Identification of Germline and Somatic Mutations Affecting the Retinoblastoma Gene

JAMES M. DUNN, ROBERT A. PHILLIPS, ANDREW J. BECKER, BRENDA L. GALLIE

Retinoblastoma (RB) is a malignant tumor of developing retina that arises when abnormalities resulting in loss of function affect both alleles of the gene at the retinoblastoma locus (RB1) on chromosome 13q. The majority of RB tumors do not show gross alterations in a 4.7-kb fragment (4.7R), which is a candidate RB1 gene. To search for more subtle mutations, the ribonuclease protection method was used to analyze 4.7R messenger RNA from RB tumors. Five of 11 RB tumors, which exhibit normal 4.7R DNA and normal-sized RNA transcripts, showed abnormal ribonuclease cleavage patterns. Three of the five mutations affected the same region of the messenger RNA, consistent with an effect on splicing involving an as yet unidentified 5′ exon. The high frequency of mutations in 4.7R supports the identification of 4.7R as the RB1 gene. However, the unusual nature of some of the abnormalities of 4.7R alleles indicates that the accepted sequence of genetic events involved in the genesis of RB may require reevaluation.

RETINOBLASTOMA (RB) IS THE BEST characterized of several human cancers inherited with an autosomal dominant phenotype. The mutation has been mapped to a single locus, the RB1 locus on chromosome 13q. Analysis of RB tumor DNA indicates that the mutation is recessive and that tumors develop only when the function of both alleles is destroyed. Patients with heritable RB carry one germline mutation such that all cells have one normal and one mutant allele; a tumor develops when the normal allele of a retinal cell is lost or acquires a mutation. Most individuals with germline RB1 mutations develop bilateral RB tumors because mutations arise in several retinal cells in both eyes. Tumors can also develop if a single retinal cell acquires somatic alterations of both RB1 alleles; in these cases, patients develop a single, unilateral tumor, and the phenotype is not heritable. In the majority of RB tumors, the second event appears to involve a chromosomal mechanism leading to homozygosity for most of the long arm of chromosome 13 (*1*). On the basis of the demonstration of reduction to homozygosity for specific chromosomal regions in tumor tissue, a similar two-hit model has been proposed for many other heritable human cancers (*2, 3*).

Retinoblastoma is the only hereditary human cancer for which a candidate gene, 4.7R, has been cloned (*4*). In the absence of a functional assay, the assessment of the validity of this gene as RB1 depends on the analysis of mutations in 4.7R in RB tumors and patients. The majority of RB tumors do not have gross deletions or rearrangements in 4.7R and express transcripts of normal size and abundance (*5*). However, these tumors may contain mutations in 4.7R undetected by RNA and DNA blotting. In the ribonuclease (RNase) protection assay, RNases A and T1 cleave single-stranded RNA at base pair mismatches in RNA:RNA or RNA:DNA hybrids (*6, 7*). The degree of cleavage is affected by the specific mismatch and by the flanking sequences. Some mismatches are totally resistant to RNase digestion, whereas others are cleaved with less than 100% efficiency. As a result, RNase protection identifies only 50% of single base pair mutations (*6, 7*).

Lee *et al.* have published a 4757-bp sequence of the 4.7R cDNA (*8, 9*). The sequence contained an open reading frame of 2922 bp; the proposed start of the coding region was the methionine at nucleotide 139, giving a coding region of 2784 bp and a polypeptide of 928 amino acids. The cDNA used in our studies (*4*) starts at nucleotide 204 according to the map of Lee *et al.* (*9*).

For use in the RNase protection assay, we produced short subclones of 4.7R (Fig. 1). The initial screen was performed with the three probes 11H, 17H, and 19Bg, which covered all but 89 bp of the proposed coding region. There were 66 bp missing between the putative start site and the end of probe 11H, and 23 bp were missing between probes 17H and 19Bg. The localization of mutations was deduced from the cleavage patterns of overlapping probes: 12H, 23H, 18Ba, and 19S.

The 11 RB tumors used in this study exhibited normal genomic patterns when DNA blots were probed with the 4.7R cDNA (*5*). RNA blot analyses of the polyadenylated RNA from four of the RB tumors (RB355, RB405, RB427, and RB522) indicated a normal length 4.7R transcript (*4*). The limiting amount of material from the seven other RB tumor samples prevented their examination on RNA blots. All normal cells and tissues tested expressed 4.7R (*5*). In the present study, we used fibroblasts, normal peripheral blood mononuclear cells, or Epstein-Barr virus (EBV)–transformed lymphocytes (*10*) to examine expression in unaffected tissue.

Five of the 11 RB tumors showed abnormal RNase cleavage patterns. In two RB tumors, RB412 (unilaterally affected) and RB543 (bilaterally affected), we identified two protected fragments that together equaled the length of the fragment normally protected, suggesting that the mutations are

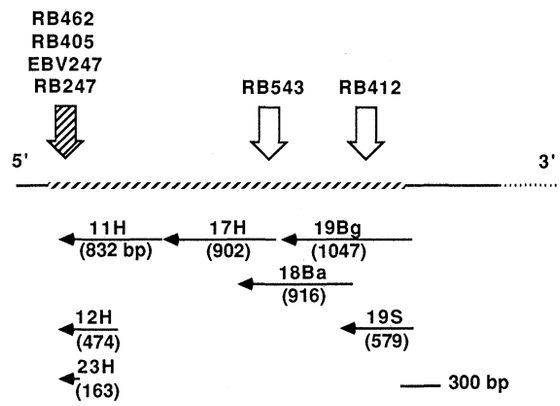

Fig. 1. Schematic diagram of mutations affecting 4.7R mRNA in five RB tumors, and the probes used to localize them. The proposed coding region (hatched line) and the 5′ and 3′ untranslated sequences (single lines) of the 4.7R mRNA are indicated. Above the mRNA are arrows showing the approximate location of each mutation and the cells in which it was detected. Below the mRNA are representations of the antisense RNA probes used, including their designations and the lengths of protected sequence in nucleotides. To make the probes, the 4.7R cDNA was cleaved at the internal Eco RI site into 3652-bp (3′) and 834-bp (5′) fragments. These were subcloned into Bluescript SK (Stratagene) and pTZ18R (Pharmacia) plasmid vectors, respectively. Overlapping plasmids were generated either by unidirectional deletion (*26*) of the 3′ end, or by digestion with the appropriate restriction endonucleases and religation. The ends of all subclones were determined by dideoxy sequencing (*27*) and the lengths were deduced from the published sequence (*9*). Radioactively labeled RNA probes were synthesized from linearized plasmid DNA with phage T7 or T3 RNA polymerase as described (*6*).

J. M. Dunn, R. A. Phillips, B. L. Gallie, Hospital for Sick Children Research Institute, 555 University Avenue, Toronto, Canada M5G 1X8.
A. J. Becker, Department of Medical Genetics, University of Toronto, Toronto, Canada M5S 1A8.

single base pair mismatches at the site of cleavage (Fig. 2, A and B). In tumor RB412, the site of cleavage mapped approximately to nucleotide 2655 and in tumor RB543 to nucleotide 1890 of the sequence published by Lee *et al.* (*9*). Hybridization with restriction fragment length polymorphic (RFLP) DNA probes (*1, 11, 12*) indicated that RB1 in these tumors had reduced to homozygosity (*1, 13*). The mutations were detected only in the tumors and not in the lymphocytes of either patient.

Three of the five tumors (RB247, RB405, and RB462) exhibited identical RNase cleavage patterns with the overlapping probes 11H, 12H (Fig. 2C), and 23H (Fig. 3). The small 23H probe (163 bp) protected both a full-length and a 111-bp fragment in each tumor, locating the abnormality in each of the three tumors precisely 52 bp from the 5′ end of the probe (Figs. 1 and 3). In none of the three tumors was the expected 52-bp 5′ fragment of 23H detected. In each tumor a fully protected band of 163 bp was also evident in addition to the 111-bp band, indicating either the presence of transcripts from the other allele or incomplete digestion of the RNA hybrids protected by the mutant allele. The possibility that these identical mutations were derived from a common origin was eliminated both by the results of RFLP analysis of tumor and constitutional DNAs with a highly polymorphic DNA probe (*14*) and by analysis of their family histories, which indicated that the patients were unrelated.

Because patient 247 had bilateral RB, unaffected parents, and two unaffected siblings, it was likely that this child had a new germline abnormality of the RB1 gene, which arose in the germ cells of one parent or early in the embryo. This prediction was confirmed by the observation that lymphocytes of the child (EBV247) showed the same mutation as the tumor (Fig. 2C); lymphocytes from both parents did not show the mutation. Patient 462 had bilateral RB, inherited from his affected father. The mutation detected in the tumor RB462, however, was evident neither in the patient's fibroblasts (F462) nor in the father's lymphocytes (EBV468) (Fig. 3). Patient 405 had unilateral RB, no family history of RB, and, as expected, the mutation was not detected in his lymphocytes (Fig. 2C).

All patients with bilateral RB must carry one mutant RB1 gene. In screening EBV lines from 12 such patients for RNase protection abnormalities of 4.7R, we detected a mutation in only one, patient 247.

The above data provide strong support for the identification of 4.7R as the RB1 gene. In our series, RNase protection detected one germline mutation and mutant alleles in 45% (5 of 11) of RB tumors. Since this method detects only about half of known mutations (*6*), our data are consistent with the hypothesis that all RB tumors contain mutations of 4.7R.

However, the data set includes several unexpected results. One puzzle is the finding that three of five mutations detected appeared to have deletions extending 5′ of nucleotide 257. The first two nucleotides 3′ of the deletion are an AG at position 257 to 258. Since the dinucleotide AG is a feature of 3′ splice junctions (*15*), one explanation for the deleted mRNA in the three tumors is that upstream mutations have inactivated an existing splice acceptor, leading to the use of this AG as a cryptic 3′ splice junction. Such mutations have been documented in β-thalassemia (*16*), phenylketonuria (*17*), hemophilia B (*18*), and Tay-Sachs disease (*19*). However, the AG at position 257 to 258 is in the putative first exon of RB1 (*20*). Since the 5′ end of the 4.7R message has not been

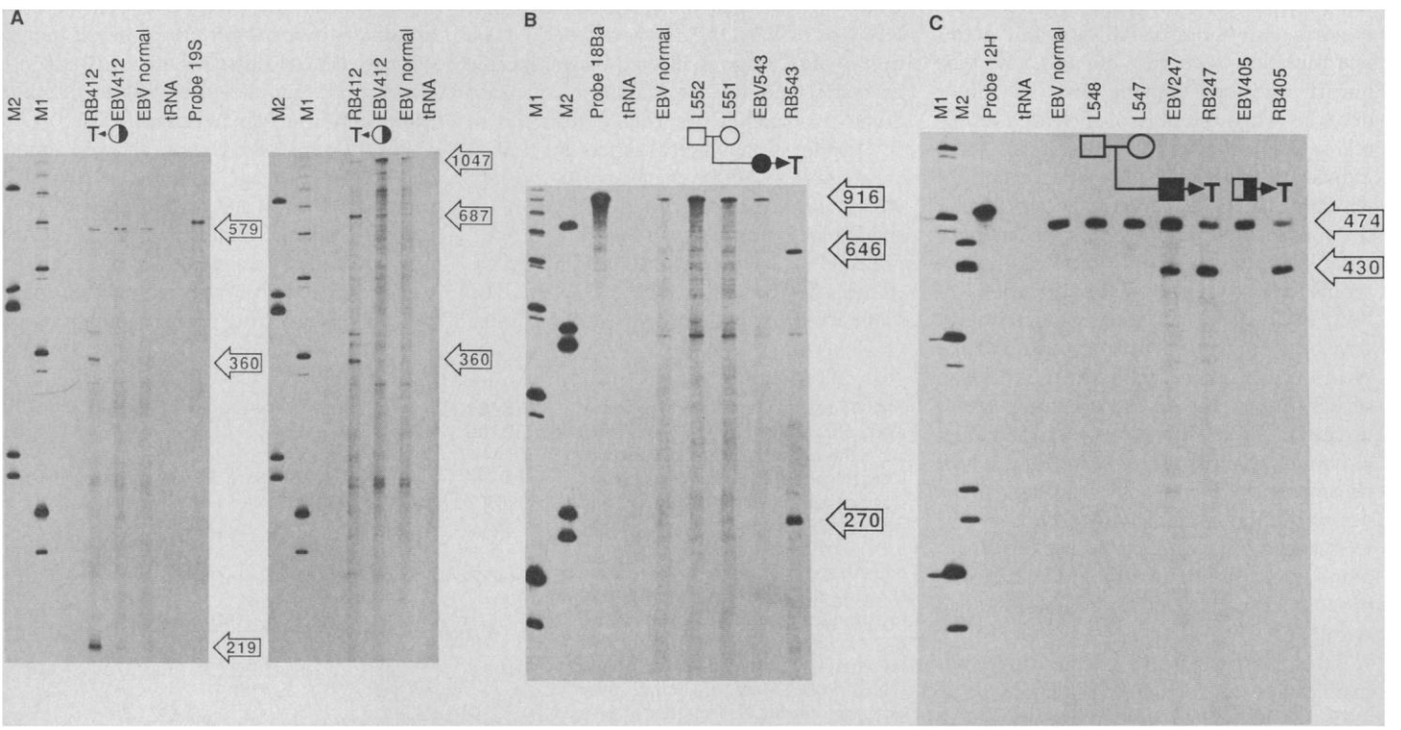

Fig. 2. RNase cleavage patterns of 4.7R transcripts from RB tumor samples, RB cell lines, and constitutional cells (see below). In the pedigrees, circles and squares represent the constitutional cells and "T" represents the tumor. Open, half-filled, and completely filled symbols represent unaffected individuals, and unilateral and bilaterally affected patients, respectively. Total RNA, 5 to 20 μg, from the tumor or constitutional cells indicated, were hybridized to 10^5 cpm of radioactively labeled antisense RNA probes (**A**) 19Bg (right) and 19S (left), (**B**) 18Ba, and (**C**) 12H. In the lanes marked tRNA, 20 μg of unlabeled yeast transfer RNA was used as a control. EBV normal indicates the lanes with RNA from a normal individual. Markers are M1, the BRL 123-bp ladder, and M2, pTZ18R digested with Hae III. The numbers on the sides enclosed in arrows indicate the size (in nucleotides) of the protected fragments. RNA was prepared by the guanidinium isothiocyanate method as described (*28*). Constitutional RNA was isolated from EBV-transformed lymphoblastoid cell lines (EBV) (*10*) or conconavalin A–stimulated lymphocytes (L). RNA probes were hybridized to total cellular RNA and digested with RNase A (5 μg/ml) and T1 (100 U/ml) as described (*6*). Protected fragments were analyzed by denaturing polyacrylamide gel electrophoresis followed by autoradiography.

unequivocally identified, it is possible that one or more 5′ exons remain to be discovered. Another, but less likely, explanation of the clustering of RNA deletions at the 5′ end is that the mutations in these three patients lead to a stable secondary RNA structure, such as a stem loop, so that the mutated mRNA cannot protect the probes in the region of the secondary structure.

Perhaps the most surprising results bear on the two-hit model of retinoblastoma (*1, 2*). The mutations in patients 405 and 247 follow the expected pattern. RB405 is a unilateral tumor from a patient with no family history; the tumor retains both the paternal and maternal chromosome 13's (*1*), and presumably was induced by two independent mutations in the RB1 alleles. Ribonuclease protection detects only one of the mutations. RB247 is a tumor that retained both chromosome 13's (*1, 13*) from a bilaterally affected patient, who must carry one mutant RB1 allele in his germline. The detection of the mutation in both the tumor and lymphocytes is consistent with this conclusion.

However, patients 462 and 543 do not appear to follow the expected pattern of inheritance of mutations detected by RNase protection. Patient 462 inherited RB from his affected father; the tumor, RB462, was previously shown to have reduced to homozygosity for the paternal chromosome by the mechanism of mitotic recombination (*21*). Therefore, we expected the mutation detected in RB462 to be the germline mutation inherited from the father. The absence of detectable mutant transcripts in normal cells of patient 462 or his father seems inconsistent with the pattern of inheritance, especially since the similar abnormality of RB247 was detectable in EBV247. The bilaterally affected patient 543 must also carry a germline mutation in the RB1 gene. RB543 reduced to homozygosity for chromosome 13 RFLP probes, including two within 4.7R (*12, 13*). RB543 was different from the other tumors in that it lacked a fully protected band and showed only two mutant bands, the sum of which equaled the fully protected band normally seen (Fig. 2B). If the mutation detected by RNase protection arose before the development of homozygosity of chromosome 13, consistent with the two-hit model for RB, the detected mutation should be the germline mutation. We were unable to detect such a germline mutation in lymphocytes (EBV543) of the patient.

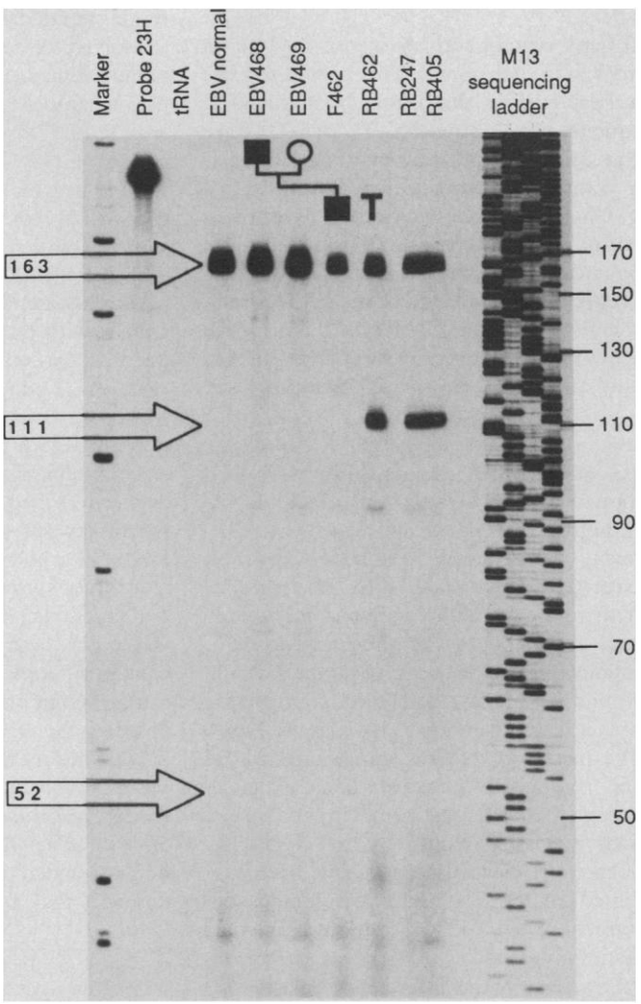

Fig. 3. Analysis of the family and tumor of patient 462, and sizing of protected fragments in the three RB cell lines RB247, RB405, and RB462 by RNase protection. RB247 and RB405 are run in immediately adjacent lanes for size comparison. Probe 23H was hybridized and digested as in Fig. 2. F462 are normal fibroblasts isolated from the bilaterally affected patient 462. The markers are: Bluescript SK digested with Eco RI and Hae III and M13 single-stranded DNA sequenced by the dideoxy method with Sequenase T7 DNA polymerase (USB) according to the manufacturer's specifications. The full-length protected fragment 163 nucleotides (nt), the mutation fragment of 111 nt, and the location of the expected fragment of 52 nt (not detected) are marked (arrows).

Several explanations could account for the apparently anomalous results obtained from patients 462 and 543. (i) The expression of some mutant transcripts may be suppressed by the normal transcript, or the mutant transcript may be unstable in EBV-transformed lymphocytes and normal fibroblasts. If the deletions observed in RB247, RB405, and RB462 reflect splice-junction mutations, the mutations responsible may be 5′ of the region probed and could be different in each tumor. Stability of the mRNA or suppression of expression in lymphocytes and fibroblasts may vary with the individual mutation. (ii) The mutations of 4.7R may have occurred as tertiary events on one chromosome after the development of homozygosity. Such tertiary mutations might selectively survive if the initial, germline mutations (not detected by RNase protection) failed to completely inactivate RB1 gene function. (iii) Since homozygosity of chromosome 13 in RB462 arose by mitotic recombination (*21*), the mutation may also have occurred during the recombination process. Further studies with intragenic RFLP probes would define the relation of the mutation to the site of mitotic recombination in this tumor. (iv) Although mosaicism cannot account for the observations in the family of RB462, since the mutation has passed through two generations, it may explain our failure to detect a mutation in the lymphocytes of other bilaterally affected patients. Motegi has reported that many new germline mutations in the RB gene first appear as mosaics (*22*). (v) It is possible that the mutation in RB543, detected by RNase protection, was the second, somatic event and that homozygosity for the somatic, and not the germline, mutation occurred as a tertiary event. If the somatic mutation resulted in a more severe loss of RB1 function than the germline mutation, homozygosity for the somatic mutation could have conferred a selective advantage on the cells. If rearrangements of RB1 occur frequently as tertiary events in RB tumors, the identification of mutations in tumors or identification of the chromosome retained after reduction to homozygosity may not be useful in genetic counseling. Instead, the germline mutation must be identified in constitutional cells.

Recent observations link abnormalities in RB1 to breast cancer (*23*) and small-cell lung cancer (*24*), tumors that have not been closely linked to RB epidemiologically. Others have demonstrated an interaction of the

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. Breathnack 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. Gütler, 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. Motegi, *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. Coulsen, *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)$_n$·d(A-G)$_n$: 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)$_n$·d(A-G)$_n$, are known to undergo a transition in plasmids to an underwound state that is hypersensitive to single strand–specific nucleases; this transition occurs under conditions of moderately acid *p*H 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)$_{18}$·d(A-G)$_{18}$ as a function of supercoiling and *p*H, and thus analyze the structure in this region. Hyperreactivity was seen in the center and at one end of the (C-T)$_n$ tract, and continuously from the center to the same end of the (A-G)$_n$ tract, in the presence of supercoiling and *p*H ≤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)$_n$·d(G-A)$_n$, henceforth designated (C-T)$_n$. S1-hypersensitive (C-T)$_n$ 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)$_{18}$ 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 *p*H and is aided by negative supercoiling, although the need for supercoiling diminishes with increasing length of the (C-T)$_n$ sequence (*8*). Several models have been suggested for the S1-sensitive, unwound conformation of (C-T)$_n$ 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_4), which reacts preferentially with thymines at B-Z junctions (*10, 21*). DEP (*22*), hydroxylamine (*23*), and OsO_4 (*24*) also react with cruciform loops. DEP favors structures, including single-stranded regions or Z-DNA, in which the reactive N^6 and N-7 positions of adenine and the N-7 position of guanine are especially accessible. The pyrimidine-specific reagents hydroxylamine and OsO_4, 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.