

This paper is dedicated to Professor David Shugar on his 80th birthday

Methoxyamine attack on cytosine produces ambivalent base pairing properties of the modified base*

Zofia Gdaniec^a, Lawrence C. Sowers^b and G. Victor Fazakerley^{c**}

^aInstitute of Bioorganic Chemistry, Polish Academy of Sciences,
Z. Noskowskiego 12/14, 61-704 Poznań, Poland,

^bDivision of Pediatrics, City of Hope National Medical Center,
Duarte, California 91010, U.S.A.

^cService de Biochimie et de Génétique Moléculaire, DBCM, CE Saclay,
91191 Gif-sur Yvette, France

Key words: DNA, mutagenesis, tautomerism, NMR, N^4 -methoxycytosine

We report the solution structure of two heptanucleotides each containing a central N^4 -methoxycytosine, in one case with adenine on the opposite strand and in the other with guanine. For the N^4 -methoxycytosine-adenine pair only the imino form of the N^4 -methoxycytosine residue is observed and base pairing is in Watson-Crick geometry. However, rotation of the methoxy group about the N-OCH₃ bond is not constrained to a particular orientation although it must be *anti* to the N3 of N^4 -methoxycytosine. The slow exchange on a proton NMR time scale between the single strand and double strand forms is attributed to the strong preference of the *syn* conformation of the OCH₃ group in the single strand which inhibits base pair formation. For N^4 -methoxycytosine base paired with guanosine we observe the N^4 -methoxycytosine base in the amino form in Watson-Crick geometry and a slow exchange of this species with an imino form base paired in wobble geometry. The amino form is predominant at low temperature whereas the imino form predominates above 40°C. Our results point to preferential replacement of dTTP by N^4 -methoxycytosine in primer elongation.

The reaction of methoxyamine with nucleic acids results in conversion of cytosine bases to N^4 -methoxycytosine (mo^4C). Substitution of one hydrogen atom of the cytosine amino group by an electron withdrawing methoxy group can induce formation of the imino tautomer form with tautomeric equilibrium constant orders of a magnitude lower than those for the standard DNA bases.

Studies on the stability of oligonucleotide duplexes containing mo^4C [1-2] have revealed that mo^4C , which is ambivalent in its hydrogen bonding potential, forms stable base pairs with both adenine and guanine. It was predicted that mo^4C would base pair with G in the amino form and with A in the imino form. In fact, an NMR study [3] has shown that mo^4C when paired with A is predominantly in the

*In part this work was supported by the National Institutes of Health Grant GM41336 and CA33572 which are gratefully acknowledged.

**Author to whom correspondence should be addressed at CE Saclay, SBGM, DBCM.

Abbreviations: NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; TOCSY, total correlation spectroscopy; COSY, correlation spectroscopy; 2QF, double quantum filtered correlated spectroscopy; mo^4C , N^4 -methoxycytosine; a, i, denote protons of amino and imino forms, respectively.

imino conformation and in Watson-Crick geometry.

An X-ray crystal structure [4] of an oligonucleotide containing a $\text{mo}^4\text{C}\cdot\text{G}$ base pair showed that the base pair adopts a wobble conformation with the mo^4C base in the imino form. The oligonucleotide was observed to be a Z-DNA form. Recent NMR studies [5–6] proposed wobble and Watson-Crick structures in equilibrium. In one of these [6] a further equilibrium between *syn* and *anti* conformations of the methoxy group of $\text{mo}^4\text{C}\cdot\text{G}$ in the imino form was invoked to explain the NMR spectra.

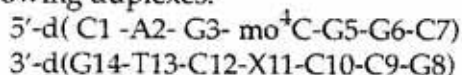
We recently reported [5] preliminary NMR studies on two heptamer duplexes with mo^4C incorporated into the central position with G or A on the opposite strand. On the basis of NMR spectra of exchangeable protons we have shown that, when paired with A, mo^4C is in the imino configuration and in Watson-Crick geometry. However, for the $\text{mo}^4\text{C}\cdot\text{G}$ duplex two structurally distinct configurations were observed in equilibrium with each other. The analysis of NMR spectra showed that both imino and amino configurations of mo^4C base pair with G. When mo^4C is in the amino form it pairs with G in Watson-Crick geometry. In the second species, mo^4C was found in the imino configuration paired with G in wobble geometry.

In this paper, we present further NMR data which confirms and extends our previous structural assignments on mo^4C base pairs in DNA. Based upon interactions between nonexchangeable protons, we have determined the influence of mo^4C substitution on local and global helix structure. Further, examination of proton exchange kinetics allows us to determine the relative rate of exchange of the mo^4C between different base pairing environments.

MATERIALS AND METHODS

N^4 -methoxycytosine was incorporated into the oligonucleotides as previously described [5].

Duplex annealing was monitored by one dimensional NMR experiments. The appropriate pairs of oligonucleotides were heated to 80°C followed by slow cooling to form the following duplexes:



where (1) X11 = A11 or (2) X11 = G11.

The duplexes, 4 mM in strand concentration, were dissolved in 10 mM phosphate buffer, 150 mM NaCl and 0.2 mM EDTA.

NMR spectra were recorded on either AMX500 or AMX600 Bruker spectrometers in either 99.99% D_2O or 90% $\text{H}_2\text{O}/10\%$ D_2O .

NOESY spectra were recorded in the phase sensitive mode [7] with mixing times of 50 and 400 ms for $\text{mo}^4\text{C}\cdot\text{A}$ and 30 and 400 ms for $\text{mo}^4\text{C}\cdot\text{G}$. The NOESY spectra in H_2O were recorded with 150 or 200 ms mixing times. The time domain data sets consisted of 1024 points in the t_2 dimension and 256 or 512 increments in the t_1 dimension. After zero filling the data were multiplied by a slightly shifted sine bell function in both dimensions except for the short mixing time experiments. For these, the data were multiplied by a $\pi/2$ shifted sine bell prior to Fourier transformation.

For spectra recorded in H_2O the observation pulse was replaced by a jump and return sequence [8] and the pulse maximum was placed at 15 ppm. TOCSY spectra [9] were recorded in the phase-sensitive mode with 25 and 70 ms mixing times.

Pure absorption 2QF-COSY spectra [10–12] were obtained with a time proportional phase incrementation scheme. 4K data points in the t_2 dimension and 256 free induction decays were collected.

For the ROESY spectra [13–14] a 200 ms spin lock was applied during the mixing period.

RESULTS

Duplex $\text{mo}^4\text{C}\cdot\text{A}$

The assignment of the exchangeable proton resonances was obtained from the analysis of a NOESY spectrum recorded with a mixing time of 150 ms at 1°C and pH 5.9. The spectrum was recorded at lower pH than those in D_2O (see below) in order to slow down the exchange with bulk solvent. We monitored the non-exchangeable proton shifts as a function of pH and observed no changes in the pH range 5–8 from which we can conclude that there was no conformational change in this pH range.

Three regions are shown in Fig. 1. The lower part shows imino-imino interactions. Starting from the only A·T base pair imino proton we

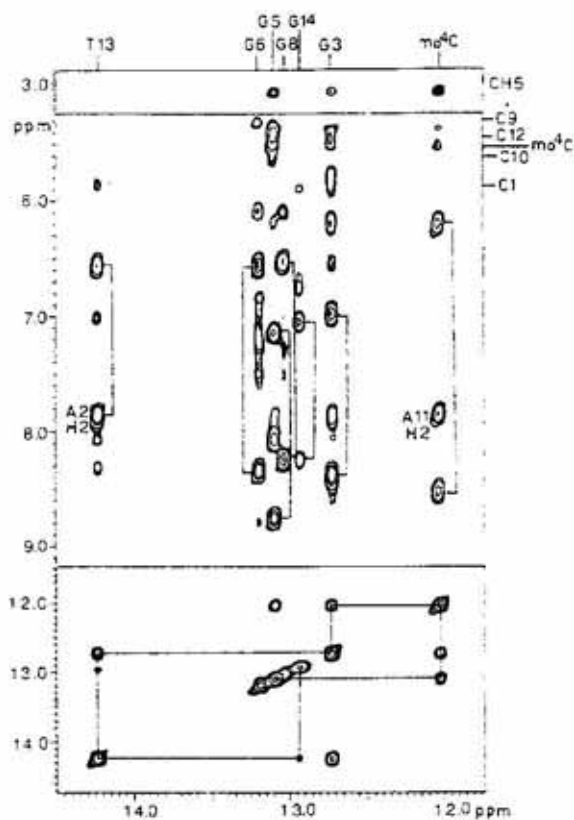


Fig. 1. Three regions of the NOESY spectrum of the N^4 -methoxycytosine·adenine duplex in 90% H_2O /10% D_2O recorded at pH 5.9 and $1^\circ C$ with mixing time 150 ms.

The assignment of the imino protons is shown on the horizontal axis. Lower part shows imino/imino interactions. The middle section shows interactions between imino and amino, CH5, AH2 protons. Pairs of C and A amino proton resonances are connected by solid lines for the intra-base interactions. The upper part shows interactions with the mo^4C methyl resonance.

can follow the interbase imino connectivity through to the G5 imino proton. Although the resonances of the terminal G residues are strongly attenuated by exchange with solvent the crosspeak corresponding to the interaction between the T13 and G14 imino protons is still visible. The remaining non attenuated imino proton is assigned to G6 for which a weak crosspeak is observed with the C10 hydrogen bonded amino proton (cf. the middle part of Fig. 1). This region corresponds to imino/amino/H2/H5 interactions and confirms the chain of connectivities observed for imino/imino interactions. Pairs of C and A amino crosspeaks are connected by solid lines.

The resonance at 12.01 ppm corresponds to an exchangeable proton of the mo^4C ·A base pair. This proton shows strong crosspeaks with the

amino protons of A11. Observation of separate resonances for the amino protons of A11 strongly indicates that this group is involved in hydrogen bonding. When non-hydrogen bonded, an adenosine amino group rotates rapidly and gives rise to a single resonance. For A·T base pairs the resonance of the hydrogen bonded proton is typically found in the range of 7.7–8.1 ppm. For the mo^4C base pair it is found at 8.49 ppm. This implies that hydrogen bonding is with a nitrogen acceptor. If the acceptor was an oxygen, as in Fig. 2a, we would expect a significant upfield shift relative to that observed for A·T base pairs. The resonance at 12.01 ppm also shows a very strong crosspeak with a non-exchangeable proton at 7.84 ppm which must be the A11 H2. The intensity of this cross peak is typical for that of imino-H2 intrabase pair interactions of A·T pairs. This excludes the possibility of an amino form of mo^4C in base pairing (Fig. 2b) for which the proton would be too far from A11 H2 to give such a crosspeak. Together with the observations for the A11 amino resonances only one model (Fig. 2c) in which the mo^4C ·A base pair adopts Watson-Crick geometry with the mo^4C base in the imino form fits all the data. This indicates that the N-4 methoxy group has to be *anti* relative to the N-3 nitrogen because only the *anti* isomer of mo^4C is able to participate in hydrogen bonding with A with two hydrogen bonds as indicated by the data.

The upper part of Fig. 1 shows interactions with the mo^4C methyl group (assignment, see below). The methyl group shows NOEs to the imino protons of the adjacent base pairs and to its own imino N3 proton resonance. Examination of the cross-section through the mo^4C -(OCH₃) resonance shows also strong NOEs to the amino protons of A11 and to amino protons of both C10 and C12. The above results suggest that the methoxy group does not occupy a fixed position but may be oriented in either the 5' or the 3' direction.

Assignment of the nonexchangeable proton resonances was obtained from analysis of the NOESY spectra recorded with 400 and 50 ms mixing times and the TOCSY spectra in D_2O .

The region corresponding to interactions between the base H6/H8/H2 protons and the H1'/H5 protons of the 400 ms NOESY spectrum recorded at $25^\circ C$ is shown in Fig. 3. Six strong cross-peaks (marked X) correspond to

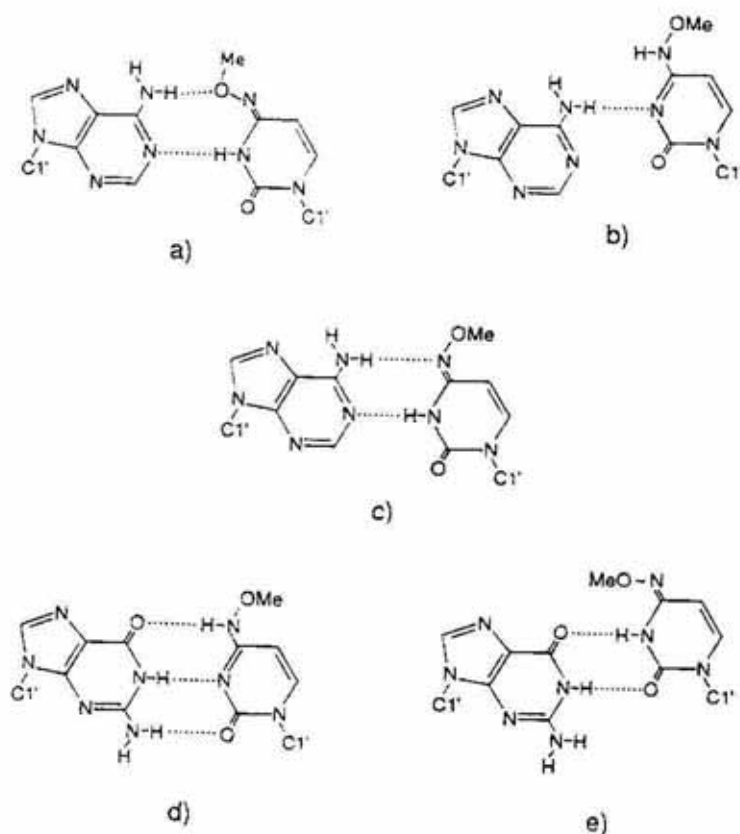


Fig. 2. Possible base pairing forms for N^4 -methoxycytosine-adenine and N^4 -methoxycytosine-guanine.

interactions between H5-H6 protons of cytidine residues. Starting from the 5'-terminal cytidine at 7.65 ppm the sequential connectivities can be followed without ambiguity up to C7. Similarly, on the other strand the connectivities can be followed from G8 to G14.

Analysis of the region of interactions between the base H8/H6 protons and H2'/H2''/CH₃ protons (not shown) confirms the assignment of the base protons shown in Fig. 3. The NOESY experiment with a short mixing time allows an unambiguous discrimination between H2' and H2'' protons, as the H2''-H1' NOE is always larger than that for H2'-H1'.

The A2 H2 and A11 H2 resonances each exhibit a weak crosspeak with their own H1'. The NOEs of the H2 resonances with their adjacent G3 H1' and C12 H1' resonances overlap resulting in a strong crosspeak at 7.83 ppm. Crosspeaks corresponding to interbase interactions H8/H6-H5 are labelled A-E in Fig. 3. The assignment of the H3' and H4' resonances was obtained by the analysis of the TOCSY spectra. The observed chemical shifts are given in Table 1.

In Fig. 3 all inter and intraresidue interactions expected for right-handed B DNA are ob-

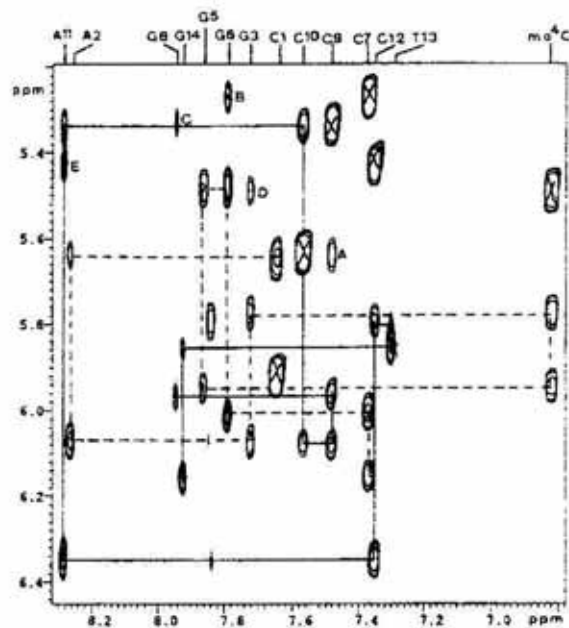


Fig. 3. Part of the 400 ms NOESY spectrum of the N^4 -methoxycytosine-adenine duplex recorded at 25°C, pH 7.

The crosspeaks marked with an X correspond to CH6-CH5 interactions. Interbase crosspeaks labelled A-E are described in the text.

served. The cross-section of the short mixing time NOESY spectrum taken through the methoxy methyl shows interactions with G3 H8, mo^4C H5 and C10 H5 protons. All of them are weak and of similar intensity. From the spectrum in H_2O we know that the N-4 methoxy group is *anti* with respect to N-3 and that the methyl group can point either towards the 3'- or 5'-direction.

The presence of the methoxy group apparently introduces some changes in the local geometry of the helix. We have checked, by model building, that in a normal B-helix the distance between the methyl group and the C10 H5 proton would be too long to observe any NOE effect. The only explanation for this interaction is that the mo^4C -A base pair is displaced towards the major groove and the helical twist increases for the step G3-C12- mo^4C -A11. This arrangement of the base pairs can result in the interactions observed in the NOESY spectra.

In order to determine whether the above distortion influences the sugar conformations we have examined the relative intensity of the intraresidue crosspeaks between the base H6/H8 protons and the H2' *versus* H3' protons. This ratio gives a good indication of the sugar

pucker [15] and we have found that for all non-terminal residues, within experimental error, it corresponds to a predominantly C2'-*endo* conformation. Additionally, these results were confirmed by analysis of the 2QF-COSY spectrum (not shown), for which the sum of the JH1'-H2' and JH1'-H2'' coupling constants is greater than 14.6 Hz for all non-terminal residues.

The only significant deviation observed in the 50 ms mixing time NOESY spectrum from that of a normal B-DNA structure is the ratio of the crosspeak volumes corresponding to inter-residue interactions between the G5 H8 and mo^4C H2',H2'' protons. This ratio is typically about 10 with short mixing times except for the above interaction where the ratio is observed to be 1.5. This can be accounted for by an unusually small helical twist between the base pairs G5-C10 and mo^4C -A11 of 15-20° to better accommodate the methoxy group.

Duplex mo^4C -G

On the basis of NOESY spectra in 90% H_2O /10% D_2O we have previously shown [5] that mo^4C when paired with G exists in an equilibrium between imino and amino configurations. The amino and imino tautomers of

Table 1
Chemical shifts of nonexchangeable protons at 25°C and of exchangeable protons at 1°C for the N^4 -methoxycytosine-adenine duplex.
The chemical shift of methoxy methyl of mo^4C is 3.18 ppm.

	H8/H6	H5/H2 CH3	H1'	H2'	H2''	H3'	H4'	NH	NH2
C1	7.65	5.92	5.64	1.88	2.35	4.68	4.02	-	8.28/7.04
A2	8.26	7.84	6.06	2.78	2.93	5.03	4.39	-	8.08/6.58
G3	7.72	-	5.77	2.51	2.51	4.95	4.39	12.71	-
mo^4C	6.82	5.49	5.94	1.97	2.41	4.92	4.25	12.01	-
G5	7.86	-	5.48	2.65	2.65	4.98	4.31	13.09	8.07/5.40
G6	7.79	-	6.01	2.54	2.72	4.97	4.41	13.18	-
C7	7.38	5.27	6.14	2.20	2.20	4.50	4.02	-	8.26/6.52
G8	7.94	-	5.96	2.64	2.75	4.82	4.25	13.01	-
C9	7.48	5.33	6.07	2.12	2.47	4.83	4.25	-	8.35/6.56
C10	7.57	5.63	5.33	2.20	2.39	4.86	4.12	-	8.78/7.15
A11	8.28	7.83	6.33	2.76	2.98	5.06	4.46	-	8.49/6.19
C12	7.35	5.43	5.79	1.91	2.40	4.76	4.34	-	8.39/6.99
T13	7.30	1.65	5.86	1.96	2.34	4.85	4.11	14.22	-
G14	7.92	-	6.15	2.63	2.38	4.68	4.17	12.92	-

the mo^4C base pair with G in Watson-Crick and wobble geometry are shown in Fig. 2d, e, respectively.

Complete assignment of the exchangeable protons based on their connectivities to other exchangeable and nonexchangeable protons has been carried out and the observed chemical shifts are given in Table 2.

N4 substituted cytidines generally have a preferred conformation for the substituent group *syn* relative to N3 [16–19]. In this conformation of the base the amino form of mo^4C could only form one hydrogen bond with guanine, between the G N1 imino proton and

the mo^4C C2 carbonyl. On the other hand, the imino form of the base can form two hydrogen bonds in which the bases are in wobble geometry (Fig. 2e) and this is what we observe. In this wobble geometry no constraints are imposed on the methoxy group. We observe a strong NOE between the methyl group and the N3 imino proton of mo^4C at 10.00 ppm confirming that the methoxy group remains *syn* (Fig. 4A). In the *anti* form this NOE would not be expected. Additionally, we observe NOEs between the methyl group and the C10 amino protons. By spin diffusion weak NOEs are also observed to the G11 and G5 imino protons.

Table 2
Chemical shifts of nonexchangeable protons at 10°C and of exchangeable protons at 1°C for N^4 -methoxycytosine-guanine duplex.

The chemical shifts for methoxy methyl of mo^4C are 3.29 ppm and 3.47 ppm for the amino and imino form, respectively. The non-assigned resonances are marked a).

	H8/H6	H5/H2 CH3	H1'	H2'	H2''	H3'	H4'	NH	NH2
C1a	7.61	5.86	5.52	1.87	2.30	4.66	4.02	–	8.18/7.00
A2a	8.24	7.77	6.01	2.76	2.90	5.02	4.37	–	8.00/6.61
A2i	8.23	7.77	5.97	2.76	2.89	5.02	4.37	–	7.88/6.62
G3a	7.70	–	5.74	2.50	2.60	4.97	4.39	12.85	–
G3i	7.69	–	5.83	2.48	2.55	4.94	a)	12.62	–
mo^4Ca	7.47	5.71	5.76	1.83	2.37	4.78	4.18	11.25	–
mo^4Ci	6.64	5.31	5.46	1.72	1.98	4.77	3.99	10.00	–
G5a	7.85	–	5.55	2.70	2.70	4.95	4.40	13.11	–
G5i	7.86	–	5.51	2.74	2.74	4.97	a)	13.40	–
G6a	7.71	–	5.96	2.49	2.71	4.96	a)	13.11	–
G6i	7.78	–	5.94	2.53	2.69	4.99	4.40	13.11	–
C7	7.20	4.95	6.06	2.22	2.22	4.48	3.96	–	8.18/6.42
G8	7.90	–	5.91	2.62	2.73	4.81	4.28	12.95	–
C9a	7.46	5.24	6.07	2.15	2.51	4.86	4.24	–	8.29/6.40
C9i	7.74	5.25	6.03	2.09	2.43	4.83	a)	–	8.29/6.40
C10a	7.46	5.58	5.58	2.10	2.41	4.86	4.40	–	8.56/6.97
C10i	7.54	5.59	5.47	2.28	2.39	4.86	a)	–	8.58/6.82
G11a	7.95	–	5.94	2.67	2.67	4.50	a)	12.80	–
G11i	7.88	–	6.06	2.58	2.79	4.99	a)	10.76	–
C12a	7.45	5.41	5.95	2.00	2.46	4.77	4.20	–	8.29/6.75
C12i	7.42	5.48	5.87	1.90	2.40	4.77	a)	–	8.27/6.95
T13a	7.35	1.66	5.81	1.98	2.34	4.85	4.13	14.19	–
T13i	7.35	1.69	5.81	1.98	2.34	4.85	4.13	14.22	–
G14	7.91	–	6.13	2.64	2.37	4.69	4.18	12.95	–

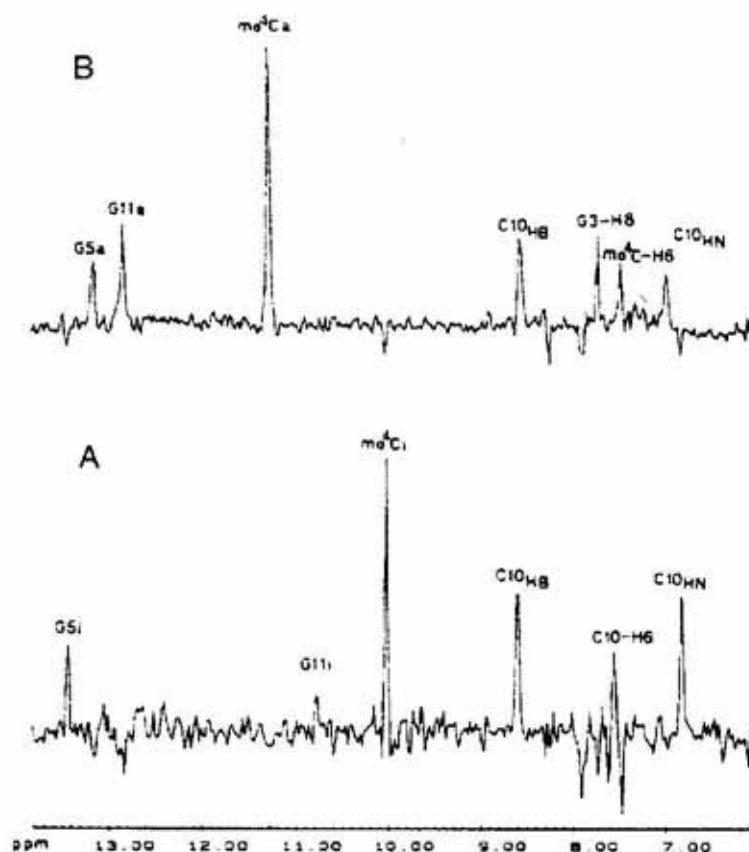


Fig. 4. The crosssection of the 200 ms NOESY spectrum recorded in 90% $H_2O/10\% D_2O$ at pH 6.2 and $1^\circ C$. Rows are taken through the methoxy methyl of the A) imino and B) amino forms of N^4 -methoxycytosine-guanine. Indexes HB and HN correspond to the hydrogen bonded and non hydrogen bonded amino protons of C10, respectively.

For mo^4C in the amino form the methoxy group has to rotate *anti* to N3 for stable hydrogen bonding to occur. Three hydrogen bonds can be formed between the bases which are in Watson-Crick geometry (Fig. 2d). We observe NOEs from the mo^4C methyl group to both amino protons of C10 and also to the G3 H8 proton (Fig. 4B) which suggests that there is rotational freedom about the N-OCH₃ bond and that the methyl group can be close to either the G3·C12 or G5·C10 base pairs.

NOESY spectra for the duplex $mo^4C \cdot G$ were recorded in D_2O at 10, 15, 20 and $30^\circ C$. Only for the spectra recorded at $10^\circ C$ and $15^\circ C$ could we follow the sequential H6/H8-H1' proton pathway without ambiguity. At these temperatures all of the H1' and H6/H8 resonances could be identified.

Figure 5 shows a region of the NOESY spectrum recorded at $10^\circ C$. For clarity, the sequential H8/H6-H1' assignments for the amino form, $mo^4Ca \cdot G$ and the imino form, $mo^4Ci \cdot G$ are shown separately. On the basis of the TOCSY spectrum we can attribute the cross-

peaks marked X to the intranucleotide CH5-CH6 interactions.

The most convenient starting point for the sequential assignment is from the resonance at 6.64 ppm which can be attributed to mo^4Ci H6 on the basis of its chemical shift. Without any ambiguity we can follow the connectivities in the 5' direction to C1i (Fig. 5A). In the other direction the connectivities are not unambiguous due to poor spectral resolution. This difficulty can be resolved by an analysis of the region corresponding to the interactions H6/H8-H2'/H2'' (not shown).

At $10^\circ C$ the amino form is more populated than the imino form. This helps us to discriminate between these two forms as we expect the crosspeaks from the imino form to be of lower intensity. With this in mind the connectivities on the other strand can be followed from G14i up to G11i. The ambiguity arising from the overlapping G11i H8 and G5i H8 resonances can be resolved referring again to the H8/H6-H2'/H2'' region and the sequential pathway can be traced as shown on Fig. 5A.

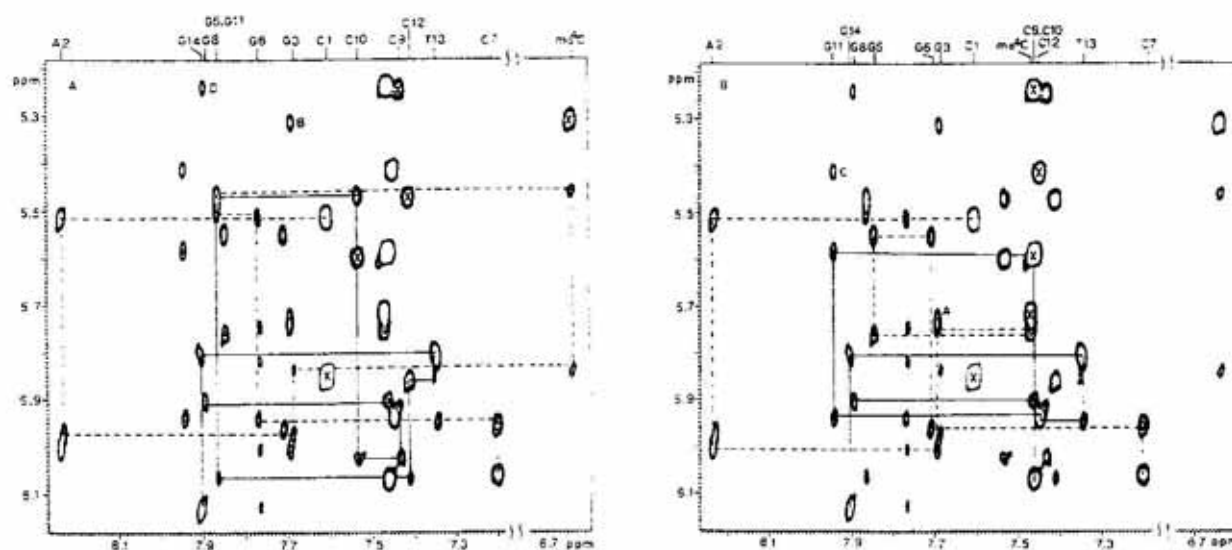


Fig. 5. Part of the 400 ms NOESY spectrum of the N^4 -methoxycytosine-guanine duplex recorded at 10°C . The region corresponds to interactions between base protons and H1'/H5 protons. Sequential assignment for the imino and amino form of N^4 -methoxycytosine-guanine duplex are traced separately in panels A and B, respectively. Cross peaks marked with an X correspond to CH6-CH5 interactions.

For the amino form we start with the terminal 5'-C1a and follow the sequential connectivities up to G3a (Fig. 5B). From the other end of the strand we can follow the chain up to mo^4Ca . The ambiguity in tracing the G3a- mo^4Ca -G5a chain arises from the overlapping of the interbase mo^4Ca H5-G3a H8 crosspeak with the G3a H8-G3a H1' crosspeak. Also both the intra and internucleotide mo^4Ca H6-H1' NOEs overlap. The H1'/H5-H2'/H2'' interactions, (not shown) enable precise assignment of the chemical shifts of the H1' resonances where overlap occurs and we can trace the pattern of sequential NOEs for this strand as shown in Fig. 5.

For the other strand the connectivities can be followed, with the aid of other regions of the spectrum, without interruption although the intraresidue C10 H1' peak is coincident with the H5-H6 crosspeak and the H6 protons of C9 and C10 are coincident. We do not expect large chemical shift differences for the H5 protons of the C residues between the two forms. We can thus assign the CH5-CH6 crosspeak next to that of C9i to that of C9a and the other at 7.46 ppm to that of C10a.

Four crosspeaks are observed with a resonance at 7.77 ppm (Fig. 5) corresponding to interactions of A2a H2 with the H1' protons of G3a, G14a, T13a and its own H1'. We cannot assign the A2i H2 on the basis of the NOESY

spectrum. However, it was found at 7.87 ppm in the NOESY spectrum in H_2O .

In order to determine the stacking we have examined the aromatic region of the NOESY spectrum and we have not found any significant perturbation in either duplex.

In Fig. 5 four interbase cross-peaks corresponding to H8/H6-H5 interactions, typical for right-handed B-DNA, are observed (peaks A-D). Three of these interbase crosspeaks arise from the central base pairs: C4a-G3a, C4i-G3i and C12a-G11a, peaks A-C, respectively. Peak D corresponds to C9a,i-G8.

The H2' and H2'' resonances were assigned from analysis of the NOESY spectrum recorded with a 30 ms mixing time. We were not able to assign all the H4' resonances because of strong signal overlap.

Detailed analysis of 2QF-COSY spectra of the mo^4C -G duplex was impossible because of strong resonance overlapping. Also some of signals from the imino form were too low in intensity to give precise values of spin-spin coupling constants. However, we could measure the sum of $J_{\text{H1}'\text{-H2}'}$ and $J_{\text{H1}'\text{-H2}''}$ for the mo^4C residues in both forms. This was found to be greater than 14.6 Hz in both cases indicating a predominantly C2' *endo* conformation. The ratio of the aromatic proton to H2'/H3' NOE also gives a measure of the C2' to C3' *endo* conformations. With the exception

of the C7 residue the NOE with the H2' proton was found to be very much larger than with the H3' proton which confirms the presence of a predominantly C2' *endo* sugar conformations.

We have measured the internucleotide NOEs in the 30 ms NOESY spectrum and we observe that these are all in agreement with a globally B form DNA in which the mo^4C ·G bases are intrahelical and base paired.

Table 2 summarizes the observed chemical shifts for mo^4Ci ·G and mo^4Ca ·G duplexes.

The cross section of the 30 ms mixing time NOESY spectrum taken through methyl group of mo^4Ci ·G shows only an interaction with the H5 proton of C10i. This interaction is possible only when the methoxy group adopts the *syn* conformation relative to N3. The cross section through the methyl group of mo^4Ca reveals the occurrence of NOEs with the C4a H5 and the C3a H8 protons. This is in agreement with the previous data that, when mo^4C pairs with G in the amino form, the methoxy group is *anti*.

Additional crosspeaks were observed in the aromatic-aromatic and H1'/H5-H1'/H5 regions of NOESY spectra recorded above 20°C. We have recorded ROESY spectra in order to determine whether they arise from cross-relaxation or from chemical exchange. All these new crosspeaks are in phase with the diagonal and must therefore arise from chemical exchange between the two forms. Below 20°C we do not observe crosspeaks from exchange between the mo^4Ci ·G and mo^4Ca ·G duplexes as the exchange is too slow.

Melting of duplexes

In order to determine the melting temperature of the mo^4C ·A and mo^4C ·G duplexes we have recorded 1H NMR spectra in D₂O as a function of temperature. For the mo^4C ·A duplex we were not able to follow the chemical shift changes of many of the proton resonances between 35°C and 55°C due to excessive line broadening. At temperatures above 35°C many new resonances appeared, indicative of slow exchange between the helix and coil forms.

However, the mo^4C H6 proton resonance is well separated from the other base protons resonances. At 37°C a new signal at 6.68 ppm appeared and was attributed to the mo^4C H6 coil resonance. To characterize the helix-coil transition we have measured the ratio of integrals between helix and coil resonances for

mo^4C (H6) proton. We find that at 47°C these two species are equally populated.

All resonances of the imino form of mo^4C ·G show typical sigmoid profiles which correspond to the double strand–single strand transition. The apparent t_m was defined as the mid-point of the sigmoid melting curves and was found to be 39°C. Resonances corresponding to the amino form disappear above 40°C. At higher temperatures the imino, amino and coil forms are in rapid exchange.

DISCUSSION

We have examined the base-pairing properties of mo^4C incorporated into the central position of heptamer duplexes with A or G on the opposite strand. Our data demonstrate that, when paired with A, mo^4C is in the imino form in Watson-Crick geometry. This requires that the methoxy group is in the less favoured *anti* conformation. All the NOESY data show that the duplex adopts a regular B form conformation with only one interaction, that between the G5 H8 and the mo^4C H2'/H2'' protons being unusual. This may be due to a changed helical twist to better accommodate the methoxy group. This result is in agreement with a previous report [3] except that its authors concluded, on the basis of chemical shift arguments, that the Watson-Crick form was in rapid equilibrium with a wobble structure. Such a structure would significantly modify the interactions and we find no evidence of its existence in the sequence that we have studied. As we have observed strong NOEs from the methoxy group to protons on both the adjacent base pairs, it would appear that the methoxy group does not have a single highly preferred orientation but that rotation occurs about the N–OCH₃ bond.

The melting curves for this duplex show slow exchange between the duplex and coil forms. Two factors can slow down the exchange process: tautomerization and/or *syn-anti* isomerization of the methoxy group. However, it has been reported [16] that mo^4C is present in the imino form for the monomer and thus no proton migration is necessary for base pair formation. We can thus conclude that the *syn-anti* isomerization is the rate limiting step. For the mo^4C ·G duplex two different tautomeric

forms of mo^4C are observed in slow exchange with each other. The amino form of mo^4C pairs with G in Watson-Crick geometry and the imino tautomer forms a wobble type base pair with G. The ratio of the amino to imino form of $\text{mo}^4\text{C}\cdot\text{G}$ is temperature dependent. At low temperature the amino form is predominant. In this form the methoxy methyl group gives strong NOEs to both the C10 NH2 protons and to the G3 H8 proton which suggests that there is a certain rotational freedom about the N-OCH3 bond. At 45°C only the imino form was detected in solution although at this temperature rapid exchange with the single strand species is also present. For this form the predominant interaction observed is with the mo^4C NH which shows that the *syn* conformation is predominant. On the other hand, we do not observe a strong NOE between the methoxy methyl group and the CH5 proton which shows that the *anti* form of the wobble structure is absent or has only a very minor population in the duplex that we have studied. In the crystal structure analysis of a Z-form duplex containing $\text{mo}^4\text{C}\cdot\text{G}$ base pairs [4] only the imino form of mo^4C paired with G was observed. The base pairs are in the wobble conformation and the methoxy group adopts the *syn* conformation.

These data indicate that two factors influence the tautomeric and conformational equilibrium of mo^4C : the base, which is on the opposite strand, and the temperature.

For the $\text{mo}^4\text{C}\cdot\text{G}$ duplex typical sigmoid curves are observed for duplex melting. In the $\text{mo}^4\text{C}\cdot\text{G}$ duplex both *syn* and *anti* conformers of mo^4C can take part in base pair formation. With the mo^4C base in the preferred imino form, even with the methoxy group *syn* to N3 there is little or no steric hindrance to formation of a wobble base pair. After formation of a wobble pair a slow rearrangement to the Watson-Crick amino form can take place. These two forms are in slow equilibrium on a proton NMR time scale. As described above the slow rearrangement of the methoxy group into the *anti* conformation explains the observation of separate resonances for the two forms up to 40°C.

Studies concerning the specificity of mo^4C incorporation during polymerisation on natural templates have shown that mo^4dCTP can replace dTTP during DNA synthesis. The

mo^4dCTP analogue can also replace dCTP during primer elongation but with lower efficiency [20]. However, the frequency at which both $\text{mo}^4\text{C}\cdot\text{A}$ and $\text{mo}^4\text{C}\cdot\text{G}$ pairs were formed was markedly influenced by the sequence of the template. The *syn-anti* isomerism of mo^4C was considered as a factor of crucial importance in base-pair formation.

Our data, may also help to explain why mo^4C replaces mainly dTTP during the primer elongation. Formation of $\text{mo}^4\text{C}\cdot\text{A}$ pairs requires that the methoxy group adopts the *anti* conformation but the resulting base pair is in Watson-Crick geometry and the mo^4C base is in the preferred imino form. Base pairing with G can occur for the imino form and the *syn* methoxy conformation but only in a less favourable wobble conformation. For a base pair in Watson-Crick geometry it is necessary that the base adopts the amino form and that the methoxy group turns *anti*, both of which are disfavoured.

REFERENCES

1. Anand, N.N., Brown, D.M. & Salisbury, S.A. (1987) The stability of oligodeoxyribonucleotide duplexes containing degenerate bases. *Nucleic Acids Res.* **15**, 8167–8176.
2. Kong Thoo Lin, P. & Brown, D.M. (1989) Synthesis and duplex stability of oligonucleotides containing cytosine-thymine analogues. *Nucleic Acids Res.* **17**, 10373–10383.
3. Stone, M.J., Nedderman, A.N.R., Williams, D.H., Kong Thoo Lin, P. & Brown, D.M. (1991) Molecular basis for methoxyamine initiated mutagenesis. 1H nuclear magnetic resonance studies of base-modified oligodeoxynucleotides. *J. Mol. Biol.* **222**, 711–723.
4. Van Meervelt, L., Moore, M.H., Kong Thoo Lin, P., Brown, D.M. & Kennard, O. (1990) Molecular and crystal structure of d(CGCG mo^4CG): N⁴-methoxycytosine-guanine base-pairs in Z-DNA. *J. Mol. Biol.* **216**, 773–781.
5. Fazakerley, G.V., Gdaniec, Z. & Sowers, L.C. (1993) Base-pair induced shifts in the tautomeric equilibrium of a modified DNA base. *J. Mol. Biol.* **230**, 6–10.
6. Nedderman, A.N.R., Stone, M.J., Williams, D.H., Kong Thoo Lin, P. & Brown, D.M. (1993) Molecular basis for methoxyamine-initiated mutagenesis: 1H nuclear magnetic resonance studies of oligonucleotide duplexes containing

- base-modified cytosine residues. *J. Mol. Biol.* **230**, 1068–1076.
7. Bodenhausen, G., Kogler, H. & Ernst, R.R. (1984) Selection of coherence-transfer pathways in NMR pulse experiments. *J. Magn. Reson.* **58**, 370–388.
 8. Plateau, P. & Guéron, M. (1982) Exchangeable proton NMR without base-line distortion, using new strong-pulse sequences. *J. Am. Chem. Soc.* **104**, 7310–7311.
 9. Davis, D.G. & Bax, A. (1985) Assignment of complex 1H NMR spectra via two-dimensional homonuclear Hartmann-Hahn spectroscopy. *J. Am. Chem. Soc.* **107**, 2820–2821.
 10. Rance, M., Sørensen, O.W., Bodenhausen, G., Wagner, G., Ernst, R.R. & Wüthrich, K. (1983) Improved spectral resolution in COSY 1H NMR spectra of proteins via double quantum filtering. *Biochem. Biophys. Res. Commun.* **117**, 479–485.
 11. Shaka, A.J. & Freeman, R. (1983) Simplification of NMR spectra by filtration through multiple-quantum coherence. *J. Magn. Reson.* **51**, 169–173.
 12. Piantini, U., Sørensen, O.W. & Ernst, R.R. (1982) Multiple quantum filters for elucidating NMR coupling networks. *J. Am. Chem. Soc.* **104**, 6800–6801.
 13. Bothner-By, A.A., Stephens, R.L., Lee, J., Warren, C.D. & Jeanloz, R.W. (1984) Structure determination of a tetrasaccharide: transient nuclear Overhauser effects in the rotating frame. *J. Am. Chem. Soc.* **106**, 811–813.
 14. Bax, A. & Davis, D.G. (1985) Practical aspects of two-dimensional transverse NOE spectroscopy. *J. Magn. Reson.* **63**, 207–213.
 15. Cuniasse, Ph., Sowers, L.C., Eritja, A., Kaplan, B., Goodman, M.F., Cognet, J.A.H., Le Bret, M., Guschlbauer, W. & Fazakerley, G.V. (1989) Abasic frameshift in DNA. Solution conformation determined by proton NMR and molecular mechanic calculations. *Biochemistry* **28**, 2018–2026.
 16. Brown, D.M., Hewlins, M.J.E. & Schell, P. (1968) The tautomeric state of N(4)-hydroxy- and N(4)-amino-cytosine derivatives. *J. Chem. Soc. (C)* 1925–1929.
 17. Shugar, D., Huber, C.P. & Birnbaum, G. I. (1976) Mechanism of hydroxylamine mutagenesis. Crystal structure and conformation of 1,5-dimethyl-N⁴-hydroxycytosine. *Biochim. Biophys. Acta* **447**, 274–284.
 18. Engel, J.D. & von Hippel, P.H. (1974) Effects of methylation on the stability of nucleic acid conformations: studies at the monomer level. *Biochemistry* **13**, 4143–4158.
 19. Fazakerley, G.V., Kraszewski, A., Téoule, R. & Guschlbauer, W. (1987) NMR and CD studies on an oligonucleotide containing N⁴-methylcytosine. *Nucleic Acids Res.* **15**, 2191–2201.
 20. Reeves, S.T. & Beattie, K.L. (1985) Base-pairing properties of N⁴-methoxydeoxycytidine 5'-triphosphate during DNA synthesis on natural templates, catalyzed by DNA polymerase I of *Escherichia coli*. *Biochemistry* **24**, 2262–2268.