Volume 219, number 2, 464–468 FEB 04951 July 1987

# Sequence-dependence of the conformational changes induced by the 5-methyl cytosine in synthetic RNA oligomers

G. Bloch, J.M. Neumann, F. Babin* and T. Huynh-Dinh*

*Service de Biophysique, Centre d'Etudes Nucléaires de Saclay, 91191 Gif-sur-Yvette* and **Unité de Chimie Organique, UA CNRS 487, Institut Pasteur, Département de Biochimie et Génétique Moléculaire, 28, rue du Docteur Roux, 75724 Paris Cedex 15, France*

Received 19 May 1987; revised version received 10 June 1987

The RNA hexamer containing a 5-methyl cytosine (m5C) r(CGUAm5CG) was studied by $^{1}H$ and $^{31}P$ NMR at 500 MHz and 121 MHz, respectively. In contrast to r(CGm5CGCG) which exhibits an atypic duplex structure [(1987) J. Am. Chem. Soc. 109, 2539–2541], r(CGUAm5CG) adopts a classical A-type conformation. This result demonstrates that the influence of the m5C on the conformation of RNA hexamers is sequence-dependent.

$^{1}H$ NMR; $^{31}P$ NMR; NOESY spectrum; Atypic ribonucleotide; RNA oligomer polymorphism; 5-Methyl cytosine

## 1. INTRODUCTION

Recently, preliminary $^{1}H$ and $^{31}P$ NMR results have shown that the self-complementary hexamer containing a 5-methyl cytosine (m5C) r(CGm5CGCG) can adopt, in low salt solution, two duplex conformations in slow equilibrium, and that the duplex form predominant at room temperature markedly differs from an A-type double helix [1]. In parallel we studied, in low salt solution, by $^{1}H$ and $^{31}P$ NMR another methylated self-complementary hexamer including an A=U base pair: r(CGUAm5CG). The data concerning this second sequence are presented in this paper and demonstrate that the influence of the m5C on the duplex conformation of RNA oligomers is sequence-dependent.

Correspondence address: J.M. Neumann, Service de Biophysique, Centre d'Etudes Nucléaires de Saclay, 91191 Gif-sur-Yvette, France

## 2. MATERIALS AND METHODS

The RNA fragment r(CGUAm5CG) was synthesized in solution by the phosphotriester method. The purity of the hexamer was checked on a Nucleosil C18 analytical column and by polyacrylamide gel electrophoresis under denaturing conditions. Enzymatic digestion by snake venom phosphodiesterase verified that r(CGUAm5CG) had the expected 3′–5′ phosphodiester linkage.

r(CGUAm5CG) was dissolved in $^{2}H_2O$ containing 5 mM $Na_2HPO_4$ and 0.1 M NaCl, divalent ions being trapped by addition of EDTA (~0.1 mM). The pH was adjusted to 6.9 by adding small amounts of $NaO^{2}H$. The samples were lyophilized twice in $^{2}H_2O$ (in $H_2O$ for exchangeable proton studies) and finally redissolved in $^{2}H_2O$ (in 20% $^{2}H_2O$ + 80% $H_2O$ for exchangeable proton studies) to a final concentration of 5 mM.

$^{1}H$ NMR 500 MHz spectra were recorded on a



*Published by Elsevier Science Publishers B.V. (Biomedical Division)*

Bruker WM-500 spectrometer and referenced relative to internal 3-(trimethylsilyl)$^2H_4$-propionic acid (TMP). The two-dimensional NOESY spectrum was recorded using the standard pulse sequence (90°–$t$1–90°–$t$m–90°–$t$2) [2] with a mixing time $t$m of 0.3 s and upon irradiation of water peak except during the acquisition time. Exchangeable proton spectra were recorded by using a two pulse sequence (45°–$T$–45°) [3,4].

The $^{31}P$ NMR spectrum (121 MHz) was recorded on a Bruker MSL-300 spectrometer, $^{31}P$ chemical shifts were determined from phosphoric acid as an external reference.

## 3. RESULTS AND DISCUSSION

### 3.1. *Duplex structure at room temperature*

The imino proton regions of one-dimensional spectra recorded at various temperatures are shown in fig.1. At room temperature two resonances with equivalent intensities are observed. Saturation of the low field resonance gives a negative nuclear Overhauser effect (NOE) on the peak assigned to the adenine H2 proton (see below), indicating that the low field imino proton resonance corresponds to the central uridine N3H proton. The high field signal is thus the internal guanine imino resonance. A third resonance, readily assigned to the external guanine N1H proton, is detected at lower temperatures. The observation of two imino proton resonances at room temperature reflects the duplex structure of r(CGUAm5CG).

Fig.2 shows the 500 MHz NOESY spectrum of r(CGUAm5CG) in 0.1 M NaCl solution recorded at 25°C. Three sets of cross peaks are observed: (i) connections between H8/H6/H2 base protons and

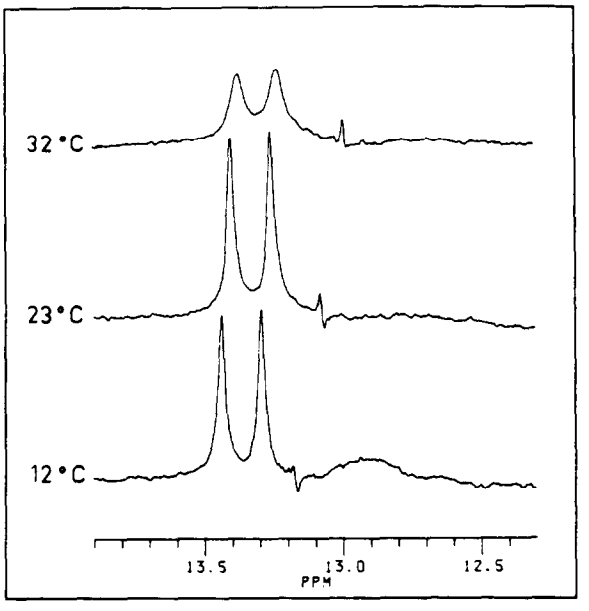


Fig.1. Exchangeable proton spectra (imino protons region) of r(CGUAm5CG) recorded at various temperatures in 80% $H_2O$ + 20% $^2H_2O$, 0.1 M NaCl, 5 mM $Na_2HPO_4$, pH = 6.9.

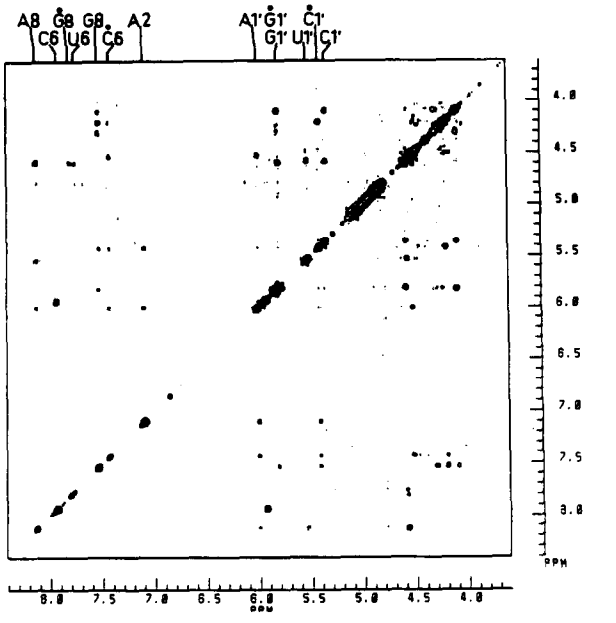


Fig.2. 500 MHz NOESY spectrum ($t$m = 0.3 s) of a 5 mM solution of r(CGUAm5CG) in $^2H_2O$, 0.1 M NaCl, 5 mM $Na_2HPO_4$, pH = 6.9, $t$ = 25°C, recorded upon irradiation of water peak except during the acquisition time. The following cross peaks labelling corresponds to the connectivities on fig.3: column A8: a = A8-U2′, b = A8-U1′, c = A8-A1′; column C6: d = C6-C5; column Ġ8: e = Ġ8-C2′; column U6: f = U6-Ġ2′; column G8: g = G8-G2′, h = G8-Ċ2′, i = G8-G3′, j = G8-Ċ1′, k = G8-G1′; column Ċ6: l = Ċ6-Ċ2′, m = Ċ6-A2′, n = Ċ6-Ċ1′, o = Ċ6-A1′; column A2: p = A2-Ċ1′, q = A2-A1′; column A1′: r = A1′-A2′; column Ġ1′/G1′: s = G1′-G2′, t = G1′-G3′, u = Ġ1′—Ġ2′; column U1′: v = U1′-U2′; column Ċ1′: w = Ċ1′-Ċ2′; column C1′: x = C1′-C4′ (tentative assignment); y = C1′-C2′. The columns are labelled according to a simplified rule defined as follows: C(1)G(2)U(3)A(4)C(5)C(6) = CGUACG and e.g. G(2)H8 = G8.

H1′/H5 protons; (ii) connections between H8/H6/H2 base protons and non-H1′ sugar protons; (iii) connections between H1′ protons and non-H1′ sugar protons. The assignment of r(CGUAm5CG) base, H1′ and H2′ non-exchangeable proton resonances at 25°C is achieved by the standard procedure using the NOE connectivities described in fig.3. The chemical shifts (relative to TMP) at 25°C of the assigned proton resonances are indicated in fig.3.

Despite the relative lack of two-dimensional NMR data concerning RNA oligomers an example of a standard A helix is available: the RNA fragment r(CGCGCG) adopts at room temperature and in 0.1 M NaCl solution an A-type conformation [5,6]. It is therefore useful to examine the r(CGUAm5CG) duplex structure in comparison with that of r(CGCGCG).

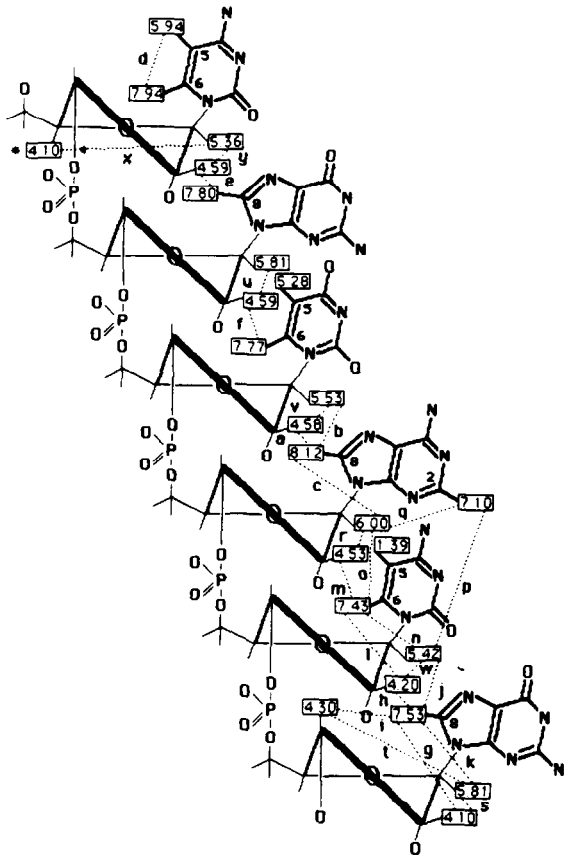

Fig.3. Schematic representation of one strand of the r(CGUAm5CG) duplex. The chemical shifts (in ppm relative to TMP) of the assigned proton resonances are given in the small rectangles. Each dotted line labelled with a small letter corresponds to a cross peak in the NOESY spectrum (fig.2). The resonance labelled with an asterisk is only tentatively assigned.

The r(CGUAm5CG) NOE pattern deduced from the NOESY spectrum recorded at 25°C (fig.2) is quite similar to that observed for r(CGCGCG) [5,6]: each H8/H6 base proton (except of course that of the C(1) residue) displays an internucleotide NOE connection with the H2′ proton of the 5′-neighbouring sugar. Moreover the cross peak corresponding to the internucleotide base proton-H2′ proton connection is more intense than that corresponding to the intranucleotide base proton-H1′ proton connection, indicating that the internucleotide distance is shorter than the intranucleotide distance. These features are characteristic of a right-handed A-type double helix, with an anti-type orientation of all the bases and a C3′-endo conformation of all the sugar rings.

The sugar puckers could be directly deduced from the resolution enhanced one-dimensional spectrum recorded at 25°C (not shown) [7]: all the residues but one display a $J$(H1′–H2′) scalar coupling constant lower than 1 Hz, characteristic of a C3′-endo conformation of the sugar ring. The higher $J$(H1′–H2′) value of the C(6) residue (2.5 Hz) corresponds to an intermediate conformation and is related to an end effect [5].

The A-type conformation of the r(CGUAm5CG) duplex is also confirmed at the sugar-phosphate backbone level: the similarity of r(CGUAm5CG) (fig.4) and r(CGCGCG) $^{31}$P spectra [5], recorded at room temperature in 0.1 M NaCl solution, indicates that both duplexes have similar sugar-phosphate backbone organizations.

In conclusion the $^{1}$H NOESY spectrum, $^{1}$H and $^{31}$P one-dimensional spectra of r(CGUAm5CG) recorded at room temperature demonstrate that the r(CGUAm5CG) duplex adopts, as r(CGCGCG) duplex, an A-type double helical conformation in 0.1 M NaCl solution.

### 3.2. *Duplex to single-strand transition*

The chemical shift versus temperature profiles of base and H1′ proton resonances (deduced from one-dimensional spectra recorded at various temperatures) are presented in fig.5. On increasing the temperature from 19 to 85°C most of the

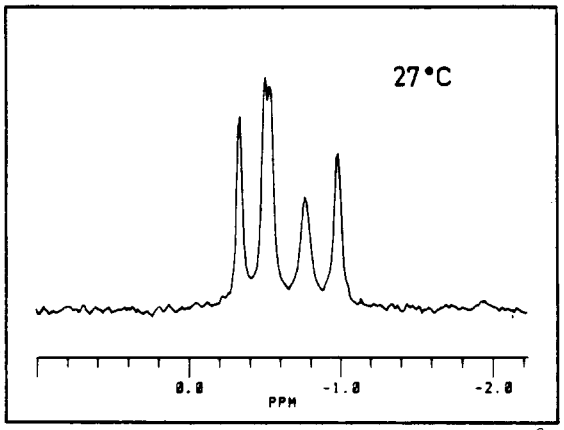


Fig.4. 121 MHz proton decoupled $^{31}P$ NMR spectrum of r(CGUAm5CG) (5 mM, 0.1 M NaCl, 5 mM $Na_2HPO_4$, pH = 6.9) in aqueous solution recorded at 27°C. Phosphoric acid was used as an external reference.

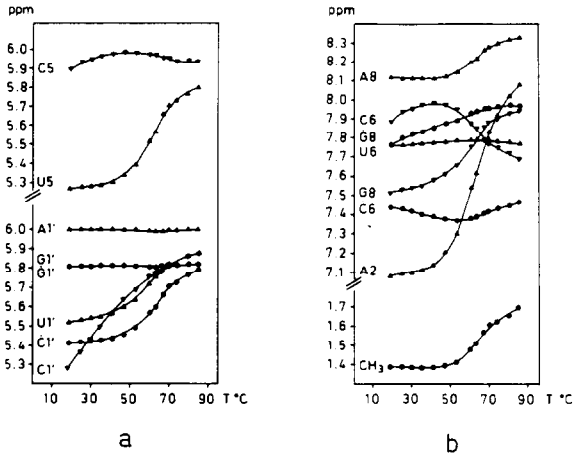


Fig.5. Temperature dependence of the proton chemical shifts (relative to TMP) of r(CGUAm5CG) (5 mM, 0.1 M NaCl, 5 mM $Na_2HPO_4$, pH = 6.9) in aqueous solution. (a) H1′ and H5 protons. (b) H8, H6, H2 and CH3 protons.

resonances undergo a sigmoidal downfield shift, which reflects the duplex to single-strand transition. From the most significant chemical shift versus temperature profiles relative to internal and central residues an average melting temperature of about 63°C is found. At 25°C the duplex proportion of r(CGUAm5CG) is practically 100%.

As expected, the r(CGUAm5CG) proton chemical shift versus temperature profiles are quite similar to those of the corresponding protons of r(CGCGCG) [5]. The melting temperatures $T_m$ of r(CGUAm5CG) (63°C) and r(CGCGCG) (65°C) in 0.1 M NaCl solution are also comparable. This feature needs to be examined since the first duplex contains two A=U base pairs, whereas the second duplex contains only C≡G base pairs. It is well known that in B-DNA hexamers the replacement of two C≡G by two A=T base pairs leads to a dramatic drop of the melting temperature, as illustrated in table 1. Nevertheless the presence of m5C increases the $T_m$ value of DNA duplexes whatever the sequence (table 1). Our data suggest that a similar and more intense effect exists in A-RNA hexamers.

### 3.3. *Sequence-dependence of methylated RNA oligomers duplex conformation*

Preliminary results indicated the existence, in 0.1 M NaCl solution, of two duplex conformations of r(CGm5CGCG) in slow exchange on the NMR time scale and showed that the duplex predominant at room temperature exhibits a structure different from an A-helix [1].

The polymorphism of r(CGm5CGCG) is obviously induced by the methyl group of the m5C, since the non-methylated sequence r(CGCGCG) only adopts the standard A-type duplex conformation. r(CGm5CGCG) atypic duplex conformation was interpreted in terms of propagation of a local distorsion, due to the methyl group, via the phosphate backbone and the base stacking. However the present data show that the methylated hexamer r(CGUAm5CG) does not ex-

Table 1

Average melting temperatures (internal and central base pairs) of DNA and RNA fragments in 0.1 M NaCl solution

| Oligomer | Reference | Concentration (mM) | Conformation | Melting temperature (°C) |
|---|---|---|---|---|
| d(CGCGCG) | [8] | 11 | B | 67 |
| d(m5CGCGm5CG) | [9] | 2 | B | 73 |
| d(GCATGC) | a | 3 | B | 50 |
| d(Gm5CATGm5C) | a | 3 | B | 56 |
| r(CGCGCG) | [5] | 7 | A | 65 |
| r(CGUAm5CG) | b | 5 | A | 63 |

(a) Neumann, unpublished work. (b) This work

hibit a conformational transition similar to that of r(CGm5CGCG). Moreover, the r(CGUAm5CG) duplex is found in a basic A-type conformation (even in 2 M NaCl solution or at low temperature (not shown)). Our results show that the influence of the m5C on RNA oligomers conformation is sequence-dependent and thus is governed by complex molecular mechanisms whose understanding needs careful comparison with other methylated RNA fragments.

## REFERENCES

[1] Ceolin, F., Babin, F., Huynh-Dinh, T., Igolen, J., Bloch, G., Tran-Dinh, S. and Neumann, J.M. (1987) J. Am. Chem. Soc. 109, 2539–2541.
[2] Macura, S. and Ernst, R.R. (1980) Mol. Phys. 41, 95–117.
[3] Plateau, P. and Guéron, M. (1982) J. Am. Chem. Soc. 104, 7310–7316.
[4] Hore, P.J. (1983) J. Magn. Reson. 54, 539–542.
[5] Westerink, H.P., Van der Marel, G.A., Van Boom, J.H. and Haasnoot, C.A.G. (1984) Nucleic Acids Res. 12, 4323–4338.
[6] Haasnoot, C.A.G., Westerink, H.P., Van der Marel, G.A. and Van Boom, J.H. (1984) J. Biomol. Struct. Dyn. 2, 345–360.
[7] Altona, C. (1982) Recl. Trav. Chim. Pays-Bas 101, 413–433.
[8] Patel, D.J. (1976) Biopolymers 15, 533–558.
[9] Tran-Dinh, S., Taboury, J., Neumann, J.M., Huynh-Dinh, T., Langlois d'Estaintot, B. and Igolen, J. (1984) Biochemistry 23, 1362–1371.