Characterization of Blocked Nucleosides by Mass Spectrometry. Benzyl Derivatives (1)

H. T. Cory, K. Yamaizumi, David L. Smith, D. R. Knowles, Arthur D. Broom and James A. McCloskey

Department of Medicinal Chemistry, University of Utah, Salt Lake City, Utah 84112

Received September 22, 1978

The mass spectra of eleven model monobenzylated nucleosides were studied using low and high resolution mass spectrometry. Structural assignments to the major ions were made and several decomposition mechanisms proposed, with the goal of establishing the uses and limitations of mass spectrometry for the characterization of benzylated nucleosides. Mass spectra generally permit determination of the extent and site of benzylation, with particular regard to base *vs.* sugar substitution, *O*-2′ *vs.* *O*-3′ or *O*-5′, and in some cases *O*-5′ *vs.* other isomers.

J. Heterocyclic Chem., **16**, 585 (1979).

Earlier studies have shown the usefulness of mass spectrometry for characterization of unblocked nucleosides (2-4), providing their polarity is sufficiently low for thermal vaporization without significant decomposition (5).

Detailed studies of the mass spectra of a number of nucleoside derivatives have been reported, including for example, trimethylsilyl (9) and other alkylsilyl (10), trifluoroacetyl (11,12), and *N*,*O*-permethyl derivatives (13) of nucleosides, and of trimethylsilyl (14-17) and other derivatives (16,17) of anhydronucleosides. In addition, shorter surveys of the mass spectrum of several other derivatives have appeared (4,18). However, aside from the work by Westmore and his collaborators on alkylsilyl derivatives (10), little systematic attention has been paid to the mass spectra of blocked nucleosides commonly used in synthesis. In the present paper we report the results of a study of the mass spectra of benzyl derivatives of nucleosides, with a goal of establishing the uses and limitations of mass spectrometry for their structural characterization.

The use of the benzyl ether for protection of the ribonucleoside 2′-hydroxyl group prior to oligonucleotide synthesis was first proposed by Khorana (19). The use of 2′-*O*-benzyluridine in the synthesis of UpU was described by Reese (20). Since that time, a number of groups (21-24) have explored various approaches to the synthesis of benzylated ribonucleosides. The most general approach has proved to be a stannous chloride-catalyzed benzylation of unprotected ribonucleosides with phenyldiazomethane, a technique which gives reasonable yields of both 2′- and 3′-*O*-benzyl nucleosides (24).

Previous studies of the mass spectrum of N^6-benzyladenine have been made by Shannon and Letham (25), and Leonard and Henderson (26), and of N^6-benzyladenosine by Robins and Trip (27).

Of the numerous combinations of benzylated nucleosides that could be examined, monosubstituted derivatives of uridine, cytidine, adenosine, guanosine and inosine were chosen as the principal models.

Results and Discussion of Mass Spectra.

Mass spectra of the four isomeric benzylated adenosines and of 2′-*O*-benzylcytidine are shown in Figures 1-5, and of the remaining compounds studied by electron impact in Table 1. The field desorption mass spectrum of 3′-*O*-benzylinosine is represented in Figure 6.

Determination of Molecular Weight.

In all cases of electron impact spectra, the abundances of molecular ions (M) are exceptionally low due to the favored formation of the highly stable tropylium ion, m/e 91, and to a lesser extent the McLafferty rearrangement

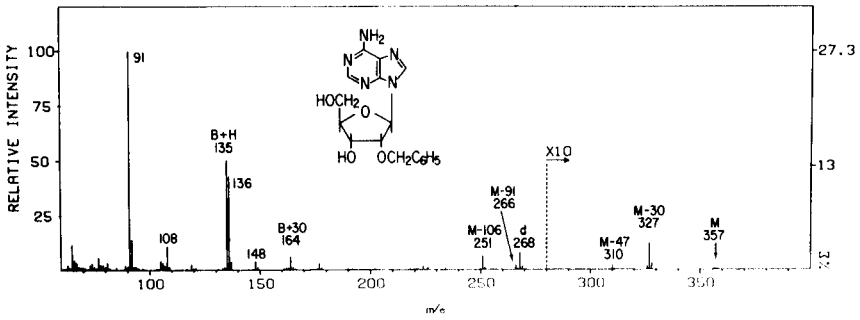


Figure 1. Mass Spectrum of 2′-*O*-Benzyladenosine

0022-152X/79/030585-06$02.25

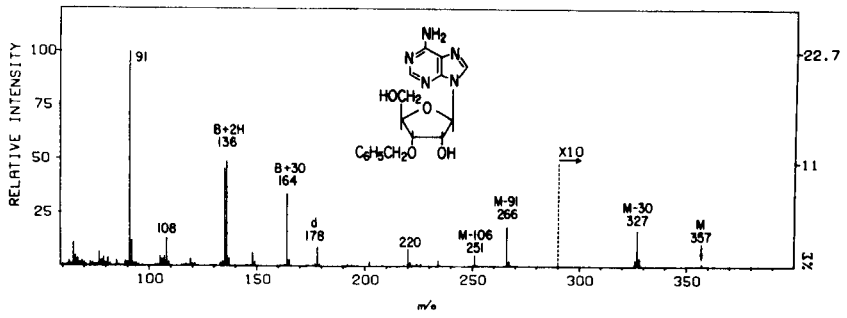


Figure 2. Mass Spectrum of 3′-*O*-Benzyladenosine

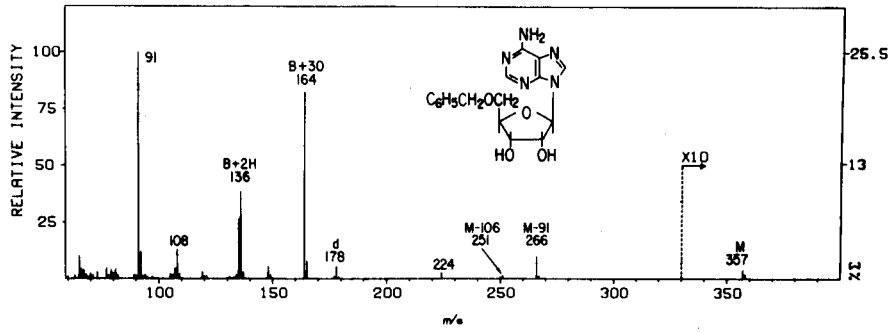


Figure 3. Mass Spectrum of 5′-*O*-Benzyladenosine

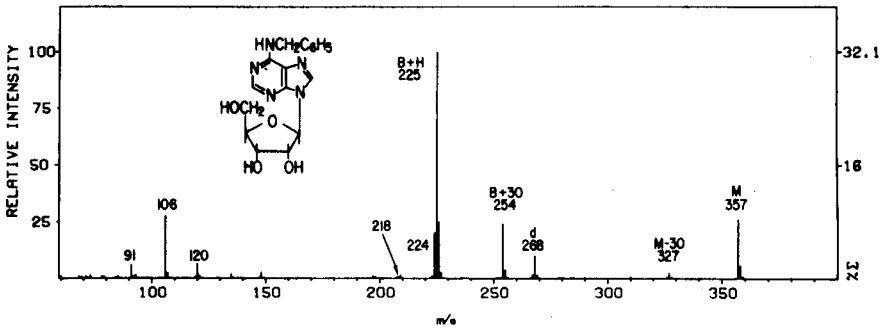


Figure 4. Mass Spectrum of N^6-Benzyladenosine

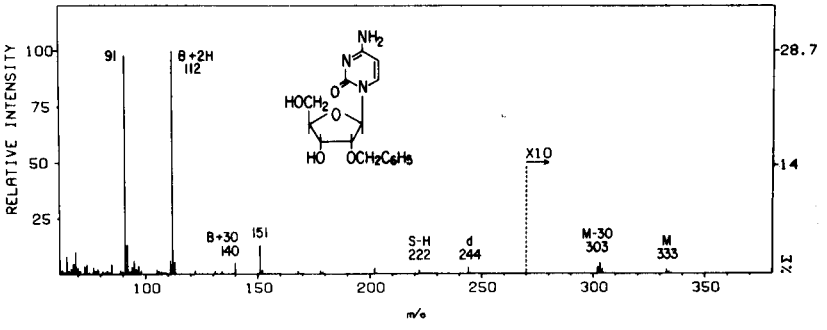


Figure 5. Mass Spectrum of 2′-*O*-Benzylcytidine

product (C_7H_8), m/e 92 (28). Molecular ions were observed in every case but their detection could be ambiguous if samples are insufficiently pure. The least favorable cases of the compounds studied were 3′-*O*-benzylguanosine and 3′-*O*-benzylinosine (Table 1). If necessary the molecular weight can be established, or an

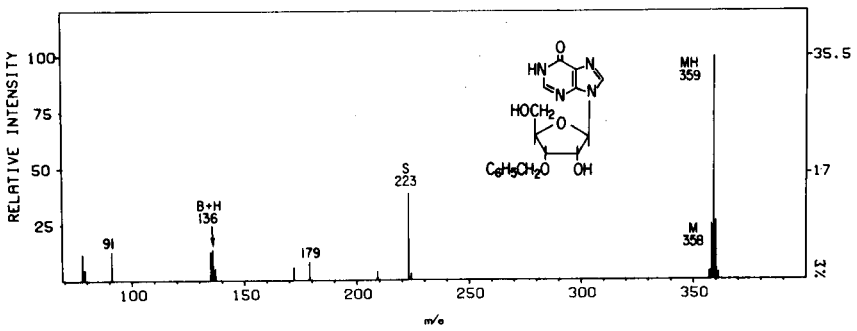
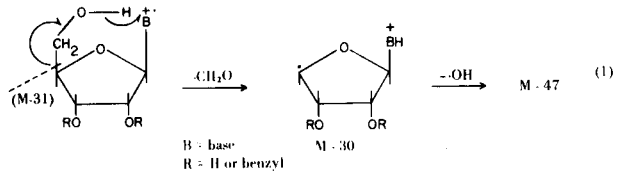


Figure 6. Field Desorption Mass Spectrum of 3′-*O*-Benzylinosine

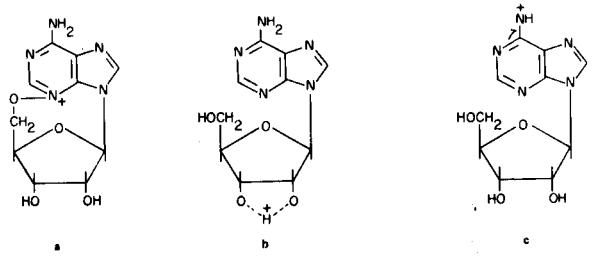
electron impact-derived value corroborated, by more gentle means of ionization, such as chemical ionization (29), field ionization (30), or field desorption (31). A typical example in the present study is illustrated by the field desorption mass spectrum of 3′-*O*-benzylinosine in Figure 6. Peaks in the molecular ion region are greatly enhanced relative to the electron impact spectrum (M^+ = 0.1%) although a certain degree of subjectivity is involved in assignment of M^+ *vs.* MH^+, not an uncommon problem when using the field desorption technique. In the present case the sugar fragment, m/e 223, dictates the molecular ion to be m/e 258. A small peak at m/e 179, and its isotope peak at 179.5, represents the doubly charged molecular ion. The assignment can also be made on the basis of the nitrogen rule (32), assuming the nitrogen content of the base is known, or by a characteristic $(M + Na)^+$ ion when traces of sodium are present.

Using electron impact ionization, the molecular weight can effectively be established by a series of three ions that are directly related to the molecular ion and were observed to be present in most cases. Loss of CH_2O or of CH_2OH from the 5′ moiety follows the behavior of free nucleosides, giving peaks in the spectrum corresponding to M-30 or M-31 (2). The latter process occurs by simple cleavage while M-30 formation has been proposed to involve hydrogen transfer to the charged base (33) (Equation 1). Following the behavior of free nucleosides,



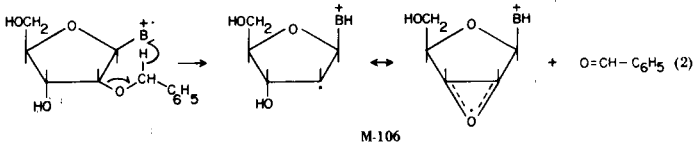
greater predominance of charge in the purine *vs.* pyrimidine nucleus leads to greater tendency to form M-30 in the purine derivatives and M-31 in pyrimidines. In some cases (N^6-benzyladenosine, 2′-*O*-benzyladenosine) further loss of a hydroxyl radical from M-30 results in a small peak at M-47. Presence of M-30 or M-31 peaks gives a facile means to establish that *O*-5′ is blocked, while their absence has no significance.

A second general fragmentation process related to the molecular ion is produced by loss of a benzyl group (91 a.m.u.) observed in all of the compounds studied except 2′-*O*-benzylcytidine and 2′-*O*-benzylguanosine. Although some driving force for the reaction may be the stability of the neutral benzyl radical, the resulting charge on oxygen is not well stablized. Comparison of M-91 peak intensities in the adenosine series (Figures 1-4) show the base-substituted isomer to be lowest, which suggests that stablization of the forms shown as **a** and **b** may be important. Stabilization of the type implied by structure **c**, as would be generated from N^6-benzyladenosine, is



not expected to be important, based on the low abundance of this ion in the mass spectra of N^6-alkylated adenosines (34,27), and absence from N^6-benzyladenine (25).

A third member of the molecular ion series produced from nine of the twelve models studied, is M-106, shown by exact mass measurements to be due to loss of 106.0418, or C_7H_6O. The prominence of M-106 in the spectra of 2′- and 3′-*O*-benzyladenosine but not that of the N^6-isomer (27) leads to the reaction shown as Equation 2 as an attractive mechanism of formation. Space-filling CPK



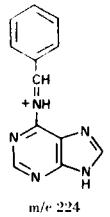
models show that in the 3′-substituted derivatives sugar ring opening must occur in order for methylene hydrogens to come within bonding distance of the base. In either

case transfer of an active benzylic hydrogen to the charge-localized base follows a rationale similar to that proposed for transfer of hydroxyl hydrogens in free nucleosides (33) and of a methyl hydrogen in the case of 2′-*O*-methyladenosine (13).

In the adenosine series the small peak at m/e 244 is derived from the m/e 251 species by loss of HCN as shown by measurement of exact mass (224.0942, fd.) (35). The process is therefore analogous to the formation of m/e 109 from protonated adenine (m/e 136) in the mass spectrum of adenosine (33).

Determination of the Site of Benzylation.

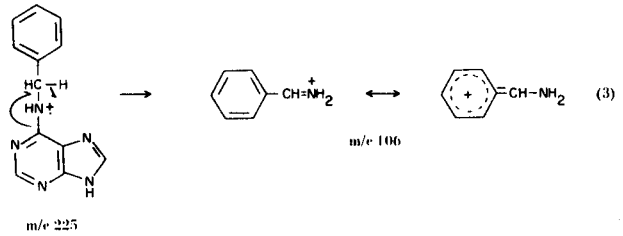
The location of substituents in base *vs.* sugar is most directly made through the ubiquitous base (B) + H and B + 2H ions which are generated by extraction of (preferentially) labile, or skeleton sugar hydrogens by the base. The B + 2H species is generally formed from B + CH_2O precursors (2,33) described below. In the present study B + 2H tends to predominate in the case of pyrimidines while B + H is relatively more intense in the spectra of most purines, following a similar trend in unblocked nucleosides (4). In adenine-containing derivatives further decomposition of B + H (m/e 135) leads to m/e 108 by loss of HCN (25), or to m/e 218 in Figure 4. The ion of nominal mass 108 also corresponds to benzyl alcohol (mass 108.0575) which might be formed through an elimination reaction. However, the two possibilities are readily distinguished at R > 8,000 (C_7H_8O, 108.0575 *vs.* $C_4H_4N_4$, 108.0436). Measurement of exact mass confirms the peak to correspond to B + H - HCN to the extent of 84-92% in Figures 1-3, the remainder being C_7H_8O. The ions of m/e 120 (Figures 1-4) are products of the adenine nucleus: $C_5N_4H_4$ (25,36). The unusual abundance of the m/e 224 ion, corresponding to the base fragment, is notable in Figure 4. Ions due to the base fragment are usually unimportant in nucleoside mass spectra (*cf.*, Figures 1-3). In the present case the ion in question is attributed to the structure shown below, in



m/e 224

correspondence with an analogous peak (M - H) and assignment in the spectrum of N^6 benzyladenine (25). The m/e 224 species is in part derived from the B + H ion, as shown by a metastable peak at m/e 223.0.

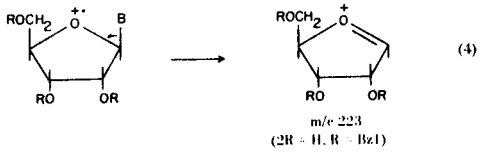
The intense peak at m/e 106 in the mass spectrum of N^6-benzyladenosine (25,27) is due to the highly stabilized aminobenzyl cation. Its mechanistic precursor was established as m/e 225 through the metastable ion observed by scanning the accelerating voltage with fixed electric sector voltage and fixed magnetic field. The formation of the mass 106 ion is rationalized in Equation 3. Recent



m/e 225 m/e 106 (3)

work by Leonard and Henderson involving ^{15}N labeling has shown that the major population (87%) of m/e 106 ions results from direct cleavage of the C-6 - N^6-bond, while a smaller population (13%) contains atom N-1, therefore formed by a mechanistically more complex process (26).

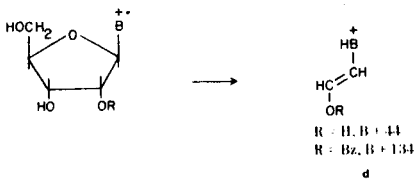
The number of substituents on the sugar is indicated by the sugar ion (S, m/e 223) or sugar-H (S-H) ions formed by direct cleavage of the glycosidic bond (Equation 4).



m/e 223 (2R = H, R = Bzl) (4)

As seen in Figures 1-4 and Table 1, these peaks were not present in the adenosine and several other derivatives, but is prominent in the field desorption spectrum shown in Figure 6. When observed, S-H is more likely to be formed than the S species when *O*-2′ is substituted. The same effect was noted in the mass spectra of 2′-*O*-methylated nucleosides (33,37). A number of the spectra examined exhibit a peak of m/e 191, which was shown by the high resolution spectrum of 3′-*O*-benzylguanosine (191.0708, fd.) to correspond to loss of CH_2OH from m/e 222 (S-H). The ion is absent in the spectrum of 5′-*O*-benzyladenosine, in support of *C*-5′ as the group which is lost.

Placement of substituents on *O*-2′ can be made by presence or absence of peaks corresponding to the ion shown below as *d*. The diagnostic ion *d* occurs widely in

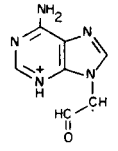


R = H, B + 44
R = Bz, B + 134
d

nucleoside mass spectra, and was observed in every spectrum in the present study. This observation is of particular importance since it offers a facile means of unambiguously determining whether a single isomer obtained from a reaction is the generally more important

2′-*O*-benzyl derivative. Although this may be accomplished by uv or nmr techniques, the former requires both 2′ and 3′ (or 5′) isomers and the latter much larger samples (24).

In the spectrum of 2′-*O*-benzyladenosine (Figure 1) a peak related to ion **d** occurs, which from measurement of exact mass (177.0652, fd.) has an elemental composition in support of structure **e**. Examination of the high resolution mass spectrum of 2′-*O*-methyladenosine shows an ion of the same exact mass (38) thus supporting the structure of **e**.



e, m/e 177

f, B + 30

Of less importance for placement of groups in the sugar is the B + 30 ion, whose structure can be represented as **f** (33) consisting of the base plus *C*-1′, *O*-4′ and rearranged hydrogen. Intensity of the B + 30 ion varies widely, as shown by comparison of Figures 2 and 3, but along with B + H and B + 2H it can be used effectively to detect impurities due to base-substituted components. The peak at m/e 148 in the adenosine series corresponds to the base plus CH_2 (33), presumably from *C*-1′.

EXPERIMENTAL

The 2′- and 3′-*O*-benzyl derivatives of adenosine, cytidine, guanosine, inosine and uridine were prepared as previously described (24). 5′-*O*-Benzyladenosine was a gift from Dr. W. Pfleiderer, University of Kontanz; its synthesis was earlier reported (39). N^6-Benzyladenosine was purchased from Heterocyclic Chemical Corp., Harrisonville, Missouri.

Mass spectra shown in the Table and Figures 1-5 were recorded using an LKB 9000S instrument with samples introduced by direct probe; ionizing energy 70 eV, ion source temperature 270°. Exact mass and metastable ion measurements were made with a Varian MAT 731 double focusing mass spectrometer.

The field desorption mass spectrum of 3′-*O*-benzylinosine was acquired using the Varian instrument. The tungsten emitter was prepared from benzonitrile in the conventional fashion according to the manufacturers instructions; desorption occurred in the range 12-14 ma of emitter current.

Table I

Principal Ions from the Mass Spectra of Benzylated Nucleosides

m/e/Relative Intensity

Compound	M	M-30 or M-31	M-91	M-106	B + 1	B + 2	B + 30	d	S or S-H	m/e 191	% Σ%/ RI
2′-*O*-Benzyluridine	334/1.4	303/0.2	243/0.5	228/0.9	112/15	113/77	141/12.5	245/0.4	222/4.9	-	0.42
3′-*O*-Benzyluridine	334/0.2	303/0.2	243/2.0	228/-	112/45	113/41	141/7.0	155/7.0	223/6.0	5.5	0.37
3′-*O*-Benzylcytidine	333/-	303/-	242/0.8	227/2.0	111/11	112/45	140/7.0	154/1.5	-	0.48	0.35
2′-*O*-Benzylguanosine	373/0.5	343/0.2	282/-	266/0.5	151/18	152/4.0	180/-	284/0.8	223/1.5	-	0.35
3′-*O*-Benzylguanosine	373/0.1	343/-	282/0.2	267/-.	151/18	152/2.0	180/0.5	-	223/0.5	0.9	0.38
2′-*O*-Benzylinosine	358/0.2	328/0.2	267/0.3	252/1.3	136/16.5	137/16.5	165/1.3	269/0.8	222/0.7	1.0	0.42
3′-*O*-Benzylinosine	358/0.1	328/-	267/0.2	252/0.6	136/38	137/8.0	165/1.6	179/1.0	-	1.0	0.31

REFERENCES AND NOTES

(1) The authors acknowledge generous support from the National Cancer Institute, U. S. Public Health Service, through grants CA 11935 (A.D.B.) and CA 18024 (J.A.M.).

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

(3) C. Hignite, in "Biochemical Applications of Mass Spectrometry," G. R. Waller, Ed., Wiley-Interscience, New York, N. Y., 1972, chapter 16.

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

(5) Mass spectra of many nucleosides that are involatile by conventional vaporization can also be obtained by field desorption (6), ^{252}Cf (plasma) desorption (7), and vaporization from activated emitters prepared for field desorption, under chemical ionization conditions (8).

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

(7) R. D. Macfarlane and D. F. Torgerson, Science, **191**, 920 (1976).

(8) D. F. Hunt, J. Shabanowitz and F. K. Botz, *Anal. Chem.*, 49, 1160 (1977).

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

(10) K. K. Ogilvie, S. L. Beaucage, D. W. Entwistle, E. A. Thompson, M. A. Quilliam and J. B. Westmore, *J. Carbohydro. Nucleosides, Nucleotides*, 3, 197 (1976).

(11) W. A. König, L. C. Smith, P. F. Crain and J. A. McCloskey, *Biochemistry*, **10**, 3968 (1971).

(12) W. A. König, K. Zech, R. Uhmann and W. Voelter, *Chem. Ber.*, **105**, 262 (1972).

(13) D. L. von Minden and J. A. McCloskey, *J. Am. Chem. Soc.*, 95, 7480 (1973).

(14) S. Tsuboyama and J. A. McCloskey, *J. Org. Chem.*, 37, 166 (1972).

(15) D. Lipkin and J. A. Rabi, *J. Am. Chem. Soc.*, 93, 3309 (1971).

(16) J. B. Westmore, D. C. K. Lin, K. K. Ogilvie, H. Wayborn and J. Berestiansky, *Org. Mass Spectrom.*, **6**, 1234 (1972).

(17) D. C. K. Lin, L. Slotin, K. K. Ogilvie and J. B. Westmore, *J. Org. Chem.*, **38**, 1118 (1973).

(18) J. J. Dolhun and J. L. Wiebers, *Org. Mass Spectrom.*, **3**, 669 (1970).

(19) M. Smith, D. H. Rammler, I. H. Golderg and H. G. Khorana, *J. Am. Chem. Soc.*, **84**, 430 (1962).

(20) B. E. Griffin, C. B. Reese, G. F. Stevenson and D. R. Trentham, *Tetrahedron Letters*, 4349 (1966).

(21) K. Kikugawa, F. Sato, T. Tsurao, N. Imura and T. Ukita, *Chem. Pharm. Bull.*, **16**, 1110 (1968).

(22) H. U. Blank, D. Frahne, A. Myles and W. Pfleiderer, *Ann. Chem.*, 742, 1, 16 (1970).

(23) W. Hytzenlareb and W. Pfleiderer, *Chem. Ber.*, **106**, 665 (1973).

(24) L. F. Christensen and A. D. Broom, *J. Org. Chem.*, **37**, 3498 (1972).

(25) J. S. Shannon and D. S. Letham, *New Zealand J. Sci.*, **9**, 833 (1966).

(26) N. J. Leonard and T. R. Henderson, *J. Am. Chem. Soc.*, 97, 4990 (1975).

(27) M. J. Robins and E. M. Trip, *Biochemistry*, **13**, 2179 (1973).

(28) H. Budzikiewicz, C. Djerassi and D. H. Williams, "Mass Spectrometry of Organic Compounds," Holden-Day, San Francisco, Ca., 1967, p. 247.

(29) M. S. Wilson and J. A. McCloskey, *J. Am. Chem. Soc.*, 97, 3436 (1975).

(30) P. Brown, G. R. Pettit and R. K. Robins, *Org. Mass Spectrom.*, 2, 521 (1969).

(31) H. R. Schulten and H. D. Beckey, *ibid.*, 7, 861 (1973).

(32) Molecules having odd numbers of nitrogen atoms have odd molecular weights, while those with zero or even numbers of nitrogens have even molecular weights.

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

(34) S. M Hecht, A. S. Gupta and N. J. Leonard, *Anal. Biochem.*, 30, 249 (1969).

(35) The observed exact mass value also permits a second reasonable composition, $C_{12}H_{10}N_5$, which corresponds to the base fragment from N^6-benzyladenosine (see Figure 4) which could be present as an impurity. This latter possibility was excluded by the absence of a more abundant m/e 225 ion as shown in Figure 4 as well as repetitive scanning experiments that suggested sample homogeniety.

(36) J. M. Rice and G. O. Dudek, *J. Am. Chem. Soc.*, **89**, 2719 (1967).

(37) H. A. Howlett, M. W. Johnson, A. R. Trim, J. Eagles and R. Self, *Anal. Biochem.*, 39, 429 (1971).

(38) D. M. Desiderio and J. A. McCloskey, unpublished results.

(39) A. Myles and W. Pfleiderer, *Chem. Ber.*, **105**, 3327 (1972).