Binding of EBNA-1 to DNA Creates a Protease-Resistant Domain That Encompasses the DNA Recognition and Dimerization Functions

WARIS A. SHAH,¹ RICHARD F. AMBINDER,^{1†} GARY S. HAYWARD,¹ AND S. DIANE HAYWARD^{2*}

Virology Laboratories, Department of Pharmacology and Molecular Sciences,¹ and Department of Neurology,² Johns Hopkins School of Medicine, 600 North Wolfe Street, Baltimore, Maryland 21205

Received 11 December 1991/Accepted 20 February 1992

The Epstein-Barr virus nuclear antigen EBNA-1 is essential for replication of the viral DNA during latency. EBNA-1 binds as a dimer to palindromic recognition sequences within the plasmid origin of replication, ori-P. In this study, proteinase K susceptibility has been used to further characterize the DNA-binding domain of EBNA-1. Limited protease digestion of EBNA-1 (amino acids 408 to 641) generated a smaller DNA-binding species that had a degree of inherent protease resistance. When EBNA-1 was preincubated with a specific DNA probe, the protease resistance of the smaller binding species increased 100-fold, suggesting that the conformation of EBNA-1 changes on binding. The protease-resistant species comprised an 18-kDa polypeptide that was further cleaved at high levels of protease to 11- and 5.4-kDa products. A model of the proposed protease-resistant domain structure is presented. Constructions carrying serial, internal deletions across the 18-kDa domain were created. Each of the deletions perturbed dimerization ability and abolished DNA binding. These studies suggest that the DNA-binding and dimerization motifs of EBNA-1 lie within a conformationally discrete domain whose overall integrity is necessary for EBNA-1-DNA interaction.

In latently infected B cells, the Epstein-Barr virus (EBV) genome is maintained as an episome that replicates in synchrony with cell division (39), and viral gene expression is limited (reviewed in reference 12). One of the latency proteins, EBNA-1, plays a key role in the maintenance of latent infection. EBNA-1 is functionally pleiotropic. A contribution to latency gene regulation is implied from its ability to transactivate the latency C promoter (32) and from the positioning of the region III EBNA-1-binding sites immediately downstream of the recently identified latency F promoter (11, 27, 29, 31). EBNA-1 also has an established role in DNA replication and is the sole virus-encoded protein required for replication of the episomal form of the EBV genome (40). EBNA-1 mediates its activity through direct interaction with the viral DNA and binds with high affinity to palindromic recognition sequences within the latency origin of plasmid replication, ori-P (11, 24). The two structural domains of ori-P, the family of repeats and the dyad symmetry region, contain 20 and 4 copies of the EBNA-1-binding site, respectively (24, 26, 37). EBNA-1 binding to the family of repeats provides an enhancer function that is required for active replication (25, 36). The dyad symmetry region is the site of initiation of replication (7), and the contribution of EBNA-1 binding at this location is poorly understood. Likely roles include facilitating DNA conformational changes and recruitment of cellular replication proteins.

The EBNA-1-binding site has been extensively characterized (2). High-affinity binding requires a 16-bp palindromic recognition sequence with the consensus GA/GTAGCATA TGCTAC/TC. The relative spacing between the halves of the recognition sequence is critical, and binding is abolished by the addition or removal of bases at the center of the

palindrome. Examination of all possible single-base substitutions within the consensus half-binding site revealed that positions 3 through 8 of the recognition sequence had the most stringent requirements, with transversions at these positions reducing or eliminating binding. Methylation interference studies also indicated contacts between EBNA-1 and bases within the major groove in this segment of the binding site (2, 14).

The EBNA-1 protein comprises 641 amino acids (aa), and information is accumulating on the functional domains within EBNA-1. One-third of EBNA-1 (aa 90 to 325) consists of a repetitive array of glycine and alanine residues which can largely be deleted without affecting function in transient replication and plasmid maintenance assays (22, 38). EBNA-1 is found in the nuclear compartment of the cell, and aa 379 to 387 mediate nuclear localization (1). The active DNA-binding form of EBNA-1 is a dimer, and both DNA-binding and dimerization functions map to the carboxy-terminal one-third of EBNA-1 (1). Since EBNA-1 function is integrally linked to DNA binding, priority has been given to characterization of the DNA-binding domain of the protein. We initially performed DNase I footprinting studies and demonstrated that a bacterial fusion protein containing aa 450 to 641 of EBNA-1 [EBNA-1(aa450-641)] bound specifically to target sequences within ori-P (24). Subsequently, mobility retardation assays were used to monitor the DNA-binding properties of serial amino-terminal and carboxy-terminal deletions of this aa 450 to 641 domain. These studies revealed that retention of wild-type levels of DNA binding required aa 459 to 607 (1). Low levels of DNA binding (<10% of the wild-type level) were observed with some additional constructions, for example, EBNA-1(aa470-641) and EBNA-1(aa450-584). However, it was unclear whether the decreased binding ability demonstrated by these polypeptides was due to encroachment on specific DNA recognition and dimerization motifs or rather to disruption of native polypeptide conformation. In this study, the protease

* Corresponding author.

† Present address: Department of Oncology, Johns Hopkins School of Medicine, Baltimore, MD 21205.

susceptibility of DNA-bound EBNA-1 has been used to gain further insight into the structure of the DNA-binding domain. The generation of a protease-resistant polypeptide mapping within EBNA-1 aa 459 to 617 and the inability of this region to tolerate internal deletions provides evidence that the DNA recognition and dimerization motifs of EBNA-1 lie within a conformationally discrete domain whose integrity is necessary for the interaction of EBNA-1 with DNA.

MATERIALS AND METHODS

Plasmid constructions. The plasmids used for in vitro transcription and translation of EBNA-1 were as follows. EBNA-1(aa408-641) was expressed from pRA362, which has been described previously (1). EBNA-1(aa459-641), pWS61, was generated from pRA362 by using oligonucleotide primers and the polymerase chain reaction (PCR). The internal deletion series were created in two steps. First, oligonucleotide-based site-directed mutagenesis was used to introduce a unique *Xba*I restriction site at EBNA-1 codons 479 and 480 (pRA324), 499 and 500 (pRA330), 521 and 522 (pRA333), 555 and 556 (pRA325), and 615 and 616 (pRA331). Second, PCR was used to amplify pairs of products in which the *Xba*I site served as either the 5' or 3' anchor sequence. The PCR fragment pairs were then cleaved with *Xba*I, ligated, and cloned into the in vitro transcription vector pBD7 (5). In this way, we created pRA374 (Δ 480-521), pRA376 (Δ 500-521), pRA378 (Δ 525-532), pWS52A (Δ 537-553), pWS43 (Δ 554-576), pWS46 (Δ 577-615), and pWS47 (Δ 599-615). All deletions were confirmed by DNA sequencing (3). The control plasmid for the cross-linking experiments, pWS17, contained a truncated Zta gene terminating at the *Pst*I site in codon 198. The protein expressed from this plasmid lacks the Zta dimerization domain (4).

Plasmids carrying 30-mer consensus EBNA-1 binding sites were pGH66 and pGH65, which have one and two copies, respectively (11), and pWS10, which has one copy cloned into the *Xba*I site of the vector pBEND2 (13). Plasmid pWS13 is similar to pWS10 except that the insert carries a 2-bp deletion at the center of the palindrome that abolishes EBNA-1 binding (2).

In vitro transcription-translation. Plasmid DNA was linearized downstream of the coding sequence of EBNA-1 and incubated with T7 RNA polymerase to prepare capped mRNA, using an in vitro transcription kit (Stratagene, La Jolla, Calif.). In vitro translation was carried out by using rabbit reticulocyte lysate (Promega, Madison, Wis.). A standard in vitro translation reaction contained 1 to 2 μ g of mRNA in a 50- μ l reaction mixture and 50 μ Ci of [³⁵S]methionine (800 Ci/mmol) or [³⁵S]cysteine (1 Ci/mmol) (New England Nuclear, Wilmington, Del.). Labeled proteins were stored at -70°C.

Mobility shift assay and proteinase K digestion. A 156-bp specific DNA probe was prepared by *Bgl*II cleavage of pWS10, and a comparable 154-bp nonspecific probe was generated by *Bgl*II cleavage of pWS13. Monomer (90-bp) and dimer (120-bp) binding-site probes were prepared by *Eco*RI and *Xba*I cleavage of plasmids pGH66 and pGH65, respectively. Radiolabeled probes were prepared by end-labeling with [³²P]dATP and Klenow polymerase. In a typical binding assay, EBNA-1 was incubated with probe (1 fmol) for 30 min at 4°C in 25 μ l of binding buffer [25 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid (HEPES; pH 7.5), 1 mM dithiothreitol, 0.05% Nonidet P-40, 8 μ g of poly(dI-dC) · poly(dI-dC) per ml. The reaction mixtures were

loaded onto a 5% polyacrylamide gel in TBE (0.05 M Tris, 0.05 M boric acid, 1 mM EDTA) and electrophoresed in TBE at 10 V/cm for 2 h at room temperature.

Digestion with proteinase K (from *Tritirachium album*; Sigma, St. Louis, Mo.) was performed in binding buffer after preincubation without DNA or with nonspecific or specific DNA probes. After electrophoresis, the positions of the DNA-bound complexes were determined by autoradiography and the complexes were excised. The protein was eluted by the method of Hager and Burgess (9) and analyzed by electrophoresis through a sodium dodecyl sulfate (SDS)-15% polyacrylamide gel (15). Gels were dried and exposed for autoradiography.

Tryptic peptide fingerprinting. A modification of the method of Gibson et al. (8) was used for two-dimensional analysis of tryptic peptides. Following electrophoresis, the gel was soaked in 500 ml of water containing 10 g of mixed bed resin (RG 501-X8; Bio-Rad) for 30 min. The gel was then dried and exposed for autoradiography. The desired bands were excised and rehydrated for 20 min in formic acid-methanol (4:1). One-fifth volume of performic acid was added, and the gel slices were soaked for 2 h at -8°C. After repeated lyophilization to remove acid, the gel slices were rehydrated for 20 min in 0.05 M NH₄HCO₃. The hydrated gel was fragmented by passage through a steel mesh and incubated at room temperature overnight with 40 μ l of TPCK-trypsin (1 mg/ml; Worthington). The acrylamide was removed by passage through two disks of a Millipore prefilter (AP25 04200). Hydrolyzed protein was lyophilized and resuspended in peptide electrophoresis buffer (acetic acid-formic acid-water, 78:25:897). The peptides were spotted onto a corner of a thin-layer cellulose plate, the plate was dampened with peptide electrophoresis buffer, and electrophoresis was carried out at 800 V for 30 min under pressure (100 lb/in²). The plates were then dried, chromatographed in the second dimension in peptide chromatography buffer (butanol-pyridine-acetic acid-water, 90:75:15:60) for 6 h, and exposed for autoradiography.

Protein cross-linking and immunoprecipitation. In vitro translated ³⁵S-labeled proteins were diluted to 90 μ l with 10 mM potassium phosphate buffer (pH 8.0), and 10 μ l of freshly diluted 0.1% glutaraldehyde was added. After 1 h at room temperature, 400 μ l of TSET buffer (150 mM NaCl, 50 mM Tris-HCl, 0.1 mM EDTA, 2% Triton X-100, pH 8.0) and rabbit anti-EBNA-1 (19) or rabbit anti-Zta (4) antiserum were added, and the mixture was incubated at 4°C for 90 min. Protein A-Sepharose was added, and after 1 h of incubation at 4°C, the precipitate was pelleted and washed three times in TSET. The sample was resuspended in buffer (62.5 mM Tris [pH 6.8], 10% glycerol, 4% SDS, 0.001% bromophenol blue) and boiled for 5 min before electrophoresis on an SDS-10% polyacrylamide gel.

RESULTS

DNA binding decreases susceptibility of EBNA-1 to protease digestion. In a previous study (1), a series of 5' and 3' deletions of the EBNA-1 open reading frame was used to map the location of the DNA-binding domain. Wild-type levels of DNA-binding activity were observed in the 5' series with EBNA-1 polypeptides commencing at or before aa 459 and in the 3' series with an EBNA-1 polypeptide terminating at aa 607. These two constructions also expressed stable polypeptides, as determined by immunoblot analyses. However, incursion inside the aa 459 and aa 607 boundaries significantly reduced the stability of the bacterially ex-

FIG. 1. Binding to DNA decreases the protease susceptibility of EBNA-1. A mobility retardation assay was performed to examine the ability of in vitro-translated EBNA-1(aa408-641) to bind to a radiolabeled, consensus binding-site probe after treatment with proteinase K (0.01 to 5 µg). Protease digestion was performed either before addition of DNA (lanes 1 to 3), after addition of a mutated sequence incapable of binding to EBNA-1 (lanes 4 to 6), or after addition of a consensus binding site (lanes 7 to 11). Lanes 12 to 15 contain control reactions consisting of mutated nonbinding probe (lane 12), mutated probe plus EBNA-1 (lane 13), consensus DNA probe (lane 14), and consensus probe plus EBNA-1 (lane 15).

pressed EBNA-1 polypeptides. This observation suggested that the region between aa 459 and 607 might represent a structural domain. To further examine this aspect of EBNA-1-DNA interaction, proteolytic digestion was used as a means of identifying structurally stable domains of the protein.

In vitro-translated EBNA-1 (aa408-641) was digested with 0.01, 0.1, or 1.0 µg of proteinase K and then tested for specific DNA binding in a mobility shift assay (Fig. 1). Digestion with 0.01 and 0.1 µg of proteinase K resulted in the generation of polypeptides that continued to bind to a specific DNA probe but were smaller in size, as indicated by the increased mobility of the shifted complex relative to the parental EBNA-1 (Fig. 1, lanes 1 and 2). Incubation with 1.0 µg of proteinase completely eliminated DNA binding (lane 3). To determine whether the proteinase sensitivity of EBNA-1 was altered by interaction with DNA, EBNA-1(aa408-641) was preincubated with either a nonspecific or a specific DNA probe prior to proteinase K digestion. Preincubation with nonspecific DNA had no effect on proteinase sensitivity (lanes 4 to 6). Proteinase digestion after incubation with a specific 156-bp DNA probe resulted in the generation of a dominant binding species with the same mobility in the gel shift assay as the smallest binding species generated in the absence of DNA or in the presence of nonspecific DNA. However, after preincubation with specific probe, this DNA-binding species demonstrated greatly increased resistance to proteinase digestion, and DNA binding remained undiminished even after exposure to 5 µg of proteinase K (lanes 7 to 11).

Protease resistance identifies an 18-kDa DNA-binding domain. The generation of a protease-resistant DNA-binding species indicates that the DNA-binding function of EBNA-1 lies within a physically and conformationally discrete domain. To determine the size and approximate boundaries of this domain, the protected species generated from three different EBNA-1 constructions were compared. Methionine-labeled EBNA-1(aa408-641), EBNA-1(aa459-641), and EBNA-1(aa459-617) were subjected to proteinase K diges-

FIG. 2. Protease digestion generates an 18-kDa DNA-binding domain that maps within the aa 459 to 617 segment of EBNA-1, as determined by gel electrophoretic analysis of the protease-resistant EBNA-1 polypeptides present in DNA-bound complexes. [³⁵S]methionine-labeled, in vitro-translated EBNA-1(aa408-641), EBNA-1(aa459-641), and EBNA-1(aa459-617) were incubated with proteinase K in the presence of unlabeled consensus DNA probe. The DNA-bound complexes were separated from unbound protein and DNA by electrophoresis through a native 5% polyacrylamide gel and isolated, and the polypeptides present were analyzed on an SDS-15% polyacrylamide gel. Results are shown for EBNA-1(aa408-641) minus (lane 1) and plus (lane 2) protease, EBNA-1(aa459-641) plus (lanes 3 and 6) and minus (lanes 4 and 5) protease, and EBNA-1(aa459-617) plus (lane 7) and minus (lane 8) protease.

tion after preincubation with specific probe. The DNA-bound species were separated from unbound protein by using a mobility shift assay. The shifted complex was isolated from the gel, and the sizes of the protected polypeptides in the complex were then determined by electrophoresis through a denaturing polyacrylamide gel (Fig. 2). The proteinase-resistant species generated from EBNA-1(aa408-641) consisted of a dominant polypeptide of 18 kDa and a less abundant 11-kDa polypeptide (Fig. 2, lane 2). The proteinase-resistant species generated from EBNA-1(aa459-641) (lanes 3 and 6) and EBNA-1(aa459-617) (lane 7) gave rise to polypeptides of the same size as did the larger EBNA-1(aa408-641) form. Thus, the 18-kDa protease-resistant domain lies within the 158-aa region bounded by aa 459 and 617. The size of this region is compatible with the 139-aa size estimated for an 18-kDa polypeptide.

Binding of EBNA-1 to adjacent sites does not modify protease resistance. Each of the three EBNA-1-binding loci in the EBV genome consists of multiple adjacent binding sites. ori-P contains two subdomains: the family of repeats which includes 20 EBNA-1-binding sites tandemly repeated at 30-bp intervals, and the dyad symmetry locus, which consists of two paired EBNA-1-binding sites spaced 21 bp apart. The physically distant Q locus also contains two overlapping EBNA-1-binding sites with their centers 25 bp apart (2). We previously examined EBNA-1 binding in competition assays (11) and did not find any evidence for cooperative binding when two or four consensus 30-bp sites were present. The observation that protease resistance changes upon DNA binding potentially provides another way to probe for cooperative EBNA-1 interactions. Methionine-labeled EBNA-1(aa408-641) was preincubated with DNA containing either monomer or dimer 30-bp binding sites and then digested with proteinase K. In the mobility

FIG. 3. Binding of EBNA-1 to adjacent consensus sites does not alter protease susceptibility. In vitro-translated, methionine-labeled EBNA-1(aa408-641) was incubated with unlabeled DNA fragments containing 1 or 2 consensus binding sites. After protease digestion, the DNA-bound complexes were separated by electrophoresis through a native polyacrylamide gel. (A) A monomer binding-site probe was used with increasing amounts of protease (lanes 1 to 4) and in the absence of protease (lane 5). A dimer binding-site probe was used with increasing protease (lanes 6 to 9) and in the absence of protease (lane 10). (B) The polypeptides present in the excised DNA-bound complexes were analyzed by SDS-polyacrylamide gel electrophoresis. No difference was observed between peptides present in the monomer and dimer complexes. Lanes: 1, monomer complex from lanes 2 to 4 of panel A; 2, complex from the dimer probe with a single site occupied, lanes 7 to 9 of panel A; 3, complex from the dimer probe with both sites occupied, lanes 7 to 9 of panel A; 4, mixed monomer and dimer complexes without protease digestion, lanes 5 and 10 of panel A.

shift assay (Fig. 3A), the monomer probe gave a single shifted complex (lanes 1 to 5) and the dimer probe produced two complexes representing the probe DNA with one site filled (dimer 1) or with both sites filled (dimer 2) (lanes 6 to 10). These shifted complexes were isolated, and the sizes of the protease-resistant polypeptides were examined on a denaturing acrylamide gel (Fig. 3B). The protease-resistant species isolated after preincubation with the monomer binding-site probe gave the expected 18-kDa and less abundant 11-kDa polypeptides (lane 1). Exactly the same polypeptides were released from the protease-resistant complexes that had been preincubated with the dimer binding-site probe, and this was true of the complexes representing one filled site (lane 2) and both sites filled (lane 3). Therefore, at this level of analysis, the presence of protein bound to adjacent sites at 30-bp intervals does not affect the formation of the protease-resistant domain, nor does it alter the size of the protected polypeptides.

The 18-kDa domain is susceptible to specific internal cleavage at high levels of protease. In addition to the dominant 18-kDa polypeptide, analysis of the protease-resistant complex revealed a less abundant 11-kDa species. To determine

FIG. 4. The 18-kDa polypeptide is cleaved to an 11-kDa product at high levels of protease. (A) Electrophoretic separation on a native 5% acrylamide gel of DNA-bound complexes formed after incubation of [³⁵S]methionine-labeled EBNA-1(aa408-641) with an unlabeled consensus DNA probe, followed by exposure to increasing amounts of proteinase K. Lanes: 1, EBNA-1(aa408-641) minus protease; 2, DNA probe added after protease digestion; 3 to 7, incubation with increasing amounts of protease. (B) SDS-polyacrylamide gel electrophoretic analysis of the polypeptides present in the DNA-bound complexes. Lanes 1 to 7 show the polypeptides present in the complexes from the corresponding lanes in panel A.

the derivation of this smaller polypeptide, methionine-labeled EBNA-1(aa408-641) was subjected to digestion with increasing amounts of proteinase K ranging from 2 to 100 µg of enzyme. As shown in Fig. 4A, a DNA-binding species was obtained even after digestion with 100 µg of protease, and the mobility of the bound complex remained similar to that formed after digestion with 2 µg of proteinase K. When the polypeptides present within these DNA-bound complexes were displayed on a denaturing polyacrylamide gel, a different picture emerged. Digestion with 2 µg of protease produced the previously observed pattern of a dominant 18-kDa species and a less abundant 11-kDa species (Fig. 4B). However, with increasing amounts of protease, the relative abundance of the two polypeptides changed until, at 100 µg of proteinase K, only the 11-kDa band was observed.

The generation of the 11-kDa polypeptide at the expense of the 18-kDa polypeptide suggested a precursor-product relationship. That the 11-kDa polypeptide does indeed represent a subfragment of the 18-kDa polypeptide was confirmed by a comparison of methionine-labeled tryptic peptides (Fig. 5). The identical nature of the methionine-labeled tryptic peptides combined with the asymmetric distribution of methionine residues in EBNA-1(408-641) (Fig. 6B) indicated that the labeled 11-kDa polypeptide was derived from the carboxy terminus of the 18-kDa species. The 11-kDa polypeptide could be generated either by continued proteolytic digestion sequentially removing amino acids from the amino-terminal end of the 18-kDa polypeptide or by specific

FIG. 5. Tryptic peptide analysis confirms that the 11-kDa polypeptide is derived from the 18-kDa species. After DNA binding, [35 S]Met-labeled EBNA-1(aa408-641) was digested with proteinase K, the resistant complex was isolated, and the polypeptides present were separated on a denaturing 15% polyacrylamide gel. The 18- and 11-kDa peptides were excised from the gel and incubated with trypsin, and the tryptic digestion products were subjected to a two-dimensional analysis on thin-layer cellulose plates. The identity of the tryptic digests confirms the relationship between the 11- and 18-kDa species.

cleavage at susceptible internal sites within the 18-kDa polypeptide. Since the relative mobility of the native EBNA-1-DNA complex does not change appreciably during the conversion of the 18-kDa form to the 11-kDa form, the latter seems more likely, but these alternatives cannot be addressed by using methionine-labeled EBNA-1. The four available methionines are located between aa 543 and 613 (Fig. 6B), and therefore any digestion products terminating prior to aa 543 would not be labeled and would not be detectable.

To obtain more information on the mechanism by which the 11-kDa polypeptide was generated, EBNA-1(aa459-641) was translated *in vitro* in the presence of [35 S]cysteine. The labeled protein was subjected to the same protocol of proteinase K digestion and analysis of the proteinase-resistant binding species on denaturing polyacrylamide gels. There are four cysteine residues in the 18-kDa polypeptide, and their distribution is similar to that of methionine except that the most amino-terminal cysteine is located at aa 530, as opposed to aa 542 for the equivalent methionine (Fig. 6B). Size analysis of the cysteine-labeled proteinase-resistant species revealed both the 18- and 11-kDa bands plus a new band at 5.4 kDa (Fig. 7). The presence of the additional smaller cysteine-labeled polypeptide is taken to indicate that the 11-kDa species is generated primarily by internal protease cleavage at a specific susceptible site within the 18-kDa polypeptide. The 5.4-kDa fragment was detectable by using radiolabeled cysteine but not methionine. The differential recovery of Cys-530 but not Met-542 within the 5.4-kDa species suggested that the protease cleavage event that generates the 11-kDa species may occur within the aa 531 to 541 segment of EBNA-1 (Fig. 6B). Cleavage in this region should generate the 11- and 5.4-kDa polypeptides with an apparent 3:1 molar ratio, since they contain three and one cysteine residues, respectively. However, the 5.4-kDa fragment has a considerably lower abundance than predicted. Since the cleavage event would occur very close to cysteine

FIG. 6. Diagrammatic representation of the location and proposed structure of the protease-resistant DNA-binding domain of EBNA-1. (A) EBNA-1(aa408-641), EBNA-1(aa459-641), and EBNA-1(aa459-617) each give rise to the 18-kDa protease-resistant polypeptide, indicating that this polypeptide maps within the aa 459 to 617 segment of EBNA-1. (B) The 11-kDa polypeptide is generated by further protease cleavage of the 18-kDa species, and analysis of methionine-labeled tryptic digestion products places the 11-kDa peptide at the carboxy terminus of the 18-kDa species. The detection of a 5.4-kDa protease-resistant polypeptide with cysteine- but not methionine-labeled EBNA-1 suggests that the cleavage event generating the 11-kDa species occurs between cysteine 530 and methionine 542 in a proline- and glycine-rich segment of the protein. ●, methionine; ○, cysteine. (C) Model illustrating the conformational changes that may occur when EBNA-1 binds DNA. The DNA-binding/dimerization domain becomes protease resistant, with a central segment (●) of this domain being the most susceptible to further cleavage at high protease concentrations.

530, it is possible that subsequent proteolytic activity results in loss of label from a percentage of the 5.4-kDa polypeptides or, as will be discussed later, that the 5.4-kDa fragment is preferentially released from the DNA-bound complex. The data obtained on the protease-resistant domain are summarized in Fig. 6, which also shows a model illustrating possible conformational features predicted by these experiments.

Deletions within the 18-kDa domain destroy DNA binding and perturb EBNA-1 dimerization. The 18-kDa protease-resistant fragment is located between aa 459 and 617 of EBNA-1 and appears to represent a discrete structural domain. The DNA-binding activity of this region and previous analyses of bacterially expressed EBNA-1 (1) indicate that the domain must contain both DNA recognition and EBNA-1 dimerization functions. The generation of the 11-kDa super-resistant polypeptide at high levels of protease raised the possibility that these two functions might be located in physically distinguishable subdomains. As a step toward examining this question, a series of internal deletions was created across the domain in an effort to identify any nonessential sequences. The deletions encompassed between 7 and 41 aa (Fig. 8). Although all of the deleted constructions yielded stable *in vitro*-translated products

FIG. 7. Labeling with [35 S]cysteine identifies an additional 5.4-kDa protease-resistant polypeptide. An SDS-polyacrylamide gel electrophoretic analysis of the polypeptides present in a DNA-bound protease-resistant complex formed by using [35 S]cysteine-labeled EBNA-1(aa459-641) with (lane 1) and without (lane 2) protease.

(data not shown), none of them proved capable of binding to DNA in a mobility retardation assay (Fig. 9A).

Since both DNA recognition and EBNA-1 dimerization are required for DNA binding, each of the deletions that we created could be intruding into one or other of these subdomains, with a resultant loss of DNA-binding activity. DNA recognition by EBNA-1 cannot be assessed in the absence of EBNA-1 dimerization. However, EBNA-1 dimerization occurs in solution in the absence of DNA, and hence dimerization ability can be measured as an independent function. Consequently, we performed glutaraldehyde cross-linking experiments on the same set of EBNA-1 deletions. Cross-linking of the parental EBNA-1(aa408-641) polypeptide gave a strong band at the dimer position (Fig. 9B and C). However, each of the deletions severely compromised dimerization ability. Since it is unlikely that all six deletions impinge directly on the dimerization motif, this result suggests that any disruption of the conformational integrity of the 18-kDa domain adversely affects the ability of the protein to dimerize and to interact with DNA.

DISCUSSION

We have used the susceptibility of EBNA-1 to protease digestion to examine aspects of the interaction between EBNA-1 and its DNA-binding site. Previous studies using

FIG. 8. Diagrammatic representation of internal deletions created within the 18-kDa domain. The locations of predicted alpha-helical regions (10) are indicated.

FIG. 9. Deletions within the 18-kDa protease-resistant domain abolish DNA binding and perturb dimerization. (A) Mobility retardation assay using a 32 P-labeled DNA fragment containing a consensus binding site. The parental EBNA-1(aa408-641) bound the probe, but none of the deletion mutants retained DNA-binding activity. (B and C) Glutaraldehyde cross-linking of [35 S]methionine-labeled, deleted polypeptides. A band at the dimer position in the SDS-polyacrylamide gel is readily detectable with the parental EBNA-1(aa408-641). However, only marginal dimerization is observed with the deleted polypeptides.

deleted amino- and carboxy-terminal variants of EBNA-1 had demonstrated that wild-type levels of DNA binding required an EBNA-1 domain located between aa 459 and 608 (1). EBNA-1 binds DNA as a dimer, and the dimerization motif also resides within this segment (1). Limited protease digestion of EBNA-1(aa408-641) in solution generated a

smaller polypeptide that bound DNA in a mobility retardation assay, indicating that the DNA-binding domain of EBNA-1 has a degree of inherent resistance to protease. EBNA-1 exists in solution almost exclusively in the dimer form (1, 6), and the partial protease resistance may be created by dimerization. The leucine zipper dimerization motif common in transcriptional activator proteins forms a helical coiled coil (16, 20, 23). Exposure of the leucine zipper protein C/EBP to trypsin in solution generates a protease-resistant fragment that encompasses the leucine zipper dimerization domain (30). However, the basic DNA recognition motif of free C/EBP is not protected, whereas the EBNA-1 recognition motif apparently is protected since DNA binding can still be demonstrated with the protease-treated EBNA-1.

When EBNA-1 is bound to DNA prior to exposure to protease, the protease resistance of the DNA-binding domain is increased at least 100-fold. This indicates a change in the EBNA-1 protein conformation on binding. The transcriptional activators Jun/Fos and GCN4 show an increase in alpha helicity on DNA binding (21, 34), and this conformational change may be necessary for stable protein-DNA contacts. The change in conformation of TFIID on binding has also been suggested to be necessary for stable binding (17). It is possible that conformational changes also unmask protein domains and in this way contribute to biological function. The change in conformation undergone by the yeast transcriptional activator PTRF on binding to a responsive site is believed to unmask an activator domain on that protein (33). In the case of EBNA-1 binding, induced changes could result in the exposure of a transcriptional activation domain or in unmasking sites for interactions between EBNA-1 and cellular replication proteins. The exact role played by EBNA-1 in episomal EBV DNA replication remains unclear. A working hypothesis is that EBNA-1 induces or stabilizes an accessible DNA structure at the origin and may then provide a scaffolding for assembly of the active DNA replication initiation complex. If the EBNA-1 domains involved in the initial steps of complex assembly were available only on DNA-bound EBNA-1 and not on free protein, this would provide a mechanism for minimizing nonproductive protein interactions. An alternative induced protein-protein interaction could involve EBNA-1 to EBNA-1 contacts. The existence of higher-order EBNA-1 DNA-bound complexes is a possibility (18).

Earlier studies had determined that wild-type levels of DNA binding required an EBNA-1 domain encompassing aa 459 to 608. We introduced a set of six internal deletions into this region and found that each of the deletions abolished DNA binding. Dimerization ability was also severely impaired by each of the deletions, and the inability to form a stable dimer is likely to have contributed to the observed loss of DNA binding. The dimerization and DNA-binding motifs of EBNA-1 have yet to be identified experimentally. Inoue et al. (10) noted a significant amino acid sequence homology between the basic-helix-loop-helix (HLH) domain of proteins such as c-Myc, MyoD, and E12/E47 and EBNA-1 in the region of aa 467 through 583. However, in the case of EBNA-1, the basic-helix 1 and helix 2 domains (aa 467 to 488 and 571 to 583) are separated by an extended loop of 83 aa, compared with only 10 to 12 aa in MyoD. This arrangement has some similarities to the basic-helix-span-helix (HSH) DNA recognition/dimerization motif of the transcriptional activator AP-2, in which two helix motifs are separated by a span of 82 aa (35). In the case of AP-2, all deletions across the HSH motif, including one in the span region, abolished

DNA binding and dimerization (35). This observation is reminiscent of our results with the deleted EBNA-1 polypeptides. Only two of our six deletions, $\Delta 480-521$ and $\Delta 556-576$, would impinge on the critical helix domains (Fig. 8) but all deletions affected both DNA binding and dimerization.

At high levels of protease, the 18-kDa protease-resistant polypeptide was clipped into an 11-kDa and a 5.4-kDa species. On the basis of the differential labeling of the 5.4-kDa fragment with cysteine and methionine, the cleavage at the junction with the 11-kDa fragment was mapped to a region between aa 531 and 541 that is rich in proline and glycine residues (Fig. 6B). A 5.4-kDa fragment would contain approximately 42 aa, which coincides with the distance between the cleavage in the proline-glycine region (aa 531 to 541) and the 3' end of the predicted helix 1 (aa 493). One interpretation of the protease data is that at high levels of protease, a second cleavage occurs near the junction of helix 1 with the intervening span DNA and that the 5.4-kDa fragment represents cleaved-out span sequences. Such a model would predict a loss of the 5.4-kDa fragment from the DNA-bound complex, and this could explain the relatively low molarity observed experimentally for this fragment. The protease experiments also raise an interesting point regarding the physical interaction between the DNA and EBNA-1. Both helix 1 and helix 2 of an HLH or HSH protein are required for dimerization and DNA binding. However, the data suggest that once binding has occurred, helix 2 and perhaps helix 1 can maintain a degree of association with the DNA even when the primary covalent linkage between them has been lost.

It is also interesting to note that protease sensitivity may in certain circumstances have biological relevance. In phage lambda, the lambda repressor is synthesized from the lyso-genic phage DNA, and its function is to stimulate its own synthesis while repressing transcription of the lytic phage genes. Lambda repressor binds specifically, with high affinity as a dimer to 17-bp operator sites. Inactivation of lambda repressor occurs through cleavage of the repressor polypeptide in a reaction that is partly autocatalytic and is facilitated by RecA protein. This cleavage event destabilizes the repressor dimer, which reduces operator binding and consequently allows initiation of the lytic phage growth cycle (28). The protease clipping of bound EBNA-1 may simply reflect the exposure of a relatively unstructured segment within a more densely packed alpha-helical region of the protein. However, the lambda repressor model has intriguing implications for the switch in EBV from latent ori-P-based DNA replication to ori-Lyt replication.

On the basis of the existence of a protease-resistant domain encompassing the DNA-binding and dimerization functions and the drastic effects of internal deletions within this region, we conclude that the EBNA-1 DNA recognition and dimerization motifs are intimately linked and form a highly structured domain that has features in common with both the HLH and HSH motifs. On the other hand, it should be noted that differences also exist. For example, the pattern of methylation protection of bases in the EBNA-1 consensus binding site differs significantly from that exhibited by the characterized MyoD HLH proteins.

ACKNOWLEDGMENTS

We thank Wade Gibson for advice and help with the tryptic digests, Mabel Chiu for technical assistance, and Sarah Heaggans and Pamela Wright for preparation of the manuscript.

This work was funded by Public Health Service grants RO1-CA42243 (S.D.H.), RO1-CA22130 (G.S.H.), and K11AI00648 (R.F.A.). W.A.S. was supported by a Pakistan Ministry of Science and Technology scholarship.

REFERENCES

- Ambinder, R. F., M. Mullen, Y. N. Chang, G. S. Hayward, and S. D. Hayward. 1991. Functional domains of Epstein-Barr virus nuclear antigen EBNA-1. *J. Virol.* **65**:1466-1478.
- Ambinder, R. F., W. A. Shah, D. R. Rawlins, G. S. Hayward, and S. D. Hayward. 1990. Definition of the sequence requirements for binding of the EBNA-1 protein to its palindromic target sites in Epstein-Barr virus DNA. *J. Virol.* **64**:2369-2379.
- 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. Sequin, P. S. Tuffnell, and B. G. Barrell. 1984. Organization of the B95-8 Epstein-Barr virus genome. *Nature (London)* **310**:207-211.
- Chang, Y.-N., D. L.-Y. Dong, G. S. Hayward, and S. D. Hayward. 1990. The EBV Zta transactivator: a member of the bZip family with unique DNA binding specificity and a dimerization domain that lacks the characteristic heptad leucine zipper motif. *J. Virol.* **64**:3358-3369.
- Dasmahapatra, B., E. J. Rozhon, and J. Schwartz. 1987. pBD7, a novel cell free expression vector with efficient translation initiation signal. *Nucleic Acids Res.* **15**:3933.
- Frappier, L., and M. O'Donnell. 1991. Overproduction, purification, and characterization of EBNA-1, the origin binding protein of EBV. *J. Biol. Chem.* **266**:7819-7826.
- Gahn, T. A., and C. L. Schildkraut. 1989. The Epstein-Barr virus origin of plasmid replication, ori-P, contains both the initiation and termination sites of DNA replication. *Cell* **58**:527-535.
- Gibson, W. D., A. I. Marcy, J. C. Comolli, and J. Lee. 1990. Identification of precursor to cytomegalovirus capsid assembly protein and evidence that processing results in loss of its carboxy-terminal end. *J. Virol.* **64**:1241-1249.
- Hager, D. A., and R. R. Burgess. 1980. Elution of proteins from SDS-PAGE, removal of SDS and renaturation of enzymatic activity: results with δ subunit of *E. coli* RNA polymerase, DNA topoisomerase and other enzymes. *Anal. Biochem.* **109**:76-86.
- Inoue, N., S. Harada, T. Honma, T. Kitamura, and K. Yanagi. 1991. The domain of Epstein-Barr virus nuclear antigen 1 essential for binding to oriP region has a sequence fitted for the hypothetical basic-helix-loop structure. *Virology* **182**:84-93.
- Jones, C. H., S. D. Hayward, and D. R. Rawlins. 1989. Interaction of lymphocyte-derived Epstein-Barr virus nuclear antigen EBNA-1 with its DNA binding sites. *J. Virol.* **63**:101-110.
- Kieff, E., and D. Liebowitz. 1990. Epstein-Barr virus and its replication, p. 1889-1920. *In* B. N. Fields, D. M. Knipe, et al. (ed.), *Virology*, 2nd ed. Raven Press, New York.
- Kim, J., C. Zwieb, C. Wu, and S. Adhya. 1989. Bending of DNA by gene regulatory proteins: construction and use of a DNA bending vector. *Gene* **85**:15-23.
- Kimball, A. S., G. Milman, and T. D. Tullius. 1989. High resolution footprints of the DNA-binding domain of Epstein-Barr virus nuclear antigen-1. *Mol. Cell. Biol.* **9**:2738-2742.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**:680-685.
- Landschulz, W. H., P. F. Johnson, and S. L. McKnight. 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science* **240**:1759-1764.
- Lieberman, P. M., M. C. Schmidt, C. C. Kao, and A. J. Berk. 1991. Two distinct domains in the yeast transcription factor IID and evidence for a TATA box-induced conformational change. *Mol. Cell. Biol.* **11**:63-74.
- 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.
- Milman, G., A. L. Scott, M.-S. Cho, S. C. Hartman, D. K. Ades, G. S. Hayward, P. G. Ki, J. T. August, and S. D. Hayward. 1985. Carboxy-terminal domain of the Epstein-Barr virus nuclear antigen is highly immunogenic in man. *Proc. Natl. Acad. Sci. USA* **82**:6300-6304.
- O'Shea, E. K., R. Rutkowski, and P. S. Kim. 1989. Evidence that the leucine zipper is a coiled coil. *Science* **243**:538-542.
- Patel, L., C. Abate, and T. Curran. 1991. Altered protein conformation on DNA binding by Fos and Jun. *Nature (London)* **347**:572-575.
- Polvino-Bodnar, M., J. Kiso, and P. A. Schaffer. 1988. Mutational analysis of Epstein-Barr virus nuclear antigen 1 (EBNA-1). *Nucleic Acids Res.* **16**:3415-3435.
- Rasmussen, R., D. Benvenuto, E. K. O'Shea, P. S. Kim, and T. Alber. 1991. X-ray scattering indicates that the leucine zippers is a coiled:coil. *Proc. Natl. Acad. Sci. USA* **88**:561-564.
- Rawlins, D. R., G. Milman, S. D. Hayward, and G. S. Hayward. 1985. Sequence specific DNA-binding of EBNA-1 to clustered sites in the plasmid maintenance region. *Cell* **42**:859-868.
- Reisman, D., and B. Sugden. 1986. Transactivation of an Epstein-Barr viral transcriptional enhancer by the Epstein-Barr viral nuclear antigen 1. *Mol. Cell. Biol.* **6**:3838-3846.
- 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.
- Sample, J., L. Brooks, C. Sample, L. Young, M. Rowe, C. Gregory, A. Rickinson, and E. Kieff. 1991. Restricted EBV protein expression in Burkitt lymphoma is due to a different Epstein-Barr nuclear antigen 1 transcriptional initiation site. *Proc. Natl. Acad. Sci. USA* **88**:6343-6347.
- Sauer, R. T., R. Jordan, and C. O. Pabo. 1990. λ Repressor: a model system for understanding protein-DNA interactions and protein stability. *Adv. Protein Chem.* **40**:1-61.
- Schaefer, B. C., M. Woisetschlaeger, J. L. Strominger, and S. H. Speck. 1991. Exclusive expression of EBV nuclear antigen 1 in Burkitt lymphoma arises from a third promoter, distinct from the promoters used in latently infected lymphocytes. *Proc. Natl. Acad. Sci. USA* **88**:6550-6554.
- Shuman, J. D., C. R. Vinson, and S. L. McKnight. 1990. Evidence of changes in protease sensitivity and subunit exchange rate on DNA binding by C/EBP. *Science* **249**:771-774.
- Smith, P. R., and B. E. Griffin. 1992. Transcription of the Epstein-Barr virus gene EBNA-1 from different promoters in nasopharyngeal carcinoma and B-lymphoblastoid cells. *J. Virol.* **66**:706-714.
- Sugden, B., and N. Warren. 1989. A promoter of Epstein-Barr virus that can function during latent infection can be transactivated by EBNA-1, a viral protein required for viral DNA replication during latent infection. *J. Virol.* **63**:2644-2649.
- Tan, S., and T. J. Richmond. 1990. DNA-binding-conformational change of the yeast transcriptional activator PRTF. *Cell* **62**:367-377.
- Weiss, M. A., T. Ellenberger, C. R. Wobbe, J. P. Lee, S. C. Harrison, and K. Struhl. 1990. Folding transition in the DNA binding domain of GCN4 on specific binding to DNA. *Nature (London)* **247**:575-578.
- Williams, T., and R. Tjian. 1991. Characterization of a dimerization motif in AP-2 and its function in heterologous DNA binding proteins. *Science* **251**:1067-1071.
- Wysokenski, D. A., and J. L. Yates. 1989. Multiple EBNA-1 binding sites are required to form an EBNA-1-dependent enhancer and to activate a minimal replicative origin within oriP of Epstein-Barr virus. *J. Virol.* **63**:2657-2666.
- 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.
- Yates, J. L., and S. M. Camiolo. 1988. Dissection of DNA replication and enhancer activation functions of Epstein-Barr virus nuclear antigen 1. *Cancer Cells* **6**:197-205.
- Yates, J. L., and N. Guan. 1991. Epstein-Barr virus-derived plasmids replicate only once per cell cycle and are not amplified after entry into cells. *J. Virol.* **65**:483-488.
- Yates, J. L., N. Warren, and B. Sugden. 1985. Stable replication of plasmids derived from Epstein-Barr virus in a variety of mammalian cells. *Nature (London)* **313**:812-815.