

COMMUNICATION

Kinetic Discrimination in the Folding of Intramolecular Triple Helices

Richard W. Roberts^{1*} and Donald M. Crothers^{2*}

¹Department of Molecular Biology, Massachusetts General Hospital, Boston MA 02114, USA

²Department of Chemistry Yale University, New Haven CT 06511, USA

We report a study of the physical properties of oligonucleotide intramolecular Pyr·Pur·Pyr triplexes modeled after *H*-form DNA. The experiments utilized a set of palindromic pyrimidine strands which form triplexes when combined with complementary purine strands. Triplexes of this nature have two possible isomers, one where the 3' half of the pyrimidine strand acts as the third strand (Y3) or one where the 5' end does (Y5). Kinetic studies of these triplexes revealed that the Y3 isomer folded 10 to 50 times faster than the corresponding Y5 isomer. Despite these kinetic differences, the complexes display relatively similar equilibrium stabilities, with seven of eight falling within a 1.1 kcal range. Addition of non-pairing sequence to the ends of the purine strand both reverses the kinetic bias (slowing Y3 formation >200 fold) and destabilizes the Y3 isomer 1.4 kcal/mol relative to Y5. Three features appear to lie at the source of both the kinetic and thermodynamic variability seen: (1) the prenucleation geometry of the triplexes prior to formation; (2) the accessibility of the major groove to the third strand; and (3) the nature of the triplex loop formed. Based on the data we propose a model for formation of *H*-form DNA that explains the biases observed for one isomer over the other in different situations. The conclusions have general implications for the tertiary folding of nucleic acids.

© 1996 Academic Press Limited

*Corresponding authors

Keywords: DNA; triplex; stopped flow; melting curves; kinetics

One of the most important questions facing molecular biologists is understanding the determinants of tertiary structure in nucleic acids. A clear picture already exists for the formation of ordered secondary structures both in terms of kinetics of helix formation (Pörschke & Eigen, 1971; Craig *et al.*, 1971; Riesner & Romer, 1973) and the stability of the resulting duplex (Freier *et al.*, 1986; Breslauer *et al.*, 1986). However, the variety of tertiary folds possible, especially in RNA, makes development of systematic rules and broad generalizations difficult. The problem is most evident when looking at large RNA structures. Even though the stability of an isolated helix can be predicted with good accuracy, it is often unclear whether such a helix will have a chance to form in the process of folding following transcription. Indeed, phylogenetic comparisons remain the most reliable method for prediction of secondary (e.g. see Michel *et al.*, 1982; Noller *et al.*, 1981, 1988) and tertiary (e.g. see Gutell *et al.*, 1986;

Michel *et al.*, 1990; Michel & Westhoff, 1990; Green & Szostak, 1994) interactions.

Our work has focused on the tertiary interactions associated with Pyr·Pur·Pyr triple helix formation. There, a pyrimidine "third strand" binds to the floor of the major groove of a Pur·Pyr duplex parallel to the duplex purine strand via "Hoogsteen" hydrogen bonding to form either T·AT or C⁺·GC base triples (reviewed by Maher *et al.*, 1991; Cheng & Pettitt, 1992; Nguyen & Hélène, 1993). Formation of the C⁺·GC triple requires protonation of third strand C bases above their normal pK_a value, causing triplexes of this nature to have a pH-dependent stability, a fact we have taken advantage of in our equilibrium and kinetic studies both to isolate the triplex denaturation from that of the duplex and to form or dissociate the triplex by pH drop or jump experiments. One unique feature of intramolecular triplexes is the directional nature of the zipping reaction, either 5' → 3' or 3' → 5'. The differences between these two processes have never been studied before, since the 2-fold

Abbreviation used: SAS, solvent-accessible surface.

symmetric duplex necessarily zippers in both directions at the same time.

H-form DNA and the "isomer paradox"

Studies of intramolecular triplexes have largely centered on *H*-form DNA, an intramolecular triplex discovered in circular plasmids (Lyamichev *et al.*, 1985, 1986, Christophe *et al.*, 1985). The structure forms under negative superhelical stress, low pH, or both, and is likely responsible for many cases of S_1 nuclease sensitivity observed in double-stranded plasmids (Mirkin *et al.*, 1987; for a review, see Frank-Kamenetskii, 1990). Further studies revealed an initially puzzling result, that one of the two possible isomers, H-y3, was greatly preferred over the other, H-y5 (Voloshin *et al.*, 1988; Johnston, 1988; Htun & Dahlberg, 1988; Hanvey *et al.*, 1988). This result was clarified by Htun & Dahlberg (1989, 1990), who demonstrated that the bias was likely to result from inequivalent topology of nucleation between the two isomers. This inequivalence energetically favored H-y3 under high levels of negative (unwinding) super-helical stress. In addition, they demonstrated that reduction of this stress reversed the bias, favoring H-y5. They proposed that steric constraints cause H-y5 to be preferred in the absence of a torsional driving force (Htun & Dahlberg, 1989, 1990).

In spite of Htun and Dahlberg's solution to the "isomer paradox" in *H*-DNA, reports persisted of cases where the distribution of isomers and conformations could not be easily explained with this model (Parniewski *et al.*, 1989; Shimizu *et al.*, 1990; Kang *et al.*, 1992a,b; Kang & Wells, 1992). In a paper by Shimizu and co-workers (Shimizu *et al.*, 1993) these inconsistencies seem to have been resolved. They demonstrated that G + C-rich sequences at the dyad and/or divalent metal ions, both caused a preference for the H-y5 isomer, whereas A + T-rich sequences and monovalent ions favored H-y3 formation. The implication from their work was twofold: (1) that the nucleation event was the critical feature in determining the preference for one isomer over the other (just as previously argued by Htun & Dahlberg, 1989, 1990); and (2) that the nucleation of the two isomers was mechanistically different, with denaturation or distortion of the helix being necessary for H-y3 formation to a greater extent than in the H-y5 case.

Previous oligonucleotide studies

Modeling, thermodynamic, and structural studies have been conducted on oligonucleotides that

are also important in our understanding of intramolecular triplex formation. Initially, intramolecular triplexes were used because they proved both more tractable for NMR studies of triplexes than their intermolecular homologs (Sklendar & Feigon, 1990) and because they provided oligo homologs to *H*-form DNA studied in plasmids (Xodo *et al.*, 1990). Subsequently, intramolecular triplexes have been used for NMR structural studies of both canonical and non-canonical interactions within Pyr-Pur-Pyr triplexes (e.g. see Macaya *et al.*, 1992; Radhakrishnan *et al.*, 1993) and detailed studies of the triplex loops (Prakash & Kool, 1992; Booher *et al.*, 1994; Wang *et al.*, 1994). However, neither thermodynamic nor structural studies of model oligonucleotide triplexes have demonstrated why one isomer should be preferred in the plasmid case. Indeed the abundance of data available indicate that where the loops are the same, the two triplex isomers have very similar stabilities, providing little driving force to favor one over the other (Häner & Dervan, 1990; Booher *et al.*, 1994; Wang *et al.*, 1994).

Modeling studies have provided insight into this problem, for example in the studies by Harvey *et al.* (1988) on the loop folding of *H*-form ends. Based on the notion by Haasnoot *et al.* (1986), that the most stable loops will preserve the normal stacking on one of the strands in the stem, Harvey *et al.* (1988) produced models for loops containing three to six T residues. There, the prediction was that the most stable structures should have the majority of the loop bases stacked on the 3' end of the stem, and with the optimal loop size of about four bases.

Both NMR and thermodynamic studies on oligonucleotide intramolecular triplexes support this basic idea (Macaya *et al.*, 1992; Prakash & Kool, 1992). NMR studies demonstrated that the first two bases of the T_3 triplex loop stack on the 3' end of the duplex pyrimidine strand (Macaya *et al.*, 1992). Here, the three base loop may be too short, as the loop-proximal triple appears disrupted. Using a slightly different system, experiments by Prakash & Kool (1992) point to an optimal loop size of five. Kool and co-workers have investigated the thermodynamics of interactions extensively, both within the triplex loop and between the loop and extensions of the duplex purine strand (Wang *et al.*, 1994; Booher *et al.*, 1994).

Here, we present a physical study of the tertiary folding of short DNA oligonucleotides capable of forming intramolecular Pyr-Pur-Pyr triple helices. The two possible triplex isomers (denoted Y3 and Y5) differ only in which end of the pyrimidine strand acts as the third strand, the 3' or 5', respectively (Figure 1)†. Despite this seemingly minor difference, the two complexes display large differences in their formation kinetics. Surprisingly, there are often compensating differences in dissociation kinetics, and hence only modest overall differences in equilibrium stabilities. Our results indicate that both the prenucleation

† This follows the convention adopted by Htun & Dahlberg (1989) used for *H*-DNA where the isomer names were H-y3 and H-y5. The Y3 isomer corresponds to a 5' loop and the Y5 isomer corresponds to a 3' loop as described by Wang *et al.* (1994) and Booher *et al.* (1994).

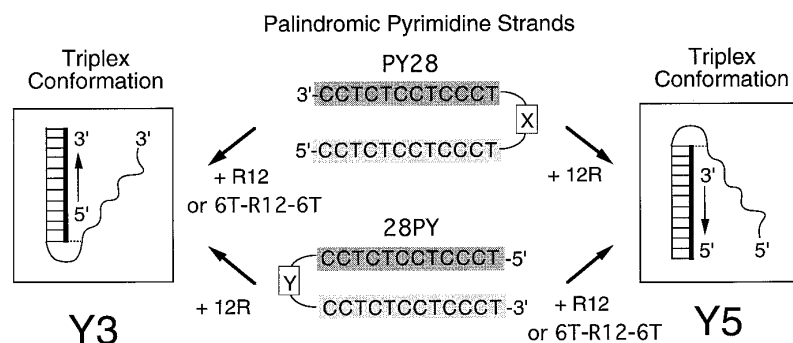


Figure 1. Geometry of intra-molecular triplexes. The two varieties of palindromic pyrimidine strands (PY28 and 28PY) and two intramolecular triplex isomers they can form (Y3 and Y5) are shown (see also Table 1). The direction of third strand “zippering” accompanying triplex formation and the triplex end of the pyrimidine strands are indicated. PY28 and 28PY sequences differ from one another in the position, length, and composition of their triplex loops (X and Y respectively).

ively). R12 and 12R purine strands both fully base-pair to the pyrimidine strands used. The triplexes formed from 12R versus R12 are not perfectly isomeric because purine strands are not palindromic but have identical sequence content. 6T-R12-6T contains six “non-pairing” thymidine residues at either end, creating both a 3' and 5' overhang (purine extension) at its ends.

geometry and accessibility of the major groove play pivotal roles in the kinetic biases observed and give insight into the nucleation process accompanying the formation of *H*-DNA in plasmids. In addition, the structural biases observed may also help to understand and predict the folding of other nucleic acid structures.

Sequence design

The oligonucleotide sequences used were modeled after the triplex portion of *H*-form DNA and are shown in Table 1 and Figure 1. The complexes

consisted of a purine strand, which formed a triple helix when combined with one of the palindromic pyrimidine strands to give either the Y3 or Y5 isomer. We refer to triplexes constructed from the R12 and 12R purine strands as “blunt” and those containing the oligo 6T-R12-6T are referred to as “purine extensions” due to the extra T residues extending from both its 3' and 5' ends. Combination of the various pyrimidine strands with one of the purine strands (R12, 12R or 6T-R12-6T) allowed us to make 16 intramolecular triplexes with varying conformation (Y3 versus Y5), triplex loop size (4, 6, 8, and 20), triplex loop sequence, nucleating base (the first pyrimidine base after the loop sequence) and end geometry.

Table 1. Sequences and triplex conformations

A. Sequences		
Purine strands		
R12	3'-GGAGAGGAGGGA-5'	
12R	5'-GGAGAGGAGGGA-3'	
6T-R12-6T	3'-TTTTTGGAGAGGAGGGATTTTT-5'	
PY28 pyrimidine strands		
PY28	5'-CCTCTCCTCCCTTATATCCCTCCTCTCC-3'	
4T-PY28	5'-CCTCTCCTCCCT-T ₄ -TCCCTCCTCTCC-3'	
8T-PY28	5'-CCTCTCCTCCCT-T ₈ -TCCCTCCTCTCC-3'	
20T-PY28	5'-CCTCTCCTCCCT-T ₂₀ -TCCCTCCTCTCC-3'	
28PY pyrimidine strands		
28PY	5'-TCCCTCCTCCTATACCTCTCCTCCCT-3'	
4T-28PY	5'-TCCCTCCTCCTCC-T ₄ -CCTCTCCTCCCT-3'	
6T-28PY	5'-TCCCTCCTCCTCC-T ₆ -CCTCTCCTCCCT-3'	
B. Conformations		
Pyrimidine strand	+ R12 (or 6T-R12-6T)	+ 12R
PY28	Y3	Y5
4T-PY28	Y3	Y5
8T-PY28	Y3	Y5
20T-PY28	Y3	Y5
28PY	Y5	Y3
4T-PY28	Y5	Y3
6T-PY28	Y5	Y3

All DNA used was synthesized on an Applied Biosystems 380B DNA synthesizer with cyanoethyl phosphoramidite chemistry and standard procedures. Deprotected samples were purified by denaturing polyacrylamide gel electrophoresis, electroeluted from the gel slices with an Elutrap (Schleicher and Schuell), and desalted with Sephadex NAP-25 columns (Pharmacia).

Y3 and Y5 isomers have similar stability

We began by looking at the relative stability of a Y3 and Y5 isomer derived from the R12 purine strand (R12 + PY28 versus R12 + 28PY). Absorbance melting curves (Figure 2A) demonstrated that their stabilities were essentially identical, differing by only 1 deg.C in melting temperature, t_m (0.25 kcal/mol in free energy (ΔG°)), which is within the error range (± 0.5 kcal/mol) of the experiment. As mentioned previously this is what would be expected based on the data available, which suggest the two isomers should have roughly equivalent stability if the loops are the same (Wang *et al.*, 1994; Booher *et al.*, 1994).

Y3 isomers fold faster than Y5 isomers

In contrast to the similarity of the stability of the isomers, kinetic studies demonstrated large differences in their folding rates (Figure 2B). There the Y3 isomer (R12 + PY28) forms about tenfold faster than the Y5 isomer (R12 + 28PY). This was initially surprising as a difference in the k_{on} value this large would be expected to produce a drop in the t_m of about 5.5 deg.C. (This calculation assumes that the two dissociation rates are the same, which it turns out that they are not.) In an attempt to understand

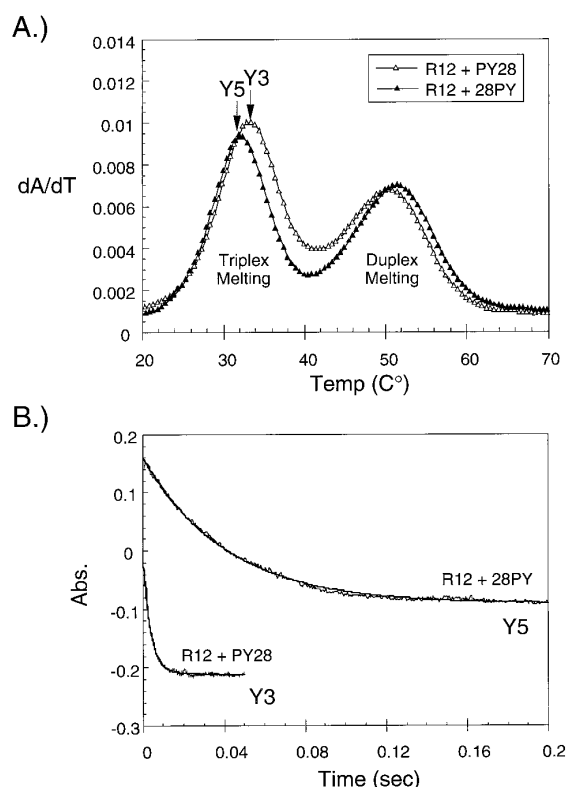


Figure 2. Comparison of stability versus kinetics of Y3 versus Y5 triplex isomers. A, The relative stability of Y3 versus Y5 isomers as determined by absorbance melting curves (260 nm). The plot shows the derivative (dA/dT) melting curves of R12 + PY28 (Δ) the Y3 isomer and R12 + 28PY (\blacktriangle) the Y5 isomer taken at pH 6.5, 100 mM Na⁺ cacodylate. In both cases the low temperature transition corresponds to the melting of the triplex and the higher transition to that of the duplex. The t_m values of the Y3 and Y5 isomers are marked.

Determination of triplex stability: the stability of all the intramolecular triplexes and bimolecular duplexes studied here were determined by van't Hoff analysis of absorbance melting curves as described (Roberts, 1993; Roberts & Crothers 1991). Briefly, all samples were equilibrated by cooling from 85°C to room temperature. Melting experiments were performed by heating from low temperature to high temperature at ~ 0.5 deg. C/min and data were taken every 0.5 deg. C (t_m values were reproducible ± 0.5 deg. C and independent of ramp rate). The derivative melting data were treated with a statistical mechanical approach just as described by Roberts & Crothers (1991). Here, the fraction of triplex, Θ_3 , is given by:

$$\Theta_3(t) = \frac{K_2 K_3 (C_t/2) (1 - \Theta_2(t))}{[K_2 K_3 (C_t/2) (1 - \Theta_2(t)) + K_2 (C_t/2) (1 - \Theta_2(t)) + 1]} \quad (1)$$

and the fraction of duplex, Θ_2 , is given by:

$$\Theta_2(t) = \frac{K_2 K_3 (C_t/2) (1 - \Theta_2(t)) + K_2 (C_t/2) (1 - \Theta_2(t))}{[K_2 K_3 (C_t/2) (1 - \Theta_2(t)) + K_2 (C_t/2) (1 - \Theta_2(t)) + 1]} \quad (2)$$

K_2 and K_3 are the equilibrium formation constants for the bimolecular duplex and intramolecular triplex respect-

ively. C_t is the total concentration of oligonucleotide and $C_t/2$ is the concentration of the third strand. To determine the values of ΔH° van't Hoff and t_m , experimental derivative melting curves were compared to calculated curves and the parameters (ΔH_2° , ΔH_3° , t_{m2} , and t_{m3}) were varied to give the best fit. Trial derivative melting curves were generated by solving equations (1) and (2) numerically over the temperature range of the melting curve. ΔG° and ΔS° were then determined for each of the triplexes and duplexes studied using the ΔH° van't Hoff and t_m data. The approximate uncertainty in ΔH° is ~ 5 kcal/mol, producing an uncertainty in ΔG° of ± 0.5 kcal/mol. A complete table of the thermodynamic parameters (ΔG° , ΔH° , ΔS° , and t_m) determined using this analysis are available from the authors upon request.

B, Relative kinetics of Y3 and Y5 formation as determined by absorbance detected stopped flow. The graph shows representative kinetic measurements of triplex annealing resulting from a pH drop experiment (pH 7.5 to 6.0) on the complex displayed in (A).

Determination of folding kinetics: the kinetics of formation and dissociation for all the intramolecular triplexes studied were determined by absorbance detected stopped flow using a Spectra Kinetic DX.17MV workstation. The temperature was controlled by a Neslab digital circulating bath and monitored by a thermistor in the sample head. For both association and dissociation reactions, one to four experiments were performed at each temperature, and the data averaged.

Standard association experiments were carried out by pH drop. Annealed duplexes (a palindromic pyrimidine strand + purine strand, 2 mM each) in 100 mM NaCl, 2.5 mM Na⁺ HPO₄/H₂PO₄, 0.025 mM EDTA (pH 7.5) were mixed with 100 mM Na⁺ cacodylate buffer of the desired final pH, 6.0. Dissociation experiments were performed by pH jump. Annealed triplexes (2 μ M concentration) in acidic buffer (100 mM Na⁺ acetate/acetic acid, 1 mM EDTA (pH 5.0)) were mixed with an equal volume of 100 mM Na₂HPO₄, producing a buffer with pH = 7.2. The Na⁺ concentration was 100 mM in all experiments before and after the jump.

The formation (k_{on}) and dissociation (k_{off}) rate constants for the triplex were determined by non-linear least-squares fitting of experimental data to a first-order rate equation. The time dependence of the concentration is given by:

$$[A] = [A]_0 e^{-kt} \quad (3)$$

Values for both k_{on} and k_{off} were determined from the data over the range of temperature where the reaction occurred. Using these data, the Arrhenius activation energy (E_a) of association and dissociation was determined for all the complexes studied (our unpublished observation). Finally, the pH dependence and proton reaction order were determined for a representative triplex of both the Y3 and Y5 isomers by varying the final pH of the association and dissociation reactions. A complete table of the kinetic parameters (k_{on} , k_{off} , $E_{a,on}$, $E_{a,off}$) determined using this analysis is available from the authors upon request.

the underlying sources of Y3 group's faster folding kinetics, the folding was studied as a function of pH and temperature for both isomers in order to determine the mechanism of the reaction. The results of these studies revealed that the tenfold kinetic bias was maintained over the range

examined and that formation was mechanistically similar for both isomers (Roberts, 1993). Specifically, the pH and temperature dependence of the rate constant implied that formation of two to three triples was the rate-limiting step in helix formation as formation has a small negative activation energy and requires ~ 1.5 protons in both cases (Roberts, 1993; see also Pörschke & Eigen, 1971; Craig *et al.*, 1971). Determination of the dissociation rate constants of both isomers revealed a faster off rate for the Y3 isomer than Y5, resulting in the negligible overall stability difference.

Kinetics versus stability

We wanted to understand how complexes with such similar stabilities and (apparently) similar geometries could have such marked differences in their formation kinetics. To do this, we examined a number of other triplexes with variations in the first triple formed (nucleation region), loop sequence, loop size, and end geometry. The triplex formation rate constant and equilibrium stability were determined (kcal/mol) for each complex studied and then plotted as a single x, y "phase point" to compare with the other triplexes (Figure 3). On the left and right of each plot is shown a schematic view of the expected geometry of the Y3 and Y5 triplex isomers prior to nucleation of the third strand.

Just as with the initial example, small energetic differences were seen between the Y3 and Y5 isomers containing the same sized loops. Figure 3A (y axis) demonstrates that seven out of eight of the four base loop triplexes fall within 1.1 kcal (about 4.5 deg.C difference in t_m) of each other. This is significantly smaller than ~ 9 deg.C range of t_m values the observed 50-fold difference in formation kinetics would be expected to produce. The similarity of Y3 and Y5 stability is also seen for loops of six (Figure 3B) four base loops with purine extensions (Figure 3C), loops of eight and 20 (not shown).

Another feature which can be seen in Figure 3A is the effect of loop sequence on the stability and kinetics of triplex formation. For both the Y3 and Y5 isomers, the loop sequence has a small but measurable effect on both, with the 5' TATA loop being more stable (by 0.1 to 0.8 kcal/mol) than the 5'-T4 loop (Figure 3A) and forming slightly faster as well (compare with the line in Figure 3A to C which has a slope of 1). Both the stability and kinetic increases are likely due to the better stacking expected in the TATA loop relative to the T4 loop. It should be noted that the formation rate of the slowest four base Y3 loop studied is still about tenfold faster than the fastest Y5 loop irrespective of the loop sequence.

One possible source of the kinetic biases seen is the sequence of the nucleation sites, which varies among the 4 nt loop triplexes, depending on the purine strand used. This appears not to be the case.

Y3 triplexes made using 12R show similar kinetics with those made from R12 even though the nucleation sites are 5'GGA... versus 5'AGG..., respectively. Similarly, Y5 triplexes made from 12R show little kinetic difference between those made from R12. Thus, we must look elsewhere to explain the differences seen in the formation kinetics.

Prenucleation geometry

One possible source of the faster formation rate in Y3 isomers with a four base loop is its pre-nucleation geometry, that is, the geometry of the complex prior to formation of the first triple (see schematic in Figure 3A). There the first pyrimidine in the third strand is poised directly over the purine site of helix nucleation (Figure 3A, Y3 isomer). In contrast, the third strand of the Y5 isomer (four base loop) lies over the major groove and would have to overwind and bend in order to form the first triple (Figure 3A, Y5 isomer).

If pre-nucleation geometry were an important source of discrimination, then addition of two bases to the triplex loop (loop = six total) should speed Y5 formation relative to that of Y3. This is indeed the case. Although the kinetic bias for Y3 is not fully reversed, addition of two bases to the Y3 loop does slow formation ~ 2.3 -fold, whereas Y5 formation is virtually unaffected (Figure 3B), confirming the importance of pre-nucleation geometry to the kinetics.

Further confirmation of the importance of pre-nucleation geometry is seen in the cases where the purine strand contains non-pairing 3' and 5' extensions. Although changing the size of the loop does not reverse the kinetic bias for Y3 over Y5, addition of extra sequence to the ends of the purine strand does (Figure 3C). The extra sequence does this predominantly by slowing Y3 formation more than 200-fold and destabilizing this isomer by 2.4 kcal/mol relative to the case without the purine extension. A much smaller kinetic effect is seen in the Y5 isomer where the formation rate is decreased only fivefold and the stability by 1.0 kcal/mol. The general destabilization resulting from the presence of a non-pairing extension is in good agreement with the work on similar systems by Kool and co-workers (Prakash & Kool, 1992; Wang *et al.*, 1994; Booher *et al.*, 1994).

This differential kinetic effect is most likely due to the steric interaction of the strands rather than a pairing interaction. The schematics in Figure 3C indicate how this could occur. In the case of the Y3 isomer, the extension of the purine strand blocks its path to the floor of the major groove (Figure 3C, Y3 isomer), whereas the path of Y5 is relatively unobstructed by the extension. A pairing interaction would be expected to produce a similar effect at both ends of the strand because the sequences and strand polarity are the same in each case. Finally, the decrease in the Y5 rate indicates

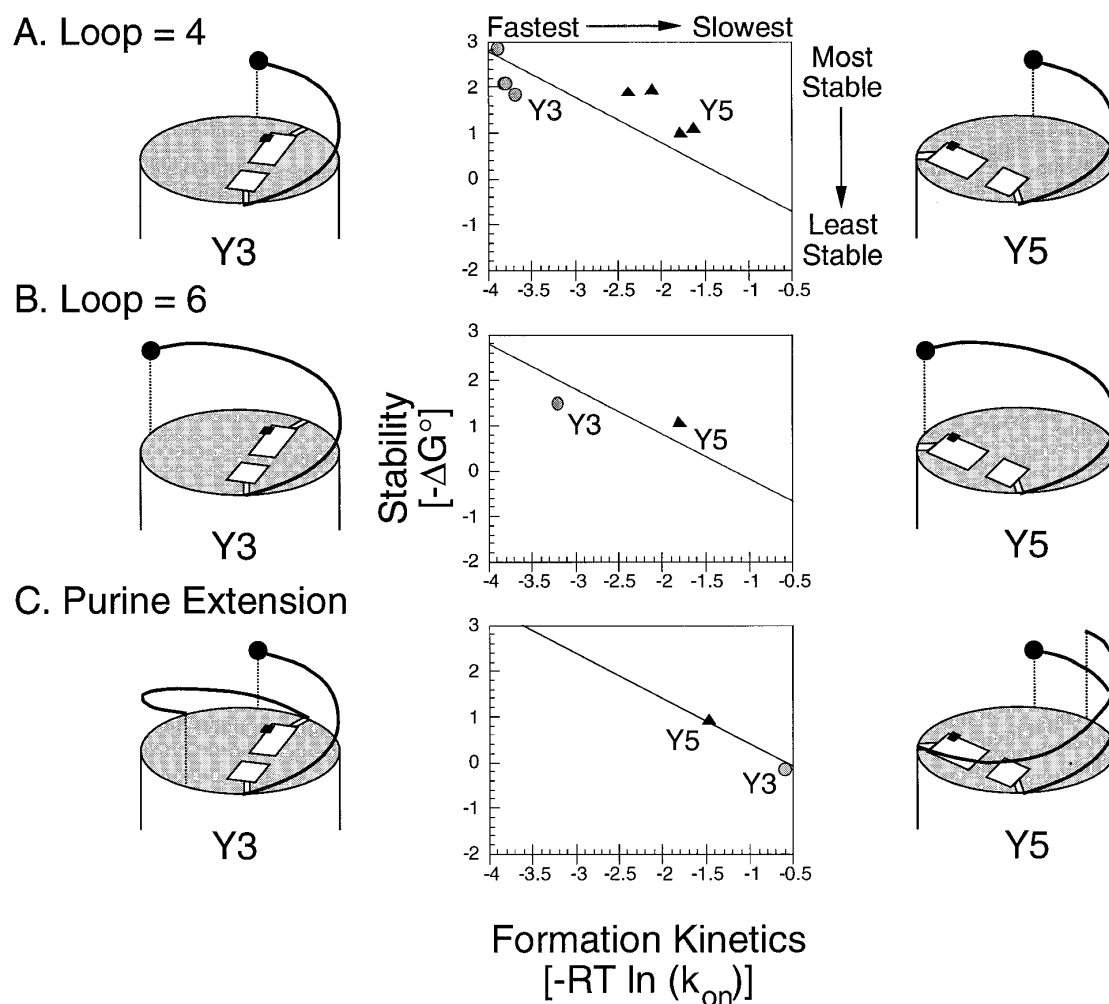


Figure 3. The effect of prenucleation geometry on the kinetics and stability of intramolecular triplexes studied at 25°C. *Graphs:* each point in the graphs represents data for a single intramolecular triplex. The x coordinate of each point represents the formation rate constant in kcal/mol for the triplex at 25°C, as determined by stopped flow measurements (error ± 0.4 kcal/mol). The y coordinate represents the equilibrium stability (here plotted as $-\Delta G^\circ$) of the complex at 25°C in kcal/mol as determined by absorbance detected melting curves (error ± 0.5 kcal/mol). All Y3 isomers are represented (●) and the Y5 isomers (▲) such that the most stable, fastest forming are at the upper left and the least stable, slowest forming are seen at the lower right. In each case a line corresponding to a slope of 1 has been included in the graph for reference. *Schematics:* the drawings beside each graph represent the expected geometry at the end of the helix prior to nucleation for the Y3 and Y5 isomers of different loop sizes. In all six drawings, the important features in formation of the first base triple are highlighted: (1) the black circle representing the first pyrimidine base in the triplex, the broken line indicating the projection of the first base onto the end of the helix, and the black square represents N-7 of the first purine. In all cases the drawings indicate the position of the strands corresponding to a right-handed 10.5-fold helix. A, Loop = 4. Both Y3 and Y5 isomers demonstrate quite similar stability (y axis). In contrast the kinetics of the two complexes are quite different. The prenucleation geometry of the Y3 isomer (left side) indicates the third strand is positioned directly over the site of nucleation whereas the third strand of the Y5 isomer (right side schematic) lies over the major groove. Eight complexes are shown. The two fastest forming Y3 and Y5 oligonucleotides both have 5'TATA loops whereas the slower forming ones of each isomer both have 5'TTTT loops. B, Loop = 6. The Y3 isomer forms about tenfold faster than the Y5 isomer of the same loop size in spite of the fact that the first pyrimidine base on the end of the helix is predicted to lie directly over the nucleation site. C, Purine Extension. Loop = 4, extension = T6. Here the Y5 isomer forms fivefold faster than the Y3 isomer. The Y3 schematic indicates that the T6 extension of the purine strand lies directly in the path required for nucleation. In Y5, the disposition of the strands provides no such steric blockage.

some interaction between the third strand and the extension even in this case. It is not known if this is due to a steric effect or a non-Watson-Crick pairing interaction between the third strand and the extension.

Dependence on loop size

In addition to having an important effect on the prenucleation geometry of the triplex, the size of the triplex loop has an important effect on both the

stability and kinetics of the intramolecular triplexes studied (Figure 4). We began by determining both the free energy for adding a triplex loop, $\Delta G_{\text{loop}}^{\circ}$, and the effective concentration of the third strand around the nucleation site, the J factor (Shore *et al.*, 1981), for several sizes of loops (Figure 4A). In general, as one inserts larger and larger loops, the J factor, the stability of the triplex, and the formation kinetics (Figure 4B) all decrease markedly. However, just as before the two isomers have somewhat different behavior, with Y3 loops being slightly more stable at small sizes (4 and 6 nt), but less stable than Y5 isomers at larger loop sizes (≥ 8). This crossover is not pronounced and may be due to subtle differences between the two such as (1) the fact that the Y5 loops must span a larger distance prior to nucleation (as previously mentioned), or (2) how the two loops are stacked on the ends of the helix after triplex formation.

Small triplex loops (4 nt) are very stabilizing to both isomers (Figure 4A). They generally provide an equal or greater stabilization to the triplex than most RNA loop sequences (Papanicolaou *et al.*, 1984) including the U_4 loop described by Groebe & Uhlenbeck (1988) and both the GNRA and UUCG tetraloops (SantaLucia *et al.*, 1992; Antao *et al.*, 1991) under slightly different conditions. This implies that the effective concentration (J) of the third strand around the duplex (~ 30 mM) is greater than the corresponding strands around each other prior to forming a hairpin. This is likely due to the higher degree of preorganization in the triplex prior to nucleation than a single-stranded RNA prior to formation of a hairpin.

Comparing Figure 4A with B, it is clear that the k_{on} value plays the primary role in distinguishing complexes of differing stability. This is particularly true for the Y3 triplexes where the range of k_{on} values matches the range of $\Delta G_{\text{loop}}^{\circ}$. In addition, k_{off} for Y3 shows little variance with loop size and is essentially equal to the intermolecular triplex dissociation rate in all cases studied (Figure 4C). k_{on} also carries the bulk of discrimination for Y5 isomers. However, some variance is seen in the k_{off} values for Y5 (Figure 4C), implying that dissociation is hindered by the presence of the triplex loop, especially at small loop size. Model building studies indicate that this could be due to the loop stacking on the triplex PY strand in the Y5 isomers and the duplex PY strand in the Y3 isomers. Thus, the Y3 isomer triplex strand would be free to dissociate from the duplex (just as in the intermolecular case) whereas the Y5 dissociation would require unstacking of the loop to get free.

The observation that k_{on} (and not k_{off}) plays the central role in kinetic discrimination of complexes was somewhat surprising, as for intermolecular duplexes exactly the opposite is observed. There, the value of k_{on} is relatively insensitive to the overall stability of the complexes, falling over a range of only about tenfold, with k_{off} playing the primary role in discrimination (Pörscke & Eigen, 1971; Craig *et al.*, 1971; Riesner & Romer, 1973). The

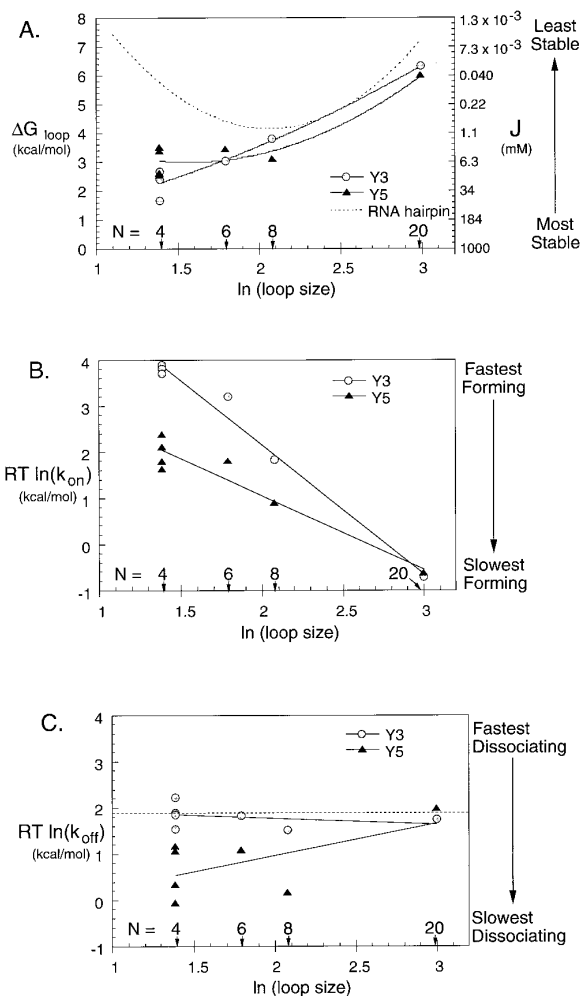


Figure 4. Effect of loop size on the equilibrium stability and kinetic behavior of intramolecular triplexes. In each case, the size of the triplex loop is indicated at the bottom of the graph. A, Free energy increment for connecting a third strand by a loop of size N ($\Delta G_{\text{loop}}^{\circ}$) and effective concentration of the third strand around the nucleation site (J factor) for Y3 (●) and Y5 (▲) triplexes. The broken line (---) indicates the expected value of $\Delta G_{\text{loop}}^{\circ}$ (or J) for RNA hairpins as predicted by Papanicolaou *et al.* (1984). The value of $\Delta G_{\text{loop}}^{\circ}$ was calculated as the difference free energy between the intermolecular complex and the intramolecular complex at the same pH and temperature ($\Delta G_{\text{intermolecular}}^{\circ} = -4.5$ kcal/mol at pH 6.5, 298°C (Roberts, 1993)). The value of J was calculated as the ratio of equilibrium constants, $J = K_{\text{intramolecular}}/K_{\text{intermolecular}}$. The stability of the Y3 isomer displays a nearly linear dependence on the $\ln(\text{loop size})$ even down to loops of four nucleotides. The Y5 isomer dependence shows more curvature at even moderate loop sizes, with loops of four, six and eight having similar stability. The slope of the Y3 isomer loops is roughly the same as the isomer with loops ≥ 8 , falling 2.5 versus 3.2 kcal/mol respectively. B, Formation rate constants (k_{on} , kcal/mol) for Y3 and Y5 isomers as a function of loop size at 25°C, pH 6.0. Both isomers show a roughly linear dependence on the $\ln(\text{loop size})$ with the Y3 isomer decreasing 2.8 kcal/mol and the Y5 isomer kinetics falling 1.6 kcal/mol as the loop is increased a factor of e . C, Dissociation rate constants (k_{off}) for Y3 and Y5 isomers as a function of loop size. The broken line indicates the dissociation rate constant for the intermolecular triplex for reference.

intermolecular triplexes studied by Rougée *et al.* (1992) provide an intermediate case with both k_{on} and k_{off} providing significant discrimination between complexes. Despite these differences, physically, our results make good sense. As the loop size increases, both the effective concentration of the third strand around the nucleation site (simply the J factor, Figure 4A) and the formation rate, should decrease. Thus, the overall effect of attaching the triplex by a loop is to increase the formation rate by several orders of magnitude (e.g. by a factor of $\sim 10^5$ at $1 \mu\text{M}$ total strand concentration) compared to the bimolecular case (Maher *et al.*, 1990; Rougée *et al.*, 1992; Roberts, 1993)).

Major groove accessibility

The fact that the relative kinetics of the Y3 and Y5 isomers cannot be swapped simply by changing the loop size (and in fact the two only become equivalent at large loop sizes: Figure 4B) implies the presence of another factor providing kinetic discrimination. The most likely candidate is the higher accessibility of the major groove in the Y3 isomer (Figure 5). There, the first triple is formed at the 5' end of the purine strand, which is more highly solvent exposed than the 3' end, as indicated by model building studies and solvent-accessible surface (SAS) calculations (performed using the program from Connolly, 1983 (as by Kagawa *et al.* (1989), but calculated on a 5' or 3'-terminal G rather than internal bases) (Figure 5). SAS calculations show the O-6 and N-7 of a G residue at the 5' end of a *B*-form DNA helix have 30 to 80% more contact area than the corresponding atoms at the 3' end of the helix. Specifically 5' guanine has contact areas of 8.97 \AA^2 for N-7 and 6.44 \AA^2 for O-6, whereas 3' guanine N-7 has a contact area of 5.0 \AA^2 and O-6, 5.0 \AA^2 . This difference results from the right-handed nature of DNA (and RNA) as the helix wraps away from the base at the 5' end (Figure 5, Y3 end), leaving its major groove face largely solvent exposed. On the 3' end, the opposite is true with the helix wrapping smoothly around the purine, decreasing its accessibility to the third strand, and likely slowing the kinetics of triplex formation (Figure 5, Y5 end).

Bearing on *H*-DNA

The differences seen in the formation kinetics between oligonucleotide Y3 and Y5 isomers provide mechanistic insight into the preferences observed in *H*-DNA formation. The logic is as follows. When conditions favor triplex formation over the Pyr·Pur duplex, nucleation of either isomer can occur. In the presence of small amounts of (–) superhelical stress, the duplex is not grossly distorted or underwound, favoring *H*-y5 formation by a simple kink (Figure 6A). Formation of the *H*-y3 isomer in this situation is unlikely, due to steric

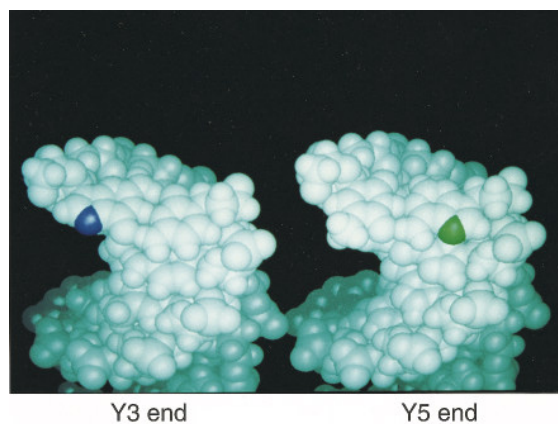


Figure 5. Accessibility of the major groove in Y3 and Y5 isomers. In each case a view down into the major groove of a *B*-form DNA helix is shown. The view corresponds to the end of the helix where nucleation is likely to occur in each isomer. Y3 end, The 5' end of the purine strand (N-7 colored purple) where nucleation occurs in the Y3 isomer. The purine base is highly solvent-exposed as the helix wraps away from it. Y5 end, The 3' end of the purine strand (N-7 colored green) where nucleation occurs in the Y5 isomer. Here, the purine base is somewhat occluded due to the fact that the DNA wraps around it in a right-handed fashion. Both models were constructed on a Silicon Graphics workstation using the Insight II software package and photographed from the screen.

hindrance from the duplex purine strand, just as the purine extension blocks formation of the Y3 isomer in the oligonucleotide case.

In order for the *H*-y3 isomer to form, the path to the major groove must be cleared. The presence of higher amounts of negative superhelical stress may do just that, by transiently unwinding and denaturing the strands out of the cylinder defining the helix at the dyad (Figure 6B). The bulging of the strands effectively clears a path to the floor of the major groove for *H*-y3 nucleation that was not present prior to denaturation. In this state, the kinetic differences which favor the Y3 isomers in the oligonucleotide system would probably also favor formation of the *H*-y3 isomer. The necessity for path clearing in the formation of *H*-y3 explains why conditions that increase duplex stability at the dyad (G + C-rich sequences and divalent metal ions) disfavor formation of *H*-y3 isomers (Shimizu *et al.*, 1993) and why a shift from *H*-y5 to *H*-y3 is seen as the negative superhelical stress is increased (Htun & Dahlberg, 1989, 1990).

RNA folding and a Y5 rule

In addition to *H*-DNA, we believe our results also have general implications for the formation of RNA tertiary interactions. This is because many of the same features responsible for the biases seen in the folding of intramolecular DNA triplexes are also present in RNA. First, because RNA and DNA possess many of the same features that determine

the relative folding of the Y3 and Y5 oligonucleotide isomers (prenucleation position of the strands relative to the floor of the major groove, and higher accessibility of the 5' end of the helix (Weeks & Crothers, 1993)). Second, even though long triplexes are unknown in functional RNA molecules (rRNA, tRNA, etc.), the rate-limiting step measured here involves formation of only two to three base triples, which may be quite common in the folding of large RNA molecules. Thus, by analogy with the DNA case we would predict that a "Y5 rule" would also apply to RNA. That is to say that when both 5' and 3' unpaired regions extend from a helix, the 5' strand would have the propensity to interact with the floor of the major groove just as in the case of the Y5 isomer with the purine extension (see Figures 3C and 7A).

In particular, this prediction has implications for the kinetics of tertiary folding at the junctions of two or more helical stems. Junctions of this sort are quite common: 16 S rRNA has at least 18, while 23 S rRNA has >28 (Noller *et al.*, 1981, 1988). Do any known structures correspond to such a rule? Two

come close. The first is the TAR RNA in the presence of argininamide (Figure 7B). There a two-way junction folds in the presence of a cofactor to give a single base triple between a pyrimidine and a duplex purine at the 5' end of a helix (Puglisi *et al.*, 1992). The second, tRNA, contains one four-way junction that also appears to conform to a Y5 type folding as shown in Figure 7C (reviewed by Rich & RajBhandary, 1976). There unpaired bases at the 5' end of the D stem fold into the major groove of the D helix, interacting with the floor of its major groove and the D loop. Although the pairing interactions in both cases are not the same as in the triplex, the general topology (5' loop interacting with the major groove) is similar to the Y5 looped end. One important exception to our proposal is the major and minor groove interactions seen at the P4/P6 junction in the group I intron, where close cohelical stacking of the strands places the 5' unpaired region into the minor groove of P6 and the 3' unpaired region into the major groove of P4 (Michel & Westhoff, 1990; Green & Szostak, 1994; Chastain & Tinoco, 1992, 1993).

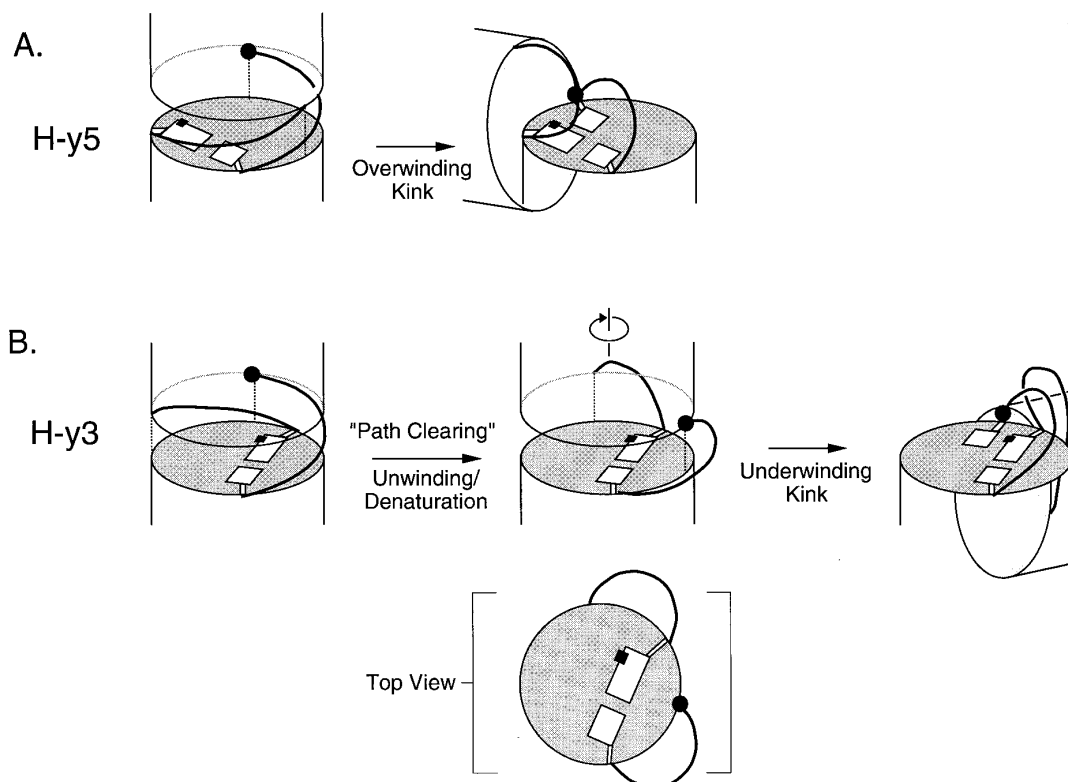


Figure 6. Proposed mechanism for H-y3 and H-y5 formation. The features of the schematic (first triplex pyrimidine, projection on the end of the helix, and purine N-7, are presented as for Figure 3C). A, Model for H-y5 formation. The first pyrimidine base (●) has an unobstructed path to the floor of the major groove where the first purine lies. Thus, triplex formation can occur by an overwinding kink toward the purine base. B, Model for H-y3 formation. Nucleation can occur by either of two paths. The first path (shown) proceeds first through helix unwinding/denaturation (1/4 to 1/2 turn) which acts to clear a path between the first pyrimidine base (●) and the floor of the major groove not previously present (see inset). Triplex formation would then occur by an underwinding kink toward the purine base. A second alternative (not shown) would be for nucleation to occur by a 90° tilting of the helix toward the purine strand at the top of the helix (here the 5' end).

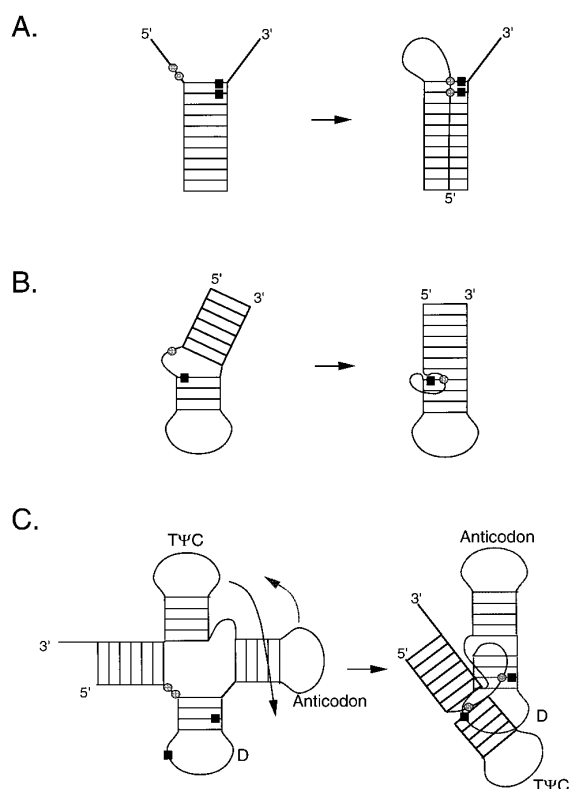


Figure 7. Comparison of the folding of the Y5 triplex, HIV TAR RNA, and tRNA^{Phe}. In the formation of Y5, the unpaired bases at the 5' end of the helix (●) fold into the major groove to interact with duplex purine residues (■) on the floor of the major groove. B, In the HIV TAR RNA, addition of argininamide causes U23 (●) at the 5' end of a three-base bulge to fold into the major groove and interact with G26 (■), which lies at the 5' end of a helix. C, In tRNA, folding of the cloverleaf into the three-dimensional structure is accompanied by two unpaired bases at the 5' end of the D stem (●) interacting with the floor of its major groove and the D loop (■).

Acknowledgements

The authors thank Jon Lapham for help in modeling the ends of the helix and Professor Pui Shing Ho of Oregon State University for help in calculating accessible surface areas of model duplexes. This work was supported by grant GM 21966 from the National Institutes of Health.

References

Antao, V. P., Lai, S. Y. & Tinoco, I. J. (1991). A thermodynamic study of unusually stable RNA and DNA hairpins. *Nucl. Acids Res.* **19**(21), 5901-5905.

Booher, M. A., Wang, S. & Kool, E. T. (1994). Base-pairing and steric interactions between pyrimidine strand bridging loops and the purine strand in DNA pyrimidine-purine-pyrimidine triplexes. *Biochemistry*, **33**, 4645-4651.

Breslauer, K. J., Frank, R., Blöcker, H. & Marky, L. (1986). Predicting DNA duplex stability from the base sequence. *Proc. Natl Acad. Sci. USA*, **83**, 3746-3750.

Chastain, M. & Tinoco, I. J. (1992). A base-triple structural domain in RNA. *Biochemistry*, **31**, 12733-12741.

Chastain, M. & Tinoco, I. J. (1993). Nucleoside triples from the group I intron. *Biochemistry*, **32**, 14220-14228.

Cheng, Y. K. & Pettitt, M. B. (1992). Stabilities of double- and triple helical nucleic acids. *Prog. Biophys. Mol. Biol.* **58**, 225-257.

Christophe, D., Cabrer, B., Bacolla, A., Targovnik, H. V. P. & Vassart, G. (1985). An unusually long poly-(purine)-poly(pyrimidine) sequence is located upstream from the human thyroglobulin gene. *Nucl. Acids Res.* **13**(14), 5127-5144.

Connolly, M. L. (1983). Solvent-accessible surfaces of proteins and nucleic acids. *Science*, **221**, 709-713.

Craig, M. E., Crothers, D. M. & Doty, P. (1971). Relaxation kinetics of dimer formation by self-complementary oligoribonucleotides. *J. Mol. Biol.* **62**, 383-401.

Frank-Kamenetskii, M. D. (1990). Protonated DNA structures. In *Nucleic Acids and Molecular Biology* (Eckstein, F. & Lilley, D. M. J. eds), vol. 4, pp. 1-8, Springer-Verlag, Berlin.

Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T. & Turner, D. H. (1986). Improved free-energy parameters for predictions of RNA duplex stability. *Proc. Natl Acad. Sci. USA*, **83**, 9373-9377.

Green, R. & Szostak, J. W. (1994). *In vitro* genetic analysis of the hinge region between helical elements P5-P4-P6 and P7-P3-P8 in the sun Y group I self-splicing intron. *J. Mol. Biol.* **235**, 140-155.

Groebe, D. R. & Uhlenbeck, O. C. (1988). Characterization of RNA hairpin loop stability. *Nucl. Acids Res.* **16**, 11725-11735.

Gutell, R. R., Noller, H. F. & Woese, C. R. (1986). Higher order structure in ribosomal RNA. *EMBO J.* **5**, 1111-1113.

Haasnoot, C. A. G., Hilbers, C. W., van der Marel, G. A., van Boom, J. H., Singh, U. C., Pattabiraman, N. & Kollman, P. A. (1986). On loopfolding in nucleic acid hairpin-type structures. *J. Biomol. Struct. Dynam.* **3**, 843-857.

Häner, R. & Dervan, P. B. (1990). Single-strand DNA triple-helix formation. *Biochemistry*, **29**, 9761-9765.

Hanvey, J. C., Shimizu, M. & Wells, R. D. (1988). Intramolecular DNA triplexes in supercoiled plasmids. *Proc. Natl Acad. Sci. USA*, **85**, 6292-6296.

Harvey, S. C., Luo, J. & Lavery, R. (1988). DNA stem-loop structures in oligopurine-oligopyrimidine triplexes. *Nucl. Acids Res.* **16**(24), 11795-11809.

Htun, H. & Dahlberg, J. E. (1988). Single strands, triple strands, and kinks in H-DNA. *Science*, **241**, 1791-1796.

Htun, H. & Dahlberg, J. E. (1989). Topology and formation of triple-stranded H-DNA. *Science*, **243**, 1571-1576.

Htun, H. & Dahlberg, J. E. (1990). The structure and formation of triple-stranded H-DNA. In *Structure and Methods: DNA and RNA* (Sarma, R. H. & Sarma, M. H., eds), vol. 3, pp. 185-205, Adenine Press, New York.

Johnston, B. H. (1988). The S1-sensitive form of d(C-T)_nd(A-G)_n: chemical evidence for a three-stranded structure in plasmids. *Science*, **241**, 1800-1804.

Kagawa, T. F., Stoddard, D., Zhou, G. & Ho, P. S. (1989). Quantitative analysis of DNA secondary structure from solvent-accessible surfaces: the B- to Z-DNA transition as a model. *Biochemistry*, **28**, 6642-6651.

- Kang, S. & Wells, R. D. (1992). Central non-Pur-Pyr sequences in oligo(dG-dC) tracts and metal ions influence the formation of intramolecular DNA triplex isomers. *J. Biol. Chem.* **267**, 20887-20891.
- Kang, S., Wohlrab, F. & Wells, R. D. (1992a). GC-rich flanking tracts decrease the kinetics of intramolecular DNA triplex formation. *J. Biol. Chem.* **267**, 19435-19442.
- Kang, S., Wohlrab, F. & Wells, R. D. (1992b). Metal ions cause the isomerization of certain intramolecular DNA triplex isomers. *J. Biol. Chem.* **267**, 1259-1264.
- Lyamichev, V. I., Mirkin, S. M. & Frank-Kamenetskii, M. D. (1985). A pH-dependent structural transition in the homopurine-homopyrimidine tract in superhelical DNA. *J. Biomol. Struct. Dynam.* **3**(2), 327-338.
- Lyamichev, V. I., Mirkin, S. M. & Frank-Kamenetskii, M. D. (1986). Structures of homopurine-homopyrimidine tract in superhelical DNA. *J. Biomol. Struct. Dynam.* **3**(1), 667-669.
- Macaya, R., Wang, E., Schultze, P. & Feigon, J. (1992). Proton nuclear magnetic resonance assignments and structural characterization of an intramolecular DNA triplex. *J. Mol. Biol.* **225**, 755-773.
- Maher, L. J. I., Dervan, P. B. & Wold, B. J. (1990). Kinetic analysis of oligodeoxyribonucleotide-directed triple-helix formation on DNA. *Biochemistry*, **29**, 8820-8826.
- Maher, L. J. I., Wold, B. & Dervan, P. B. (1991). Oligonucleotide-directed DNA triple helix formation: an approach to artificial repressors? *Anti-sense Res. Devel.* **1**(3), 227-281.
- Michel, F. & Westhoff, E. (1990). Modeling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J. Mol. Biol.* **216**, 585-610.
- Michel, F., Jacquier, A. & Dujon, B. (1982). Comparison of fungal mitochondrial introns reveals extensive homologies in RNA secondary structure. *Biochimie*, **64**, 867-881.
- Michel, F., Ellington, A. D., Couture, S. & Szostak, J. W. (1990). Phylogenetic and genetic evidence for base-triples in the catalytic domain of group I introns. *Nature*, **347**, 578-580.
- Mirkin, S. M., Lyamichev, V. I., Drushlyak, K. N., Dobrynin, V. N., Filippov, S. A. & Frank-Kamenetskii, M. D. (1987). DNA H form requires a homopurine-homopyrimidine mirror repeat. *Nature*, **330**, 495-497.
- Nguyen, T. T. & Hélène, C. (1993). Stereospecific detection and modification of double helix DNA by oligonucleotides. *Angew. Chem. Int. Ed. Engl.* **105**, 666-690.
- Noller, H. F., Kop, J., Wheaton, V., Brosius, J., Gutell, R., Kopylov, A. M., Dohme, F. & Herr, W. (1981). Secondary structure model for 23S ribosomal RNA. *Nucl. Acids Res.* **9**, 6167-6189.
- Noller, H. F., Stern, S., Moazed, D., Powers, T., Svenson, P. & Changchien, L. (1988). Studies on the architecture and function of 16S rRNA. *Cold Spring Harbor Symp. Quant. Biol.* **211**, 897-906.
- Papanicolaou, C., Gouy, M. & Ninio, J. (1984). An energy model that predicts the correct folding of both the tRNA and the 5S RNA molecules. *Nucl. Acids Res.* **12**, 31-44.
- Parniewski, P., Galazaka, G., Wilk, A. & Klysik, J. (1989). Complex structural behavior of oligopurine-oligopyrimidine sequence cloned within the supercoiled plasmid. *Nucl. Acids Res.* **17**, 617-629.
- Pörschke, D. & Eigen, M. (1971). Co-operative non-enzymatic base recognition. III. Kinetics of the helix-coil transition of the oligoribouridylic-oligoriboadenylic acid alone at acidic pH. *J. Mol. Biol.* **62**, 361-381.
- Prakash, G. & Kool, E. T. (1992). Structural effects in the recognition of DNA by circular oligonucleotides. *J. Am. Chem. Soc.* **114**, 3523-3527.
- Puglisi, J. D., Tan, R., Calnan, B. J., Frankel, A. D. & Williamson, J. R. (1992). Conformation of the TAR RNA-arginine complex by NMR spectroscopy. *Science*, **257**, 76-80.
- Radhakrishnan, I., Patel, D. J., Priestley, E. S., Nash, H. M. & Dervan, P. B. (1993). NMR structural studies on a nonnatural deoxyribonucleoside which mediates recognition of GC base pairs in pyrimidine-purine-pyrimidine DNA triplexes. *Biochemistry*, **32**, 11228-11234.
- Rich, A. & RajBhandary, U. L. (1976). Transfer RNA: molecular structure, sequence, and properties. *Annu. Rev. Biochem.* **45**, 805-860.
- Riesner, D. & Romer, R. (1973). Thermodynamics and kinetics of conformational transitions in oligonucleotides and tRNA. In *Physio-chemical Properties of Nucleic Acids* (Duchesne, J., ed.), vol. 2, pp. 237-318, Academic Press, New York.
- Roberts, R. W. (1993). Physical chemistry of nucleic acid triple helices. PhD thesis, Yale University, CT.
- Roberts, R. W. & Crothers, D. M. (1991). Specificity and stringency in DNA triplex formation. *Proc. Natl Acad. Sci. USA*, **88**, 9397-9401.
- Rougée, M., Faucon, B., Mergny, J. L., Barcelo, F., Giovannangeli, C., Garestier, T. & Hélène, C. (1992). Kinetics and thermodynamics of triple-helix formation: effects of ionic strength and mismatches. *Biochemistry*, **31**, 9269-9278.
- SantaLucia, J. J., Kierzek, R. & Turner, D. H. (1992). Context dependence of hydrogen bond free energy revealed by substitutions in an RNA hairpin. *Science*, **256**, 217-219.
- Shimizu, M., Hanvey, J. C. & Wells, R. D. (1990). Multiple non-B-DNA conformations of polypurine-polypyrimidine sequences in plasmids. *Biochemistry*, **29**, 4704-4713.
- Shimizu, M., Kubo, K., Matsumoto, U. & Shindo, H. (1993). The loop sequence plays crucial roles for the isomerization of intramolecular DNA triplexes in supercoiled plasmids. *J. Mol. Biol.* **235**, 185-197.
- Shore, D., Langowski, J. & Baldwin, R. L. (1981). DNA flexibility studied by covalent closure of short fragments into circles. *Proc. Natl Acad. Sci. USA*, **78**, 4833-4837.
- Sklenar, V. & Feigon, J. (1990). Formation of a stable triplex from a single DNA strand. *Nature*, **345**, 836-838.
- Voloshin, O. N., Mirkin, S. M., Lyamichev, V. I., Belotserkovskii, B. P. & Frank-Kamenetskii, M. D. (1988). Chemical probing of homopurine-homopyrimidine mirror repeats in supercoiled DNA. *Nature*, **333**, 475-476.
- Wang, S., Booher, M. A. & Kool, E. T. (1994). Stabilities of nucleotide loops bridging the pyrimidine strands in DNA pyrimidine-purine-pyrimidine triplexes: special stability of the CTTTG loop. *Biochemistry*, **33**, 4639-4644.

Weeks, K. M. & Crothers, D. M. (1993). Major groove accessibility of RNA. *Science*, **261**, 1574–1577.

Xodo, L. E., Manzini, G. & Quadrifoglio, F. (1990).

Spectroscopic and calorimetric investigation on the DNA triplex formed by d(CTCTTCTTTCTTTCTT-TCTTCTC) and d(GAGAAGAAAGA) at acidic pH. *Nucl. Acids Res.* **18(12)**, 3557–3564.

Edited by P. E. Wright

(Received 9 April 1996; accepted 18 April 1996)