MOLECULAR AND CELLULAR BIOLOGY, June 1989, p. 2738–2742
0270-7306/89/062738-05$02.00/0

Vol. 9, No. 6

High-Resolution Footprints of the DNA-Binding Domain of Epstein-Barr Virus Nuclear Antigen 1

AMY S. KIMBALL,[1] GREGORY MILMAN,[2] AND THOMAS D. TULLIUS[1,3]*

Department of Biology and the McCollum-Pratt Institute,[1] and Department of Chemistry,[3] The Johns Hopkins University, Baltimore, Maryland 21218, Department of Biochemistry, School of Hygiene and Public Health, The Johns Hopkins University, Baltimore, Maryland 21205[2]

Received 7 November 1988/Accepted 9 March 1989

The DNA-binding domain of Epstein-Barr virus nuclear antigen 1 was found by hydroxyl radical footprinting to protect backbone positions on one side of its DNA-binding site. The guanines contacted in the major groove by the DNA-binding domain of Epstein-Barr virus nuclear antigen 1 were identified by methylation protection. No difference was found in the interaction of the DNA-binding domain of Epstein-Barr virus nuclear antigen 1 with tandemly repeated and overlapping binding sites.

The Epstein-Barr virus genome is maintained as a low-copy-number plasmid in latently infected host cells. The control of plasmid copy number requires 1,800 base pairs (bp) of the viral genome (*oriP*) and the 80-kilodalton protein Epstein-Barr virus nuclear antigen 1 (EBNA-1) (18). Two regions within *oriP* are essential for activity (Fig. 1A): the tandem repeat region, consisting of 20 30-bp-long tandemly repeated sequences, and the dyad symmetry region, made up of four homologous overlapping repeats (12) (Fig. 1C). *oriP*-dependent DNA replication requires the presence of 12 of the 20 30-bp repeats in the tandem repeat region and the 3 homologous repeats included in the 65 bp of dyad symmetry (11, 17).

Rawlins et al. (10) showed that a 25-kilodalton fusion protein containing the carboxyl-terminal one-third of EBNA-1 (we call this portion of EBNA-1 EBNA in this paper) is capable of binding to the tandem repeat region, to the dyad symmetry region, and to a third site in the Epstein-Barr virus genome outside of *oriP*. All three of these regions contain homologous sequences, but in the tandem repeat region, the centers of the sequences are spaced at 30-bp intervals, while in the dyad symmetry region, the centers of the overlapping repeats are only 21 bp apart (Fig. 1C).

In this paper, we characterize the interaction of EBNA with its DNA-binding sites by using two chemical techniques, hydroxyl radical footprinting (15) and methylation protection (14).

Plasmids, DNA molecules, and EBNA. Plasmids pR1, containing the single EBNA-1 consensus binding site (Fig. 1B), and pR3, containing three tandemly repeated binding sites (7), were purified as previously described (16). DNA fragments (Fig. 1D and E) labeled with ^{32}P were generated by 5′ or 3′ labeling of the *Hin*dIII site (5) followed by *Eco*RI digestion and gel purification (6).

The native dyad symmetry and tandem repeat regions of Epstein-Barr virus are contained in the 2,185-bp pHEBO-1 restriction fragment of Epstein-Barr virus (10). This fragment was cloned into the *Hin*cII site of pGEM blue (Promega Biotec) to form pGEMori, and the plasmid was CsCl purified (5). The dyad symmetry region was labeled at a *Hin*dIII site, digested with *Oxa*NI, and gel purified (Fig. 1C).

A fusion protein made up of 191 amino acids from the carboxyl terminus of EBNA-1 and 36 amino acids from the bacteriophage λ N protein and linker was expressed in *Escherichia coli* and purified (8). The EBNA fusion polypeptide was originally named 28K-EBNA on the basis of the predicted molecular weight (10). However, the fusion polypeptide appears to start at the second methionine of the open reading frame, giving a predicted molecular weight of 24,532. Here we refer to this protein as EBNA.

Hydroxyl radical footprinting. EBNA was bound to DNA by combining 60 to 200 fmol (30 to 100 kcpm) of end-labeled DNA with 200 to 400 pmol of EBNA in a total volume of 35 μl in a buffer consisting of 10 mM Tris hydrochloride (pH 8.0), 5 mM $MgCl_2$, 200 mM NaCl, 1 mM dithiothreitol, 12.5 μg of bovine serum albumin per ml, and 20 μg of sonicated calf thymus DNA per ml. The mixture was incubated at room temperature for 30 to 90 min. Hydroxyl radical cleavage reactions were done as described previously (15).

Methylation protection. EBNA was bound to DNA by combining 30 to 100 fmol of DNA (15 to 50 kcpm) with 100 to 200 pmol of EBNA in a total volume of 10 μl in the buffer described above and incubated at room temperature for 30 min before addition of 200 μl of the Maxam and Gilbert (6) guanine-specific sequencing reaction buffer. The methylation reaction was initiated at 4°C by adding 1 μl of dimethyl sulfate and was terminated after 1 min (5).

Samples were resolved by electrophoresis through an 8% denaturing gel. The gel was autoradiographed and scanned as previously described (15).

The EBNA-1 synthetic consensus binding site is shown in Fig. 1B, with each nucleotide position numbered so that symmetry-related positions have the same number; positions on the bottom strand are denoted with primes. The 29-bp-long site is made up of a central palindrome (nucleotides 9 through 20, with the dyad axis between nucleotides 14 and 15) and nonsymmetrical outer sequences. The hydroxyl radical footprint of EBNA bound to a cloned, synthetic single consensus binding site (Fig. 2A and B) shows four symmetry-related groups of strongly protected nucleotides. The centers of the two protected sites on each strand are 10 nucleotides apart and are flanked by nucleotides that are cleaved as readily as in naked DNA. These four stretches of

* Corresponding author.

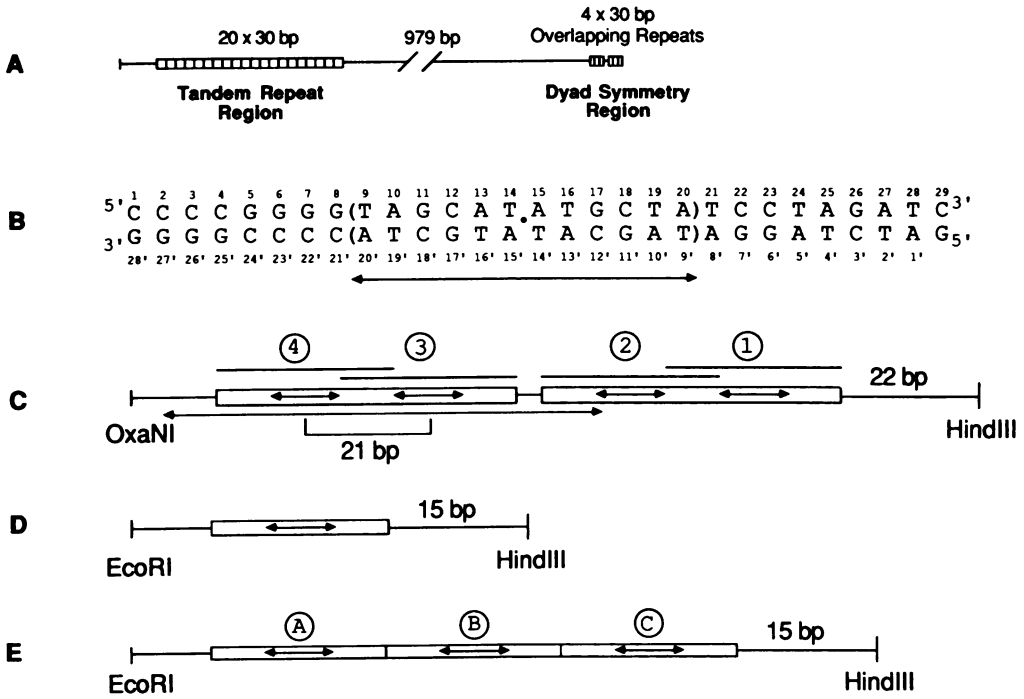

FIG. 1. The 1,800-bp *oriP* fragment and the DNA molecules used in these experiments. (A) The DNA fragment shown, called *oriP*, is the minimum amount of the Epstein-Barr virus genome required in *cis* for wild-type replication. Wild-type replication also requires the virally encoded protein EBNA-1, which binds to the tandem repeat and dyad symmetry regions. (B) The 29-bp synthetic EBNA-1 consensus binding sequence contains a central 12-bp palindrome at nucleotides 9 through 20, as indicated by the double-headed arrow. The numbering system specified here for the top and bottom strands is used throughout the text. Positions on the bottom strand are denoted with primes. (C) The native dyad symmetry region was studied in the illustrated fragment. In this region, pairs of consensus repeats (indicated by circled numbers above the diagram) overlap by 9 bp. The 65 bp with dyad symmetry are indicated by the double-headed arrow below the diagram. Single (D) and triply tandemly repeated (E) synthetic consensus sites were studied here by 5′ (bottom strand) and 3′ (top strand) labeling at the *Hin*dIII site.

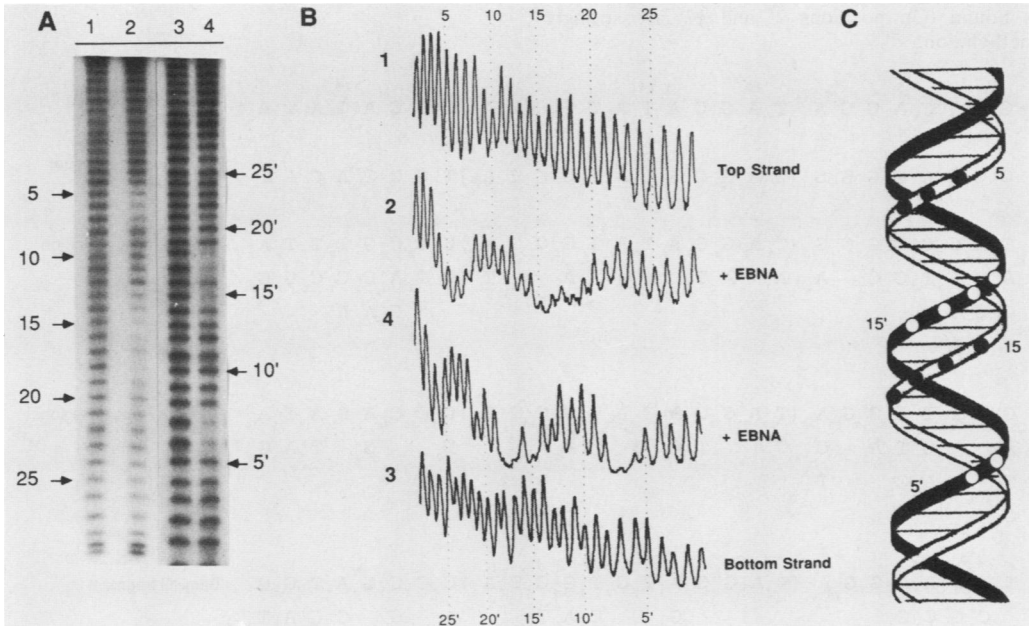

FIG. 2. Hydroxyl radical footprints of EBNA bound to a single synthetic consensus sequence. Autoradiograph (A) and densitometer tracings (B) of a gel show the hydroxyl radical cleavage pattern of a single synthetic consensus sequence (Fig. 1B and D) in the absence (lanes 1 and 3) and presence (lanes 2 and 4) and EBNA. Lanes 1 and 2, Hydroxyl radical cleavage products derived from the top strand; lanes 3 and 4, cleavage products derived from the bottom. (C) Positions on a DNA helix (10.5 bp per turn) that are protected from hydroxyl radical attack by EBNA bound to a single synthetic consensus binding site are indicated by dots on the DNA strands. The white strand is the top strand.

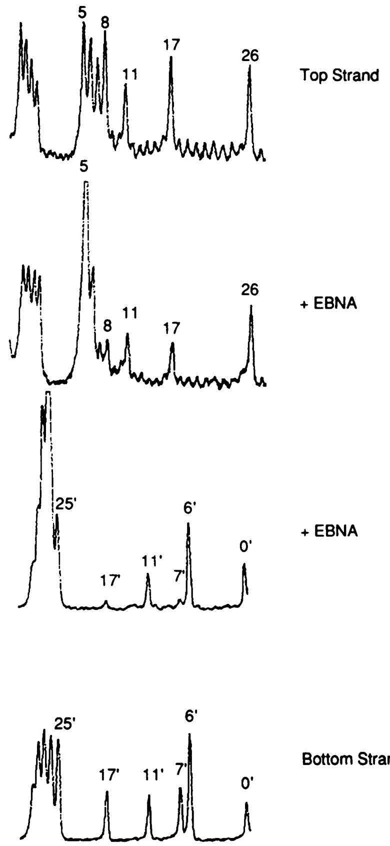

FIG. 3. Methylation protection by EBNA. (A and B) Densitometer tracings of lanes on an autoradiograph that show the methylation pattern of the top strand of the single consensus site (Fig. 1B and D) without (A) and with (B) bound EBNA. Positions 7, 8, and 17 are strongly protected from methylation by EBNA, as is seen by comparing the cleavage intensities of these residues with those for the flanking guanines at the left of the figure. Position 5 is methylated more when EBNA is bound. Numbering is as in Fig. 1B. (C and D) Methylation data for the bottom strand of the single consensus site. When EBNA is bound (C), positions 7′ and 17′ are strongly protected from methylation.

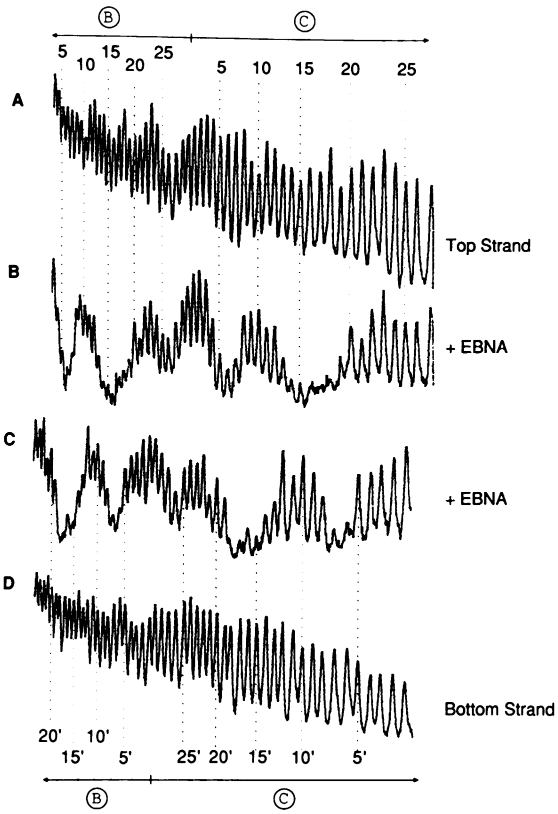

FIG. 5. Hydroxyl radical footprint of EBNA bound to the triply tandemly repeated synthetic consensus sequence. Densitometer scans of autoradiographs are shown for cleavage by hydroxyl radical of the top strand of the triply tandemly repeated consensus sequence (see Fig. 1E) in the absence of bound protein (A), the top strand with bound EBNA (B), the bottom strand with bound EBNA (C), and the naked bottom strand (D). In each of these sites, as in the single site, the greatest protections are seen at positions 5 through 7, 15 through 18, 6′ through 7′, and 15′ through 18′.

				5					10				14					19							25			
G	A	T	T	A	G	G	A	(T	A	G	C	A	T	A	T	G	C	T	A)	C	C	C	A	G	A	T	A	Natural Consensus Sequence
C	C	C	C	G	G	G	G	(T	A	G	C	A	T	A	T	G	C	T	A)	T	C	C	T	A	G	A	T	Synthetic Consensus Sequence
T_{2}	A_{2}	C_{1}	C_{2}	T_{3}	G_{3}	G_{4}	G_{2}	(T_{3}	A_{4}	G_{3}	C_{3}	A_{4}	T_{4}	A_{4}	T_{4}	G_{3}	C_{4}	T_{4}	A_{3})	C_{3}	C_{2}	C_{2}	G_{2}	T_{2}	T_{2}	T_{2}	A_{3}	Dyad Symmetry Region
A_{1}	G_{1}	G_{1}	T_{2}	C_{1}	C_{1}		A_{2}	A_{1}		A_{1}	T_{1}					A_{1}			T_{1}	T_{1}	T_{2}	A_{1}	C_{1}	C_{1}	C_{2}	G_{2}	G_{1}	
C_{1}	C_{1}	T_{1}																				G_{1}	A_{1}	A_{1}				
		A_{1}																										
G_{13}	A_{20}	T_{20}	T_{18}	A_{18}	G_{19}	G_{20}	A_{18}	(T_{19}	A_{20}	G_{20}	C_{19}	A_{16}	T_{20}	A_{20}	T_{20}	G_{15}	C_{19}	T_{20}	A_{20})	C_{18}	C_{20}	C_{20}	A_{17}	G_{20}	A_{19}	T_{19}	A_{19}	Tandem Repeat Region
A_{3}			G_{1}	T_{1}	A_{1}		G_{1}	C_{1}			T_{1}	C_{4}				A_{5}	A_{1}			T_{2}			G_{3}		T_{1}	A_{1}	G_{1}	
T_{3}			A_{1}	C_{1}			T_{1}																					
C_{1}																												
T_{2}	G_{1}	T_{1}	A_{1}	A_{1}	G_{2}	G_{2}	A_{2}	(T_{2}	A_{2}	G_{2}	C_{2}	A_{1}	T_{2}	G_{2}	T_{1}	G_{1}	C_{1}	T_{2}	A_{2})	C_{2}	C_{2}	C_{1}	G_{2}	A_{1}	T_{1}	G_{1}	G_{1}	BamHQ fragment
	C_{1}	G_{1}	C_{1}	G_{1}								G_{1}			C_{1}	A_{1}	T_{1}					G_{1}		C_{1}	C_{1}	A_{1}	T_{1}	

FIG. 4. Conservation of sequence in the three areas of the Epstein-Barr virus genome where EBNA binds. The palindromic sequence of the binding site is enclosed in parentheses. The frequencies of occurrence of a particular nucleotide in the dyad symmetry, tandem repeat, and the non-*oriP* (*Bam*HQ fragment) EBNA-1-binding sites are indicated by subscripts. The highly conserved nucleotides at positions 10 and 19 are shaded (see text).

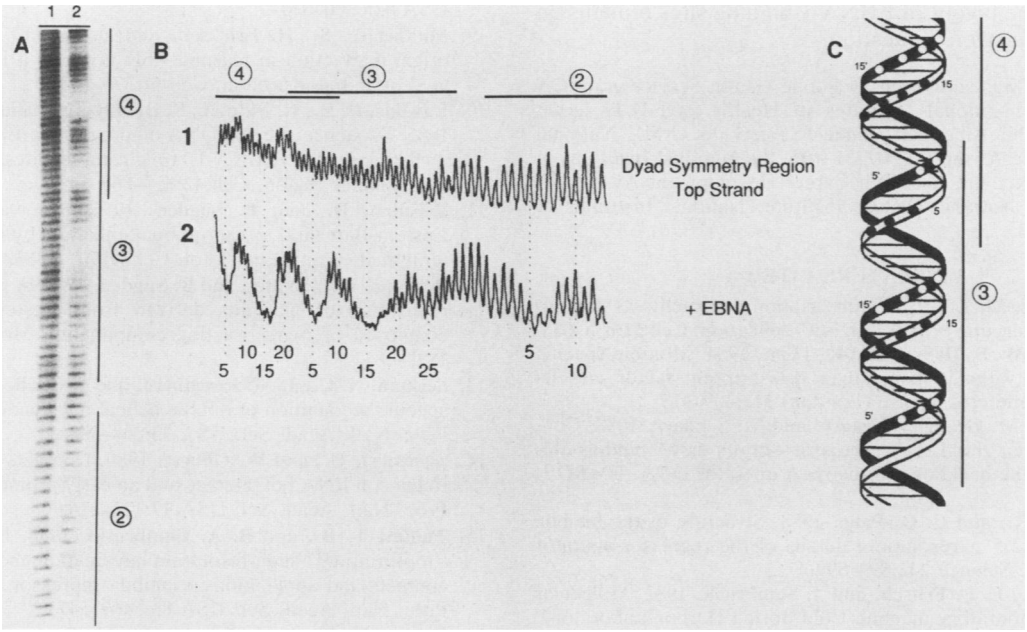

FIG. 6. EBNA binding to the dyad symmetry region analyzed by hydroxyl radical footprinting. Autoradiograph (A) and densitometer scans (B) of the cleavage products of the top strand of the dyad symmetry region of *oriP* (see Fig. 1C) in the absence (lane 1) and presence (lane 2) of EBNA. The numbered regions to the right of the autoradiograph correspond to the overlapping EBNA-1-binding sites as illustrated in Fig. 1C. (C) Positions on a DNA helix (10.5 bp per turn) that are protected from hydroxyl radical attack by EBNA bound to the dyad symmetry region are indicated by dots on the DNA backbone. Two of the four overlapping repeats are shown. The dark strand is the bottom strand. (Data for this strand were collected but are not shown in panel A or B.)

protected nucleotides fall on one face of a canonical B-DNA helix (Fig. 2C). The LexA protein and the λ repressor and cro proteins also bind to palindromic sequences and strongly protect four patches of nucleotides on one face of the helix from hydroxyl radical attack (3, 15).

The cocrystal structure of the λ repressor bound to DNA (4) shows that the recognition alpha helices of the repressor dimer sit in two neighboring symmetry-related major grooves, making sequence-specific contacts with five consecutive base pairs. This crystal structure helps us to relate the hydroxyl radical-protected nucleotides to the nucleotides that are specifically contacted by protein. With λ repressor and cro, the most protected nucleotides border those that are contacted in the major groove (15). The nucleotides that provide the base pairs in the major groove that make contact with the protein have their sugar residues on the back side of the protein-DNA complex and so are not protected from hydroxyl radical attack (Fig. 2C).

The idea that EBNA binds in these two major grooves is supported by methylation protection data. The N7s of guanines 7, 8, 11, 17, 7′, 11′, and 17′ face into these grooves and all but 11 and 11′ are strongly protected from dimethyl sulfate modification by bound EBNA (Fig. 3). Enhancement of methylation is seen at positions 5 and 26′. Similar methylation enhancement has been observed in other systems at guanines that neighbor protected guanines (2). These data suggest that EBNA is located on the major groove side of bases 7 through 12 and 17 through 22.

Comparison of the sequences of the multiple naturally occurring EBNA-1-binding sites identifies nucleotides that might be important for binding (Fig. 4). Absolute conservation of sequence is seen at positions 10, 14, and 19. Positions 10 and 19 are symmetry equivalent AT base pairs in the middle of the major grooves implicated in EBNA binding. We suggest that EBNA interacts directly with this base pair, perhaps by forming two hydrogen bonds to each adenine (13).

EBNA bound to a DNA fragment containing three tandemly repeated 29-bp synthetic consensus sites (Fig. 1E) protects the same backbone positions in each repeated site from hydroxyl radical attack as it does when bound to an isolated site (compare Fig. 2B and 5). The characteristic shapes of the footprints for the triply repeated site are the same as the shapes of the footprints for a single site and are also the same as the footprints for the 3 of 30 native tandem repeats (repeats 17 through 19) that we studied (data not shown). The borders of adjacent strong footprints are 11 bp apart. Over these intervening regions, we see no change in hydroxyl radical reactivity compared with the regions bordering a single site.

The contacts of EBNA with the dyad symmetry region (Fig. 1C) are also exactly analogous to those seen in the footprint of the single consensus site (compare Fig. 2B and 6). Here, because of the 9-bp overlap of adjacent consensus repeats, the strong contacts from adjacent repeats are brought closer together. All eight strongly protected regions within one set of two overlapping repeats fall on the same side of a 10.5-bp-turn DNA helix (Fig. 6C). The strong contacts on two adjacent overlapping repeats are one minor groove width apart, leaving open the possibility that EBNA occupies both simultaneously.

Why are two distantly situated EBNA-1-binding sites necessary for control of DNA replication in the Epstein-Barr virus? Several lines of evidence suggest that *oriP* and EBNA-1 might control replication by forming a loop of DNA held together at its base by protein (1, 7, 9). We have detected no difference in the way the binding domain of EBNA-1 contacts the two arrangements of binding sites that occur at the ends of *oriP*. The functional role of this

difference in the layout of EBNA-1-binding sites remains an interesting unanswered question.

This research was supported by Public Health Service grant CA 37444 from the National Institutes of Health to T.D.T.; grant MV-287 from the American Cancer Society to G.M.; National Research Service Award GM 07231 from the National Institutes of Health to A.S.K.; and Research Career Development Award CA 01208 from the National Cancer Institute, National Institutes of Health to T.D.T.

LITERATURE CITED

1. **DePamphilis, M. L.** 1988. Transcriptional elements as components of eukaryotic origins of DNA replication. Cell **52**:635–638.
2. **Gidoni, D., W. S. Dynan, and R. Tjian.** 1984. Multiple specific contacts between a mammalian transcription factor and its cognate promoters. Nature (London) **312**:409–413.
3. **Hurstel, S., M. Granger-Schnarr, and M. Schnarr.** 1988. Contacts between the LexA repressor—or its DNA-binding domain—and the backbone of the recA operator DNA. EMBO J. **7**:269–275.
4. **Jordan, S. R., and C. O. Pabo.** 1988. Structure of the lambda complex at 2.5 Å resolution: details of the repressor-operator interactions. Science **242**:893–899.
5. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
6. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
7. **Milman, G., and E. S. Hwang.** 1987. Epstein-Barr virus nuclear antigen forms a complex that binds with high concentration dependence to a single DNA-binding site. J. Virol. **61**:465–471.
8. **Milman, G., A. L. Scott, M.-S. Cho, S. C. Hartmen, D. K. Ades, G. S. Hayward, P.-F. Ki, J. T. August, and S. D. Hayward.** 1985. Carboxyl-terminal domain of the Epstein-Barr virus nuclear antigen is highly immunogenic in man. Proc. Natl. Acad. Sci. USA **82**:6300–6304.
9. **Mukherjee, S., H. Erickson, and D. Bastia.** 1988. Enhancer-origin interaction in plasmid R6K involves a DNA loop mediated by initiator protein. Cell **52**:375–383.
10. **Rawlins, D. R., G. Milman, S. D. Hayward, and G. S. Hayward.** 1985. Sequence-specific DNA binding of the Epstein-Barr virus nuclear antigen (EBNA-1) to clustered sites in the plasmid maintenance region. Cell **42**:859–868.
11. **Reisman, D., and B. Sugden.** 1986. *trans* activation of an Epstein-Barr viral transcriptional enhancer by the Epstein-Barr viral nuclear antigen 1. Mol. Cell. Biol. **6**:3838–3846.
12. **Reisman, D., J. Yates, and B. Sugden.** 1985. A putative origin of replication of plasmids derived from Epstein-Barr virus is composed of two *cis*-acting components. Mol. Cell. Biol. **5**: 1822–1832.
13. **Seeman, N. C., J. M. Rosenberg, and A. Rich.** 1976. Sequence-specific recognition of double helical nucleic acids by proteins. Proc. Natl. Acad. Sci. USA **73**:804–808.
14. **Siebenlist, U., and W. Gilbert.** 1980. Contacts between Escherichia coli RNA polymerase and an early promoter of phage T7. Proc. Natl. Acad. Sci. USA **77**:122–126.
15. **Tullius, T. D., and B. A. Dombroski.** 1986. Hydroxyl radical "footprinting": high-resolution information about DNA-protein contacts and application to lambda repressor and cro protein. Proc. Natl. Acad. Sci. USA **83**:5469–5473.
16. **Tullius, T. D., B. A. Dombroski, M. E. A. Churchill, and L. Kam.** 1987. Hydroxyl radical footprinting: a high-resolution method for mapping protein-DNA contacts. Methods Enzymol. **155**:537–558.
17. **Yates, J., N. Warren, D. Reisman, and B. Sugden.** 1984. A cis-acting element from the Epstein-Barr viral genome that permits stable replication of recombinant plasmids in latently infected cells. Proc. Natl. Acad. Sci. USA **81**:3806–3810.
18. **Yates, J. L., N. Warren, and B. Sugden.** 1985. Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells. Nature (London) **313**:812–815.