Sequence Analysis of Oligodeoxyribonucleotides by Mass Spectrometry. 1. Dinucleoside Monophosphates†

J. L. Wiebers* and J. A. Shapiro

ABSTRACT: The base components of underivatized oligodeoxynucleotides can be determined qualitatively by mass-spectral analysis at the nanogram level. The thermal and electron-impact conditions of the spectrometer allow the cleavage of the phosphodiester bonds of the oligonucleotide chain, resulting in fragments (I), whose mass identifies the base, and other fragments (II), which contain the purine or pyrimidine base plus portions of the deoxyribose and the phosphate moieties. A study of 16 dinucleoside monophosphates indicates that the relative intensities of the *m/e* values of the type II fragments are significantly and reproducibly different for sequence isomers. From the complex spectra of dinucleoside monophosphates, specific ions for each mononucleotide residue have been selected which reveal the location (5′ or 3′ terminus) of the bases in the dinucleoside monophosphate. These ions appear in spectra of deoxyribonucleic acid fragments as well as in model compounds. A simple computer program has been devised which utilizes ion ratio values to determine sequence. The method is applicable to oligonucleotides of longer length.

Successful methods for the determination of the base sequences of DNA fragments have principally employed combinations of enzymatic, chromatographic, and radioactive tracer techniques. The objective of the work presented here has been to develop a physical sequencing method utilizing mass spectrometry, which would complement the existing methods for sequence analysis. The exploitation of the mass-spectrometric approach in this area has been pursued by us because of several distinctive characteristics of the technique, one of which is sensitivity. Since oligonucleotides derived from biological sources are always in limited supply, the sensitivity of any sequence method is paramount. The sensitivity of a low resolution mass spectrometer allows less than 0.001 optical density units of the nucleic acid bases to be detected, a value considerably beyond the limit of the ultraviolet absorption spectrometer. Detection by radioactivity is somewhat more sensitive than the mass spectrometric technique; however, the latter has the added advantage of allowing the detection and identification of unusual bases or chemically modified bases that may be present in an oligonucleotide. In addition, with respect to radioactive techniques, considerable difficulties can be encountered in preparing highly labeled messenger RNA or RNA transcripts or DNA fragments derived from copying mechanisms. A sequence method applicable to small amounts of such materials in the unlabeled state should be advantageous. Other attributes of the mass-spectrometric method are the speed of analysis and the unequivocal identification of the components present in the sample. The latter can be of particular value for confirmation of the components of an oligonucleotide which previously has been assigned a structure based upon electrophoretic or chromatographic mobility. Another useful application might be for the sequence analysis of large oligonucleotides that have been chemically modified for selective enzymatic attack to obtain smaller fragments. The regions that are blocked and thus are not susceptible to further enzymatic degradation can easily be analyzed by mass spectrometry.

Initial attempts for the development of mass-spectrometric methods for sequence analysis have been impeded by the problem of low volatility of nucleic acids. Derivatization procedures, which permit the volatilization of nucleotide compounds, result in mass values which in the case of oligonucleotides soon become too high for analysis. In spite of this limitation, this type of approach has allowed the sequence determination of ribodinucleotides (Hunt et al., 1968; Dolhun and Wiebers, 1969), and a chemico-mass spectrometric method was developed for the sequence analysis of any tri- or tetraribonucleotide (Wiebers, 1970).

A procedure employing a different strategy was developed for the sequence determination of oligodeoxyribonucleotides (Wiebers, 1973) which resulted from the observation (Charnock and Loo, 1970) that deoxyoligonucleotides behave much differently from ribooligonucleotides when analyzed in the mass spectrometer. Underivatized oligodeoxyribonucleotides of any chain length are cleaved (by electron impact and by pyrolysis) at every phosphodiester linkage, and diagnostic *m/e* peaks of relatively low mass value for the bases and for an "elimination" derivative from the nucleosides are obtained in the spectrum. This permits a qualitative analysis for the components of a deoxyribooligonucleotide of any size. The sequence method, which was essentially an end-group analysis, involved the trifluoroacetylation of the 3′ and 5′ terminals of an oligonucleotide. A mass spectrum of the product showed ions which were diagnostic for the 3′ and 5′ terminals of the molecule and other ions which identified all the purine and pyrimidine bases present in the interior of the molecule. The sequence of the internal residues, as well as the identity of the terminal nucleosides, was determined by the application of the technique to the isolated products (which are oligonucleotides of decreasing chain length) of limited venom phosphodiesterase digestion of the oligonucleotide. The identification of the 3′-terminal nucleoside of each $n - 1$ oligonucleotide permitted the reconstruction of the parent molecule. Model compounds of up to 12 residues were successfully sequenced by this method.

A weakness of the end-group analysis procedure, however, is that it depends upon the separation of the $n - 1$ chain-length oligonucleotides after the phosphodiesterase reaction and this

† From the Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907. *Received May 14, 1976.* Supported by National Science Foundation Research Grant BMS-74-22213.

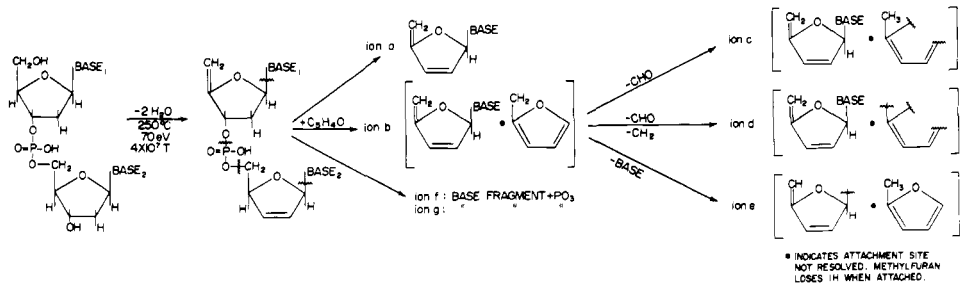

FIGURE 1: Suggested mass-spectral fragmentation pattern for deoxydinucleoside monophosphates.

is not feasible when very small amounts of the original oligonucleotide fragment are used. For this reason, a different approach (Shapiro and Wiebers, 1975) has been investigated and is reported here in which the strategy has been the direct structural analysis of intact, underivatized, oligonucleotides by allowing the mass spectrometer itself to cleave the large molecule into smaller units for analysis. This communication presents evidence that the base components of underivatized oligodeoxynucleotides can be determined qualitatively by mass spectral analysis at the nanogram level. Furthermore, the thermal and electron-impact conditions of the spectrometer cleave oligonucleotides of any chain length, yielding fragments which contain the purine or pyrimidine base and structures composed of the base plus portions of the deoxyribose and/or the phosphate moieties. The relative intensities of the *m/e* values of these fragments are significantly and reproducibly different in sequence isomers of dinucleoside monophosphates. From the complex spectra of the 16 dinucleoside monophosphates commonly found in DNA, certain ions specific for each mononucleotide residue have been selected which reveal the location (5′ or 3′ terminus) of the bases in the dinucleoside monophosphate. A simple computer program has been devised which utilizes data from the ratios of these ions to each other to determine sequence. The following paper in this issue indicates how this same approach has been applied to oligonucleotides of longer chain length, and how computerized pattern-recognition techniques have been applied to the analysis of the more complex spectral data.

Materials and Methods

Materials. Deoxydinucleoside monophosphates, deoxydinucleotides and deoxymononucleotides, and dinucleotides with protecting groups, were obtained from Collaborative Laboratories, Waltham, Mass. Data sheets from Collaborative Laboratories supplied information on the purity of the compounds based upon chromatographic mobility in two solvent systems. Confirmation of the purity was done by thin-layer chromatography in our laboratory. Carrier-free [^{32}P]deoxyadenosine 5′-monophosphate was synthesized and purified by New England Nuclear, Boston, Mass. and had a specific activity of 117.7 Ci/mM. The single-stranded oligonucleotide from calf thymus DNA was isolated and purified by Dr. Nancy Ho after DNase I digestion of the DNA.

Sample Preparation. Water solutions containing 0.1–1.0 optical density units of the compounds were introduced into capillary sample tubes (12-mm length) and taken to dryness in vacuo in a desiccator containing P_2O_5. For the experiment comparing the spectra of [^{32}P]deoxyadenosine 5′-monophosphate with deoxyadenosine 5′-phosphate, 0.03 optical density unit of both the labeled and unlabeled compounds was used.

Mass-Spectral Analysis. Samples were introduced into the spectrometer (DuPont 21-490B, low resolution) by direct probe and slowly heated to about 250 °C at 70 eV, 3×10^{-6} Torr, source temperature, 200 °C. Spectra were taken at the point at which the maximum number of ions were generated as indicated by the ion monitor. Multiple analyses were done on each compound for statistical analysis of reproducibility of the spectra. Elemental composition analysis of specific ions was done on a high-resolution mass spectrometer (CEC-21-110 B) in the Chemistry Department, Purdue University.

Computer Program. A PDP-8/L 12-bit digital computer with a peripheral ASR/35 teletype was used to write the program in FOCAL. A flow diagram indicating how the program was designed and a print-out of the program appear in the Added in Proof section of this paper.

Results

Mass-Spectral Analysis of Deoxydinucleoside Monophosphates. A proposed scheme for the mass-spectral fragmentation of dinucleoside monophosphates is outlined in Figure 1. In this scheme, two molecules of water are eliminated at the 5′- and 3′-hydroxyl sites, resulting in a structure which further fragments to yield (1) the purine or pyrimidine base; (2) ion a, whose structure has been previously characterized (Charnock and Loo, 1970); (3) ions b, c, and d, which are similar in structure to those described in a high-resolution pyrolysis field desorption mass-spectral study of deoxyribonucleic acid (Schulten et al., 1973) and which are apparently formed in the mass spectrometer under pyrolytic conditions by attachment of methylfuran moieties to ion a; (4) ion e, which is composed of two methylfuran moieties and (5) ions f and g, which are fragments from the purine or pyrimidine bases with PO_3 attached through a phosphoramidate linkage at the exocyclic nitrogen of those bases that possess exocylic amino groups.

A typical spectrum of a dinucleoside monophosphate is that of deoxycytidylylguanosine monophosphate (Figure 2). The peaks at *m/e* 111 and 151 represent cytosine and guanine, respectively, and they are the base plus H ions derived from the fragmentation of the base from the nucleoside (McCloskey, 1974). In the case of the cytosine-containing moiety of the compound, ion a appears at *m/e* 191; ions b, c, and d at *m/e* 271, 242, and 228; and ions f and g at *m/e* 148 and 162, respectively. Corresponding structures from the guanine-containing moiety are ion a at *m/e* 231; ions b, c, and d, at *m/e* 311, 282, and 268; and ions f and g at 188 and 202, respectively. Thus, there are fragments from any dinucleoside monophosphate which are specific for the component nucleotide moieties. Table I summarizes the mass values for ions a, b, c, d, f, and g for all four of the nucleotide components derived from dinucleoside monophosphates. These same ions also can be found in mass spectra of oligonucleotides from natural origin, as can be seen in the spectrum (Figure 3) of an oligonucleotide approximately 25 residues in length that was isolated from calf

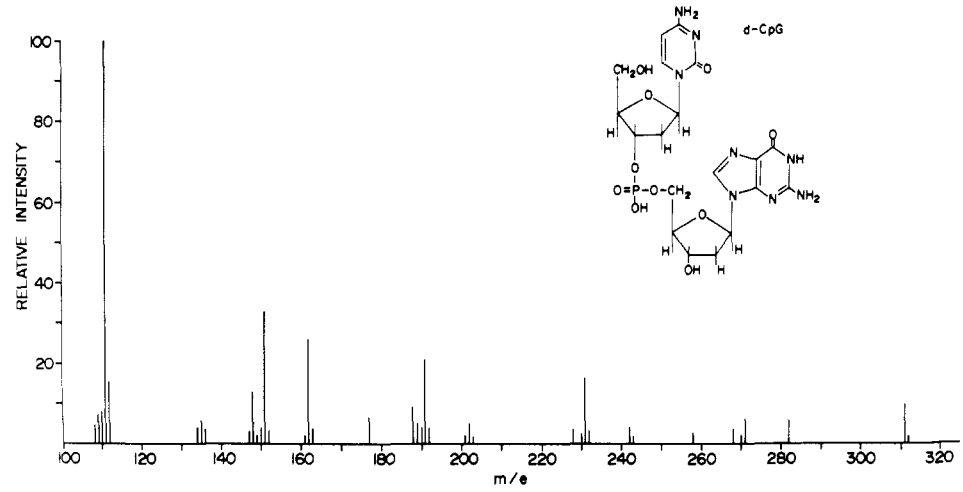

FIGURE 2: Mass spectrum of deoxycytidylylguanosine monophosphate.

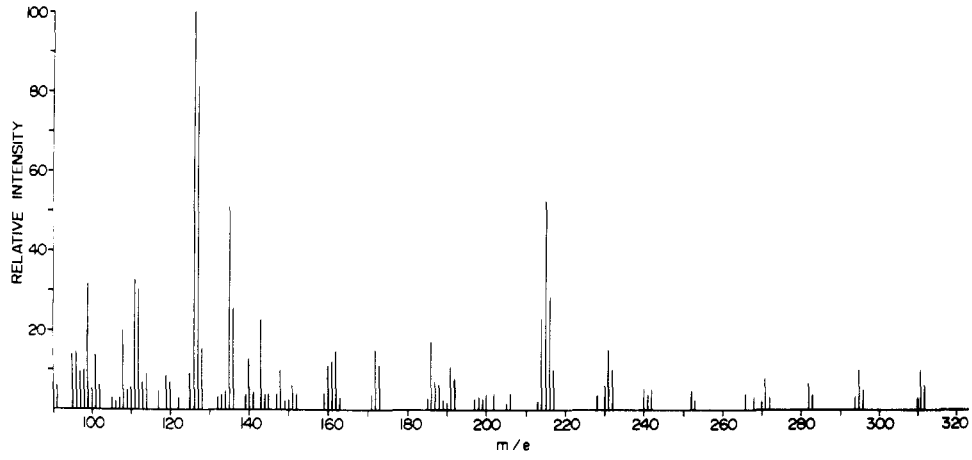

FIGURE 3: Mass spectrum of an oligonucleotide fragment approximately 25 residues in length from calf thymus DNA.

thymus deoxyribonucleic acid which contains all four nucleotide components. It can be noted that all of the ions listed in Table I are present in this spectrum.

To ascertain the probable structure of ions a, b, c, d, e, f, and g, a high-resolution mass-spectrometric elemental composition analysis was carried out on deoxyadenylyladenosine monophosphate and the results (Table II) support the suggested ions diagrammed in Figure 1. Ions f and g are of particular interest, since this is the first time that they have been characterized. The attachment of a PO_3 moiety to part of the purine or pyrimidine base appears to occur through a phosphoramidate link at the amino substituent at the 4 position in cytosine, the 6 position in adenine, and the 2 position in guanine, yielding values for the f and g ions to be *m/e* 148 and 162 for cytosine, 172 and 186 for adenine, and 188 and 202 for guanine. There are no corresponding ions found from thymidylylthymidine monophosphate, or from deoxyuridine 5′-monophosphate, presumably because of the lack of an exocyclic nitrogen in these molecules.

Additional proof for the attachment site of PO_3 was obtained by the mass-spectral analysis of the dinucleotides 5′-*O*-phosphorylthymidylyl-(3′-5′)-2′-deoxycytidine (d-pTpC) and 5′-*O*-phosphorylthymidylyl-(3′-5′)-N^4-anisoyl-2′-deoxycytidine (d-pTpC^{An}). A comparison of the spectra for these two compounds (Figures 4 and 5) clearly shows that ions f and g are major ions in d-pTpC (*m/e* 148 and 162), but that they are nearly absent in d-pTpC^{An}, presumably because any PO_3 that is released upon cleavage of the compound cannot attach to the protected amino group in cytosine. In the spectrum of d-pTpC^{An}, the anisoyl substituent appears at *m/e* 135, the cytosine fragment plus anisoyl at *m/e* 245 and 246, and ion a at *m/e* 325. Similar studies on 5′-*O*-phosphorylthymidylyl-(3′-5′)-2′-deoxy-N^2-isobutyrylguanosine and on 5′-*O*-phosphoryl-N^6-benzoyl-2′-deoxyadenylyl-(3′-5′)-thymidine yielded corresponding results.

Further confirmation of the attachment of PO_3 was afforded by comparing a spectrum of carrier-free [^{32}P]deoxyadenosine 5′-monophosphate with a spectrum of unlabeled deoxyadenosine 5′-monophosphate. A comparison of the pertinent ions from the two spectra (Figure 6) shows a shift of one mass unit

TABLE I: Mass Values for the Fragment Ions from Components of Dinucleoside Monophosphates.

Ion Type	Mass			
	Ado	Guo	Thd	Cyd
a	215	231	206	191
b	295	311	286	271
c	266	282	257	242
d	252	268	243	228
f	172	188		148
g	186	202		162

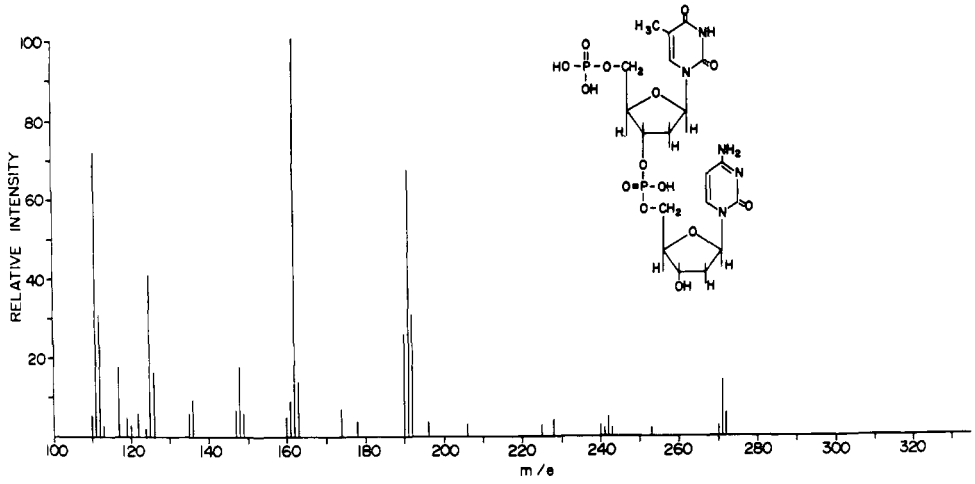

FIGURE 4: Mass spectrum of 5′-*O*-phosphorylthymidylyl-(3′-5′)-2′-deoxycytidine.

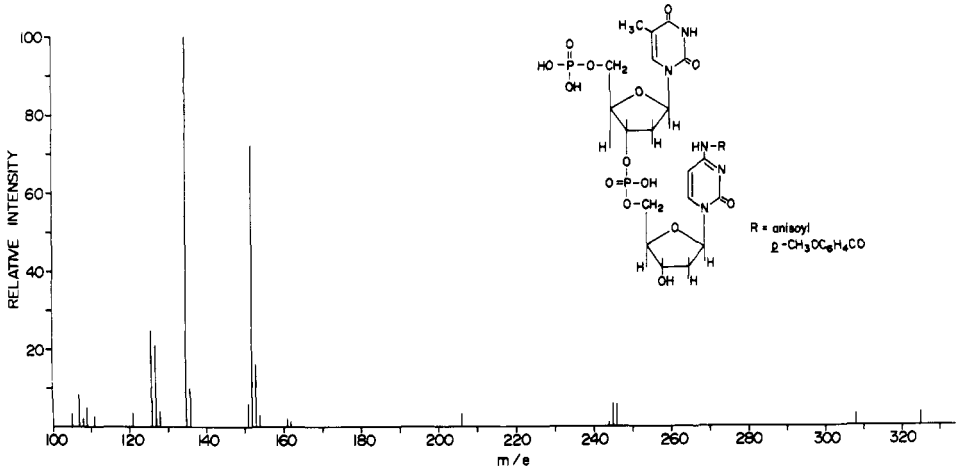

FIGURE 5: Mass spectrum of 5′-*O*-phosphorylthymidylyl-(3′-5′)-N^4-anisoyl-2′-deoxycytidine.

TABLE II: High-Resolution Data from the Major Ions from the Mass Spectrum of Deoxyadenylyladenosine Monophosphate.

m/e	Ion Structure Type	Composition
295	b	$C_{15}H_{13}N_5O_2$
266	c	$C_{14}H_{12}N_5O$
252	d	$C_{13}H_{10}N_5O$
215	a	$C_{10}H_9N_5O$
186	g	$C_5H_3N_2O_4P$
172	f	$C_4H_3N_3O_3P$
161	e	$C_{10}H_9O_2$

for structures containing the radioactive phosphorus. For example, the relative intensity of *m/e* 98, which is H_3PO_4, and of *m/e* 99 are different for the labeled and unlabeled compounds. It can be seen that a similar shift from *m/e* 172 (ion f) to 173 and from *m/e* 186 (ion g) to 187 occurred. Other ions in the spectrum of the radioactive compound did not show this difference with the exception of the ones shown in Figure 6 and these ions have not been further characterized.

Determination of the Base Components of a Deoxydinucleoside Monophosphate. The base components of any dinucleoside monophosphate can be determined readily by examination of the mass spectrum. The fragment ions of *m/e* 111, 126, 135, and 151 are outstanding in intensity and represent the purine or pyrimidine bases (base + H) derived from the nucleoside moieties and are identified as cytosine, thymine, adenine, and guanine, respectively (McCloskey, 1974).

Mass Spectra of Sequence Isomers of Deoxydinucleoside Monophosphates. Spectra of deoxyadenylylthymidine monophosphate (d-ApT) and deoxythymidylyladenosine monophosphate (d-TpA) are superficially similar in appearance with respect to the types of fragmentation (Figure 7). However, it may be observed that the relative intensities of the major peaks from the two compounds differ markedly and this observation suggested that these differences could be used for sequence determination. Multiple spectra of each compound were obtained to verify that the pyrolytic and electron-impact fragmentation process was reproducibly different for the two compounds. The results of this study (Figure 8) indicated that when peak height ratios of selected ions are calculated reproducible differences are apparent. Similar studies were carried out on the six pairs of sequence isomers commonly found in DNA with the results (Figure 9) that reproducible differences can be demonstrated between isomers for each pair when the appropriate ion ratios are calculated. It may be noted that the ion ratios used to show these differences were derived from the ion types a, b, c, d, f, and g, whose structures were discussed above. In the case of thymine-containing compounds which do not give ions f and g, other ions of outstanding intensity were selected.

Computer Program for the Sequence Determination of Deoxydinucleoside Monophosphates. A simple computer

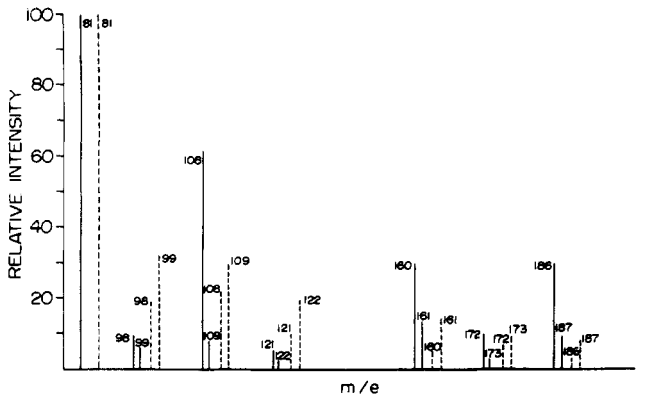

FIGURE 6: Comparison of selected ions from spectra of deoxyadenosine 5′-monophosphate (—) and [^{32}P]deoxyadenosine 5′-monophosphate (- - -).

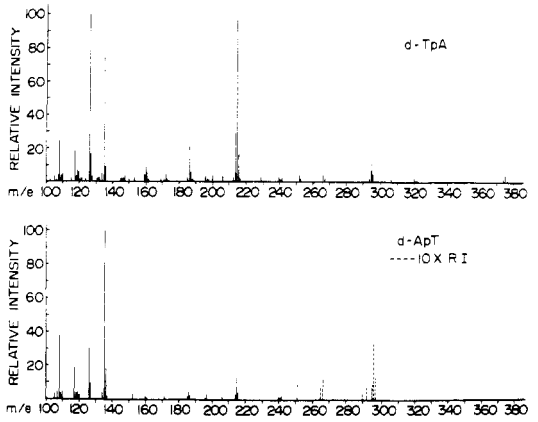

FIGURE 7: Mass spectra of sequence isomers deoxythymidylyladenosine monophosphate and deoxyadenylylthymidine monophosphate.

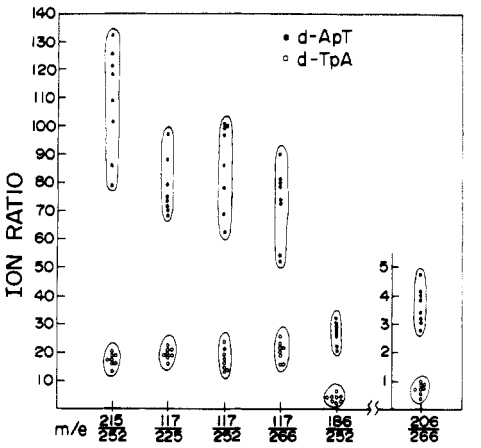

FIGURE 8: Reproducibility of ion ratio values from eight spectra of deoxyadenylylthymidine monophosphate (●) and from deoxythymidylyladenosine monophosphate (○).

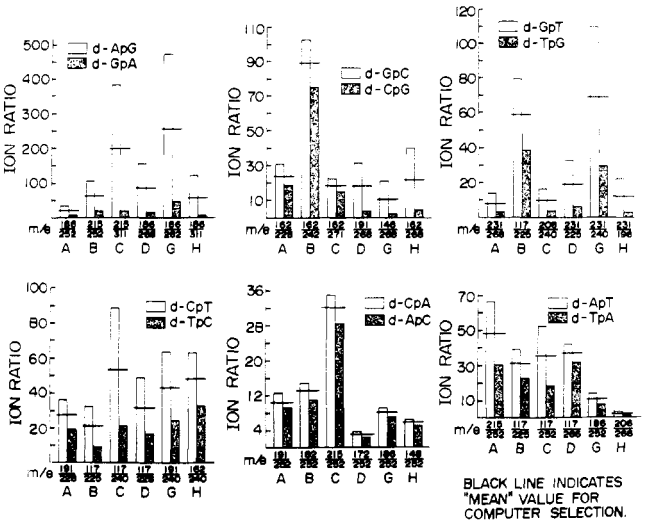

FIGURE 9: Differences in ion ratio values for pairs of deoxydinucleoside monophosphate sequence isomers.

program, which incorporated the type of data given in Figure 9, was devised to determine the sequence of any dinucleoside monophosphate from its mass spectrum. From the results of multiple analyses, a specific value (which is indicated by the horizontal line on the bar graphs in Figure 9) was selected midway between the lowest ion ratio value for isomer d-XpY (where X and Y = the purine or pyrimidine base) and the highest ion ratio value of isomer d-YpX, and this value was used to arbitrarily determine the sequence of an unknown compound. For example, if the ion ratio values from the spectrum of the unknown were above this selected value, the isomer would be designated as XpY and below the value, as YpX. Six different ion ratio values from different *m/e* ions were used for each analysis with the criterion that the results from all six must agree in order to define the sequence of the isomer.

The sequence determination procedure is summarized in Figure 10. A mass spectrum of the unknown dinucleoside monophosphate is obtained and the major purine or pyrimidine base components are recorded. For example, in the case of deoxyguanylylcytidine, (d-GpC), *m/e* ions 111 and 151 indicate that "G" and "C" moieties make up the compound. The operator then requests the computer to determine the order of these moieties. The peak heights (in millimeters) for the appropriate ions for "G" and "C", which in this case are the peak heights for *m/e* 162, 191, 228, 242, 268, and 271, are entered I through V, as can be seen in the example of the computer print-out in Figure 10. The computer calculates the appropriate ratios from the peak height values and makes the determination that the values for all six ratios, A through H on the print-out, are above (or below) the arbitrary cutoff value, and therefore, in this example, that "C" is at the 3′ position of d-GpC. The program has been tested on multiple samples of all 16 dinucleoside monophosphates from spectral data taken at different times and on varying amounts of sample from 0.1 to 1.0 optical density unit and has proved to yield reliable results in all cases.

Discussion

The concept that oligonucleotides are not amenable to extensive mass-spectral analysis because of their size and low volatility must be discarded with respect to deoxyribooligonucleotides. The combined effects of heat, vacuum, and electron impact of the mass spectrometer permit a principal cleavage at the phosphodiester linkages of deoxyribooligonucleotides. This does not occur in the case of ribooligonucleotides which are not susceptible to mass-spectral analysis without prior derivatization to increase volatility. Indeed, ribooligonucleotides are nearly inert under the conditions described above. Presumably, this different behavior is due to the presence of the 2′-hydroxyl group in the ribo compounds. The scheme for the cleavage of deoxypolynucleotides (Charnock

PROCEDURE FOR SEQUENCE ANALYSIS OF A DINUCLEOSIDE PHOSPHATE

I. MASS SPECTRUM: NOTE BASES PRESENT. MEASURE PEAK HEIGHTS OF SELECTED IONS.

II. COMMAND COMPUTER TO DO A SPECIFIC DINUCLEOTIDE PAIR.

AT	PAIRS	,	T	POSITION
AC	"	,	A	"
GC	"	,	C	"
AG	"	,	G	"
GT	"	,	T	"
CT	"	,	T	"

III. COMPUTER ASKS FOR PEAK HEIGHT VALUES FOR THE SELECTED PAIR. ENTER VALUES.

IV. COMPUTER CALCULATES RATIO OF PEAK HEIGHTS AND DETERMINES WHETHER RATIO VALUE INDICATES 3' OR 5' POSITION OF THE BASE.

V. EXAMPLE OF COMPUTER PRINT OUT FOR d-GpC.

```
*D 1.15
GC PAIR, C POSITION
I,: 280
J,: 9
K,: 280
L,: 2
M,: 280
N,: 7
O,: 220
P,: 3
S,: 118
T,: 3
U,: 280
V,: 3
A: 3'
B: 3'
C: 3'
D: 3'
G: 3'
H: 3'
```

FIGURE 10: Summary of the procedure for the sequence determination of deoxydinucleoside monophosphates.

```
C-FOCAL, 1969

01.05 S AA=82;S BB=101.7;S CC=58.7;S DD=102.3; D 1.07
01.07 S GG=91.9;S HH=18.4;T "AT PAIR,T POSITION",!; D 2.0
01.10 S AA=31.8;S BB=127;S CC=60.7;S DD=116.9; D 1.12
01.12 S GG=128.1;S HH=32.1;T "AC PAIR,A POSITION",!; D 2.0
01.15 S AA=24.6;S BB=89.2;S CC=18.6;S DD=18.3; D 1.17
01.17 S GG=11.3;S HH=22;T "GC PAIR,C POSITION",!; D 2.0
01.20 S AA=272.4;S BB=835;S CC=199.2; S DD=84.1; D 1.22
01.22 S GG=257.5;S HH=61.5;T "AG PAIR,G POSITION",!; D 2.0
01.50 S AA=8.4;S BB=59.2 ;S CC=10;S DD=19.7;D 1.52
01.52 S GG=69.6; S HH=13; T "GT PAIR,T POSITION",!; D 2.0
01.55 S AA=27.6; S BB=21; S CC=54.3;S DD=32.7; D 1.57
01.57 S GG=44.2 ;S HH=48.7; T "CT PAIR,T POSITION",!; D 2.0

02.02 A ?I,J,K,L,M,N,O,P,S,T,U,V?; S A=I/J;S B=K/L;S C=M/N;S D=O/P
02.03 S G=S/T; S H=U/V
02.05 T "A"; I (A-AA) 2.13,2.12,2.11
02.11 T "; 3'",!; G 2.15
02.13 T "; 5'",!; G 2.15
02.15 T "B"; I (B-BB) 2.18,2.17,2.16
02.16 T "; 3'",!; G 2.2
02.18 T "; 5'",!; G 2.2
02.20 T "C"; I (C-CC) 2.23,2.22,2.21
02.21 T "; 3'",!; G 2.25
02.23 T "; 5'",!; G 2.25
02.25 T "D"; I (D-DD) 2.28,2.27,2.26
02.26 T "; 3'",!; G 2.3
02.28 T "; 5'",!; G 2.3
02.30 T "G"; I (G-GG) 2.33,2.32,2.31
02.31 T "; 3'",!; G 2.35
02.33 T "; 5'",!; G 2.35
02.35 T "H"; I (H-HH) 2.38,2.37,2.36
02.36 T "; 3'",!; Q,!
02.38 T "; 5'",!; Q,!
•
```

FIGURE 11: Computer program for the sequence determination of deoxydinucleoside monophosphates.

d-ApT–d-TpA SEQUENCE DETERMINATION: COMPUTER FLOW DIAGRAM

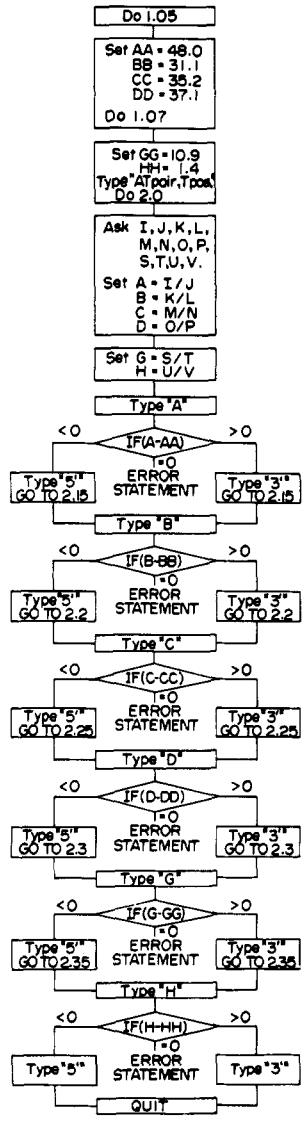

FIGURE 12: Computer program flow diagram for the sequence determination of A–T pairs of deoxydinucleoside monophosphates.

and Loo, 1970), which proposed that fragmentation involves extrusion of the 3′- and 5′-phosphodiester entities from the polynucleotide and that extrusion of the oxygen functions in the 3′ and 5′-positions is more energetically favored in deoxyribo compounds than in ribo compounds, appears to be a reasonable, if not a complete explanation of the process. In the case of dinucleoside monophosphates, we have shown that the major fragmentation products are the purine and pyrimidine bases, fragments of these bases plus part of the phosphate moiety, the bases plus fragments of the deoxyribose, and the base plus part of deoxyribose plus methylfuran. Similar results have been reported (Schulten and Beckey, 1973; Schulten et al., 1973) when deoxynucleosides and deoxynucleotides were examined by high-resolution field desorption mass spectrometry and when deoxyribonucleic acid was subjected to pyrolysis field desorption mass spectrometry. These authors have also suggested the possible use of such fragment ions for sequence analysis.

It would appear then that, since it can be demonstrated that pyrolytic and electron-impact fragmentations of deoxyoligonucleotides yield reproducible structures (or ions) that are informative with regard to sequence, direct structural analysis on these oligonucleotides is possible without prior enzymatic or derivatization procedures. It can be envisaged that an oligonucleotide of some length can be broken into its nucleotide or dinucleotide components, or fragments thereof, and an analysis of the intensities of these structures in relation to each other will yield sequence information. This study on dinucleoside monophosphates is a first step in this type of analysis.

The interest in a sequence method for dinucleotides is, of course, minimal, since there are other simple procedures for this; however, the application of this approach to the sequence analysis of longer oligonucleotides is of value. The following paper of this issue demonstrates how this strategy can be exploited for the analysis of oligodeoxyribonucleotides of greater chain length.

Added in Proof

A print-out of the computer program used in this paper is given in Figure 11. The values for the constants AA, BB, CC, DD, GG, and HH are different for each pair of dinucleotide sequence isomers. As described in the text, these values were derived from multiple mass-spectral analyses and they represent the "mean" ratio values between the highest value in a set of multiple analyses for one isomer and the lowest value in a set of multiple analyses for the other isomer. The mass values that were used for the ratios for the six pairs of sequence isomers are those presented in Figure 9 in the text of the paper.

The print-out of the computer program includes a DO statement number for each of the six pairs of sequence isomers: A–T at 1.05; A–C at 1.10; G–C at 1.15; A–G at 1.20; G–T at 1.50; and C–T at 1.55.

An initial survey of the mass spectrum of a dinucleotide of unknown sequence indicates the two components. Subsequently, the correct DO statement is selected for that particular pair of dinucleotides. The computer then asks for the appropriate peak height values for the selected pair and these values (in millimeters) are entered at I, J, K, L, M, N, O, P, S, T, U, and V by the operator. The computer calculates the ratios of the specific peak heights, determines on which side of the "mean" value such ratio values fall, and states whether this value indicates a 3′ or 5′ position of a particular nucleoside in the dinucleotide. An example of the process is the flow diagram in Figure 12, which indicates the sequence determination of an A–T pair.

Acknowledgments

The authors thank Will Perry for the high-resolution mass-spectrometric measurements and are grateful to Dr. Graham Cooks for advice on the interpretation of that data. Appreciation is also extended to Polly Wise for expert technical assistance.

References

Charnock, G. A., and Loo, J. L. (1970), *Anal. Biochem. 37,* 81.

Dolhun, J. J., and Wiebers, J. L. (1969), *J. Am. Chem. Soc. 91,* 7755.

Hunt, D. F., Hignite, C. E., and Biemann, K. (1968), *Biochem. Biophys. Res. Commun. 33,* 378.

McCloskey, J. A. (1974), in Basic Principles in Nucleic Acid Chemistry, Vol 1, Ts'o, P. O. P, Ed., New York, N.Y., Academic Press, Chapter 3.

Schulten, H. R., and Beckey, H. D. (1973), *Org. Mass Spectrom. 7,* 861.

Schulten, H. R., Beckey, H. D., Boerboom, A. J. H., and Meuzelaar, H. L. C. (1973), *Anal. Chem. 45,* 2358.

Shapiro, J. A., and Wiebers, J. L. (1975), *Fed. Proc., Fed. Am. Soc. Exp. Biol. 34,* 607.

Wiebers, J. L. (1970), *Adv. Mass Spectrom. 5,* 757.

Wiebers, J. L. (1973), *Anal. Biochem. 51,* 542.