Mass Spectrometry of Nucleic Acid Components. Trimethylsilyl Derivatives of Nucleotides[1]

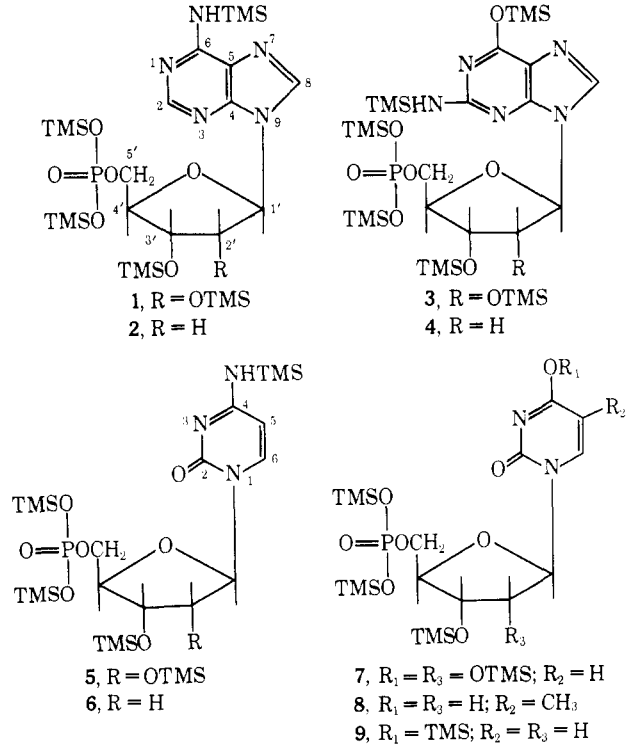
A. M. Lawson, R. N. Stillwell, M. M. Tacker, K. Tsuboyama, and James A. McCloskey*[2]

Contribution from the Institute for Lipid Research and Department of Biochemistry, Baylor College of Medicine, Houston, Texas 77025. *Received August 26, 1970*

Abstract: Trimethylsilylation of mononucleotides yields derivatives which are sufficiently volatile for mass spectrometry and for sample introduction by gas chromatograph. Structure correlations and principal fragmentation pathways have been determined for the trimethylsilyl (TMS) derivatives of the major 5′-nucleotides from RNA and DNA, 2′-deoxyuridine 5′-phosphate, and adenosine 3′,5′-cyclic monophosphate (cyclic AMP), with the aid of high-resolution mass spectrometry and derivatives deuterated in the TMS moiety. Major ions consist of the intact base plus certain portions of the sugar skeleton, and of fragments derived from the phosphate ester. The fragmentation behavior of phosphate-bound TMS was distinguished in some cases from those attached to the base or sugar by deuterium labeling specifically in phosphate TMS. Isomerization of adenosine 2′- and 3′-monophosphate derivatives after electron impact, and under some conditions during silylation, led to spectra which were essentially indistinguishable, but predictably different from that of the 5′ isomer. The mass spectrum of cyclic AMP was found to retain all important features of nucleotide spectra, but clearly reflected the cyclic nature of the molecule.

Although mass spectrometry has shown considerable promise for the structure elucidation of modified nucleosides,[3] its application to nucleotides has been precluded by their high polarity and hence low volatility. The great potential advantages of mass spectrometry, most notably sensitivity and utility in dealing with structural problems, prompted us to seek new approaches in determining mass spectra of mononucleotides. In 1966 Hashizume and Sasaki reported[5] the preparation of trimethylsilyl (TMS) derivatives of mononucleotides, which were sufficiently volatile for gas chromatography.[6] Their work led us to undertake a detailed examination of the mass spectra of these compounds. Our results, reported in a preliminary communication,[7] provide the first mass spectra of nucleotides, and demonstrate the application of combination gas chromatography–mass spectrometry, which facilitates the recording of spectra of individual components of mixtures. Biemann and coworkers have studied the mass spectra of dinucleotide TMS derivatives,[8] and Dolhun and Wiebers have reported the spectra of trimethylsilylated phenylboronates of dinucleotides,[9] both from the viewpoint of sequence determination.

The present article deals principally with TMS derivatives of the four major nucleotides from RNA and DNA,[10] and several structurally related compounds, including cyclic AMP. Structure relationships of the principal fragment ions have been determined, based on results from deuterium labeling and high-resolution mass spectra, permitting comparison of their basic modes of fragmentation with those of carbohydrate TMS ethers reported by DeJongh, Sweeley, and their



1, R = OTMS
2, R = H
3, R = OTMS
4, R = H
5, R = OTMS
6, R = H
7, $R_1 = R_3$ = OTMS; R_2 = H
8, $R_1 = R_3$ = H; $R_2 = CH_3$
9, R_1 = TMS; $R_2 = R_3$ = H

(1) Presented in part at the Sixteenth Annual Conference on Mass Spectrometry and Allied Topics, Pittsburgh, Pa., May 1968, p 228.

(2) To whom correspondence should be addressed.

(3) See for instance (a) S. Hanessian, D. C. DeJongh, and J. A. McCloskey, *Biochim. Biophys. Acta*, **117**, 480 (1966); (b) K. Biemann, S. Tsunakawa, J. Sonnenbichler, H. Feldmann, D. Dütting, and H. G. Zachau, *Angew. Chem., Int. Ed. Engl.*, **5**, 590 (1966); (c) W. J. Burrows, D. J. Armstrong, F. Skoog, S. M. Hecht, J. T. A. Boyle, N. J. Leonard, and J. Occolowitz, *Science*, **161**, 691 (1968). For a comprehensive list of references, see ref 4.

(4) S. J. Shaw, D. M. Desiderio, K. Tsuboyama, and J. A. McCloskey, *J. Amer. Chem. Soc.*, **92**, 2510 (1970).

(5) T. Hashizume and Y. Sasaki, *Anal. Biochem.*, **15**, 199 (1966).

(6) For leading references to the trimethylsilylation and gas chromatographic properties of mononucleotides and related compounds, see: (a) C. W. Gehrke, D. L. Stalling, and C. D. Ruyle, *Biochem. Biophys. Res. Commun.*, **28**, 869 (1967); (b) R. L. Hancock, *J. Gas Chromatogr.*, **6**, 431 (1968); (c) F. Eisenberg, Jr., and A. H. Bolden, *Anal. Biochem.*, **29**, 284 (1969).

(7) J. A. McCloskey, A. M. Lawson, K. Tsuboyama, P. M. Krueger, and R. N. Stillwell, *J. Amer. Chem. Soc.*, **90**, 4182 (1968).

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

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

(10) Names of compounds and abbreviations used in this paper: ribonucleic acid, RNA; deoxyribonucleic acid, DNA; the following 5′-nucleotides: adenosine monophosphate, AMP; guanosine monophosphate, GMP; cytidine monophosphate, CMP; uridine monophosphate, UMP; the 2′-deoxyribotides are denoted by the letter d, as dAMP, except thymidine monophosphate, TMP; adenosine 3′,5′-cyclic monophosphate, cyclic AMP.

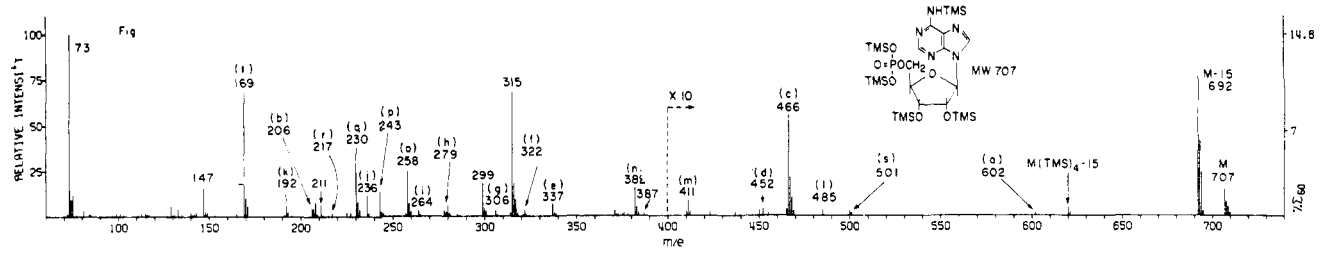


Figure 1. Mass spectrum of 5′-AMP-(TMS)$_5$ (**1**).

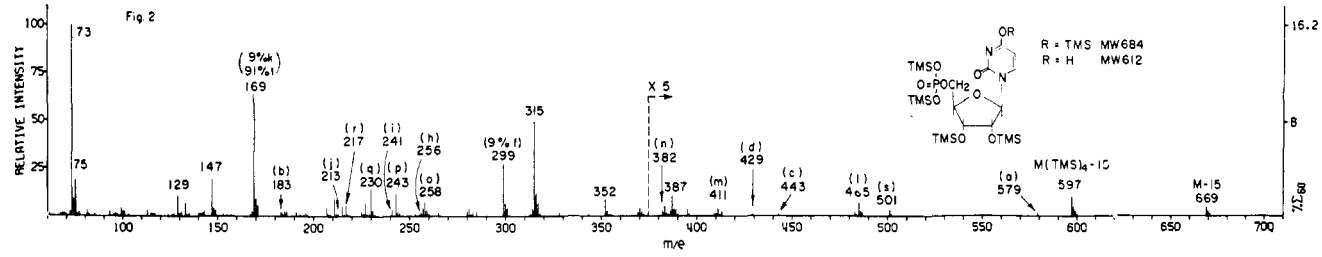


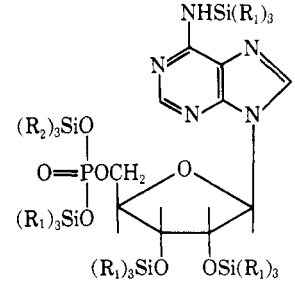
Figure 2. Mass spectrum of 5′-UMP-(TMS)$_5$ (**7**).

collaborators,[11] and of trimethylsilylated sugar phosphates studied by Zinbo and Sherman.[12]

Trimethylsilyl Derivatives. TMS derivatives have found wide use in gas chromatography[13] and mass spectrometry because of their ability to reduce hydrogen bonding interactions and thus increase volatility. In our experience with nucleic acid components, these derivatives were found to possess several inherent difficulties, but the advantages justified their use. Although we have successfully prepared and gas chromatographed all of compounds **1–9**, relatively poor reproducibility with regard to yields was observed. The principal derivatives formed were those indicated in structures **1–9**. This represents the first successful attempt to chromatograph a TMS derivative of 5′-CMP.[5,6] From 20 to 40% of the total yield of **3** and **7** was (TMS)$_5$ derivatives, while the yield of minor derivatives for other compounds was generally low, and caused no confusion in the interpretation of the spectra. The AMP derivative **1**, Figure 1, usually gave the lowest yield of lower derivative (*m/e* 635, 620), while UMP (**7**) represented the least favorable case, shown in Figure 2 (*m/e* 612, 597). Multiple derivatives such as those of AMP were found to very nearly cochromatograph on nonpolar liquid phases (1% SE-30), but to separate slightly on 1% OV-17. The condition of the gas chromatographic column was found to play a major role in determining the apparent yield of the reaction. If the column had not been recently used for nucleotide TMS derivatives, it was often necessary to inject the sample several times to obtain a full and reproducible detector response. In the case of **5** tailing sometimes occurred, and peak area was found to decrease markedly with increasing retention time, although no consistent relationship could be found.

In spite of these difficulties, the silylation reaction is simple to carry out on a microscale, gives no undesirable side products, and permits the usual advantages of gas chromatography–mass spectrometry: rapid sample introduction and clean separation from other nucleotides and impurities. Using the jet-orifice carrier gas separator,[14] no significant tailing, decomposition, or adsorption of samples occurring during transfer between the two instruments was apparent. Model studies showed that the lower sample limit for useful low-resolution mass spectra was approximately 50 ng in terms of underivatized AMP by either probe or gas chromatographic introduction. We point out that this value represents the amount of material introduced into the instrument, not the amount ultimately reaching the ion source, and does not represent other problems of sample isolation and derivatization which could result from an actual hydrolysate of RNA.

The use of TMS-d_9 derivatives[15] such as **1a** and the corresponding derivatives of **2** and **7** (**2a** and **7a**) were of substantial importance in the interpretations presented in the Discussion. Exact mass values from high-resolution mass spectra were of little value in many instances because of ambiguities in elemental composition associated with the possible presence of five or six elements in many ions. Indirect determination of the number of Si atoms in a fragment ion from mass shifts[15] from the



1a, $R_1 = R_2 = CD_3$
b, $R_1 = CH_3$; $R_2 = CD_3$

(11) D. C. DeJongh, T. Radford, J. D. Hribar, S. Hanessian, M. Bieber, G. Dawson, and C. C. Sweeley, *J. Amer. Chem. Soc.*, **91**, 1728 (1969).
(12) M. Zinbo and W. R. Sherman, *ibid.*, **92**, 2105 (1970).
(13) A. E. Pierce, "Silylation of Organic Compounds," Pierce Chemical Co., Rockford, Ill., 1968.
(14) R. Ryhage, *Ark. Kemi*, **26**, 305 (1967).
(15) J. A. McCloskey, R. N. Stillwell, and A. M. Lawson, *Anal. Chem.*, **40**, 233 (1968).

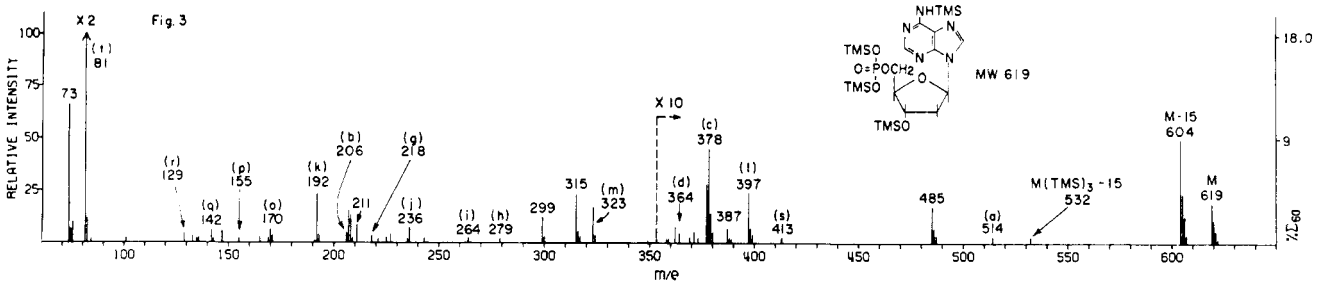


Figure 3. Mass spectrum of 5′-dAMP-(TMS)$_4$ (**2**).

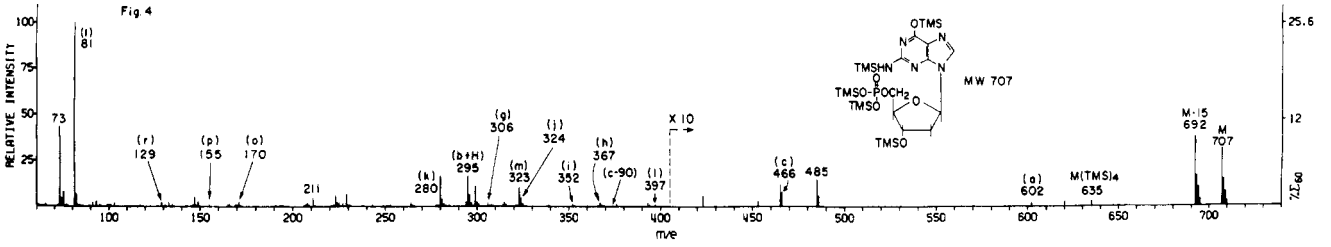


Figure 4. Mass spectrum of 5′-dGMP-(TMS)$_5$ (**4**).

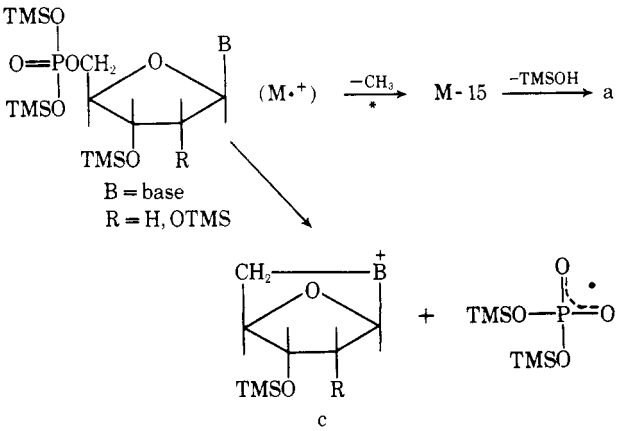
TMS-d_9 derivative usually permitted unambiguous assignment of elemental composition.[16]

Previous results from this laboratory have demonstrated the unique and informative use of partially deuterium-labeled TMS derivatives in the interpretation of their mass spectra.[15] In the present case, preparation of the labeled compound **1b** (mixed in the ratio 4:5 with **1**) provided an unusual opportunity to distinguish and assess the fragmentation behavior of different TMS groups in the same molecule. Although we have not made extensive studies of reaction conditions to prepare differentially labeled TMS derivatives of these compounds, the use of trimethylchlorosilane-d_9[15] in conjunction with other reagents[13] such as hexamethyldisilazane and TMS-imidazole should yield rewarding results as a general approach for a variety of compounds.

Discussion of Mass Spectra

Fragment ions have been classified in terms of their structural relationships to the parent molecule. Ions of the base series include the intact purine or pyrimidine base plus a fragment or hydrogen from the sugar skeleton, while the sugar phosphate series ions contain the phosphate moiety plus various fragments of the sugar skeleton without the base.

Ions Closely Related to the Molecular Ion. In a system of relatively restricted molecular structures such as components of nucleic acids, the molecular ion assumes added importance since it greatly restricts the number of probable structures. Molecular ions were not observed from **6** and **7**, but in every case the characteristic[17] M − 15 peak due to loss of a TMS methyl radical was clearly evident, as indicated in Figures 1–4 and Table I. These ions are generally more abundant in the purine ribotides, corresponding to high stabilization of the molecular ion by the purine nucleus. Further elimination of TMSOH from M − 15 produces the minor but characteristic ion a (*e.g.*, *m/e* 579 in Figure 2). The M and M − 15 ions of pyrimidine deoxynucleotide derivatives are substantially more abundant than those of their deoxynucleoside[18] counterparts. This fact may have bearing on the choice of a product for DNA hydrolysis (*i.e.*, nucleoside *vs.* nucleotide) if the hydrolysate is to be examined for modified components by gas chromatography–mass spectrometry.



Although the loss of a methyl radical from a TMS function upon electron impact is a ubiquitous process, no information has been available concerning the relative tendency for this process to occur from nonequivalent TMS groups in the same molecule. Compound **1b**, containing a TMS-d_9 group in the phosphate moiety, provides a means of distinguishing phosphate methyls from those associated with the base or sugar. M and M − 15 ions from the **1** component (*cf.* Table II) are different in mass from the **1b** ions M (*m/e* 716), M − CH_3 (*m/e* 701), and M − CD_3 (*m/e* 698), and so do not interfere. Random loss of methyl from *m/e* 716 would

(16) For example, from the experimentally determined value *m/e* 411.1243 from **1** (ion m, to be discussed), there are 38 possible elemental compositions falling within ±10 ppm, using composition limits of the molecular ion ($C_{25}H_{54}O_7N_5Si_5P$). Use of the labeled derivative **1a** showed the presence of three silicon atoms in *m/e* 411, reducing the number of arithmetically possible compositions to eight, from which $C_{14}H_{32}O_6Si_3P$ was chosen, based on structural and other factors.

(17) H. Budzikiewicz, C. Djerassi, and D. H. Williams, "Mass Spectrometry of Organic Compounds," Holden-Day, San Francisco, Calif., 1967, p 471.

(18) Unpublished results; see also ref 7.

Table I. Selected Ions from the Mass Spectra of Nucleotide TMS Derivatives

Compound[a]	(m/e)/relative intensity M	M − 15	e	i	b + 2H	k	s	t
Adenosine 5′-monophosphate-(TMS)$_5$ (**1**)	707/1.7	692/9.5	337/10	264/6.0	208/14[b]	192/26	501/0.42	169/82
2′-Deoxyadenosine 5′-monophosphate-(TMS)$_4$ (**2**)	619/0.96	604/2.5	249/0.0	264/1.3	208/6.6[b]	192/12	413/0.14	81/100
Guanosine 5′-monophosphate-(TMS)$_6$ (**3**)	795/8.5	780/9.8	425/1.7	352/2.8	296/7.6	280/11	501/0.16	169/87
2′-Deoxyguanosine 5′-monophosphate-(TMS)$_5$ (**4**)	707/3.2	692/4.0	337/0.0	352/1.7	296/6.5	280/17	413/0.0	81/100
Cytidine 5′-monophosphate-(TMS)$_5$ (**5**)	683/0[d]	668/3.3	313/3.5	240/5.1	184/8.0	168/9.3	501/0.35	169/26
2′-Deoxycytidine 5′-monophosphate-(TMS)$_4$ (**6**)	595/0.5	580/0.6	225/0.0	240/12	184/18	168/29	413/0.0	81/100
Uridine 5′-monophosphate-(TMS)$_5$ (**7**)	684/0.0	669/0.9	314/0.0	241/3.2	185/2.9	169/5.6[c]	501/0.6	169/58[c]
Thymidine 5′-monophosphate-(TMS)$_3$ (**8**)	538/0.4	523/0.1	168/0.0	183/1.2[b]	127/0.0	111/0.0	413/0.15	81/100
2′-Deoxyuridine 5′-monophosphate-(TMS)$_4$ (**9**)	596/1.8	581/1.2	226/1.3	241/5.3	185/4.6	169/9.1	413/1.5	81/100
Adenosine 3′-monophosphate-(TMS)$_5$ (**11**)	707/0.4	692/6.3	337/2.7	264/2.4	208/12	192/9.5	501/7.5	169/10
Adenosine 3′,5′-cyclic monophosphate-(TMS)$_3$ (**13**)	545/4.9	530/36	337/25	264/3.5	208/37[b]	192/31	339/0.0	169/20

[a] Lesser amounts of other derivatives also present; see text. [b] Doublet; intensity uncorrected. [c] Doublet; intensity corrected. [d] M + H (*m/e* 684) from ion–molecule reactions usually observed; 0.69% relative intensity in mass spectrum used for this table.

Table II. Molecular Distribution and Location of TMS-d_9 in **1b**

Ion	Relative intensity[a] Unlabeled (*m/e*)	TMS-d_9	(TMS-d_9)$_2$
M (TMS)$_5$	50 (707)	40	10
M − 15 and M − 18 (TMS)$_5$	48 (692)	43	9
M − 15 and M − 18 (TMS)$_4$	50 (620)	41	8
k[b]	95 (192)	5	
c (TMS)$_3$	90 (466)	9	<1
c (TMS)$_2$	93 (394)	5	<2
o	96 (258)	2	2

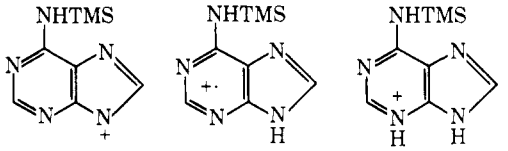
[a] ±30% error estimated for values below 10 due to low intensities on recorded mass spectrum. [b] Relative intensity corrected for contributions from ion i derived from (TMS)$_4$ molecular species.

yield a *m/e* 701:698 (CH_3:CD_3) ratio of 4:1. The observed value was 7.2:1, indicating that 1 in 3.6 methyls lost originate from the phosphate silyl group. This substantially lower tendency for participation of phosphate TMS was also found in the analogous AMP-(TMS-d_9)$_4$ species, in which the base was not silylated. The observed CH_3:CD_3 ratio (*m/e* 629:626) was 7:1, compared with a statistical availability of 3:1, assuming absence of secondary isotope effects.

An important and diagnostic ion c arises from loss of the phosphate moiety, formally leaving the charge on C-5′. The analogous ion is absent in the mass spectra of nucleoside TMS derivatives[7] (M − OTMS) as well as those of sugar phosphates[12] (M − OPO(OTMS)$_2$). A combination of factors therefore apparently plays a role, including the high stability of the phosphate radical which is lost. Examination of Corey–Pauling–Koltun (CPK) space-filling models[19] indicates that the minimum distance between N-3 of the purine or O-2 of pyrimidine bases and C-5′ is 2–2.5 1 Å (center to center), therefore permitting stabilization of a charge at 5′ by the base in either case by slight bending of the glycosidic bond. Interaction between the base and O-5′ in the mass spectra of adenine nucleosides has previously been demonstrated.[4]

(19) W. L. Koltun, *Biopolymers*, **3**, 665 (1965).

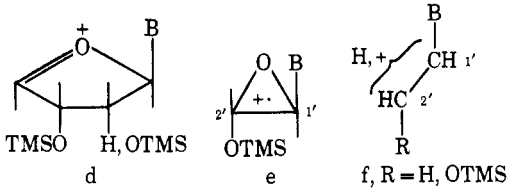
Fragment Ions of the Base Series. Resistance of the stable aromatic purine or pyrimidine residue toward fragmentation results in numerous ions which contain the intact base plus a portion of the sugar skeleton. The simplest of these is the base fragment itself (b) arising by simple cleavage of the glycosidic bond. As in the mass spectra of free nucleosides[4,20] b is usually accompanied by ions b + H and b + 2H which contain hydrogens rearranged from the neutral fragment which is lost. The structures can be represented as shown for adenine, but numerous other sites for attachment of the rearranged hydrogens are available. Mass spectra of the labeled compounds such as **1a** reveal that unlike the case of free nucleosides, the source of rearranged hydrogens is exclusively the ribose skeletal carbons.



b *m/e* 206 | b + H *m/e* 207 | b + 2H *m/e* 208

Simple cleavage of C-4′,5′ results in a minor ion, d, with charge stabilization provided by the 4′-ether oxygen. Of considerable diagnostic utility in the detection of ribotide sugar modification are two ions occurring 131 and 116 mass units above the base, b. Measurements of exact mass support the compositions base + $C_5H_{11}O_2Si$, and base + $C_5H_{12}OSi$, respectively, indicating the presence of C-1′ and -2′ and leading to representations e and f. Ion f includes a skeletal hydrogen rearranged from C-3′, 4′, or 5′, and is sometimes accompanied by its unrearranged counterpart one mass unit lower. In the spectra of 2′-deoxyribotide derivatives ion e (base + C_2H_3O) is absent, probably reflecting a role for O-2′ in the cleavage of C-2′,3′. As expected from the charge stabilizing role of O-2′, absence of the 2′-heteroatom reduces the occurrence of f in the deoxy

(20) K. Biemann and J. A. McCloskey, *J. Amer. Chem. Soc.*, **84**, 2005 (1962).

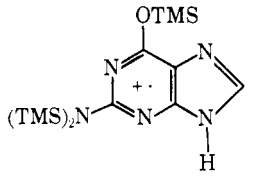


d e f, R = H, OTMS

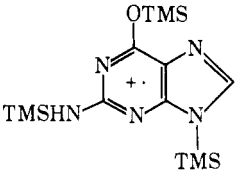
derivatives, **6** being the sole example, while the companion ion one mass unit lower corresponds to peaks in the spectra of **4** and **6**.

A small but characteristic peak (g) 16 mass units below f is evident in most spectra except the pyrimidine ribotides **5** and **7**. Deuterium labeling shows the f–g difference to include CH_3 from a TMS group, but no metastable peaks were observed to indicate its path of formation. Although elimination of methane from f is a logical possibility (*m/e* 322 → 306 in Figure 1), metastable ion evidence in the case of nucloside TMS derivatives[7] shows loss of a methyl radical, which would correspond to *m/e* 321 → 306 in Figure 1.

A useful ion in confirming the identity of the base arises from cleavage of the glycosidic bond with rearrangement of a TMS group to the base, resulting in a structure equivalent to h in the case of compounds **3** and **4** (Figure 4). The equivalent b + H ion resulting



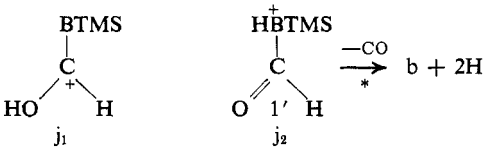
b + H (**3**, **4**)
m/e 367



h (**3**, **4**)
m/e 367
* −CH_3
i
m/e 352

from the presence of a higher derivative cannot be rigorously excluded in every case, but must not be a major contributor since under normal conditions of derivatization a higher derivative is not formed and no other supporting ions are observed. In some cases such as the UMP derivative (**7**, Figure 2), ion h is observed for the completely derivatized component ($(TMS)_5$), which militates against any other possibility except TMS migration, a process for which there is ample precedent.[21] Ion h is accompanied in most cases by ions one mass unit lower and higher (base + 72 and base + 74 mu). The latter ion is therefore analogous to b + 2H, with the rearranged hydrogen derived from the ribose skeleton. The spectrum of the deuterium-labeled compound **1a** (but not **2a** or **7a**) reveals that *m/e* 278 in Figure 1 principally contains 17 rather than 18 TMS hydrogens, indicating probable loss of a TMS hydrogen from ion h (*m/e* 279) after TMS migration. Further loss of a TMS methyl radical from h affords i, a characteristic ion found in the mass spectra of **1**–**9** which provides an additional means of identifying the base.

A general mode of fragmentation in free nucleosides produces a usually abundant and diagnostic ion consisting of the base plus C-1′ and O-4′ of the sugar.[4,20] In the mass spectra of nucleotide TMS derivatives the same ion type is observed, but numerous other available routes of fragmentation reduce its abundance to that of a minor ion. Either representation j_1[20] or j_2[4] appears valid, since most of the factors which govern its mechanism of formation in free nucleosides[4] (which support j_2) do not pertain to TMS derivatives. The pathway j → b + 2H which is common in most nucleoside spectra[4,7] is supported in the present study by an appropriate



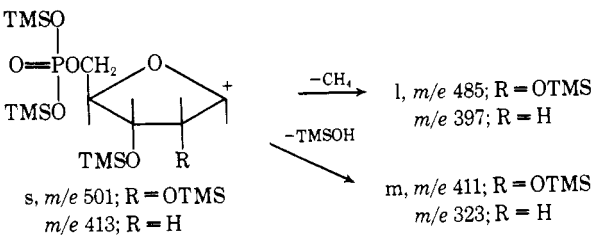
j_1 j_2

metastable peak only in the spectrum of **1**.

The b + H ion, which corresponds to the mass of a partially silylated free base, undergoes further decomposition by loss of a TMS methyl group. The resulting ion k is generally abundant as shown in Table I, and is therefore one of the best means for establishing the mass of the base from fragment ions in the mass spectrum.

Fragment Ions Derived from the Sugar Phosphate Moiety. Breakage of the glycosidic bond in the molecular ion leads to the sugar phosphate fragment (ion s) and a series of related ions which can be used to confirm the mass and identity of the sugar portion of the molecule. Separate assessment of the behavior of ribose- and phosphate-bound TMS groups in fragment ions of this series is possible in some cases using the partially labeled model **1b**. The principal ions used to determine the location and distribution of the label are listed in Table II. The molecular ion and M − 15 of both $(TMS)_4$ and $(TMS)_5$ species show **1b** to constitute about 40% of the total sample. Concentration of the label primarily in the phosphate is shown by the low abundance of TMS-d_9 species in ions c and k, and fragment o (to be discussed), which contains the ribose but not the base and phosphate moieties. A consideration of fragment ions which can arise from either **1** or **1b** requires that the ratio **1**:**1b** be taken into account to determine the statistical population of TMS-d_9 groups present in the total sample. When an ion of a given type exhibits different mass values in compounds **1** and **1b**, such as M − 15 discussed previously, the presence of **1** in the sample is of no consequence when considering ions of **1b**.

For the nucleotides from RNA unmodified by methylation at O-2′ ion s occurs at *m/e* 501, or at *m/e* 413 for monophosphate derivatives from DNA. A characteristic ion one mass unit lower was frequently observed in both series, and either ion was found in every case except the 2′-deoxyribotides **4** and **6**. However, elimination of methane from s involving a TMS methyl and hydrogen from ribose forms an ion (l) found in the spectra of all ribotides **1**–**9**. The low abundance of l precludes reliable determination of the preferred source of



s, *m/e* 501; R = OTMS
m/e 413; R = H

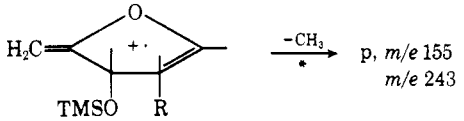
(21) See, for example, G. H. Draffan, R. N. Stillwell, and J. A. McCloskey, *Org. Mass Spectrom.*, **1**, 669 (1968), and references therein.

the methyl which is lost in the formation of this ion in **1b**.

Elimination of a molecule of trimethylsilanol from the sugar phosphate fragment s gives ion m, present in spectra of both ribotides and 2′-deoxyribotides, but of higher abundance in the latter. Deuterium labeling (**1a**) shows the source of hydrogen to be the ribose skeleton. To determine whether the source of TMSO is primarily the sugar or phosphate, the overall distribution of the label in **1b** must be considered. Taking into account the approximately 1:1 ratio of $(TMS)_4$ to $(TMS)_5$ species in the sample (which both yield ion s of *m/e* 510), and correcting the molecular distribution of TMS-d_9 for that present in the base (5%, Table II), the sugar phosphate distribution of the label is approximately 52% unlabeled, 40% TMS-d_9, and 8% $(TMS\text{-}d_9)_2$. In the formation of ion m, random elimination of TMSOH from the four possible sites would produce a *m/e* 411:*m/e* 420 ratio of 1.8 in the spectrum of the partially labeled derivative, where *m/e* 411 represents s − TMSOH from **1** and s − TMSOH-d_9 from **1b**, and *m/e* 420 is principally s − TMSOH from **1b**. No participation of the phosphate TMS groups would yield a ratio of approximately 1.3. The found value for *m/e* 411:420 was 1.1. Since 1.3 is in principle the minimum possible value the difference is attributed to errors in measuring the distribution of labeled species in the sugar phosphate fragment. The data indicate however that elimination of TMSOH from s proceeds largely from the ribose moiety, a result which is mechanistically most plausible because of the availability of ribose skeletal hydrogens.

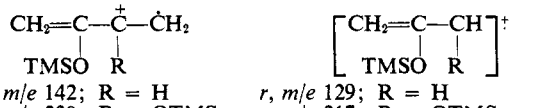
Spectra of ribotide derivatives (Figures 1 and 2) all show a peak 119 mass units lower than the sugar fragment, at *m/e* 382 (ion n). Measurement of exact mass and the spectrum of **1a** show the elements of TMSOH and COH to be involved, the latter logically C-1′ and O-4′. The calculated ratios for random availability of TMS for TMSOH:TMSOH-d_9 from **1b** are the same as for ion m. The observed ratio of 1.2 (*m/e* 383:392) reveals the greater participation of ribose TMS groups in the formation of n. The absence of this ion from the 2′-deoxy compounds (*e.g.*, Figure 3) suggests the involvement of 2′- rather than 3′-OTMS.

Several relatively abundant fragment ions[7,11,12] are observed in nucleotide–TMS mass spectra which contain none of the original skeletal atoms of the base or phosphate. Mass 258 (ion o) is generally prominent in ribotide spectra, and the analogous *m/e* 170 in the deoxy compounds (Figures 1–4). A shift for o of 18 mass units in the spectrum of **1a** and measurement of exact mass supports the composition $C_{11}H_{22}O_3Si_2$, *i.e*, two hydrogens less than the sugar fragment (C-1′ through -5′). Further loss of a TMS methyl radical produces ion p (*m/e* 243 or 155) as shown by metastable peaks in the spectra of **1–9**. Detailed structures for o and p cannot be postulated in the absence of knowledge as to which skeletal hydrogens are lost in the formation of o, although the latter will be isomeric with the structure shown.



o, R = H; *m/e* 170
R = OTMS; *m/e* 258

Extensive fragmentation and probably rearrangement of the ribose skeleton lead to *m/e* 230 and 217 in ribotide spectra, which can be given the generalized structures q and r. These ions, particularly the latter, which have been observed in the spectra of carbohydrate TMS ethers[11,22–28] and sugar phosphate derivatives,[12] have less abundant analogs in some of the deoxynucleotides (*m/e* 142 and 129, Figure 2). The presence of other unrelated peaks in the spectrum of **1b** prevents

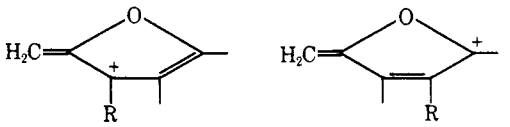


q, *m/e* 142; R = H
m/e 230; R = OTMS

r, *m/e* 129; R = H
m/e 217; R = OTMS
m/e 226; R = OTMS-d_9

conclusions regarding the identities of TMS groups in *m/e* 230. The ratio *m/e* 217:*m/e* 226 is greater than 10, showing the presence of little, if any, phosphate TMS.

Two ions of characteristically high abundance in the spectra of **1–9** are *m/e* 169 ($C_8H_{13}O_2Si$) from ribotides and *m/e* 81 (C_5H_5O) from deoxyribotides. These ions are isomeric with structures t, and serve as primary indicators to differentiate nucleotides of RNA and DNA. Both ions may be present in any given spectrum but with very large differences in abundance, as shown in Figures 1 and 2. The high abundance of t is derived from the considerable stabilization of the positive charge avail-



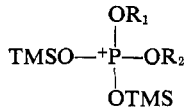
t, *m/e* 81; R = H
m/e 169; R = OTMS

able in either structure, involving the electrons of two double bonds and the 4′-oxygen. It is of interest that *m/e* 169 is essentially absent in the spectra of hexose-TMS derivatives reported by DeJongh and Sweeley,[11] and of all sugar phosphate derivatives studied by Sherman except that of β-D-ribose 5-phosphate.[12]

General Fragment Ions Associated with the Trimethylsilyl and Phosphate Groups. Numerous fragment ions are formed from the sugar phosphate portion of the molecule which are also found in the spectra of TMS derivatives of carbohydrates[11] and of sugar phosphates.[12] These ions are of generally little value for the determination of nucleotide structure, but represent a variety of interesting processes of basic importance in the mass spectrometry of phosphate- and trimethylsilyl-containing compounds. Of low abundance but perhaps the most interesting is *m/e* 387,[1] recently reported by Sherman.[12] This ion bears four TMS groups, two of which are rearranged from the sugar and base.[29] For

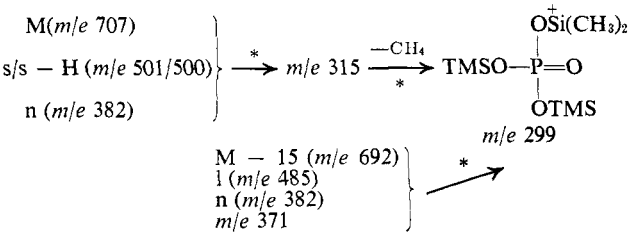
(22) O. S. Chizhov, N. V. Molodtsov, and N. K. Kochetkov, *Carbohyd. Res.*, **4**, 273 (1967).
(23) G. Petersson, O. Samuelson, K. Anjou, and E. von Sydow, *Acta Chem. Scand.*, **21**, 1251 (1967).
(24) N. K. Kochetkov, O. S. Chizhov, and N. V. Molodtsov, *Tetrahedron*, **24**, 5587 (1968).
(25) G. Petersson and O. Samuelson, *Sv. Paperstid.*, **71**, 77 (1968); *Chem. Abstr.*, **69**, 19444 (1968).
(26) H.-C. Curtius, J. A. Völlmin, and M. Müller, *Z. Anal. Chem.*, **243**, 341 (1968).
(27) J. Kärkkäinen and R. Vihko, *Carbohyd. Res.*, **10**, 113 (1969).
(28) J. Kärkkäinen, *ibid.*, **11**, 247 (1969).

nucleotides containing a total of four TMS groups (*e.g.*, **2**, Figure 3) both available TMS functions are transferred to the phosphate, while formation of the ion is blocked when only one TMS is available for transfer, as in thymidylic acid-$(TMS)_3$ (**8**).



m/e 387; $R_1 = R_2 =$ TMS
m/e 315; $R_1 =$ H; $R_2 =$ TMS
m/e 243; $R_1 = R_2 =$ H

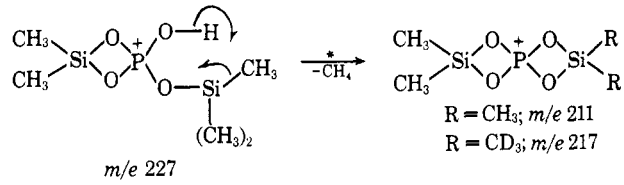
A ubiquitous peak characteristic of the phosphate TMS ester function was observed at *m/e* 315.[7,12] This usually prominent ion is structurally analogous to *m/e* 387 in that it contains both a TMS group and a non-TMS hydrogen abstracted from the remainder of the molecule. The metastable focussing technique[30] using compound **1** indicated *m/e* 315 to be derived from the molecular ion, ion s, and/or s − H, and the phosphate-containing fragment n. An abundant metastable ion in all spectra marked the further elimination of CH_4 involving the lone non-TMS hydrogen to yield *m/e* 299. Results of metastable focussing on *m/e* 299 revealed



other sources of *m/e* 299, most notably M − 15. The multiple sources of *m/e* 315 and 299 reflect the considerable stabilities of these ions, and indicate their probable general occurence in the mass spectra of phosphate-TMS esters.[31-33]

A minor ion of composition $C_6H_{20}O_4Si_2P$ occurs in most nucleotide spectra as the lower mass member of a doublet at *m/e* 243 with ion p ($C_{10}H_{19}O_2Si_2$), discussed previously. This ion contains two rearranged skeletal hydrogens and is structurally analogous to *m/e* 315 and 387. It should be noted that neither *m/e* 243 ion which we report here corresponds to that characteristically found in spectra of sugar phosphate TMS derivatives, which is a sugar fragment containing two intact TMS functions but no phosphate.[12]

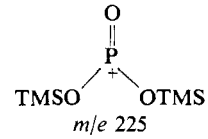
The interesting ion of mass 211[12] retains the four oxygens of the phosphate ester plus fragments of the two trimethylsilyl functions. In the spectrum of **1b**, *m/e* 211 shifts entirely to *m/e* 217; absence of +3 (*m/e* 214) and +9 (*m/e* 220) shifts demonstrates that the two TMS methyls which have been lost are from different silyl groups. Metastable ion evidence from most compounds studied shows *m/e* 211 to be derived from *m/e* 227, $C_5H_{16}O_4Si_2P$. The reaction can be envisioned as leading to the highly stabilized symmetrical *m/e* 211 as shown. An analogous transition is observed in the mass spectra of TMS sugar phosphates, which was



m/e 227

postulated to lead to an asymmetric form of *m/e* 211.[12]

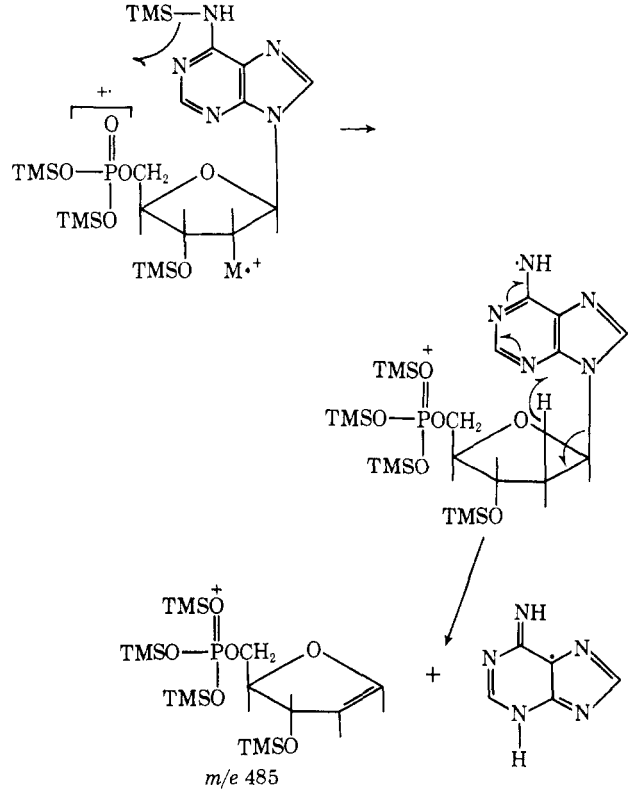
Measurements of exact mass and deuterium-labeling results show that the commonly observed ion *m/e* 225 corresponds to a fragment of the phosphate moiety, in which stabilization is provided by unshared electrons of three adjacent oxygens. The spectrum of **1b** suggests



m/e 225

the presence of two TMS groups originally bound to the phosphate in the molecular ion. The ion therefore does not evidently contain O-5′, which makes *m/e* 225 potentially useful for experiments using ^{18}O labels which require differentiation between O-5′ and the three phosphate oxygens.

The spectra of dAMP (**2**) and dGMP (**4**) derivatives show a prominent ion, *m/e* 485, not present in other deoxynucleotide spectra. Measurement of exact mass for **2** and a shift of 36 mass units in the spectrum of **2a** indicate that this ion consists of the sugar phosphate moiety which has added a TMS group and lost a ribose skeletal hydrogen. We envision a process in which the base donates a TMS function to the phosphate ester group in the molecular ion, followed by transfer of



m/e 485

(29) The occurrence of fragment ions containing two rearranged TMS functions is not without precedent. Two cases were reported by DeJongh, *et al.*,[11] including the symmetrical *m/e* 279: $^+C(OTMS)_3$.
(30) P. Schulze and A. L. Burlingame, *J. Chem. Phys.*, **49**, 4870 (1968), and references therein.
(31) M. Zinbo and W. R. Sherman, *Tetrahedron Lett.*, **33**, 2811 (1969).
(32) W. C. Butts, *Anal. Lett.*, 3, 29 (1970).
(33) C. B. Hirschberg, A. Kisic, and G. J. Schroepfer, Jr., *J. Biol. Chem.*, **245**, 3084 (1970).

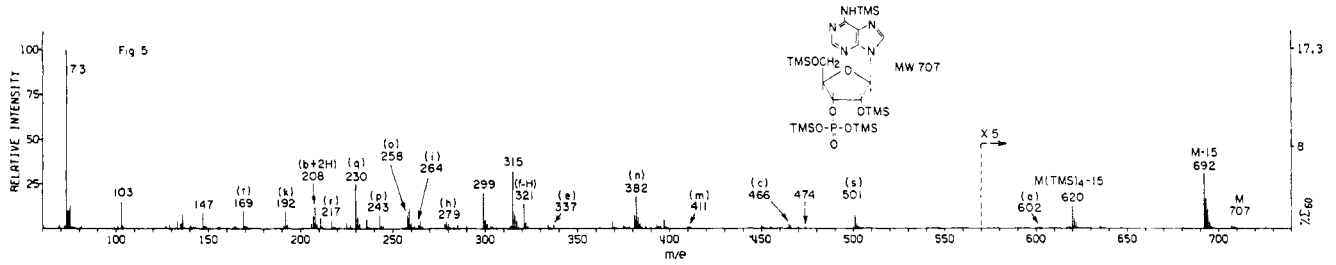
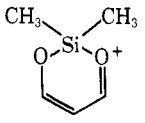


Figure 5. Mass spectrum of 3′-AMP-(TMS)$_5$ (**11**).

hydrogen to the base and breakage of the glycosidic bond. Since H-2′ comes in such close proximity to the base that it hinders free rotation about the glycosidic bond in CPK models,[19] C-2′ is considered a likely source of the rearranged hydrogen.

All spectra examined showed the ubiquitous *m/e* 73, 75, and 147.[17] The mass spectrum of **1b** revealed no marked preference for phosphate *vs.* sugar and base

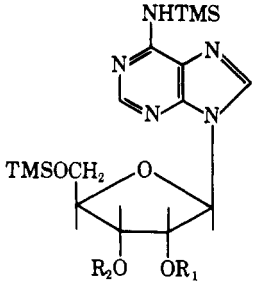
$(CH_3)_3\overset{+}{Si}$ *m/e* 73 $\quad$ $(CH_3)_2Si{=}\overset{+}{O}H$ *m/e* 75 $\quad$ $(CH_3)_3Si{-}\overset{+}{O}{=}Si(CH_3)_2$ *m/e* 147



m/e 129 $CH_2{=}CHCH{=}\overset{+}{O}TMS$

TMS functions in these ions. The minor peak at *m/e* 129, previously shown to be a doublet in the case of carbohydrate[11] and sugar phosphate TMS derivatives,[12] is present in nucleotide–TMS spectra. Both ions were generally found to be present, as shown by high-resolution spectra and deuterium-labeling results.

Derivatives from 2′- and 3′-Adenosine Monophosphate. Since some fragment ions serve to reflect structural modification at C-2′ and -5′, the mass spectra of 2′- or 3′-nucleoside phosphate derivatives should differ from each other and from compound **1**. Silylation of 2′- and 3′-AMP yielded derivatives (**10**, **11**)



10, R_1 = PO(OTMS)$_2$; R_2 = TMS
11, R_1 = TMS; R_2 = PO(OTMS)$_2$

which produced essentially the same mass spectra (Figure 5) but which were clearly different from that of the 5′ isomer (Figure 1).

Several experiments were carried out to determine whether the unexpected similarity of the spectra of **10** and **11** was due to isomerization during derivatization or phosphate migration after electron impact. Gas chromatography is inconclusive since **10** and **11** apparently exhibit similar retention times.[6b] Heating of 2′- and 3′-AMP in H_2O at 60° for 1 hr left both compounds unchanged, as determined by paper chromatography. Samples of 2′- and 3′-AMP were silylated at room temperature and their mass spectra obtained to confirm derivatization and then evaporated to dryness under nitrogen. The samples were hydrolyzed by heating at 60° in H_2O for 1 hr. Paper chromatography of the products clearly indicated no isomerization had occurred during silylation or hydrolysis. Silylation at elevated temperatures, such as 100°, was found to isomerize some 2′- to 3′-AMP, but caused only a trace of isomerization of the 3′ isomer. From these results we conclude that similarity of mass spectra of **10** and **11** is due to isomerization resulting from trimethylsilyl and phosphate migration between C-2′ and -3′ after electron impact, although isomerization during derivatization can also occur upon heating.

Although the mass spectra cannot distinguish phosphorylation at C-2′ *vs.* -3′, they are predictably different from that of the 5′-phosphate, **1**. As shown in Figure 5, the molecular ion and M − 15 are unchanged in mass compared with Figure 1, but ion a (M − 15 − TMSOH) is absent. Loss of the phosphate group gives an ion of type c, which is substantially reduced in abundance compared with Figure 1 (0.74% *vs.* 0.31% Σ_{40}) even though the resulting ion would bear the charge on a secondary carbon (C-2′ or -3′) rather than primary as in the case of **1**. Loss of the 5′ group as TMSO is likewise not observed as in the case of nucleosides,[7] supporting our earlier conclusion that formation of ion c from nucleotide–TMS derivatives depends on the identity of the radical which is lost as well as stabilization of the charge afforded by the base. The most important changes are reflected in ions e and f, which contain C-2′ but not C-3′. The ion e expected from a 3′-phosphate (*m/e* 337) is present, while that from a 2′ component (*m/e* 489) is not. However, in addition to the normal ion of type f (minus hydrogen, *m/e* 321), the rearranged 2′-phosphate analog is also observed at *m/e* 474. The presence of a primary TMS ether function at C-5′ is shown by the prominent ion of *m/e* 103 (5′-CH_2OTMS), which is characteristic in the mass spectra of nucleoside TMS derivatives. Ion t (*m/e* 169), which requires elimination of the 5′ group, is substantially reduced compared with Figure 1, similar to nucleoside derivatives which also lack the phosphate moiety at C-5′.

Effects from Chemical Opening of the Ribose Ring. Selective chemical modification of terminal nucleotide residues in polyribonucleotides is often carried out for the determination of end groups. One of the more successful approaches involves periodate oxidation of the 2′,3′ bond, followed by reduction of the resulting dialdehyde to a diol.[34] Since mass spectrometry is potentially useful for the identification of modified components of this nature, we have examined the spec-

(34) U. L. RajBhandary, *J. Biol. Chem.*, **243**, 556 (1968).

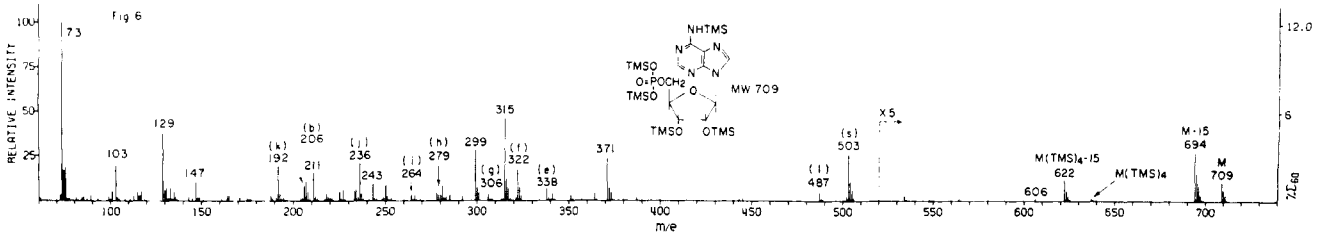


Figure 6. Mass spectrum of the TMS derivative of the ring-opened product of 5′-AMP (**12**).

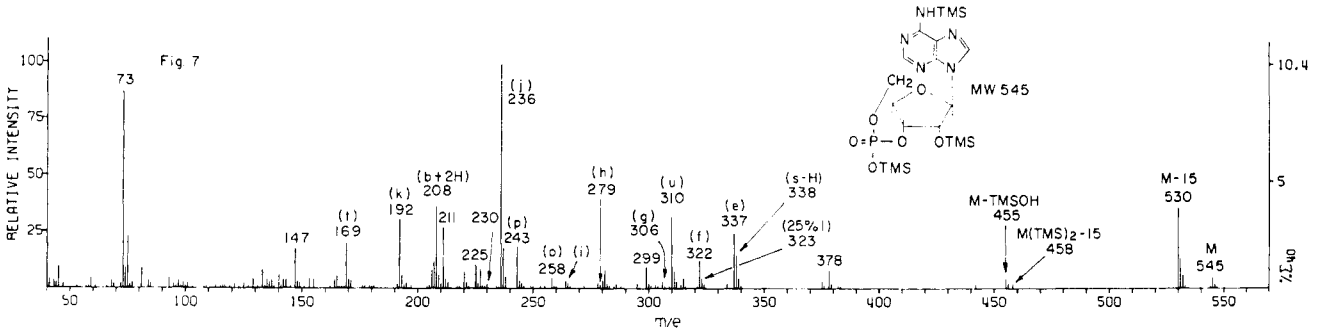
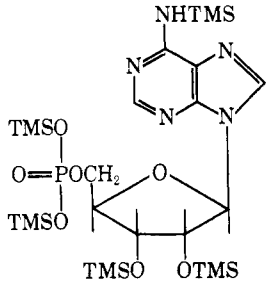


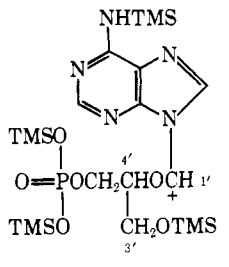
Figure 7. Mass spectrum of cyclic AMP-(TMS)$_3$ (**13**).

trum of the TMS derivative of the ring-opened product of 5′-AMP, **12** (Figure 6).

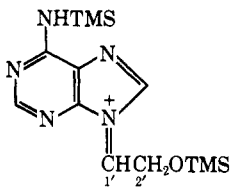


12

The net addition of two hydrogens is clearly shown by the shift of M, M − 15, and the sugar fragment s, to *m/e* 709, 694, and 503, in comparison with Figure 1. The most marked changes in the spectra are the absence of ions q (normally *m/e* 230) and t (*m/e* 169), as expected from the absence of the important 2′,3′ bond. Also blocked from formation is ion n (*m/e* 382, Figure 1), since CO cannot be expelled from C-1′ and O-4′ without additional loss of the C-2′ ether group. Less expected was the absence of ion c, evidently due to the greater conformational mobility of **12** compared with **1**. The availability of two primary TMS ether groups in **12** leads to the prominent *m/e* 103 (CH_2=O$^+$TMS), as well as the small but diagnostic *m/e* 606 which at least in part must come from loss of C-2′ with charge stabilization at C-1′ by O-4′ and N-9. Ion e, which as previously discussed contains C-1′ and -2′, is shifted one mass unit higher to *m/e* 338. Interestingly, the even-electron ion f is formed by simple cleavage and does not require hydrogen rearrangement, so that its mass is unshifted (*m/e* 322) from its cyclic counterpart **1**.



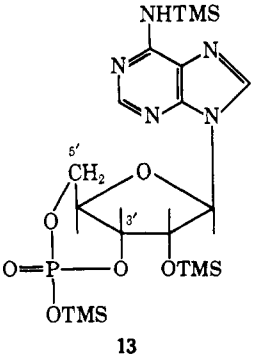
m/e 606



f, *m/e* 322

A prominent ion in Figure 6 is *m/e* 371, which is observed as a minor ion of composition $C_{11}H_{28}O_6Si_3P$ in most nucleotide spectra. This ion therefore probably contains the C-3′ through 5′ portion of the molecule (including O-4′), minus the elements of CH_4. Formulation of an ion structure cannot be attempted in the absence of knowledge as to which skeletal and TMS hydrogens are present.

The Mass Spectrum of the TMS Derivative of Adenosine 3′,5′-Cyclic Monophosphate. The spectrum of the cyclic AMP derivative **13** (Figure 7) retains all important structural features of nucleotide spectra but clearly reflects the cyclic nature of the molecule. Increased molecular ion stability due to cyclization is shown by the

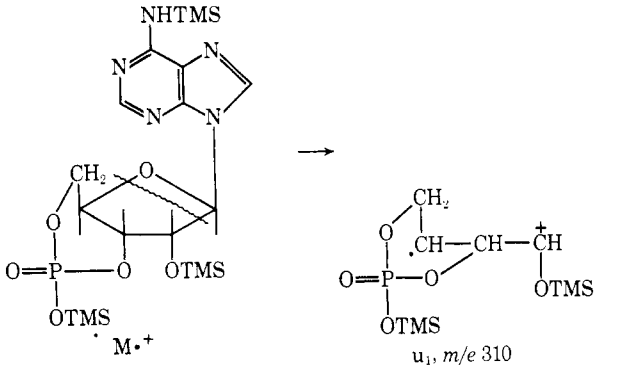


13

fivefold greater abundance of M and M − 15 compared with **1** (Figure 1). The phosphate-3′ bond precludes the presence of two TMS functions, thus lowering the molecular mass to 545. Ions c and d, which involve breakage of 5′-O and 4′–5′, respectively, are absent in Figure 7, because these fragmentations do not generate ions of lower mass in the cyclic molecule.

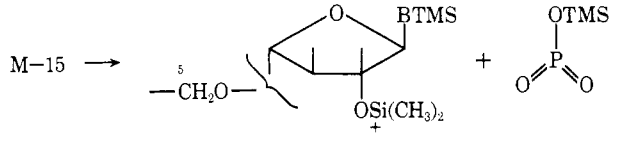
The cyclic structure also results in decreased abundances of some ions which must undergo more extensive and energy-consuming rearrangements to form. These include ion t (*m/e* 169), which must lose the entire cyclic phosphate moiety, and the 4-carbon fragment q (*m/e* 230) whose formation requires abstraction of TMS, presumably from phosphate. The presence of *m/e* 315 and its daughter ion *m/e* 299 reflects the considerable stability of these ions since they require migration of TMS from both the base and O-2′ to the phosphate group. Similarly, rearrangement processes lead to other well-stabilized phosphate ions which also occur in normal nucleotides: *m/e* 227, 211, 225, 258 (o), and both members of the *m/e* 243 doublet.

A peak at *m/e* 310 is observed in Figure 7 (ion u) which is absent in other nucleotide spectra. The composition $C_{10}H_{23}O_5Si_2P$ was obtained by measurement of exact mass after the deuterium-labeled TMS derivative showed the presence of two silyl groups. Its structure therefore must include carbons 2′ through 5′, which can be represented as u_1 or the rearranged highly conjugated



form u_2. Driving force for the formation of this ion can be attributed to either the release of C-1′,2′ and C-1′, O-4′ bonds, or the high stability of u_2.

An additional ion apparently unique to the 3′,5′-cyclic structure is *m/e* 378, $C_{15}H_{24}N_5O_3Si_2$. The mass spectrum of the TMS-d_9 derivative shows a shift for this ion of 15 mass units to *m/e* 392, requiring the presence of two silyl functions minus one methyl group. The ion is therefore related to M − 15 by loss of the elements of TMS metaphosphate from the cyclic moiety.



m/e 378

Experimental Section

Low-resolution mass spectra were recorded on an LKB 9000 instrument, with sample introduction through the gas chromatographic inlet; 1/4 in. diameter, 6 ft (glass) 1% OV-17 or 9 ft 1% SE-30, temperature programmed at 7.5°/min from 220°; carrier gas separator and ion source temperatures 250°; accelerating voltage 3.5 kV; ionizing energy 70 eV.

High-resolution mass spectra of **1**, **2**, **5**, **6**, **9**, and **13** were photographically recorded using a CEC 21-110B mass spectrometer, with sample introduction by direct inlet after removal of reagents and solvent in the direct inlet vacuum lock; probe and ion source temperatures 190–250°; accelerating voltage 6.2 or 8.3 kV; ionizing energy 70 eV.

The 5′-phosphates **1**–**9** and 3′,5′-cyclic AMP (**13**) were purchased from Calbiochem, Los Angeles, Calif., and Sigma Chemical Co., St. Louis, Mo., as either acids, Na_2, or $(NH_4)_2$ salts. Compounds were checked for purity by gas chromatography of their TMS derivatives.[4,5a,b] 2′- and 3′-AMP (**10** and **11**) were obtained from Schwarz Bioresearch, Inc., Mount Vernon, N. Y. Freedom of each isomer from contamination by the other was demonstrated by chromatography on Eastman chromatography sheets with saturated $(NH_4)_2SO_4$–1 *M* NaOAc–isopropyl alcohol (40:9:1). Paper chromatography of **10** and **11** was carried out on grade no. 54 Whatman paper, developed with the same reagents.

The 2′,3′ ring-opened derivative of 5′-AMP (12) was prepared by a modification of the procedure of RajBhandary for tritium labeling.[34] **1** (25 μg) was mixed with 200 μl of 0.01 *M* NaOAc and 10 μl of 0.1 *M* $NaIO_4$ and allowed to stand 30 min in the dark. To the resulting solution containing the 2′,3′-dialdehyde was added 10 μl of 0.15 *M* ethylene glycol, followed after 20 min by 10 μl of 0.2 *M* KBH_4. After 30 min the solution was evaporated under vacuum, dried over P_2O_5, and trimethylsilylated to yield **12**. The mass spectrum of **12** (Figure 6) showed the expected molecular mass of 709 and was free of peaks due to **1** or the intermediate dialdehyde.

N,O-Bis(trimethylsilyl)trifluoroacetamide (BSTFA) and trimethylchlorosilane (TMCS) were purchased from Pierce Chemical Co., Rockford, Ill., and Regis Chemical Co., Chicago, Ill. *N,O*-bis(trimethylsilyl)acetamide-d_{18} (BSA-d_{18}) and TMCS-d_9 were purchased from Merck Sharp and Dohme of Montreal, Canada; early experiments used BSA-d_{18} prepared in our laboratory.[15]

Preparation of Trimethylsilyl Derivatives. The most generally satisfactory procedure for all nucleotides except those of cytidine (**5**, **6**) and cyclic AMP (**13**) consisted of a solution of 30 μg of the dried acid or its salt in pyridine (30 μl), BSTFA (30 μl), and TMCS (1 μl), followed by standing for 30 min at room temperature or at 150°, in a tightly sealed screw-cap vial. After opening and resealing, the samples remained good for days and sometimes weeks, with slow conversion to the next lower derivative. Reheating would not generate the fully silylated nucleotide.

Optimum conditions for preparation of **5** and **6** consisted of heating the acid or its salt (200 μg) with BSA (35 μl) and TMCS (5 μl) in acetonitrile (16 μl) at 70° for 60–90 min. The yield was 20–30% of the other major nucleotides treated under the same conditions. An alternative procedure[7] using BSA and TMCS without acetonitrile gave the $(TMS)_4$ derivative of 5′-CMP, which could be introduced into the mass spectrometer by probe but not by gas chromatograph.

Adenosine 3′,5′-cyclic monophosphate-$(TMS)_3$ (**13**) was formed by solution of 50 μg of the acid as its disodium or diammonium salt in pyridine–BSTFA–TMCS (9:90:1), followed by heating in a sealed vial at 100° for 3 hr, or until little or no $(TMS)_2$ derivative remained.

The TMS-d_9 derivatives **1a**, **2a**, and **7a** were prepared by dissolving the parent acid or salt in pyridine (10 μl), BSA-d_{18} (10 μl), and TMCS-d_9 (0.5 μl) in a melting point capillary which was then sealed and allowed to stand 30 min or heated gently.

1b was formed by an exchange reaction in the gas chromatograph coupled to the (LKB) mass spectrometer. About 5 μg of **1** was injected into a 3-ft 1% SE-30 column (170°) which had been pretreated by repeated injections of BSA-d_{18} and TMCS-d_9 (20:1); the column was temperature programmed at 7.5°/min, **1b** eluting with unchanged **1** at 232°. See Table II for distribution of the label.

Acknowledgments. The work was supported by the National Institutes of Health (GM 13901) and the Robert A. Welch Foundation (Q-125). Computer facilities were provided by the Baylor College of Medicine Computer Facility (NIH Grant No. RR 259) and the Common Research Computer Facility of the Texas Medical Center (NIH RR 254). A. M. L. received postdoctoral support from the Robert A. Welch Foundation (1966–1968). We are grateful to Mr. K. J. Lyman and Miss P. F. Crain for technical assistance, and to Dr. E. White for helpful discussions.