Research Capsule

Epstein-Barr Virus Transcription Factors

Alison J. Sinclair1 and Paul J. Farrell
Ludwig Institute for Cancer Research, St. Mary’s Hospital Medical School, Norfolk Place, London W21PG, United Kingdom

Background

EBV2 is estimated to infect 90% of the world’s population and is by far the best example of a common human virus that efficiently immortalizes human cells. The principal target cells for EBV infection are human B-lymphocytes and oropharyngeal epithelial cells. The virus immortalizes B-lymphocytes and persists in those cells predominantly as a nonproductive, latent infection (reviewed in Ref. 1). Typically, 10–30 copies of the 172-kb circular episomal virus DNA genome are maintained in the nucleus of the latently immortalized B-lymphocyte. The viral DNA is assembled into nucleosomes and replicates once per cell cycle in S phase using the maintenance origin of replication ori-P and the virus protein EBNA-1. Virus gene expression uses mainly the cell polymerase II and III transcription apparatus, and virus mRNAs are polyadenylated and spliced like cell mRNAs. EBV also encodes a few transcription regulatory proteins itself, and these are the topic of this review. So far, the virus proteins most clearly shown to be transcription regulatory molecules are EBNA-1 and EBNA-2 in the latent (immortalizing) cycle and BZLF1 and BRLF1 in the productive virus cycle. The location of these genes in the viral genetic map and the nomenclature of EBV genes have been described elsewhere (2, 3).

Activation of the EBV Productive Cycle

About 60 EBV genes are expressed during the EBV productive cycle, and these are organized into a cascade; the three groups in the cascade that can be recognized easily are usually called immediate early, delayed early, and late genes. Late genes are expressed after productive DNA replication, and early genes begin expression before productive replication. Immediate early genes are transcribed in response to the inducer of the productive cycle without any intervening requirement for protein synthesis. The original insight into the genes involved in activation of the productive cycle came from analysis of rearranged defective EBV genomes in the P3HR1 cell line (4–6) and identified BZLF1 and BRLF1 as the important genes. The activation of the EBV productive cycle in lymphocytes is now most studied in the latently infected Akata BL cell line (7, 8). These cells express surface IgG and respond rapidly to cross-linking the surface immunoglobulin with anti-immunoglobulin antibody. This technique is used to simulate the binding of the (unknown) cognate antigen to the surface immunoglobulin and results in many changes in the cell, including efficient activation of the EBV productive cycle. The promoters for BZLF1 and BRLF1 (7, 9, 10) respond to the cross-linking without any intervening protein synthesis (so they are the immediate early promoters), and the mechanism of this promoter activation may be related to a natural mechanism of reactivation of EBV in lymphocytes (11). BZLF1 and BRLF1 are the immediate early genes and are the key activators of the productive cycle.

Target Genes for BZLF1 and BRLF1

Several target genes for activation by the immediate early BZLF1 and BRLF1 genes in the delayed early group have been identified. These include the promoters for BSLF2 + BMFL1 (EB2) (12–14), BMRF1 (15), the DR promoter, and the DL promoter (6, 16–20). No late promoters activated directly by BZLF1 or BRLF1 have yet been reported, but a careful study of this has not yet been made. The BZLF1 protein also activates its own promoter (21) and the promoter for BRLF1 (22). Deletion and site-directed mutagenesis experiments in several of the promoters regulated by BZLF1 (14, 15, 20, 21, 23, 24) and BRLF1 (14, 18, 19, 25–27) showed that their effects are mediated through short DNA sequence elements. BZLF1 is a promoter factor and BRLF1 is an enhancer factor. Some promoters have been shown to be regulated negatively by BZLF1 [for example, the Cp (BC-R2) (28, 29) and Wp (BW-R1) (28) promoters and the promoter for the LMP gene of EBV (28)]. BZLF1 and BRLF1 can also up-regulate the human immunodeficiency virus long terminal repeat (30, 31), but the significance of this is unclear. All of the delayed promoters activated by BZLF1 that have been studied in detail so far are also targets for BRLF1, and there are some clear cases of cooperation between the two factors (15, 18, 32).

BZLF1 is a b-Zip Transcription Factor

The BZLF1 protein is also sometimes known as ZEBRA, EB1, Zta, and Z. The protein from the prototype B95-8 strain of EBV has a predicted molecular weight of 27,000 but migrates at 38 kd on SDS gels. This may partly reflect modification of the protein, but BZLF1 translated in vitro in reticulocyte lysate comigrates on SDS gels with BZLF1 from cells infected with B95-8 EBV (13), so any modification is probably quite restricted. There is some evidence for phosphorylation, but the sites and stoichiometry of modification have not been determined accurately (33). There is also some minor size variation of BZLF1 between EBV strains, but there is no evidence linking this to altered protein activity. The 1.0 kb mRNA for BZLF1 has been cDNA cloned and contains three exons (5, 13, 34). A second mRNA 2.8 kb in length (encoding the BRLF1 protein) overlaps the BZLF1 mRNA, but there is contradictory evidence on whether BZLF1 is

Received 4/20/92.

1 To whom requests for reprints should be addressed.

2 The abbreviations used are: EBV, Epstein-Barr virus; kb, kilobase(s); kd, kilodalton(s); SDS, sodium dodecyl sulfate; cDNA, complementary DNA; EBNA, Epstein-Barr nuclear antigen.

that is clearly a member of the b-Zip family (37) of proteins, although it has some unusual features, particularly with regard to its dimerization domain. The first specific DNA binding site (36) has been identified. The second specific DNA binding site has been found to be identical to the c-jun/c-fos AP-1 site, but subsequent analysis showed that BZLF1 can bind to a range of DNA sites, and its specificity is similar to but not identical to that of c-jun/c-fos (19, 24, 38–41). Individual alteration of nucleotides within the BZLF1 target and measurement of binding affinity produced the consensus shown in Fig. 1 (42). The most frequent variant high affinity site, which is located within the BZLF1 promoter, will also bind the C/EBP transcription factor (42). Chimeric proteins between BZLF1 and other bZip proteins show that the sequence specificity of DNA binding is determined by the basic motif part (Fig. 2) of the DNA binding domain (32, 41, 42). The only sites mapped for BZLF1 in the cell genome are in the c-fos promoter (43), but one might expect BZLF1 to find many binding sites in cell promoters.

**Dimerization of BZLF1**

Cross-linking studies and gel retardation analysis show that the active form of BZLF1 is a homodimer, and deletion studies map the dimerization domain of the COOH terminus, downstream of the DNA binding domain (24, 39–42, 44). Heterodimerization of BZLF1 with other b-Zip proteins does not seem to occur. The activation domain is at the NH₂ terminus (39, 45, 46). These three domains of function correspond very closely to the three-exon structure of the BZLF1 gene, and the exon boundaries of the middle exon of BZLF1 containing the "basic motif" DNA contact domain are almost identical to the boundaries of the third exon of human c-fos, which also contains its DNA binding element (36). There are short but significant sections of protein sequence homology between the first two exons of BZLF1 and parts

---

**Fig. 1.** A, the consensus BZLF1 binding site. B, alignment of DNA sequence of BRLF1 binding sites. Conserved sequences are underlined. C, the DNA sequence of the most frequent EBNA-1 binding site in the EBV genome (in the ori-P family of repeats). Arrows mark the position of the inverted repeat.
of c-fos (36) and between the third exon of BZLF1 and the dimerization domain of C/EBP (32, 42).

The striking feature of the dimerization domain of BZLF1 is the absence of the heptad leucine repeat found in other members of the b-Zip family of transcription factors. The dimerization domain of BZLF1 is functionally equivalent to the zipper dimerization domains of GCN-4 and C/EBP in the sense that chimeras can be made of the DNA binding and dimerization domains, and these still bind DNA with good affinity (42). There is also some evidence that the part of the BZLF1 dimerization domain (amino acids 200–222) closest to the DNA binding region forms a coiled coil structure analogous to the leucine zipper. Point mutagenesis of residues in the positions equivalent to the leucine heptad repeat interferes with dimerization, whereas some other point mutations do not prevent dimerization (42, 44). The BZLF1 residues equivalent to the heptad leucine repeat are Tyr, Lys, Leu, and Met, and these are probably able to make the hydrophobic interactions necessary for the dimerization. The interheptad residues clearly are also important for determining dimerization in leucine zipper structures in c-jun/c-fos (47), and further mutagenesis and direct physical analysis will be required to confirm the structure of the BZLF1 dimerization domain. Although kinked α-helical structures containing proline residues are known (48), it still seems likely that proline 223 (Fig. 2) would cause a termination of the α-helix. Sequences COOH-terminal of proline 223 further stabilize dimerization in moderate salt conditions (24), so the data are most consistent with dimerization via a relatively short stretch of α-helical coiled coil further stabilized by interaction of the most COOH-terminal sequences. The ability of a monoclonal antibody to BZLF1 to recognize both monomeric and dimeric forms of the protein although binding to an epitope close to proline 223 (49) suggests a structure there which is unaffected by the dimerization, but this could still be compatible with the model.

Trans-activation

The NH2-terminal part of the protein contains the trans-activation domain. Deletion analysis and mutagenesis (24, 39, 45, 46) have shown amino acids 28–78 to be the most important for trans-activation, but this region does not show an obvious similarity to other acidic or very proline rich activation domains. The clearest insight into the mechanism of the activation domain has come from a demonstration of interactions between BZLF1 and TBP, part of the cell TATA box binding transcription factor (46). BZLF1 binds recombinant TFII D (TBP), but the TFII D associated proteins (TAFs) are required to confer BZLF1 responsiveness to TBP in standard in vitro transcription reactions.* So, in summary, it is likely that BZLF1 functions as a dimer, binding a defined DNA target and directly contacting TBP in the transcription complex. Synergy between BRLF1 and BZLF1 (15, 18, 32) and c-myc and BZLF1 (35) raises the possibility of other important interactions of BZLF1, but none has yet been defined in detail. Less of the trans-activation region (amino acids 25–42) seems to be required for synergy with BRLF1 than for simple trans-activation by BZLF1 (32). Some proteins that may bind specifically to BZLF1 have been described (50), but these require further characterization.

BRLF1 Is an Enhancer Factor

BRLF1 (also known as Rta and R) is a 605-amino acid protein with no obvious homology to other proteins. Despite a predicted molecular weight of 66,000, the protein migrates with an apparent molecular weight of 94,000/98,000 when isolated from EBV positive B-lymphocytes (51) or when expressed in an in vitro translation system (26, 34). The structure of the mRNA for BRLF1 has been determined by cDNA cloning and S1 mapping. It contains several exons, but the entire coding sequence is within the second exon (5, 22, 23).3

In vitro analysis of BRLF1 function revealed that BRLF1 is a sequence specific DNA binding protein which interacts with sequence elements within the DR enhancer (26, 52) and the BMLF1 enhancer (27). The three R response elements (RRE) (Fig. 1) consist of two regions of partial homology separated by six to eight bases. The BRLF1 response element has not been identified in the other targets of BRLF1. Mutational analysis of the protein (Fig. 3) identified the DNA binding domain (52); amino acids 1–320 give full DNA binding, and 1–280 retain some DNA binding activity. BRLF1 dimerizes in solution and interacts with DNA as a homodimer (52). The dimerization domain (amino acids 1–232) has also been analyzed by mutation (52); dimerization is essential for DNA binding but can also occur without DNA binding (52). The carboxy-terminal half of the protein is required for trans-activation. The trans-activation domain has been divided into two regions by mutagenesis (52), one is a proline rich region (amino acids 352–515), and the other is acidic (amino acids 515–605). BRLF1 is a powerful enhancer factor that clearly functions through its defined DNA target element and binds that sequence specifically (26, 27, 52). However, its binding to DNA is much weaker and less stable than that of BZLF1, and the suspicion remains that other unidentified factors will contribute to its DNA binding stability in the cell.

---

* Q. Zhou, P. Lieberman, and A. Berk, personal communication.
EBV Latent (Immortalizing) Cycle

EBV infection of primary B-lymphocytes results in the expression of a subset of viral genes, establishment of viral latency, and immortalization of the cells to generate lymphoblastoid cell lines, a process accompanied by many changes to B-lymphocyte surface markers. Of the six viral nuclear proteins (EBNAs) expressed during immortalization, EBNA-2 and EBNA-1 most clearly affect transfection, but there is some evidence that EBNA-3C (53) and EBNA-LP also have gene regulatory effects.

EBNA-2

EBNA-2 is one of the earliest EBV genes to be expressed after infection, and it is essential for the immortalizing phenotype (54, 55). It is a nuclear phosphoprotein which appears to be associated with large complexes in cells (56). There are two recognized virus types in EBV, and this type variation is reflected in EBNA-2. EBNA-2 type A (type 1) from cells infected with EBV migrates at 82/87 kD on SDS gels, and type B (type 2) runs at 75 kD. EBNA-2 has some effects on cell phenotype including transactivation of the B-lymphocyte activation antigens CD21 (53, 57), CD23 (53, 57, 58), and the cell protooncogene c-igr (59). EBNA-2 is also required for the switch in viral promoter usage from the promoter used to express EBNA-2 during the first stage of infection, Wp, to the promoter used to express all of the EBNA genes in established cell lines, Cp (60). The trans-activation of the EBNA genes through Cp (60–62) results in an auto-stimulatory loop in EBNA gene expression during latency. EBNA-2 is also responsible for trans-activation of the promoters for the latent viral genes LMP1, LMP2A, and LMP2B (LMP2A and 2B were originally called TP1 and TP2) (63–68). Part of the role of EBNA-2 in B-lymphocyte immortalization assays may be via the trans-activation of LMP1, since LMP1 is itself active in rodent fibroblast transfection assays, and LMP1 up-regulates many B-lymphocyte activation markers (53, 69).

Extensive mutational analysis (Fig. 4) of type A (type 1) EBNA-2 revealed that most of the protein is essential for the immortalizing phenotype (55, 66, 70, 71), but two small regions (amino acids 200–234 and 463–483) are dispensable. Large deletions show concordance between loss of immortalizing activity and loss of trans-activation of LMP1 (70), suggesting that trans-activation of viral and cell genes by EBNA-2 plays an important role in immortalization. A small part of the protein (amino acids 426–462) acts as a very powerful transcription activation domain when linked to a gal4 DNA binding domain (71), and this section is likely to be an important determinant of the trans-activation properties of EBNA-2, even though EBNA-2 has not been demonstrated to bind to DNA.

Analysis of the promoters trans-activated by EBNA-2 has identified B-lymphocyte specific, EBNA-2 responsive elements between the promoters for LMP1 and LMP2B (62, 65–67), within the promoters for CD23 (58), and in Cp (60, 61). These regions span several hundred base pairs, and no specific EBNA-2 response sequence element has yet been identified. The two types of EBNA-2 trans-activate LMP1 and CD23 differentially, (53, 58, 66), which may reflect important differences in the mechanism of transcriptional activation of these two genes. Deletion analysis of the LMP1 promoter indicates that, in this case, EBNA-2 works by overcoming the action of a cell repressor (63, 67). Deletion of the target sequence for the putative repressor results in constitutive promoter activity. EBNA-2 does not seem to bind DNA specifically, and the identity of the repressor and potential interaction with EBNA-2 remain to be elucidated.

EBNA-1 Activates an Enhancer in ori-P

The EBNA-1 protein of EBV has mostly been studied because of its role in maintenance replication of the EBV genome. It binds to the ori-P replication origin (72) and is the only EBV protein required for maintenance replication (73). EBNA-1 is a high affinity, sequence specific DNA binding protein, and the most common high affinity palindromic core binding site in EBV is shown in Fig. 1. Considerable variation from this sequence is possible while retaining good binding (74, 75). The ori-P comprises two sections, the dyad symmetry element and the family of repeats, both of which bind EBNA-1 (74–78). The dyad symmetry element is where replication is initiated (79), and the family of repeats consists of a tandem array of about 20 copies of a 30-base pair repeat containing the EBNA-1 binding site. In common with some other viral DNA replication origins, the family of repeats acts as a powerful transcription enhancer (79–81). This activity is dependent on the binding of EBNA-1, and at least seven copies of the repeat are required for efficient trans-activation. The Cp (BC-R2) promoter, which is located about 3 kb away from the ori-P enhancer, is up-regulated by EBNA-1/ori-P (82), and this provides a feedback loop of activation, since the Cp promoter is responsible for EBNA-1 expression in B-cell lines immortalized with EBV.

Deletion analysis of the EBNA-1 protein (Fig. 5) has shown that the COOH terminus (amino acids 459–607) is required for DNA binding (83–85) and that DNA binding is required for trans-activation. EBNA-1 binds DNA as a dimer (84). A well defined nuclear localization signal lies in the region 379–386 (85). About one-third of the EBNA-1 protein sequence (the IR-3 segment amino acids 90–328) is composed of a simple repeat of Gly-Gly-Ala, and most of this can be deleted without major effect (84). There is some confusion about which sequences other than the DNA binding domain are required for the trans-activation. In one system measuring trans-activation in Vero cells, a truncated protein containing only amino acids 450–641 seems to work better than the full length protein (85), but other work testing the effects of small deletions in the NH2 terminus of the protein in CV-1 cells shows an additional role for part of the NH2 terminus of the protein (84).

Controlling the Decision between Immortalization and Virus Production

EBV is a herpesvirus, and a common strategy of herpesviruses is to combine active replication in one tissue type with long term latent persistence in another cell type. Most EBV replication and shedding are thought to occur in diffuse epithelial sites within the oropharynx. This epithelial replication, which can be observed readily in oral hairy leukoplakia, does not seem to be accompanied by conventional latent gene expression, so this may be a simple productive infection (49). Lymphocytes, in contrast, always seem to proceed first to the latent (immortalizing) virus life cycle with only occasional spontaneous induction to virus production. It appears that the host
cell transcription factor environment must determine the course of infection. Expression of EBNA-2 from Wp is arranged to occur very early in infection of lymphocytes, and EBNA-2 has a central role in coordinating the establishment of immortalization through trans-activation of LMP1, LMP2B, Cp, CD23, and perhaps c-igr. In some respects, it is surprising that the productive cycle genes can be kept switched off so effectively during latent immortalization. The silencer element upstream of the BZLF1 promoter (86) may be important in this regard, but a lack of positive cell factors associated with activation may also be relevant. Although the effects of changes in cell transcription factors on the BZLF1 and BRLF1 promoters are probably the immediate mechanism of virus reactivation, autoactivation of the promoters by BZLF1 strongly accelerates the promoter activity (21, 22). This is the essence of the switch between immortalization and the virus productive cycle during reactivation. The rising levels of BZLF1 stimulate further BZLF1 transcription (21) and simultaneously inhibit transcription of the EBNA genes involved in immortalization (29). The down-regulation of EBNA-as relies on the dependence of Cp on up-regulation by glucocorticoids acting through a glucocorticoid response element with the promoter (87). BZLF1 interferes with this up-regulation (29). The mechanism of this effect is indirect and seems to work via activation of the c-fos promoter by BZLF1 (43). The resulting increase in c-fos levels (29) may cause inhibition of the activity of the glucocorticoid receptor (88).

In general, herpesviruses are composed of a core set of delayed early and late genes that are common to many herpesviruses combined with genes unique to each individual virus. The unique genes in EBV are mainly the immediate early transcription factors and the genes involved in immortalization. These reflect adaptation of the virus to its host B-lymphocytes and epithelial cells. Understanding the mechanisms of these virus control systems will provide a unique insight into cell type specific control of gene expression and the regulation of a common human pathogen.

References


