Topology and Formation of Triple-Stranded H-DNA

HAN HTUN AND JAMES E. DAHLBERG

Repeating copolymers of $(dT\text{-}dC)_n\cdot(dA\text{-}dG)_n$ sequences ($TC\cdot AG_n$) can assume a hinged DNA structure (H-DNA) which is composed of triple-stranded and single-stranded regions. A model for the formation of H-DNA is proposed, based on two-dimensional gel electrophoretic analysis of DNA's with different lengths of $(TC\cdot AG)_n$ copolymers. In this model, H-DNA formation is initiated at a small denaturation bubble in the interior of the copolymer, which allows the duplexes on either side to rotate slightly and to fold back, in order to make the first base triplet. This nucleation establishes which of several nonequivalent H-DNA conformations is to be assumed by any DNA molecule, thereby trapping each molecule in one of several metastable conformers that are not freely interconvertible. Subsequently, the acceptor region spools up single-stranded polypyrimidines as they are released by progressive denaturation of the donor region; both the spooling and the denaturation result in relaxation of negative supercoils in the rest of the DNA molecule. From the model, it can be predicted that the levels of supercoiling of the DNA determine which half of the $(dT\text{-}dC)_n$ repeat is to become the donated third strand.

WHEN EXPOSED TO NEGATIVE SUPERCOILING OR LOW *p*H, $(dT\text{-}dC)_n\cdot(dA\text{-}dG)_n$ repeats [referred to as $(TC\cdot AG)_n$] can assume an unusual structure (*1–18*) called H-DNA. Such sequences are present in the genomes of both vertebrates and invertebrates (*1, 2, 9*). Although several models have been proposed for the structure of H-DNA (*1–8*), the one that is supported by enzymatic (*1–5, 9–12*), chemical (*4, 5, 12–16*), and hybridization and protection (*16*) data postulates that this conformation contains both single-stranded and triple-stranded regions. In this model, half of the Watson-Crick duplex is disrupted and the released polypyrimidine strand is folded back, enabling it to progress down the major groove of the other half of the repeat (Fig. 1A). This third strand (which we call the donated strand) associates with purines in the helical half of the repeat (the acceptor region) by Hoogsteen base pairs, forming dT-dA·dT and dC-dG·dC^+ base triplets. The structure introduces a flexible kink into DNA, which serves as a hinge between the flanking duplexes (*16*). Originally the structure was called H-DNA (*3*), primarily because of its requirement for protons (*1–4, 7, 8, 10*), but the name is also appropriate for hinged DNA (*16*).

Either the 3′ or the 5′ half of the polypyrimidine strand could be used as the donated strand in H-DNA (*16*) (Fig. 1B). We refer to the conformer in which the 3′ half of the polypyrimidine repeat is the donated region as H-y3 DNA and we call the other isomer, in which the 5′ half of the pyrimidine strand is donated, the H-y5 DNA conformer. Although both H-y3 and H-y5 conformers appear equivalent when drawn in two dimensions (Fig. 1B), a strong unexplained bias toward the H-y3 conformer has been observed (*4, 11–19*). We show here that the two forms are not equivalent, since more negative supercoils are relaxed by formation of H-y3 than H-y5 DNA. However, in spite of their nonequivalence topologically and hence energetically, the two conformers can coexist in noninterconvertible, metastable forms.

We propose a model for H-DNA formation in which nucleation of the Hoogsteen base pairing near the middle of the repeat determines the ultimate structure. Subsequently, the two duplexes flanking the triplex undergo mutual rotations in which the donor region produces an unpaired polypurine strand and polypyrimidines that are spooled onto the acceptor duplex, thereby extending the triplex. This model predicts that the direction of rotation of the flanking duplexes just prior to nucleation dictates whether the 5′ or the 3′ region of the pyrimidine repeat is to be the donor; moreover, the degree of supercoiling influences the sense of this rotation. Thus, $(TC\cdot AG)_n$ repeats in highly supercoiled DNA should primarily form H-y3 DNA, and repeats in less supercoiled DNA should also form H-y5 DNA. This prediction is borne out experimentally, so that we can explain the bias toward the H-y3 conformer and thus establish the validity of our model.

Relaxation of negative supercoils during H-DNA formation. Two-dimensional gel electrophoresis of a population of topoisomers of DNA from two different U1 RNA genes containing $(TC\cdot AG)_n$ repeats (Fig. 2) showed that H-DNA formation caused relaxation of negative supercoils. Molecules containing a locally underwound structure (in the first dimension) had less writhe and hence migrated more slowly than normal. The number of negative supercoils relaxed as a result of the new conformation was revealed as the difference in linking numbers (determined after electrophoresis in the second dimension) of two topoisomers that have the same mobility when one of them contains this structure (*20*). In this context, the alternative structure itself is assumed to have no significant effect on the overall mobility of the plasmid in the porous gels used here.

Because $(TC\cdot AG)_5$ is too short to form H-DNA at normal levels of negative supercoiling (*2, 21*), any transition in mobility must derive from structures elsewhere in the plasmid. Topoisomers of plasmids containing $TC\cdot AG)_5$, with a linking difference of -13 or more (relative to the most relaxed molecule in the first dimension) have an anomolously slow mobility in the first dimension (Fig. 2A). This transition results from formation of a cruciform at a 39-bp (base pair) inverted repeat in pBR322 sequences (*22, 23*), so that the plasmid migrates at the rate of a molecule containing 3.6 to 3.9 fewer negative supercoils (Fig. 2A).

The authors are in the Department of Physiological Chemistry, University of Wisconsin–Madison, Madison, WI 53706.

A population of DNA topoisomers of plasmid DNA containing $(TC{\cdot}AG)_{15}$ rather than $(TC{\cdot}AG)_5$ exhibits an additional mobility (and hence structural) transition when the linking difference of the molecule is −8 (Fig. 2B); formation of this structure in $(TC{\cdot}AG)_{15}$ relaxes 3.4 negative supercoils. As a result of this relaxation, cruciform formation in these molecules requires a linking difference of −16 rather than −13; thus the mobility change in molecules with a linking difference of −8 is due to unwinding of the double helix (*3, 4*) rather than the simple bending of the DNA, as proposed by others (*24*). This new mobility transition is due to formation of H-DNA since nuclease S1 cleaves molecules in the TC·AG repeat, but only when they have an altered mobility (and not just a high level of negative supercoiling) (*21*).

The length of the TC·AG repeat influences the two-dimensional gel pattern of the topoisomers in several ways. Longer repeats support the generation of multiple conformers, and they require less

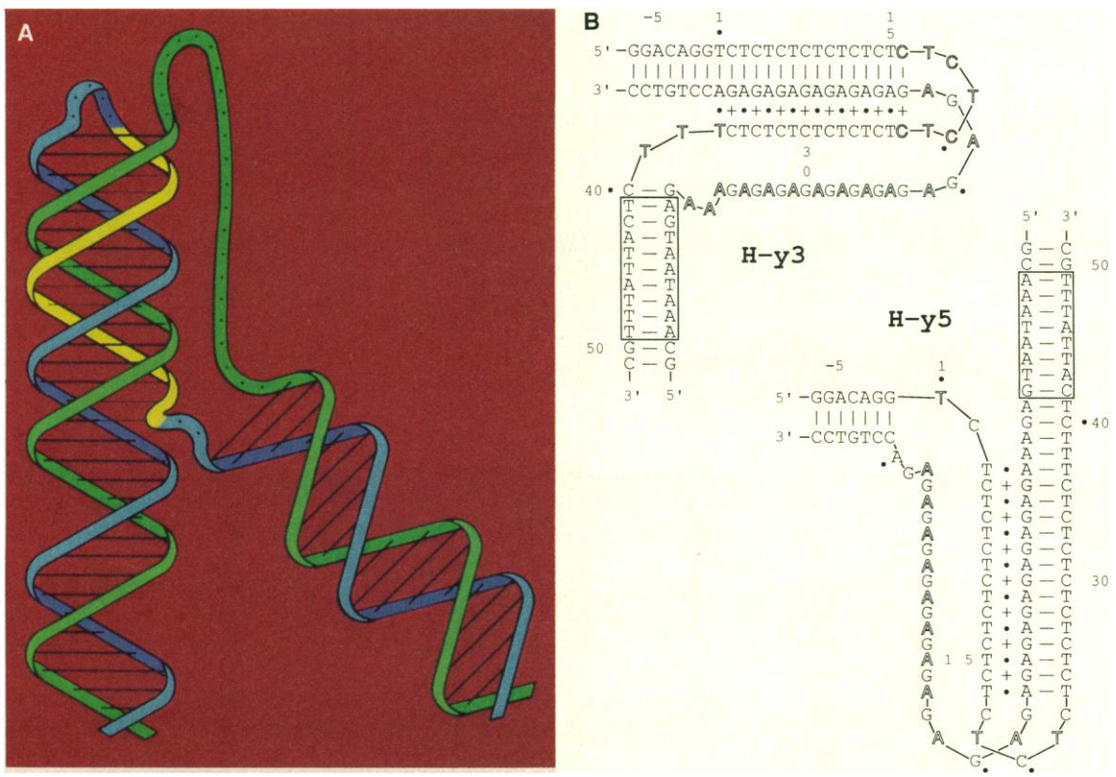


Fig. 1. Structure of H-DNA. (**A**) Three-dimensional projection of H-DNA in $(TC{\cdot}AG)_{18}$, with the 3′ half of the polypyrimidine (dT-dC) repeat donated to the triplex, forming the H-y3 conformer. The 5′ half of this repeat, plus the complementary 3′ half of the $(dA\text{-}dG)_n$, polypurine repeat, act as the acceptor helix in this conformation. The two halves of the polypyrimidine strand in the triplex (yellow and blue) are antiparallel (*38*). Watson-Crick and Hoogsteen base pairs are represented by thick and thin lines, respectively. (**B**) Alternative use of the 3′ or 5′ half of the (dT-dC) repeat as the donated strand, to form H-y3 or H-y5 conformers of H-DNA. Watson-Crick base pairs (in the flanking and acceptor helices) are shown as lines and Hoogsteen base pairs between the acceptor purines and uncharged [T] or protonated [C^+] pyrimines are shown by · and +, respectively. Residue numbers and the box denoting a flanking region sequence are shown. Large letters indicate nucleotides that are reactive to diethyl pyrocarbonate (A), osmium tetroxide (T), or methoxylamine (C) when the DNA is in the H conformation (*16*) (Fig. 6).

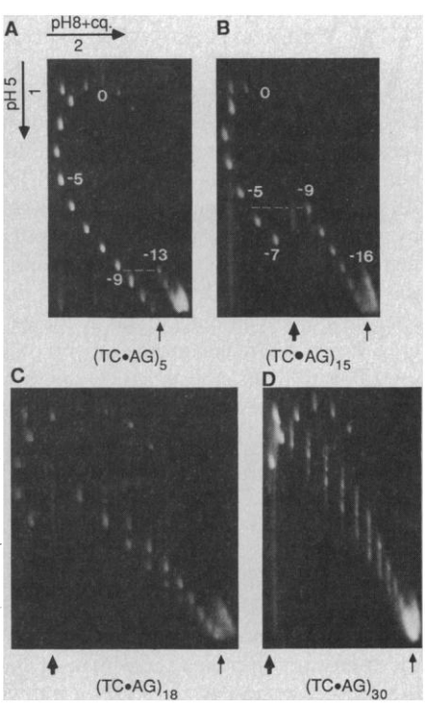


Fig. 2. Two-dimensional gel electrophoretic analysis of $(TC{\cdot}AG)_n$–containing topoisomers that differ in their repeat length and degree of negative supercoiling. (**A**) Molecules containing $(TC{\cdot}AG)_5$, which is unable to form H-DNA at normal levels of negative supercoiling (*2, 21*). (**B**) Molecules containing $(TC{\cdot}AG)_{15}$, which forms H-DNA when subjected to a moderate degree of negative supercoiling. (**C**) Molecules containing $(TC{\cdot}AG)_{18}$. (**D**) Molecules containing $(TC{\cdot}AG)_{30}$ (but with a C·G → T·A transition at the 25th dinucleotide repeat). All of these copolymers were derived from a region 1.8 kb downstream of human U1 RNA genes (*2*). In the first dimension (*p*H 5), there is enough negative supercoiling to support the formation of alternative structures that partially relax the supercoiling in the rest of the molecule; this relaxation is manifested as a reduction in mobility, resulting in a discontinuity in the pattern of spots (arrows). The absence of conditions that would support alternative structures in the second dimension provides a basis for determining the linking difference of each topoisomer by counting the number of forms separating that spot from topoisomer "O." The magnitude of relaxation due to H-DNA formation is given by the relative linking differences of two topoisomers with the same mobility in the first dimension when one of them contains the alternative structure (dashed lines) (*20*). The structural transition due to formation of H-DNA is noted by a thick arrow and the transition due to formation of a cruciform at a 39-bp repeat (*21, 22*) in the vector is noted by a thin arrow. Mixtures of DNA topoisomers were separated in 1.5 percent agarose gel by electrophoresis at *p*H 5 (top to bottom), which readily permits H-DNA formation in negatively supercoiled DNA's (*16*) (20 m*M* acetic acid and 1 m*M* magnesium acetate adjusted to *p*H 5.0 with tris-base at 13° to 17°C), followed by electrophoresis at *p*H 8 (left to right) which does not support H-DNA (*16, 21*) (40 m*M* tris-base, 25 m*M* sodium acetate, 0.9 m*M* sodium salt of EDTA, and chloroquine diphosphate (cq) (0.5 μg/ml), *p*H 8.0, at 14° to 18°C).

negative supercoiling to promote formation of H-DNA (Fig. 2, C and D). Furthermore, the maximal amount of negative supercoiling that can be relaxed by H-DNA formation increases with longer repeats (Fig. 3).

The presence of multiple topoisomers at any one linking difference (arrayed in vertical lines of spots in Fig. 2, C and D) indicates the existence of multiple H-DNA conformers, each relaxing a characteristic number of negative supercoils. The discrete nature of the spots (Fig. 2C) suggests that the conformers are not interconvertible but are trapped in local energy minima. In support of this concept is the changed pattern of spots when the DNA is subjected to various reversible treatments before electrophoresis (*21*). Thus each spot contains a single H-DNA conformer that is trapped in a local energy minimum. As we discuss below, the conformation assumed by each DNA molecule is determined by the kinetics of H-DNA nucleation, but the stabilities of the conformers are determined by their length, with longer structures being more stable than shorter ones (*21*).

The multiple metastable H-DNA conformers result from differences in the alignment and orientation of the triplex within the $(TC{\cdot}AG)_n$ repeat (Fig. 2, C and D). Longer repeats are more likely than short ones to allow extension of an H-DNA structure to a length that is stable, when the structure is located away from the middle of the repeat. Thus, it is not surprising that, as the repeat length increases, increasing numbers of stable conformers are observed. Also, the greater stability of these longer structures (*21*) reduces the level of negative supercoiling needed to support them; that is shown by the shift in transitions to lower levels of supercoiling with longer repeats.

The strands of H-DNA are locally unlinked. The extent to which various lengths of $(TC{\cdot}AG)_n$ repeats could relax negative supercoils is summarized in Fig. 3. To allow direct comparison of molecules with multiple H-DNA forms, we present the maximum extent of supercoil relaxation, measured at linking differences only slightly greater than those needed to achieve H-DNA formation. For repeats longer than $(TC{\cdot}AG)_{15}$, the data fall on a straight line, the slope of which shows that one negative superhelical turn in plasmid DNA is relieved for every 11 nucleotides of $(TC{\cdot}AG)_n$ repeat that can be converted from duplex to H-DNA conformation. Therefore, the entire $(TC{\cdot}AG)_n$ repeat, and not just the donor half, becomes topologically unlinked. We demonstrated earlier (*16*) that this would occur as the donated strand progressed down the major groove of the acceptor half of the repeat; being folded back, the polypyrimidine strand would effectively unwind itself from around the polypurine strand in the acceptor. The result is a paranemic knot with no net winding over the length of the triplex.

Extrapolation of the line in Fig. 3 to a repeat length of zero shows that incorporation of repeat sequences into the triplex does not account for all of the relaxation of negative supercoiling. Another component, equivalent to about 0.7 turn, is introduced by "end" effects at the tip (where the single-stranded pyrimidines fold back onto the acceptor helix) and at the hinge region (the flexible kink) present at the base of the structure (*16*). Our model-building studies (*21*) show that about one half turn can be accommodated by the single-stranded nucleotides at the base, which may account for the displacement of the data for $(TC{\cdot}AG)_9$ and $(TC{\cdot}AG)_{11}$ above the line in Fig. 3. Also, this single-stranded region can be increased from 2 nucleotides to 12 by high levels of negative supercoiling (*16*).

A model for H-DNA formation. The existence of multiple metastable conformers of H-DNA demonstrated that the pathway to formation of H-DNA is critical. That led us to consider how the donated pyrimidine strand could become wound around the acceptor helix. We propose that the key event in H-DNA formation is the generation of the first base triplet (nucleation) in the interior of the repeat, to create the tip of the triplex structure [Fig. 4, H-y3(1) or H-y5(1)]. Subsequently, the triplex is propagated outward toward the ends of the repeat, with a minimal number of unpaired nucleotides at any stage (Fig. 4, lower structures). This is accomplished by turning the donor and acceptor helices in concert, in the same plane but in opposite directions. This counterrotation of the two helices (curved arrows in Fig. 4) allows the acceptor helix to spool up pyrimidines as they are released from the donor helix. Because both of these rotations reduce the negative supercoiling in the molecule, two turns are relaxed for every 10 to 11 pyrimidines that are transferred from Watson-Crick base pairs in the donor duplex to Hoogsteen base pairs in the triplex. In the process, the single-stranded polypurines form a stacked loop on one side of the triplex, rather than winding around it. The rotations continue until one or the other duplex runs out of TC·AG repeats. The turning of the two helices, as one of them progresses down the other, resembles a combination of a rack-and-pinion gear and a worm gear.

For clarity, the kinks in all structures in Fig. 4 have arbitrarily been drawn at a constant 45°, although this angle is likely to vary as a result of flexibility of the single strands at the base of the triplex (*16*). Also, because the displaced polypurine chain acts as a tether between the tip of the triplex and the end of the donor duplex, it could influence the angle between these two helices. In particular, when the triplex is turned so that the polypurine strand emerges from it on the side away from the donor duplex [as in H-y3(12) or H-y5(12)], the tether would tend to pull the donor helix up, toward the tip of the triplex. We expect this effect to be more pronounced with shorter H-DNA structures, which have relatively short polypurine chains. Furthermore, ionic conditions that affect the rigidity of the stacked purines in this strand could change the effective length of this tether; the accessibility of these purines to DEPC (diethyl pyrocarbonate) is altered by the presence of Mg^{2+} [compare figures 2 and 4 from (*16*) with Fig. 6 (see below)] showing that the single-stranded polypurine chain can assume several conformations.

Prenucleation events at the tip of the triplex affect the formation of H-DNA (Fig. 5). The generation of a small denaturation bubble of six nucleotides would unwind the duplex by about one half turn (Fig. 5, A and B). The two duplexes must *again* be rotated by about one half turn relative to each other to place the pyrimidines that are released at the end of the donor duplex into the major groove of the acceptor duplex (Fig. 5C). Simply folding the denatured pyrimidine chain directly back on itself would place the donated strand about half of one turn away from the point where it must be to form a Hoogsteen base pair with the acceptor purine (*25, 26*).

The prenucleation rotation that orients the ends of the two duplexes could occur in either a positive or a negative sense. With negative rotation (Fig. 5C, left), formation of the tip of H-DNA contributes approximately one full turn to the overall negative supercoil relaxation; about one half turn results from denaturation of the 6 bp at the tip, and another half from the rotation itself. In contrast, with positive rotation, formation of the tip has little or no net effect on the overall amount of negative supercoiling; in this case, the half turn resulting from denaturation is offset by the rotation in the opposite direction. Thus the structures on the left halves of Fig. 5, C and D, are topologically distinct from those on the right, with the former relaxing one more negative supercoil than the latter.

Upon orientation of the ends of the flanking helices and kinking of the DNA so that one helix meets the other, the first Hoogsteen base pair is formed, an event we call nucleation of the H-DNA triplex (Fig. 5D). It is possible to draw four nucleation structures by donating either the 5′ or the 3′ side of the polypyrimidine strand to the major groove of DNA that had undergone either negative or positive prenucleation rotation. Two of those structures would force

the single-stranded polypurine chain to pass through the loop made by the polypyrimidine chain as it folds back onto the acceptor helix (the two central structures in Fig. 5D). However, this local linking would be energetically very costly and hence unlikely to occur. Thus, to avoid local linking of the strands at the tip, the 3′ half of the polypyrimidine chain must be the donor in DNA that has undergone negative rotation; the result is the H-y3 conformer. Likewise the 5′ half must be used after positive rotation, resulting in the H-y5 conformer. Because these two conformers arise from intermediates that differ by one negative supercoil (Fig. 5C), they should be separable by gel electrophoresis.

Influence of supercoiling on H-y3 and H-y5 conformer formation. The degree of supercoiling of the DNA should determine whether the H-y3 or H-y5 conformers are favored. Because nucleation of the H-y5 conformer requires no net rotation, it can proceed simply by disrupting 6 bp as the duplex kinks back on itself, with minimal disruption of base stacking. Thus in DNA's that are relaxed or show little negative supercoiling, the H-y5 conformer is more easily nucleated. In contrast, H-y3 conformers should prevail when negative supercoiling is moderately high, which would promote denaturation and full (negative) rotation of the two abutting helices.

We tested these predictions directly by determining which half of the polypurine repeat in $(TC{\cdot}AG)_{18}$ was single-stranded at high and low levels of negative supercoiling. The 3′ half of the polypurine strand was accessible to diethyl pyrocarbonate, and hence single-stranded, in most of the molecules with low levels of negative supercoiling (Fig. 6, lane 1). Thus, the complementary 5′ half of the polypyrimidine strand was the third strand in the triplex, confirming the prediction that small amounts of negative supercoiling would promote the formation of H-y5 conformers. When more highly negatively supercoiled plasmids were assayed (Fig. 6, lane 2), the reactive region switched to the other end of the repeat (the H-y3 DNA), as predicted.

This shift from H-y5 or H-y3 conformers with increasing levels of negative supercoiling coincides with a progressive increase in the proportion of topoisomers that are in the slower migrating, more

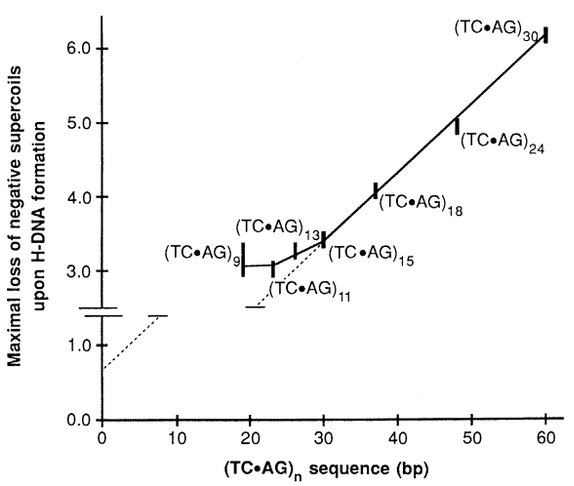


Fig. 3. Negative supercoil relaxation as a function of repeat length. The maximum number of negative supercoils relaxed by H-DNA formation was determined for TC·AG repeats of various lengths by two-dimensional gel electrophoresis (Fig. 2). Relaxation was determined when supercoiling was slightly in excess of that required for H-DNA formation (dashed lines in Fig. 2B). The slope of the line between $(TC{\cdot}AG)_{15}$ and $(TC{\cdot}AG)_{30}$ shows that one supercoil is relaxed for every 11 bp of repeat. The deviation of shorter repeats from this line—$(TC{\cdot}AG)_9$ through $(TC{\cdot}AG)_{13}$—probably results from flexibility of the hinge at the base of the triplex (*16*) under conditions of higher negative supercoiling.

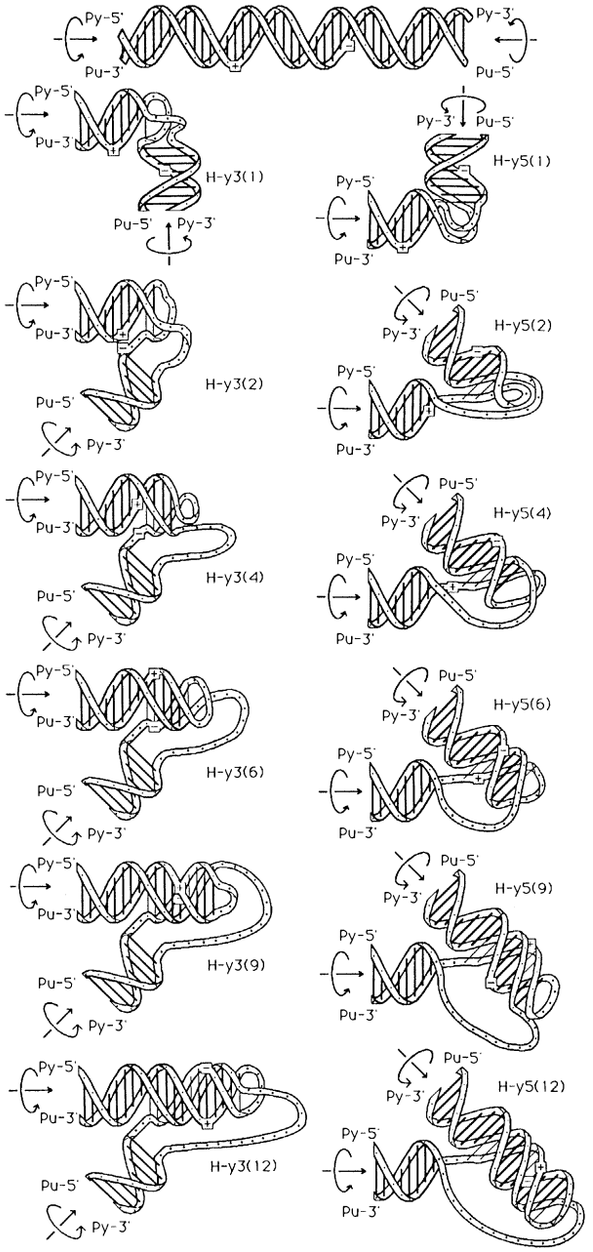


Fig. 4. A model for the formation of H-DNA. Left and right sides, the formation of H-y3 and H-y5 conformers, respectively; parentheses, the lengths of the triplex; curved arrows, direction of rotation that relaxes negative supercoiling; straight arrows, the direction that nucleotides move as they are recruited into the H-DNA structure during its elongation; thick lines, Watson-Crick base pairs; thin lines, Hoogsteen pairs; plus (+) and a minus (−) located 13 nucleotides apart on the phosphodiester backbone of the polypyrimidine strand illustrate the progression of nucleotides from the flanking region duplex into the body of the triplex. The nucleotides in the single-stranded polypurine loop are stacked in pairs along one side of the triplex (*2, 5*). All helices have arbitrarily been drawn with 11 bp or triplets per turn. H-DNA formation is initiated by breathing within the middle of the TC·AG repeat, allowing the duplex to kink back on itself. Nucleation occurs by alignment of a pyrimidine at the end of one duplex (the donor) in the major groove of the other (the acceptor), allowing formation of the Hoogsteen pair at the tip of the triplex. Elongation occurs when the two helices rotate coordinately, allowing the acceptor helix (shown above) to spool up pyrimidines as they are released by denaturation of the donor helix (below); both rotations are propelled by the force associated with negative supercoiling. The structure is thereby elongated until either the donor or the acceptor helix no longer has nucleotides that can be incorporated into the triplex. The single-stranded nucleotides at the end of the donor helix (*16*) provide for a small amount of supercoiling relaxation in the absence of further denaturation.

relaxed form in the first dimension (Fig. 2C). Thus, as expected, one more negative supercoil is relaxed when molecules contain H-y3 DNA than H-y5 DNA of comparable length. These two conformers are therefore topologically, and energetically, nonequivalent.

The influence of the level of negative supercoiling on the distribution of H-y3 and Hy5 conformers can be explained only if local linking at the tip (Fig. 5D, central two structures) were not allowed; otherwise, rotation in either direction prior to nucleation would give rise to both conformers. The absence of such linking supports the direction of H-DNA formation proposed in Fig. 4, because elongation toward the tip would have resulted in local linking at the tip.

The above results and earlier ones (*16*) make it unlikely that several alternative structures for H-DNA that have been considered by others are correct. The lack of chemical reactivity at the base of the acceptor helix, regardless of whether the DNA is in the H-y3 (*4, 12–19*) or H-y5 (*16*) conformation, is strong evidence against any discontinuity between the triplex and its adjacent duplex. Thus the triplex is likely to be a right-handed helix, resembling an A-like right-handed duplex with an extra strand, in agreement with the results of fiber-diffraction studies on the triplex dT_n-dA_n·dT_n (*27*). Our topological arguments (Fig. 5) demonstrate that the use of a 3′ donor relaxes more negative supercoils in a right-handed helix than does use of a 5′ donor, when the third strand is in the major groove

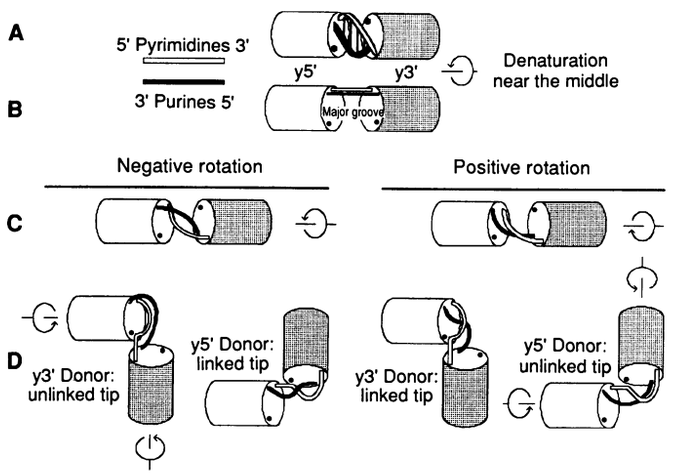
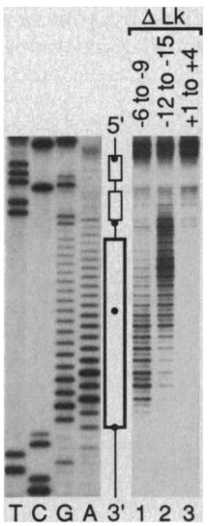


Fig. 5. The role of prenucleation rotation in the determination of donor and acceptor polarities in H-DNA. The two flanking duplexes are represented by cylinders and the major groove by a spot. Prior to nucleation (Fig. 4), the two helices (**A**) undergo two half turn rotations: denaturation of nucleotides to make the loop and first donated pyrimidine (**B**) and orientation of a donor pyrimidine opposite an acceptor major groove (**C**). The two flanking duplexes containing the 5′ and 3′ halves of the polypyrimidine repeat are represented by white and gray cylinders, respectively. For clarity, only one of the two flanking duplexes is rotated. While the rotation associated with denaturation occurs only in the negative sense (relaxing negative supercoils in the DNA) (B), rotation to reorient the pyrimidines at the ends of the duplexes opposite the major groove of the other duplexes can be in either the negative or positive sense (C). The sense (direction in space) of this latter rotation determines whether formation of the nucleation structure will be topologically significant ($-1/2$ plus $-1/2 \approx -1$ turn) or neutral ($-1/2$ plus $+1/2 \approx 0$ turn). Furthermore, it dictates which half of the polypyrimidine is to be donated as the third strand. To avoid local linking of strands at the tip (where the polypurine would be forced to pass through the polypyrimidine loop), only two of the four possible structures drawn in (**D**) can be used to nucleate H-DNA formation: negative prenucleation rotation results in use of the 3′ region of the polypyrimidine as donor whereas positive prenucleation rotation results in use of the 5′ region. Extensive negative supercoiling would be expected to foster negative prerotation (and give rise to H-y3 DNA), whereas low *p*H or little negative supercoiling, which allow the DNA duplex to kink directly back on itself, would support positive rotation and give rise to H-y5 DNA.

Fig. 6. Effect of the amount of supercoiling on the use of 3′ or 5′ donors in H-DNA. The distribution of $(TC{\cdot}AG)_{18}$–containing topoisomers in the H-y3 or H-y5 conformation was monitored by reaction with diethyl pyrocarbontate under conditions in which only single-stranded purines are reactive (*16*). Three different populations of topoisomers were probed, with average linking differences (ΔLk) of −7 or −8 (range to −6 to −9), −13 or −14 (range −12 to −15) and +2 or +3 (range +1 to +4) (assayed as Fig. 2). The locations of adducts on the single-stranded polypurines were mapped (*16*) and compared to the sequence of the repeat (heavy box) and flanking region (light boxes) (*16*) (Fig. 1B). The lanes on the left orient the reactive nucleotides within the sequence. 5′ and 3′ show the polarity of the polypurine strand in the repeat. Reactivity of the 3′ region of the polypurines at low levels of negative supercoiling (lane 1) and of the 5′ region of the polypurines at high levels (lane 2) indicates that the 5′ and 3′ halves of the polypyrimidine strands are the donated regions, respectively.



of the acceptor helix. Therefore, our data lend strong support for the structure of H-DNA presented in Fig. 1A. The binding of Z-DNA–specific antibodies to plasmids containing TC·AG repeats (*4*) must result either from recognition of other structures in the DNA, lack of specificity of the antibodies for only Z-DNA, or recognition of an additional structure induced by the antibody itself.

Our model for H-DNA formation can be generalized to triplex structures derived from other sequences regardless of whether the donated strand is polypyrimidine (H-y) or polypurine (H-r). The repeating polymer dC_n·dG_n has been shown to assume either H-y3 or H-r3 conformations when negatively supercoiled, making it very likely that these structures resemble the ones formed by the TC·AG repeats studied here. Because of differences in the location of the donated strand in the triplexes of H-y and H-r structures (*19*), we expect that H-r3 DNA would relax slightly less (about one-fourth of a turn) negative supercoiling than does H-y3 DNA (*28*). Also, we predict that when there is little negative supercoiling this repeat should assume H-y5 and H-r5 conformations.

The multiple conformations at any one linking difference (Fig. 2D) demonstrate that multiple nucleation events can result in a population of metastable H-DNA's. Some of these forms differ in being H-y3 or H-y5 conformers of comparable length (Fig. 2C), whereas others also differ in the amount of repeat that is incorporated into H-DNA (Fig. 2D). Although we have been aware of these metastable forms for several years (*29*), we were unable to explain their existence adequately until we developed the model described above. While this model explains these processes qualitatively, it is not yet known how various conditions affect the energetics of nucleation and elongation.

The existence of H-DNA in vivo is still a matter of conjecture (*30, 31*); repeats of TC·AG long enough to hybridize to $(dT\text{-}dC)_{18}$ are present in human DNA about once every 140 to 150 kb (*21, 32*), raising the possibility that H-DNA may be widespread in genomic DNA. Sequences capable of forming H-DNA (*33, 34*) are present in regions that function in transcription, replication, or recombination (*16, 21, 32, 35*). As noted earlier, the kink and single or triple strands of H-DNA could all influence DNA-protein interactions (*16*). Also, because H-DNA formation relaxes negative supercoiling, it could act as a buffer for relaxation of surges in negative supercoiling such as those generated by transcription (*36, 37*). As a consequence, transcription may activate DNA functions that require an H-DNA structure. Finally, depending on transient conditions at the moment of nucleation, the same DNA sequence may assume either an active

or inactive conformation because of the nonequivalence of the H-y3 and H-y5 conformations.

REFERENCES AND NOTES

1. C. C. Hentschel, *Nature* **295**, 714 (1982).
2. H. Htun, E. Lund, J. E. Dahlberg, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 7288 (1984).
3. V. I. Lyamichev, S. M. Mirkin, M. D. Frank-Kamenetskii, *J. Biomol. Struct. Dyn.* **3**, 327 (1985).
4. D. E. Pulleyblank, D. B. Haniford, A. R. Morgan, *Cell* **42**, 271 (1985).
5. T. Evans and A. Efstratiadis, *J. Biol. Chem.* **261**, 14771 (1986).
6. G. Yagil and J. L. Sussman, *EMBO J.* **5**, 1719 (1986).
7. J. S. Lee, M. L. Woodsworth, L. P. Latimer, A. R. Morgan, *Nucleic Acids Res.* **12**, 6603 (1984).
8. V. I. Lyamichev, S. M. Mirkin, M. D. Frank-Kamenetskii, *J. Biomol. Struct. Dyn.* **3**, 667 (1986).
9. G. C. Glikin, G. Gargiulo, L. Rena-Descalzi, A. Worcel, *Nature* **303**, 770 (1983).
10. H. Htun, E. Lund, G. Westin, U. Pettersson, J. E. Dahlberg, *EMBO J.* **4**, 1839 (1985).
11. J. B. Margot and R. C. Hardison, *J. Mol. Biol.* **184**, 195 (1985).
12. J. C. Hanvey, J. Klysik, R. D. Wells, *J. Biol. Chem.* **263**, 7386 (1988).
13. M. Vojtiskova *et al.*, *FEBS Lett.* **234**, 295 (1988).
14. J. C. Hanvey, M. Shimizu, R. D. Wells, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6292 (1988).
15. B. H. Johnston, *Science* **241**, 1800 (1988).
16. H. Htun and J. E. Dahlberg, *ibid.*, p. 1791.
17. D. A. Collier, J. A. Griffin, R. D. Wells, *J. Biol. Chem.* **263**, 7397 (1988).
18. O. N. Voloshin, S. M. Mirkin, V. I. Lyamichev, B. P. Belotserkovskii, M. D. Frank-Kamenetskii, *Nature* **333**, 475 (1988).
19. Y. Kohwi and T. Kohwi-Shigematsu, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 3781 (1988).
20. J. C. Wang, L. J. Peck, K. Becherer, *Cold Spring Harbor Symp. Quant. Biol.* **47**, 85 (1982).
21. H. Htun and J. E. Dahlberg, unpublished results.
22. D. M. J. Lilley, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 6468 (1980).
23. C. K. Singleton, *J. Biol. Chem.* **258**, 7661 (1983).
24. F. Wohlrab, M. J. McLean, R. D. Wells, *ibid.* **262**, 6407 (1987).
25. J. S. Lee, D. A. Johnson, A. R. Morgan, *Nucleic Acids Res.* **6**, 3073 (1979).
26. K. Hoogsteen, *Acta Cryst.* **16**, 907 (1963).
27. S. Arnott and E. Selsing, *J. Mol. Biol.* **88**, 509 (1974).
28. As we demonstrated above (Figs. 3 and 4) and earlier (*16*), the strands in H-DNA–like structures are locally unlinked so the total number of relaxed supercoils is determined by the number of base pairs per turn and the length of the duplex used, rather than the number of base triplets per turn of the resulting triplex. Furthermore, the size of the loop is irrelevant to this relaxation if the nucleotides are to be in either the loop or the triplex. In contrast, both the orientation of the donated strand within the base triplet and the sense of rotation prior to nucleation contribute to a baseline of end effects. An additional component due to flexibility of the hinge at the base (*16*) may also contribute to the relaxation of negative supercoils, particularly in highly underwound molecules.
29. Two-dimensional gel analyses of H-DNA, in support of the structure shown in Fig. 1A, have been informally presented: H. Htun and J. Dahlberg, "Biological Effects of DNA Topology," Cold Spring Harbor, NY (September 1986); "Unusual DNA Structures," Gulf Shores Symposium, Gulf Shores, AL (April 1987); and "Biomolecular Stereodynamics," Fifth Conversation, Albany, NY (June 1987).
30. J. S. Lee, G. D. Burkholder, L. J. P. Latimer, B. L. Haug, R. P. Braun, *Nucleic Acids Res.* **15**, 1047 (1987).
31. B. Hoffman-Liebermann, D. Liebermann, A. Troutt, L. H. Kedes, S. N. Cohen, *Mol. Cell. Biol.* 6, 3632 (1986).
32. H. Manor, B. Sridhara Rao, R. G. Martin, *J. Mol. Evol.* **27**, 96 (1988).
33. J. H. Miller and H. M. Sobell, *Proc. Natl. Acad. Sci. U.S.A.* **55**, 1201 (1966).
34. S. M. Mirkin *et al.*, *Nature* **330**, 495 (1987).
35. R. D. Wells, D. A. Collier, J. C. Hanvey, M. Shimizu, F. Wohlrab, *FASEB J.* **2**, 2939 (1988).
36. C. R. Cantor and A. Efstratiadis, *Nucleic Acids Res.* **12**, 8059 (1984).
37. G. N. Giaever and J. C. Wang, *Cell* **55**, 849 (1988); Y.-P. Tsao, H.-Y. Wu, L. F. Liu, *ibid.* **56**, 111 (1989).
38. E. Moser and P. Dervan, *Science* **238**, 645 (1987); D. Praseuth *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 1349 (1988).

39. Supported by grants from NSF (PCM 8309618) and NIH (GM 30220). We thank P. Tregloan for typing the manuscript and D. Horowitz, F. Johnston, E. Lund, G. Q. Pennabble, and J. H. White for useful discussions.

28 December 1988; accepted 12 February 1989

A Multiubiquitin Chain Is Confined to Specific Lysine in a Targeted Short-Lived Protein

VINCENT CHAU, JOHN W. TOBIAS, ANDREAS BACHMAIR,* DAVID MARRIOTT, DAVID J. ECKER,† DAVID K. GONDA, ALEXANDER VARSHAVSKY

The ubiquitin-dependent degradation of a test protein β-galactosidase (βgal) is preceded by ubiquitination of βgal. The many (from 1 to more than 20) ubiquitin moieties attached to a molecule of βgal occur as an ordered chain of branched ubiquitin-ubiquitin conjugates in which the carboxyl-terminal Gly^{76} of one ubiquitin is joined to the internal Lys^{48} of an adjacent ubiquitin. This multiubiquitin chain is linked to one of two specific Lys residues in βgal. These same Lys residues have been identified by molecular genetic analysis as components of the amino-terminal degradation signal in βgal. The experiments with ubiquitin mutated at its Lys^{48} residue indicate that the multiubiquitin chain in a targeted protein is essential for the degradation of the protein.

UBIQUITIN, A 76-RESIDUE PROTEIN, IS PRESENT IN EUkaryotes either free or covalently joined, through its carboxyl-terminal glycine residue, to various cytoplasmic, nuclear, and integral membrane proteins (*1*). The coupling of ubiquitin to other proteins is catalyzed by a family of ubiquitin-conjugating enzymes (also called E2 enzymes) (*1, 2*). In the yeast *Saccharomyces cerevisiae*, two of the approximately six such enzymes present have been identified as products of the genes *RAD6*, whose

V. Chau and D. Marriott are in the Department of Pharmacology, Wayne State University School of Medicine, Detroit, MI 48201. J. W. Tobias, A. Bachmair, D. K. Gonda, and A. Varshavsky are in the Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139. D. J. Ecker is in the Department of Molecular Pharmacology, Smith, Kline, and French Laboratories, Philadelphia, PA 19101.

*Present address: Max-Planck-Institut für Züchtungsforschung, Abteilung Genetische Grundlagen der Pflanzenzüchtung, Egelspfad, D-5000 Köln 30, Federal Republic of Germany.

†Present address: ISIS Pharmaceutical, Carlsbad, CA 92008.