Volume 14 Number 10 1986 Nucleic Acids Research

Hairpin and duplex formation of the DNA octamer d(m^5C-G-m^5C-G-T-G-m^5C-G) in solution. An NMR study

Leonard P.M.Orbons, Gijs A.van der Marel, Jacques H.van Boom and Cornelis Altona*

Gorlaeus Laboratories, State University of Leiden, PO Box 9502, NL 2300 Leiden, The Netherlands

Received 3 April 1986; Accepted 1 May 1986

ABSTRACT

The partly self-complementary DNA octamer d(m^5C-G-m^5C-G-T-G-m^5C-G) was investigated by NMR spectroscopy in solution. It is demonstrated that this peculiar DNA fragment, under suitable conditions of concentration, salt and temperature, exclusively prefers to adopt a monomeric hairpin form with a stem of three Watson-Crick type base pairs and a loop of two residues. At high single strand concentration (8 mM DNA) and low temperature (i.e. below 295 K) the hairpin occurs in slow equilibrium with a B-dimer structure. At high ionic strength ($\geq$ 100 mM Na+) and/or in the presence of methanol a third species appears, which is assigned to a Z-like dimer. In the B form, as well as in the Z dimer, the two central base pairs form G.T wobble pairs with the bases as major tautomers.

INTRODUCTION

Studies regarding the effects of mismatches, i.e. non-Watson-Crick complementary base pairs, or mispairs, in the double helix, on the thermodynamical and structural features of DNA and RNA duplexes may provide more insight into the consequences of DNA/RNA mutations. For example, it is known from X-ray crystallography of single crystals (3,4) that the G.T mismatch in A, B and Z helices forms a hydrogen-bonded pair with the bases as major tautomers and in the anti orientation around the glycosyl bond. Similar results were obtained by Patel et al. (5), who investigated the NMR spectra of a B-DNA dodecamer containing the G.T mismatch. The vertical base stacking of the G.T wobble bases upon their nearest neighbours appears to be affected to some extent. Tibanyenda et al. (6) showed that replacement of a G.C base pair by a G.T wobble pair decreases the duplex stability. In order to obtain more information regarding the effects of G.T mismatches

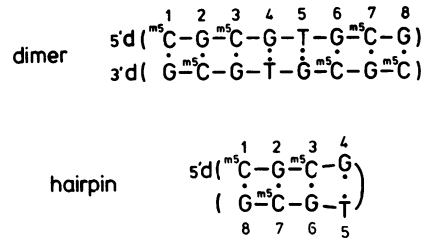

Fig. 1 Secondary structures accessible to the DNA octamer d(m^5C-G-m^5C-G-T-G-m^5C-G).

on the structure and stability of B and Z DNA, an NMR study of the partly self-complementary octamer d(m^5C-G-m^5C-G-T-G-m^5C-G) was initiated.

In the present report it will be shown that this peculiar octamer, under suitable conditions of concentration and ionic strength, exclusively prefers to adopt a monomeric hairpin (Fig. 1) in which G(4) and T(5) form a loop of two residues. At high single-strand concentration (8 mM) the major hairpin occurs in slow equilibrium with a minor B-DNA dimer structure with two central G.T wobble base pairs (Fig. 1). At high salt concentrations ($\geq$ 100 mM Na+) a third species is observed, which is stabilized by methanol. This new form, which occurs in slow exchange with the hairpin and with the B-DNA dimer, is identified with a Z-like duplex structure with two central G.T wobble pairs.

MATERIALS AND METHODS

The DNA octamer was synthesized via an improved phosphotriester method (7,8) and was treated with Dowex cation exchange resin (Na+ form) to yield the sodium salt.

Two NMR samples, which contained 0.4 and 8 mM DNA single-strand concentration respectively, were lyophilized three times from 99.75 % ^2H_2O and finally taken up in 99.95 % ^2H_2O; the pH was adjusted to 7.0 (meter reading); DSS was added as an internal reference and EDTA in order to neutralize effects due to paramagnetic impurities. In order to observe the exchangeable protons, a second set of samples was prepared in a $H_2O/^2H_2O$ (90/10) mixture. The DNA concentrations, again, were 0.4 and 8 mM.

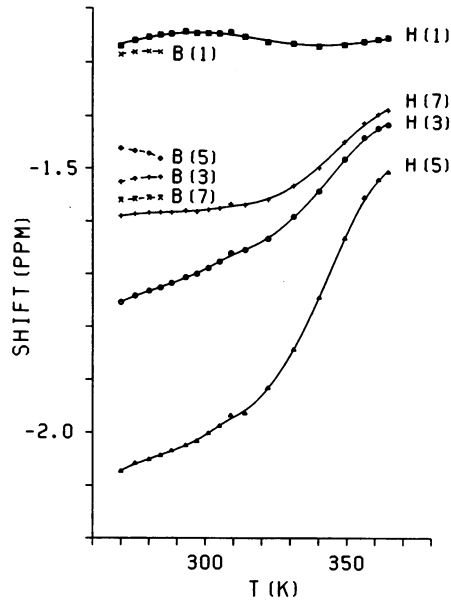

Fig. 2 Chemical shifts versus temperature profiles of the methyl-proton resonances of cytosine 1, 3, 7 and thymine 5 in d(m^5C-G-m^5C-G-T-G-m^5C-G); pH 7.0, 8 mM single-stranded DNA; H, methyl resonances of the hairpin form; B, methyl resonances of the B dimer.

Imino-proton spectra were recorded with a time-shared pulse as described elsewhere (9). ^{1}H-NMR spectra were recorded on a Bruker WM-500 and on a WM-300 spectrometer, both instruments were interfaced with an ASPECT-2000 computer.

RESULTS AND DISCUSSION

Non-exchangeable protons

Fig. 2 shows the chemical shifts of the methyl protons of the cytosine residues 1,3, 7 and of the thymine residue 5, as a funtion of the temperature (total single strand concentration 8 mM). Apart from the 5'-terminal residue, all methyl protons show sigmoidal chemical shift/temperature profiles, which can be analyzed in terms of a two-state helix-to-coil equilibrium; an average T_m of 332 K was calculated from the experimental curves. At lower temperatures (i.e. below 300 K) a second species B is observed (Fig. 2), which is in slow exchange on the NMR time

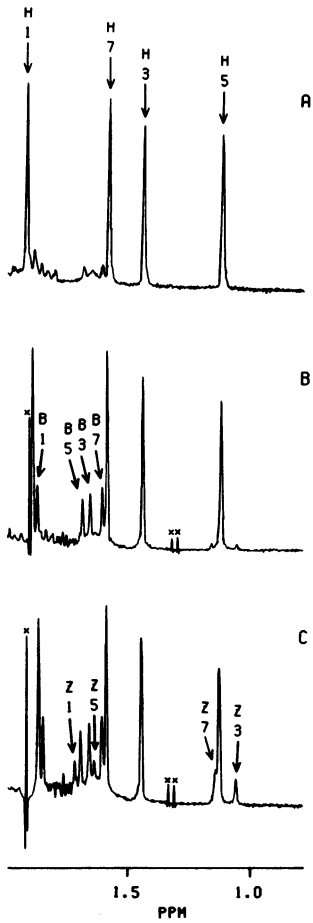

Fig. 3 High-field part of the ^{1}H-NMR spectra of d($m^{5}C$-G-$m^{5}C$-G-T-G-$m^{5}C$-G) in $^{2}H_2O$ at 284 K, showing the methyl resonances; A, 500 MHz spectrum, no salt added; B, 300 MHz spectrum, 100 mM NaCl; C, 300 MHz spectrum, 200 mM NaCl. The signals marked H, B and Z correspond to the methyl-resonances of the hairpin, B and Z dimer respectively. Impurity peaks are marked with an asterisk.

scale with the major component. In the remainder of this paper the major component will be designated H. The fraction of species B increases to approximately 20 % at 275 K. At a twenty-fold lower DNA concentration (0.4 mM) the methyl resonances of the thymine and cytosine residues show a chemical shift/temperature behaviour, which is virtually identical with that dis-

played by the major species H at high DNA concentration; an average T_m of 333 K was derived. However, at this low DNA concentration - even at 270 K - the second species B was not observed.

Addition of 100 mM NaCl to the 8 mM DNA sample causes an increase of the second form B to about 35 % at 275 K (Fig. 3B). This minor species B, which occurs specifically at higher DNA concentration and appears stabilized by increase of the ionic strength, can be identified safely with a DNA-dimer structure in the B form. The properties mentioned accord with those reported by Wemmer et al. (10) and Marky et al. (11). The major species H, T_m of which is not affected by the single-strand concentration, will be referred to as the hairpin structure. Upon further increase of the NaCl concentration, i.e. above 100 mM, a third form (Z) appears (Fig. 3C), which disappears on lowering the single-strand concentration. This third species Z is stabilized by addition of methanol and is assigned to a Z-like dimer. At this point it should be mentioned that all base proton, 5-CH_3, H-1', H-2', H-2", H-3' and H-4' resonances of both the hairpin form (H) and of the Z form (Z) have been assigned by means of standard COSY and NOESY experiments. The B form proved to be more elusive, but its base proton and 5-CH_3 resonances could be assigned beyond doubt. These assignments, together with conformational arguments concerning the details of the three different species discussed above and model building, will be described elsewhere (L.P.M. Orbons et al., manuscript in preparation). For example, the NOESY spectrum of the Z form clearly indicated the syn position of the G bases, whereas the C and T bases remain anti.

Exchangeable protons

Fig. 4A, B and C show the low field part of the ^{1}H-NMR spectra of the 8 mM sample at various pH values (no salt added), recorded in $H_2O/^2H_2O$ (90/10). Again two molecular species existing side by side are observed at low temperatures. The weaker signals of the B species (Fig. 4A, indicated by arrows) disappear upon lowering the nucleic acid concentration (Fig. 4D). Aside from the three normal Watson-Crick hydrogen-bonded imino protons displayed by the B-DNA species B, which resonate around 13.5 ppm,

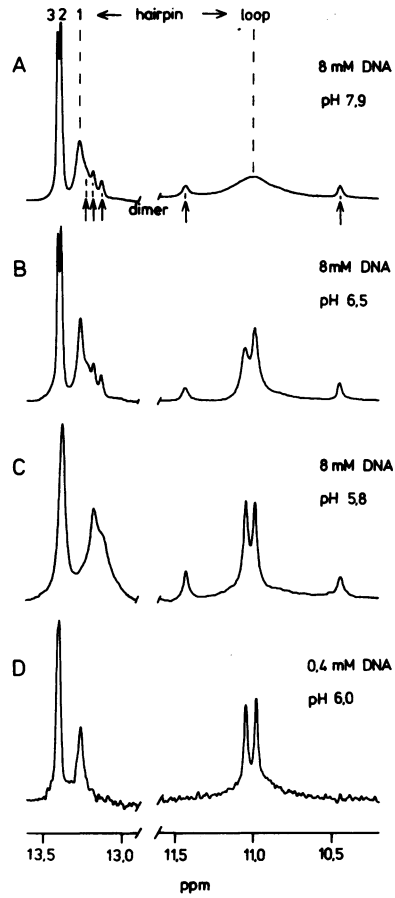

Fig. 4 500 MHz imino-proton spectra of d(m^5C-G-m^5C-G-T-G-m^5C-G) in $H_2O/^2H_2O$ (90/10 v/v) at 272 K; A, 8 mM DNA, pH 7.9; B, 8 mM DNA pH 6.5; C, 8 mM DNA, pH 5.8; D, 0.4 mM DNA, pH 6.0. Arrows indicate the B-type dimer. Numbers refer to the assignment of the imino protons of the hairpin.

two other imino-proton resonances belonging to this species are observed at higher field (10.44 ppm and 11.42 ppm). The chemical shifts of these protons agree well with earlier observations concerning G.T wobble pairs; 10.58 and 11.78 ppm, Patel et al. (5); 10.0 and 11.6 ppm, Tibanyenda at al. (6). Following Uesugi et al. (12), the resonance at 11.42 ppm is assigned to the imino proton of T(5) and the resonance at 10.44 ppm to the imino proton of G(4), respectively.

The imino protons of base pairs 1, 2 and 3 of the major hairpin form H resonate at approximately 13.5 ppm (Fig. 4A). By means of a NOESY experiment these resonances were assigned via interbase-pair NOEs between the imino protons. It may be noted that the imino proton resonance of residue 1 is broadened due to fraying effects (13). All three imino protons show NOEs to the amino protons of their base-paired cytosine residues, which at their turn exhibit intranucleotide NOEs to their methyl protons. These facts indicate that all three imino protons are hydrogen-bonded in Watson-Crick type base pairs. The remaining two imino protons of the hairpin H, assigned to T(5) and G(4), resonate at higher field (11.05 ppm and 10.98 ppm). A specific assignment of both resonances is not possible at this stage. The line widths of these peculiar proton resonances show a pronounced pH dependence, in contrast to the behaviour of the remaining imino signals. On increasing the pH from 5.8 to 7.9 the line width of these two signals at $\sim$11 ppm strongly increases (Fig. 4A, B and C), which fact indicates that the rate of exchange with water and not the life time of the molecular species itself governs the line width of these imino-proton resonances (14).

Recently Haasnoot et al. (15) have shown that one of the major characteristics of DNA hairpins - aside from the concentration independence of T_m - is constituted by the relatively fast exchange of the imino protons located in the hairpin loop, which is limited by the exchange with water as is observed here. In contrast, the imino protons of the fully duplexed structure usually display an opening-limited behaviour. The resonance positions of the two imino protons at 11.05 and 10.98 ppm also agree with expectations based upon earlier report concerning DNA hairpins (15,16,17).

Fig. 5 shows the imino-proton resonances in a $H_2O/C^2H_3O^2H$ (55/45 v/v) solution containing 8 mM DNA and 5 mM $MgCl_2$ at 270 K. Under these conditions the title compound essentially adopts a pure Z-like conformation (approximately 5 % hairpin (H) and less than 1 % B-type dimer (B) are still present). The resonances at 13.48, 12.25 and 13.04 ppm are assigned by means of a NOESY experiment similar to the one described for the hairpin (vide infra). Again, these three downfield signals are identified with

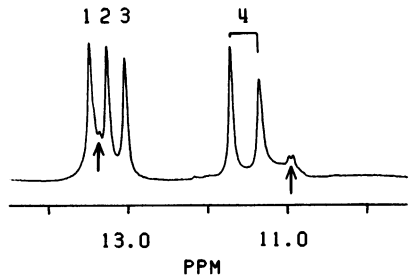

Fig. 5 500 MHz imino-proton spectrum of d(m^5C-G-m^5C-G-T-G-m^5C-G) in $H_2O/C^2H_3O^2H$ (55/45 v/v) at 270 K; 8 mM DNA, 5 mM $MgCl_2$; pH 7.0. Numbers refer to the assignment of the imino protons in the Z form. Arrows indicate imino-proton resonances arising from the hairpin structure.

imino protons that are hydrogen bonded in normal Watson-Crick type base pairs. The more upfield shifted imino resonances at 11.73 and 11.36 ppm are designated to the wobble pairs T(5) and G(4). However, a specific resonance assignment can not be presented at the present stage. Both imino protons show a strong NOE to each other, similar to observations on the G.T wobble pair in the B dimer (not shown). It should be noted that none of the five imino-proton resonances of the Z form display line broadening upon increasing the pH. It is suggested that in solution, the G.T mismatch in the Z-form duplex adopts a G.T wobble pair with the bases as the major tautomers, similar to the B form discussed above. The chemical shift difference between the imino-proton resonances of the G.T wobble pairs in the B and Z dimers (Fig. 4 and 5) appears as a consequence of the different base stacking in both DNA forms.

Conclusion

Aside from the single-stranded random coil (high temperature form) the DNA octamer d(m^5C-G-m^5C-G-T-m^5C-G), under suitable conditions of concentration, ionic strength and methanol, can adopt either one of three different molecular species, which occur in slow exchange with each other. At low ionic strength the fragment prefers to adopt a hairpin structure with a stem of three Watson-Crick type base pairs and a loop of only two residues. This in contrast with an earlier postulation of Tinoco et al. (18) and recent results of Haasnoot et al. (19). Haasnoot et al.

showed that the optimal loop for DNA consists of four to five residues, even if this must be accomplished by breakage of an A.T base pair in the stem. For the title compound a loop of four residues would imply breakage of a G.C pair in the stem, but apparently this third base pair in combination with a loop of two residues is more favoured. However, model building studies reveal a certain resemblance, as far as stacking continuity is concerned, between a loop of four residues (C.A.G. Haasnoot et al., private communication) and one which consists of two residues and is closed by a G.C base pair (L.P.M. Orbons et al., manuscript in preparation). At 8 mM single strand concentration the hairpin appears in slow equilibrium with a B-type dimer (80 / 20 at 275 K, no salt added). After addition of methanol and $MgCl_2$ a third species predominates, which has Z-type characteristics.

ACKNOWLEDGEMENT

This research was supported by the Netherlands Foundation of Chemical Research (SON) with financial aid from the Netherlands Organisation for the Advancement of Pure Research (ZWO). NMR spectra were recorded at the Dutch National HF-NMR Facility Nijmegen (P.A.W. van Dael and J.J.M. Joordens) and at the NMR facility at the university of Leiden (C. Erkelens). We wish to thank Dr. C.A.G. Haasnoot for stimulating discussions.

*To whom correspondence should be addressed

REFERENCES

1. This paper is number 44 in a series Nucleic Acid Constituents from this laboratory; for part 43 see Rinkel, L.J., Sanderson, M.R, van der Marel, G.A., van Boom, J.H. and Altona C. (1986) Eur. J. Biochem., submitted for publication
2. Abbrevations used: NOE, Nuclear Overhauser Effect; NOESY, two dimensional nuclear Overhauser spectroscopy; COSY, correlation spectroscopy; DSS, 4,4-dimethyl-4-sila-pentanesulphonic acid sodium salt.
3. Brown, T., Kennard, O., Kneale, G. and Rabinovich, D. (1985) Nature 315, 604-606.
4. Kennard, O. (1985) J. Biomol. Struct. Dyns. 3, 205-226.
5. Patel, D.J., Kozlowski, S.A., Marky, L.A., Rice, J.A., Broka, C., Dallas, J., Itakura, K. and Breslauer, K.J. (1982) Biochemistry 21, 437-444.
6. Tibanyenda, N., de Bruin, S.H., Haasnoot, C.A.G., van der Marel, G.A., van Boom, J.H. and Hilbers, C.W. (1984) Eur. J. Biochem. 139, 19-27.

7. Van der Marel, G.A., van Boeckel, C.A.A., Wille, G. and van Boom, J.H. (1981) Tetrahedron Lett. 22, 3887-3890.
8. Van der Marel, G.A., Wille, G., Westerink, H., Wang, A.H.-J., Rich, A., Mellema, J.R., Altona, C. and van Boom, J.H. (1982) J. Royal Netherlands Chem. Soc. 101, 77-78.
9. Haasnoot, C.A.G. and Hilbers, C.W. (1983) Biopolymers 22, 1259-1266.
10. Wemmer, D.E., Chou, S.H., Hare, D.R. and Reid, B.R. (1985) Nucl. Acids Res. 13, 3755-3772.
11. Marky, L.A., Blumenfeld, K.S., Kozlowski, S. and Breslauer, K.J. (1983) Biopolymers 22, 1247-1257.
12. Uesugi, S., Nakagawa, E., Ohtsuka, E., Ikehara, M., Watanabe, M., Kobayashi, Y., Kyogoku, Y. and Kainosho, M. (1982) J. Am. Chem. Soc. 104, 7340-7341.
13. Patel, D.J. and Hilbers, C.W. (1975) Biochemistry 14, 2651-2656.
14. Hilbers, C.W. in Biological Applications of Magnetic Resonance, Ed. Shulman, R.G., Academic Press, New York (1979), 1-43.
15. Haasnoot, C.A.G., de Bruin, S.H., Berendsen; R.G., Janssen, H.G.J.M., Binnendijk, T.J.J., and Hilbers, C.W. (1983) J. Biomol. Struct. Dyns. 1, 115-129.
16. Haasnoot, C.A.G., den Hartog, J.H.J., de Rooij, J.F.M., van Boom, J.H. and Altona, C. (1979) Nature 281, 235-236.
17. Haasnoot, C.A.G., den Hartog, J.H.J., de Rooij, J.F.M., van Boom, J.H. and Altona, C. (1980) Nucl. Acids. Res. 8, 169-181.
18. Tinoco, I., Uhlenbeck, O.C. and Levine, M.D. (1971) Nature 230, 362-367.
19. Hilbers, C.W., Haasnoot, C.A.G., de Bruin, S.H., Joordens, J. J.M., van der Marel, G.A. and van Boom, J.H. (1985) Biochimie 67, 685-695.