JOURNAL OF VIROLOGY, Aug. 1997, p. 6136–6143 Vol. 71, No. 8
0022-538X/97/$04.00+0

Sp1 Binds to the Precise Locus of End Processing within the Terminal Repeats of Epstein-Barr Virus DNA

REN SUN,[1,2] TAMMY A. SPAIN,[1] SU-FANG LIN,[1] AND GEORGE MILLER[1,3]*

Department of Molecular Biophysics and Biochemistry,[1] Department of Genetics,[2] and Departments of Pediatrics and Epidemiology and Public Health,[3] Yale University School of Medicine, New Haven, Connecticut 06520-8064

Received 25 February 1997/Accepted 24 April 1997

Interconversion between the linear genome of Epstein-Barr virus (EBV) present in virions and intracellular circular EBV DNA is a novel DNA recombination process. A previously characterized DNA binding activity called terminal repeat or tandem repeat binding protein (TRBP) was found to recognize several G-rich recombinogenic sequences in the EBV genome and in cellular DNA. TRBP was also found to be an autoantigen recognized by sera from certain patients with undifferentiated connective-tissue disorders. Here the transcription factor Sp1 has been identified as a component of TRBP and has been shown to be an autoantigen. Sp1 bound to recombination junctions of EBV DNA, such as those in the terminal repeats and in the large internal repeats, as well as to recombinogenic regions of cellular DNA, such as variable-number tandem repeats and switch regions of the immunoglobulin genes. We defined the ends of the linear EBV genome present in virions and showed that Sp1 binds to the sequence (GGGGTGGGGCATGGG) within EBV terminal repeats at the precise locus of interconversion of linear and circular viral DNA. Sp1 may be involved in DNA recombination.

Although the detailed mechanism of DNA recombination in higher eukaryotes is unclear, certain sequence elements promote recombination. The human genome contains dispersed minisatellite sequences or variable-number tandem repeat (VNTR) sequences which are recombinogenic (15, 16, 22, 30, 52). Minisatellite and VNTR sequences often contain the motif GGNNGTGGGG. Immunoglobulin heavy-chain class switch regions also consist of highly repetitive GC-rich sequences (8, 17, 18, 28, 31, 32, 34, 39, 55). The switch region of gamma immunoglobulin genes (Sγ) contains repeats of 49-mers within which the sequence GGGGAGCTGGGG is often found (38, 39). Chromosome telomere sequences, composed of TTGGGG or TTAGGG repeats, are also recombinogenic (3, 7, 35, 53). Furthermore, VNTR sequences, Sγ repeats, and telomere sequences have been found to form stable four-stranded structures in vitro, called G4 DNA, G quartet, or G quadruplex (1, 23, 38). Such structures may be intermediates of recombination or may promote stabilization of the ends of DNA (24).

Similar to the cellular genome, the Epstein-Barr virus (EBV) genome contains repeated GC-rich recombinogenic sequences (5). The number of these repeats varies remarkably in different virus isolates, even after only one passage of the same isolate (references 36 and 37 and unpublished results). The double-stranded DNA viral genome is linear in virions but circular in the nuclei of infected cells (2, 12, 19). The circularization occurs 16 to 20 h postinfection and requires that the infected cells be activated into the G_1 phase of the cell cycle (12). Furthermore, EBV infection activates the immunoglobulin gene class switch recombination mechanism in B cells (48). The EBV terminal repeats (TR), which are the sites of linearization and circularization, are also loci of occasional integration of EBV DNA into cellular DNA (27). TR of other herpesviruses play a role in recombination events that are involved in isomerization (29). The termini of some herpesviruses are similar to cellular telomeres (9, 21). Internal repeats (IR) in EBV are also recombinogenic. During coinfection with two different EBV strains, hybrid viruses containing crossovers within the first IR were frequently isolated (unpublished results).

Our laboratory has previously identified a recombination event between squirrel monkey retrovirus (SMRV) and EBV (45). The recombination junction in EBV DNA was in the first IR at the same position as one junction of the deletion in strain P3J-HR-1 which removed the *EBNA2* gene (14). The sequence (GGGTGG) was present at the junctions of both recombination events (43). The sequences around the recombination sites in EBV are homologous to the common sequence of the Sγ repeats and VNTRs. Similar sequences are found in EBV TR.

A specific binding activity, called terminal repeat or tandem repeat binding protein (TRBP), was identified (46). TRBP bound to G-rich recombinogenic sequences in EBV and cellular DNA, it formed a specific footprint on the EBV TR at the site of terminal processing, it was present in the nuclear extracts of all mammalian cells tested, and it was a human autoantigen. Here we identify Sp1 as a component of this binding activity and show that it binds precisely at the locus of EBV terminal processing.

MATERIALS AND METHODS

Cell culture. Human and marmoset B lymphocytes [Raji, BJAB, BJAB-B1, P3HR-1/CL16(HH514-16), and B95-8] were grown in RPMI 1640 medium supplemented with 8% fetal calf serum. HeLa cells were cultured with RPMI 1640 medium supplemented with 6% calf serum. All media contained penicillin, amphotericin B, and streptomycin.

Nuclear extracts. Nuclei of cultured cells were prepared as described previously (46); nuclei were extracted with 0.4 M KCl–0.2 mM EDTA (pH 8.0)–20 mM HEPES (pH 7.9)–20% glycerol–0.5 mM dithiothreitol–0.5 mM phenylmethylsulfonyl fluoride. The protein concentrations of nuclear extracts, determined by the Bio-Rad protein assay, ranged from 3 to 10 mg/ml.

DNase I footprinting. The probe was a 282-bp *Bst*EII/*Nar*I restriction fragment which corresponds to bp 170538 to 170819 of the EBV sequence. The probe was radiolabelled at the *Bst*EII end with [γ-^{32}P]ATP by using T4 polynucleotide kinase. DNA binding and DNase I digestion were performed under previously described conditions (4). Nuclear extract from P3HR-1/CL16(HH514-16) cells

* Corresponding author. Mailing address: 420 LSOG, 333 Cedar St., New Haven, CT 06520. Phone: (203) 785-4758. Fax: (203) 785-7194. E-mail: George_Miller@quickmail.yale.edu.

TABLE 1. Sequences of double-stranded oligonucleotides[a]

Oligonucleotide	Sequence
TRBS1	CGAGATCGGGGTGGGGCATGGGGGATCCCG
Tm1A	CGAGATCGGGGTGAGGCATGGGGGATCCCG
Tm1B	CGAGATCTCGGGGTGGGGCATGAGGGATCCATATC
Tm1C	CGAGATCTCGCGGTGGGGCATGGGGGATCCATATC
Tm2	CGAGATCTCGGGGAGGGGCAAGGGGGATCCATATC
Tm3	CGAGATCGCGGTGAGGCATGAGGGATCCCG
SP1	ATTCGATCGGGGCGGGGCGAG
W	CGAGATCTCTGGGGGCTAGTCTGGGTGGGGGATCCCG
H	CGCTGGCCCCACCCTGGGGGTCGCCGCCGGATCCATATC
Sγ3c	TGAGCTGGGGGAGCTGGGGTGGGGATCCATATC
VNTR	ACAGCTCTGGGGACGTGGGGTGGGATCCATATC

[a] Mutated sequences are underlined.

or purified human Sp1 protein purchased from Promega (Madison, Wis.) was incubated with the labelled probe before DNase I digestion. A 24-mer oligonucleotide (5′ GTGACCCAGCCAAGCGTGACCAAG 3′), corresponding to the 5′ end of the probe, was used as a primer in a DNA sequencing reaction. The sequencing reaction and the DNase I protection assay were run in adjacent lanes to determine the protected sequences.

Synthetic DNA oligonucleotides. The sequences of double-stranded oligonucleotides used in the electrophoretic mobility shift assay (EMSA) are shown in Table 1. Single-stranded oligonucleotides were annealed and radiolabelled by filling in recessed 3′ ends with Klenow fragment and [α-^{32}P]dCTP (4).

EMSA. Binding reactions were carried out in 10 μl of binding buffer [10 mM HEPES (pH 7.9), 100 mM KCl, 1 mM dithiothreitol, 0.5 mM EDTA (pH 8.0), 20% (vol/vol) glycerol, 1 μg of poly(dI-dC)]. ^{32}P-labelled oligonucleotides with 2.5×10^4 cpm at a final concentration of 5 nM and specific competitor DNAs were used as indicated below. The binding reactions were initiated by adding 5 to 10 μg of nuclear extract or 0.1 footprint unit (FPU) of pure Sp1 protein to the initial mixture and incubating it for 10 min at room temperature. Bound and free probes in the EMSA were separated on a 4% polyacrylamide gel.

Antibodies and supershift assay. Rabbit polyclonal anti-Sp1 antibody was purchased from Santa Cruz Biotechnology (Santa Cruz, Calif.). Coded human sera from patients with autoimmune diseases were obtained from Yale New Haven Hospital. Antibody supershift assays were performed by adding 1 μl of autoimmune serum to the EMSA reaction mixture and incubating it for an additional 10 min at room temperature. Bound and free probes in the EMSA and supershift assays were separated on a 4% polyacrylamide gel.

Western immunoblotting. After resolution by electrophoresis on sodium dodecyl sulfate (SDS)-polyacrylamide gels, proteins were electrotransferred to nitrocellulose filters in a solution containing 0.25 M Tris, 0.15 M glycine, 0.1% SDS, and 25% methanol with a Bio-Rad Transblot apparatus for 2 h at 200 mA (4). Nitrocellulose filters were blocked in 5% nonfat dry milk for over 1 h before incubation with antibodies. Human sera were used at a 1:200 dilution, and rabbit sera were used at a 1:1,000 dilution. The bound primary antibodies were detected by ^{125}I-labelled protein A.

Primer extension with virion DNA. Primers were radiolabelled by T4 polynucleotide kinase at the 5′ end (4). EBV virion DNA was prepared as described previously (45). One hundred nanograms of EBV virion DNA was denatured and annealed with ^{32}P-labelled primer B (5′ ATCCCCGGAACGTCCGCCGC CATC) or primer C (5′ GTGACCCAGCCAAGCGTGACCAAGG) to map the 5′ nucleotides of the left and right ends, respectively, of the linear genome. Primer extension reactions were performed with *Taq* DNA polymerase. Extended products were run next to a DNA sequencing reaction mixture which served as a molecular weight marker.

High-resolution Southern blotting. Virion DNA digested with a restriction endonuclease was run on a 6% acrylamide gel adjacent to a ^{35}S-labelled DNA sequencing reaction mixture which served as a molecular weight marker. *Apa*I was used to map the right end, and *Ava*II was used to map the left end. The DNA was transferred from the gel to nitrocellulose filters with a Bio-Rad gel dryer without heating. The DNA was cross-linked to the filters with UV light and hybridized with oligonucleotides radiolabelled with T4 polynucleotide kinase. Oligonucleotides H (TGACACAGGCAACCCTGACAAAGG) and I (GCCGT AGCGCCGCTCTGTGCG), respectively, were used to determine the 5′ ends of the termini. Then the filters were stripped and hybridized with kinase-labelled oligonucleotides J (CGGGGGTCTTTCCTGGGGGCCTTTGTCAGGGTTG) and K (CATTCCTGGAAAAAGTGGAGGGGGCGTGGCCTTCC), respectively, to determine the 3′ ends of the termini.

Ligation-mediated PCR. To generate adapters, oligonucleotide F (5′ GTAA ACGACGGCCAGTAAGAATTCCCC) was annealed to one of three oligonucleotides (5′ GGGAATTCTTAC, 5′ GGGGAATTCTTAC, or 5′ GGGGGAA TTCTTAC); oligonucleotide G (5′ GTAAAACGACGGCCAGTAAGAATTC GGG) was also annealed to one of three oligonucleotides (5′ CCGAATTCTT AC, 5′ CCCGAATTCTTAC, or 5′ CCCCGAATTCTTAC). These adapters contained either one 3′ protruding nucleotide, a blunt end, or one 5′ protruding nucleotide. In each ligation reaction, 50 ng of EBV virion DNA was incubated overnight with 5 ng of annealed adapter. PCRs were performed with one of the two oligonucleotides, pB and pC, derived from EBV TR and with one of the two oligonucleotides, F and G, from the adapters. The PCR products were cloned into pBluescript KS(−) and sequenced.

RESULTS

Sp1 binds specifically to the EBV TR sequences. Preliminary experiments that explored the biochemical properties of TRBP, including its DNA binding specificity, its requirement for Zn^{2+} in DNA binding, its binding to wheat germ agglutinin, and its behavior on gel filtration columns, suggested that TRBP might be related to the Sp1 family of transcriptional activators (43). The behaviors of TRBP and Sp1 were directly compared in two DNA binding assays (Fig. 1). DNase footprinting showed that nuclear extract protected three regions on the EBV TR designated TRBS1, TRBS2, and TRBS3 (Fig. 1A). Sp1 protein formed a footprint on the TRBS1 and TRBS2 sites but did not bind TRBS3. TRBS1 and TRBS2 bound the same protein. They competed with each other for binding in footprinting assays and in EMSAs (data not shown). TRBS3 seemed to bind a different protein. It did not compete with TRBS1 or TRBS2 for the binding of TRBP.

Oligonucleotide competitions confirmed that TRBP and Sp1 bound related sequences (Fig. 1B). With nuclear extract as a source of binding activity, three complexes, designated A, B, and C, formed on oligonucleotide TRBS1 derived from the EBV TR. Complexes of mobility identical to these formed by nuclear extracts from Raji cells were formed with nuclear extracts from every mammalian cell tested, including P3HR-1/CL16(HH514-16), Daudi, B95-8, BJAB, and a variety of EBV-negative Burkitt lymphoma cell lines. Complexes A and B were competed by an oligonucleotide containing a consensus Sp1 site from the simian virus 40 (SV40) promoter (Fig. 1B, lanes 3 to 7). Similarly, complexes A and B formed by nuclear extract on the SV40 Sp1 site were competed by TRBS1 (Fig. 1B, lanes 19 to 23). Based on these competitions, the affinity of TRBP for TRBS1 was about twofold greater than that for the Sp1 site from SV40. In competition EMSA experiments with purified Sp1 protein with labelled TRBS1 or Sp1 DNA probes, TRBS1 competed for the binding of Sp1 more efficiently than did the Sp1 site (data not shown). Complexes C and D appear to consist of factors that bind DNA nonspecifically (Fig. 1B and 2A). Complex C, which was not effectively competed by Sp1 oligonucleotide, was shown to contain the Ku autoantigen (40a).

To pinpoint the sequence within TRBS1 essential for the binding of TRBP or Sp1, a series of mutant oligonucleotides were used in EMSAs (Fig. 2). Mutations that decreased TRBP binding decreased Sp1 binding concordantly. Mutants Tm1A, Tm1B, Tm1C, and Tm3, containing interruptions in the G tracts of TRBS1, all had reduced affinity for TRBP (Fig. 2A). No complex A was formed with Tm1C or Tm3. Mutant Tm2, in which two T nucleotides were replaced with A's, retained wild-type affinity for complexes A and B. Figure 2B shows that Sp1 protein bound TRBS1 (T); however, only complex A was formed. Mutations of TRBS1 caused similar effects on the binding of TRBP and purified Sp1 (compare Fig. 2A and 2B). Rabbit antibody against Sp1 supershifted complex A formed on the wild type and the Tm2 mutant as well as the less abundant complex A formed on Tm1A and Tm1B (Fig. 2C). However, anti-Sp1 antibody did not recognize the unrelated complexes that formed on Tm1C and Tm3.

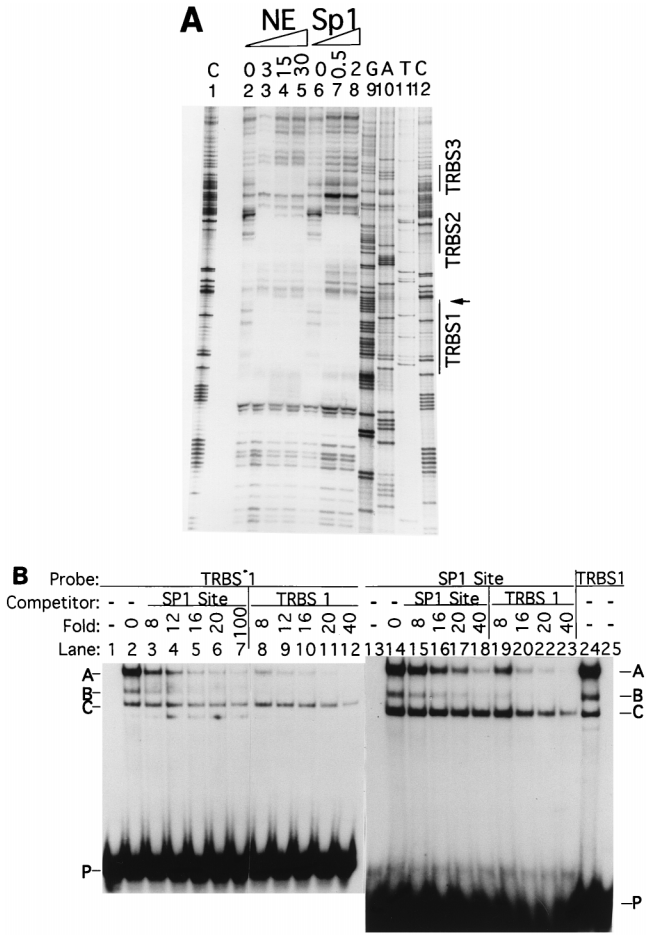


FIG. 1. Binding of Sp1 to EBV TR sequences. (A) DNase I footprinting on EBV TR sequences by nuclear extract (NE) or purified Sp1 protein. Lanes 2 and 6, DNase I digestion of the probe in the absence of nuclear extract or Sp1 protein; lanes 3 to 5, incubation of the probe with 3, 15, and 30 μl of nuclear extract before DNase digestion; lanes 7 and 8, incubation of the probe with 0.5 and 2 FPU of pure Sp1 protein before digestion. Nuclear extract for this experiment was prepared from P3HR-1/CL16(HH514-16) cells. The probe corresponds to EBV nt 170538 to 170819. An oligonucleotide corresponding to the 5′ end of the probe was used as a primer in a DNA sequencing reaction. The sequencing reaction and the DNase I protection assay were run together to determine the protected sequences (lanes 1 and 9 to 12). The protected sequences (TRBS1, TRBS2, and TRBS3) are shown in Fig. 7. In lanes 4 and 5, a hypersensitive site corresponding to the right end of the EBV genome is indicated by an arrow. (B) Competition for binding of TRBP by an Sp1 binding site or TR binding site 1 (TRBS1). Lanes 1 and 25, ^{32}P-labelled TRBS1 probe alone (EBV nt 170628 to 170644) (5); lane 2, labelled TRBS1 probe with no specific competitor added; lanes 3 to 7, ^{32}P-labelled TRBS1 competed with increasing amounts of unlabelled oligonucleotides containing the Sp1 binding site; lanes 8 to 12, ^{32}P-labelled TRBS1 probe competed with unlabelled TRBS1; lane 13, ^{32}P-labelled Sp1 probe alone; lane 14, labelled Sp1 probe with no specific competitor added; lanes 15 to 18, ^{32}P-labelled Sp1 probe competed with unlabelled Sp1 oligonucleotide; lanes 19 to 23, ^{32}P-labelled Sp1 probe competed with unlabelled TRBS1 oligonucleotide; lane 24, labelled TRBS1 probe with no specific competitor added. "Fold" indicates the ratio of excess cold competitor to the labelled probe added into the binding reaction mixture. Shifted complexes from the top of the gel to the bottom are labelled A, B, and C. P, free probe. Nuclear extract for this experiment was prepared from Raji cells.

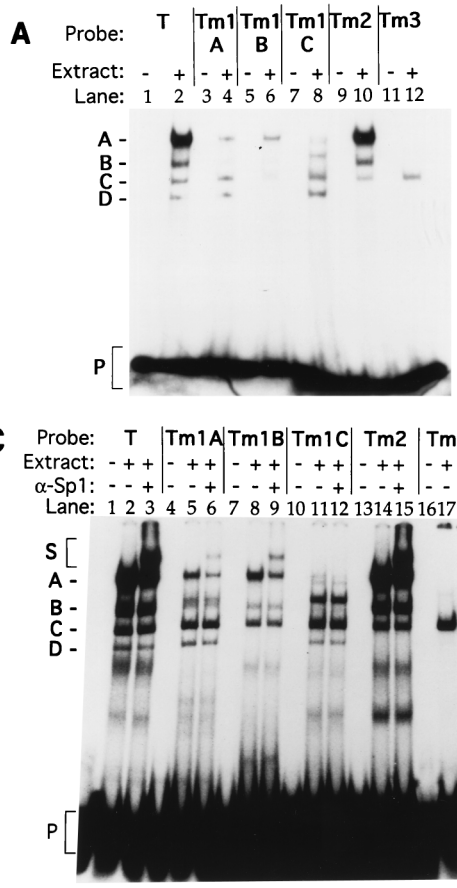
Sp1 binds to other EBV and cellular recombinogenic sequences. The binding of TRBP and pure Sp1 was compared by using oligonucleotides T (TRBS1), W, and H, which represent recombinogenic sequences in the EBV genome, and oligonucleotides Sγ3c and VNTR, which represent recombinogenic sequences in the cellular genome (Table 1; Fig. 3). W and H represent the sites of recombination that generate the deletion in P3J-HR-1 DNA. Sp1 bound all the viral and cellular sequences that were bound by TRBP. In each instance, Sp1 protein formed a complex that comigrated with complex A formed by TRBP. Anti-Sp1 antibody supershifted TRBP complex A on all the viral and cellular sequences.

Sp1 is an autoantigenic component of TRBP. Eight autoimmune sera from different patients recognized TRBP; all of them also reacted with Sp1 in a supershift EMSA and in immunoprecipitation experiments (data not shown). For example, Fig. 4A shows that autoimmune serum 229, which caused a supershift of TRBP in nuclear extract (lanes 7 and 17), also supershifted purified Sp1 (lanes 2 and 12). This autoantibody supershifted TRBP or Sp1 whether an Sp1 binding site or TRBS1 was used as a probe in the EMSA. Figure 4B shows that this autoantibody, which was reactive with TRBP and Sp1 in EMSA, also recognized purified Sp1 protein on an immunoblot (lane 8). In cell extracts (Fig. 4B, lanes 3, 4, and 7), this antibody recognized a 95-kDa protein corresponding in size to Sp1 (lanes 6 and 8).

Mapping the ends of the EBV linear genome. We reasoned that if Sp1 played a role in EBV terminal processing, the cleavage site should be close to, or at, the binding sites for Sp1. At the time we initiated these experiments, the cleavage site for EBV linearization had not been precisely mapped (43, 44, 46). Therefore, we mapped the ends of the EBV linear genome, initially using primer extension. Virion DNA was obtained from viral particles purified on sucrose gradients. After denaturation, viral DNA was annealed with primers located at several positions on the EBV TR. Primer C, which was used to map the right end of the genome, generated an unambiguous pattern. The major extended product was 114 nucleotides (nt), and a minor product was 113 nt (Fig. 5A). The longest extended product from primer B, which was used to map the left end, was 171 nt (Fig. 5B). Based on the results of additional experiments, described below, the shorter fragments were likely the result of premature stops of polymerization.

The EBV ends were also mapped by Southern blotting (Fig. 5C and D). Virion DNA was digested with restriction endonucleases which cut within the terminal repeat. *Apa*I was used to map the right end, and *Ava*II was used to map the left end. The digested products were separated on 6% denaturing polyacrylamide gels, transferred to nitrocellulose paper, and probed with strand-specific oligonucleotides which were located between the restriction enzyme cleavage site and the putative site of terminal processing of EBV DNA. The right termini were again found to be heterogeneous, since doublets were detected on the Southern blots. The 3′ end of the right terminus was 86 or 85 nt away from the *Apa*I site (Fig. 5C). The 5′ end was 89 or 88 nt away from the *Apa*I site. Considering that *Apa*I digestion generates a 4-nt 3′ overhang, the right end of the EBV genome appears to have one extra nucleotide at the 3′ end. This estimate was based on the assumption that the migration rate of the DNA fragments was dependent solely on the number of nucleotides and was not affected by base composition. A single band of 117 nt represented the length between the 5′ end of the left terminus and the *Ava*II site (Fig. 5D). The distance between the 3′ end of the left terminus and the *Ava*II site was 120 nt. Since *Ava*II digestion generates a 3-nt 5′ overhang, the left terminus may be blunt.

Unusual sequence at the EBV right terminus. The results of both primer extension and high-resolution Southern blotting indicated that there was approximately 20 bp of extra sequence present in the linear termini that was not present in the circular genome. A ligation-mediated PCR method was used to identify these extra sequences in the EBV termini. Unlike other



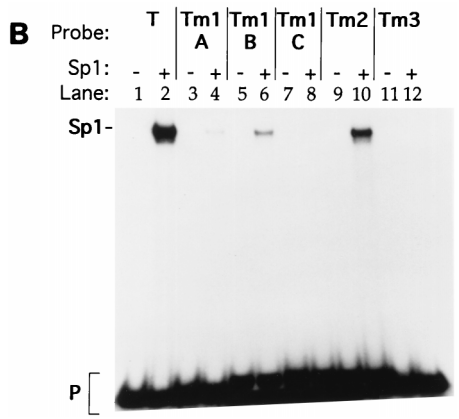


FIG. 2. Electrophoretic mobility shifts with TRBS1 mutant oligonucleotides. (A) Shift of wild-type and mutant oligonucleotides with a nuclear extract from Raji cells. (B) Shift of mutant oligonucleotides with 0.1 FPU of purified Sp1 protein. (C) Supershift of complexes formed between Raji cell nuclear extract (3.25 μg of total protein) and mutant oligonucleotides with 1 μl of rabbit anti-Sp1. Probe T is an oligonucleotide containing the TRBS1 sequence. A, B, C, and D denote complexes from the top of the gel to the bottom. S, supershifted complex; P, free probe.

methods of DNA cloning, this method does not require modification of the ends. Three types of adapters were ligated to the left or right terminus of virion DNA: one contained a 1-nt 3′ overhang, one was blunt, and one contained a 1-nt 5′ overhang. Separate PCRs were performed with each ligation, using one primer in the TR and another primer in the adapter. Three clones with lengths consistent with the results of primer extension and Southern blotting were obtained from PCR products resulting from ligation of the right terminus, with the adapter containing 1 nt (cytosine) protruding at the 3′ end of the bottom strand. These three independent clones contained identical inserts, the sequence of one of which is shown in Fig. 6. The sequence revealed the identity of the extra 20 bp found by primer extension and Southern blotting.

Linearization and circularization of the EBV genome are novel DNA recombination processes. The termini of the encapsidated linear EBV genome in the virion were recently characterized by end labelling and by direct cloning and sequencing (56). The three different methods we used, primer extension, high-resolution Southern blotting, and ligation-mediated PCR, confirmed these results (44). Linearization of the viral genome is not due to simple cleavage of the circular or concatemeric form; instead, the conversion of the linear to the circular form and the reverse involved a novel form of DNA processing (Fig. 7). Formation of the right terminus of linear EBV DNA was associated with the insertion of 9 bp (GTGT TGGCG) and the deletion of 1 bp (T) immediately adjacent to the insertion. The right terminus of the EBV genome was coterminal with the right end of TRBS1, one of the two major binding sites for Sp1 (Fig. 1). The left terminus was in the middle of TRBS1. According to previously published results, the sequence of the left terminus of EBV DNA contains 11 bp (GGGGCATGGGG) that are also present at the right end of the genome (56).

DISCUSSION

We show that Sp1, a well-characterized transcriptional activator, binds to recombination junction sequences in viral and cellular DNA. Sp1 binds to TRBS1, the site of EBV terminal processing. Sp1 binds to other recombinogenic sequences in EBV and cellular DNA which contain the common motif GGGTGG. The sequence TRBS1 contains three G_n tracts, all

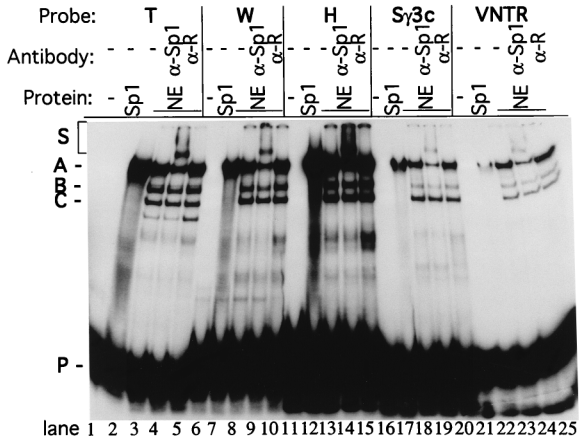


FIG. 3. Complexes between TRBP or Sp1 and EBV or cellular recombination junctions are supershifted by antibody to Sp1 (α-Sp1). Double-stranded oligonucleotides T (TRBS1) (lanes 1 to 5), W (2907 to 2927 in EBV *Bam*HI W) (lanes 6 to 10), H (3586 to 3613 in EBV *Bam*HI H) (lanes 11 to 15), Sγ3c from the immunoglobulin heavy-chain γ3 switch region, and VNTR common sequences (lanes 21 to 25) were labelled with ^{32}P. Probes were shifted with 0.1 FPU of Sp1 protein (lanes 2, 7, 12, 17, and 22) or with 3.25 μg of total protein of Raji cell nuclear extract (NE) (lanes 3 to 5, 8 to 10, 13 to 15, 18 to 20, and 23 to 25). One microliter of anti-Sp1 (lanes 4, 9, 14, 19, and 24) was added to binding reactions to supershift Sp1-containing complexes on the various probes. Rabbit antibody against the EBV immediate-early protein BRLF1 (α-R) was added as a negative control (lanes 5, 10, 15, 20, and 25).

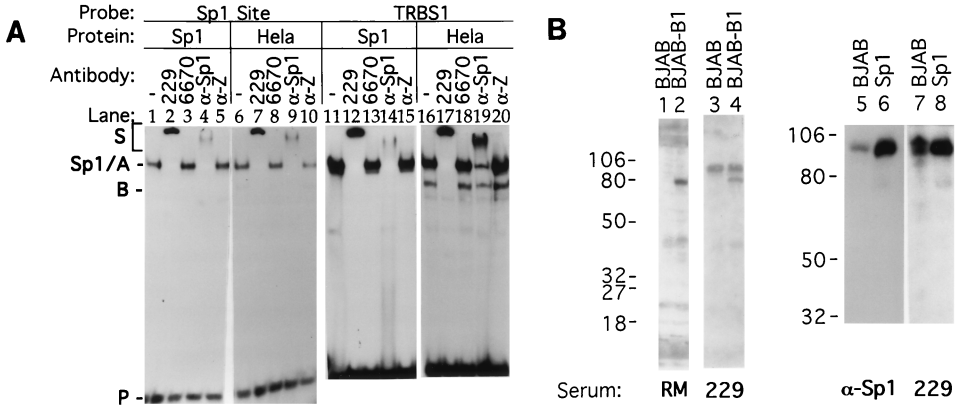


FIG. 4. Sp1 is an autoantigenic component of TRBP. (A) Supershifts of TRBP-DNA complexes and Sp1-DNA complexes by autoimmune sera. Lanes 1 to 10 contain the ^{32}P-labelled Sp1 binding site probe. Lanes 11 to 20 contain the ^{32}P-labelled TRBS1 probe. Lanes 1 to 5 and 11 to 15 contain 0.1 FPU of Sp1 protein in each binding reaction mixture. Lanes 6 to 10 and 16 to 20 contain approximately 1.9 μg of total protein from a HeLa cell nuclear extract. Lanes 1, 6, 11, and 16, no antibody; lanes 2, 7, 12, and 17, 1 μl of human autoimmune serum 229 (anti-TRBP positive); lanes 3, 8, 13, and 18, 1 μl of human autoimmune serum 6670 (anti-TRBP negative); lanes 4, 9, 14, and 19, 1 μl of rabbit anti-Sp1 (α-Sp1); lanes 5, 10, 15, and 20, 1 μl of rabbit anti-EBV immediate-early protein BZLF1 (α-Z). Sp1/A indicates the Sp1 shifted band and the top band in nuclear extract binding reactions. In lanes 6 to 10, complex B is visible on a longer exposure of the autoradiograph. (B) Recognition of Sp1 protein by an autoimmune serum on an immunoblot. Cellular extract from EBV-negative B-cell lymphoma line BJAB (lanes 1, 3, 5, and 7), the corresponding EBV (B95-8) converted cell line BJAB-B1 (lanes 2 and 4), and purified Sp1 protein (lanes 6 and 8) were electrophoresed on an 8% denaturing SDS-polyacrylamide gel. After transfer, the filters were probed with human serum RM (lanes 1 and 2), autoimmune serum 229 (lanes 3, 4, 7, and 8), and anti-Sp1 serum (lanes 5 and 6). The 80-kDa protein in the BJAB-B1 lanes is EBV nuclear antigen 1. The numbers at left are molecular masses, in kilodaltons.

of which are required for optimal binding by Sp1. Additional mutagenesis confirmed this conclusion (40a). Binding of Sp1 to recombinogenic sequences suggests a novel function for the protein in the process of DNA recombination.

The impetus for this study was the observation of a recombination event between EBV and SMRV (43, 46). The sequence at the recombination crossover junction between EBV and SMRV was TGGGTGG. A similar sequence, GGGTG GG, has recently been found to promote recombination in a functional assay with avian spleen necrosis retrovirus (33).

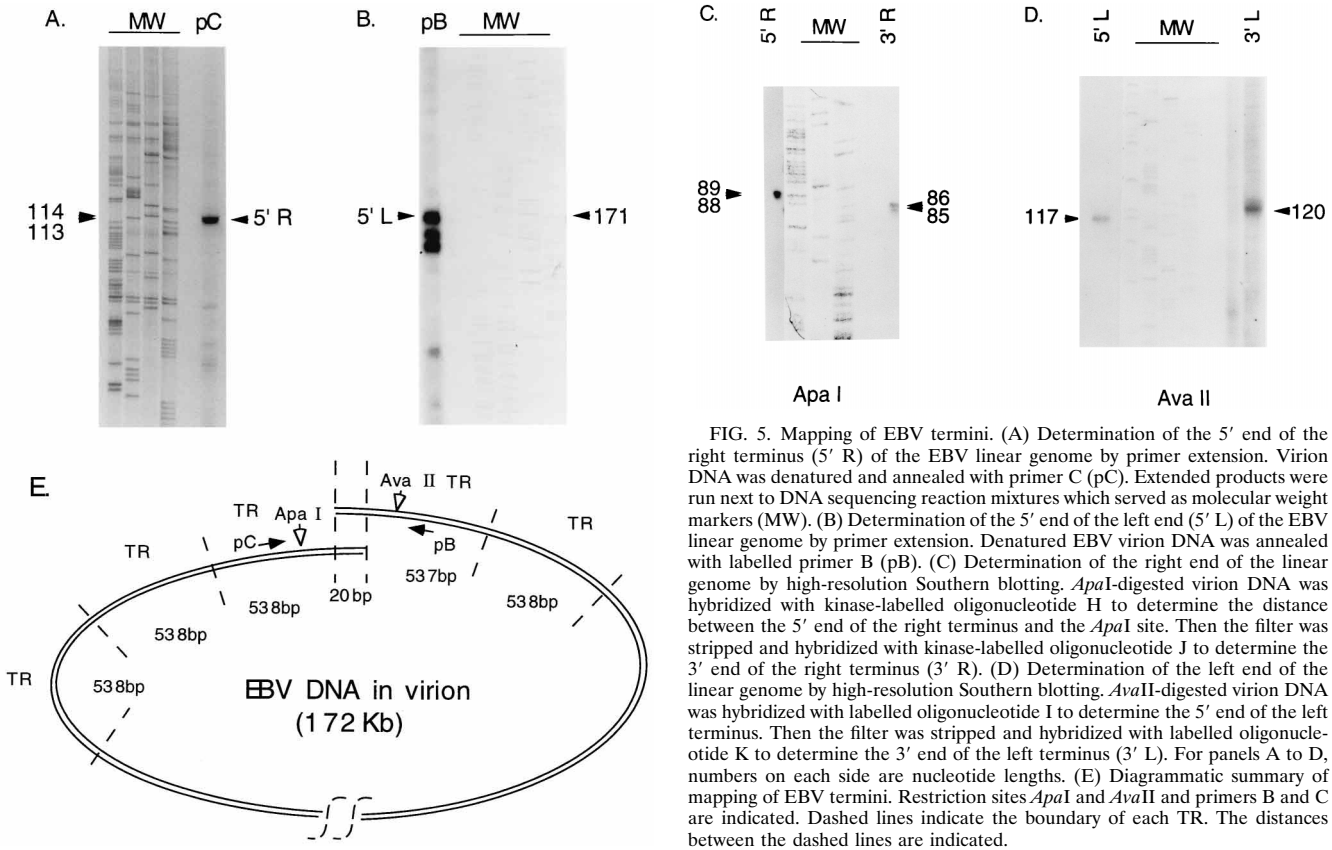


FIG. 5. Mapping of EBV termini. (A) Determination of the 5′ end of the right terminus (5′ R) of the EBV linear genome by primer extension. Virion DNA was denatured and annealed with primer C (pC). Extended products were run next to DNA sequencing reaction mixtures which served as molecular weight markers (MW). (B) Determination of the 5′ end of the left end (5′ L) of the EBV linear genome by primer extension. Denatured EBV virion DNA was annealed with labelled primer B (pB). (C) Determination of the right end of the linear genome by high-resolution Southern blotting. *Apa*I-digested virion DNA was hybridized with kinase-labelled oligonucleotide H to determine the distance between the 5′ end of the right terminus and the *Apa*I site. Then the filter was stripped and hybridized with kinase-labelled oligonucleotide J to determine the 3′ end of the right terminus (3′ R). (D) Determination of the left end of the linear genome by high-resolution Southern blotting. *Ava*II-digested virion DNA was hybridized with labelled oligonucleotide I to determine the 5′ end of the left terminus. Then the filter was stripped and hybridized with labelled oligonucleotide K to determine the 3′ end of the left terminus (3′ L). For panels A to D, numbers on each side are nucleotide lengths. (E) Diagrammatic summary of mapping of EBV termini. Restriction sites *Apa*I and *Ava*II and primers B and C are indicated. Dashed lines indicate the boundary of each TR. The distances between the dashed lines are indicated.

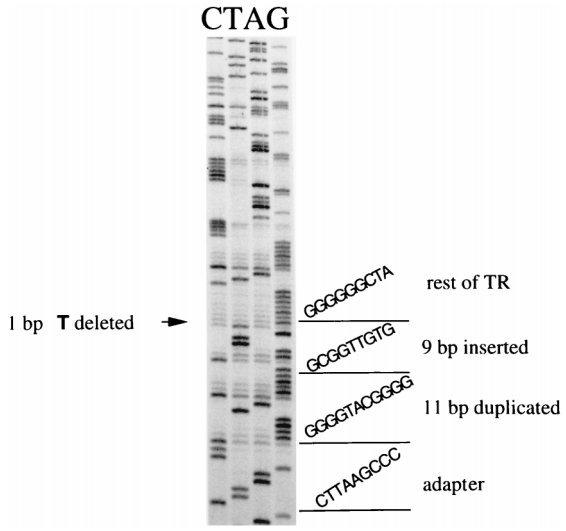


FIG. 6. Sequence of the right terminus of the EBV genome. The sequence of the ligation-mediated-PCR product shows the unusual processing at the right end of the EBV genome. The adapter sequence was linked to the duplicated sequence, followed by the insertion of 9 bp. A base pair (T) is deleted before the sequence continues with the rest of the TR.

Thus, recombination junctions of a primate and an avian retrovirus contain related G-rich elements. Moreover, similar sequences are found at internal recombination junctions in EBV, such as the recombination junction of the *EBNA2* deletion in EBV strain P3J-HR-1 and the EBV termini (Fig. 5). The appearance of this motif in recombinogenic loci of unrelated viruses supports the hypothesis of a functional role for such sequences and suggests that they may interact with similar cellular proteins.

Formation of the ends of EBV DNA is not a simple cleavage and ligation reaction but an intricate process (Fig. 5 to 7). The total number of TR may be increased or decreased in the linear virion DNA relative to their number in the circular latent DNA (19). A 9-bp sequence, GTGTTGGCG, also found in unique sequences adjacent to TR, is newly added to the right terminus. An 11-bp sequence, GGGGCATGGGG, found only once in each fused TR of circular DNA, is duplicated at both termini of linear DNA (56). Moreover, this 11-bp sequence is embedded in an 18-bp sequence which is located 37 bp from a nearly perfect inverted repeat (Fig. 7). A nucleotide, T, present in circular DNA is absent from linear DNA. All these findings point to the likely participation of DNA replication, recombination, and repair in the interconversion of the circular and linear forms of EBV DNA. The recombination events discussed here are to be distinguished from the only well-characterized eukaryotic recombination reaction in higher eukaryotes, namely, V(D)J recombination, which does not involve DNA replication (40, 51). High rates of deletions, insertions, and mutations, found close to the recombination junctions in the spleen necrosis virus, likewise implicate DNA replication and repair in this recombination process (33).

Sp1 could play a role in either circularization or linearization of the EBV genome. Both these events are likely to involve DNA recombination processes. Several functions of Sp1 in recombination can be envisioned. Sp1 has been shown to bring two regions of double-stranded DNA together, either apposing ends or forming loops (26, 41). This represents an essential step in recombination and could play a role in circularization of the input genome. Sp1 can bend DNA (13, 42). Distortion of DNA near a point of cleavage may promote recombination. For example, the binding of Sp1 protein to TRBS1 and TRBS2 together with another protein on TRBS3 may induce conformational changes to direct the cleavage on the EBV TR. Finally, Sp1 can be phosphorylated by the complex of autoantigen Ku and DNA-dependent protein kinase, which are involved in DNA repair and V(D)J recombination (6, 10, 11,

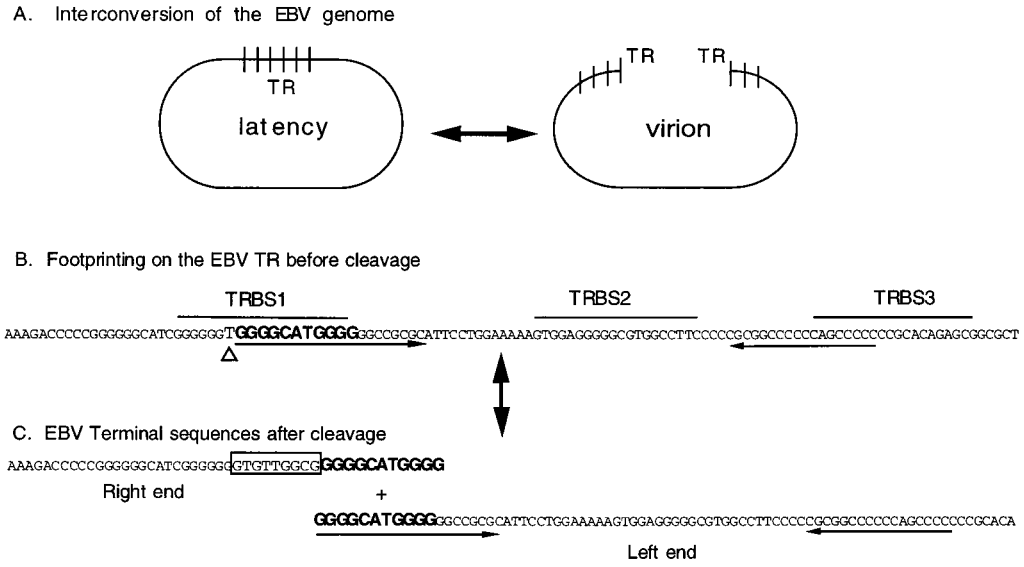


FIG. 7. Summary of EBV terminal processing. (A) Diagram of the interconversion of the EBV genome between the linear form in virions and the circular form in cells during latency. Short vertical bars represent the boundary between each direct TR. The linearization of EBV DNA results in molecules with variable numbers of TR at each end. (B) Partial sequence of EBV TR before cleavage. The footprinting areas, TRBS1, TRBS2, and TRBS3, are indicated (Fig. 1A). The arrows below the sequences indicate nearly perfect inverted repeats. One T, indicated by a triangle, is deleted during linearization. (C) EBV sequence at the termini after cleavage. The sequences in bold are present on both termini after cleavage, based on data shown in Fig. 5 and described in reference 56. The boxed sequences are introduced after cleavage and can be found at the junction between the short unique sequence and the TR (EBV nt 172232 to 172240) (5).

20, 40, 47, 49). Sp1 may attract components of the recombination machinery, such as Ku and DNA-dependent protein kinase, to sites with the (GGGTGG) recombinogenic motif. Other recent evidence implicates components of the transcription apparatus in herpesvirus genome end processing. *$TAF_{II}250$*, which is known to play a role in Sp1 transcriptional activation, has recently been found to play a role in genome end formation of herpes simplex virus (25, 50, 54). In summary, Sp1 binds to the EBV termini at a locus intimately associated with a novel form of DNA recombination.

ACKNOWLEDGMENTS

We are grateful to J. Craft for providing autoimmune sera and to P. Olson and R. Dornburg for communicating results before publication. J. Konner and S. Katz participated in this study as Yale University undergraduate students.

This work was supported by a grant from the NIH (CA16038) to G.M., a MARC predoctoral fellowship from the NIAID (GM14637) to T.A.S., and a training grant from the NIH (T32CA09159) and a Swebilius Cancer Research award to R.S.

REFERENCES

1. **Aboul-ela, F., A. I. H. Murchie, and D. M. J. Lilley.** 1992. NMR study of parallel-stranded tetraplex formation by the hexadeoxynucleotide d(TG_4T). Nature **360:**280–282.
2. **Adams, A., and T. Lindahl.** 1975. Epstein-Barr virus genomes with properties of circular DNA molecules in carrier cells. Proc. Natl. Acad. Sci. USA **72:**1477–1481.
3. **Ashley, T.** 1994. Mammalian meiotic recombination: reexamination. Hum. Genet. **94:**587–593.
4. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.).** 1995. Current protocols in molecular biology. John Wiley & Sons, New York, N.Y.
5. **Baer, R., A. T. Bankier, M. D. Biggin, P. L. Deininger, P. J. Farrell, T. J. Gibson, G. Hatfull, G. S. Hudson, S. C. Satchwell, C. Seguin, et al.** 1984. DNA sequence and expression of the B95-8 Epstein-Barr virus genome. Nature **310:**207–211.
6. **Blunt, T., N. J. Finnie, G. E. Taccioli, G. C. Smith, J. Demengeot, T. M. Gottlieb, R. Mizuta, A. J. Varghese, F. W. Alt, P. A. Jeggo, et al.** 1995. Defective DNA-dependent protein kinase activity is linked to V(D)J recombination and DNA repair defects associated with the murine scid mutation. Cell **80:**813–823.
7. **de Bruin, D., M. Lanzer, and J. V. Ravetch.** 1994. The polymorphic subtelomeric regions of Plasmodium falciparum chromosomes contain arrays of repetitive sequence elements. Proc. Natl. Acad. Sci. USA **91:**619–623.
8. **Dunnick, W., G. Z. Hertz, L. Scappino, and C. Gritzmacher.** 1993. DNA sequences at immunoglobulin switch region recombination sites. Nucleic Acids Res. **21:**365–372.
9. **Gompels, U. A., J. Nicholas, G. Lawrence, M. Jones, B. J. Thomson, M. E. Martin, S. Efstathiou, M. Craxton, and H. A. Macaulay.** 1995. The DNA sequence of human herpesvirus-6: structure, coding content, and genome evolution. Virology **209:**29–51.
10. **Gottlieb, T. M., and S. P. Jackson.** 1993. The DNA-dependent protein kinase: requirement for DNA ends and association with Ku antigen. Cell **72:**131.
11. **Hartley, K. O., D. Gell, G. C. Smith, H. Zhang, N. Divecha, M. A. Connelly, A. Admon, M. S. Lees, C. W. Anderson, and S. P. Jackson.** 1995. DNA-dependent protein kinase catalytic subunit: a relative of phosphatidylinositol 3-kinase and the ataxia telangiectasia gene product. Cell **82:**849–856.
12. **Hurley, E., and D. Thorley-Lawson.** 1989. B cell activation and the establishment of EBV latency. J. Exp. Med. **168:**2059–2075.
13. **Ikeda, K., K. Nagano, and K. Kawakami.** 1993. Possible implications of Sp1-induced bending of DNA on synergistic activation of transcription. Gene **136:**341.
14. **Jeang, K.-T., and S. D. Hayward.** 1983. Organization of the Epstein-Barr virus DNA molecule. III. Location of the P3HR-1 deletion junction and characterization of the *Not*I repeat units that form part of the template for an abundant 12-*O*-tetradecanoylphorbol-13-acetate-induced mRNA transcript. J. Virol. **48:**135–148.
15. **Jeffreys, A. J., A. MacLeod, K. Tamaki, D. L. Neil, and D. G. Monckton.** 1991. Minisatellite repeat coding as a digital approach to DNA typing. Nature **354:**204–209.
16. **Jeffreys, A. J., V. Wilson, and S. L. Thein.** 1985. Hypervariable 'minisatellite' regions in human DNA. Nature **314:**67–73.
17. **Kataoka, T., T. Kawakami, N. Takahashi, and T. Honjo.** 1980. Rearrangement of immunoglobulin γ-1-chain gene and mechanism. Proc. Natl. Acad. Sci. USA **77:**919–923.
18. **Kataoka, T., T. Miyata, and T. Honjo.** 1981. Repetitive sequences in class-switch recombination regions of immunoglobulin heavy chain genes. Cell **23:**357–368.
19. **Kintner, C. R., and B. Sugden.** 1979. The structure of the termini of the DNA of Epstein-Barr virus. Cell **17:**661–667.
20. **Kirchgessner, C. U., C. K. Patil, J. W. Evans, C. A. Cuomo, L. M. Fried, T. Carter, M. A. Oettinger, and J. M. Brown.** 1995. DNA-dependent kinase (p350) as a candidate gene for the murine SCID defect. Science **267:**1178–1183.
21. **Kishi, M., G. Bradley, J. Jessip, A. Tanaka, and M. Nonoyama.** 1991. Inverted repeat regions of Marek's disease virus DNA possess a structure similar to that of the *a* sequence of herpes simplex virus DNA and contain host cell telomere sequences. J. Virol. **65:**2791–2797.
22. **Krowczynska, A. M., R. A. Rudders, and T. G. Krontiris.** 1990. The human minisatellite consensus at breakpoints of oncogene translocations. Nucleic Acids Res. **18:**1121–1127.
23. **Laughlan, G., A. I. H. Murchie, D. G. Norman, M. H. Moore, D. M. J. Lilley, and B. Luisi.** 1994. The high-resolution crystal structure of a parallel-stranded guanine tetraplex. Science **265:**520–524.
24. **Liu, Z., and W. Gilbert.** 1994. The yeast KEM1 gene encodes a nuclease specific for G4 tetraplex DNA: implication of in vivo functions for this novel DNA structure. Cell **77:**1083–1092.
25. **Maldonado, E., R. Shiekhattar, M. Sheldon, H. Cho, R. Drapkin, P. Rickert, E. Lees, C. W. Anderson, S. Linn, and D. Reinberg.** 1996. A human RNA polymerase II complex associated with SRB and DNA-repair proteins. Nature **381:**86–89.
26. **Mastrangelo, I. A., A. J. Courey, J. S. Wall, S. P. Jackson, and P. V. Hough.** 1991. DNA looping and Sp1 multimer links: a mechanism for transcriptional synergism and enhancement. Proc. Natl. Acad. Sci. USA **88:**5670.
27. **Matsuo, T., M. Heller, L. Petti, E. O'Shiro, and E. Kieff.** 1984. Persistence of the entire Epstein-Barr virus genome integrated into human lymphocyte DNA. Science **226:**1322–1325.
28. **Mills, F. C., J. S. Brooker, and R. D. Camerini-Otero.** 1990. Sequences of human immunoglobulin switch regions: implications for recombination and transcription. Nucleic Acids Res. **18:**7305–7316.
29. **Mocarski, E. S., and B. Roizman.** 1982. Structure and role of the herpes simplex virus DNA termini in inversion, circularization and generation of virion DNA. Cell **31:**89–97.
30. **Nakamura, Y., M. Leppert, P. O'Connell, R. Wolff, T. Holm, M. Culver, C. Martin, E. Fujimoto, M. Hoff, E. Kumlin, and R. White.** 1987. Variable number of tandem repeat (VNTR) markers for human gene mapping. Science **235:**1616–1622.
31. **Nikaido, T., S. Nakai, and T. Honjo.** 1981. Switch region of immunoglobulin C-μ gene is composed of simple tandem repetitive sequences. Nature **292:**845–848.
32. **Obata, M., T. Kataoka, S. Nakai, H. Yamagishi, N. Takahashi, Y. Yamakawi-Kataoka, T. Nikaido, A. Shimizu, and T. Honjo.** 1981. Structure of a rearranged γ-1-chain gene and its implication to immunoglobulin class-switch mechanism. Proc. Natl. Acad. Sci. USA **78:**2437–2441.
33. **Olson, P., and R. Dornburg.** Capture of a recombination activating sequence from mammalian cell with homology to Chi. Submitted for publication.
34. **Petrini, J., and W. A. Dunnick.** 1989. Products and implied mechanism of H chain switch recombination. J. Immunol. **142:**2932–2935.
35. **Pluta, A. F., and V. A. Zakian.** 1989. Recombination occurs during telomere formation in yeast. Nature **337:**429–433.
36. **Raab-Traub, N., and K. Flynn.** 1986. The structure of the termini of the Epstein-Barr virus as a marker of clonal cellular proliferation. Cell **47:**883–889.
37. **Sato, H., T. Takimoto, S. Tanaka, J. Tanaka, and N. Raab-Traub.** 1990. Concatameric [sic] replication of Epstein-Barr virus: structure of the termini in virus-producer and newly transformed cell lines. J. Virol. **64:**5295–5300.
38. **Sen, D., and W. Gilbert.** 1988. Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. Nature **334:**364–366.
39. **Shimizu, A., and T. Honjo.** 1984. Immunoglobulin class switch. Cell **36:**801–803.
40. **Smider, V., W. K. Rathmell, M. R. Lieber, and G. Chu.** 1994. Restoration of X-ray resistance and V(D)J recombination in mutant cells by Ku cDNA. Science **266:**288.

40a. **Spain, T. A., R. Sun, and G. Miller.** Submitted for publication.

41. **Su, W., S. Jackson, R. Tjian, and H. Echols.** 1991. DNA looping between sites for transcriptional activation: self-association of DNA-bound Sp1. Genes Dev. **5:**820.
42. **Sun, D., and L. H. Hurley.** 1994. Cooperative bending of the 21-base-pair repeats of the SV40 viral early promoter by human Sp1. Biochemistry **33:**9578–9587.
43. **Sun, R.** 1993. Ph.D. thesis. Yale University, New Haven, Conn.
44. **Sun, R., et al.** 1994. Presented at the 6th International Symposium of EBV and Associated Diseases, Cold Spring Harbor, N.Y.
45. **Sun, R., E. Grogan, D. Shedd, A. F. Bykovsky, V. M. Kushnaryov, S. E. Grossberg, and G. Miller.** 1995. Transmissible retrovirus in Epstein-Barr virus-producing B95-8 cells. Virology **209:**374–383.

46. **Sun, R., T. A. Spain, S.-F. Lin, and G. Miller.** 1994. Autoantigenic proteins that bind recombinogenic sequences in Epstein-Barr virus and cellular DNA. Proc. Natl. Acad. Sci. USA **91:**8646–8650.
47. **Taccioli, G. E., T. M. Gottlieb, T. Blunt, A. Priestley, J. Demengeot, R. Mizuta, A. R. Lehmann, F. W. Alt, S. P. Jackson, and P. A. Jeggo.** 1994. Ku80: product of the XRCC5 gene and its role in DNA repair and V(D)J recombination. Science **265:**1442–1445.
48. **Thyphronitis, G., G. C. Tsokos, C. H. June, A. D. Levine, and F. D. Finkelman.** 1989. IgE secretion by Epstein-Barr virus-infected purified human B lymphocytes is stimulated by interleukin 4 and suppressed by interferon γ. Proc. Natl. Acad. Sci. USA **86:**5580–5584.
49. **Tuteja, N., R. Tuteja, A. Ochem, P. Taneja, N. W. Huang, A. Simoncsits, S. Susic, K. Rahman, L. Marusic, J. Chen, et al.** 1994. Human DNA helicase II: a novel DNA unwinding enzyme identified as the Ku autoantigen. EMBO J. **13:**4991–5001.
50. **Umene, K., and T. Nishimoto.** 1996. Inhibition of generation of authentic genomic termini of herpes simplex virus type 1 DNA in temperature-sensitive mutant BHK-21 cells with a mutated *CCG1*/*TAF$_{II}$250* gene. J. Virol. **70:**9008–9012.
51. **van Gent, D. C., D. A. Ramsden, and M. Gellert.** 1996. The Rag1 and Rag2 proteins establish the 12/23 rule in V(D)J recombination. Cell **85:**107–113.
52. **Wahls, W. P., L. J. Wallace, and P. D. Moore.** 1990. Hypervariable minisatellite DNA is a hotspot for homologous recombination in human cells. Cell **60:**95–103.
53. **Wang, S., and V. A. Zakian.** 1990. Telomere-telomere recombination provides an express pathway for telomere acquisition. Nature **345:**456–458.
54. **Weinzierl, R. O. J., B. D. Dynlacht, and R. Tjian.** 1993. Largest subunit of Drosophila transcription factor IID directs assembly of a complex containing TBP and a coactivator. Nature **362:**511–517.
55. **Wu, T. T., M. Reid-Miller, H. M. Perry, and E. A. Kabat.** 1984. Long identical repeats in the mouse γ2b switch region and their implications for the mechanism of class switching. EMBO J. **3:**2033–2040.
56. **Zimmermann, J., and W. Hammerschmidt.** 1995. Structure and role of the terminal repeats of Epstein-Barr virus in processing and packaging of virion DNA. J. Virol. **69:**3147–3155.