Biophysical Chemistry 17 (1983) 285–295
Elsevier

STRUCTURE OF THE PLANAR COMPLEX OF N^4-METHOXYCYTOSINE WITH ADENINE, AND ITS RELEVANCE TO THE MECHANISM OF HYDROXYLAMINE MUTAGENESIS

Borys KIERDASZUK and David SHUGAR

Department of Biophysics, Institute of Experimental Physics, University of Warsaw, 93 Zwirki Wigury St., 02-089 Warszawa, and Institute of Biochemistry and Biophysics, Academy of Sciences, 36 Rakowiecka St., 02-532 Warszawa, Poland

Received 12th August 1982
Revised manuscript received 16th December 1982
Accepted 3rd January 1983

Key words: Hydroxylamine mutagenesis; N^4-Methoxycytosine; Base-pairing; NMR; Infrared spectroscopy

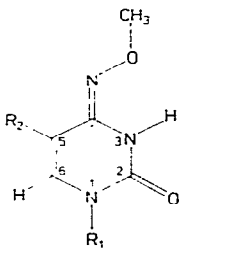
Infrared, and ^{1}H- and ^{13}C-NMR spectroscopy has been applied to a study of the planar interaction in apolar media between 1-substituted N^4-methoxycytosine (and the corresponding 5-methyl analogue) and 9-substituted adenines. In both chloroform and carbon tetrachloride solutions, the exocyclic N^4-methoxy group of N^4-methoxycytosine, and of 5-methyl N^4-methoxycytosine (which is in the oxime form under these conditions), is so oriented that it is predominantly *syn* to the ring N(3), and neither compound forms planar auto-associates. In chloroform solution, both the N(3)–H and the C(2)=O interact weakly with the solvent, the interaction being of the nature of non-typical hydrogen bonding. The ^{13}C-NMR chemical shift of the C(2) of N^4-methoxycytosine is modified during formation of hetero-associates with 9-substituted adenine, in accordance with the C(2)=O of the former being the acceptor of an adenine amino proton. The resulting planar hetero-associate is a non-Watson-Crick type of base-pair. This was further substantiated by infrared absorption studies of the carbonyl frequency during complex formation. The results are examined in the light of the mechanism of hydroxylamine (and methoxyamine) mutagenesis.

1. Introduction

N^4-Hydroxycytidine (or N^4-methoxycytidine), the product of the reaction of hydroxylamine (or methoxyamine)with cytidine in vitro and in vivo (for reviews, see refs. 1 and 2), is a promutagenic derivative considered as being responsible for point transition mutations C → U(T). N^4-Hydroxycytosine residues, when present in a natural template during in vitro transcription with RNA polymerase or Qβ replicase, behave as uracil or cytosine [3]. By contrast, in a synthetic template, such as poly(C) (8% N^4-hydroxycytosine), under competitive conditions, the N^4-hydroxycytosine residues exhibit complete ambiguity [4]. Under analogous conditions N^4-methoxycytosine residues promote incorporation of only AMP [4], and hence behave solely like uracil (or thymine). This is reflected in the fact that 1-Me-N^4-OMeCyt (1a in scheme 1) is in the oximino form in apolar medium [5], in the gas phase [6] and in the crystalline state [2,7]; and it is this form which predominates in aqueous medium [5], and forms specific hydrogen-bonded complexes only with adenine derivatives in apolar media [8,9].

The structure and stability of the planar, hydrogen-bonded complexes depend on the solvent and the donor and acceptor sites involved in hydrogen bonding. Results obtained with the aid of infrared

Abbreviations: 1-Me-N^4-OMeCyt, 1-methyl-N^4-methoxycytosine; 1-Oc-5-Me-N^4-OMeCyt, 1-octyl-5-methyl-N^4-methoxycytosine; 1,5-Me_2-N^4-OMeCyt, 1,5-dimethyl-N^4-methoxycytosine; $(OMe)_3$Ado, 2′,3′,5′-tri-*O*-methyladenosine; $(OMe)_3$Urd, 2′,3′,5′-tri-*O*-methyluridine. Note: It should be borne in mind that the choice of a 1-substituent for a pyrimidine, or a 9-substitutent for a purine, is dictated primarily by the necessity for the resulting derivative to exhibit adequate solubility in the solvents employed.



1a R_1=CH_3, R_2=H
1b R_1=R_2=CH_3
1c R_1=octyl, R_2=CH_3

Scheme 1. Analogues of N^4-methoxycytosine in the *syn*-oximino form.

spectroscopy in the region of N–H and O–H stretching frequencies [6,8,9], as well as ^{1}H-NMR spectroscopy, pointed to the utility of examining this further with the aid of ^{13}C-NMR spectroscopy, and infrared spectroscopy in the region of carbonyl stretching frequencies.

Notwithstanding that changes in chemical shifts of a ^{13}C atom vicinal to a proton acceptor are usually much lower than the shift of the ^{1}H proton involved in such hydrogen bonding [10,11], it has proven possible to demonstrate that the C(2)=O of 1-Me-N^4-OMeCyt is a proton acceptor of an amino hydrogen of $(OMe)_3$Ado and, furthermore, to confirm this by an examination of carbonyl stretching frequencies in the infrared.

2. Materials and methods

Previously described procedures were employed for the synthesis of 1-Me-N^4-OMeCyt (1a [12]), 1,5-Me_2-N^4-OMeCyt (1b [12]), $(OMe)_3$Ado [13] and $(OMe)_3$Urd [14].

2.1. 1-Oc-5-Me-N^4-OMeCyt (1c)

A solution of 554 mg of 2,4-diethoxy-5-methylpyrimidine in 15 ml octyl iodide was maintained at 90°C for 3 days. The resulting Hilbert-Johnson rearrangement led to formation of 1-octyl-O^4-ethylthymine, which was isolated on Merck PF_{254} silica gel plates by development with benzene/acetone (17:3, v/v), with $R_f = 0.6$ for the starting compound and 0.3 for the product. Identification of the latter was based on its ultraviolet spectrum and by its acid hydrolysis to 1-octylthymine. The 1-octyl-O^4-ethylthymine (15000 OD_{280} units, 2.3 mmol) was reacted with methoxyamine (30 mmol) in 15 ml methanol at 85% overnight, at which point the reaction was 80% complete. The resulting 1-Oc-5-Me-N^4-OMeCyt (12000 OD_{284} units) was isolated on Merck PF_{254} silica gel with benzene/acetone (4:1, v/v), with R_f values of 0.4 for 1-octylthymine, 0.5 for 1-octyl-O^4-ethylthymine and 0.7 for 1-Oc-5-Me-N^4-OMeCyt. Crystallization of the latter from *n*-hexane yielded 90 mg in the form of needles, m.p. 37–39°C, chromatographically homogeneous, and unequivocally identified on the basis of its ultraviolet absorption spectrum at various pH values [12].

2.2. N^6,N^6-Dimethyl-2',3',5'-tri(O-methyl)adenosine

This compound was prepared by treatment with dimethylamine of 2',3',5'-tri(*O*-methyl)-O^6-ethylinosine. The latter was obtained by nitrite deamination of 2',3',5'-tri(*O*-methyl)adenosine, and methylation of the product with diazomethane in dimethylformamide, as previously described for the synthesis of O^6-ethylguanosine [15].

A solution of the 2',3',5'-tri(*O*-methyl)-O^6-ethylinosine (300 mg, 0.9 mmol) in 5 ml of 30% ethanolic dimethylamine and 1 ml H_2O in a sealed ampoule was kept at 100–105°C for 3 days (the use of anhydrous ethanolic dimethylamine, or shorter reaction times, resulted in low reaction yield). The reaction mixture was subjected to preparative chromatography on Merck F_{254} silica gel 60 plates with isopropanol/chloroform (1:19, v/v), and the major band ($R_f = 0.30$) eluted with methanol to give the desired product in 70% yield. The eluate was treated with charcoal, refluxed, filtered and brought to dryness. Crystallization from 2 ml cyclohexane gave the desired product in the form of needles, m.p. 72.5–73.5°C. Identification was based on elementary analysis (C, H, N), melting point and its ultraviolet spectrum by comparison with a similar compound prepared by a different method [16].

Solvents for infrared and NMR spectroscopy

included C^2HCl_3 (99.7 mol% ^2H, from the Institute of Nuclear Studies, Swierk), $CHCl_3$ (POC, Gliwice) and CCl_4 (Merck, Darmstadt, F.R.G.).

The various compounds were dried in vacuo, the solvents were stored over molecular sieves, and solutions were prepared in a dry box.

Infrared spectra were recorded on a Zeiss (Jena, G.D.R.) Specord 75IR instrument fitted with a diffraction grating. Cuvettes with different path lengths (as indicated in the figures below) were fitted with Infrasil windows for the N–H stretching frequency region, and with NaCl (or KBr) windows for the carbonyl region.

^1H-NMR spectra were run at room temperature (unless otherwise indicated) on solutions in C^2HCl_3 (unless otherwise indicated) with an internal tetramethylsilane standard, using a Bruker-90 instrument operating in the Fourier transform mode.

^{13}C-NMR spectra were obtained with a Jeol PFT-90 Fourier transform system locked on deuterium, at room temperature, using C^2HCl_3 solutions with an internal tetramethylsilane standard, and with full decoupling of protons.

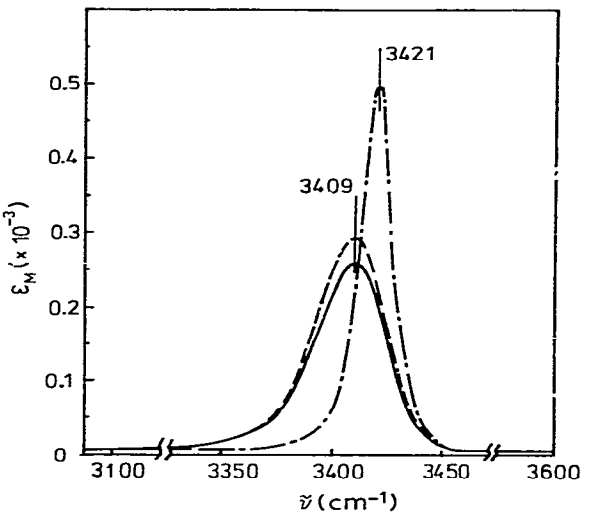


Fig. 1. Infrared absorption spectra at 20°C of: (———) 1-Me-N^4-OMeCyt, 0.0325 M in C^2HCl_3 (1-mm Infrasil cell). (- - - -) 1,5-Me_2-N^4-OMeCyt, 0.0325 M in C^2HCl_3 (1-mm Infrasil cell). (- · - · -) 1-Me-N^4-OMeCyt, 1.06×10^{-3} M in CCl_4 (10-mm Infrasil cell).

3. Results

3.1. Self-association of 1-Me-N^4-OMeCyt and 1,5-Me_2-N^4-OMeCyt

Notwithstanding that the presence of a 5-methyl substituent sterically constrains the exocyclic N^4-methoxy group to the *syn* conformation (1b in scheme 1, see ref. 2), this does not essentially affect the infrared spectra of these two compounds in the range 3100–3600 cm^{-1} (fig. 1 and table 1), with the exception of the N(3)–H stretch at 3409 cm^{-1}, the molar extinction of which is slightly higher for the 5-methyl derivative. With an increase in temperature, the shape, molar extinction and integral absorption of the N(3)–H bands are unaltered, but there is a slight shift in frequency, from 3406 to 3412 cm^{-1}, an effect typical for an N–H stretch [17]. There was no observable temperature effect on stretching frequencies in the range 2700–3100 cm^{-1}, thus excluding the existence of an intermolecular hydrogen-bonded N–H stretch, and hence of self-association, as it is in the case for uracil and adenine [18].

Over a range of concentration of 0.001–0.15 M in chloroform, the N(3)–H band of 1-Me-N^4-OMeCyt exhibited no change in frequency (3409 cm^{-1}), molar extinction (262 cm^{-1} M^{-1}), half-width (18 cm^{-1}) and integral absorption (10.1×10^3 cm^{-2} M^{-1}). The corresponding values in CCl_4 solution (3421 cm^{-1}, 495 cm^{-1} M^{-1}, 8 cm^{-1}, 10.0×10^3 cm^{-2} M^{-1}) were also unaffected in the concentration range 0.005–0.015 M. It should be noted (see fig. 1) that the N(3)–H stretch is shifted from 3409 to 3421 cm^{-1} on transfer from C^2HCl_3 ($\mu = 1.2$ debye) to CCl_4 ($\mu = 0$), and that this is accompanied by a marked reduction in the half-width of the band, with a concomitant increase in symmetry, and essentially no change in integral absorption. It is clear that, in chloroform solution, there is some type of interaction between the N(3)–H and the solvent.

Table 1

Parameters of the 3409 cm^{-1} bands of 1-Me-N^4-OMeCyt and 1,5-Me_2-N^4-OMeCyt, each at a concentration of 0.0325 M in C^2HCl_3 at 20°C

	ϵ_M (cm^{-1} M^{-1})	$\nu_{1/2}$ (cm^{-1})	I_∞ (M^{-1} cm^{-2}) ($\times 10^{-3}$)
1-Me-N^4-OMeCyt	262	18	10.07
1,5-Me_2-N^4-OMeCyt	290	19	10.96

3.2. NMR spectra: concentration dependence

The apparent absence of auto-association of N^4-methoxycytosine and its 5-methyl derivative in C^2HCl_3 (and in CCl_4) was further examined by means of ^1H- and ^{13}C-NMR spectroscopy in chloroform solution with increasing concentrations of each compound. From tables 2 and 3 it can be seen that neither the proton, or carbon, chemical shift of either compound is dependent on concentration. It follows that neither the N(3) proton, nor the C(2) carbonyl (potential donor and acceptor of protons, respectively), is involved in intermolecular hydrogen bonding.

It had earlier been reported by Brown et al. [5] that, in aqueous medium, the chemical shift of H(5) in 1,3-dimethyl-N^4-hydroxycytosine (fixed *anti*-oximino form) is 3.85 ppm, and in 1-methyl-N^4-hydroxycytosine 4.31 ppm. Such a marked effect of the fixed *anti*-oximino form is a strong argument for the existence in aqueous medium of the *syn*-oximino form (scheme 1). We have shown that methylation of C(5) in 1-Oc-5-Me-N^4-OMeCyt, which sterically constrains the N^4-methoxy group to the *syn* conformation, is without effect on the chemical shift of the N(3) proton in this compound in chloroform solution (table 2). This further supports the existence of 1-Me-N^4-OMeCyt (and of its 5-methyl analogue) in chloroform solution in the imino form with the N^4-methoxy group *syn* to the ring N(3).

3.3. Association of 1-Me-N^4-OMeCyt and 1,5-Me_2-N^4-OMeCyt with $(OMe)_3$Ado

Such associations in non-aqueous media (C^2HCl_3, CCl_4) have previously been studied by examining possible changes in N–H and $N\!<\!^{H}_{H}$

Table 2

Dependence of proton chemical shifts (vs. internal tetramethylsilane) on concentration of 1-Me-N^4-OMeCyt, 1-Oc-5-Me-N^4-OMeCyt, $(OMe)_3$Urd, $(OMe)_3$Ado, and their 1:1 mixtures in C^2HCl_3 at 25°C

Base or base-pair	NH_2		N(3)H	
	0.02 M	0.2 M	0.02 M	0.2 M
$(OMe)_3$Ado	5.74 [a]	6.27 [a]	–	–
1-Me-N^4-OMeCyt	–	–	8.09 [b]	8.09 [b]
1-Oc-5-Me-N^4-OMeCyt	–	–	8.05 [b]	8.10 [b]
1-Me-N^4-OMeCyt:$(OMe)_3$Ado	5.63 [b]	6.39 [a]	8.11 [c]	8.71 [b]
1-Oc-5-Me-N^4-OMeCyt:$(OMe)_3$Ado	5.62 [a]	6.37 [b]	8.15 [c]	8.62 [a]
$(OMe)_3$Urd	–	–	8.51 [b]	9.78 [b]
$(OMe)_3$Urd:$(OMe)_3$Ado	6.06 [a]	6.69 [a]	10.67 [b]	12.65 [b]

Accuracies: (a) ±0.01 ppm; (b) ±0.02 ppm; (c) ±0.05 ppm.

Table 3

Dependence on concentration of the ^{13}C-NMR chemical shifts (in ppm vs. internal tetramethylsilane, with complete decoupling of protons) of 1-Me-N^4-OMeCyt, $(OMe)_3$Ado, and their 1 : 1 mixture

^{13}C [a]	$(OMe)_3$Ado		1-Me-N^4-OMeCyt		1 : 1 mixture	
	0.2 M	0.04 M	0.2 M	0.04 M	0.2 M	0.04 M
$(OMe)_3$Ado						
C6	155.68	155.36			155.81	155.36
C2	153.01	153.07			153.07	153.07
C4	149.49	149.55			149.35	149.58
C8	139.17	139.30			139.10	139.30
C5	120.11	120.30			119.98	120.24
C1′	86.95	87.01			86.88	86.95
C′, C3′	82.25	82.25			82.18	82.18
	81.07	81.14			81.07	81.07
C4′	77.75	77.75			77.81	77.68
C5′	71.41	71.41			71.41	71.35
O′CH_5	59.27	59.27			59.27	59.27
	58.62	58.62			58.55	58.55
	58.23	58.23			58.16	58.16
1-Me-N^4-OMeCyt						
C2 [b]			149.81	149.84	150.27	149.94
C4			145.31	145.33	145.37	145.31
C6			136.04	136.05	135.97	136.04
C5			97.52	97.57	97.72	97.59
N^4-OCH_3			61.56	61.64	61.49	61.56
1CH_3			35.25	35.28	35.25	35.25

[a] Accuracies +0.03 ppm.

[b] Identification of C(2) and C(4) made on basis of selective decoupling from the group of protons N(1)–CH_3.

stretches in the region 3100–3600 cm^{-1} [8,9]. It has in fact been shown that the observed absorption of these bands in 1 : 1 mixtures of 9-substituted adenine and 1-substituted N^4-methoxycytosine differs from the algebraic sum of the two components, testifying to their hydrogen bonding in the form of planar complexes.

In view of the fact that the N^4-methoxy group of the pyrimidine component is *syn* with respect to the ring N(3), thus sterically precluding Watson-Crick hydrogen bonding, the foregoing was checked by examining the spectrum of a 1 : 1 mixture of 1-Me-N^4-OMeCyt and N^6,N^6-dimethyl-2′,3′,5′-tri(*O*-methyl)adenosine. The latter is clearly incapable of hydrogen bonding. The spectrum of the mixture proved to be identical with that of 1-Me-N^4-OMeCyt alone in the region 3100–3600 cm^{-1}, thus confirming that the previously observed interaction of adenine with N^4-methoxycytosine is due to hydrogen bonding. Engel and Von Hippel [19] have similarly shown, with the aid of ^{1}H-NMR spectroscopy, that there is no interaction in chloroform solution between 1-substituted uracil and 9-substituted N^6,N^6-dimethyladenine.

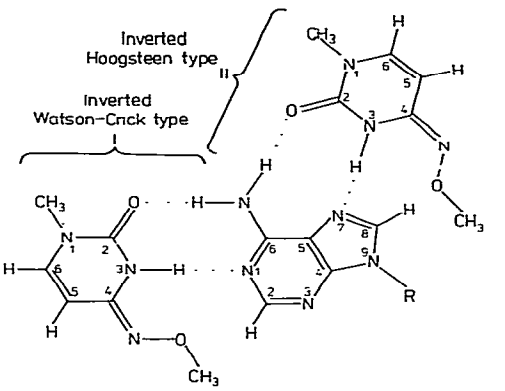
To test further the validity of the above results, recourse was had to ^{1}H-NMR spectroscopy. An increase in concentration of a 1 : 1 mixture of 1-Me-N^4-OMeCyt (or its 5-methyl derivative) with $(OMe)_3$Ado, which should lead to enhanced association, led to deshielding of the protons involved in hydrogen bonding, as evidenced by an increase in their chemical shifts (table 2). The chemical shifts of the adenine amino protons are modified to the same extent as in a 1 : 1 complex of $(OMe)_3$Ado with $(OMe)_3$Urd. The change in

chemical shift of the N(3)–H on mixing 1-Me-N^4-OMeCyt (or its 5-methyl analogue) with $(OMe)_3$Ado is less than that with a 1 : 1 mixture of $(OMe)_3$Urd with $(OMe)_3$Ado. With the two concentrations (0.2 and 0.02 M), the difference in chemical shifts of the N(3)–H and the amino protons between 1 : 1 mixtures and the individual components in chloroform (table 2) is much more pronounced for the complex $(OMe)_3$Urd-$(OMe)_3$Ado (2.87 and 0.42 ppm at 0.2 M, and 2.16 and 0.32 ppm for 0.02 M) than for the complex 1-Me-N^4-OMeCyt : $(OMe)_3$Ado (0.62 and 0.12 ppm for 0.2 M, and 0.02 and −0.11 ppm for 0.02 M). Notwithstanding that the concentration dependence of chemical shifts was not determined over a broader range of concentrations (to determine more accurately association constants), the results for the two concentrations (table 2) do permit a qualitative evaluation of the relation between them. This relation confirms the data obtained by infrared spectroscopy, which show that the association constant between 1-Me-N^4-OMeCyt and $(OMe)_3$Ado is 14 ± 1 M^{-1} [9] and between $(OMe)_3$Urd and $(OMe)_3$Ado 152 ± 13 M^{-1}.

Furthermore, the chemical shifts of the amino and ring N(3) protons for 1 : 1 mixtures of $(OMe)_3$Ado with 1-Me-N^4-OMeCyt, or with 1-Oc-5-Me-N^4-OMeCyt, were identical. This confirms earlier results obtained by means of infrared spectroscopy, which yielded identical association constants for the two complexes, $K = 14$ M^{-1} [9].

3.4. *Structure of planar complex of 1-Me-N^4-OMeCyt with $(OMe)_3$Ado*

The concentration dependence of the ^{13}C-NMR chemical shifts in 1 : 1 mixtures of $(OMe)_3$Ado with 1-Me-N^4-OMeCyt in C^2HCl_3 (table 3) permits some inferences regarding the structure of this complex. An increase in chemical shift in the mixtures, relative to the individual components and with an increase in concentration derives from the deshielding resulting from formation of hydrogen bonds. The significantly larger increase in chemical shift of C(2), as compared to C(4), in the pyrimidine component testifies to involvement of O^2 as a hydrogen bond acceptor of one of the



Scheme 2. The two possible hydrogen-bonded planar associates formed between 1-Me-N^4-OMeCyt and $(OMe)_3$Ado.

amino protons of $(OMe)_3$Ado. The chemical shift of the N^4-methoxy carbon is unaffected with increase in concentration, so that the oxygen atom of this group, a potential proton acceptor, does not participate in hydrogen bonding.

Consequently, 1-Me-N^4-OMeCyt (or its 5-methyl analogue) and $(OMe)_3$Ado, in chloroform solution form 1 : 1 planar complexes [8], the hydrogen bonding of which (see scheme 2) may be either

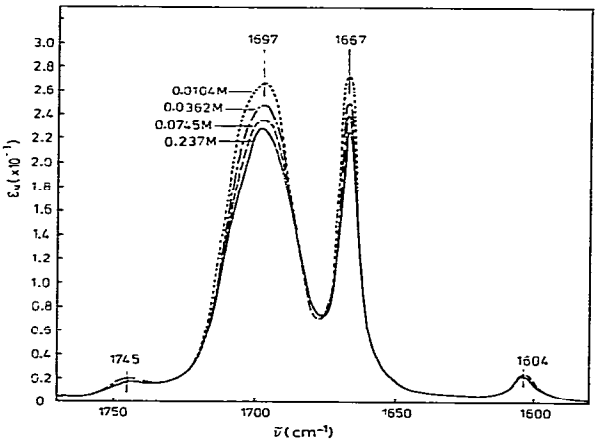


Fig. 2. Infrared absorption spectra of 1-Me-N^4-OMeCyt at various concentrations in C^2HCl_3 at 20°C: (———) 0.237 M (0.050-mm KBr cell). (- - - - -) 0.0745 M (0.050-mm KBr cell). (—·—·—) 0.0362 M (0.253-mm NaCl cell). (—··—··—) 0.0104 M (0.253-mm NaCl cell).

(I) of the inverted Watson-Crick type, or (II) an inverted Hoogsteen base-pair. The minimal effect of planar hetero-association on the chemical shifts of C(2), C(5) and C(8), by comparison with that for C(6) in $(OMe)_3$Ado (table 3), and the absence of any effect on the chemical shift of ^{1}H(2) and ^{1}H(8), 8.24 and 8.35 ppm, respectively, do not permit an unequivocal conclusion as to which form is predominant in chloroform solution. Independent studies, with the aid of ^{15}N-NMR (in preparation) have shown that the chemical shifts of ^{15}N(7) of $(OMe)_3$Ado are not affected, while those of ^{15}N(1) and ^{15}N of the amino nitrogen are modified, on association with 1-Me-N^4-OMeCyt. This excludes formation of reversed Hoogsteen planar hydrogen bonding (II in scheme 2) and points to formation of reversed Watson-Crick pairing.

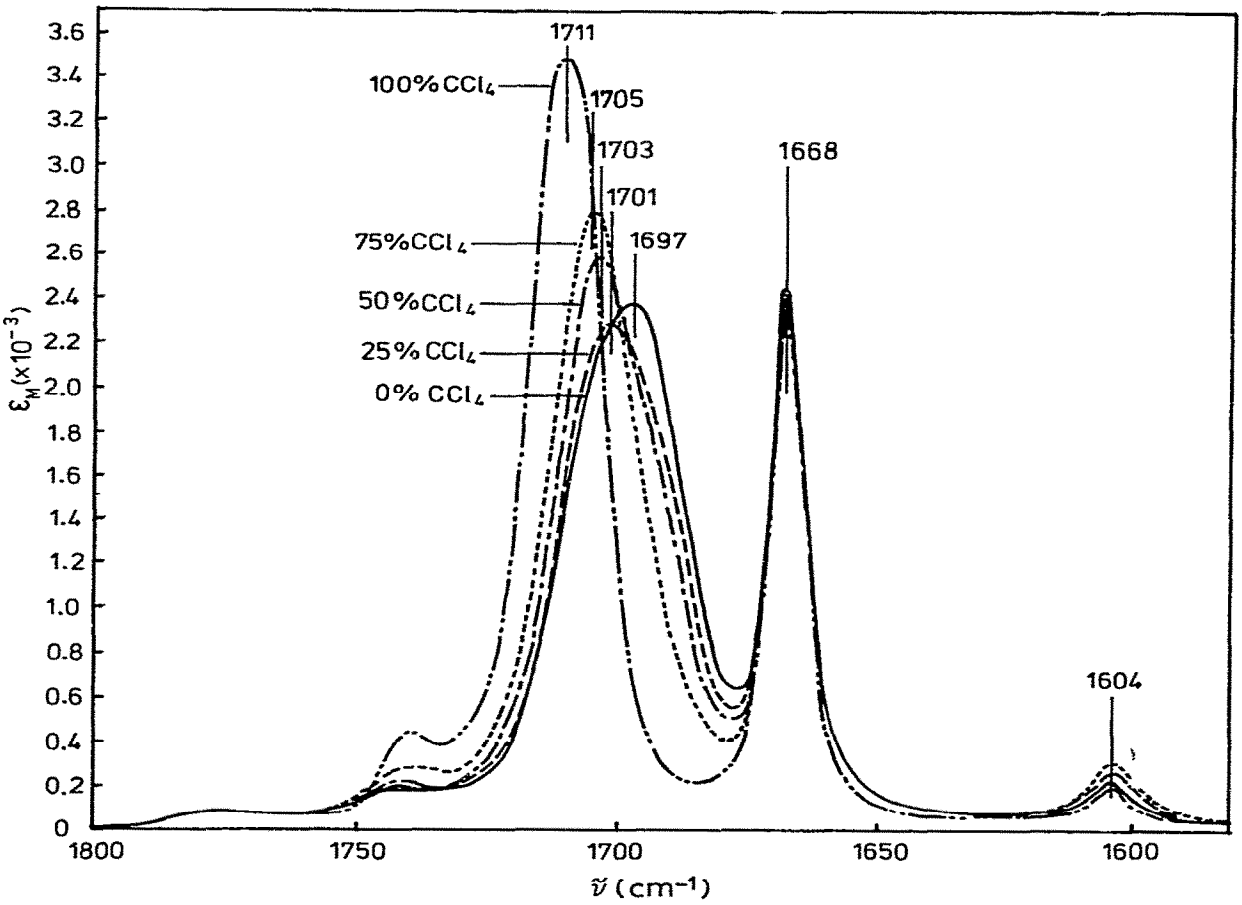


Fig. 3. Infrared absorption spectra of 1-Me-N^4-OMeCyt, 0.00320 M in 0.631-mm KBr cell, at 20°C in C^2HCl_3 containing various proportions of CCl_4 as follows: 0% (———), 25% (- - - - - -), 50% (·—·—·), 75% (· · · · · ·), 100%, i.e., CCl_4 alone (—·—·—·—).

3.5. *Infrared absorption spectrum of 1-Me-N^4-OMeCyt in the region 1500–1800 cm^{-1}*

This spectrum was found to be significantly concentration dependent in chloroform solution (fig. 2), a rather surprising result, since it was shown (see above) that 1-Me-N^4-OMeCyt does not form auto-associates. It consequently appeared likely that it interacts with the solvent. When the spectrum of the analogue was examined in chloroform solution containing increasing concentrations of CCl_4, the band at 1697 cm^{-1} in C^2HCl_3 was most sensitive to addition of CCl_4 (fig. 3). In CCl_4 alone it is located at 1711 cm^{-1}, and in ^2H_2O at 1663 cm^{-1}. The polarities of C^2HCl_3 and of the C(2)=O group suggest that the interaction observed in chloroform solution is due to non-typical hydrogen bonding.

The frequencies of the bands at 1667 and 1604 cm^{-1} are solvent independent, and the structures of these bands are also unaffected following 90% deuteration of the N(3)–H, so that they are not conjugated with the N(3)–H. On replacement of the ring N(3) proton by deuterium, the locations of these bands undergo shifts typical for purine and pyrimidine bases [20,21], viz. from 1697, 1667 and 1604 to 1694 cm^{-1}, 1664 and 1594 cm^{-1}. Taking account of other literature data [19–21], the assignments of these bands in chloroform solution are as follows: 1697 cm^{-1}, C(2)=O stretching vibration; 1667 and 1604 cm^{-1}, ring stretching vibrations.

Although the interaction of the carbonyl group and the ring N(3)–H of 1-Me-N^4-OMeCyt with chloroform competes with the process of complex formation of this analogue with $(OMe)_3$Ado, the change in frequency of the N(3)–H stretch on transfer from CCl_4 to C^2HCl_3 (13 cm^{-1}) is on

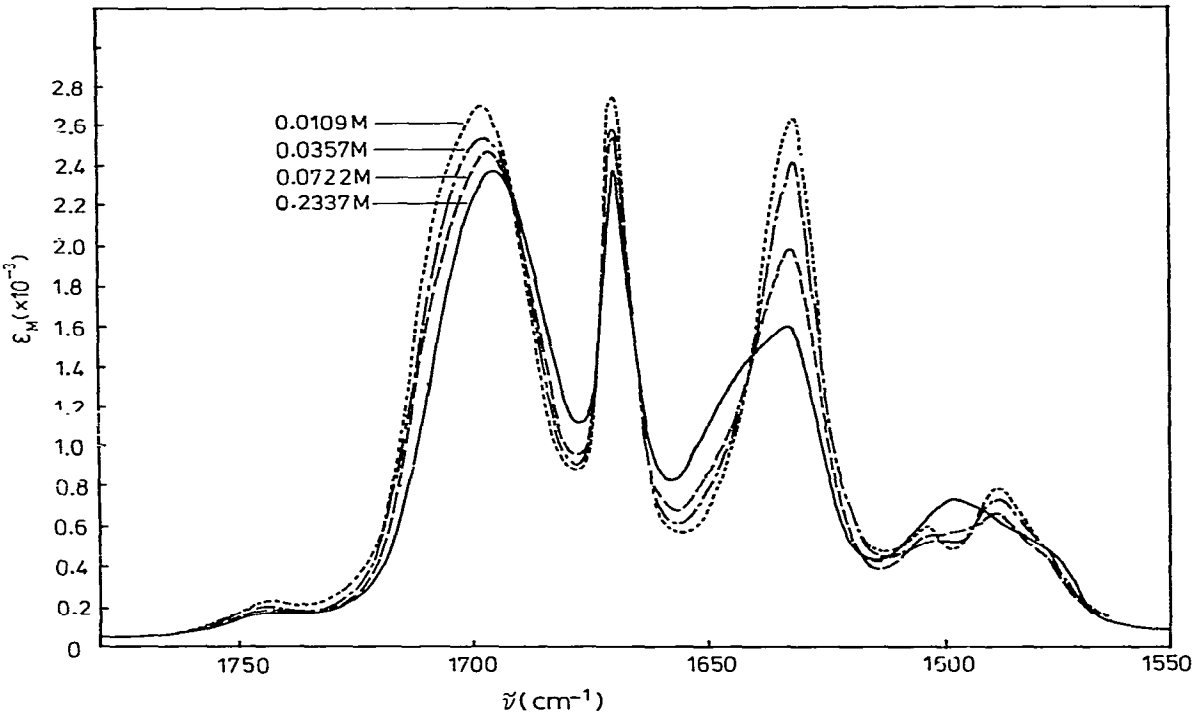


Fig. 4. Infrared absorption spectra of a 1:1 mixture of 1-Me-N^4-OMeCyt and $(OMe)_3$Ado in C^2HCl_3 at 20°C and with various equimolar concentrations of each constituent.

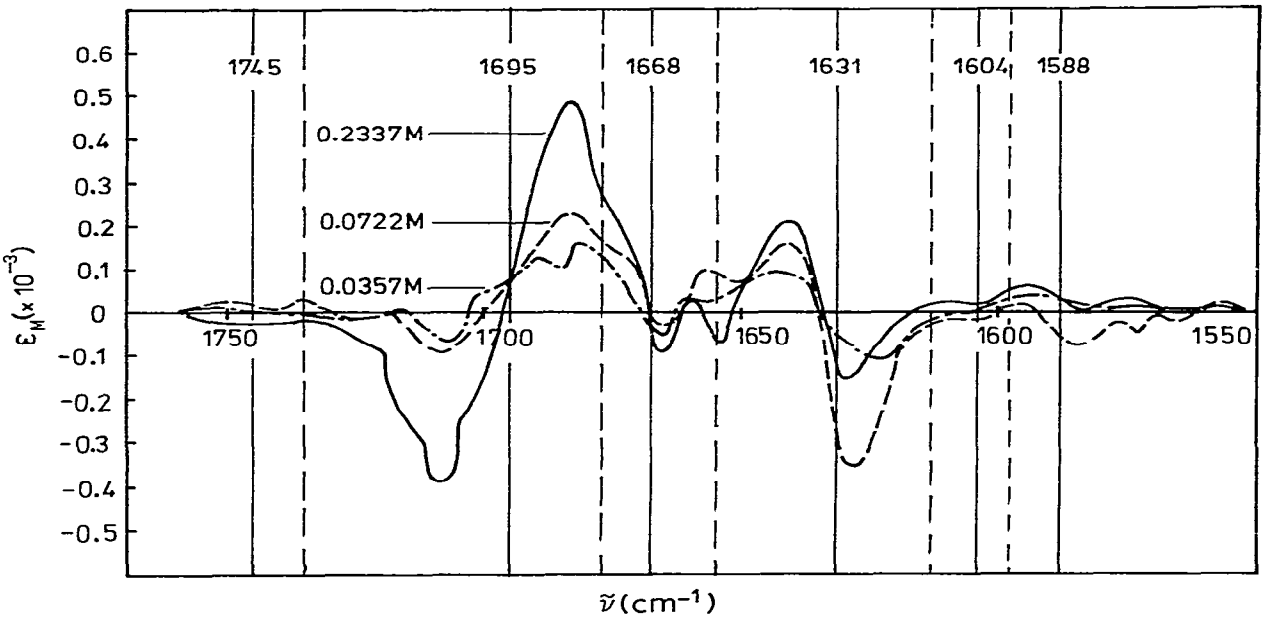


Fig. 5. Infrared difference spectra of a 1 : 1 mixture of 1-Me-N^4-OMeCyt and $(OMe)_3$Ado, relative to the arithmetic sum of the two components, in C^2HCl_3 at 20°C, and for three different concentrations as indicated (cell path lengths as in Fig. 2); 0.2337 M (———), 0.0722 M (- - - - -), 0.0357 M (·—·—·).

average 5-fold lower than that which occurs as a result of formation of complementary base-pairs in chloroform [19].

In the planar complex of 1-Me-N^4-OMeCyt with $(OMe)_3$Ado in chloroform, it was shown above that the carbonyl group of the former is the acceptor of one of the amino protons. We now see that such hydrogen-bonded base-pairing modifies the intensities and frequencies of the bands of the mixture in the range 1500–1800 cm^{-1} as a function of the concentration (fig. 4). The C_2=O band, located at 1679 cm^{-1}, lies outside the range of the $(OMe)_3$Ado bands, so that its behaviour may be readily followed during complex formation.

The extent of involvement of the carbonyl group is the difference between the absorption of the 1 : 1 mixture and the algebraic sum of the components, both measured under the same conditions. The resulting difference spectrum is then dependent on the concentration of the equimolar mixture of the components (fig. 5), i.e., the intensities of the individual bands decrease with decreasing concentration. This is in agreement with expectations, since a decrease in concentration decreases the extent of association of the two components. Thus, this procedure further supports the involvement of the carbonyl group in the planar interaction of 1-Me-N^4-OMeCyt with $(OMe)_3$Ado.

4. Discussion

4.1. Structure of planar complex

As pointed out above, the nature and stability of the hydrogen-bonded planar complex of 1-Me-N^4-OMeCyt with $(OMe)_3$Ado necessarily depend on the tautomeric equilibrium, and the orientation of the exocyclic N^4-methoxy group of the former. Notwithstanding that earlier findings [2,5–7], as well as those presented here, point to the predominance (if not exclusively) of the oximino form,

with donor and acceptor sites similar to those of uracil (or thymine), the steric hindrance of the N^4-methoxy group in the *syn* conformation would be expected to prevent Watson-Crick pairing with $(OMe)_3$Ado. And, in fact, the observed modifications, as a result of complex formation, of the signals due to $^{13}C(4)$, to ^{13}C and ^1H of the methyl in the N^4-methoxy group (tables 2 and 3), and the ^{15}N of N^4 (unpublished data, in preparation), exclude this fragment of the molecule from participation in complex formation. The change in chemical shift of C(2) accompanying complex formation is reflected, independently, in the displacement (and change in absorption) of the C(2)=O band in the infrared (figs. 4 and 5). Together with the results of ^{15}N-NMR spectroscopy, the complex is established as one involving inverted Watson-Crick pairing (I, scheme 2), similar to one of the dimers, 9-ethyladenine: 1-cyclohexyluracil, also in chloroform solution [11]. Furthermore, whereas the N^4-methoxy substituent in the *syn* form does not sterically interfere with formation of such an inverted Watson-Crick pair, it would be expected to do so in inverted Hoogsteen pairing (II, scheme 2) because of the presence of the H(8) of the adenine moiety.

The presently observed involvement of a uracil (or thymine) O^2 (instead of O^4) as a hydrogen bond acceptor is by no means limited to the monomer level. Such pairing has been observed in both natural and synthetic polynucleotides, e.g., in the crystal structure of tRNA [22] and is the only logical means of interpretation of the hairpin double-helical structure of poly(U) in solution [23,24].

The conclusion regarding inverted Watson-Crick pairing is further reinforced by the fact that the structure and stability of the 1-Me-N^4-OMeCyt: $(OMe)_3$Ado complex are unaffected by introduction into the N^4-methoxycytosine components of a 5-methyl substituent, which further accentuates the steric hindrance to the existence of the *anti* form [2]. This, in turn, poses a problem of some interest in relation to the finding of Singer and Spengler [4] that N^4-methoxycytosine residues in a copolymer lead to incorporation, during transcription, of only adenine residues. It must therefore be concluded that, at the polymer level, the N^4-methoxy group may be constrained to the *anti* form, most likely by the RNA polymerase. It would, consequently, be of obvious interest to examine the behaviour, in such a polymerase system, of a copolymer containing N^4-methoxycytosine residues with a 5-methyl substituent, where the energy barrier for transition to the *anti* form must be higher. Pertinent to this is the observation that 1,5-N^4,N^4-tetramethylcytosine in the solid state exists with the N^4-dimethylamino group in the plane of the ring [25], a most unexpected result, rendered possible by expansion of the bond angles at C(4), C(5) and N(4). Conceivably, such a process at the polymer level, due to the action of the polymerase, could account for the incorporation of adenine residues at sites of N^4-OMeC residues.

4.2. Solvent effects

The observed interaction of the N(3)–H and C(2)=O of 1-Me-N^4-OMeCyt with chloroform (see figs. 1–3), albeit apparently weak, may compete in the interaction with $(OMe)_3$Ado. Because of the absence of auto-association of 1-Me-N^4-OMeCyt in chloroform and carbon tetrachloride, the observed effect on transfer from chloroform ($\epsilon = 4.8$, $\mu = 1.2$) to carbon tetrachloride ($\epsilon = 2.2$, $\mu = 0$) must be due largely to the interactions with chloroform, which fulfils the role of donor to O^2 and/or acceptor from the N(3)–H.

Such interactions with solvent in non-aqueous media are not uncommon [17,26,27]; and a similar effect of solvent polarity on the location of the N–H stretching band was noted for 1,5-dimethyl-N^4-hydroxycytosine and 1,5-Me_2-N^4-OMeCyt [6], and for 1-substituted uracils and 9-substituted adenines [18,28]. Interaction of one or both of the components with the solvent may affect the association constant, and probably accounts for the fact that the association constant of 1-Me-N^4-OMeCyt (or 1,5-Me_2-N^4-OMeCyt) with $(OMe)_3$Ado in chloroform (14 M^{-1}) is so much lower than that with 5′-acetyl-2′,3′-*O*-isopropylene adenosine in a mixture of chloroform and carbon tetrachloride (120 M^{-1}) [8].

Acknowledgments

We are indebted to Dr. Hanna Sierzputowska-Gracz for the ^{13}C-NMR spectra; to Dr. Irena Ekiel and Dr. Ryszard Stolarski for assistance in interpretation of NMR spectra; and to Mrs. Liliana Ciesielska for meticulous assistance in preparation of the manuscript. This investigation was supported by the Polish Academy of Sciences (09.3.7), the Polish National Cancer Research Program (Pr-6) and The Wellcome Trust.

References

1 E.I. Budowsky, Prog. Nucleic Acid Res. Mol. Biol. 16 (1976) 125.
2 D. Shugar, C.P. Huber and G.I. Birnbaum, Biochim. Biophys. Acta 447 (1976) 274.
3 R.A. Flavell, D.L. Sabo, E.F. Bandle and C. Weissmann, J. Mol. Biol. 89 (1974) 255.
4 B. Singer and S. Spengler, Biochemistry 20 (1981) 1127.
5 D.M. Brown, M.J.E. Hewlins and P. Schell, J. Chem. Soc. C. (1967) 1925.
6 K. Kulińska, A. Psoda and D. Shugar, Acta Biochim. Pol. 27 (1980) 57.
7 G.I. Birnbaum, T. Kulikowski and D. Shugar, Can. J. Biochem. 57 (1979) 308.
8 D.M. Brown and M.J.E. Hewlins, Nature 221 (1969) 656.
9 A. Psoda, B. Kierdaszuk, A. Pohorille, M. Geller, J.T. Kuśmierek and D. Shugar, Int. J. Quantum Chem. 20 (1981) 543.
10 I. Morishima, T. Inubushi, T. Yonezawa and Y. Kyogoku, J. Am. Chem. Soc. 99 (1977) 4299.
11 H. Iwahashi and Y. Kyogoku, J. Am. Chem. Soc. 99 (1977) 7761.
12 C. Janion and D. Shugar, Biochem. Biophys. Res. Commun. 18 (1965) 617.
13 Z. Kazimierczuk, E. Darzynkiewicz and D. Shugar, Biochemistry 15 (1976) 2735.
14 J.T. Kuśmierek, J. Giziewicz and D. Shugar, Biochemistry 12 (1973) 194.
15 J.T. Kuśmierek and D. Shugar, Nucleic Acids Res., Special Publ. 4 (1978) 73.
16 L.B. Townsend, R.K. Robins, R.N. Loeppky and N.J. Leonard, J. Am. Chem. Soc. 86 (1964) 5320.
17 A.K. Covigton and P. Jones, Hydrogen bonded solvents (Academic Press, London, 1968).
18 Y. Kyogoku, R.C. Lord and A. Rich, J. Am. Chem. Soc. 89 (1967) 496.
19 J.D. Engel and P.H. von Hippel, Biochemistry 13 (1974) 4143.
20 M. Tsuboi, Appl. Spectrosc. Rev. 3 (1969) 45.
21 M. Tsuboi, S. Takahashi and I. Harada, in: Physico-chemical properties of nucleic acids, ed. J. Duchesne (Academic Press, London, 1973) p. 91.
22 S-H. Kim, in: Topics in nucleic acid structure, ed. S. Neidle (MacMillan, London, 1968) p. 83.
23 A. Rabczenko and D. Shugar, Acta Biochim. Pol. 18 (1971) 387.
24 D. Bode, M. Heinke and U. Scherman, Biochem. Biophys. Res. Commun. 52 (1973) 1234.
25 J.K. Dattagupta, W. Saenger, K. Bolewska and I. Kułakowska, Acta Crystallogr. B33 (1977) 85.
26 G.C. Pimental and A.L. Mac Clellan, The hydrogen bond, ed. L. Pauling (W.H. Freeman, San Francisco, 1960) p. 197.
27 P. Schuster, G. Zundel and C. Sandorby, The hydrogen bond (North-Holland, Amsterdam, 1976).
28 E. Küchler and J. Derkosh, Z. Naturforsch. 21b (1966) 209.