

transforming proteins of adenovirus and SV40 with the RB1 gene product (25). These results suggest that mutations in RB1 may be associated with initiation or progression of the malignant phenotype in many tumors. The germline mutation of RB1 identified in constitutional cells reported here, confirms the initiating role of RB1 in RB tumors.

#### REFERENCES AND NOTES

1. W. K. Cavenee *et al.*, *Nature* **305**, 779 (1983).
2. A. G. Knudson, *Proc. Natl. Acad. Sci. U.S.A.* **68**, 820 (1971).
3. M. F. Hansen *et al.*, *ibid.* **82**, 6216 (1985); A. Koufos *et al.*, *Nature* **309**, 5964 (1984); A. Koufos *et al.*, *ibid.* **316**, 330 (1985); C. Lundberg, L. Skoog, W. K. Cavenee, M. Nordenskjold, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 2372 (1987); E. R. Fearon, A. P. Feinberg, S. H. Hamilton, B. Vogelstein, *Nature* **318**, 377 (1985); I. U. Ali, R. Lidereau, C. Theillet, R. Callahan, *Science* **238**, 185 (1987); E. Solomon *et al.*, *Nature* **328**, 616 (1987); B. R. Seizinger *et al.*, *Science* **236**, 317 (1987); C. G. P. Mathew *et al.*, *Nature* **328**, 527 (1987); K. Kok *et al.*, *ibid.* **330**, 578 (1987).
4. S. H. Friend *et al.*, *Nature* **323**, 643 (1986).
5. A. Goddard *et al.*, *Mol. Cell. Biol.* **8**, 2082 (1988).
6. M. M. Myers, Z. Larin, T. Maniatis, *Science* **230**, 1242 (1985).
7. E. Winter, F. Yamamoto, C. Almoguera, M. Perucho, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7575 (1985); R. A. Gibbs and C. T. Caskey, *Science* **236**, 303 (1987); K. Forrester *et al.*, *Nature* **327**, 298 (1987).
8. W.-H. Lee *et al.*, *Science* **235**, 1394 (1987).
9. W.-H. Lee *et al.*, *Nature* **329**, 642 (1987).
10. L. Winger *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 4484 (1983).
11. J. J. Wiggs *et al.*, *N. Engl. J. Med.* **318**, 151 (1988).
12. A. D. Goddard, unpublished data.
13. X. Zhu, unpublished data.
14. J. M. Dunn *et al.*, unpublished data.
15. R. Breathnach and P. Chambon, *Annu. Rev. Biochem.* **50**, 349 (1981).
16. R. Treisman *et al.*, *Nature* **302**, 591 (1983).
17. A. G. DiLella, J. Marvit, A. S. Lidsky, F. Güntler, S. L. Woo, *ibid.* **322**, 799 (1986).
18. D. J. Rees, C. R. Rizza, G. G. Brownlee, *ibid.* **316**, 643 (1985).
19. E. Arpaia *et al.*, *ibid.* **333**, 85 (1988).
20. T. P. Dryja, personal communication.
21. W. K. Cavenee *et al.*, *Science* **228**, 501 (1985).
22. T. Moteji, *Hum. Genet.* **58**, 168 (1981).
23. E. Y.-H. P. Lee *et al.*, *Science* **241**, 218 (1988).
24. J. W. Harbour *et al.*, *ibid.*, p. 353.
25. P. Whyte *et al.*, *Nature* **334**, 124 (1988); X. DeCaprio *et al.*, *Cell*, in press.
26. S. Henikoff, *Gene* **28**, 351 (1984).
27. F. Sanger, S. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463 (1977).
28. J. M. Chirgwin, A. E. Przybyla, R. J. MacDonald, W. J. Rutter, *Biochemistry* **18**, 5294 (1979).
29. We thank T. P. Dryja for generously providing the p4.7R cDNA clone and the 3.8 and 0.9 subcloned fragments. We are grateful for the encouragement of the entire Toronto RB Group, especially M. Canton, H. Balakier, and E. Reyes, who grew the cells and isolated the RNA. Supported by a Terry Fox Programme Project Grant of the National Cancer Institute of Canada and by the Medical Research Council of Canada. J.M.D. is a recipient of a Steve Fonyo Studentship Award from the National Cancer Institute of Canada. B.L.G. is a research associate of the Ontario Cancer Foundation.

19 May 1988; accepted 2 August 1988

## The S1-Sensitive Form of d(C-T)<sub>n</sub>d(A-G)<sub>n</sub>: Chemical Evidence for a Three-Stranded Structure in Plasmids

BRIAN H. JOHNSTON

Homopurine-homopyrimidine sequences that flank certain actively transcribed genes are hypersensitive to single strand-specific nucleases such as S1. This has raised the possibility that an unusual structure exists in these regions that might be involved in recognition or regulation. Several of these sequences, including d(C-T)<sub>n</sub>d(A-G)<sub>n</sub>, are known to undergo a transition in plasmids to an unwound state that is hypersensitive to single strand-specific nucleases; this transition occurs under conditions of moderately acid pH and negative supercoiling. Chemical probes were used to examine the reactivity of a restriction fragment from a human U1 gene containing the sequence d(C-T)<sub>18</sub>d(A-G)<sub>18</sub> as a function of supercoiling and pH, and thus analyze the structure in this region. Hyperreactivity was seen in the center and at one end of the (C-T)<sub>n</sub> tract, and continuously from the center to the same end of the (A-G)<sub>n</sub> tract, in the presence of supercoiling and pH ≤ 6.0. These results provide strong support for a triple-helical model recently proposed for these sequences and are inconsistent with other proposed structures.

**A** NUMBER OF ACTIVELY TRANSCRIBED eukaryotic genes in chromatin are hypersensitive to nucleases that normally only attack single-stranded DNA, such as S1 (1-3). The S1-sensitive sites have been mapped to homopurine-homopyrimidine sequences, and many are found in promoter regions at the 5' ends of the genes (2, 3), suggesting that they may play a role in transcriptional control.

One homopurine-homopyrimidine sequence common in eukaryotic DNA is the alternating sequence d(C-T)<sub>n</sub>d(G-A)<sub>n</sub>, henceforth designated (C-T)<sub>n</sub>. S1-hypersensitive (C-T)<sub>n</sub> sequences have been found in genes encoding sea urchin histones (4), *Drosophila* heat-shock proteins (5), and human U1 (6) and U2 RNAs (7), for example. In actively transcribed members of the human

U1 gene family, the sequence (C-T)<sub>18</sub> is located 1.8 kb downstream of the RNA-coding region (6). When cloned into plasmids, this sequence has been shown to undergo a transition to an S1-sensitive state with a concomitant loss of up to four supercoils (8, 9). Thus the altered structure is topologically equivalent to a melted, unwound region or a cruciform, contributing no net helical winding to the plasmid DNA. The transition requires somewhat acid pH and is aided by negative supercoiling, although the need for supercoiling diminishes with increasing length of the (C-T)<sub>n</sub> sequence (8). Several models have been suggested for the S1-sensitive, unwound conformation of (C-T)<sub>n</sub> sequences. To help distinguish among these models, the chemical reactivity of the altered conformation has

been examined using reagents which were previously shown to be sensitive probes for non-B-DNA structures (10). The results of these experiments support a recent model (11, 12) in which the normal double helix reapportions itself into triple-stranded and single-stranded regions. This conclusion is reinforced by recent experiments on other homopurine sequences (13-16) and by experiments in which the same sequence was analyzed (17) with a somewhat different technique (18). The present results further suggest that the length of the triple helix in the altered conformation may vary with superhelical density.

The chemical probes used include diethyl pyrocarbonate (DEP), which reacts preferentially with purines in Z-DNA that are in the usual *syn* conformation (10, 19); dimethylsulfate (DMS), which is hyperreactive with purines in the unusual *anti* conformation in Z-DNA (10); hydroxylamine, which reacts specifically with cytosines at junctions between B- and Z-DNA (10) and at junctions between out-of-phase blocks of Z-DNA (20); and osmium tetroxide (OsO<sub>4</sub>), which reacts preferentially with thymines at B-Z junctions (10, 21). DEP (22), hydroxylamine (23), and OsO<sub>4</sub> (24) also react with cruciform loops. DEP favors structures, including single-stranded regions or Z-DNA, in which the reactive N<sup>6</sup> and N-7 positions of adenine and the N-7 position of guanine are especially accessible. The pyrimidine-specific reagents hydroxylamine and OsO<sub>4</sub>, on the other hand, appear to recognize single-stranded regions or dislocations in the helix that allow the reagents to attack

Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139.

their respective target bases from out of the plane of the pyrimidine ring (10). These four reagents, together with the hydroxylamine derivative methoxyamine, were used to chemically modify pGEM2-SB-OF, a plasmid that contains a restriction fragment from a human U1 gene bearing the (C-T)<sub>18</sub> sequence. After the reaction, sites of chemical modification were identified as previously described (10). Briefly, a restriction fragment from the modified plasmid containing the sequence of interest was prepared with a radioactive label at one end, the backbone was cleaved at modified sites by treatment

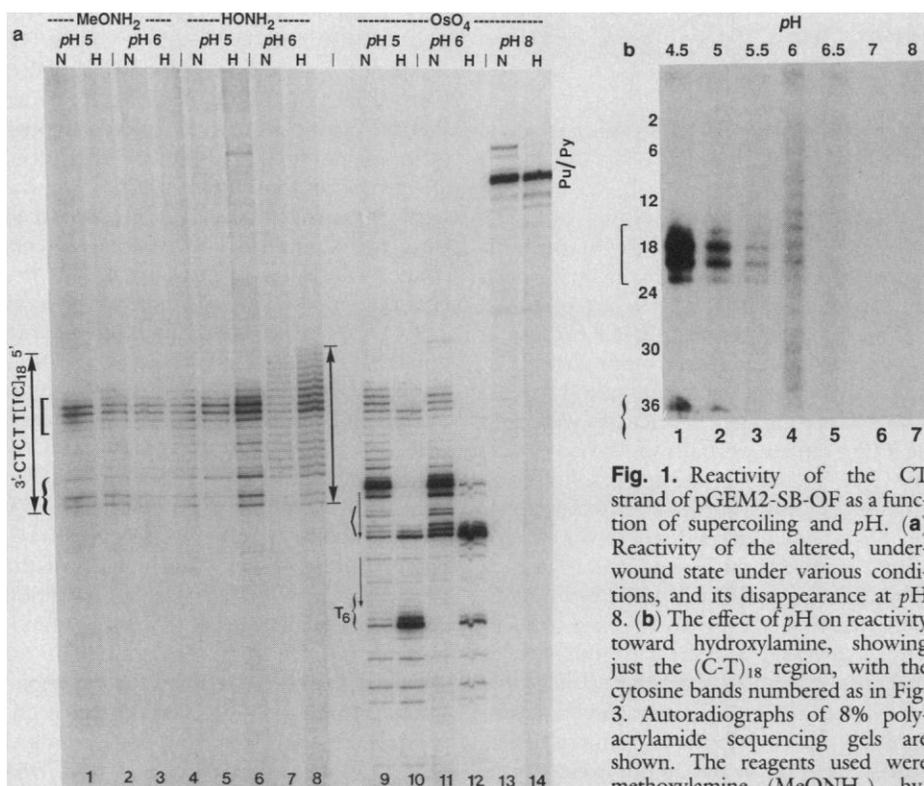
with piperidine, and the resulting fragments were resolved on a sequencing gel.

Autoradiographs of gels indicating the reactivity of bases on the 3' end-labeled CT-containing strand and the 5' end-labeled AG-containing strand are shown in Figs. 1 and 2. The sequence is shown in Fig. 3, with base pairs numbered from the 5' end of the (C-T)<sub>n</sub> sequence. The double-headed arrows in Figs. 1a and 2 extend from nucleotides (nt) 1 to 42, covering the entire homopyrimidine sequence. Lanes 1 to 8 of Fig. 1a show the reactivity patterns exhibited by the altered form toward methoxyamine and hy-

droxylamine for two pH levels (5.0 and 6.0) and two levels of negative supercoiling. These levels were the "native" level (N) characteristic of plasmids isolated from *Escherichia coli*, with specific linking difference  $\sigma \cong -0.05$ , and a much higher level (H) of  $\sigma \cong -0.12$ . Under all these conditions, cytosines in the center of the CT block (square bracket; nt 16 to 22) and at its 3' end (heavy curved bracket; nt 36, 40, and 42) were strongly reactive. For hydroxylamine, the hyperreactivity of these two regions declined as the pH was increased above 4.5; virtually no reactivity was detected above pH 6 at the native level of supercoiling (Fig. 1b). In the case of hydroxylamine at pH 6 (Fig. 1a, lanes 7 and 8, and Fig. 1b, lane 4), a moderate hyperreactivity throughout the CT block appeared to be superimposed on the bimodal distribution of reactivity seen at lower pH.

When OsO<sub>4</sub> was used to probe thymines, reactivity was again seen in the center and the 3' end of the CT block at the native level of supercoiling and pH 5 or 6 (Fig. 1a, lanes 9 to 12). Additional hyperreactivity was seen just to the 3' side of the CT block, in an ATTATTT sequence (nt 44 to 49; Fig. 1a, lanes 9 and 11, angled bracket). However, at the higher superhelical density and pH 5, the peak of hyperreactivity in the center of the CT block (square bracket) was shifted slightly but reproducibly in the 3' direction, and the hyperreactivity at the 3' end of the CT block (nt 35 to 41) was greatly reduced (Fig. 1a, compare lanes 9 and 10). In addition, the 3' end of the adjacent AT sequence (nt 48 and 49) as well as the 3' end of a nearby T<sub>6</sub> sequence (nt 63 and 64) became highly reactive (Fig. 1a, lanes 10 and 12). [At pH 6 and high superhelicity (Fig. 1a, lane 12) the reduced reactivity in the middle of the CT tract was variable and may be partly due to heavy reaction at nt 48 and 49, which would lead to depletion of longer fragments.] At pH 8 all the hyperreactivities seen at the lower pHs were abolished, and new hyperreactivity was seen at an alternating purine-pyrimidine sequence located 76 bp to the 5' side of the CT region (Fig. 1a, lanes 13 and 14, Pu/Py). Apparently the high pH disallows the unwinding transition at the CT sequence and the unrestrained helical tension helps to convert the alternating purine-pyrimidine sequence to left-handed Z-DNA, which is reactive with OsO<sub>4</sub> at its junctions with B-DNA (10).

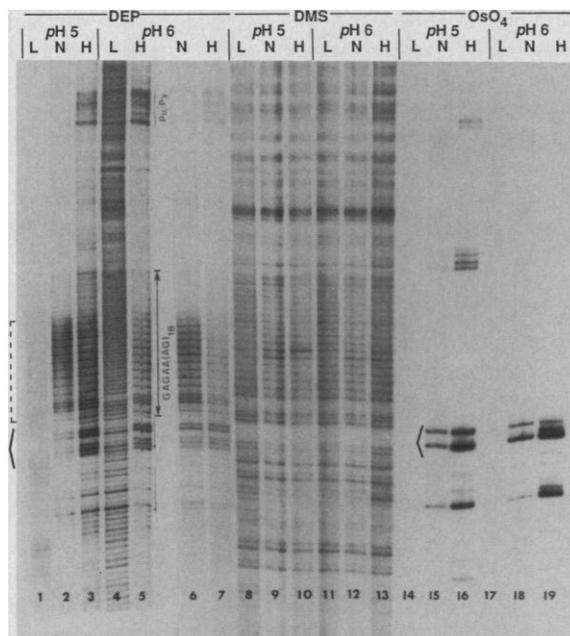
In Fig. 2, the AG-containing strand is 5' end-labeled so that sequences toward the bottom in Figs. 1 and 2 are toward the right in Fig. 3. The reaction of supercoiled, low pH samples with DEP showed hyperreactivity at adenines in the entire 5' (right-hand) half of the (A-G)<sub>n</sub> sequence (Fig. 2,



**Fig. 1.** Reactivity of the CT strand of pGEM2-SB-OF as a function of supercoiling and pH. (a) Reactivity of the altered, underwound state under various conditions, and its disappearance at pH 8. (b) The effect of pH on reactivity toward hydroxylamine, showing just the (C-T)<sub>18</sub> region, with the cytosine bands numbered as in Fig. 3. Autoradiographs of 8% polyacrylamide sequencing gels are shown. The reagents used were methoxyamine (MeONH<sub>2</sub>), hydroxylamine (HONH<sub>2</sub>), and osmium tetroxide (OsO<sub>4</sub>).

um tetroxide (OsO<sub>4</sub>). Superhelical density is indicated by N (native; characteristic of plasmids isolated from *E. coli*) and H (high;  $\sigma \cong -0.12$ ). The double-headed arrow exactly spans the homopyrimidine sequence (nt 1 to 42), the single-headed arrows indicate the short AT-rich direct repeats, and Pu/Py indicates an alternating purine/pyrimidine sequence. Hyperreactive regions in the middle and at the lower (3') end of the (C-T)<sub>18</sub> sequence are designated by heavy square and curved brackets, respectively. The adjacent hyperreactive AT sequence is indicated by angled brackets. Preparation of supercoiled DNA, modification reactions, and subsequent enzymatic and electrophoretic steps were as described (17) except as follows: all reaction mixtures were preincubated in the appropriate buffer plus 1 mM EDTA for 15 min at 37°C prior to the addition of the reagent. The reaction buffers, titrated to the indicated pHs, were: 50 mM sodium acetate for pH 4.5 to 5.5; 50 mM sodium cacodylate for pH 6 to 7, and 20 mM EDTA for pH 8. DEP reactions were stopped by the addition of NaCl to 0.2M and precipitation by ethanol, solubilization of the pellet in tris-EDTA (pH 7.4), and a second ethanol precipitation. Osmium tetroxide reactions were initiated by adding pyridine (titrated to the appropriate pH with HCl) to 4% final concentration, followed by aqueous OsO<sub>4</sub> to 1%. OsO<sub>4</sub> reactions were terminated by three extractions with carbon tetrachloride at 0° followed by one ethanol precipitation. For the hydroxylamine pH series in (b), hydroxylamine was titrated to the various pHs with diethylamine. The active form of the reagent is presumably the free base, which is present at higher levels for higher pH values. To compensate for this, the time of incubation with hydroxylamine was made shorter for higher pH values as follows: pH 4.5, 35 min; pH 5, 30 min; pH 5.5, 20 min; pH 6 and 6.5, 15 min; and pH 7 and 8, 12 min. Under these conditions, a darker exposure showed that corresponding bands outside the homopyrimidine region were of nearly equal intensity. To the extent that there were variations in intensity, they followed the pattern pH 6 > 6.5  $\cong$  8 > 5.5 > 5  $\cong$  4.5 > 7. Thus for the four lowest pH values, homopyrimidine-specific reactivity drops quickly while background reactivity is slightly increasing.

**Fig. 2.** Reactivity of the AG strand of pGEM2-SB-OF as a function of supercoiling and pH. Symbols are the same as in Fig. 1 with the addition of: DEP, diethyl pyrocarbonate; DMS, dimethyl sulfate; and L, linear (not supercoiled). The region of the homopurine sequence that is hyperreactive in the altered conformation is designated by the dashed square bracket, and the AT sequence just below the AG repeat is designated by the angled bracket.



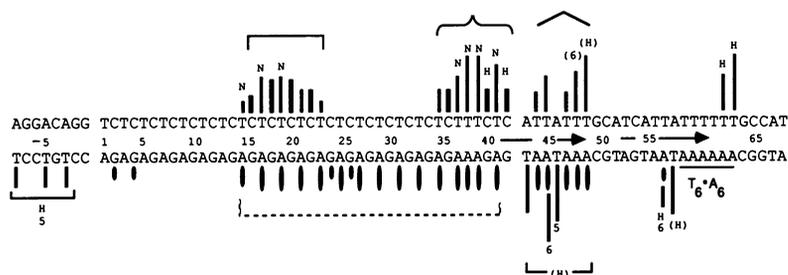
lanes 2, 3, and 5 to 7; dashed square bracket; Fig. 3, open cigar-shaped symbols on nt 15 to 41), in marked contrast to the localized, bimodal reactivity seen on the CT strand. However, in both strands it was the right half of the sequence shown in Fig. 3 that was mainly reactive. In the AG strand as well as in the CT strand, the AT sequence (Fig. 2, nt 44 to 49; angled bracket) was hyperreactive when supercoiled. This was particularly true at high superhelicity for both DEP reaction at adenines (Fig. 2, lanes 3, 5, and 7) and OsO<sub>4</sub> reaction at thymines (Fig. 2, lanes 16 and 19). At high superhelicity some OsO<sub>4</sub> reactivity was also seen for thymines on the opposite side of the (A-G)<sub>n</sub> sequence (Fig. 2, lane 16, just above the AG block; nt -3, -5, and -8) and at the 3' end of the A<sub>6</sub> sequence that is complementary to the hyperreactive T<sub>6</sub> sequence (nt 57 and 58; lowest bands in Fig. 2, lanes 16 and 19). [Hyperreactivity in the alternating purine-pyrimidine sequence at high superhelicity (Fig. 2, lanes 3, 5, 16, and 19) indicates that torsional strain can be sufficient to convert this sequence to Z-DNA even in the presence of the altered conformation at the (C-T)<sub>n</sub> sequence.] As expected, the DEP reactivity of linear DNA at pH 6 is characteristic of normal B-DNA, with uniform reaction inside and outside the (A-G)<sub>n</sub> repeat (Fig. 2, lanes 4 and 17), although guanines were slightly more reactive than adenines inside the repeat (Fig. 2, lane 4). Linear DNA at pH 5 appeared to weakly favor the altered conformation, consistent with the observation that this DNA can convert to the underwound state in the absence of torsional strain at low pH (6, 7). This absence of strain evidently allows the

adjacent AT sequence (nt 44 to 49) to remain base-paired, since no hyperreactivity was seen in that region (Fig. 2, lanes 1 and 14).

The transition to the altered conformation results in only rather subtle changes in reactivity of the homopurine strand toward DMS. Some changes in reactivity were seen at four guanines within the (A-G)<sub>n</sub> repeat, nt 1, 3, 24, and 26 (Fig. 3; solid cigar-shaped symbols). These were hyperreactive in the altered form, but the exact pattern varied with pH and superhelical density (Fig. 2, lanes 9, 10, 12, and 13). In addition, densitometer traces of lanes 8 to 13 in Fig. 2 revealed a decreased reactivity of guanines in the 3' (left-hand) half of the (A-G)<sub>n</sub> sequence (dashed square bracket). Ignoring the contributions from the hyperreactive nt 1, 3, 24, and 26, the average ratio of band intensities for the 3' half (nt 6 to

16) relative to the 5' half (nt 24 to 32) of the (A-G)<sub>n</sub> repeat decreased from  $0.85 \pm 0.05$  for the presumed B-DNA structure (Fig. 2, lane 11; linear DNA, pH 6) to  $0.5 \pm 0.05$  for the altered structure (Fig. 2, lanes 9, 10, 12, and 13).

These results are consistent with, and strongly favor, the partially three-stranded structure shown in Fig. 4a. In this model, the right half of the (C-T)<sub>n</sub> repeat has effectively melted, and the pyrimidines of the melted region have folded back into the major groove of the left half of the repeat to form Hoogsteen base pairs with the purines. This blocks the N-7 position of the purines, thus accounting for the partial protection of the left half of the purine repeat from DMS. The reactive bases at the center of the CT strand are assigned to the single-stranded loop. The pyrimidines involved in either Watson-Crick pairing (nt 1 to 15) or Hoogsteen pairing (nt 23 to 34) of the triple-helix are protected, but those which form the loop region (nt 16 to 22) are reactive to all three pyrimidine reagents. [These reagents are normally unreactive toward uniform helical regions (10).] The adenines of the right half of the purine repeat (nt 17 to 41) are unpaired and therefore reactive to DEP. Hyperreactivity extends progressively into the region beyond the triplex (nt 35 to 49) with increasing supercoiling, presumably reflecting the melting of this AT-rich sequence under the unwinding stress of supercoiling. Triplex models of this sort have been proposed recently (11, 12) on the basis of the pH dependence of the unwinding transition and earlier evidence that the linear homopolymer poly(dC-dT)·poly(dA-dG) can associate with poly(dC-dT) to form a triple-stranded structure (25). The degree of unwinding expected from the model, dubbed "H-DNA" by Lyamichev *et al.* (12) was approximately equivalent to that of a melted



**Fig. 3.** Summary of reactivity data. Solid bars represent OsO<sub>4</sub> reactivity and open bars hydroxylamine (except at pH 6; see text for discussion) and methoxylamine reactivity. The height is roughly proportional to the intensity of the reaction as determined from densitometer traces of the autoradiograms in Figs. 1 and 2, except for the tallest solid bars, which underrepresent the true OsO<sub>4</sub> reactivity. Solid and open cigar-shaped symbols represent DMS and DEP hyperreactivity, respectively. H, N, 5, or 6 next to a symbol indicates that hyperreactivity was seen only at high or native superhelical density, or only at pH 5 or 6, respectively. (H) indicates that the intensity shown refers to high superhelical density, and the reactivity at native supercoiling was less than, but still above, background. Only changes induced by supercoiling or acid pH (or both) are shown. The short direct AT-rich repeats are indicated by arrows, and the brackets are as in Figs. 1 and 2.

or a cruciform structure, consistent with two-dimensional gel data (9). The requirement that the cytosines of Hoogsteen C:G base pairs be protonated explained the stabilizing influence of low pH on the altered structure (12). More recently, oligopyrimidines have been shown to associate with complementary homopurine-homopyrimidine duplex regions of restriction fragments (26), and evidence that only mirror-symmetric homopurine sequences can form the "H-DNA" structure, as required by the model, has been presented (27). It now appears that a wide variety of these kinds of sequences can form this structure (13-16).

Under the low salt conditions of the OsO<sub>4</sub> reaction, the structure at pH 5 appears to change in going from native to high superhelical density. The shift of the maximal OsO<sub>4</sub> hyperreactivity in the center of the CT block from nt 19 to 21 (for example, Fig. 1a, compare lanes 9 and 10 at the square bracket) and the loss of hyperreactivity of thymines in the TCTTT sequence at the 3' side of the CT block (nt 35 to 39; same lanes, curved bracket) may be due to the formation of the structure shown in Fig. 4b, in which the length of the triplex region has been extended to include the TCTTT sequence. This allows more unwinding in response to the increased torsional stress, at the cost of forming a mismatched C:G:T base triplet at the middle thymine of the TTT sequence (Fig. 4b).

The key feature of the data that the triplex model explains, and with which other proposed models have difficulty, is the asymmetry in the reactivity patterns seen in the differences between the two halves of the homopurine repeat as well as in the differences between the two strands. According to the triplex model, this arises from the fact that both halves of the pyrimidine strand participate in the triple helix, while only the left half of the purine strand does so. However, two conformers are possible for a triplex structure, depending on which half of the repeat retains Watson-Crick base-pairing (12). If both were present in equal proportions, no asymmetry would be seen in the two halves of the repeat. The asymmetry that is seen indicates that the conformers shown in Fig. 4, in which the 3' half of the homopyrimidine sequence loses its Watson-Crick base pairs, are strongly favored. Only at high superhelical density is there some slight hyperreactivity to the left of the homopurine sequence (especially noticeable in Fig. 2, lane 16, above the AG repeat), indicating that the alternate isomer does exist in a minor proportion. While the AT-rich sequence on the 3' edge of the homopyrimidine sequence may facilitate melting from the 3' side, recent studies of other mirror-

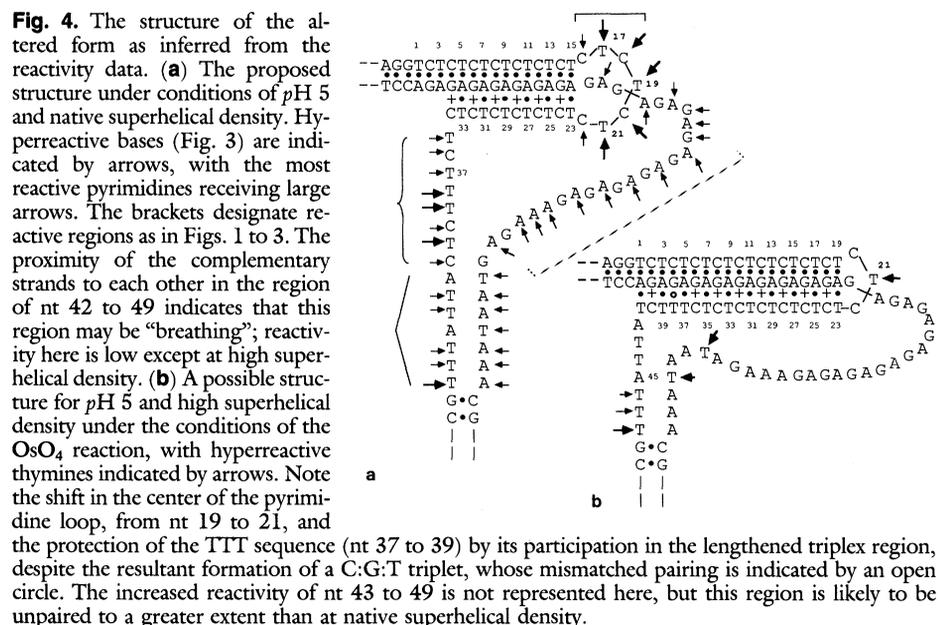
symmetric homopurine sequences indicate that opening of the 3' half of the pyrimidine strand is a common feature (13-16).

Alternate proposals to explain the S1-sensitivity of (C-T)<sub>n</sub> sequences include slippage loops (4), a left-handed duplex (28), a duplex in which bases alternate between Watson-Crick and Hoogsteen base-pairings (29), and a heteronomous duplex (30). Slippage loops should result in a reactive region on each strand (the loop) opposite a largely unreactive sequence on the other strand; the center of the repeat would be paired and hence unreactive. However, the left side of the sequence is unreactive on both strands, and the middle is especially reactive. Thus slippage loops could not account for the reactivity data. [For the same reason, they could not account for the hyperreactivity at the AT sequences (nt 41 to 49 and 53 to 61), which form a pair of short direct repeats (Figs. 1 to 3; single-headed arrows).] The other proposed models are all some sort of duplex structure. A left-handed double helix (28) appears to be ruled out by data on the extent of unwinding accompanying the transition (8, 9). Pulleyblank *et al.* (29) reported that the S1-sensitive conformation of a plasmid containing (T-C)<sub>22</sub>(A-G)<sub>22</sub> partially protects about half of the (A-G)<sub>n</sub> sequence from methylation by DMS in a manner similar to that seen in Fig. 2. However, they interpreted those results in terms of a uniform duplex structure in which all the guanines are Hoogsteen base-paired to the complementary cytosines, which are protonated. Evans and Efstratiadis (30) used a variety of techniques to study a (T-C)<sub>38</sub>(A-G)<sub>38</sub> sequence, which becomes S1-sensitive at low pH when either linear or

supercoiled. They found no protection with DMS and observed hyperreactivity with DEP in the S1-sensitive conformation. Although they described hyperreactivity over the entire insert at pH 4.5 for both linear and supercoiled plasmids, their data (figure 7b) indicate that, for the linear samples, only the adenines in about half of the homopurine sequence are hyperreactive, as has been observed here (Fig. 2). The more uniform hyperreactivity across the repeat they observed for the supercoiled samples might be due to the formation of both potential isomers of triplex structure under the added stress of supercoiling, while the weaker driving force for the transition in linear DNA seeks out the most stable isomer. They also found that alternating CT primers (but not GA primers) annealed to the S1-sensitive conformation could be extended with DNA polymerase, a result entirely consistent with the triplex model, although they favored instead a heteronomous duplex of unknown structure.

Each of these duplex models, and indeed any duplex model, has difficulty explaining how a uniform sequence, (C-T)<sub>n</sub>, in a presumably uniform duplex could result in the strongly nonuniform reactivity that is seen along the sequence. While a heteronomous duplex model in which the two strands have different helical characteristics could account for differences in reactivity between the two strands, the highly nonuniform reactivity along each strand cannot be accounted for in any simple way. Thus all these alternative models are unlikely in view of the reactivity data.

The hyperreactivity that is seen farthest from the CT repeat (Fig. 1a, lanes 10 and



12, and Fig. 2, lanes 16 and 19; nt 57, 58, 63, and 64) occurs mainly at thymines that lie at either end of a (dT)<sub>6</sub>(dA)<sub>6</sub> sequence. Because oligo[(dA)·(dT)] sequences can adopt a non-B structure having a high propeller twist (P-DNA), there is rather poor stacking at junctions between B-DNA and P-DNA regions (31, 32). Hence the helix may "buckle" under high torsional strain at these sites, allowing accessibility to OsO<sub>4</sub>.

While most probes yield similar results at pH 5 and 6, hydroxylamine exhibits at pH 6 a moderate reactivity over most of the homopyrimidine sequence which is superimposed upon the localized hyperreactivity seen at pH 5 (Fig. 1a, lanes 7 and 8, and Fig. 1b, lane 4). Because the hydroxylamine reaction at pH 6 takes place under conditions of high salt (10) and nearly neutral pH, both of which somewhat disfavor the altered structure (6, 7), there may be an equilibrium mixture of B-DNA, the altered conformation, and intermediate structures which are substantially unpaired and therefore reactive over much of their length. It is not clear why this anomalous reactivity at pH is seen with hydroxylamine but not with methoxylamine.

#### REFERENCES AND NOTES

1. A. Larsen and H. Weintraub, *Cell* **29**, 609 (1982).
2. J. M. Nickol and G. Felsenfeld, *ibid.* **35**, 467 (1983).
3. S. C. R. Elgin, *Nature* **209**, 213 (1984).
4. C. C. Hentschel, *ibid.* **295**, 714 (1982).
5. E. Siegfried, G. H. Thomas, U. M. Bond, S. C. R. Elgin, *Nucleic Acids Res.* **14**, 9425 (1986).
6. H. Htun, E. Lund, J. E. Dahlberg, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 7288 (1984).
7. H. Htun, E. Lund, G. Westin, U. Pettersson, J. E. Dahlberg, *EMBO J.* **4**, 1839 (1985).
8. H. Htun and J. E. Dahlberg, personal communication.
9. V. I. Lyamichev, S. M. Mirkin, M. D. Frank-Kamenetskii, *J. Biomol. Struct. Dynam.* **3**, 327 (1985).
10. B. H. Johnston and A. Rich, *Cell* **42**, 713 (1985).
11. D. Christophe *et al.*, *Nucleic Acids Res.* **13**, 5127 (1985).
12. V. I. Lyamichev, *J. Biomol. Struct. Dynam.* **3**, 667 (1986).
13. J. C. Hanvey, J. Klysik, R. D. Wells, *J. Biol. Chem.* **263**, 7386 (1988).
14. D. A. Collier *et al.*, *ibid.*, p. 7397.
15. O. N. Voloshin, S. M. Mirkin, V. I. Lyamichev, B. P. Belotserkovskii, M. D. Frank-Kamenetskii, *Nature* **333**, 475 (1988).
16. Y. Kohwi and T. Kohwi-Shigamatsu, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 3781 (1988).
17. H. Htun and J. E. Dahlberg, *Science* **241**, 1791 (1988).
18. H. Htun, B. H. Johnston, J. E. Dahlberg, in preparation.
19. W. Herr, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 8009 (1985).
20. B. H. Johnston *et al.*, in preparation.
21. K. Nejedly, M. Kwinkowski, G. Galazka, J. Klysik, E. Palecek, *J. Biomol. Struct. Dynam.* **3**, 467 (1985).
22. P. M. Scholten and A. Nordheim, *Nucleic Acids Res.* **14**, 3981 (1986).
23. B. H. Johnston, in preparation.
24. D. M. Lilley and B. Kemper, *Cell* **36**, 413 (1984).
25. A. R. Morgan and R. D. Wells, *J. Mol. Biol.* **37**, 63 (1968); J. S. Lee, D. A. Johnson, A. R. Morgan, *Nucleic Acids Res.* **6**, 3073 (1979).
26. H. E. Moser and P. B. Dervan, *Science* **238**, 645 (1987).

27. S. M. Mirkin *et al.*, *Nature* **330**, 495 (1987).
28. C. R. Cantor and A. Efstratiadis, *Nucleic Acids Res.* **12**, 8059 (1984).
29. D. E. Pulleyblank, D. B. Haniford, A. R. Morgan, *Cell* **42**, 271 (1985).
30. T. Evans and A. Efstratiadis, *J. Biol. Chem.* **261**, 14771 (1986).
31. H. C. M. Nelson *et al.*, *Nature* **330**, 221 (1987).
32. M. Coll, C. A. Frederick, A. H.-J. Wang, A. Rich, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 8385 (1987).
33. I thank A. Rich, in whose laboratory this work was

performed, for support and encouragement, H. Htun and J. Dahlberg for generously supplying the plasmid and encouraging this work, T. Kohwi-Shigamatsu for disclosing data before publication, and the Charles King Trust for a postdoctoral fellowship. Supported by grants from the National Science Foundation, NIH, NASA, the American Cancer Society, and the Office of Naval Research to A. Rich.

28 April 1988; accepted 18 August 1988

## Changing the Acceptor Identity of a Transfer RNA by Altering Nucleotides in a "Variable Pocket"

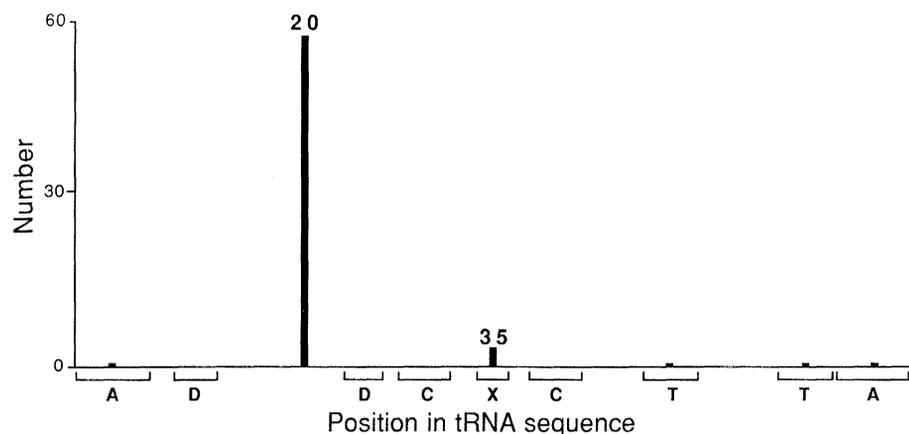
WILLIAM H. McCLAIN AND K. FOSS

The specificity of tRNA<sup>Arg</sup> (arginine transfer RNA) for aminoacylation (its acceptor identity) were first identified by computer analysis and then examined with amber suppressor tRNAs in *Escherichia coli*. On replacing two nucleotides in tRNA<sup>Phe</sup> (phenylalanine transfer RNA) with the corresponding nucleotides from tRNA<sup>Arg</sup>, the acceptor identity of the resulting tRNA was changed to that of tRNA<sup>Arg</sup>. The nucleotides used in the identity transformation occupy a "variable pocket" structure on the surface of the tRNA molecule where two single-stranded loop segments interact. The middle nucleotide in the anticodon also probably contributes to the interaction, since an amber suppressor of tRNA<sup>Arg</sup> had an acceptor identity for lysine as well as arginine.

**T**RANSFORMATION OF THE GENETIC CODE by tRNA molecules requires two specificities, one for amino acids and the other for nucleotide sequences. The specificity of a tRNA for an amino acid is referred to as its acceptor identity (1). The acceptor identity is read by 1 of 20 aminoacyl-tRNA synthetase enzymes, which attach the proper amino acid only to its cognate tRNA species. Computer algorithms can effectively locate the nucleotides in tRNA that determine acceptor identity (2, 3).

The acceptor identity of a tRNA results from two types of interactions, the productive interaction of the tRNA with its cognate aminoacyl-tRNA synthetase and the nonproductive interaction of the tRNA with all other aminoacyl-tRNA synthetases (4, 5). The acceptor identity of tRNA appears to be determined by multiple nucleotides whose locations are not conserved between different molecules. This plurality may sharpen aminoacyl-tRNA synthetase

Department of Bacteriology, University of Wisconsin, Madison, WI 53706.



**Fig. 1.** Frequency of positions in discriminators of tRNA<sup>Arg</sup> (2). Frequency on the vertical axis is a function of position in the tRNA on the horizontal axis. Cloverleaf stem segments on the horizontal axis are indicated as A, acceptor; D, dihydrouridine; C, anticodon (X marks the anticodon); and T, TΨ. The data set contained 65 tRNA sequences. There were 58 one-position discriminators involving position 20 and four two-position discriminators involving positions 4 and 35; 35 and 51; 35 and 63; and 35 and 69.