Effects of Translocations on Transcription from PVT

E. SHTIVELMAN* AND J. M. BISHOP

The G. W. Hooper Research Foundation, Department of Microbiology and Immunology, University of California Medical Center, San Francisco, California 94143

Received 17 November 1989/Accepted 10 January 1990

We have previously described a transcription unit on human chromosome 8, designated as PVT, that is consistently disrupted by the minority forms of translocations (t(2;8) and t(8;22)) in Burkitt's lymphoma. PVT begins 57 kilobase pairs downstream of the proto-oncogene MYC and is more than 200 kilobase pairs in length. In order to explore the pathogenic impact of translocations affecting PVT, we have characterized further the structure and transcription of the locus. In normal cells, PVT is transcribed into a variety of RNAs, the diversity of which remains unexplained. Alleles of PVT affected by translocations give rise to additional RNAs. These RNAs arise from a fusion of the first exon of PVT on chromosome 8 to the constant region of an immunoglobulin light chain on either chromosome 2 or chromosome 22. We have found no evidence that any of the normal or abnormal transcripts of PVT give rise to a protein. Our results suggest that the pathogenic effects of the variant translocations in Burkitt's lymphoma are not executed by a gene situated in a vicinity of the chromosomal breakpoints. Instead, our data leave open the possibility that the effects of the translocations may be mediated by activation of the relatively distant MYC gene.

Activation of the proto-oncogene MYC occurs in a variety of human and animal tumors (1). The best-documented example is the involvement of MYC in the development of B-cell malignancies such as Burkitt's lymphomas in humans and plasmacytomas in mice (2). Consistent chromosomal abnormalities associated with these tumors typically result in the translocation of MYC into the immunoglobulin heavy-chain locus and inappropriate transcriptional activation of MYC. However, in about 20% of the tumors, translocations involve the immunoglobulin light-chain (IGLC) loci. In these variant lymphomas, MYC remains on its normal chromosome, whereas parts of IGLC loci (usually including the transcriptional enhancer and the constant-region exon) are translocated to chromosome 8 downstream of MYC (4-6, 12, 18). Although the breakpoints can occur at substantial distances downstream of MYC, the transcriptional activation of MYC appears to be an inevitable accompaniment of the variant translocations (4-6, 12, 18).

A DNA locus known as pvt-l located more than 100 kilobase pairs (kbp) downstream of myc is consistently involved in variant translocations in mouse plasmacytomas (3, 20). The same locus is a recurrent site for retroviral insertions in T lymphomas of mice and rats (8, 14, 19). Sequences within pvt-l are conserved in the human genome and map 260 kbp downstream of MYC (7, 9, 15); at least one translocation breakpoint of Burkitt's lymphoma with t(2;8) occurs in the human pvt-l sequences (15). It has been proposed that pvt-l harbors a gene whose function is disturbed through rearrangements of this locus and thus contributes to the process of transformation (7). Until now, however, there has been no evidence that pvt-l or sequences in its vicinity are transcribed.

We have recently reported the identification of a human transcription unit, PVT, which starts 57 kbp downstream of MYC, extends more than 200 kbp in the telomeric direction from MYC, and encompasses the human pvt-l sequences (16, 17). PVT is transcribed by chromosome 8 breakpoints in a number of Burkitt's lymphomas carrying variant translocations. Rearrangements of PVT are accompanied by the appearance of short PVT transcripts expressed at relatively high levels (17).

Here we report the results of our analysis of transcription from normal PVT and PVT rearranged in Burkitt's lymphoma. We show that the transcripts from the PVT locus probably do not give rise to proteins. The analysis of the anomalous PVT transcripts in Burkitt's lymphomas reveals that they are products of splicing between the first exon of PVT and the constant-region exon of IGLC loci translocated to chromosome 8q+. The functional significance of transcription from PVT remains a puzzle, and disturbance of the locus by translocations is unlikely to play a direct role in tumorigenesis.

Analysis of RNAs transcribed from PVT. We have demonstrated previously that PVT is expressed in a variety of human cell lines as several RNA species of extremely low abundance (16). The PVT cDNA described in our previous work was isolated from the colon carcinoma cell line COLO320-DM, in which the PVT locus is highly amplified and possibly rearranged (16). Thus, we attempted to clone normal PVT from a cDNA library derived from human placenta. Screening of about 106 cDNA clones with a probe for the first three exons of PVT (16) produced only one positive clone, designated HP14. The 1.6-kbp insert of HP14 was used as a probe for repeated screening of a placental cDNA library without producing additional PVT clones.

The insert of clone HP14 was subjected to sequencing analysis, and the results are presented in Fig. 1B. The 5' end of HP14 contains sequences of the first exon of PVT defined previously (16, 17). The two longest open reading frames identified within the HP14 sequence are 80 and 82 amino acid residues long and start at ATG codons at positions 210 and 420, respectively (Fig. 1B). The sequences in the immediate vicinities of both putative translation initiation sites represent poor contexts for translation initiation (13). Attempts to translate in vitro RNA made from HP14 did not result in proteins that we could attribute specifically to PVT (data not shown).

We compared the sequence of PVT cDNA clone HP14 with that of clone Y2 from COLO320-DM described previ-
The method used to construct the cDNA libraries in λ bacteriophage vector λ10 was described previously (16). The cDNAs isolated ranged in size from 1.5 to 5.5 kbp. Ten cDNA clones were analyzed by restriction enzyme mapping, hybridization with the exon-specific PVT probes, genomic Southern blot hybridization, and partial sequence analysis. Sequences of PVT exons I, II, and III were found to compose the 5' regions of all cDNAs examined. The various lengths of sequences farther downstream in all clones contained repetitive elements. We failed to identify substantial protein-coding regions within regions of PVT cDNAs free of repetitive sequences. Thus, we are left with no evidence that the PVT transcription unit may encode a protein.

Aberrant PVT RNA in variant Burkitt's lymphomas. We demonstrated previously that the PVT transcription unit is frequently transacted by breakpoints of chromosome 8 in Burkitt's lymphomas carrying variant translocations 2:8 or 8:22 (17). We observed that in six of nine cell lines of Burkitt's lymphoma examined, there is accumulation of short PVT-specific RNAs 0.8 to 1 kilobases long; these RNAs are expressed at substantially higher levels than normal PVT RNAs (17). The aberrant PVT transcripts hy-
bridged to a probe for the first exon of PVT but not to other regions of PVT and thus contained unidentified nucleotide sequences.

To study the structure of lymphoma-specific PVT RNAs, we cloned corresponding cDNAs from libraries constructed with polyadenylated RNAs from two Burkitt's lymphoma cell lines, BL21 (carrying t(2;8)) and MWIKA (carrying t(8;22)). The libraries were screened with probes for the first and the second exon of PVT. Several cDNA clones hybridizing to the probe of the first PVT exon only were analyzed from each library. The partial nucleotide sequences of two representative cDNAs from each library are shown in Fig. 2B. The 3' ends of the cDNAs consist of full copies of exons from the constant region of IGLC genes. cDNAs Bs27 and Bs40 from cell line BL21 contain the C_κ exon followed by remnants of poly(A) tail, whereas cDNAs MW8 and MW10 from cell line MWIKA contain the C_λ exon. All cDNAs analyzed contain the previously defined first exon of PVT (designated Ia in Fig. 2) at their 5' ends. These data indicate that the aberrant PVT RNAs are in fact products of transcription starting at the first exon of PVT and continuing in light-chain sequences brought to chromosome 8q- through variant translocation.

Analysis of the clone MW8 revealed that the first exon of PVT and the C_λ exon are joined through normal splice signals, since all the exon sequences were preserved (Fig. 2B). However, additional sequences were found between the PVT and IGLC exons in other aberrant cDNAs. For example, clone MW10 contained an additional sequence of 115 nucleotides between the first exon of PVT and C_λ. The same sequence fragment of 115 nucleotides (hereafter designated Ib) was found in another cDNA clone, Bs40, isolated from cell line BL21 (Fig. 2A and B). Bs40 contained sequences Ia and Ib at the 5' end, followed by a new sequence of 164 base pairs (bp) (Ic) and only then the full C_κ exon. Another cDNA clone from BL21, designated Bs27, contained only the Ic sequence fragment between the first exon of PVT and the IGLC sequences (Fig. 2).

We conclude that translocation of the constant region of the IGLC loci downstream of the first exon of PVT results in production of fused transcripts between PVT and immunoglobulin sequences. The fused RNAs are produced in relatively great abundance, perhaps because the IGLC enhancer region is present on chromosome 8q-. Two additional sequence fragments of unknown origin, designated here as Ia and Ic, are on occasion included in PVT-IGLC fused transcripts, apparently as a result of alternative splicing (see below; Fig. 3).

Analysis of the sequences of chimeric cDNAs derived from the cell line MWIKA demonstrated that within their 5' end sequences, there are no translation initiation codons in frame with the C_κ protein coding sequence. Thus, the fused transcripts that contain C_κ are presumably sterile. The cDNAs Bs27 and Bs40 from cell line BL21 contain an ATG codon in the Ic sequence 43 base pairs upstream of the C_κ exon in the correct reading frame (Fig. 2). Translation of this open reading frame should result in production of a protein that consists of 14 amino acids encoded by Ic and then the full κ constant region. In vitro translation of Bs27 RNA produced a 16-kilodalton protein, in agreement with the
predicted size. This protein could be specifically precipitated from in vitro products with antibodies against the human IGLC of type $\kappa$ (data not shown). However, all attempts to detect an analogous protein in lysates of BL21 cells were unsuccessful (data not shown).

**Origin of internal sequences found in PVT-IGLC chimeric transcripts.** Since the sequence designated Ib was found in chimeric RNAs derived from lines carrying both types of variant translocations, we assumed that it is most likely to be encoded within chromosome 8, between the first exon of $PVT$ and the translocation breakpoint. In an attempt to define the genomic localization of Ib, we hybridized its probe to cloned genomic fragments from the $PVT$ locus. We found that Ib specifically hybridizes to the cosmid clone cosY14 described previously (17), which spans the DNA region between 45 and 70 kbp downstream of $MYC$. This cosmid clone contains the first exon of $PVT$, which maps 57 kbp downstream of $MYC$ (17). Restriction mapping and hybridization analysis enabled us to localize the Ib genomic region to a position 1.4 kbp downstream of the first exon of $PVT$ (Fig. 3A). Sequence analysis of the Ib genomic fragment revealed that the sequence expressed within chimeric RNAs is preceded in genomic DNA by a sequence resembling an acceptor splice site (Fig. 3B). This presumably allows splicing of Ib into chimeric RNAs.

Similar analysis was done with a probe for the sequence Ic found in chimeric transcripts from cell line BL21. We found that the probe specifically hybridized to the genomic phage clone $\lambda 8q1$, whose isolation was reported previously by others (9). $\lambda 8q1$ encompasses chromosome 8 sequences localized about 120 kbp downstream of $MYC$ and maps the translocation breakpoints in cell line BL21 and in two additional Burkitt’s lymphomas with t(2;8) (9). Restriction enzyme mapping and hybridization analysis of $\lambda 8q1$ revealed that the genomic locus for Ic is positioned 1 kbp upstream of the breakpoint in BL21 (Fig. 3A). Thus, the Ic locus and the translocated C type are juxtaposed on chromosome 8q$^-$ in BL21. Ic is flanked by potential splice signals which presumably facilitate its incorporation into $PVT-C_\kappa$ fusion transcripts (Fig. 3B).

**Discussion.** The identification of a transcription unit encompassing the $pvt-1$ locus was of potential interest because the locus is affected consistently by DNA rearrangements in B- and T-cell malignancies of humans and rodents. In this work, we tried to address the relevance of $PVT$ rearrangements to lymphomagenesis through the analysis of $PVT$ transcription in normal and Burkitt’s lymphoma cells.

We report here that the multiple transcripts produced from $PVT$ are unlikely to encode a protein. Because of this finding, the functional role of transcription from $PVT$ in a variety of cells remains a puzzle. Transcription from normal $PVT$ produces an array of low-abundance RNAs, apparently through a complex pattern of splicing between exons separated by large introns. Thus, the first intron of $PVT$ is more than 100 kbp long (17). Why should a cell maintain transcription from a large and complex locus that results in production of multiple sterile transcripts? One possibility is that $PVT$ produces functional mRNAs, but only in some specialized tissue or during a restricted period of development. If so, the rare $PVT$ transcripts detected in a variety of human cell lines represent aberrant transcriptional activity from the $PVT$ promoter. Another possibility is that $PVT$ represents remnants of an ancestral gene which became obsolete and suffered mutations during the course of evolution. This view is encouraged by our failure to find a complete representation of $PVT$ in mouse DNA (see below).

The abnormal $PVT$ RNAs produced in variant Burkitt’s lymphomas are chimeric transcripts containing $PVT$ and IGLC constant-region sequences. These RNAs are abundant in all but one of the cell lines with breakpoints located within $PVT$ (17). The exception was cell line JBL2, which has lost the $C_\kappa$ sequences from 8q$^-$ (B. Henglein, personal communication). The $PVT$-IGLC fusion transcripts contain the first exon of $PVT$ and the constant-region exon of the light-chain type involved in translocation. Two sequence fragments, designated Ib and Ic in this paper and located downstream of the first exon of $PVT$, are included in the chimeric transcripts, apparently via alternative splicing.

We found that the Ib and Ic loci are not expressed in RNA from human cell lines containing unrearranged $PVT$ (data not shown). We presume that the presence of the light-chain enhancer on chromosome 8q$^-$ not only results in previously demonstrated transcriptional activation from the normal $PVT$ promoter (17) but also enables splicing of Ib and Ic.
sequences into PVT-IGLC fusion transcripts. This is presumably facilitated by the presence of sequences resembling splice signals flanking Ib and Ic fragments (Fig. 3). High levels of transcription from Ib and Ic were also found in colon carcinoma cell line COLO320-DM, where PVT is amplified (unpublished data). It is possible that Ib and Ic are retained in PVT-specific RNAs only when PVT is rearranged or amplified.

PVT is the first transcribed locus found to contain pvt-1 (17). PVT seemed to be a good candidate for a gene whose rearrangements and disturbed function might contribute to development of Burkitt’s lymphomas carrying variant translocations. It is transacted by translocation breakpoints in at least some Burkitt’s lymphomas, resulting in accumulation of abnormal PVT transcripts (17). Moreover, we reported detection of PVT sequences in the mouse genome, suggesting that PVT might be the gene affected by variant translocations and proviral insertions (16). However, subsequent analysis revealed that the detectable homology of PVT between the human and mouse genomes is limited to the sequences of the first PVT exon only and does not pertain to the rest of the PVT transcription unit (unpublished results).

The absence of a homologous transcribed locus in the vicinity of mouse pvt-1 argues against the potential role of PVT in lymphomagenesis. This view is supported by the results of our analysis of normal PVT RNAs and PVT-IGLC fused transcripts. We found no evidence that transcription from normal PVT gives rise to a protein or that it is activated in Burkitt’s lymphomas. Furthermore, the PVT-Cs transcripts are presumably sterile; the PVT-Cs transcripts might produce a truncated C protein, but we failed to detect it. Our failure to discern functional activity of PVT leaves open the possibility that the pathologically important consequence of both the variant and major types of translocations in Burkitt’s lymphomas is transcriptional activation of the proto-oncogene MYC.

We express our gratitude to Judith Levine for her help with the sequencing, to B. Henglein for the gift of phage clone λq1,1 and to J. Ebina for the human placental cDNA library.

Work described here was supported by Public Health Service grant CA 44338 from the National Institutes of Health and by funds from the G. W. Hooper Research Foundation.

LITERATURE CITED