between the two sequence ions. This modification is further restricted to the nucleoside moiety since any change in the mass of the phosphate link would give rise to peaks 98 u below the mass of both 5′- and 3′-sequence ions. Peaks appearing 35 u below the m/z 567 peak, 35 u above the proposed parent ion mass and with ~132 u difference between the peaks tentatively identified as $(M' - R')^-$, m/z 1050, and $(M' - H)^-$, m/z 1181, are indicative of the β,β,β-trichloroethyl phosphate protecting group. From these data we can establish that the 5′ termini of both components are identical, that the 3′-termini differ in mass by 98 u, and that the phosphate ester links are identical. Within the logical constraints of the synthetic procedure the impurity could therefore only differ in the nature of the 3′-terminal base or at the 3′- or 2′-terminal positions. The uracil ion is observed at m/z 111. No peak is found 98 u above this mass thereby also eliminating the base moiety as a candidate. The new levulinyl protecting group has been introduced into the structure of this dinucleoside at the 2′ position. The mass of this functional group is 99 u. Hydrolysis of the levulinyl moiety resulting in the formation of a free 2′-hydroxyl function in its place could account for the 98 u difference between the two components.

The proton magnetic resonance spectrum of this dinucleoside can be used to verify the presence of the levulinyl group at the

(3) McNeal, C. J.; Narang, S. A.; Macfarlane, R. D.; Hsiung, H. M.; Brousseau, R. *Proc. Natl Acad. Sci. U.S.A.* **1980**, *77*, 735–739.
(4) Ogilvie, K. K.; Nemer, M. J. *Can. J. Chem.* **1980**, *58*, 1389–1397.

3′-terminal position. If the levulinyl and monomethoxytrityl groups are intact the two methyl resonance signals of these groups should be observed in equivalent intensities occurring at two different chemical shifts arising from the differing environment of each isolated methyl group.[4] The NMR spectrum of the sample containing two components taken on the unused portion shortly after mass spectral analysis indicated that the levulinyl methyl resonance was substantially less intense than the monomethoxytrityl methyl resonance. However, the NMR spectrum of this sample obtained soon after the molecule was synthesized verified that these two signals were of equivalent intensity. We can only infer from this that at some time prior to obtaining the ^{252}Cf-PD mass spectrum of the first sample (spectrum 8B) that it was subjected to conditions favorable for hydrolysis of the levulinyl group.

This result serves to demonstrate the ability of ^{252}Cf-PDMS to unambiguously detect and characterize the presence of nucleotide impurities in a sample. During the course of our investigation of the application of this method to the analysis of many different protected oligonucleotides from a variety of sources we have observed in at least several cases that samples which were judged to contain a single component by thin layer or liquid chromatography actually contained multiple nucleotide components. The recent structural elucidation of side products produced when an excess of condensing agent was employed in the phosphotriester method of synthesis[5] emphasizes that a method such as ^{252}Cf-PDMS which can be used to obtain a rapid and complete characterization of the fully or partially protected synthetic intermediates could be a useful tool in identifying side reactions which might otherwise significantly affect the outcome of the oligonucleotide synthesis.

Conclusion

In this paper we have demonstrated that the base sequence and purity of ribooligonucleotides containing phosphotriester internucleotidic links can be assessed unambiguously from the ^{252}Cf-PD negative ion spectrum. The fragmentation pattern of this class of molecules gives rise to two nested sets of fragment ions which resemble the summed products of two partial enzymatic digests utilizing a 5′- and a 3′-exonuclease. The base sequence information of one set of fragment ions is thus mirrored in the complimentary set. In addition, negative molecular ions are produced thereby identifying the molecular weight of the intact oligonucleotide. Despite the low instrumental resolution, masses of the primary sequence ions can be determined with sufficient precision to meet the necessary analytical requirements. In the examples presented the β,β,β-trichloroethyl group was employed as the phosphate blocking group. This is possibly the least favorable moiety due to the numerous natural isotopic abundance peaks yet sequences could easily be verified extending to a heptamer.

Several criteria have been established for identifying the base sequence from the evaluation of numerous spectra of protected oligonucleotides containing a variety of protecting groups. Because sequence ions are produced from both termini of the oligonucleotide a relationship exists which can be expressed mathematically as MW(5′-) + MW(3′-) − MW(link) = MW. For the simplest case of the dinucleoside monophosphate depicted in Figure 1A of the preceding paper in this series it can be seen that the mass of the 5′-sequence ion at m/z 923, MW(5′-), plus the mass of the 3′-sequence ion at m/z 685, MW(3′-), minus the mass of one phosphate link, MW(link) must give the average parent molecular weight, MW = 1402 u. The link is defined as PO_3OR' which has a molecular weight of 206.5 u for R′ = TCE. For larger oligonucleotides containing N residues there are $(N - 1)$ relationships. In a pentamer, for example, the mass of the 5′-mononucleotide fragment and the 3′-tetranucleotide fragment minus MW (link) must yield the parent molecular weight. There are three more interrelationships that must also correlate to confirm one unique base sequence. In establishing the parent molecular weight a pattern of peaks must be identified which contains as a minimum the $(M - R')^-$ and $(M - H)^-$ peaks. The mass of R′ can be checked since it must give a satisfactory solution to the previously described interrelationship between the sequence ions. The presence of each base must be verified by its characteristic fragment ion. The intense PO_3^- ion peak is the signature of a phosphor triester linkage; unusual internucleotidic links such as phosphite, PO_2OR', or alkyl phosphonate, $R'PO_3$ (both of which are amenable to sequence analysis by ^{252}Cf-PDMS), are distinguished by having PO_2^- peak intensities higher than the PO_3^- peak intensity and the absence of the HPO_4^- ion. The presence of the monomethoxytrityl, TBDMS, and benzoyl group is verified by the presence of characteristic positive ion fragments which are described in the companion article of this series. The positive ion spectrum is also used to identify the parent molecular weight thereby providing a double-check of the value obtained from the negative ion spectrum. The proposed base sequence must meet all of these constraints before it is accepted. We therefore believe that the method is highly reliable because of the numerous interrelationships that must be consistent with one another.

The mode of fragmentation, e.g., rupture along the C5′–OP, C3′–OP, and P–OR′ bonds, appears to be universal for protected nucleotides. Chemically blocked deoxoligonuleotides[3] fragment in precisely the same manner as the ribolignonucleotides described in this manuscript. Oligonucleotides containing unusual linkages such as $R'PO_3$ and PO_2OR' also fragment along the C5′–OP and C3′–OP bonds producing the diagnostic sequence ions. Linkage isomers, 2′→2′, 2′→5′, and 3′→3′ linked nucleotides, also produce the expected sequence ions. Fragmentations along both the C3′–OP and C5′–OP bonds simultaneously, yielding fragments bearing two terminal phosphate groups, are not detected. Oligonucleotides containing the naturally occurring phosphodiester internucleotidic links exhibit a much more complex fragmentation pathway and molecular ion yields are severely attenuated compared to the phosphotriester analogue.

The ^{252}Cf-PDMS method offers several advanatages over conventional methods of oligonucleotide analysis. Foremost is that the base sequence of synthetic oligonucleotides can be determined from the fully protected intermediates. This avoids the time-consuming procedure of deprotection and ^{32}P labeling during intermediate stages of the synthesis. As a result the method can give a more accurate assessment of the sample composition. For example, if an oligonucleotide does not incorporate the radioactive label due to incomplete deprotection or other modification introduced during the chemical synthesis, the products would not be visualized in the chemical or enzymatic digest. The method is very rapid. The sequence of small oligomers can be verified in minutes; for larger oligomers data are accumulated for longer periods of time in the same fashion as in NMR analysis to resolve peaks of weak intensity from the background by signal averaging. Because the method is virtually nondestructive the sample can be returned without any decomposition.

In conclusion, we have presented an alternative method of sequencing oligonucleotides which are not amenable to chemical or enzymatic methods of analysis utilizing the method of ^{252}Cf-plasma desorption mass spectrometry. The method is applicable to a wide variety of chemically blocked deoxy- and ribooligonucleotides, rapid, highly reliable, and nondestructive. These spectra also demonstrate that the ^{252}Cf-PDMS method can be used to produce large negative ions. To our knowledge these are the largest negative ions that have been reported in the mass spectral literature.

Acknowledgment. This work was supported by grants from the National Institute of General Medical Sciences (GM26096), the National Science Foundation (CHE-79-04863), and the Robert A. Welch Foundation (A258), all awarded to R. D. Macfarlane, and the Natural Sciences and Engineering Research Council of Canada. We are indebted to Ronald Macfarlane for many fruitful discussions, and to Dennis Shelton, Randall Martin, and Michele Bailey for their technical assistance.

Registry No. Trimer CCA, 80049-83-8; heximer CAACCA, 80063-10-1; heptimer GCAACCA, 80105-68-6; $MTrU_p^{Si}(TCE)U_{LV}^{Si}$, 80063-12-3; $MTrU_p^{Si}(TCE)U^{Si}$, 80063-13-4.

(5) Reese, C. B.; Ubasawa, A. *Tetrahedron Lett.* **1980**, *21*, 2265–2268.