4116–4121 Nucleic Acids Research, 1995, Vol. 23, No. 20

NMR characterisation of a triple stranded complex formed by homo-purine and homo-pyrimidine DNA strands at 1:1 molar ratio and acidic pH

Sukesh R. Bhaumik, Kandala V. R. Chary, Girjesh Govil, Keliang Liu[1] and H. Todd Miles[1]

Chemical Physics Group, Tata Institute of Fundamental Research, Homi Bhabha Road, Bombay 400 005, India and [1]National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA

Received June 26, 1995; Revised and Accepted September 6, 1995

ABSTRACT

Homo-purine (d-TGAGGAAAGAAGGT) and homo-pyrimidine (d-CTCCTTTCTTCC) oligomers have been designed such that they are complementary in parallel orientation. When mixed in a 1:1 molar ratio, the system adopts an antiparallel duplex at neutral pH with three mismatched base pairs. On lowering the pH below 5.5, a new complex is formed. The NMR results show the coexistence of a intermolecular pyrimidine.purine:pyrimidine DNA triplex and a single stranded oligopurine at this pH. The triplex is stabilized by five T.A:T, four C^+.G:C and two mismatched triads, namely, C^+.G–T and T.A–C. This triplex is further stabilized by a Hoogsteen C^+.G base-pair on one end. Temperature dependence of the imino proton resonances reveals that the triplex dissociates directly into single strands around 55°C, without duplex intermediates. Parallel duplexes are not formed under any of the conditions employed in this study.

INTRODUCTION

The sequence-specific recognition of DNA duplexes by a third strand (1–12), has implications in gene-regulation and site specific cleavage of genomic DNA (13–15). There have been several NMR studies on the structural characterization of pyrimidine(Y).purine(R):pyrimidine(Y) DNA triplexes (throughout this paper, Watson and Crick base pairs are represented as R:Y, Hoogsteen as Y.R and mismatch as R–Y). In such studies, purine and pyrimidine strand concentrations of 1:2 molar ratio have been used. The second pyrimidine strand is parallel to the purine strand and forms Hoogsteen base-pairing with the standard antiparallel Watson–Crick R:Y double helical DNA (16–20). A different class of triplexes are single-stranded oligonucleotide sequences that form intramolecular triple-stranded structures with well defined strand orientations (21–23).

Felsenfeld *et al.* (1) were the first to show that poly(rA) and poly(rU) mixed in 1:2 proportions form triple stranded polyribonucleotides with the third Y strand binding in the major groove, parallel to the purine strand of Watson–Crick double helical RNA, using Hoogsteen base-pairing [U.A:U]. In an attempt to form the DNA analogue, Lee *et al.* (7) found that a 1:1 molar mixture of $(GA)_n$ and $(TC)_n$ spontaneously transformed into a triplex together with a free purine strand at lower pH. Based on S1 nuclease digestion, buoyant density measurements, hyperchromic shifts and circular dichroism studies (7), the reaction has been characterized as:

$$2[(GA)_n]:[(CT)_n] \xrightarrow{H^+} [(C^+T)_n].[(GA)_n]:[(CT)_n] + [(GA)_n]$$

The ease with which this triplex forms is attributed to the moderately high pK_a of cytosine in DNA [4.1 for free cytidine (24)]. The base–base recognition is achieved through sequence specific hydrogen bonds between C^+ and G and between T and A (Fig. 1). Thus, when C^+ recognizes C:G base pair C^+.G:C triad (Fig. 1A) is formed, while when T recognises A:T base pair T.A:T triad (Fig. 1B) is formed. Although, the triplex formation at 1:1 molar ratio has been proposed (7), there has been no structural evidence.

We have carried out NMR structural characterization of a 1:1 mixture of a homopyrimidine strand (a 12mer; d-CTCCTTTCTTCC) and basically a homopurine strand (a 14mer; d-TGAGGAAAGAAGGT). This system has perfect purine–pyrimidine base complementarity when the two strands are aligned in parallel orientation. In fact, one of the objectives of this synthesis was to experimentally look for the formation of parallel duplexes. At neutral pH, the molecular system adopts an antiparallel duplex, stabilised by five A:T, four C:G base pairs, and three mismatched base pairs, namely G–T, A–C and T–C (to be published). Here we report the structure of the above system at low pH.

MATERIALS AND METHODS

Oligonucleotides were synthesized with Applied Biosystems Model 380 B DNA synthesizer using a solid-phase cyanoethyl-phosphoramidite method. Oligonucleotides were purified on 20% denaturing polyacrylamide/bis (19:1) gel with 7.6 M urea,

* To whom correspondence should be addressed

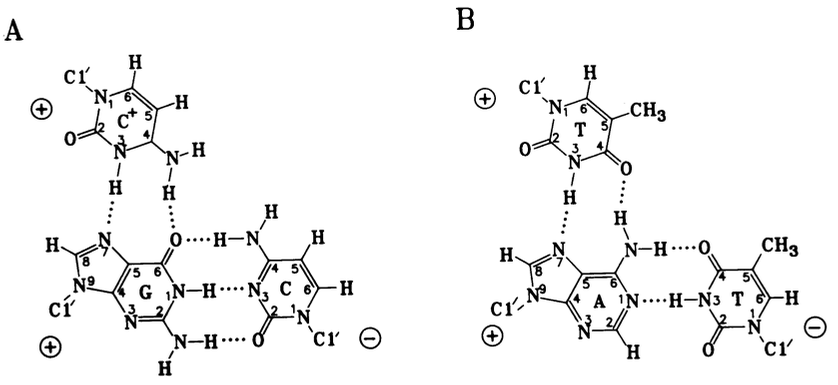

Figure 1. Schematics for hydrogen bonded base-pairing in (**A**) C^+.G:C triad and (**B**) T.A:T triad.

0.09 M Tris/boric acid buffer, pH 8.3, 0.02 M EDTA. Molar extinction coefficients were measured by phosphate analysis, as described previously (25). The sample was prepared by adding equimolar amounts of the pyrimidine strand (a 12mer; d-CTCCTTTCTTCC) and purine strand (a 14mer; d-TGAGGA-AAGAAGGT). The final composition is as follows: 5 mM duplex (~15 mg of material dissolved in 0.4 ml 99.9% ^2H_2O or in a mixed solvent consisting of 90% H_2O and 10% ^2H_2O); 0.05 M sodium phosphate buffer; pH 7.0/4.35; 0.08 M NaCl. Temperature of 10°C was used in most experiments, though some studies were carried out in the range 1–80°C.

NMR experiments were carried out on a Bruker AMX 500 spectrometer with a ^1H frequency of 500 MHz. The NMR spectra in 90% H_2O + 10% ^2H_2O included one-dimensional (1D) ^1H spectra recorded with P1$\bar{1}$ pulse sequence (26) at different pH and temperatures and two dimensional (2D) nuclear Overhausser enhancement spectroscopy (NOESY) (27) spectrum with 1$\bar{1}$ detection pulse and a mixing time of 200 ms. The 2D experiment in ^2H_2O included clean total correlation spectroscopy (clean TOCSY) (28) with a mixing time of 80 ms.

RESULTS

Evidence for the formation of new structure(s)

Figure 2 shows the pH dependence of imino and part of amino proton region in the 1D ^1H NMR spectrum of the sample in a mixed solvent of 90% H_20 and 10% ^2H_2O at 10°C.

It is known that hydrogen bonded imino protons of G, T and C^+ nucleotides give rise to resonances in the range 11–16 p.p.m. Non-hydrogen bonded imino protons usually give broad and weak signals in the range 8–11 p.p.m. The amino protons of C resonate in the range 6–8.5 p.p.m. The amino protons of C^+ and A resonate in the range 8–10 p.p.m. In the present system, there are six guanines, six adenosines, six cytosines and eight thymines (six belonging to the pyrimidine strand and two to the purine strand). Assuming standard Watson–Crick hydrogen bonded base pairing, one expects a maximum of 12 imino proton resonances in the range 12–15 p.p.m. Thymine imino protons generally resonate down-field (13.0–15.0 p.p.m.) compared to guanosine imino protons (12.0–13.5 p.p.m.); this is indeed observed in the spectrum at pH 7.0 (Fig. 2). The overhangs at the ends of the 14mer, consisting of the two Ts may either exchange fast and may not be observed or result in two extra imino proton resonances in

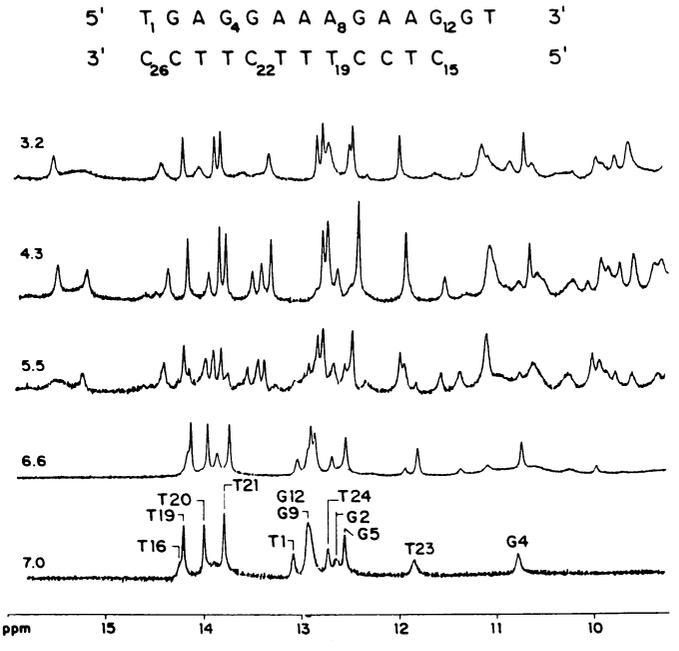

Figure 2. The pH dependence (ranging from 7.0–3.3) of imino/amino proton region at 10°C in 1D ^1H NMR spectrum of a 1:1 molar mixture of d-TGAGGAAAGAAGGT and d-CTCCTTTCTTCC dissolved in a mixed solvent of 90% H_2O and 10% ^2H_2O.

the region 10–12 p.p.m. The six cytosines are expected to show 12 amino proton resonances (six from those involved in hydrogen bonding and six from those exposed to water) in the region 6–8.5 p.p.m.

As the pH is lowered, cytosines are protonated and additional imino proton resonances are observed in the region 14–16 p.p.m. Further, the amino protons get shifted and appear in the region 8.5–10 p.p.m. As shown in Figure 2, the complexity of the imino/amino region increases manyfold with a multitude of lines between 8.5 and 16.0 p.p.m. Such a dramatic change in the 1D spectrum at pH below 5.5 and at moderate ionic strength, is a strong evidence for the existence of a new structure(s). Classification of the 1D spectra is based on the observation of resonances in the regions 8.5–10.0 and 14.0–16.0 p.p.m. which are expected to arise from the amino and imino protons of protonated cytosine,

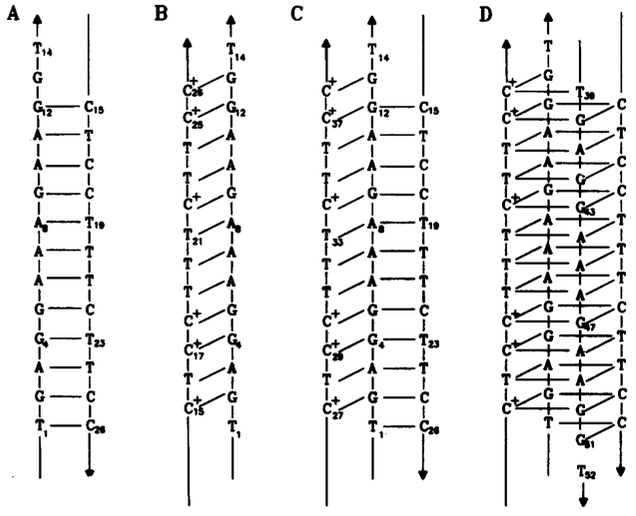

Figure 3. Possible solution conformations for the new homopurine–homopyrimidine complex. (**A**) Antiparallel Watson–Crick duplex; (**B**) parallel Hoogsteen duplex with six T.A and C^+.G Hoogsteen base pairs; (**C**) $(Y^+)_n.(R)_n:(Y)_n$ triplex; and finally, (**D**) tetraplex.

respectively. It is clear that there are more resonances than expected from a double-stranded DNA.

Characterization of molecular structure at low pH

While considering possible structures for the low pH complex we have considered schemes which have minimum number of mismatched base pairs. Possible multiplexes up to tetraplex have been considered. The schemes are: (i) dissociation of the duplex into monomeric strands. (ii) Duplexes: (a) pH driven conformational changes in the anti-parallel Watson–Crick duplex (Fig. 3A) leading to changes in base pairing and/or alteration in backbone structure; (b) transformation of the antiparallel duplex into a parallel Hoogsteen duplex with six T.A and six C^+.G Hoogsteen base-pairs (Fig. 3B); (c) coexistence of antiparallel and parallel duplexes. (iii) Triplex: coexistence of a triplex (Fig. 3C) with a monomeric purine strand. (iv) Tetraplex: a fourth $(R)_n$ strand can be accommodated in the triplex to form a tetraplex, as proposed by Johnson and Morgan (29). The tetraplex may consist of four [G(C^+.G:C)] and five [A(T.A:T)] tetrads, two mismatched [G(T.A–C)] and [A(C^+.G–T)] tetrads, and two mismatched non-canonical triads, [T(C^+.G)] and [(T–C)G] on the ends (Fig. 3D).

The imino and amino proton resonances provide vital information on the secondary structure of DNA. Figures 4 and 5 show selected regions of the NOESY spectrum of the sample in 90% H_2O + 10% ^2H_2O at 10°C and pH 4.35. As a first step in the assignment of exchangeable protons, we identify two sets of amino protons, one belonging to protonated (8–10 p.p.m.) and the other to non-protonated cytosines (6–8.5 p.p.m.). This is based on the classification of chemical shifts and observation of intranucleotide nOes such as intra amino proton nOe and those from C(H6) to C(H5), C(4NH_2) to C(H5) and C(4NH_2) to C(H6). Further, the two sets of amino protons of C and C^+ are found to show distinct cross peaks (Fig. 4A) between each other. Such peaks in a NOESY spectrum can be either due to through space dipolar interaction or due to a given cytosine moiety undergoing chemical exchange between the C and C^+ forms. The second situation can be ruled out in the present study because the observed inter-amino proton cross peaks are between C and C^+, which are at different positions along the DNA sequence. Such sequentially assigned inter-amino proton interaction can be explained only on the basis of formation of C^+.G:C triads. This is the first structural evidence for the transformation of the sample into triplex.

Generally, in a multistranded DNA every base pair has one imino proton, which is involved in the inter-base hydrogen bonding. Since the rise per residue (h) along the helix axis ranges from 0.26 nm (A-DNA) to 0.34 nm (B-DNA), one should observe nOes between the imino protons of the sequential base pairs along the helix axis except in the case of DNA intercalation (i.e. i-motif DNA) (30). This implies that one should be able to walk along the helix axis from one end to the other end by monitoring inter-imino proton nOes. Thus, with an unambiguous knowledge of the chemical shift of any one imino proton, the resonance assignment of the remaining imino protons can be completed. In the present study, intercalation of any kind is ruled out because we do not observe H1′–H1′ cross peaks (30) in the NOESY spectrum.

We have identified the imino protons of protonated cytosines and thymines from the direct observation of intranucleotide nOes, namely $C^+(NH_2)\rightarrow C^+(3NH)$ and $T(CH_3)\rightarrow T(3NH)$. Having identified the thymine imino proton resonances, we found three of them in the region 13.6–14.4, showing nOes among themselves (Fig. 4B). Such nOes can be expected if the three T units are adjacent to each other. Thus, the three 3NH protons are assigned to the trinucleotide segment -TTT- at the center of 12mer strand. The proton at 14.25 p.p.m. shows nOes to two other 3NH protons in the triplet, and can be assigned as belonging to the unit in the middle of the -TTT- segment. Such an assignment is supported by the observation of inter-imino proton nOes from 3NH at 14.05 to the one at 12.65 p.p.m., which in turn shows to yet another imino proton resonance at 11.10 p.p.m. The former nOe can arise from a through space interaction between the imino protons belonging to sequential T and C^+ or diagonally across the strand T and G. From the fact that C is protonated at pH 4.35 and its 3NH resonates in the region 14–16 p.p.m., first possibility is ruled out. Thus the resonance at 12.65 p.p.m. is assigned to guanosine diagonally across the strands in the neighbourhood of the outer thymines belonging to the -TTT- segment. This indicates that the -TTT- segment belongs to the strand in which the cytosines are non-protonated; according to Figure 3C, this is -T19-T20-T21-. The resonance at 12.65 p.p.m. is thus assigned to G5 (1NH) of G5:C22 base pair. Assignment of the imino protons of the TTT segment serves as a starting point for the sequential resonance assignment of the remaining imino protons. Sequential connectivities have been observed between the Watson–Crick base paired imino protons (including those belonging to mismatched base pairs) all the way from the G5:C22 to T1:C26 base pairs (Fig. 4B). The assignment pathway is: G5(1N<u>H</u>) (belonging to G5:C22) → G4(1N<u>H</u>) and T23(3N<u>H</u>) (G4:T23) → T24(3N<u>H</u>) (A3:T24) → G2(1N<u>H</u>) (G2:C25) → T1(3N<u>H</u>) (T1:C26). On the other side of the segement, starting at T19(3NH), we could walk from A8:T19 to G12:C15. The assignment pathway is: T19(3N<u>H</u>) (A8:T19) → G9 (1N<u>H</u>) (G9:C18) → A10(6N<u>H</u>$_2$) (A10–C17) → T16(3N<u>H</u>) (A11:T16) → G12(1N<u>H</u>) (G12:C15). It is worth mentioning that A10–C17 mismatched base pair does not involve imino proton in its base-pairing scheme; the nOes between the A10(6NH_2) and G9(1NH) and T16(3NH) (Fig. 5) help in its sequential assignment.

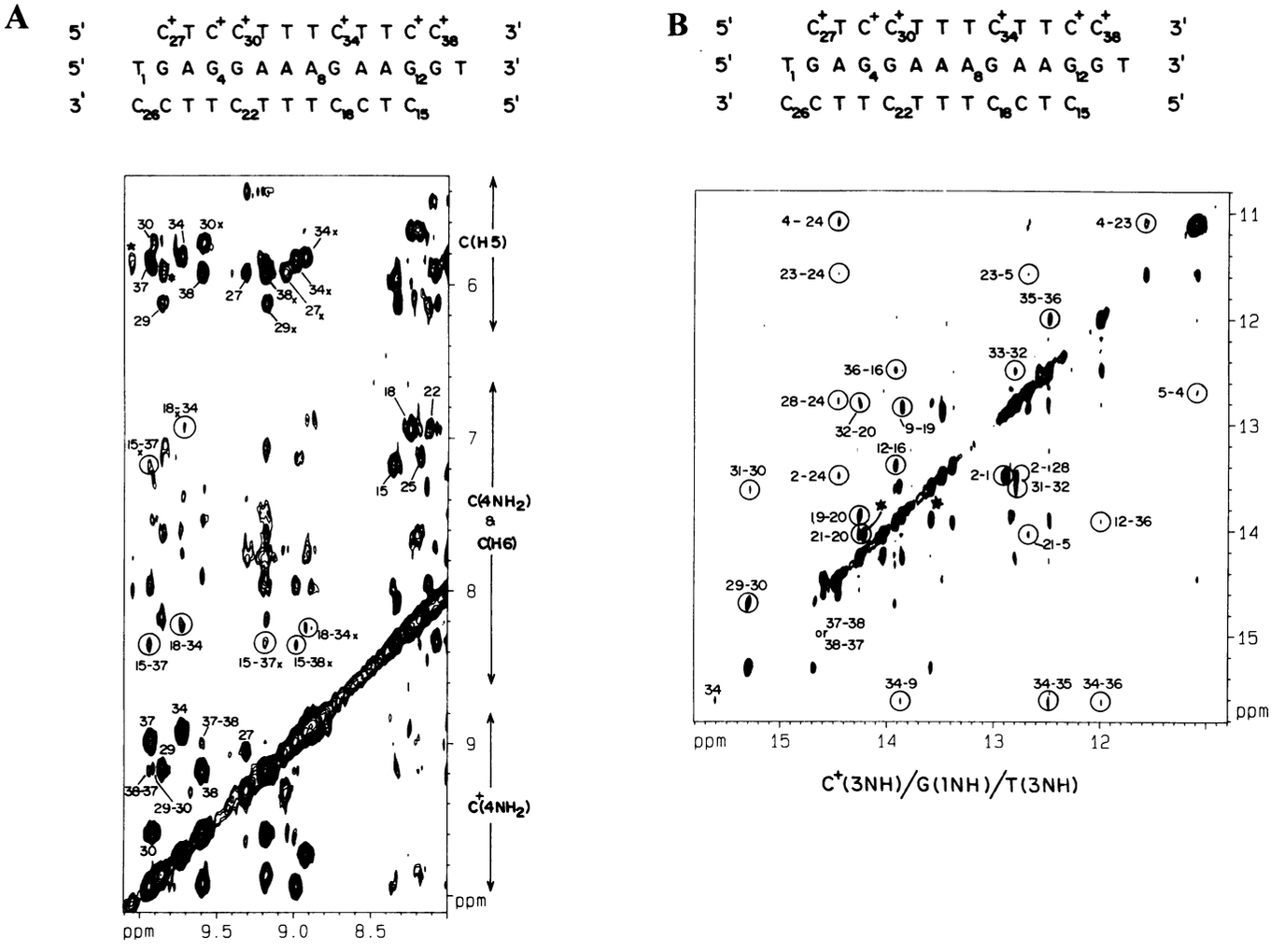

Figure 4. Selected regions of pure-absorption NOESY spectrum of 1:1 molar mixture of d-TGAGGAAAGAAGGT and d-CTCCTTTCTTCC recorded in a mixed solvent of 90% H_2O + 10% ^2H_2O at 10°C and pH 4.35. Experimental parameters were as follows: t_{1max} = 22.5 ms, t_{2max} = 205.0 ms, recycle delay = 1 s, 96 scans/t_1 increment, time-domain data points were 450 and 2048 along t_1 and t_2 dimensions, respectively. The ^1H carrier frequency was on the water resonance. The data were multiplied with sine bell window functions shifted by $\pi/4$ and $\pi/8$ along t_1 and t_2 axes, respectively, and zero-filled to 1024 data points along t_1 dimension prior to stripped 2D-FT. The digital resolution along ω_2 and ω_1 corresponds to 3.9 and 1.85 Hz/pt, respectively. (**A**) The nOe connectivities from H5/H6/C(4NH_2)/C^+(4NH_2) protons to the C^+(4NH_2) protons. The suffix 'x' to the nucleotide number identifies the exposed amino proton of the cytosine amino group. (**B**) Inter-imino proton nOe connectivities. The intranucleotide cross peaks are shown with the corresponding nucleotide number along the sequence. The internucleotide cross peaks are shown with the corresponding numbers of the nucleotide units along the sequence to which the protons belong [for example: 19–20 identifies the nOe between the T19(3NH) (ω_1 axis) and the T20(3NH) (ω_2 axis)]. Unassigned peaks are indicated with asterisks.

Starting again at the imino proton chemical shift of T20, we could assign the chemical shift of T32(3NH) from the observation of interstrand (T32-A7-T20) inter-imino proton nOe [T20(3NH) → T32(3NH)]. From T32(3NH) we could walk along the Hoogsteen base pairs up to the C^+29(3NH) following the sequential connectivity pathway: T32(3N<u>H</u>) (T32.A7) → T31(3N<u>H</u>) (T31.A6) → C^+30(3N<u>H</u>) (C^+30.G5) → C^+29(3N<u>H</u>) (C^+29.G4). We could not proceed further as C^+29(3NH) is relatively broad (may be an influence of G4–T23 mismatch base pair, which forms part of the C^+29.G4–T23 triad) and hence could not be sequentially connected to T28(3NH). However, T28(3NH) could be assigned from the previous knowledge of T24(3NH) from the intratriad inter-imino proton nOe between T24(3NH) and T28(3NH). Though the amino protons of C^+27 could be assigned by elimination process, assignment of its imino proton was not feasible. This protonated cytosine is at one end of the proposed triplex and hence its imino proton may be exposed to solvent water and exchange faster than its own amino protons.

Starting at the position of T32(3NH), we could walk in the other direction of the Hoogsteen strand up to T36(3NH) following the pathway: T32(3N<u>H</u>) (T32.A7) → T33(3N<u>H</u>) (T33.A8) → C^+34(3N<u>H</u>) (C^+34.G9) → T35(3N<u>H</u>) (T35.A10) → T36(3N<u>H</u>) (T36.A11). No sequential connectivities could be seen between T36 and C^+37. At this stage only two imino resonances are left unassigned (at 14.45 and 14.60 p.p.m.) which show nOes between themselves. These can be attributed to C^+37(3NH) and C^+38(3NH). The corresponding chemical shifts also satisfy our interpretation, thus helping us to complete the resonance assignment of the expected imino proton resonances [except C^+27(3NH)].

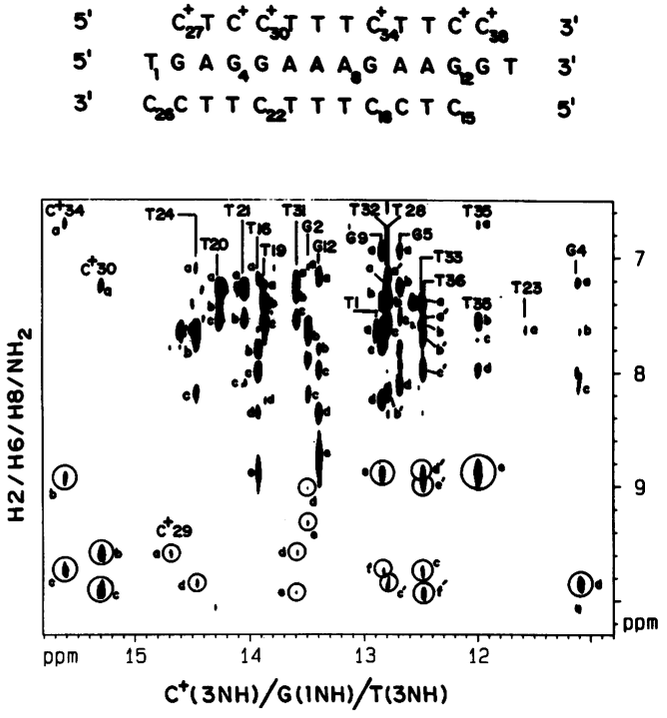

Figure 5. Selected regions of pure-absorption NOESY spectrum of 1:1 molar mixture of d-TGAGGAAAGAAGGT and d-CTCCTTTCTTCC recorded in a mixed solvent of 90% H_2O + 10% ^2H_2O at 10°C and pH 4.35. NMR parameters are the same as that shown in Figure 4. This figure shows the nOe connectivities from H2/H6/H8/NH_2 protons to the imino protons, [C^+(3NH), T(3NH) and G(1NH)]. The nOe cross peaks (a–f) in each column (identified by the imino proton) are assigned as follows with hydrogen bonded and exposed amino protons abbreviated as HB and X, respectively: C^+34(3NH) → G9(H8) (a), C^+34(4NH_2) (X, b), C^+34(4NH_2) (HB, c); C^+30(3NH) → G5(H8) (a), C^+30(4NH_2) (X, b), C^+30(4NH_2) (HB, c); C^+29(3NH) → C^+30(4NH_2) (X, a); T24(3NH) → C25(4NH_2) (X, a), A3(H2) (b), C25(4NH_2), (HB, c), C^+29(4NH_2) (HB, d); T20(3NH) → A6(H2) (a), A8(H2) (b), A7(H2) (c); T21(3NH), → A6(H2) (a), A7(H2) (b), C22(4NH_2) (E, c); T16(3NH) → C15(4NH_2) (X, a), A10(H2) (b), A11(H2) (c), C15(4NH_2) (HB, d), A10(6NH_2) (HB & X, e); T19(3NH) → C18(4NH_2) (X, a), A7(H2) (b), A8(H2) (c), C18(4NH_2) (HB, d); T31(3NH) → A6(H8) (a), G5(H8) (b), A7(H8) (c); C^+30(4NH_2) (X, d), C^+30(4NH_2) (HB, e); G2(1NH) → C25(4NH_2) (X, a), A3(H2) (b), C25(4NH_2) (HB, c), C^+27(4NH_2) (X, d), C^+27(4NH_2) (HB, e); G12(1NH) → C15(4NH_2) (X, a), A10(H2) (b), A11(H2) (c), Cl5(4NH_2) (HB, d), A11(6NH_2) (HB & X, e); T1(3NH) → C26(4NH_2) (X/HB, a); G9(1NH) → C18(4NH_2) (X, a), A8(H2) (b), A10(H2) (c), C18(4NH_2) (HB, d), A10(6NH_2) (X & HB, e), C^+34(4NH_2) (X, f); T32(3NH) → A6(H8) (a), A8(H8) (b), A7(H8) (c); T28(3NH) → C25(4NH_2) (X, a'), C25 (4NH_2) (HB, b'), C^+29(4NH_2) (HB, c'); G5(1NH) → C22(4NH_2) (X, a), A6(H2) (b), A7(H2) (c), C22(4NH_2) (HB, d); T33(3NH) → A8(H8) (a), A7(H8) (b), C^+34(4NH_2) (HB, c); T36(3NH) → A10(H8) (a'), G12(H8) (b'), A11(H8) (c'), A10(6NH_2) (HB & X, d'), C^+37(4NH_2) (X, e'), C^+37(4NH_2) (HB,f'); T35(3NH) → G9(H8) (a), A10(H8) (b), G12(H8) (c), A11(H8) (d), A10(6NH_2) (HB & X, e); T23(3NH) → A3(H2) (a); G4(1NH) → A6(H2) (a), A3(H2) (b), C24(4NH_2) (HB, c), C^+29(4NH_2) (HB, d). In the overlapped regions the two columns of peaks are distinguished by single letter a–f, with and without primes.

The assignment of the imino protons are substantiated by other nOe connectivities to these protons from amino protons and base protons (H2/H6/H8) of individual nucleotides. Such assignments are explained in Figure 5 and its caption. For example, in the T.A:T triads, a strong nOe is expected between T(3NH) and A(H2) for each A:T base pair, while for each Hoogsteen T.A base pair a strong nOe is expected from T(3NH) to A(H8) (Fig. 1). Similarly, in the case of C^+.G:C triad, for G:C base pair one strong and one medium size cross peaks are expected [G(1NH) → C(NH_2)], while for Hoogsteen C^+.G base pair a strong nOe is expected from G(H8) to C^+(3NH). One also expects nOes diagonally across the strands between any two neighbouring base pairs. Thus, we could provide further support to the assignment of the imino proton resonances and complete the assignment of all the exchangeable protons and most of the base protons (H2/H6/H8).

Intricate networks of interstrand and intrastrand NOESY cross peaks establishes the hydrogen bonded base paired network and thus the secondary structure of complex. While some of the peaks are expected from all the multiplexed models discussed earlier, the duplex models, including the Hoogsteen duplex are not consistent with the nOe data. The observation of thirteen hydrogen bonded thymine 3NH protons: six belonging to the pyrimidine strand with protonated cytosines, six more to a second pyrimidine strand with non-protonated cytosines and the last one belonging to the purine strand, may indicate coexistence of two duplex models. However, interstrand T(3NH) → T(3NH) nOes rule out such a possibility. This is further substantiated by the observation of nOes between the amino protons of the cytosines belonging to the protonated and non-protonated pyrimidine strands. Tetraplex formation can be ruled out on several grounds. The extra hydrogen bonds, [G(60)–C^+(4NH_2) and G(60)–C(4NH_2)] (29) formed by tetraplex formation would result in downfield shift of the amino protons of C and C^+, involved in such hydrogen bonds. The fourth homo-purine strand will give rise to six additional G(1NH) resonances. Further, one should also observe strong nOe's from the H5 base protons of the C^+ units to the imino protons of the G belonging to the fourth purine strand in the tetrad. Neither of the nOes nor the extra imino proton resonances have been observed in the NMR spectra.

How a complex with three strands is formed when the homopurine and homopyrimidine DNA strands are taken in 1:1 molar ratio? The answer is that the triplex formation is via the disproportionation reaction

$$2[(R)_n]:[(Y)_n] \xrightarrow{H^+} [(Y^+)_n].[(R)_n]:[(Y)_n] + [(R)_n]$$

driven by the pH. The existence of a free purine strand is confirmed by the observation of two extra CH_3–H6 cross peaks from the thymines flanked on the two ends of the purine strand, in the TOCSY spectrum (Fig. 7).

Dissociation and stability of the triplex

The stability and dissociation of the triplex has been monitored by temperature dependent 1D ^1H NMR spectra (Fig. 6). As the temperature increases, the rapidly exchanging C^+ imino protons broaden first. This is followed by the imino protons belonging to the nucleotides at the two ends of the triplex. Interestingly, even at 50°C there are several imino and amino proton resonances. Beyond 50°C, there is a sudden disappearance of the remaining exchangeable protons including those coming from amino protons belonging to C^+ units. Similar co-operative triplex-single strand transition has been observed in other triplexes (20,23).

Conclusions

NMR results provide possibly the first structural evidence for the pH driven disproportionation of the oligonucleotide sample, consisting of homo-purine and homo-pyrimidine strands in 1:1 molar ratio, into an intermolecular pyrimidine.purine:pyrimidine DNA triplex and a

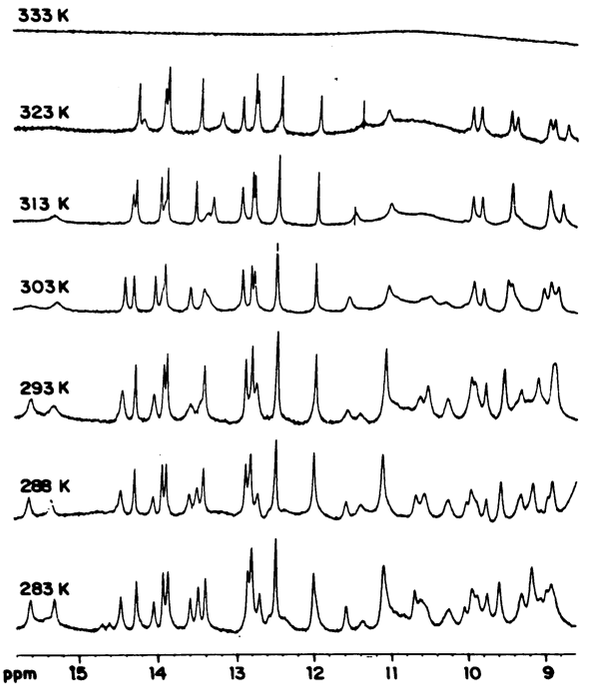

Figure 6. Temperature dependence of imino and amino proton region of the 1D ^1H NMR spectrum of 1:1 molar mixture of d-TGAGGAAAGAAGGT and d-CTCCTTTCTTCC recorded in a mixed solvent of 90% H_2O and 10% ^2H_2O at pH 4.35.

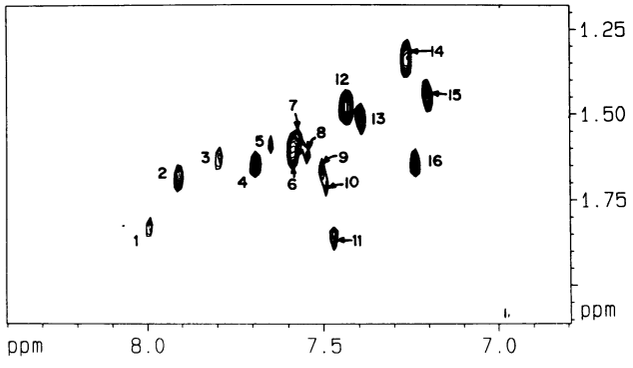

Figure 7. Selected region of clean TOCSY spectrum of 1:1 molar solution of d-TGAGGAAAGAAGGT and d-CTCCTTTCTTCC recorded in 99.9% ^2H_2O at 10°C and pH 4.35, with a mixing time of 80 ms. A total of 16 CH_3-H_6 cross peaks are visible. Fourteen belong to thymines of the triplex and the other two to thymines flanking the ends of the free purine strand.

single stranded oligopurine. The triplex is stabilized by five T.A:T, four C^+.G:C and two mismatched triads, namely, C^+.G–T and T.A–C. This triplex with 11 triads is further stabilized by a Hoogsteen C^+.G base pair on one end. The triplex is stable even at 50°C, inspite of the presence of mismatched base pairs. Temperature dependence study of the imino proton resonances reveals that the triplex dissociates directly into single strands without any duplex intermediates. This study provides some insight into the interactions stabilizing the base triplets, which include mismatched base pairs (C^+.G–T and T.A–C). This mode of recognition offers new possibilities for the targeting of mismatched sequences by oligonucleotide-directed triple-helix formation.

A surprising result is the absence of parallel stranded duplexes at either neutral or acidic pH. The sequences were designed such that there are no mismatches in case parallel duplexes are formed. Instead, the system prefers to accommodate mismatches either in antiparallel duplex (neutral pH) or triplex with a free purine strand (acidic pH). This shows the comparative lower stability of parallel duplexes, proposed by Liu *et al.* (31) and Raghunathan *et al.* (32).

ACKNOWLEDGEMENTS

The facilities provided by the High Field FT NMR National Facility supported by the Department of Science and Technology, India, Department of Biotechnology, India, Council of Scientific and Industrial Research, India and Tata Institute of Fundamental Research, Bombay, India are gratefully acknowledged.

REFERENCES

1 Felsenfeld. G., Davies, D. R., and Rich, A. (1957) *J. Am. Chem. Soc.* **79**, 2023–2024.
2 Hoogsteen, K., (1959) *Acta Cryst.* **12**, 822–823.
3 Lipsett, M. N. (1963) *Biochem. Biophys. Res. Commun.* **11**, 224–228.
4 Howard, F. B., Frazier, J., Lipsett, M. N. and Miles, H. T. (1964) *Biochem. Biophys. Res. Commun.* **17**, 93–102.
5 Felsenfeld. G. and Miles, H. T. (1967) *Annu. Rev. Biochem.* **36**, 407–448.
6 Morgan, A. R. and Wells R. D. (1968) *J. Mol. Biol.* **37**, 63–80.
7 Lee, J. S., Johnson, D. A and Morgan, A. R. (1979) *Nucleic Acids Res.* **6**, 3073–3091.
8 Arnott, S. and Selsingh, E. (1974) *J. Mol. Biol.* **88**, 509–521.
9 Moser, H. E., Dervan, P. B. (1987) *Science* **238**, 645–650.
10 Mirkin, S. M., Lyamichev, V. I., Drushlyak, K. N., Dobrynin, V. N., Filippov, S. A. and Frank-Kamenetskii, M. D. (1987) *Nature* **330**, 495–497.
11 Voloshin, O. N., Mirkin, S. M., Lyamichev, V. I., Belotserkovskii, B. P., and Frank-Kamenetskii, M. D. (1988) *Nature* **333**, 475–476.
12 Hanvey, J. C., Shimizu, M., and Wells, R. D. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 6292–6296.
13 Wells, R. D., Collier, D. A., Hanvey, J. C., Shimizu, M. and Wohlrab, F. (1988) *Fed. Amer. Soc. Exp. Biol. J.* **2**, 2939–2949.
14 Htun, H. and Dahlberg, J. E. (1988) *Science* **241**, 1791–1796.
15 Mirkin, S. M. and Frank-Kamenetskii, M. D. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 541–576.
16 Rajagopal, P. and Feigon, J. (1989) *Nature* (London) **339**, 637–640.
17 de los Santos, C., Rosen, M. and Patel, D. (1989) *Biochemistry* **28**, 7282–7289.
18 Umemoto, K., Sarma, M. H., Gupta, G., Luo, J. and Sarma, R. H. (1990) *J. Am. Chem. Soc.* **112**, 4539–4545.
19 Pilch, D. S., Levenson, C. and Shafer, R. H. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 1942–1946.
20 Kan, L. S., Callahan, D. E., Trapane, T. L. and Miller, P. S. (1991) *J. Biomol. Struct. Dynam.* **8**, 911–933.
21 Sklenar, V. and Feigon, J. (1990) *Nature* **345**, 836–838.
22 Radhakrishnan, I., Santos, C. D. L. and Patel, D. J. (1991) *J. Mol. Biol.* **221**, 1403–1418.
23 Macaya, R., Wang, E., Schultze, P., Sklenar, V. and Feigon, J. (1992) *J. Mol. Biol.* **225**, 755–773.
24 Clauwaert, J. and Stockx, J. (1968) *Natureforsch. B.* **23**, 25–30.
25 Muraoka, M., Miles, H. T., and Howard, F. B. (1980) *Biochemsitry* **19**, 2429–2439.
26 Hore, P. J. (1983) *J. Magn. Res.* **55**, 283–300.
27 Anil Kumar, Wagner, G., Ernst, R. R. and Wuthrich, K. (1980) *Biochem. Biophys. Res. Commun.* **96**, 1156–1163.
28 Griesinger, C., Otting, G., Wuthrich, K. and Ernst, R. R. (1988) *J. Am. Chem. Soc.* **110**, 7870–7872.
29 Johnson, D. A and Morgan, A. R. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 1637–1641.
30 Gehring, K., Leroy, J-L. and Gueron, M. (1993) *Nature* **363**, 561–565.
31 Liu, K., Miles, H. T., Frazier, J and Sasisekharan, V. (1993) *Biochemistry* **32**, 11802–11809.
32 Raghunathan, G., Miles, H. T and Sasisekharan, V. (1994) *Biopolymers* **34**, 1573–1581.