Polypurine/polypyrimidine hairpins form a triple helix structure at low pH

Margret M.W.Mooren, David E.Pulleyblank¹, Sybren S.Wijmenga, Marcel J.J.Blommers and Cornelis W.Hilbers*
Laboratory of Biophysical Chemistry, University of Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands and ¹Department of Biochemistry, University of Toronto, Toronto, Ontario M5S-1A8, Canada

Received August 29, 1990; Revised and Accepted October 22, 1990

ABSTRACT

1D and 2D NMR investigations of the 15 residue deoxynucleotide sequence d(TCTCTC-TTT-GAGAGA) show that above pH = 6.5 the molecule adopts a B-form hairpin conformation. As the pH is lowered below 6.5 molecules progressively associate in pairs to form a partially triple helical, partially single stranded structure in which the bases of the oligopyrimidine d(TC)₃ tract from one molecule form Hoogsteen pairs with the d(G-A)₃ tract of the other. Imino protons of protonated cytosines can be observed at very low field (= 15 ppm). The enthalpy of triplex formation was estimated by NMR techniques to be ~16 kcal mol⁻¹. Intense H6 to H3′ cross peaks from residues in all three strands suggest the presence of N-type sugars at some but not at all possible sites. Surprisingly strong cross peaks between H5′ or H5″ and non-exchangeable base protons are also observed. These suggest that certain of the 5′-C5′-C4′-C3′ phosphate backbone torsion angles (γ) are unusual.

INTRODUCTION

Polypurine/polypyrimidine tracts are abundant components of most eukaryotic genomes where they frequently bracket structural genes. In some cases these sequences may have regulatory functions [1]. An unexpected feature of these tracts is their unusual sensitivity to S1 (and related) nucleases. This sensitivity suggested that they may exist in a partially single stranded state. Other work has shown the nuclease sensitivity in these sequences to be a consequence of structural transitions from the B-form to protonated forms under the low pH conditions typically used for S1-nuclease reaction [2, 3]. In the special case of tracts of poly dGpdc₃ a non-protonated S1 sensitive form has also been found [4]. Anomalies in the electrophoretic mobilities of supercoiled plasmids show that these transitions result in unwinding of these sequences [2, 3, 5, 6, 7, 8, 9]. In supercoiled plasmids interplay between unwinding torsion and pH allow protonated nuclease sensitive states to persist under neutral pH conditions, suggesting that the protonated structures may have functional significance in living cells.

Two general classes of models have been proposed for the protonated nuclease sensitive state of the polypurine/polypyrimidine tracts. Pulleyblank et al. [3, 10] attributed the S1 nuclease sensitivity to alterations in the sugar-phosphate backbone torsion angles necessary to accomodate protonated G·CH⁺ Hoogsteen pairs in a double helix which also contained A·T Watson-Crick pairs. In a second model (H-DNA) [9, 11, 12] it was proposed that a triplex forms by folding the tract at its center so that the polypuridine strand from one half forms Hoogsteen pairs with the purine strand of the opposite half of the tract. In this model the half of the d(GA)ₙ strand not involved in the triplex is unpaired and is therefore sensitive to S1-nuclease. Because polypurine/polypyrimidine polymers had previously been shown to form triplex readily [13, 14, 15] the ‘H’-model did not require the introduction of a new form of DNA helical structure. Although several recent reports have provided experimental support for the H-model in individual cases of protonated polypurine tracts [16, 17] patterns of nuclease and chemical reactivity observed in some tracts have been difficult to reconcile with this model [10, 18, 19]. These results suggest that S1-nuclease sensitivity of nucleic acid structures depends on complex criteria and that the enzyme may cleave some nucleic acids which are not single stranded but which have unusual torsion angles in the backbone of the helix. In some cases more than one protonated form of a single plasmid insert can be observed by 2D gel electrophoresis [9, 10]. These observations suggested that the effective cross-links between strands at the ends of an embedded polypurine/polypyrimidine tract in plasmids might sufficiently stabilize a protonated duplex structure that it could compete with strand disproportionation to a triple helical structure.

While chemical and enzymatic probes do not provide sufficient information to unambiguously determine detailed molecular structure other techniques such as X-ray diffraction and nuclear magnetic resonance (NMR) which are able to provide more detailed structural information, are necessarily restricted to

* To whom correspondence should be addressed
relatively small model systems. In this paper we have investigated the behaviour of homopurine/homopyrimidine hairpin oligonucleotides as a function of pH. Hairpin loop structures have been object of numerous investigations. The cross-link between the complementary strands provided by the loop stabilizes the paired duplex by as much as 55°C, thereby allowing relatively small molecules to be studied in the duplex state. The cross-link is also analogous to the effective cross-link at the ends of polypurine/polypyrimidine sequences embedded in plasmids.

**MATERIALS AND METHODS**

Homopurine/homopyrimidine synthetic oligonucleotides were purchased in deblocked form from ’Synthetic Genetics’, La Jolla, California. Purity was checked by electrophoresis on 20% polyacrylamide gels.

For spectrophotometric analysis DNA-samples were dissolved in 1 ml of a mixture of 0.05M sodium acetate with 0.05M sodium cacodylate (extinction of 0.8 at 260 nm).

In the spectrophotometric experiments temperature control was achieved by circulating water from a cryothermostat (Mettler, WKS), first through a thermostated cell holder and then through the jacket of the cell.

In most NMR experiments DNA concentrations were about 2 mM in 0.1 M sodium deuteroacetate in water or deuterium oxide adjusted to the pH/pD indicated by addition of deuterioacetic acid. In the dilution experiments, the initial oligonucleotide concentrations were about 12 mM. For these experiments samples were prepared by dialysing DNA stock solutions against 0.1 M sodium deuteroacetate buffer pH=5.5 (meter-reading) in water for 24 hours, using a Spectra Por3 (3000 MW cutoff) dialysis membrane.

1D-proton spectra were recorded at 5.5°C, 14.0°C, 21.2°C and 29.2°C. Sample concentrations were adjusted by addition of dialysis buffer.

The 400- and 600 MHz 1H NMR 1D and 2D spectra were recorded on Bruker AM400 and AM600 spectrometers respectively, which were interfaced to an Aspect 3000 computer.

Spectra were accumulated with a time shared long pulse [20] (8K data points); prior to Fourier transformation the spectra were zero-filled to 16K data points and apodized with Lorentz-Gauss window functions.

Phase-sensitive 400 MHz NOESY spectra of the water sample were recorded at 297K with a time-shared long pulse as observation pulse [20]. The spectral width was 16000 Hz; the carrier was placed at the low-field end of the spectrum. The spectrum was acquired with 4K points in the t2 direction and 512 points, zerofilled to 1K points, in the t1 direction, with a mixing time of 0.3 s. Before Fourier-transformation each free induction decay was subjected to data shift accumulation; apodization in the t2-direction took place by applying a Lorentz-Gauss filter function and in the t1-direction by applying a squared sinebell function phase shifted by 60°.

The 600 MHz MINSY spectrum [21] of the D2O sample was recorded at 297K with a mixing time of 0.3 s. The spectral width was 6000 Hz. The spectrum was acquired with 2K data points in the t2-direction and 370 data points, zerofilled to 1K points, in the t1-direction. Before Fourier transformation a squared sinebell function phase shifted by 90° was applied in the t2-direction and phase shifted by 60° in the t1-direction. The region of the spectrum containing the H2O, H2O resonances was completely saturated during the mixing time and residual HDO was suppressed by replacing the last 90° pulse with a 11 spin-echo pulse sequence [22].

**RESULTS AND DISCUSSION**

**Formation of a triplex hairpin structure**

A family of eleven partially complementary homopurine/homopyrimidine synthetic oligonucleotides was screened for their ability to form hairpin structures at neutral pH and protonated structures at low pH. Two of these oligonucleotides were chosen for detailed study because of the relative simplicity of the imino proton spectra of their low pH forms. The results obtained for one of these molecules, d5(TCTCTC-TGGAGAGA)3', designated 'T3', are discussed in detail. UV melting studies gave a concentration independent Tm for this oligonucleotide of 63°C at pH=8.0. The concentration independence of the melting temperature was retained at NMR concentrations above 2mM. These results establish that the NMR spectra obtained under neutral to slightly alkaline pH conditions represent the hairpin form of the molecule [23, 24].

1D imino proton spectra recorded for T3 at pH=4 and at pH=8 both at a temperature of 294K are shown in Figure 2. Assignment of the spectrum recorded at pH=8 was achieved via standard procedures i.e. by utilizing NOESY spectra recorded for H2O solutions. At pH=8 the imino protons of the A·T Watson-Crick base pairs resonate near 14 ppm, while those of G·C base pairs resonate near 13 ppm, as indicated in Figure 2.

The low pH spectrum is considerably more complex than that obtained at pH=8. Several new imino proton resonances appear, while those of the structure dominant at pH=8 become less intense. This change in the spectrum reveals the appearance of a new protonated form of this oligonucleotide at low pH. In the range of 15−17 ppm one strong new resonance appears; in addition some very weak and broad resonances are seen (see also Figure 3). The very low field position (15−17 ppm) of these new resonances suggests that they arise from the imino protons of protonated cytosines. Amino proton resonances near 9 and 10 ppm are also consistent with the presence of protonated cytosines [25, 26, 27] (vide infra). New resonances between 12.5 and 13 ppm suggest that normal Watson-Crick G·C pairs are also present in the structure(s).

Possible explanations for the occurrence of both protonated and non-protonated cytosines in the low pH form of this molecule are illustrated in Figure 1. In the first (Fig. 1A) two hairpin molecules interact so that the pyrimidine strand of one binds in the major groove of the other to form a partially triple helical, partially single stranded structure. In the second model the G·C base pairs in the hairpin are protonated. In Fig. 1B two extreme situations are indicated, i.e. non-protonated and completely protonated hairpins, however intermediate states are conceivable as well and would be necessary to explain Fig. 2B. Since in the second the reaction is unimolecular with respect to oligonucleotide while in the first it is bimolecular these possibilities can be distinguished by a dilution experiment. Spectra were recorded at constant pH for a series of different temperatures and at different DNA concentrations. The results are collated in Figure 3. Comparison of the spectra (e.g. 3A, D and G) shows that the signals characteristic of the low pH form (c.f. Figure 2) diminish in intensity and are replaced by the signals characteristic of the high pH form upon dilution of the sample. The concentration dependence of the spectra under iso-ionic/thermal conditions shows that formation of the protonated structure requires the
participation of more than one molecule of oligonucleotide. The result therefore eliminate model 1B as the dominant low pH form of T3. Since, even at the highest dilution, there was no evidence for a monomeric protonated form it is unlikely that structures of the type shown in Figure 1B are stable at the given pH. Analysis of selected parts of the NOESY spectra of T3 presented in Figure 4A and B shows that the low pH form is formed through interaction of two molecules as illustrated in Figure 1A. Complete description of the assignments, together with more detailed structural analysis will be published elsewhere.

Sequential assignments of the NOESY spectra in H2O and D2O were made using cross peaks from both the imino region (12.5-16 ppm) and the aromatic/amino proton region (6.8-10 ppm). The six cytosine H5-H6 cross peaks shown in Figure 4B were particularly valuable in the assignment process. These cross peaks were identified as belonging to cytosines on the basis of their presence as strong cross peaks in a double quantum filtered COSY spectrum (not shown). We note in passing that assignments will only be discussed insofar as they are important to demonstrate the existence of the triple helical region. Cross peaks between the CH5 (and via spin diffusion CH6) to cytosine amino protons in the 7-8 ppm range or the 9-10 ppm range identified the CH5/CH6 pairs as belonging to either non-protonated or protonated cytosines respectively. Other cross peaks from the cytosine amino protons to exchangeable imino protons as well as to non-exchangeable aromatic protons were used in the sequential assignment.

The protonated cytosines of C·G·CH⁺ base triads shown in Figure 1C contribute imino proton resonances near 15-16 ppm. These are connected to individual cytosine H5-H6 resonances via spin diffusion through the amino proton resonances. The non-protonated cytosines give rise to amino proton resonances near 7 and 8.2 ppm. Guanines in Watson-Crick base pairs generate imino proton resonances near 13 ppm. Short interatomic distances between the guanine imino proton and the non-protonated cytosine amino protons result in the strong cross peaks shown in Figure 4A. The strong cross peak observed between the H8 proton of G12 at 7.3 ppm and the imino proton of protonated C19 at 15.3 ppm directly establishes the presence of a Hoogsteen dCH⁺·dG pair in the complex. Similarly, a strong NOE between the imino proton of T20 and the H8 proton of A13 establishes the presence of a Hoogsteen A·T pair. Relatively weak cross peaks between the amino protons of the non-protonated C4 and protonated C19 in Figure 4B establish the interaction of Watson-Crick C·G and Hoogsteen CH⁺·G pairs as predicted for C·G·CH⁺ base triads illustrated in Figure 1C.

Imino proton resonances from protonated cytosines other than that of C19, i.e. those of C17 and C21, are observed in the region 15-17 ppm (c.f. Figs. 2 and 3), but are very broad and of a very low amplitude. As a result no connectivities of these downfield shifted positions (marked 17 and 21) are expected. However, the downfield shifted positions of the amino protons and the connectivities, albeit weak, of the 17NH2 to the 6NH2 resonances and from the 21NH2 to the 2NH2 resonances do show
that base triples are formed for these residues. Further evidence for the protonated structure being triple helical comes from the imino protons of the thymines. Watson-Crick paired thymines of the T·A·T base triad illustrated in Figure 1C contribute imino proton resonances near 14.3 ppm, as shown below. NOE’s from the imino proton of thymine 18 to the amino protons of adenine 11 which also has NOE’s to the imino proton of the Watson-Crick and Hoogsteen paired thymines. Strong cross peaks (not shown) between the imino protons of Watson-Crick paired thymines and the imino protons of Watson-Crick paired guanines show them to be stacked in the same helix. Other strong NOE’s such as those found between the amino protons of non-protonated cytosines 2 and 4 and the imino protons of the adjacent Watson-Crick paired thymines 3 and 5 confirm this conclusion. Since the imino proton of Hoogsteen paired thymine 18 has strong NOE’s to the amino protons of protonated cytosines 17 and 19 and the imino proton of Hoogsteen base paired thymine 20 has strong NOE’s to the amino protons of protonated cytosines 19 and 21, Hoogsteen paired thymines are stacked above and below Hoogsteen paired cytosines.

**Thermodynamic and structural aspects**

The observation reported here that two hairpin molecules readily associate to form a bimolecular complex was surprising. For the bimolecular triplexes to form, one of the stable hairpin molecules involved has to unfold. Triple helical structures formed by combining a strand of poly(A) or poly(d(A) with two strands of poly(U) or poly(d(T) have been known since 1965 [28, 29] and triple helices containing G·C·CH⁺ triads and I·C·CH⁺ triads shortly thereafter [30, 31]. To our knowledge, only in the case of the combination of poly(A) with two strands of poly(U) there

---

**Fig. 2.** 600 MHz ¹H NMR spectra of d5’TCTCTC-TTT-GAGAGA3’ (T3) in 95% H₂O:5% D₂O at 294 K; 2mM of T3 was dissolved in 0.1 M sodium deuteracetate. (A) pH = 8 (B) pH = 4.5. Base numbering corresponds to that in Figure 1.

**Fig. 3.** 1D-¹H-NMR spectra of T3 at pH = 5.5 at different concentrations and temperatures. Sample concentrations were 12mM (A, B and C), 4mM (D, E and F) and 1.3mM (G, H and I) respectively. The spectra shown were measured at 14.0°C ((A), (D) and (G)). at 21.2°C ((B), (E) and (H)) and at 29.2°C ((C), (F) and (I)). Solutions were prepared as described in ‘Materials and Methods’. Spectra were acquired with a sweep width of 32000 Hz, 8K data points and 128 acquisitions. All spectra were zero-filled to 16K and multiplied by a Lorentz-Gauss filter function prior to Fourier-transformation. Assignments of imino proton peaks from the triplex and duplex forms are given in Figure 2.
have been extensive thermodynamic studies. The experiments of Krakauer and Sturtevant [28] and Ross and Scruggs [29] showed that for the reaction

\[ 2A_n \cdot U_n = A_n \cdot 2U_n + A_n \]

the enthalpy for triple helix formation per base triad amounts to +3.5 kcal mole\(^{-1}\), and therefore formation of this triplex from two existing duplexes was unlikely.

To obtain more insight into the process of triplex formation from two stable hairpins we recorded 1D NMR spectra as a function of temperature at constant pH and concentration. These spectra show that upon raising the temperature, triple stranded molecules returned to the monomeric duplex form (c.f. Fig 3). The enthalpy of the transition from duplex to triplex can be obtained from the ratio (q) of duplex to triplex calculated from the relative intensities of imino proton resonances from the two forms present in the mixture. Consider the reaction from duplex (D) molecules to form a single triplex (T), which in the case of oligonucleotide T3 also involves the binding of three protons (H\(^+\)):

\[ 2D + 3H^+ = T \] (1)

The equilibrium constant \( K_{eq} \) of this reaction can be written as

\[ K_{eq} = \frac{[T]}{[D][H^+]^3} \] (2)

With the known pH of the sample, the total DNA-concentration \([\text{DNA}]_{\text{total}}\), \((\text{DNA})_{\text{total}} = [D] + 2[T])\) and the ratio...
duplex/triplex, \( q \), \( q = [D]/[T] \) determined from the NMR spectra by integrating the peaks belonging to each form, the equilibrium constant \( K_{eq} \) can be calculated:

\[
K_{eq} = \frac{q+2}{q[H^+][DNA]_{total}}
\]

(3)

Plotting \( \ln (q+2)/q^2 \) against \( \ln [DNA]_{total} \) will yield a straight line with an intercept equal to \( \ln K_{eq} + 3 \ln [H^+] \) from which \( \ln K_{eq} \) can be derived. Subsequently, the reaction enthalpy can be derived by means of the Van’t Hoff-equation:

\[
\frac{\partial \ln K_{eq}}{\partial 1/T} = -\frac{\Delta H}{R}
\]

(4)

In Figure 5, \( \ln K_{eq} \) is plotted against \( 1/T \) from which it can be derived that \( \Delta H = -16 \pm 3 \) kcal mole\(^{-1} \). Since six base triplets are present in the complex the mean transition enthalpy per base triad is \(-2.7 \) kcal mole\(^{-1} \).

Before the bimolecular complex of T3 can form one of the two hairpins must denature. The transition enthalpy for the latter reaction is estimated to be 32.4 kcal mole\(^{-1} \) [32]. Together with the experimentally determined enthalpy value of \(-16 \) kcal mole\(^{-1} \) for the reaction \( 2D + 3H^+ \rightarrow T \) the enthalpy of triplex formation from a hairpin duplex and a single stranded pyrimidine oligonucleotide was calculated to be \(-48.4 \) kcal mole\(^{-1} \). It is clear that this triplex molecule is energetically much more favourable than the two separate hairpin structures. This is also reflected in the free enthalpy of reaction, \( \Delta G \), for formation of the triple helix T3. For the reaction \( 2D + 3H^+ \rightarrow T \) the equilibrium constant, \( K_{eq} \), is known for several temperatures and it follows (using \( \Delta G = -RT \ln K_{eq} \)) that \( \Delta G = -24 \) kcal mole\(^{-1} \) at 294K. Since the free enthalpy change of opening the hairpin into its single stranded state is calculated to be 8 kcal mole\(^{-1} \) [32], we find that \( \Delta G = -32 \) kcal mole\(^{-1} \) for the reaction of one open hairpin with one closed hairpin to yield a triplex. The entropy change, \( \Delta S \), is then equal to 85.8 cal mole\(^{-1} \) K\(^{-1} \). As a result, the mean entropy per base triplet will amount to \(-9.3 \) cal mole\(^{-1} \) K\(^{-1} \) basetriplet\(^{-1} \). Using these results and those previously obtained for hairpin formation by Breslauer [32] it is possible to estimate the free enthalpies of formation of this molecule and the internuclear interactions in the form of intermolecular interactions. For short oligonucleotides it can be shown that the reaction leading to internuclear interactions is usually more favourable than refolding the molecule to an intramolecular triplet.

Existing structural information about triple helices comes from fibre diffraction studies of a \( d(A)_n d(T)_n d(T)_n \) and a \( d(A)_n d(U)_n d(U)_n \) system [13], and from the NMR-studies of triple helices formed by \( d(GA)_4 \) and \( d(TC)_4 \) oligonucleotides [25, 26], by the 11mers \( d(CTCTCGTCTCC) \), \( d(GAGGGAAGAAG) \) and \( d(CCTCTCTCTC) \) by the 7mers \( d(CCTCTC) \), \( d(GAGGAGAAG) \) and \( d(CCTCTCTC) \) [27].

Rajagopal & Feigon reported that in the triple helix \( d(GA)_4 d(TC)_4 d(TC)_4 \) both homopyrimidine strands have N-type sugars close to the C3'-endo conformation while residues of the purine strand have S-type sugar conformations [26]. This proposal was based on the observation of NOE crosspeaks between all H6 resonances of the pyrimidines and one or more H3' resonances, which is expected for DNA with N-type sugar conformations. De los Santos et al. proposed that the oligopurine strand of the 11-mer triplex has an A-type of helical stack. This conclusion was based on the observation that the H8 protons of the adenine bases resonate approximately 1 ppm upfield of their usual position in a B-type helix as predicted by ring current calculations.

We also find that N-type sugar conformations are mixed in the triple helical complex formed by the T3-oligonucleotide. Strong cross peaks in the NOESY spectra are observed between the aromatic protons and the H3' protons (also for short mixing times). Direct cross peaks between these protons are predicted for N-type but not for S-type deoxyribose residues. However, since the short interatomic distance between aromatic protons and H2' or H2'' protons create facile spin diffusion pathways between aromatic and H3' protons it is important to distinguish between direct and indirect H3' to aromatic proton cross peaks. Therefore, MINSY spectra were obtained for the low pH T3 complex at a series of different mixing times. Saturation of the H2' and H2'' resonances during the mixing time of the MINSY experiment inactivates the spin diffusion pathway to aromatic which lie in different strands of the triplex, the assignment is indicated. Strong NOE's between H5' (or H5'') and aromatic proton resonances which are of about the same intensity as the aforementioned H6/H8/H3' cross peaks indicate that some phosphate backbone torsion angles adopt unusual values.

Fig. 6. Expanded region of a 600 MHz \(^1\)H-MINSY spectrum of the low pH form of T3, dissolved in 0.1M sodium deuteroacetate in D\(_2\)O. The region shown is the one in which cross peaks are found between the H3', H4', H5' or H5'' resonances and the aromatic proton resonances. N-type sugars are present in several parts of the triplex as witnessed by the strong NOE's between H3' and aromatic protons. For three aromatic protons which have strong cross peaks to H3' protons which lie in different strands of the triplex, the assignment is indicated. Strong NOE's between H5' (or H5'') and aromatic proton resonances which are of about the same intensity as the aforementioned H6/H8/H3' cross peaks indicate that some phosphate backbone torsion angles adopt unusual values.
all three strands of the protonated dimeric form of T3 (see assignments in Fig. 6). It appears probable that the differences between our results and those of Rajagopal and Feigon and de los Santos et al. are the result of strain in the triple helical stem structure of the T3 oligonucleotide caused by the relatively small three residue loop which is present in this complex but is not present in the linear triplexes studied by the others. Unusual folding of residues in the T3 loop is indicated by a number of strong NOE's in the NOESY spectra involving methyl protons of the thymines.

An additional surprising feature of the low pH NOESY and MINSY spectra of the T3 oligonucleotide are several strong cross peaks between aromatic proton- and H5'- or H5''- resonances of sugar residues as is shown in Figure 6. These peaks are present in MINSY spectra recorded with short mixing times and are therefore predominantly due to direct interaction between H5' or H5'' and the aromatic proton spins. In both A and B forms of DNA where the O=C5'-C4'-C3' (γ) torsion angle is in the gauche+ range the distance between H5'/H5'' protons and aromatic protons is in a range in which NOE's are either weak or absent. Even if one makes allowance for spin diffusion, in normal A-type DNA the cross peaks between the aromatic H6/H8 and the H5'/H5'' resonances remain appreciably smaller than those between H6/H8 and H3' resonances. The presence of strong cross peaks between aromatic and H5' or H5'' protons in the low pH T3 spectra with intensities approaching those of the H5''-H6 cross peaks therefore suggest that in these residues the τ torsion angle deviates from the usual gauche+ value (possibly gauche− or trans).

The differences between the observations reported here and those of Rajagopal and Feigon and de los Santos et al. suggest that protonated triple helices formed by d(T,C)₆ and d(G,A)₆ oligonucleotides should be considered to be a family of related but distinct structures which can vary in response to local strain, rather than as a single structure with a unique set of backbone torsion angles.

As noted in the introduction the outstanding feature of the low pH 'H-forms' of polypurine/polypyrimidine sequences embedded in plasmids is their sensitivity to S1-nuclease. Although this nuclease is usually regarded as being specific for single stranded nucleic acids, sites of cleavage are found in those parts of the polypurine strands which would have to participate in the triple helices of H-form structures. Reactivity profiles are extremely variable among different plasmid inserts of apparently similar polypurine/polypyrimidine sequences although they are reproducible for single plasmids. A possible reason for both the appearance and extreme variability of the S1-nuclease sensitive sites in the pyrimidine strands of triplex structures is that highly individual strains at the foldback and base ends of H-form triple helices alter the distribution of τ torsion angles. Although no crystal structure is available for a single strand specific endonuclease of the class represented by S1-nuclease, there are strong reasons for believing that an unusual torsion angle γ is likely to play a role in its mode of phosphodiester bond recognition. Neither base nor H2' substituent on the sugar play an important role in the recognition of cleavage sites by members of the S1 class of enzymes. Therefore the key determinants which permit strand cleavage by S1-nuclease are restricted to the relatively small part of the phosphodiester backbone which includes the τ torsion angle. Within helical structures other potentially important torsion angles (α, β, δ, ε, and ζ) are either restricted by steric factors to narrow ranges (β, δ, ε) or are strongly dependent on γ (α, ζ). The sugar puck determined by the δ torsion angle is also unlikely to be the key determinant since S1-nuclease cleaves both DNA and RNA. In RNA a C3'-endo conformation is preferred (δ ≈ 85°) whereas in DNA a C2'-endo conformation (δ ≈ 130°) is preferred.

ACKNOWLEDGEMENTS

NMR-spectra were recorded at the Dutch National HF-NMR Facility at Nijmegen (The Netherlands). Ing. J. Joordens is thanked for expert technical assistance. This work was supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Research (NWO) and by an operating grant from the Medical Research Council of Canada.

REFERENCES