ACQUISITION AND LOSS OF CHROMATIN MODIFICATIONS

DURING AN EPSTEIN-BARR VIRUS INFECTION

ANNE SCHMEINCK

ERKLÄRUNG

Hiermit erkläre ich, dass die vorliegende Arbeit mit dem Titel

„ACQUISITION AND LOSS OF CHROMATIN MODIFICATIONS DURING AN EPSTEIN-BARR VIRUS INFECTION"

von mir selbstständig und ohne unerlaubte Hilfsmittel angefertigt wurde, und ich mich dabei nur der ausdrücklich bezeichneten Quellen und Hilfsmittel bedient habe. Die Arbeit wurde weder in der jetzigen noch in einer abgewandelten Form einer anderen Prüfungskommission vorgelegt.

München, 28. April 2011

Anne Schmeinck

BRIGHT LIGHTS IN BLACK HOLES

CONTENT

1. INTRODUCTION

1.1   Epstein-Barr virus – discovery and basic principles

Epstein-Barr virus (EBV), also called human herpesvirus 4 (HHV 4), belongs to the family of γ-herpesviruses. EBV infects very efficiently human B cells due to the strong interaction of the viral glycoprotein gp350/220 with the complement receptor CD21, which is presented on the surface of B cells and certain epithelial cells (Nemerow et al., 1987).

EBV's discovery is closely related to an important characteristic of the virus: the association with different human tumors and lymphomas. In 1958, Denis Burkitt investigated a malignant tumor of children in equatorial Africa (Burkitt, 1958), which became to be known as Burkitt's lymphoma (BL). Because of the conspicuous coincidence of this tumor disease with the distribution area of malaria, he assumed the connection of an infectious agent spread by arthropods to this disease (Burkitt, 1962). In 1964, the virologist Anthony Epstein and his student Yvonne Barr as well as R. J. Pulvertaft were able to cultivate a B cell line out of those tumor samples (Epstein et al., 1964; Pulvertaft, 1964). Epstein was capable to prove the existence of herpesvirus-like particles in those cells by electron microscopy (Epstein et al., 1964). The classification of EBV as a new virus strain succeeded in 1966 by the work of Gertrud and Werner Henle (Henle et al., 1966), who named the new virus after its discoverers Epstein and Barr. Further serological assays identified EBV as etiological agent of infectious mononucleosis and as a common feature of another malignancy, i.e. NPC[*](Henle et al., 1968; zur Hausen et al., 1970). EBV is also associated with a number of other human tumors including B-cell malignancies such as Hodgkin's disease (HD) and lymphoproliferative

---

[*] undifferentiated nasopharyngeal carcinoma

disease, some T-cell lymphomas, and epithelial tumors such as gastric carcinoma (summarized in Young et al., 2003). All the connections of EBV to human tumor malignancies led to the decision of the WHO[*] to classify EBV as a group I carcinogen in 1997 (IARC, 1997).

Another characteristic of EBV is its success in its host: More than 95% of all adults are seropositive for EBV and carry a lifelong infection. What makes this virus so effective? The key to EBV's success in infecting and remaining in host cells is an identifying feature of γ-herpesviruses: EBV establishes a strict latency upon infection of resting human B-lymphocytes and induces the indefinite proliferation of infected cells. EBV subdivides its life cycle into a latent phase, where the virus expresses only a subset of its genes for the steady establishment of its DNA as an extrachromosomal episome in the nucleus of infected cells, and a second productive, lytic phase, during which progeny virus is produced ready for the infection of new host cells (Kieff et al., 2007; Rickinson et al., 2007). As a consequence, EBV can circumvent the host immune response during latency but still propagates and infects new cells after induction of the lytic phase.

Certain genes expressed during early infection and mainly during latency are functional analogues of cellular genes, all of them are important for cellular transformation and genome maintenance. It is a principal viral strategy to mimic cellular processes, which contributes to the very successful and lifelong infection of human B-lymphocytes with EBV.

1.1.1   EBV and cellular mimics

EBV's important characteristics are an effective infection of host cells, a lifelong maintenance in those cells, and a successful spreading of viral progeny to new target cells. These features include (i) the prevention of apoptosis and the downregulation of the immune system after infection, (ii) the growth transformation of infected cells, (iii) the gain of EBV genome stability in infected cells, and (iv) the possibility to switch to a productive infection, where newly synthesized virus ready for the infection of new B-lymphocytes is produced. EBV genes that are cellular mimics control theses concepts. This suggests the existence of cellular ancestors of EBV genes that the virus has acquired during co-evolution with its host.

(i)   Prevention of apoptosis and downregulation of the immune system early after infection

Two important mechanisms of the immune system control viral infections at an early stage:

---

[*] World Health Organization

the induction of apoptosis and the establishment of a cytotoxic T cell response. Apoptosis is a direct host cell reaction to infections that actively kills the cell prior to viral maturation. EBV encodes two mimics of the cellular anti-apoptotic BCL-2[*] protein to intervene with this process. The very early expression of the *BCL*-2 homologs *BALF1* and *BHRF1* is a mechanism to immediately evade apoptosis early after infection and is mandatory to initiate and maintain latency (Altmann et al., 2005). A second anti-apoptotic effect is obtained by another adaption of cellular principles: the expression of viral miRNAs[†]. Recently it could be shown that EBV's miRNAs from the *BHRF1* locus inhibit apoptosis and favor cell cycle progression early after infection (Seto et al., 2010).

A further mechanism to control viral infections is the activation of the host immune system to induce a cytotoxic T cell response. Infected cells degrade intracellular viral proteins proteolytically and present the viral peptides in association with the MHC[‡] class I molecule on the surface of infected cells to CD8[+] cytotoxic effector T-lymphocytes. EBV's *BCRF1* locus encodes a viral homolog to the cytokine IL10[§], which modulates and regulates the expression of the TAP transporter[**] (Zeidler et al., 1997). TAP proteins play an essential role in the transport of peptides derived from proteolytic degradation in the proteasome (Neefjes et al., 1993). The transport of loaded MHC class I molecules to the cell surface is impaired by downregulation of TAP proteins, evading the cytotoxic immune response of T cells that recognize viral antigenic epitopes. EBV has adopted a mechanism to escape the cytotoxic T cell response by mimicking the cellular cytokine IL10 (Zeidler et al., 1997).

(ii)   Growth transformation of the infected cells

EBV has to drive resting B-cells into proliferation before it establishes stable latency. One important protein, which has transforming potential and is necessary for B cell immortalization is the EBV-encoded latent membrane protein 1 (LMP1). LMP1 is a signaling molecule, which has limited homology but shares high functional similarity with cellular CD40, which is a co-stimulatory protein found on antigen presenting cells and required for their activation. LMP1 mimics many, if not all, downstream signaling cascades of CD40, transforms infected B cells, and governs the process of immortalization. One important difference between the viral and the cellular signaling molecule is the mode of activation.

---

[*] B-cell lymphoma 2
[†] microRNA
[‡] major histocompatibility complex
[§] interleukin 10
[**] transporter associated with antigen presentation

CD40 needs the protein CD40L as a ligand for the induction of signaling, which is present on helper T cells, but LMP1 is constitutively active through oligomerization of its C-terminal[*] domain by clustering of the N-terminal[†] transmembrane domain (Gires et al., 1997; Dirmeier et al., 2003).

A second signaling molecule, which EBV has adopted from human cells, is the latent membrane protein 2a (LMP2a). Like LMP1, LMP2a is constitutively active, which is again accomplished by clustering of its transmembrane domain. The cytoplasmatic tail of LMP2a resembles the signaling domains of the B cell receptor (BCR). Somatic hypermutation takes place during the germinal center reaction. Mutations are randomly inserted into the immunoglobulin genes. These mutations can lead to cells expressing BCRs with higher affinity. However, most of the inserted mutations result in cells with no or nonsense BCRs. These cells are usually negatively selected and undergo apoptosis. In order to growth transform and activate these proapoptotic B cells, EBV has to provide this missing signal in BCR-negative B cells. This is realized by the expression of LMP2a in latently infected cells (Mancao et al., 2005; Mancao et al., 2007). Recently it was shown that the very early expression of BZLF1, a functional homolog of members of the AP-1[‡] family of transcription factors, mimics the AP-1 function by the activation and G1-to-S transition of EBV infected cells during the early phase of infection (Kalla et al., 2010). As a consequence, BZLF1 contributes to the activation and growth transformation of EBV infected B cells as well.

(iii)   EBV's genomic stability in infected cells

EBV needs to ensure that its genome is efficiently and exactly copied in each round of mitosis to maintain genetic stability, once the cells are activated and proliferate. This is achieved by EBV's origin of plasmid replication (*oriP*), which is a *cis*-acting element of the viral episome essential for DNA replication, nuclear retention, and viral gene regulation. The functions of *oriP* depend on EBNA1[§], which is a mimic of cellular HMG1a[**], a chromatin component (Yates et al., 1984; Marechal et al., 1999; Sears et al., 2004). HMG1a preferentially binds to the minor groove of AT tracks in the DNA sequence and induces conformational changes. Recently it was shown, that HMG1a recruits the origin recognition complex (ORC) to artificial origins of replication, demonstrating its essential role in the local formation of a

---

[*] carboxy-terminal
[†] amino-terminal
[‡] activator protein 1
[§] EBV-encoded nuclear antigen 1
[**] high mobility group protein 1a

prereplication complex (Thomae et al., 2008). Like HMG1a, EBNA1 is able to target ORC to *oriP* and is thus essential for the maintenance of the EBV episome in latently infected B cells (Schepers et al., 2001).

(iv)   Induction of the lytic phase

EBV is compelled to possibly switch to the productive, lytic phase of infection in order to allow the horizontal spread of EBV from cell to cell or from individual to individual. The production of progeny virus allows increasing the pool of latently infected B cells, but also enables the infection of formerly EBV-negative people. The lytic switch is controlled by the AP-1 like transcription factor BZLF1, which has the extraordinary function to sequence-specifically bind to and activate promoters of essential lytic EBV genes, even when they carry methylated CpG dinucleotides (Bhende et al., 2004; Bergbauer et al., 2010). The cognate sequences are termed "meZREs" (<u>me</u>thylated B<u>Z</u>LF1-<u>r</u>esponsive <u>e</u>lement<u>s</u>) and can be found predominantly in promoters of early lytic viral genes. With this unique protein, EBV has developed a clever mechanism to escape latency, once BZLF1 is expressed (Kalla et al., 2010).

In summary, EBV has evolved to adopt many cellular mechanisms for the infection, activation, and maintenance of its DNA in host cells, using cytokines, signaling molecules, anti-apoptotic proteins, and miRNAs. In my work I wanted to concentrate on yet another host cell mechanism, which EBV uses to regulate its gene expression. I already embraced the importance of CpG-methylation for the induction of the lytic phase of EBV. I aimed to further understand how epigenetic mechanisms influence the regulation of viral gene expression. Epigenetics is a growing field and technical progress facilitated many new insights recently. Key principles of epigenetic gene regulation are introduced in the next chapter.

1.2   Epigenetics

In multicellular organisms, all cells originate from the same pluripotent stem cell. Although they carry the same genetic information, their gene expression profile changes dramatically depending on cell type, cell function, environmental influences, or growth phase. One principle question in biology is therefore: *"How do we get from one genotype to more than 200 phenotypes?"*

Conrad Hal Waddington, a developmental biologist, proposed in 1957 the concept of the epigenetic landscape. This is a metaphor for the process of cellular decision-making during

development (illustrated in Fig. 1.1). In this image, the cell is represented as a marble, which is rolling down a hill. On the way down, the marble can take specific trajectories and consequently ends in different grooves at the foot of the hill. Conrad Waddington imagined the development of differentiated cells to be equal. Specific decisions in the developmental process would lead to different cell fates (Waddington, 1957; reviewed in Goldberg et al., 2007). By coining the subject of epigenetics,
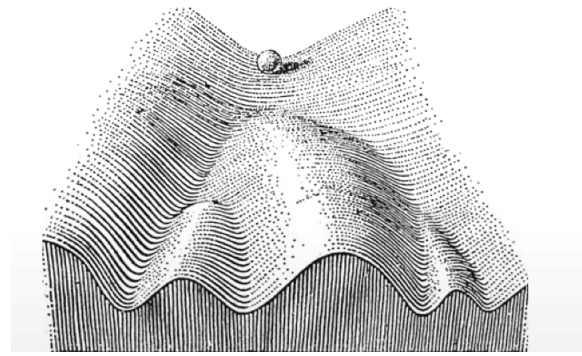
Fig. 1.1 Waddington's epigenetic landscape
The metaphor explains the development of a differentiated cell from a pluripotent stem cell. The cell is illustrated as a marble, which can take different ways on its way down the hill. The outcome or the cell fate is defined by the decision that is taken at each crossing point of the groove (Waddington, 1957).

Waddington started to link developmental biology to the discipline of genetics, but without giving an explicit molecular definition of the principle of epigenetic changes. Several decades later, specific mechanisms were proposed, which suggested that information is superimposed on DNA sequences (Holliday, 2006). Today the word epigenetics is used to describe the study of heritable changes in genome function that occur without a change in DNA sequence (The Epigenome Network of Excellence (NOE), 2004-2008). But what defines the nature of these DNA changes?

The detection of the structural details of the DNA double helix is one of the landmark discoveries in biology (Watson et al., 1953) and was the basis for the understanding of how genetic information is stored and propagated to the next generation. The fact that DNA is not freely distributed but organized in chromosomes inside the cell already implies that genetic information is not only stored, but also controlled by the structural organization of the DNA molecule. The molecular actors participating in epigenetic regulation are covalent modifications of single nucleotides and histone tails, noncoding RNAs (ncRNAs), transcription factors, and DNA binding proteins. In contrast to changes of the DNA sequence, most chromatin states are remarkably reversible and transient and are altered during aging or in response to environmental factors like physical exercise and diet. An epigenetic mark does not stand on its own, the sum of modifications defines, whether a certain chromatin segment is "active" or "poised" to be active (euchromatin) or whether the chromatin is "silent" (heterochromatin). Euchromatin largely consists of coding sequences, which only accounts for a small fraction of a mammalian genome (Lander et al., 2001). Although not necessarily

active, euchromatin is in a decompacted or open configuration and therefore accessible for transcription factors or other DNA binding proteins. Heterochromatin reflects the major part of the mammalian genome and consists mainly of noncoding and repetitive sequences. In contrast to euchromatin, heterochromatin is tightly packed and condensed, making it inaccessible for enzymes like RNA polymerase II (RNA Pol II) or DNA-binding factors like transactivators.

One can distinguish three different layers of epigenetic modifications: DNA methylation of CpG dinucleotides, the positioning of nucleosomes, and the influence of histone modifications. These layers are exemplified in Fig. 1.2 (Jones et al., 2008) and will be discussed in detail in the following sections.
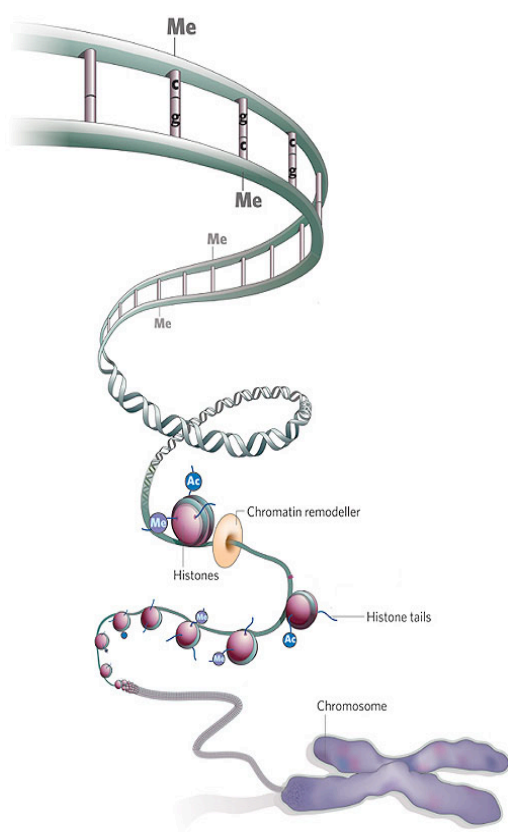
Fig. 1.2 Levels of epigenetic regulation
DNA methylation of CpG dinucleotides is the first level of epigenetic regulation. The accessibility of DNA for transcription factors is further affected by the organization of DNA into open areas (euchromatin) ore more closed patches (heterochromatin), which is determined by the position of nucleosomes. Chromatin remodelers influence their positioning, therefore affecting transcriptional regulation. Histone modifying enzymes provide a further stage of organization into active and silent chromatin by setting activating or repressing marks on the tails of histone proteins (Jones et al., 2008).

1.2.1   DNA methylation

Site-specific DNA methylation is an epigenetic mark found in organisms across all domains of life. In higher eukaryotes, DNA is frequently methylated at the C5 position of cytosines in a cytosine-guanine (CpG) dinucleotide. The methyl-group of 5-methylcytosine (5mC) is located in the major groove of the DNA, where it does not interfere with the Watson/Crick base pairing capacity of nucleotides (reviewed in Hermann et al., 2004a).

DNA methylation is usually linked to gene repression, which can be interpreted by different means. An obvious explanation is that DNA methylation interferes with the basal binding of the transcription machinery to DNA. Several transcription factor binding sites contain CpG dinucleotides in their cognate sequence. Binding to the cognate sequence is impaired by methylation of the cytosine (Watt et al., 1988). A second way of repression is the attraction of methyl-CpG-binding proteins (MBPs). This protein family comprises MeCP2, MBD1, MBD2, MBD3, and MBD4. MeCP2 is the founding member and contains a methyl-CpG-binding domain (MBD) and a transcriptional repression domain (TRD). It is capable of binding to a single, symmetrically methylated CpG dinucleotide. Silencing of transcription can occur over a distance of several hundred base pairs by the establishment of a repressive chromatin environment with the help of the TRD (Bird et al., 1999). Interestingly, the target sites of different MBPs are largely non-overlapping and seem to have sequence specificity. For example, MeCP2 prefers CpG sites, which are flanked by AT-rich regions (Klose et al., 2005).

Functions of DNA methylation

The majority of 5mCs is found in repetitive DNA elements, suggesting that cytosine methylation has evolved as a defense mechanism against transposons and parasitic elements. Transposons represent at least 35% of the genome and can be held inactive through methylation of their promoter sequences. Over time, C to T transition mutations through hydrolytic desamination at 5mCs destroys many of the intragenomic parasites (Yoder et al., 1997). Beyond that, methylation is also necessary for the monoallelic expression of imprinted genes, which are normally expressed from only one of the two alleles according to the sex of the parent that contributed that allele. Failures in imprinting at specific loci are responsible for certain human disorders (reviewed in Paulsen et al., 2001). Also X-chromosome inactivation, the process by which one of the two X-chromosomes of females is downregulated by means of dosage compensation, depends on CpG-methylation (Panning et al., 1996).

DNA from mammalian somatic tissues is heavily methylated: More than 70% of all CpG sites carry 5mCs. Most sequences are methylated according to their frequency of CpG dinucleotides, like the repetitive regions mentioned above. CpG islands are an exception from this rule. CpG islands are commonly located in promoter regions of genes, GC rich (> 55% GC content), and about 0.4 - 3kb in length. They are associated with both tissue-specific and housekeeping genes and they are unmethylated or lightly and variably methylated in all tissues. The fact that they are unmethylated in tissues with active and inactive gene expression

leads to the assumption, that their activity is not only controlled by methylation (Goll et al., 2005). But methylated CpG islands are strongly and heritably repressed.

DNA methylation patterns are highly deregulated in cancer; failures in the repression of oncogenes through demethylation or the inactivation of tumor suppressors through promoter methylation can contribute to tumorigenesis.

DNA methyltransferases implement DNA methylation

Enzymes catalyzing the addition of a methyl-group to the cytosine molecule are termed DNA methyltransferases (DNMTs). So far, four members of DNMTs have been identified in mammals: DNMT1, DNMT2, DNMT3a, and DNMT3b. Additionally, DNMT3L has been characterized as a stimulator for DNMT3a/b.

DNMT1 is called the maintenance DNMT, because it copies the methylation signature of the parental cell DNA to the daughter cell DNA in each round of replication. This is achieved by DNMT1's preference for hemimethylated DNA (Hermann et al., 2004b). The N-terminal regulatory domain of DNMT1 contains a PCNA binding domain, which can interact with PCNA[*], the processivity factor for the DNA polymerase. The association of DNMT1 with the replication machinery is proposed to be the mechanism to couple the maintenance of DNA methylation to DNA replication (Leonhardt et al., 1992; Chuang et al., 1997). The estimated kinetics of DNA methylation by DNMT1, which is approximately 3-4 times slower than the progression of DNA replication, argued against a stable interaction of DNMT1 with PCNA (Maga et al., 2003). Later it was shown that the interaction of both proteins is transient and highly dynamic (Schermelleh et al., 2007).

DNMT2 is the most widely conserved methyltransferase, with homologues in protists, plants, fungi, and animals. Its function is still enigmatic and controversially discussed. The catalytic domain of the protein clearly predicts its function as a DNA methyltransferase, the structure undoubtedly resembles the bacterial DNA methyltransferase M.HhaI[†] (Dong et al., 2001), and also the DNA binding capacity of the protein was proven. Various studies could show that DNMT2 possesses DNA methyltransferase activity, although this activity is very weak (reviewed in Jeltsch et al., 2006). DNMT2 is also present in organisms that do not possess any sign for DNA methylation at all (Gutierrez et al., 2004). This suggests that DNMT2 has additional functions not related to DNA methylation. A study in 2006 supported this

---

[*] proliferating cell nuclear antigen
[†] DNA cytosine C5 methyltransferase from Haemophilus hemolyticus

hypothesis: DNMT2 was found to have a prominent tRNA[*] methyltransferase activity (Goll et al., 2006). DNMT2 is not a canonical RNA methyltransferase, because it does not share any significant sequence similarity with known RNA methyltransferases. Merely, DNMT2 utilizes a DNA methyltransferase mechanism for the methylation of tRNA, which suggests that the enzyme belongs to the DNMT family, but might have changed its substrate during evolution (Jeltsch et al., 2006).

DNMT3a and DNMT3b are both considered as *de novo* methyltransferases. Unlike DNMT1, they do not differentiate between unmethylated and hemimethylated substrates (Okano et al., 1998; Gowher et al., 2001). Disruption of both genes in mouse ES cells[†] is lethal, which confirms their importance during development (Okano et al., 1999). Both proteins partially substitute for each other, as the individual knockouts of the genes have less severe phenotypes than the double knockout. However, they have distinct biological roles. While DNMT3b is responsible for methylation of pericentromeric satellite regions (Okano et al., 1999), DNMT3a appears to be important for the methylation of single copy genes, retrotransposons (Bestor et al., 2004), and for the establishment of genomic imprinting during germ cell development (Kaneda et al., 2004). Mutations in *DNMT3b* are responsible for the ICF syndrome[‡] in humans. Lymphoblastoid cell DNA from ICF patients shows reduced genomic methylation, specifically in repetitive DNA sequences associated with pericentromeric chromosome regions (Hansen et al., 1999; Ehrlich, 2003).

Latest discoveries in DNA methylation research

Although CpGs appear to be the main targets for methylation in differentiated cells, a recent study identified exceptions to this principle. Non-CpG-methylation was identified as a common feature of ES cells in a genome-wide screen. Nearly 25% of DNA methylation in pluripotent cells was in a non-CpG context, which disappeared after induction of differentiation. Non-CpG-methylation is probably maintained by continual *de novo* methyltransferase activity. This could be a key principal for the maintenance of cellular pluripotency (Lister et al., 2009).

Recent studies start to deny 5mC to be the exclusive covalent change of DNA with an epigenetic impact. In 2009, two studies reported the occurrence of 5-hydroxymethylcytosine (hmC) as the "6th base" in DNA in ES cells and in neuronal cells (Kriaucionis et al., 2009;

---

[*] transfer RNA
[†] embryonic stem cells
[‡] immunodeficiency, centromeric instability, and facial abnormalities

Tahiliani et al., 2009). It was shown that proteins of the TET[*] family are able to catalyze the conversion of 5mC to hmC. In mouse ES cells it was demonstrated that a knockdown of *TET1* leads to downregulation and promoter methylation of the protein NANOG, which is a protein important for ES cell self-renewability and maintenance. These results implicate a role for hmC in the regulation of the DNA methylation status. It is possible that the conversion of 5mC to hmC is the first step of a locus-specific demethylation process, but the search for an enzyme that further processes hmC into an unmethylated cytosine is still ongoing (Ito et al., 2010). Current research focuses on finding methods to distinguish between 5mC and hmC. First results on the distribution of the 6th base in a genome-wide manner have already been published (Song et al., 2010).

1.2.2   Nucleosomes: More than just beads on a string

According to textbook knowledge, packaging of DNA into higher order structures is necessary to achieve high compaction of DNA in the nucleus. But already in the 1980s the idea arose that chromatin could be inhibitory to the binding of proteins, for example transcription factors, which are needed to synthesize mRNA in the cell. It seemed likely that nucleosomes were important for a stable gene expression profile of a cell, despite of providing the DNA with structural stability and compactness.

Nucleosomes are the primary repeating unit of chromatin. They usually consist of a core of histone proteins with two copies each of the canonical histones H2A, H2B, H3, and H4. They are wrapped nearly twice by 147bp of DNA and separated by a linker region, which is usually staffed with histone H1 (Kornberg, 1974; Luger et al., 1997). The synthesis of the canonical histone proteins is tightly coupled to the S-phase of the cell cycle to assure efficient packaging of DNA into nucleosomes during replication. The fact that histone genes are encoded in clusters and in multiple copies is very likely a way to ensure the expression of proper amounts of histones in the cell during replication (Marzluff et al., 2002). Histone chaperons are responsible for the replication-coupled positioning of histones. CAF-1[†] for example is a histone chaperone that facilitates the incorporation of histone H3-H4 dimers into the DNA, which is the first step of nucleosome assembly. CAF-1 can interact with PCNA similar to the maintenance methyltransferase DNMT1, which implies a link between nucleosome assembly and the progression of DNA replication (Loyola et al., 2004). Nucleosomes can encompass several histone variants, which alter the chromatin structure and

---

[*] ten-eleven translocation protein
[†] chromatin assembly factor 1

function. For example, CENP-A[*] is a variant histone that is necessary for centromer formation, while histone H3.3 specifies active chromatin. Histone variants are usually deposited in a replication-independent manner, which involves the local eviction of a nucleosome and the replacement with a variant protein (Marzluff et al., 2002).

Principles of nucleosomal positioning: Insights from genome-wide studies

Genome-wide mapping of nucleosomes in different organisms from yeast, fly, worm and humans start to uncover the principle behind the positioning of nucleosomes (Yuan et al., 2005; Barski et al., 2007; Mavrich et al., 2008; Shivaswamy et al., 2008; Valouev et al., 2008; Lantermann et al., 2010). The most apparent outcome of these studies is the difference in the nucleosome density between regulatory regions and transcribed genes. In the model organism *Saccharomyces cerevisiae* (*S. cerevisiae*), 90% of all promoters contain stretches of DNA with very low nucleosome occupancy, and most of them own a nucleosome-depleted region (NDR) at the transcriptional start side (TSS). Downstream of the TSS, DNA is usually occupied by well-phased nucleosomes. The so-called "+1-nucleosome" next to the TSS is very accurately positioned. Further upstream and downstream, nucleosomes gradually loose their phasing; an exact position is difficult to assess (Yuan et al., 2005). The stereotypical pattern of nucleosomes in promoter regions is more or less conserved throughout eukaryotic genomes. This observation caused the question of how nucleosomes become positioned and whether the mechanism of positioning is conserved throughout the species as well. In *S. cerevisiae*, the contribution of the DNA sequence alone seems to have a very high influence on the positioning of nucleosomes (Kaplan et al., 2009): Nucleosomes intrinsically oppose poly-AT tracks. But already the nucleosome occupancy map of *S. cervisiae's* close relative *Shizosaccharomyces pombe* (*S. pombe*) revealed positioning mechanisms distinct from those of *S. cerevisiae*. Proteins like nucleosome remodelers, histone variants, and NDR boundary elements appeared to have a higher impact on the position of nucleosomes. In multicellular eukaryotes, the mechanisms must be even more difficult, as the nucleosome coverage at certain promoters can be cell type specific.

The concepts mentioned above commonly apply for constitutively active promoters, which usually encode for housekeeping genes. The analysis of nucleosome occupancies at inducible promoters, which are regulated depending on certain environmental influences, provides insights into mechanisms of gene regulation by nucleosome positioning.

---

[*] centromer protein A

Effects of nucleosome positions on the activation state of inducible promoters

The analysis of different nucleosome configurations on inducible promoters helps to understand the effect of nucleosome positioning on gene expression. The most well known example of the impact of nucleosomes on an inducible promoter is probably the phosphate-regulated *PHO5*[*] promoter of *S. cerevisiae*, which becomes activated upon phosphate starvation and results in a 50-fold increased production of secreted acid phosphatase. In the repressed state, the promoter is assembled into an array of well-positioned nucleosomes, disrupted by a NDR of 80bp in length. The NDR contains one of the two binding sites for the transcription factor PHO4. Upon activation, nucleosomes -1 to -4 disassemble from the promoter, mediated by the ATP-dependent histone chaperone ASF1[†]. As a consequence, the promoter exposes an additional binding site for PHO4 and allows the transcription machinery to access the TATA box. Interestingly, a failure of nucleosome reassembly at the *PHO5* promoter after the addition of phosphate results in a persisting expression of the gene, although PHO4 disappears from its corresponding binding sites. This indicates that the nucleosome positions are solely responsible for the repression of the promoter (Svaren et al., 1997; Adkins et al., 2004).

Another example is the *GAL1-10* promoter[‡] of *S. cerevisiae*. Four binding sites of the activator GAL4 are covered by a complex of the nucleosome-remodeling enzyme RSC[§] and an unstable nucleosome, which allows efficient binding of GAL4. This complex is flanked by an array of nucleosomes on each site, which is rapidly removed upon activation. But unlike the *PHO5* promoter, nucleosomes are not the only determinants of the expression state of the genes because the promoter is silent even in the absence of nucleosomes (Bryant et al., 2008; Floer et al., 2010).

There are certain other promoters of *S. cerevisiae*, which are known to behave similarly to the promoters described above (Bai et al., 2010). Equally well-studied examples of promoters of higher eukaryotes are rare. Nevertheless, research on human T cells during the G_0 to G_1 transition provides hints that signatures of active chromatin, including a relatively low abundance of nucleosomes in the TSS, are already set prior to induction of these genes (Smith et al., 2009). A further study on the transcriptional induction of mammalian primary response genes by Toll-like receptors in macrophages provides new insights into distinct modes of regulation in different types of inducible promoters. This study identified two major

---

[*] phosphate metabolism gene 5
[†] anti-silencing function
[‡] galactose metabolism genes 1-10
[§] remodels the structure of chromatin

functional classes of primary response genes: Class I is characterized by the abundance of CpG islands in the promoter region, which facilitates promiscuous induction from constitutively active chromatin without the requirement for SWI/SNF[*] nucleosome remodeling complexes. The high CpG content appeared to be responsible for the promoter to assemble into unstable nucleosomes, which might lead to the independence from nucleosome remodelers for the induction of genes. The second identified class of promoters does not contain CpG islands and assembles into stable nucleosomes. This class of promoters strictly depends on the remodeling of nucleosomes by SWI/SNF complexes prior to transcription and has therefore the capacity of being tightly regulated (Ramirez-Carrozzi et al., 2006; Ramirez-Carrozzi et al., 2009). The authors hypothesize that cells could benefit from CpG island promoters because of the possibilities for constitutive expression and rapid induction without the requirement for energy consuming remodelers. This parallels the fact that many of these promoters are activated by common transcription factors like NF-κB[†] and MAP kinases[‡], which are targeted by a large number of growth factors. In contrast to this, promoters, which are assembled into stable nucleosomes, are often activated by transcription factors with highly specialized function. A stable repressed chromatin configuration of the promoters could have the benefit of minimizing basal transcription that would be harmful to the cell.

An example for a human inducible promoter is the *IFN*-β[§] locus, which is activated in response to virus infection for example. Activation requires the assembly of an "enhancosome", a protein complex that recruits chromatin remodelers and modifiers. Upon recruitment, one nucleosome slides in dependency on a remodeling complex, opening the promoter for the binding of the general transcription machinery to the TATA box (Agalioti et al., 2000; Lomvardas et al., 2002).

Chromatin remodelers can jostle nucleosomes

Nucleosome positions change depending on the activation state of the gene, particularly in inducible promoters. Efficient and specific modification of nucleosome positions requires the help of chromatin remodelers. Chromatin remodelers are enzymes, which use energy from ATP hydrolysis for the repositioning, reconfiguration, or eviction of nucleosomes. Chromatin remodelers create a chromatin architecture that either allows or represses the binding of transcription factors and RNA Pol II to target sites in the promoter (reviewed in Saha et al.,

---

[*] switch/sucrose nonfermentable
[†] nuclear factor 'kappa-light-chain-enhancer' of activated B-cells
[‡] mitogen-activated protein
[§] interferon beta

2005). Eukaryotes contain at least five families of chromatin remodelers (SWI/SNF[*], ISWI[†], NURD/Mi-2/CHD[‡], INO80[§], and SWR1[**]), which fulfill distinct roles in the establishment of a given chromatin environment. The two best-studied families are the SWI/SNF and the ISWI family of chromatin remodelers.

SWI/SNF remodelers contain a conserved ATPase subunit and a set of five additional core proteins (Vignali et al., 2000). They primarily disorganize nucleosomes by sliding and ejecting them from promoters and their presence is therefore often associated with gene activation. In conjunction with this feature, SWI/SNF complexes own domains that can recognize and bind to acetylated histone tails. This modification is linked to active gene expression (Martens et al., 2003).

The ISWI family members of chromatin remodelers also share a conserved ATPase domain but have unique associated proteins depending on the function of the remodeling complex. Their purpose is mainly to organize nucleosomes into uniformly spaced arrays by evaluating the length of the linker region between two nucleosomes and sliding them until a certain distance is achieved. As a consequence, they generally contribute to gene repression (Langst et al., 2001). ISWI complexes contain two specialized domains in their C-terminus: The SANT[††] domain recognizes histone tails, which lack activation marks like acetylation and guides the remodeling complex to genes that are prone to be repressed. The SLIDE[‡‡] domain assists to determine the length of the linker region for the estimation of the distance the nucleosome has to slide (Grune et al., 2003).

Both families share mechanistic features: They (i) bind to the nucleosome at a defined position close to the dyad, (ii) induce ATP-dependent conformational changes and break proximal histone-DNA contacts to create a DNA wave on the octamer surface, (iii) move the DNA wave in a directed manner along the histone core, and (iv) propagate the DNA wave around the nucleosome by one-dimensional diffusion (Saha et al., 2005).

---

[*] switch/sucrose nonfermentable
[†] imitation swi
[‡] nucleosome remodeling factor/ chromodomain helicase DNA-binding
[§] inositol 80
[**] sick with rat8 ts
[††] switching-defective protein 3 (Swi3), adaptor 2 (Ada2), nuclear receptor co-repressor (N-CoR), transcription factor (TF)IIIB
[‡‡] SANT-like ISWI (imitation switch) domain

1.2.3   Histone modifications

Not only histone proteins influence the strength of DNA binding to the histone octamer, but also modifications, which can be introduced to their N-terminal tails. Unlike the major part of the histone, which is globular in structure, the histone tails are unstructured, very flexible, and poke out of the very compact base of the nucleosome. This feature makes them highly accessible for histone-modifying enzymes, which can dynamically and rapidly establish a huge variety of modifications to the tails and consequently influence and manipulate different chromatin functions like transcription, DNA repair, DNA replication, and chromosome condensation. The character of modifications is very diverse and include the addition of a small chemical compound (acetylation, phosphorylation, or methylation of certain amino acid residues) or the incorporation of large and bulky peptides (ubiquitination, sumoylation, and ADP-ribosylation). There are four core histone proteins, at least eight modes of modifications, the possibility to modify different amino acids in the histone tail, some of which can be singly, doubly, or triply modified. Together the mode of different modifications is very high and reflects possible chromatin states. Nevertheless, the use of modification-specific antibodies in Chromatin Immunoprecipitation (ChIP) assays makes it possible to identify many different modifications and link them to different functions.

Histone modifications and their mode of action

Histone tails can be modified by (i) acetylation, (ii) phosphorylation, (iii) methylation, and (iv) the addition of bulky products. The modifications are specified by the histone protein that is targeted, e.g. histone "H3", the amino acid residue that acquires the modification, e.g. "K4" for lysine four, and the mode and multiplicity of the modification, e.g. "me3" for trimethylation. The correct description of trimethylation at histone H3 and lysine 4 is thus H3K4me3. Modifications have diverse functions in the regulation of cellular pathways like DNA damage repair, cell cycle control, or replication. The focus of this overview will be on transcriptional regulation.

(i) Acetylation of lysine residues in histone tails commonly correlates with active transcription. Acetylation neutralizes the positively charged lysine and eases the strong binding of the basic histone proteins to DNA. The acetylation reaction is carried out by histone acetyltransferases (HATs) but can be reverted by the action of histone deacetylases (HDACs). Many coactivators and corepressors possess either HAT or HDAC activity. For

example, the yeast SAGA complex[*] comprises HAT activity and assembles TATA-box binding protein (TBP) associated factors. Thereby the protein directly links histone acetylation to transcriptional initiation (Grant et al., 1998). A special protein domain, called bromodomain, can recognize acetylated histones. Bromodomains are found in many transcriptional coactivators but also in SWI/SNF chromatin remodeling complexes (Hassan et al., 2002).

(ii) Phosphorylation of serine and threonine residues in histone tails is a modification that can influence the affinity of DNA and histones by introducing negative charges to the histone tail. The modification correlates with chromosome condensation during mitosis and meiosis (Xu et al., 2009). It is also present on promoters, where it can modulate and promote other modifications. For example, phosphorylation of histone H3T6 by PKC-β[†] in androgen-receptor controlled gene activation leads to the prevention of the erase of methylation at histone H3K4 by the demethylase LSD1[‡]. Histone H3T6 phosphorylation contributes by this to keep the promoter active (Metzger et al., 2010).

(iii) Methylation of histone tails is the most complex epigenetic mark, because it can occur on different lysine or arginine residues. It can have multiple methylated states on each residue and it includes activating and repressive functions. Methylation of lysines is set by histone lysine methyltransferases (HKMTs) and is erased by histone demethylases. Methylation at H3K4, H3K36, and H3K79 are implicated in transcriptional activation. Beyond that, H3K4 and H3K36 are likewise involved in transcriptional elongation. Trimethylation of H3K4 localizes at the 5' end of active genes and is associated with the initiated form of RNA Pol II. H3K36me3 accumulates at the 3' end of genes, and is found in association with the elongating form of RNA Pol II. One important function of H3K36me3 is the suppression of transcriptional initiation from cryptic start sites inside the gene body. This is supported by the observation that this modification represses transcription, if located in inducible promoters in their repressed state (Carrozza et al., 2005; Keogh et al., 2005). H3K79 methylation is mainly found in active promoters (Barski et al., 2007). It is an unusual modification, because it locates rather to the core of the nucleosome than in the tail. The addition of methyl-groups to this lysine residue leads to a reshaping of the nucleosomal surface near the C-terminus of histone H3, which probably effects chromatin structure and function (Lu et al., 2008).

Methylation of H3K9, H3K27, and H4K20 is connected to transcriptional repression. H3K9

---

[*] spt-ada-gcn5-acetyltransferase complex
[†] protein kinase C beta
[‡] lysine specific histone demethylase 1

methylation is involved in heterochromatin formation through the recruitment of HP1[*]. Special protein domains, called chromodomains, can read the modification, similar to bromodomains that recognize acetylated histone tails. HP1 encompasses a chromodomain and is therefore targeted to H3K9 methylated regions. H3K27 methylation is connected to three different functional states: It is found in silent euchromatic gene loci that are regulated by Polycomb proteins, in pericentromeric heterochromatin, and in the inactive X chromosome in mammals. H4K20 methylation is found in pericentromeric heterochromatin. Functionally, the monomethylated form of this modification is important for DNA repair *via* binding of the DNA checkpoint protein CrB2[†] in *S. cerevisiae* (Martin et al., 2005).

Methylation of arginines is involved in both positive and negative regulation of transcription. For example, arginine methylation is important in the *pS2* gene, which is a target of estrogen signaling. Stimulation leads to the recruitment of the estrogen receptor to the *pS2* promoter, and it brings along the two methyltransferases CARM1[‡] and PRMT1[§], which methylate residues in the tails of histone H3 and histone H4. Concomitant to the occurrence of those histone marks, active RNA Pol II can be found on the promoter. Only minutes later, methylation and RNA Pol II disappear, but a new cycle of methylation and RNA Pol II recruitment can start right away. The flexibility and the dynamics of this process is probably the key to a rapid transcription shutoff, if needed (Metivier et al., 2003).

(iv) The addition of bulky polypeptides like ubiquitin or SUMO[**] to histones is distinct from the modifications described above, because it increases the size of a histone by approximately two-thirds. Histone ubiquitination has been linked to transcriptional activation and repression, depending on the context, the multiplicity of addition, and the histone protein modified. An example of an activating function is the connection of histone H2B ubiquitination to histone H3K4 methylation, which is well studied in yeast. The ubiquitin ligase is recruited by polymerase associated factor I (PAFI) to sites of active transcription and elongation. The ubiquitination of histone H2B at lysine 123 leads to the subsequent recruitment of a histone methyltransferase, which trimethylates H3K4 (Wood et al., 2003). Whether these mechanisms are conserved in mammals is conversely discussed and still needs to be clarified (reviewed in Atanassov et al., 2010).

This section is just a rough overview about the best-studied modifications of histone tails, but

---

[*] heterochromatin protein 1
[†] crumbs homolog 2
[‡] coactivator-associated arginine methyltransferase
[§] protein arginine N-methyltransferase 1
[**] small ubiquitin-like modifier

there are many more. In order to understand the contribution of modifications to transcriptional activation without getting lost in the different modifications and systems, it is useful to have a look at functional key mechanisms.

Mechanisms and function of histone modifications

There are two described mechanisms of how histone modifications can have an influence on chromatin function: The first includes the alteration of physical contacts between the histone proteins and the nucleosomal DNA, which leads to unraveling or wrapping of nucleosomes. For example, introducing a negatively charged adduct like an acetyl-group leads to a loosening of the DNA from the histone proteins due to the repulsion from the DNA backbone. This makes the DNA more accessible to transcription factors. The second mode of function is the recruitment of non-histone proteins that can then further affect the accessibility to DNA. For example, trimethylation of H3K4 can be recognized by a domain of the nucleosome-remodeling factor (NURF), which belongs to the ISWI family of chromatin remodelers. The chromatin structure is opened locally, which improves the accessibility of the DNA to the transcription machinery (Wysocka et al., 2006). An example for gene repression by non-histone proteins is the binding of HP1 to methylated H3K9. This protein is bound by SUV39H1[*], an enzyme capable of catalyzing the methylation of H3K9, leading to a spread of this heterochromatin mark (Nakayama et al., 2001).

The function of histone modifications can be divided into two categories: (i) the establishment of a global chromatin environment and (ii) the coordination of special DNA-based tasks. The first category mainly embraces the organization of genetic information into eu- and heterochromatin. While some modifications like methylation of H3K9 or H3K27 are necessary to keep the chromatin in a condensed and repressed state, others are responsible for maintaining the chromatin active or poised-to-be active. Mammals have evolved proteins to insulate active and inactive chromatin from each other: The protein CTCF[†] for example is a boundary element found on borders between eu- and heterochromatin, preventing heterochromatic marks from spreading to active regions (Bell et al., 1999). DNA-based tasks encompass the regulation of transcription, replication, DNA repair, and chromosome condensation. For example, regulation of transcription is affected by transcription factors that initiate a cascade of modification events, which results in the end in the activation or repression of a gene. Namely, H3K4me3 localizes at the 5' end of active genes and is

---

[*] suppressor of variegation 3-9, homolog 1
[†] CCCTC-binding factor

associated with the initiated form of RNA Pol II and is thereby related to active transcription.

The association of certain modifications and functions suggested the existence of a histone code, comparable to the genetic code (Strahl et al., 2000). The simplest model would be a binary relationship between modification and function. But a single histone mark does not determine an outcome alone; it is the combination of modifications at a defined locus that dictates the effect on the chromatin state. A recent publication tried to find patterns of modifications linked to distinct functions. A dataset was analyzed that encompassed 38 different histone modifications, histone variant H2AZ, RNA Pol II, and CTCF binding. The group identified 51 distinct chromatin states (Ernst et al., 2010). But still, these findings will rather be the start than the end of understanding the instructions encoded in the combination of epigenetic modifications. Schreiber and Bernstein suggested a more general model in order to account for all possible epigenetic states a cell can have: Histone modifications could serve as mediators for a nuclear, DNA-associated signal transduction pathway that is comparable with cytoplasmatic signaling pathways, which are mainly propagated through protein phosphorylation (Schreiber et al., 2002).

1.2.4   The epigenetic memory

The classic definition of epigenetics is the study of heritable phenotypes that do not involve changes in the DNA sequence. This concept is very important for maintaining cell fate and phenotype, as discussed previously. The recognition of hemimethylated DNA by DNMT1 and its association with PCNA is an efficient mechanism to maintain an epigenetic state over many generations. However, the heritability of histone modifications is discussed more controversially. For each histone-modifying enzyme another one was identified that can erase the modification. The use of the term "heritable" in the definition of epigenetics was therefore revisited. Epigenetics is now considered to cover a research field that encompasses generic information, which DNA does not encode. But epigenetics provide a non-genetic memory with modifications that are stably transmitted to daughter cells during mitosis and sometimes even during meiosis, thereby affecting the next generation (Silva et al., 1988). First evidence for a regulator responsible for the maintenance of a specific cell identity by histone modifications stems from work in drosophila: Pam and Ed Lewis discovered the protein Polycomb, which is responsible for maintaining the segment specific HOX[*] gene expression pattern throughout fly development (Lewis, 1978). Polycomb group (PcG) proteins are today defined as regulators that repress gene expression by keeping a transcriptionally inactive state,

---

[*] homeobox

which is mediated by H3K27 trimethylation. PcG proteins can be divided into two distinct complexes, Polycomb repressive complex 1 (PRC1) and PRC2. The common view is that PRC2 acts as the "writer" of the repressed state. It establishes H3K27 trimethylation with the histone methyltransferases EZH1 or EZH2[*], which are part of the PRC2 complex. PRC1 is seen as the "reader" of the epigenetic state. It recognizes histone H3K27me3 with its chromodomain of CBX2, CBX4, or CBX8[†] and acts as a silencing complex by ubiquitination of histone H2A (Cao et al., 2005) or by chromatin compaction of defined nucleosome arrays (Francis et al., 2004). This model suggests that PRC1 acts downstream of PRC2. But this hierarchy does not necessarily apply for each and every locus. Although both PRCs share many binding sites in the chromatin, there are exceptions of this rule, suggesting that both complexes might also act independently of each other (Schoeftner et al., 2006; Ku et al., 2008).

The mechanism of maintaining the PcG-mediated repression during the cell cycle is discussed controversially. One theory suggests a replication-dependent inheritance of PRC2 repression, quite similar to maintenance of DNA methylation by DNMT1. The co-localization of PRC2 with PCNA during S-phase and its ability to bind to its own target, i.e. H3K27me3, suggests a binding to the parental DNA strand and the spread of this mark to newly incorporated histones on the daughter strand (Hansen et al., 2008). However, a second theory rebuts that histone modifications themselves mediate epigenetic inheritance because they may rather be the consequence and not the cause of inherited information. Instead, a default silenced state of PcG target sequences, so-called Polycomb Responsive Elements (PREs), is proposed. An active state would be marked by ncRNAs[‡] stemming from the PRE and could be propagated by specific chromatin-binding proteins or by H3.3 replacement (Ringrose et al., 2007).

Not only H3K27me3 has an epigenetic memory, mechanisms to recall the epigenetic state after cell division are also known for other histone methylation marks. The counter players of PcG proteins are Trithorax proteins, which memorize an activated state through H3K4me3. And the heterochromatin mark H3K9me3 is maintained by the function of HP1. In general, there has to be a tight regulation of the epigenetic state of genes, balancing between an essential stability of the epigenetic state to maintain cell identity *versus* a required flexibility of the cell's gene expression pattern for the ability to react to environmental changes.

---

[*] human enhancer of zeste 1/2
[†] chromobox protein 2/4/8
[‡] non-coding RNA

1.3 Epigenetics in EBV

The balance between epigenetic stability and flexibility is a conflict that has to be solved by EBV during its infectious cycle. On the one hand, latency guarantees a lifelong persistence in the host, but on the other hand, the lytic switch is mandatory to spread viral progeny to uninfected cells. Both aspects depend on different epigenetic states of the viral genome. It is therefore interesting to study epigenetics in EBV regarding the regulation of the viral life cycle.

1.3.1 EBV's life cycle

According to our working model, EBV's life cycle is comprised of three phases: the abortive lytic, the latent, and the lytic phase of infection (Fig. 1.3). During the abortive lytic phase, the viral genome is established as a plasmid-like minichromosome, a so-called episome, in the nucleus of infected cells. Characteristic for this phase is the *W* promoter (*Wp*)-inititated transcription of latent genes. But also a set of lytic genes and miRNAs can be identified. The gene products provide anti-apoptotic signals, pro-proliferative effects, or immune-inhibitory properties (Zeidler et al., 1997; Altmann et al., 2005; Wen et al., 2007; Katsumura et al., 2009; Kalla et al., 2010; Seto et al., 2010). Importantly, the full set of early lytic genes, which are responsible for lytic DNA replication of the virus, and genes coding for EBV's structural proteins are not expressed during the abortive lytic phase. As a consequence, viral progeny cannot be produced. The early expression of lytic genes is shut down eventually to establish a stable latent infection of the proliferating B cells. During latency, EBV's episome behaves similarly to the host cell chromosomes: It replicates once per cell cycle and is propagated during mitosis to daughter cells (Adams, 1987; Kirchmaier et al., 1995). Latently infected B cells exhibit a restricted gene expression profile. At least three different latency gene expression profiles are known, i.e. latency I, latency II, and latency III. All latency profiles share the expression of *EBNA1*, but differ in the expression of the other *EBNAs*, the *LMPs*, and the *EBERs*. EBNA1 protein and the *cis*-element *oriP* are sufficient but indispensable for maintaining the viral episome in the B cell (Yates et al., 1984; Yates et al., 1985). The majority of EBV-infected B cells reside in the latent phase, which enables the virus to hide from the immune system of the host. Lymphoblastoid cell lines from *in vitro*-infected B cells usually hold the latency III expression profile.

Eventually, the virus switches to the lytic phase of infection and progeny virus is released. Signaling cascades that are activated by the B cell receptor probably trigger expression of the

transcription factor *BZLF1* (Campbell, 1999). Once *BZLF1* is expressed, the activation of lytic genes is irreversible in latently infected cells: The transcription factor acts as molecular switch and induces the lytic phase in latently infected B cells. It activates the expression of the second EBV-encoded transcription factor *BRLF1* but also stimulates its own promoter.

infection of B cells with EBV

abortive lytic phase

packaging and release of progeny

establishment of the viral episome in the nucleus

expression of latent and certain **late lytic** genes but **block of early gene expression**

expression of all lytic genes

latent phase

amplification of EBV's DNA

outgrowth of lymphoblastoid cell lines (LCL's)

expression of BZLF1 leads to the activation of early genes

lytic reactivation

in vitro: **latency III program** EBNAs, EBERs, LMPs and ncRNA

lytic phase

Fig. 1.3 Three different phases comprise EBV's life cycle in infected B cells *in vitro*
The scheme shows our model of EBV's life cycle in infected B cells *in vitro*. The abortive lytic phase is characterized by the expression of latent genes and certain late lytic genes. Expression of early lytic genes and EBV's structural genes is blocked. During this phase, EBV's genome is established as an episome in the nuclei of infected cells. The latent phase encompasses a restricted gene expression profile. Lymphoblastoid cell lines (LCLs) exhibit the latency III expression profile with the expression of all *EBNAs*, *EBERs*, *LMPs* and *ncRNAs*. BCR crosslink and expression of BZLF1 initiate the switch to the lytic phase, where early lytic genes that encode for the lytic replication machinery are expressed. EBV's genome is massively amplified during lytic DNA replication. In a second step, structural proteins are made, the virus is packaged, and progeny is released, ready to start a new infectious cycle in uninfected B cells.

The two transcription factors induce individually and synergistically the expression of a set of so-called early lytic genes (Countryman et al., 1987; Shimizu et al., 1989; Feederle et al., 2000). Early lytic gene products code for the protein machinery that is necessary for lytic DNA replication of the EBV genome (Fixman et al., 1992, 1995). Lytic DNA replication takes place at discrete sites in so-called replication compartments in the nucleus of the infected cell and initiates at *oriLyt* (*origin of lytic replication*). Lytic replication leads to a 100- to 1000-fold amplification of the EBV genome (Tsurumi et al., 2005). Subsequent to

DNA replication, the full set of lytic genes is expressed, including genes encoding structural proteins. Newly synthesized EBV DNA is packaged into capsids and progeny virus is released, ready for the spread to uninfected B cells.

1.3.2   Important checkpoints of EBV's life cycle rely on epigenetic mechanisms

EBV's genome acquires epigenetic modifications similar to the host cell DNA by the cellular epigenetic protein machinery (Shaw et al., 1979; Dyson et al., 1985). It is well established that EBV uses epigenetic mechanisms to regulate its growth program during early infection. Initially, B cells support early viral gene expression by the latent *W* promoter. Expressed genes code for EBV's transforming genes, including *EBNA1* and *EBNA2* (Kieff et al., 2007). Later, EBNA1 protein induces the expression from the alternate EBNA promoter *Cp*, either directly or indirectly by binding of cellular proteins. The promoter switch is essential for the persistence of EBV in B cells, because *Wp* becomes inactive shortly after infection, presumably by epigenetic mechanisms (Tierney et al., 2000; Altmann et al., 2006). LCLs promote the latency III gene expression profile, which is characterized by the expression of all latent genes, including *EBNAs*, *EBERs*, *LMPs,* and miRNAs. *In vivo*, the latency III program leads to a strong immune response. DNA methylation and silencing of the *C* promoter ensures that a proportion of EBV-infected cells survive the cytotoxic T cell response, because the expression profile of the latency III program is replaced by the latency I program. The only promoter, which escapes epigenetic silencing, is the *Q* promoter (*Qp*). Transcription initiation from *Qp* results in the expression of *EBNA1* only and is described as latency I phenotype (Schaefer et al., 1995; Paulson et al., 1999).

Induction of the viral lytic phase is also controlled by DNA methylation. BZLF1 is the master regulator of EBV's lytic switch in latently infected B cells (Countryman et al., 1987). The transcription factor owns the outstanding ability to bind to methylated DNA (Bhende et al., 2004; Karlsson et al., 2008; Dickerson et al., 2009). We could previously show that CpG-methylation of promoters of early lytic genes is a prerequisite to allow induction of the lytic phase. Binding of EBV's transcription factor BZLF1 to the methylated promoter elements of certain viral genes abolishes repression and enables the expression of lytic genes and the production of progeny virus (Bergbauer et al.; Kalla et al., 2010).

1.4   Scope of my thesis work

Epigenetic modifications play an important role in gene regulation. The mode of modification is complex as there are many different epigenetic states, which depend on the combination of DNA methylation, nucleosome positions, and modifications of histone tails (Ernst et al., 2010). Key principles of epigenetics are known to be important in the regulation of EBV's life cycles as well: DNA methylation regulates the promoter usage in the early phase of infection but is additionally required during lytic reactivation. Histone modification patterns were shown to be characteristic for different latency types in different EBV-positive cell lines (Minarovits, 2006), but the knowledge about fundamental principles of EBV's epigenetic control during its life cycle was still poor. It was aim of my work to complement and complete the big picture of EBV's different epigenetic states in the three phases of infection and to link my results to the "snapshots" of epigenetics in EBV that were previously investigated by others.

The central hypothesis of my work was that ingenious epigenetic events govern the regulation of EBV's life cycle. We postulated that chromatin changes are the key events responsible for the consecutive steps during the early phase of an EBV infection. I collected information about the temporal establishment of epigenetic patterns on EBV's DNA to address this issue. Furthermore, I characterized the epigenetic nature in latently infected cells to understand, why and how lytic genes are repressed. A strict latency is the formula for EBV's success: A discrete gene expression profile abrogates the expression of immune-reactive viral genes evading an immune response by the host. Last, I wanted to identify the epigenetic events that allow escape from latency after BZLF1 expression, a mechanism that guarantees the horizontal spread of progeny virus for the infection of uninfected cells and hosts.

2. MATERIAL

2.1   Plasmids

The following list contains all plasmids that were used in this thesis. All of them stem from the collection of the department of gene vectors.

plasmid	description
p2089	wild type maxi-EBV: F plasmid (pMBO131) with eGFP and hygromycin resistance gene in B95.8 DNA (Delecluse et al., 1998)
p3862	pRTS-2-BZLF1-GFP: plasmid for the inducible expression of full length BZLF1 fused to GFP
p3927	pCMV:GFP:BZLF1: plasmid carrying the DNA binding domain of BZLF1 under the control of the CMV promoter

2.2   Antibodies

Antibodies that were used in this work are listed below.

specificity	species	distributor	application	amount/application
anti-H3	rabbit	Abcam (#1791-100)	ChIP	5µg
anti-H3K4me3	rabbit	Active Motif (#39159)	ChIP	5µg
anti-H3K27me3	rabbit	Upstate (#17-622)	ChIP	4µg
anti-H3K9me3	rabbit	Active Motif (#39161)	ChIP	5µg
anti-RNA polymerase II (N20)	mouse	Santa Cruz (#39097)	ChIP	10µg
anti-EZH2 (clone AC22)	mouse	Active Motif (#39875)	ChIP	3 µg
rabbit-IgG	rabbit	Upstate (#12-370)	ChIP	5 µl
anti-tubulin	mouse	Santa Cruz (#sc 23948)	ChIP	10 µl
Rpb1-CTD (8WG16)	rat	E. Kremmer/D. Eick	ChIP	100µl
CTD Ser5-P (3E8)	rat	E. Kremmer/D. Eick	ChIP	100µl
anti-RNAPII (1C7)	rat	E. Kremmer/D. Eick	ChIP	100µl
Pol II Rpb1 (Pol3-3)	mouse	E. Bautz/Heidelberg	Western	1:10
α-IgG mouse, HRP	goat	Promega	Western	1:2500
anti 5-methyl-Cytosine	mouse	Diagenode (MAb-600-500)	MeDIP	10µg

2.3   Oligonucleotides

Oligonucleotides used in this work were synthesized by Metabion, Munich. Their sequences are listed in the appendix.

2.4   Bacterial strains

For all cloning purposes and the preparation of plasmid DNA, the strain DH5-alpha (F-, lacl-, recA1, endA1, hsdR17, Δ(lacZYA- argF), U169, F80d lacZΔM15, supE44, thi-1, gyrA96, relA1; Hanahan, 1985) was used.

2.5   Eukaryotic cell lines

Cell lines used in this study stem from the collection of the department of gene vectors or were established in this work (*) and are recorded below.

name	description
B95.8	lymphoblastoid cell line, established by an *in vitro* infection of marmoset B-cells (Miller et al., 1972)
Bjab	human EBV-negative Burkitt-lymphoma cell line
Raji	human EBV-positive Burkitt-lymphoma cell line (Pulvertaft, 1964)
Raji/3927	Raji cells carrying a plasmid for GFP:BZLF1 (DNA binding domain only), clone 3927.4
Raji/3862*	Raji cells carrying a tet-inducible plasmid for BZLF1, clone 3862.1

2.6   Cell culture media and additives

2.6.1   Media for the cultivation of bacteria

Bacteria were either cultivated in liquid culture or on solid agarose plates.

appellation	application	distributor
LB-medium	medium for bacterial culture (1% (w/v) tryptone, 1% (w/v) NaCl, 0.5% (w/v) yeast extract)	Invitrogen GmbH
select agar	LB-medium with 1.5% (w/v) select agar for the cultivation of bacterial clones	Invitrogen GmbH

2.6.2 Media for the cultivation of eukaryotic cells

The following list contains media and additives for eukaryotic cell culture used in this work.

appellation	application	distributor
RPMI-1640	cell culture medium for cultivation of all cell lines used in this study	Invitrogen GmbH
Opti-MEM I	minimal culture medium for transfections	Invitrogen GmbH
alpha-thioglycerole and bathocupproine-disulfonic acid	antioxidant agents (additive for RPMI-1640)	Sigma Aldrich
fetal calf serum (FCS)	nutritive substance (additive for RPMI-1640)	Chromatech
sodium pyruvate	antioxidant agent (additive for RPMI-1640)	Invitrogen GmbH
sodium selenite	antioxidant agent (additive for RPMI-1640)	Sigma Aldrich
penicillin/streptomycin	antibiotics (additives for RPMI-1640)	Invitrogen GmbH
hygromycin	antibiotic for the selection of mutant cell lines	Invitrogen GmbH

2.7 Chemicals and enzymes

The following list shows the used chemicals and enzymes and their distributors.

appelation	distributor
RNAse A, Lysozyme	Applichem, Darmstadt
Bradford-Solution	Bio-Rad, Munich
PBS	Biochrome AG, Berlin
hygromycin B, L-alpha-lyso-lecithin (egg yolk)	Calbiochem GmbH, Frankfurt (Main)
phenol, chloroform, bovine serum albumbin fraction V (BSA) powder	Carl Roth GmbH, Karlsruhe
bacto-agar, yeast extract, trypton	Difco Laboratories, Detroit, USA
marker for DNA electrophoresis, restriction enzymes	Fermentas, Leon-Roth
$(\gamma\text{-P}^{32})$-dCTP	Hartmann Analytik, Braunschweig
Hoechst-Dye H33342	Hoechst, Frankfurt
agarose, Luria broth base, trypsin-EDTA, UltraPure TM 20x SSC	Invitrogen, Karlsruhe
ammonium sulfate, boric acid (H_3BO_3), calcium chloride (CaCl), cesium chloride (CsCl), acetic acid, ethanol, ethidium bromide, ethylene diamine tetra-acetic acid (EDTA), glucose, glycine, glycerol, isoamyl-ethanol, isopropanol, potassium acetate ($CH_3COO^-K^+$), potassium chloride (KCl), lithium chloride (LiCl), magnesium chloride ($MgCl_2$), magnesium sulfate ($MgSO_4$), sodium acetate ($CH_3COO^-Na^+$), sodium chloride (NaCl), sodium hydroxide (NaOH), sodium dodecyl sulfate (SDS), paraformaledhyde (PFA), hydrochloric acid (HCl) sucrose, Triton X-100	Merck-Eurolab GmbH
Immobilon NY$^+$ charged nylon transfer membrane	Millipore
restriction enzymes	New England Biolabs, Frankfurt
G418-Sulfate (Neomycin)	PAA Laboratories GmbH, Cölbe
RNAsin plus RNAse inhibitor	Promega, Mannheim
Complete Mini Protease Inhibitor Cocktail tablets, desoxy-nucleotides, LightCycler 480 SYBR Green I Master, High Prime	Roche Diagnostics GmbH, Mannheim
ampicillin, beta-mercaptoethanol, bromophenol blue, chloramphenicol, dimethyl sulfoxide (DMSO), dithiothreitol (DTT), doxycycline, Igepal CA-630 (NP-40), calf thymus DNA, Proteinase K, RedTaq Polymerase	Sigma-Aldrich, Munich

2.8 Buffers and solutions

The following lists records the buffers and solutions that were used in this work.

buffer/solution	ingredients	application
agarose loading dye	1mg/ml bromophenol blue, 1mg/ml xylene cyanol, 50% sucrose	general lab techniques
ampicillin stock solution	100mg/ml ampicillin in 20% ethanol, sterile filtered	
PBS, pH 7.3	10mM Na_2HPO_4, 1.8mM KH_2PO_4, 140mM NaCl, 2.7mM KCl	
SSC buffer	150mM NaCl, 15mM trisodium citrate, pH 7.0 with HCl	
TAE	40mM Tris-acetate, 1mM EDTA	
TBE	89mM Tris, 89mM boric acid, 2mM EDTA	
TE, pH 8.0	1mM EDTA, 10mM Tris-HCl	
TENS buffer	10mM Tris-HCl, 1mM EDTA, 100mM NaOH, 0.5% SDS	
MeDIP buffer	10mM sodium phosphate buffer (pH 7.0), 140mM NaCl, 0.05% Triton X-100	MeDIP
1 M sodium-phosphate, pH 7.0	0.78M NaH_2PO_4, 1.22M Na_2HPO_4	
Proteinase K digestion buffer	50mM Tris-HCl, pH 8.0, 10mM EDTA, 0.5% SDS	
wash buffer I	2x SSC + 0.1% SDS	EBV-microarray
wash buffer II	1x SSC	
wash buffer III	0.5x SSC	
depurination solution	0.64% HCl	Southern Blot
denaturation solution	1.5M NaCl, 0.5M NaOH	
church buffer	0.5M Na_2HPO_4, 1mM EDTA, 7% SDS, 1% BSA, pH 7.2 (adjust with ortho-phoshoric acid)	
SB wash buffer	0.2x SSC, 1% SDS	
lysis buffer	10mM Tris-HCl, pH 7.9, 10mM NaCl, 3mM $MgCl_2$, 0.5% NP-40	ChIP
sonication buffer	10mM Tris-HCl, pH 7.9, 200mM NaCl, 3mM $MgCl_2$, 1mM $CaCl_2$, 4% NP-40, 3mM EGTA, 1% SDS, 1x "Complete" proteinase inhibitor cocktail (Roche)	
IP dilution buffer	20mM Tris-HCl, pH 7.9, 2mM EDTA, pH 8.0, 1% TritonX-100, 150mM NaCl, 1x "Complete" proteinase inhibitor cocktail (Roche)	
low salt buffer	20mM Tris-HCl, pH 8.0, 0.1% SDS, 1% Triton X-100, 2mM EDTA, pH 8.0, 150mM NaCl	
high salt buffer	20mM Tris-HCl, pH 8.0, 0.1% SDS, 1% Triton X-100, 2mM EDTA, 500mM NaCl	
LiCl buffer	10mM Tris-HCl, pH 8.0, 0.25M LiCl, 0.5% NP-40, 0.5% deoxycholic acid, 1mM EDTA, pH 8.0	
elution buffer	25mM Tris-HCl, pH 8.0, 10mM EDTA, pH 8.0, 0.5% SDS	
2x Tris/SDS pH 8.8	22.68g Tris/Base, 2.5ml SDS (20%), ad 250ml (H_2O), pH 8.8 with HCl	western blot immunodetection
2x Tris/SDS pH 6.8	7.56 g Tris/Base, 2.5ml SDS (20%), ad 250ml (H_2O), pH 6.8 with HCl	
running buffer (10x)	60.4g Tris/Base, 288g glycine, 5ml SDS (20%), ad 2L (H_2O)	
Laemmli buffer (2x)	2% SDS, 100mM DTT, 10mM EDTA, 10% glycerol, 60mM Tris/HCl (pH6.8), 0.01% bromophenol blue, 1mM PMSF	
separation gel (6.5%)	8.6ml polyacrylamide (30%), 20ml Tris/SDS pH 8.8, 11ml H_2O, 334µl APS, 20µl TEMED	
stacking gel (3%)	1.5ml polyacrylamide (30%), 7.5ml 2x Tris/SDS pH 6.8, 5.9ml H_2O, 90µl APS, 20µl TEMED	

buffer/solution	ingredients	application
transfer buffer	60.4 g Tris/Base, 288 g glycine, 5ml SDS (20%), 200ml methanol, ad 2L (H_2O)	western blot immunodetection
blocking solution	10% (v/v) TBS, 0.1% Tween 20, 5% milk powder	
permeabilizing buffer	150mM sucrose, 50mM Tris-HCl, pH 7.9, 50mM NaCl, 2mM $CaCl_2$	micrococcal nuclease digestion
TNESK buffer	20mM Tris-HCl, pH 7.9, 200mM NaCl, 2mM EDTA, 2% SDS, 20µg/ml Proteinase K	

2.9 Commercial kits

Commercial kits that were used are listed below.

appellation	distributor
illustra Nick Columns	GE-Healthcare, Munich
Jet-Star DNA Maxi Kit	Genomed, Bad Oeynhausen
BioPrime Total Genomic Labeling System, SuperScript III First Strand Synthesis SuperMix, pCR 2.1 TA Cloning Kit	Invitrogen GmbH, Karlsruhe
NucleoBond PC 500, NucleoSpin Extract II Kit, NTB Buffer	Macherey-Nagel, Düren
QIAamp DNA Mini Kit, QIAprep Spin Miniprep Kit, RNeasy MiniKit, QiaShredder	Qiagen, Hilden
GenomePlex Complete WGA2	Sigma Aldrich, Munich
EZ DNA Methylation-Gold Kit	Zymo Research, Irvine, USA

2.10 Software

Software tools that contributed to this work are listed below.

appelation	distributor
Adobe-CS3	Adobe Systems Inc., San Jose (USA)
AIDA Alias Software	Raytest Isotopenmessgeräte GmbH, Straubenhardt
Bioconducter	Fred Hutchinson Cancer Research Center, Seattle (USA)
BiQ Analyzer HT 0.9 (beta test version)	MPI for Informatics and Saarland University, Saarbrücken
CLC Genomics Workbench	CLC bio A/S, Katrinebjerg, Denmark
DNA baser Multi-Fasta to Fasta Converter	Heracle BioSoft S.R.L., Romania
EndNote X2	Thomson Reuters, New York (USA)
FlowJo 9.1	TreeStar Inc., Ashland (USA)
FreeHand MX11.0	Adobe Systems Inc., San Jose (USA)
Genepix Pro 6.0	Molecular Devices, Sunnyvale (USA)
MacVector 11.0	Accelrys, Cambridge (UK)
Microsoft Office 2008	Microsoft, Redmond (USA)
Prism 5	Graphpad Software, La Jolla, USA
R	The R Foundation for Statistical Computing Vienna, Austria

2.11   Devices and consumables

The list below comprises devices and consumables that were used in this work.

appelation	distributor
CP100 developer machine	AGFA, Cologne
CoulterCounter Z1, ultracentrifuge L8-70M/XL-70, centrifuge Avanti J-25 and Avanti-26XP	Beckman, Heidelberg
FACS-Calibur, FACS-Aria II, cell culture flasks, 6-/12-/24-/48- and 96-well plates	Becton Dickinson GmbH, Heidelberg
Gene-Pulser II electroporation device	Bio-Rad, Munich
Branson Sonifier S-250D	Branson, Danbury, USA
Bioruptor (UCD-200)	Diagenode, Liège, Belgium
BioPhotometer, PCR machine MasterCycler Personal, reaction tubes, desk centrifuge 5415, thermomixer comfort	Eppendorf, Hamburg
cover slips, glass slides, forceps, whatman paper, Neubauer cell chamber	Hartenstein, Würzburg
H Scan-Jet G4050	Hewlett Packard (USA)
Bovivet veterinary needle	Henry Schein Vet GmbH, Hamburg
sonifier nozzle for Branson sonifier, 5 mm	G. Heinemann, Schwäbisch Gmünd
DyNAQuant Fluorimeter	Hoefer Scientific Instruments (USA)
Phosphoimager FLA 5100, film cassettes, intensifier screen	FujiFilm, Kleve
plastic consumables for cell culture and laboratory work	Greiner Bio One, Solingen
X-ray film	Kodak Company, New York (USA)
water deionizing device Milli-RO 60 PLUS	Millipore, Schwalbach
Axon GenePix 4100A Microarray Scanner	Molecular Devices (USA)
vacuum centrifuge univapo H150	Montreal Biotech Inc., Montreal, Canada
cell culture dishes, cryotubes 1.8 ml	Nunc GmH, Wiesbaden
EBV-microarray, hybridization buffer	Ocimum Biosolutions (Netherlands)
chambers for agarose gel electrophoresis, NanoDrop ND-1000, 4 mm electroporation cuvettes, reaction tubes	Peqlab, Erlangen
LightCycler 480 Real Time PCR System 454 Genome Sequencer FLX System	Roche Diagnostics GmbH, Mannheim
glass ware	Schott, Mainz
HS 4000 Pro Hybridizing Station	Tecan, Crailsheim
Heraeus Multifuge 3L-R, orbital shaker, incubators	Thermo Scientific (USA)
phase contrast microscope Axiovert 40C	Zeiss, Göttingen
glass pipettes 5 ml, 10ml and 25 ml	Brand, Wertheim
0.8 μm and 1.2 μm filter	Schleicher & Schüll, Dassel
DNA-Sequencing	Sequiserve
Roboycler Gradient 96	Stratagene, Heidelberg

3. METHODS

3.1 Bacterial culture

Escherichia coli (*E.coli*) suspension cultures were grown in LB medium at 37°C and 200rpm in orbital shakers.

For isolation of single cell clones, bacteria were streaked with an inocculating loop on LB agar plates and incubated at 37°C over night. Antibiotics (100µg/ml ampicillin, 30µg/µl chloramphenicol, or 30µg/ml kanamycin) were added to autoclaved LB medium to select for resistance genes.

For long-term storage, 25% glycerol stocks of over night cultures were frozen at -80°C.

3.2 Eukaryotic cell culture

3.2.1 Cell culture conditions

Eukaryotic cells were cultivated in 37°C incubators in an atmosphere of 5% CO_2 and 95% air humidity. All working steps were carried out in lamina hoods. Cells were centrifuged at 300g. Washing steps were performed with PBS unless stated otherwise.

Cells were grown in RPMI-1640 medium with 10% fetal calf serum, 100µg/ml streptomycin, 100 I.U.[*]/ml penicillin, 1mM sodium pyruvate, 100nM sodium selenite, 20µM bathocuproin disulfonic acid, and 0.0004% α-thioglycerole (full medium).

---

[*] international unit

Cells were counted with the CoulterCounter Z1 (Beckman Coulter) according to the instructions of the company or alternatively with a Neubauer cell chamber.

All suspension cells were cultivated at an average density of 3-5x 10^5 cells/ml. The cell lines Raji 3927 and Raji 3862 were selected with 1.5mg/ml G418 sulfate or 300μg/ml hygromycin B, respectively.

3.2.2   Storage of eukaryotic cells

Cells were frozen in 10% DMSO and 90% FCS in the gas phase of liquid nitrogen at a concentration of 1x10^7 cells/ml for long-term storage.

For gentle freezing, cells were transferred to a 2ml CryoTube (Nunc) and slowly cooled in a "Nalge Nunc Cryo 1°C Mr. Frosty Freezing Container" (Nunc) with a cooling rate of -1°C/min.

For thawing, cells were warmed to 37°C, washed with 50ml of full medium, and taken up in 5ml full medium. If required, selection pressure was applied the day after thawing.

3.2.3   Electroporation of eukaryotic cells

Raji cells were electroporated with the gene pulser II electroporator (Biorad). 5x10^6 cells were spun down and resuspended in 250μl Opti-MEM I. 5-10μg plasmid DNA were added and cells were incubated on ice for 15min. Electroporation was performed in 4mm cuvettes at 230V and 975μF. After pulsing, cells were supplied with 400μl FCS, transferred to 5ml full medium, and cultivated at 37°C.

3.2.4   Establishment of stable cell lines

Electroporated cells (3.2.3) were grown in bulk culture for two days. For the establishment of single cell clones, cells were diluted in 96 well plates and cultivated under selection for four weeks. Medium was changed if necessary, and outgrowing cells were expanded continuously. Expression of the transgene was evaluated using flow cytometry.

3.2.5   Isolation, separation, and infection of human B cells

Adenoids were repeatedly cut and washed with PBS to isolate human primary mononuclear cells (PBMCs). Human T cells were depleted from PBMCs by resuspending in 30ml PBS and 500μl sheep blood and subsequent Ficoll-Hypaque density-gradient centrifugation at 1900rpm at 10°C for 30min. B cells were collected from the interface and washed three times with

PBS. Cells were infected with B95.8 virus at an MOI of one. Cells were spun down and transferred to fresh full medium one day *post* infection.

3.2.6   Collection of B95.8 virus stocks

For the preparation of wildtype virus stocks, B95.8 cells were cultivated in high density. Virus is continuously released to the medium because a proportion of cells spontaneously undergoes lytic reactivation. Virus containing medium was collected and stored at 4°C. 1ml of supernatant contains approximately $1x10^6$ infectious particles (experienced data from the department of gene vectors).

3.2.7   Flow cytometry

Raji 3862 cells were washed with PBS and analyzed in a FACS Calibur (Beckton Dickinson) to detect and count GFP-positive cells. Data was evaluated with the FlowJo 9.1 software.

3.2.8   Sorting of GFP expressing cells

Raji 3862 cells were sorted using a FACS Aria II device according to the instructions of the manufacturer to get pure populations of cells expressing the transgene GFP. Briefly, Raji cells were resuspended in PBS with 2% FCS and 2mM EDTA to a final concentration of $3x10^7$ cells/ml and subsequently filtered to obtain single cells. Sorting criteria were (i) living cells, (ii) exclusion of duplets, and (iii) GFP positivity. Cells were sorted with a 100μm nozzle and a velocity of 6000 events/second with a sorting mask of "yield 32" and "purity 16".

3.3   Nucleic acid techniques

General methods like phenol/chloroform extraction, DNA precipitation with ethanol, electrophoretic separation of DNA, enzymatic restriction hydrolysis, or ligation of free DNA ends were performed according to standard protocols (Sambrook et al., 2001).

3.3.1   DNA purification from E.coli

(i) small-scale purification of plasmids <20 kb

Small-scale purification was performed for the evaluation of plasmid DNA after cloning. 2ml of a bacterial culture in the stationary phase were harvested through centrifugation at 4500rpm for three minutes and resuspended in 300µl TENS buffer (10mM Tris-HCl, 1mM EDTA, 100mM NaOH, 0.5% SDS). After addition of 150µl sodium acetate (3M), cells were centrifuged at maximal speed for eight minutes. DNA was precipitated with 900µl 100% ethanol and centrifugation at maximal speed for five minutes. The pellet was dried and resuspended in 50µl TE containing 10µg/ml RNase A.

(ii) high-scale purification of plasmids <20 kb

Quantitative purification of plasmids for transfection and cloning was carried out using the JetStar Maxi-Prep-Kit (Genomed) according to the manufacturer's instructions with 400ml bacterial culture in the stationary growth phase.

(iii) high scale purification of plasmids >20 kb (maxi EBV DNA)

Maxi EBV DNA was purified using the Nucleobond PC 500 Kit (Macherey-Nagel) according to the manufacturer's instructions with 400ml bacterial culture in the stationary growth phase.

3.3.2   DNA purification from eukaryotic cells

Isolation of genomic DNA was performed using the QIAamp DNA Mini Kit (Qiagen) following the manual of the company.

3.3.3   Purification of DNA from PCR products and agarose gels

Purification of DNA from PCR products was carried out with the NucleoSpin Extract II Kit (Macherey-Nagel) following the manufacturer's instructions. SDS containing samples were purified with five volumes of NTB buffer instead of the standard binding buffer.

3.3.4 Polymerase chain reaction (PCR)

Polymerase chain reaction (PCR) amplifies defined nucleic acid fragments exponentially (Mullis et al., 1992). The components of a PCR reaction and a typical PCR program are listed below.

component	Sigma Red Taq	Go Taq Polymerase
template	1ng up to 50ng	1ng up to 50ng
forward and reverse primer	100pm each	100pm each
buffer (5 x)	/	4μl
nucleotides	/	5μM each
preMix (2 x)	25μl	/
H_2O	ad 50μl	ad 20μl

step	temperature	time	cycles
denaturation	98°C	5min	1
denaturation	98°C	1min	
primer annealing	50-60°C, depending on primer pair	1min	30-40, depending on the application
elongation	72°C	500bp/min	
final elongation	72°C	10 min	1

Amplified DNA was separated on a 1% TBE agarose gel to control the size of the PCR product or for preparative gel purification (see 3.3.3).

3.3.5 Quantitative real time PCR

Immunoprecipitated DNA or cDNA was quantified with a Roche LightCycler 480 device. The detection method is based on the incorporation of the fluorescence dye SYBR-Green I into the newly synthesized DNA double helix, which is measured after each elongation cycle (Higuchi et al., 1993).

Immunoprecipitated DNA was analyzed with "absolute quantification". This method allows the calculation of the sample concentration by comparing the crossing points (Cp) of the unknown sample, in this case the immunoprecipitated DNA, with a defined standard curve, which encompasses different dilutions of input DNA. The analysis was done automatically with the LightCycler 480 software according to the second derivative maximum method. The amount of immunoprecipitated DNA was displayed as an absolute value (% input).

Evaluation of RNA expression levels was done with "relative quantification". The crossing points of different target genes were compared to the crossing point of a reference gene in different samples. The reference gene is a constitutively expressed gene or housekeeping gene, which is not regulated in all conditions tested. The analysis was done automatically with

the LightCycler 480 software in the advanced modus. The result was displayed as the ratio of the expression level of the gene of interest to the reference gene, which is a dimensionless number. Both methods included the PCR efficiencies by entrainment of internal or external standard curves. The PCR mix and the PCR program for qPCR are listed below.

component	amount
template	1µl immunoprecipitated DNA or cDNA
primer	5pm each
SYBR Green I Master (2 x)	5µl
H$_2$O	3.5µl

program	target temperature (°C)	hold (s)	acquisition mode	RampRate (°C/s)	Cycles	Analysis mode
pre-incubation	95	600	none	4.4	1	none
amplification	95	1	none	4.4		
	62	10	none	2.2	45	quantification
	72	10	none	4.4		
	75	3	single	4.4		
melting curve	97	1	none	4.4		
	67	10	none	2.2	1	melting curve
	97	/	continuous	0.11		
cooling	37	15	none	2.2	1	none

3.3.6   Isolation of RNA from cells

Isolation of RNA from cells was carried out with the RNase Mini Kit (Qiagen). 1×10^7 cells were used in each sample. The lysate was homogenized with QiaShredder columns (Qiagen). All subsequent steps were performed according to the instructions of the manufacturer. RNA was eluted in 80µl of RNase-free water and stored at -80°C.

3.3.7   Reverse transcription of RNA

RNA was transcribed into cDNA to determine expression levels of certain genes of interest. Prior to cDNA synthesis, contaminating DNA was removed from the RNA preparation with the enzyme DNase I[*] (Invitrogen). 2µg of extracted RNA were incubated for 90 min at 37°C in the presence of 2U DNase, 40U RNase inhibitor and 1x DNase buffer in a 20µl approach. Remaining DNase was heat inactivated by incubating at 65°C for ten minutes. The efficiency of DNase treatment was controlled by PCR. Reverse transcription of RNA was performed with the SuperScript III First Strand Synthesis SuperMix Kit (Invitrogen) according to the manufacturer's instructions. Quantification of cDNA was either done semiquantitatively using PCR (3.3.4) or quantitatively using qPCR (3.3.5).

---

[*] desoxyribonuclease I

3.3.8   Transfer of DNA to membranes (Southern blot)

Southern blotting was performed according to standard protocols (Sambrook et al., 2001).

DNA was separated on a 1% TAE agarose gel. The gel was shaken for 15min in depurination

buffer to facilitate the transfer of large DNA fragments to the membrane and afterwards

denatured for 30min in denaturation buffer. Southern blotting was performed *via* an upward

capillary transfer of the DNA from the agarose gel to a positively charged nylon membrane

(Immobilon NY$^+$, Millipore). The gel was placed upside down on a plastic wrap and capped

air bubble free with a membrane of the same size. Two pieces of thin whatman paper with

1cm overhang at each side were positioned on top and overlaid with several layers of paper.

The paper was changed if wet and the blot was allowed to stand for approximately three

hours. Afterwards, the membrane was washed in 2x SSC and further processed.

3.3.9   Radioactive labeling of DNA

DNA was radioactively labeled with the High Prime mix (Roche). 100ng of denatured DNA

were labeled at 37°C for one hour in a 20µl approach containing 4µl 5x High Prime mix and

100µCi α-P^{32} dCTP. The labeled probe was separated from contaminating free nucleotides by

gravity-flow chromatography with illustra NICK Columns (GE Healthcare), which are

prepacked with Sephadex G-50. The instructions of the company were followed. The purified,

radioactive labeled probe was denatured at 95°C for 10min before use.

3.3.10   DNA-DNA hybridization

DNA blots were pre-incubated in Church buffer at 60°C for one hour. The denatured, α-P^{32}

dCTP labeled probe was added and hybridized at 60°C over night with constant rotating in a

hybridization oven. The membrane was washed at least two times with Southern blot washing

buffer (0.2x SSC, 1% SDS) at 58°C for ten minutes.

3.4 Methylated DNA immunoprecipitation (MeDIP)

Methylated DNA Immunoprecipitation (MeDIP) is a method to specifically enrich DNA containing methlylated cytosine-guanine pairs (CpG dinucleotides) with an antibody recognizing methylated CpG dinucleotides. Quantitative real-time PCR or microarray hybridizations (MeDIP-on-Chip) were readouts to evaluate the enrichment of DNA as compared to input DNA.

3.4.1 Immunoprecipitation of methylated DNA (MeDIP)

DNA was prepared using the QIAamp DNA Mini Kit (Qiagen) according to the instructions of the manufacturer. MeDIP was performed as described previously (Weber et al., 2005) with minor modifications. 4µg of DNA were sheared using the Biorupter UCD-200 for five minutes, 30s on/off with 320W. After purification with the NucleoSpin Extract II Kit (Macherey-Nagel), DNA was denatured and immediately cooled on ice. Immunoselection of 4µg of this so called "input" DNA was carried out using 10µl of monoclonal 5-methylcytosine antibody (Diagenode) in a final volume of 500µl IP buffer (10mM sodium phosphate, pH 7.0, 140mM NaCl, 0.05% Triton X-100) for three hours at 4°C by overhead shaking. Immunopreciptiation was performed using 40µl of preblocked M-280 sheep anti-Mouse IgG Dynabeads (Invitrogen) for two hours at 4°C by overhead shaking. After three washing steps with 750µl IP buffer, DNA was Proteinase K treated for three hours and recovered by phenol/chloroform extraction and ethanol precipitation. DNA was resuspended in 20µl H_2O.

3.4.2 Quantification of MeDIP DNA by real time PCR (qPCR)

The DNA content MeDIP and input samples was quantified using quantitative real time PCR with the Roche LightCycler 480 system (see 3.3.5). Stepwise dilutions of total input DNA ranging from 0.01% to 10% and MeDIP-DNA were analyzed with primers for *Cp*, *EBER*, *Lmp1p*, *OriP*, *Qp*, *Wp*, *BBLF4,* and *BZLF1*. The content of MeDIP DNA was determined as compared to input DNA for each locus. Oligonucleotide sequences are listed in the appendix.

3.4.3 Genome-wide analysis of MeDIP DNA by microarray hybridization (MeDIP- on-ChIP)

DNA was amplified prior to microarray hybridization with the GenomePlex Complete Whole Genome Amplification (WGA) Kit (Sigma-Aldrich) according to the manufacturer's

instructions with minor modifications. The fragmentation step in the protocol was omitted because MeDIP-DNA was already sheared. After the reaction, DNA was checked for linear amplification by qPCR using the same primer pairs as described before (3.4.2). Labeling of enriched and input DNA was performed using the BioPrime Total Genomic Labeling System (Invitrogen). 2µg of MeDIP and input DNA were labeled with Alexa5 or Alexa3 dye, respectively. The efficiency of fluorochrome incorporation was determined at 647nm and 555nm. Labeled samples were combined, evaporated and resuspended in 130µl hybridization Buffer (Ocimum Biosolutions). The labeled samples were analyzed using our custom made EBV microarray (Kalla, 2007). One slide of the microarray contains four copies of a complete set of 285 separate, partially overlapping PCR fragments with a length of 500-700bp covering the complete genome of the prototype B95.8 EBV strain.

Hybridization of the slides was performed using a Tecan HS 4000 Pro Hybridization Station at 60°C with gentle agitation for 16h. After three washing steps with 2x SSC and 0.1x SDS, 1x SSC and 0.5x SSC at 30°C for each one minute, slides were dried with nitrogen and immediately scanned with a GenePix Personal 4100A Scanner (Axon). Analysis of visual data was performed using the GenePix Pro 6 software. The signal ratios of precipitated *versus* input DNA of the four copies of each fragment were averaged. Single spots with a standard error >0.15 were excluded from analysis. For normalization, the calculated median ratio was set at one.

For the analysis of the methylation pattern in the cell line Raji, ratios of three independent MeDIP experiments were averaged.

To analyze the kinetics of methylation pattern formation in primary B cells, the adjusted data were processed as the ratio to the constant point of reference, *Cp*, which was set at one in each experiment. The *C* promoter region is free of methylated CpGs even in established lymphoblastoid cell lines (Tierney et al., 2000). To validate the data, two biological replicates with two technical replicates each were conducted.

3.5 Bisulfite sequencing

3.5.1 Bisulfite modification of DNA

Bisulfite modification of DNA was carried out using the EZ Methylation Gold Kit (Zymo) according to the instructions of the manufacturer with minor modifications. DNA of Raji cells was sonicated prior to modification for three seconds at 10% in a Branson Sonifier to open secondary structures. 500-1000ng DNA was used in each reaction. The incubation time at 65°C was prolonged to five hours to improve the conversion rate. After purification of the bisulfite modified DNA, samples were eluted in 12µl of the supplied elution buffer.

3.5.2 PCR of bisulfite modified DNA

PCR amplification of bisulfite modified DNA was carried out using the Sigma Red Taq Polymerase in 50µl reactions with 5µl DNA and 50pmol of each primer pair. Typically, the reaction started with an initial denaturation step for 5min at 95°C. DNA was amplified in 40 cycles with one minute denaturation at 95°C, one minute annealing at 56°C, and one minute elongation at 72°C. PCR fragments were purified and directly sequenced by the Sanger method and a commercial service provider (http://www.sequiserve.de).

3.5.3 Deep bisulfite sequencing

60 PCR products of bisulfite modified DNA with 100ng each and an average product size of 463bps were pooled for deep sequencing on a Roche Genome Sequencer FLX system ("454 sequencing"). Sequencing was performed in collaboration with Microsynth GmbH (Switzerland). Library preparation and DNA sequencing was carried out according to the standard protocols of the company.

3.5.4 Data analysis

For alignment of sequences, I used the CLC genomics workbench software with parameters for high throughput sequencing in the menu "reference assembly". The local alignment was performed with parameters for long run reads (mismatch cost = 2, insertion cost = 3, deletion cost = 3, length fraction = 0.5 and similarity of 0.9). Sequence reads were aligned to a modified version of the B95.8 wildtype sequence, in which I converted all cytosines in a non-CpG context to thymines. Cytosines of CpG dinucleotides were kept unconverted. The parameters of the sequence alignment were set to tolerate RY-mismatches in the context of

CpG dinucleotides, allowing mismatches of purines and pyrimidines, which are expected if cytosines are unmethylated. Non-specific matches (situations, where a read aligns to more than one sequence, which is expected in overlapping PCR products) were placed randomly. Each contig was saved separately and the included sequences were extracted in multifasta-format. Methylation analysis was carried out using the BiQ analyzer HT 0.9/beta-test version (Bock et al., 2005). For this, multifasta-files were extracted with the "Multi-Fasta to Fasta Converter" of the "DNA baser" software. A Matlab script was developed to extract the information of the BiQ analyzer output file to identify the total number of methylated CpG dinucleotides at each position and the EBV-genome location of the CpG-dinucleotide. Details of the Matlab script are listed in the appendix. Data were graphically analyzed and visualized in the Prism software.

3.6   Analysis of nucleosome occupancy

The enzyme micrococcal nuclease (MNase) preferentially cuts the linker region of two nucleosomes. DNA, which is part of nucleosomal structures, is protected from hydrolysis if the amount of MNase is carefully titrated. After cleavage of chromatin, protected DNA can be purified and used as a sample for quantitative real-time PCR or for microarray hybridization.

3.6.1   MNase digestion of chromatin

$1x10^7$ cells were harvested, washed once with PBS, and resuspended in 250µl permeabilizing buffer (150mM sucrose, 50mM Tris-HCl, pH 7.9, 50mM NaCl, 2mM $CaCl_2$). Cells were incubated for two minutes at 37°C in the presence 0.1% lysolecithine to destabilize membranes and allow access of MNase to chromatin. MNase digestion was carried out using 50U MNase for five minutes at 37°C, which typically resulted in a pattern of about 90% mononucleosomal DNA. After MNase treatment, cells were lysed in TNESK buffer (20mM Tris-HCl, pH 7.9, 200mM NaCl, 2mM EDTA, 2% SDS, 20µg/ml Proteinase K) and incubated at 50°C for five hours. DNA was phenol/chloroform purified, precipitated with ethanol and sodium acetate, and loaded quantitatively on an agarose gel to purify the mononucleosomal DNA. Gel purification was carried out using the NucleoSpin Extract II Kit (Macherey-Nagel). For the preparation of input DNA, DNA was extracted from whole cells using the QIAamp DNA Mini Kit. DNA was sheared 15 times for five minutes in a biorupter device (Diagenode), to obtain DNA fragment sizes similar to MNase digested DNA.

3.6.2 Labeling of DNA for microarray hybridization

DNA was chemically labeled with the Universal Linkage System (ULS) of Kreatech diagnostic. The ULS molecule consists of a monofunctional platinum complex, which is coupled to the detection dyes Cy3 or Cy5. The platinum atom forms a coordinative bond with the DNA by binding to nitrogen seven of guanine. In contrast to Klenow labeling, this method is independent of any enzymatic step and therefore suitable for labeling of small-sized DNA. Labeling of DNA was carried out in cooperation with Imagenes.

3.6.3 Microarray hybridization

A custom designed Nimblegen 385k array with overlapping oligonucleotides of a length of 50 nucleotides, a spacing of ten, and an offset of five nucleotides between upper and lower DNA strand was designed. Labeling and hybridization was done in cooperation with Imagenes according to their standard protocols.

3.6.4 Data analysis

Raw data from the custom Nimblegen microarrays (.PAIR format) were analyzed with the software "R" (http://www.r-project.org) and the software "Bioconductor" (http://www.bioconductor.org). All functions were called using default parameters if not indicated otherwise. Raw data signals of all replicates were "scale normalized" to compensate for potential biases introduced during the manufacturing process (Yang et al., 2002). The $\log_2$-ratios (enriched/input) of the normalized signals were determined and averaged for each replicate. The ZREs that had been identified in the EBV genome (Bergbauer et al., 2010) were reviewed for the following criteria for a subsequent bioinformatical analysis: Clusters of ZREs and meZREs or single BZLF1 binding sites had to be at least 1500bp apart from each other in the EBV genome. In regions that encompass two or more closely spaced ZREs in a cluster, the position with the strongest binding of BZLF1 was selected and included in the analysis. 32 ZREs and meZREs fulfilled the criteria and were selected for the analysis (see Tab. A5 in the appendix). Average nucleosome occupancies were determined separately for the three different conditions: Raji cells, Raji 3927 cells, and lytically induced Raji 3862 cells. A window of ±2000bp, centered at the start position of each ZRE, was chosen. The $\log_2$-ratios (enriched/input) were averaged in sliding sub-windows of 150bp, which is approximately the length of one nucleosome, and a step size of 10bp. To identify differences between the latent and lytic phase, subtractions of $\log_2$-ratios (enriched/input) of lytically induced Raji 3862 cells and parental Raji cells were calculated for each individual ZRE and visualized in a heat

map. A hierarchical cluster analysis using the Ward's minimum variance method was performed on the same heat map matrix with a window of ±350bp, centered at the start position of each ZRE to probe for functional groups among ZREs. Average nucleosome occupancy profiles of each group identified by the cluster analysis were presented separately. As a second step, subtraction of $\log_2$-ratios (enriched/input) and cluster analysis was performed on ZREs that comprised a difference between the latent and lytic states with the datasets Raji 3862 and Raji 3927 to determine the role of BZLF1's activation domain on nucleosome occupancies. 32 randomly chosen positions that lack ZREs were chosen and analyzed similar to the ZREs as a control.

3.7   Chromatin Immunoprecipitation (ChIP)

Chromatin Immunoprecipitation (ChIP) is a technique to study interactions between proteins and DNA. The protein/DNA complex can be isolated with specific antibodies. The output can be analyzed with quantitative real time PCR upon DNA purification. Usually the chromatin is crosslinked with formaldehyde prior to immunoselection to stabilize the protein-DNA interactions. Formaldehyde reversibly links primary aminogroups of proteins with nitrogen atoms of nearby proteins or DNA.

3.7.1   Chromatin preparation

4×10^7 cells were harvested by centrifugation at 1200rpm for seven minutes and washed twice with ice-cold PBS to remove any residual serum from full medium. Crosslinking was carried out in a total volume of 20ml PBS (room temperature) and 1% formaldehyde for seven minutes at room temperature with constant rotating. Addition of 2.5ml 1M glycine and incubation for five minutes at room temperature stopped the reaction. Cells were pelleted and washed with ice-cold PBS in three centrifugation steps at 500g and 4°C for five minutes. The cells were resuspended three times in 10ml of ice-cold lysis buffer (10mM Tris-HCl, pH 7.9, 10mM NaCl, 3mM MgCl$_2$, 0.5% NP-40) and centrifugation at 300g for 10min to isolate nuclei. Nuclei were resuspended in 2ml sonication buffer (10mM Tris-HCl, pH 7.9, 200mM NaCl, 3mM MgCl$_2$, 1mM CaCl$_2$, 4% NP-40, 3mM EGTA, 1% SDS, 1x "Complete" proteinase inhibitor cocktail). Chromatin was fragmented by sonication with a BRANSON sonifier W-250 D (amplitude = 35%, total pulse time = 180s with 36 cycles of 5s pulse and 25s pause). The sonicated samples were transferred into 1.5ml Eppendorf tubes and centrifuged at maximum speed at 4°C for ten minutes. Samples were analyzed for their

protein content with the Bradford method, subsequently frozen on dry ice, and stored at -80°C. The fraction size of the chromatin was analyzed after treatment of 10µg chromatin with Proteinase K at 65°C over night. After DNA purification with the NucleoSpin Extract kit II (Macherey-Nagel), samples were loaded on a 1% TBE agarose gel to visualize the DNA size.

3.7.2 Chromatin immunoprecipitation and purification of ChIP DNA

100µg of chromatin in a total volume of 100µl sonication buffer was adjusted to 1ml with IP dilution buffer (20mM Tris-HCl, pH 7.9, 2mM EDTA, pH 8.0, 1% TritonX-100, 150mM NaCl, 1x "Complete" proteinase inhibitor cocktail) to obtain a final SDS concentration of 0.1%. After addition of the antibody, the chromatin was incubated at 4°C over night with constant overhead rotating. 100µl of protein G or protein A resin (50% slurry) were pre-equilibrated with IP dilution buffer and incubated at 4°C over night with overhead shaking. The next day, beads were spun down at 2000rpm for two minutes, added to the antibody-chromatin mixture, and incubated at 4°C for two hours by overhead rotating. Beads were washed for five minutes each with (i) low salt buffer, (ii) high salt buffer, (iii) LiCl buffer, and (iv) TE buffer to remove unspecific proteins and DNA. DNA/protein complexes were eluted off the beads with 125µl elution buffer at 65°C and constant shaking for 15min. The supernatant was transferred to a new tube and the elution step was repeated. The two supernatants were combined, supplied with 40µg of Proteinase K and incubated at 65°C over night to reverse the formaldethyde crosslink. DNA was purified with the NucleoSpin Extract II Kit (Macherey-Nagel) following to the instructions of the manufacturer and eluted in 60µl elution buffer. ChIP samples were immediately analyzed or stored at -20°C.

3.7.3 Quantification of ChIP DNA by real time PCR (qPCR)

ChIP and input samples were analyzed by quantitative real time PCR with the LightCycler 480 system (see 3.3.5) to quantify the DNA content. Stepwise dilutions of total input DNA ranging from 0.01% to 10% and ChIP DNA were analyzed with primer pairs specific for different loci of latent, early lytic, and late lytic EBV regions. Oligonucleotide sequences are listed in the appendix.

3.8   Indirect endlabeling

Indirect endlabeling is a method to map nucleosomes in a region of interest. Chromatin is treated with MNase; DNA is purified and cleaved with a restriction endonuclease with a restriction site close to the region of interest. Southern blotting and hybridization with a probe that binds in close proximity to the restriction start site visualize the digestion pattern of the region of interest. Each band in the resulting autoradiogram represents the boundary of a nucleosome or a hypersensitive site.

MNase digestion was performed similarly to the sample preparation for microarray hybridization but with minor changes. The digestion of chromatin with MNase has to be mild to achieve a partial digestion, only. Buffer volumes were increased to 1ml to ensure proper isolation of DNA after MNase treatment and MNase digestion was performed with 0.1-1.5 U MNase in each sample, only. Chromatin was Proteinase K treated and purified by phenol/chloroform extraction and ethanol precipitation. Cleavage of MNase treated DNA was carried out using 30µg of DNA and 100U of the specific restriction endonuclease at 37° C for two hours. Samples were purified and loaded on a 1% TAE agarose gel. Fragments were separated at 50V over night. Transfer of the fragments to a membrane and DNA-DNA hybridization was done as described in 3.3.8-10.

3.9   Western blot immunodetection of RNA Pol II

2.5×10^6 cells were harvested by centrifugation. Cell pellets were resuspended in 2x Laemmli buffer (200µl each) and immediately heat denatured for five minutes at 100°C. 20µl of each sample were size separated in a 7% SDS polyacrylamid gel electrophoresis for 2.5h and 30mA. Transfer to a nitrocellulose membrane was performed and 450mA for 90min. The membrane was washed with water and incubated at room temperature in blocking solution for 60min. The membrane was incubated at 4°C over night with a 1:10 dilution of an antibody specific for the amino-terminal part of RNA polymerase II (Pol 3-3) in 5ml blocking solution. The membrane was washed three times in TBST for five minutes. Secondary antibody binding was performed at room temperature for two hours. After several washing steps (three times with water, three times with TBST, three times with water, and five minutes for each washing step), antibodies were visualized by incubation with ECL solution.

4. RESULTS I

EPIGENETIC REGULATION IN THE EARLY PHASE OF AN EBV INFECTION

Upon infection of primary B cells, EBV establishes a latent infection as a rule. Eventually, the virus switches to the lytic phase of infection, and progeny virus is released, ready for the horizontal spread to infect other cells. BZLF1 is an EBV-encoded transcription factor responsible and sufficient for the switch to the lytic phase. *BZLF1* is expressed in EBV infected cells very early after infection, but virus release cannot be detected before day twelve *post* infection (pi), with a gradual increase of released progeny virus until day 21 pi (Kalla et al., 2010). This contradiction was not understood, but the discovery that DNA methylation at CpG sites is essential for the expression of crucial lytic genes was the key to solve it. Certain viral genes carry a class of BZLF1 responsive elements (ZREs), which are exclusively or preferentially bound by BZLF1, if the central CpG dinucleotide in their cognate sequence carries a methyl-group (Bergbauer et al., 2010; Kalla et al., 2010). These special cognate sequences are termed meZREs (for methylated BZLF1 responsive elements). Virion DNA is epigenetically unmodified or "naked" at the time of infection, but acquires modifications similar to human chromatin after the establishment of the episome in the nucleus of infected cells. Accordingly, expression of essential lytic genes that encompass meZREs in their promoters is not possible in the early phase of infection. But once the methylation pattern is instituted, these promoters are ready to become activated. The dynamics of CpG-methylation of EBV's DNA, which is the precondition to induce EBV's lytic phase, still had to be solved. It was part of my work to study the kinetics of DNA methylation in primary infected B cells,

and to compare it with the occupation of EBV's DNA with another epigenetic mark, i.e. the nucleosomes. I could show that DNA methylation is slow, starting not earlier than one week pi. A characteristic LCL methylation profile was evident three weeks pi. In contrast, my experiments demonstrated that incorporation of EBV's DNA into nucleosomes is a fast and gradual process.

4.1 Early epigenetic events can be monitored in infection experiments with primary B cells *in vitro*

It is nearly impossible to study epigenetic events on EBV's DNA *in vivo* with human cells, because the percentage of cells infected by EBV is very low. Genome-wide studies require scores of cells, which precludes this cell source for many experimental approaches. However, the infection of B cells can be simulated in cell culture. Adenoids of young children provide a source of uninfected, primary B lymphocytes. They can be isolated and separated from T cells with Ficoll-Hypaque density gradient centrifugation in the presence of sheep erythrocytes. Upon infection with EBV, cells start to proliferate and grow out to lymphoblastoid cell lines (LCLs), which can be kept in cell culture indefinitely and are a perfect source for material like chromatin, DNA, or RNA from stably and predominantly latently infected B cells (Fig. 4.1).
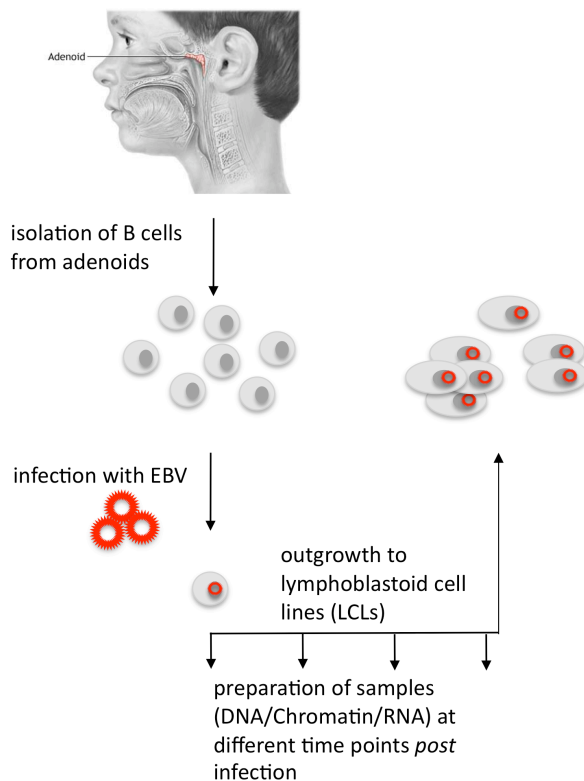
Fig. 4.1 Source of primary B cells and schematic view of infection experiments
One source of primary B cells are adenoids, which are surgically removed from the upper throat of young children. The B cells were infected to study epigenetic modifications at EBV's viral DNA in a time course infection experiment. EBV transforms or immortalizes the primary B cells to yield latently infected, proliferating lymphoblastoid cell lines (LCLs). Preparation of DNA, Chromatin or RNA was performed at several time points *post* infection.

4.2   EBV's DNA methylation is a slow, but precise process and can be followed *in vitro*

Tracking of DNA methylation *in vitro* requires a robust, effective and economic technique to identify methylated DNA in a genome-wide manner. Methylated DNA can be enriched in so-called "methylated DNA immunoprecipitation" (MeDIP) experiments with an antibody that specifically recognizes methylated CpG dinucleotides in single stranded DNA (Weber et al., 2005). Genomic DNA was isolated and sheared to an average length of 500bp by sonication. After heat denaturation, the single stranded DNA molecules were incubated over night with a monoclonal 5-methylcytosine antibody (Diagenode). The DNA-immuno-complexes were precipitated with sheep anti-mouse IgG1 magnetic beads. Proteinase K digestion followed by DNA purification resulted in a sample ("MeDIP") that was enriched in methylated DNA compared to the "input" DNA (Fig. 4.2 A). Input DNA is total genomic DNA from the very same DNA sample, but which was not immunoselected and immunoprecipitated. It is representative for the DNA content of the analyzed cells and enables comparison of samples with different genomic loads. MeDIP DNA can be further analyzed locus specifically with quantitative real-time PCR (qPCR) or EBV-genome-wide with microarray technology (MeDIP-on-Chip).

4.2.1   MeDIP experiments in the cell line Raji

The human Burkitt-lymphoma cell line Raji was used to implement MeDIP with EBV-positive cells. The Raji genome is highly methylated, and the methylation pattern of certain latent gene promoters is well characterized. For example, the *C* promoter is highly methylated in Raji cells, while the *Q* promoter, the *LMP1* promoter, and the *EBER* locus completely lack methylation (Minarovits et al., 1992; Salamon et al., 2001; Fernandez et al., 2009). With this information at hand, it was possible to assess the results of this approach.

Locus specific qPCR analysis revealed an overall high enrichment of MeDIP DNA compared to input DNA at the tested promoters. Only the *EBER* locus was not enriched in the MeDIP fraction (Fig. 4.2 B). As pointed out above, the *C* promoter is highly methylated in Raji cells, while the *EBER* locus is free of methylation, which is in line with the MeDIP results. However, the *LMP1* promoter and the *Q* promoter were highly enriched in the MeDIP sample, which is controversial to published data. These observations uncover the restriction of MeDIP experiments: The resolution is limited, because the fragments have an average length of 500bp. The resolution might be too low for specific and shorter promoter sections. The

hypomethylated region in the *EBER* locus exceeds 1000bp and can be detected with MeDIP. In contrast, the *Q* promoter is located in close proximity to the very densely methylated *F* promoter. Accordingly, it cannot be identified by MeDIP and a false-positive result is obtained.
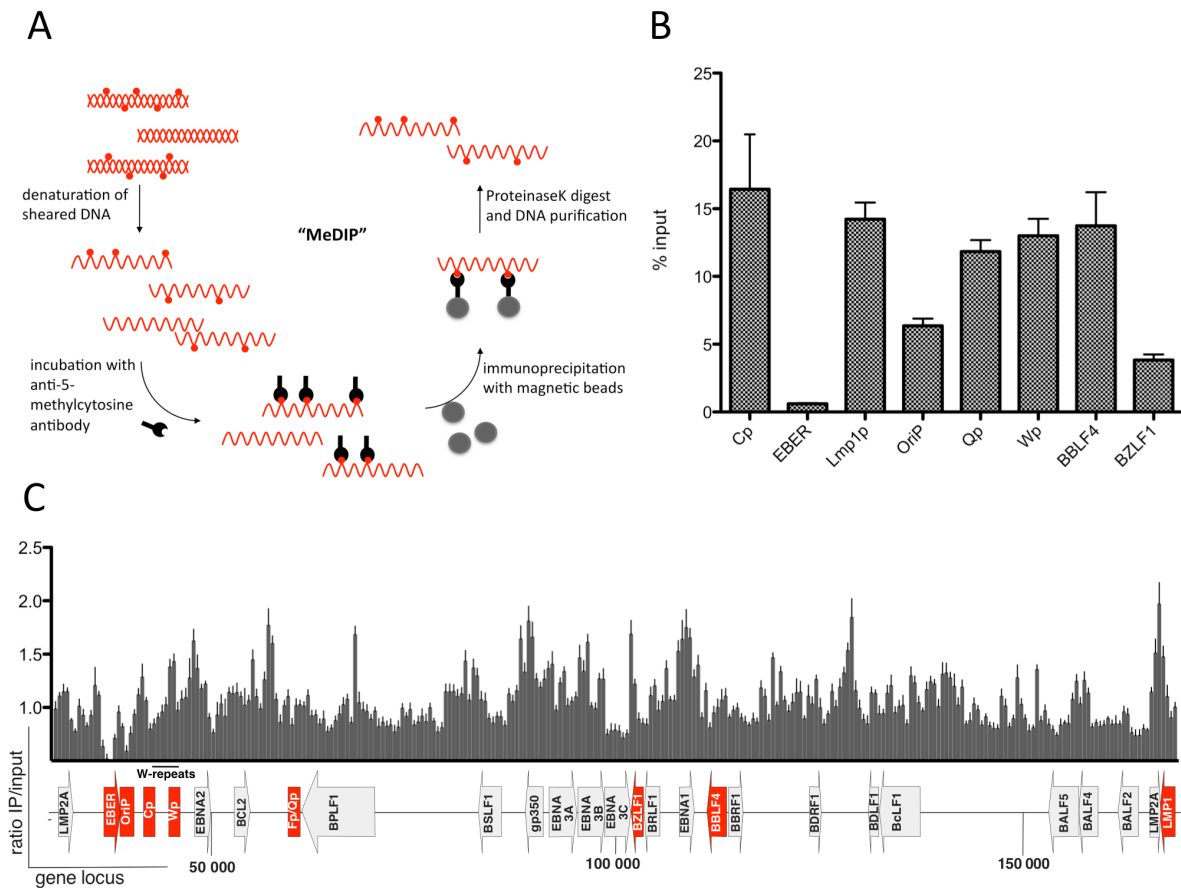
Fig. 4.2 MeDIP (methylated DNA immunoprecipitation) experiments in Raji cells
(A) For MeDIP experiments, Raji DNA was isolated, sheared to an average length of 500bp by sonication and heat denatured. To enrich for the methylated fraction, DNA was incubated over night with an antibody, which specifically recognizes methylated CpG dinucleotides. After immunoprecipitation of antibody/DNA complexes with anti-IgG1 magnetic beads, samples were Proteinase K digested and DNA was purified for further analysis by qPCR (B) or microarray hybridization (C).
(B) Analysis of CpG-methylation of selected EBV promoter regions by qPCR. MeDIP-enriched DNA of three independent experiments was quantified for precipitated DNA compared to input DNA in a Roche LightCycler 480 real-time PCR system at eight selected promoter regions. While the *C* promoter, the *LMP1* promoter, the *Q* promoter, the *W* promoter and the *BBLF4* promoter show a high degree of CpG-methylation (exceeding 10% of input DNA), the *BZLF1* promoter and *OriP* show a lower degree of methylation (about 5% of input DNA). The *EBER* locus seems to be spared from methylation (below 1% of input DNA).
(C) Genome-wide microarray analysis of the methylation profile of EBV DNA in Raji cells. MeDIP DNA and input DNA of three independent experiments were hybridized to a custom-made EBV-tiling array. The array consists of four identical sets of 285 PCR fragments with an average length of 500bp, covering the entire genome of B95.8 strain of EBV. Data was normalized such that the median of signal ratios was set to one. The ratio of MeDIP enriched DNA (IP) *versus* input DNA from the same cell sample is depicted on the y-axis, the genome location is shown on the x-axis. Selected features of EBV's genome are illustrated below the x-axis and locations that were also analyzed in qPCR (B) are highlighted in red. The methylation profile indicates an overall high degree of CpG-methylation in Raji cell DNA. An exception is the *EBER* locus, which is free of methylation in a region of >1kb (two spots on the microarray/two bars in the diagram), which is in line with the qPCR results in (B).

I switched to MeDIP-on-Chip to obtain a genome-wide methylation profile of the EBV genome in Raji cells. The EBV-genome-wide MeDIP-on-Chip analysis was performed with a custom-made EBV microarray. The microarray consists of 285 probes of PCR fragments with an average length of 500bp, covering the whole EBV genome (strain B95.8) in tiling, i.e. partially overlapping fragments (Kalla, 2007). MeDIP DNA and input DNA were amplified with the whole genome amplification kit (Invitrogen), labeled with Alexa5 and Alexa3 fluorochromes, respectively, and hybridized to the EBV microarray. The fluorescence signal of hybridized DNA was measured in a GenePix Personal 4100A scanner. The ratio of MeDIP/input signal reflects the methylation state of EBV's DNA.

Fig. 4.2 C shows the results of the microarray analysis. The normalized ratio of MeDIP/input is depicted on the y-axis and the x-axis indicates the EBV genome location. The experiment revealed differentially methylated regions. The *EBER* locus was completely spared from methylation, and the *C* promoter was enriched in methylated DNA. Unfortunately, the resolution of the analysis was again too low to detect the unmethylated region in the *Q* promoter or the *LMP1* promoter.

Summing up the experiment, (i) the establishment of MeDIP experiments in EBV-positive cell lines was possible, (ii) differently methylated DNA regions could be detected by qPCR and MeDIP-on-Chip, i.e. the *EBER* locus and the *C* promoter, (iii) the experiment is robust and reproducible, as experiments were done in triplicates, but (iv) the resolution of MeDIP was limited. This experimental setting is unfeasible to study discrete methylation patterns of small loci (<500bp). But a high resolution is not necessary to analyze the temporal changes of DNA methylation in primary infected B cells. For this, the overall methylation profile is needed, but not the methylation pattern at single CpG sites. Thus, MeDIP-on-Chip experiments fulfill all criteria listed at the beginning of this chapter: It is a reliable, robust, and economic technique for the analysis of the EBV-genome-wide change of DNA methylation over time.

4.2.2 Kinetics of DNA methylation in primary infected B cells

B cells were isolated from adenoids as depicted in Fig. 4.1 and infected with B95.8 wildtype virus at an MOI of 1.0 to analyze EBV's methylation profile in the early phase of infection. MeDIP-on-Chip experiments were conducted with infected cells one, two, three, nine, and twelve weeks pi. The cells had different EBV genome loads, especially in the early phase of infection. Thus, the ratio of a constant point of reference was always set to one to normalize and compare the different datasets.
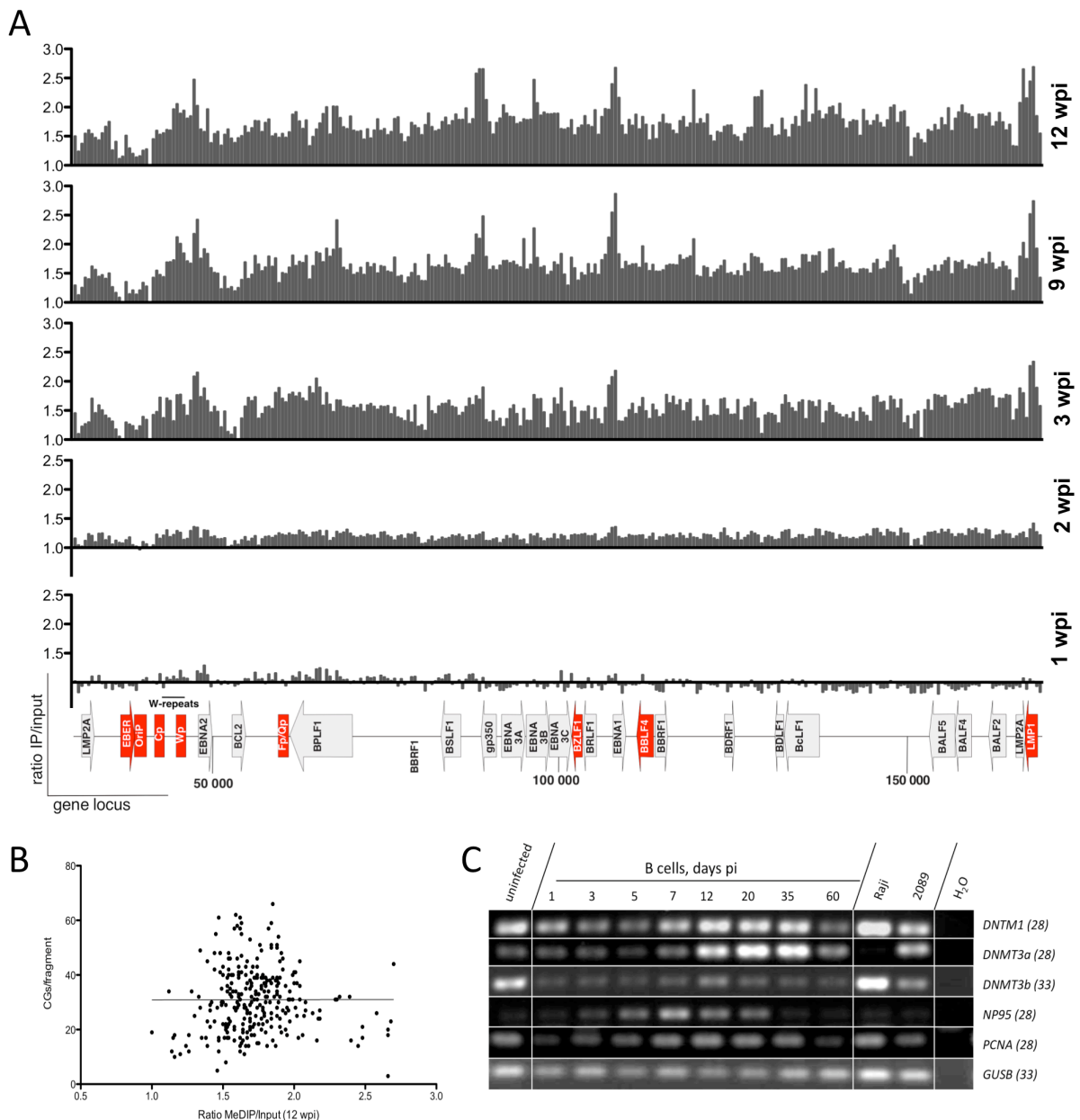
Fig. 4.3 Kinetics of EBV methylation in primary infected B cells
(A) MeDIP on Chip analysis of primary infected B cells at several weeks *post* infection (wpi). DNA from isolated B cells infected with B95.8 virus at an MOI of 1.0 was prepared after different time points pi. MeDIP analysis was carried out as described in Fig. 4.2, and samples were hybridized to our custom-made EBV tiling microarray. To analyze the kinetics of methylation pattern formation in samples with different genome loads, the

adjusted data were processed so that the ratio of a constant point of reference, *Cp*, was set to one in each experiment, because the *C* promoter region is free of methylated CpG dinucleotides even in established LCLs (Tierney et al., 2000), in contrast to EBV genomes in Raji cells. CpG dinucleotides in EBV DNA were entirely or mostly unmethylated early after infection but became increasingly methylated in resident EBV DNA of latently infected LCLs over time. The results of one of two biological repeats are shown.

(B) The x-axis of the diagram shows the ratio of MeDIP-enriched EBV DNA *versus* input DNA of cells twelve weeks pi as in (A). The y-axis shows the number of CpG dinucleotides that occur in each of the 285 tiling PCR fragments of the EBV microarray. The regression curve was flat, its coefficient of determination was extremely low ($R^2 = 5 \times 10^{-6}$), and the correlation coefficient was <0.003. The results indicated that there is no correlation between the degree of CpG-methylation in latently infected B cells and the occurrence of CpGs in EBV DNA.

(C) RNA isolated from primary B cells infected with B95.8 virus at an MOI of 1.0 was transcribed into cDNA and analyzed by semiquantitative PCR. The expression of DNA methyltransferases (DNMTs) and other genes involved in DNA methylation was assessed, as indicated on the right of the gel image with the number of PCR cycles in brackets. *DNMT1* and *DNMT3a* show an increase in their expression level one week pi, pointing to a role in EBV's DNA methylation. *DNMT3b* likewise revealed a slight increase, at much lower levels than the two other *DNMTs*. Expression levels of *NP95*, which is a protein involved in maintenance methylation by *DNMT1*, followed the same pattern seen for *DNMT1* and *DNMT3a*, while *PCNA*, which is a proliferation marker, showed only a minor increase in expression one week pi. Expression of the housekeeping gene β-glucoronidase (*GUSB*) was stable in all samples. The results of one representative experiments out of three biological replicates are shown.

In contrast to Raji cells, the *C* promoter is unmethylated in LCLs (Tierney et al., 2000) and the ratio of MeDIP/input at the C promoter was the lowest value in all experiments. Therefore, this locus served as the reference point. Fig. 4.3 A shows the representative MeDIP-on-Chip result of one out of two biological replicas with two technical repeats each. The genome-wide methylation profile of B cells one week pi did not display a significant enrichment of any locus. This finding indicated that the DNA was unmethylated or only slightly methylated. Two weeks pi, an enrichment of methylated DNA was visible. The most substantial increase in methylation happened between week two and three pi. A manifest DNA methylation profile was visible three weeks pi comprising highly methylated regions and DNA segments that were spared from methylation like the *C* promoter and the *EBER* locus, respectively. This profile did not further change, as the MeDIP-on-Chip profile seen nine or twelve weeks pi did not exhibit any significant difference to the profile at week three. This experiment showed that CpG-methylation of EBV DNA is a slow, but specific process that culminates in a distinct methylation profile after several weeks pi.

The amount of MeDIP DNA is not only a function of the methylation state of the fragment, but also of CpG density in the respective DNA sequence. The ratios of MeDIP/input of each microarray probe at twelve weeks pi was plotted against the number of CpG dinucleotides in the particular probes to test whether the methylation profile seen in LCLs was only dependent on the CpG density. The regression curve was flat, with a very low coefficient of determination (5×10^{-6}), and the correlation coefficient (ranging from one for high correlation to minus one for anti-correlation) was close to zero (Fig. 4.3 B). The analysis indicated no sequence bias in MeDIP enrichment. Other mechanisms but CpG density seem to be

important to determine the methylation profile of EBV's DNA in LCLs.

Semiquantitative real-time (RT)-PCR experiments were performed to test, whether *de novo* methylation of EBV's DNA might stem from the upregulation of the expression of DNA methyltransferases (DNMTs) in freshly infected B cells (Fig. 4.3 C). RNA of freshly infected B cells was prepared at different time points pi and transcribed into cDNA.

The maintenance methyltransferase *DNMT1* was expressed already at day one pi, but expression increased further, starting seven days pi. *DNMT3a* and *DNMT3b* encode for the two *de novo* methyltransferases. Both genes were weakly expressed at one day pi. The expression level started to rise seven days pi, similar to *DNMT1* expression. Previous reports demonstrated that EBV's latent membrane protein 1 (LMP1) induces the expression of all methyltransferases in nasopharyngal carcinoma cell lines (Tsai et al., 2002) and that *DNMT1* expression is upregulated by LMP1 via the JNK pathway (Tsai et al., 2006). These reports suggest that LMP1 is responsible for the upregulation of *DNMT* expression in my experiments, as it is expressed early after infection. The nuclear protein of 95 kDa (NP95) is a protein that is associated with maintenance and *de novo* DNA methylation. It recruits DNMT1 to hemimethylated CpG sites in the DNA, but also interacts with *de novo* DNMTs, histone methyltransferases, and trimethylated H3K9 and connects the DNA methylation pathway to the establishment of repressive histone marks (Meilinger et al., 2009; Rottach et al., 2009). *NP95* was upregulated in EBV infected, primary B cells seven days pi, indicating that this protein is involved in the establishment of epigenetic modifications in EBV as well. The proliferative cell nuclear antigen PCNA is a marker for proliferating cells, but is also involved in maintenance DNA methylation through interactions with DNMT1 at the replication fork. The slight upregulation in primary infected B cells was most likely a consequence of the growth transformation of these cells caused by the EBV infection. The expression of the housekeeping gene *β-glucoronidase* (*GUSB*) was stable in primary infected B cells over time.

The results of the expression analysis suggested that cellular methyltransferases introduce CpG-methylation at EBV DNA. Upregulation of important genes involved in DNA methylation paralleled the increase in DNA methylation at EBV's DNA. The slow increase in DNA methylation of EBV's DNA could probably be the reason for the lack of virion synthesis until day twelve pi (Kalla et al., 2010).

4.3 EBV's DNA is governed by nucleosomes very early after infection

Nucleosomes are the fundamental repeating unit of eukaryotic chromatin. They organize the packaging of the huge genome inside the nucleus, while still ensuring proper access to the DNA. Nucleosomes consist of eight core histones, which are wrapped 1.67 fold by approximately 147bp of DNA. This core structure is separated from the next nucleosome by a linker region with a length of up to 80bp. A fraction of DNA is part of regularly arrayed nucleosomes that resemble "beads-on-a-string" (Fig. 4.4 A, middle). Some parts of the genome are very densely packed (Fig. 4.4 A, top) or contain stretches that lack nucleosomes in a DNA segment, which exceeds the size of a normal linker region (Fig. 4.4 A, bottom). The mode of nucleosomal organization constitutes a mechanism for gene regulation by adjusting the accessibility of DNA for regulatory proteins.
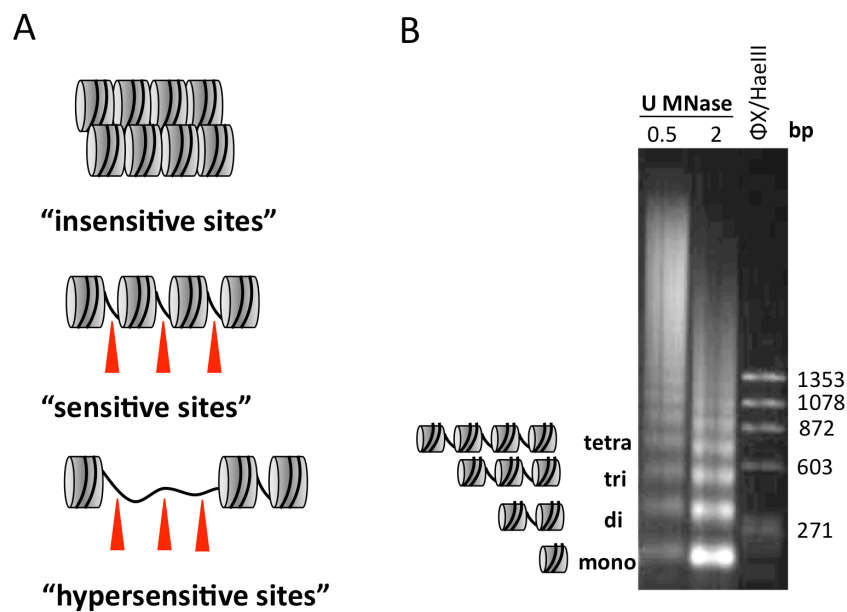
Fig. 4.4 Digestion of chromatin with micrococcal nuclease (MNase)
(A) Chromatin can be classified into different categories according to the accessibility to MNase digestion. Insensitive sites are tightly packed into nucleosomes, covering all contact surfaces for the enzyme and precluding it from DNA digestion. This chromatin moiety is usually found in heterochromatic regions. Sensitive sites have regular arrays of nucleosomes. The internucleosomal spaces are cleaved by MNase, which can reach the linker regions between regularly spaced nucleosomes easily. Hypersensitive sites contain long stretches of DNA, which are not packaged into nucleosomes and are therefore the preferred targets of MNase. Cutting sites of MNase are indicated with red arrowheads.
(B) Partial digestion of chromatin with different amounts of MNase leads to a ladder of nucleosomal DNA in an agarose gel. Chromatin was prepared from Raji cells and treated with either 0.5U or 2U of MNase at 37°C for five minutes. After Proteinase K digestion and DNA purification, DNA was loaded on a 1% TAE agarose gel with ethidium bromide for visualization of the digestion pattern. Mononucleosomal DNA has a length of 147bp, which is one nucleosome encompassing eight core histone proteins. Chromatin, which is partially digested with MNase, will show a ladder of mononucleosomal DNA fragments and multimers thereof as indicated in this image.

4.3.1 Micrococcal nuclease is a tool to study nucleosomal DNA

The accessibility of DNA can be used to study nucleosomal structures in the DNA on isolated chromatin *in vitro*. The enzyme micrococcal nuclease (MNase) is a relatively unspecific endo-exonuclease. Its cleavage efficiency depends on its accessibility to DNA, and therefore on the nucleosomal organization of the DNA. In regular nucleosomal arrays, MNase targets the histone-free linker region. DNA is nearly inaccessible to MNase in tightly packed nucleosomes. The preferred cleavage sites for MNase digestion are so-called hypersensitive sites, i.e. stretches of DNA, which lack histone proteins. In an agarose gel, partial digestions can be visualized as "MNase ladder" with mononucleosomal DNA at a length of 147bp and multimers thereof (Fig. 4.4 B). Chromatin preparation and treatment with MNase is used in a variety of methods that aim to analyze nucleosome positioning and/or occupancy on DNA.

4.3.2 A Southern blot analysis detects EBV DNA in nucleosomes nine days pi

The kinetics of nucleosome occupancy of EBV DNA can be analyzed in Southern blot experiments. Chromatin of infected primary B cells at several time points pi was treated with limited amounts of MNase, the DNA was purified, loaded in equal amounts on a 1% TAE agarose gel, and the MNase ladder was visualized by ethidium bromide staining (Fig. 4.5 A, upper part). The DNA was transferred to a nylon membrane and hybridized to a radioactively labeled probe (135.16), recognizing a specific part of EBV's genome ranging form coordinates 7315 to 56081, but containing only two W repeats (Fig. 4.5 B). Only EBV DNA that is already part of nucleosomes is protected from degradation and will result in a typical MNase ladder in an autoradiogram (Fig. 4.5A, lower panel). Raji cells served as a positive control. In contrast, MNase treated chromatin from the EBV-negative cell line BJAB, uninfected primary B cells, and infected primary B cells one and three days pi did not give rise to discrete signals. The first specific signal emerged at day nine pi, indicating that EBV's DNA was partially protected and adopted nucleosomal structures. The signals increased gradually, which could result from either an increase in the genome copy load per cell, like it is probably the case between day 16 and day 22 pi, or an increase of the fraction of EBV's DNA protected from digestion in the MNase treatment over time.
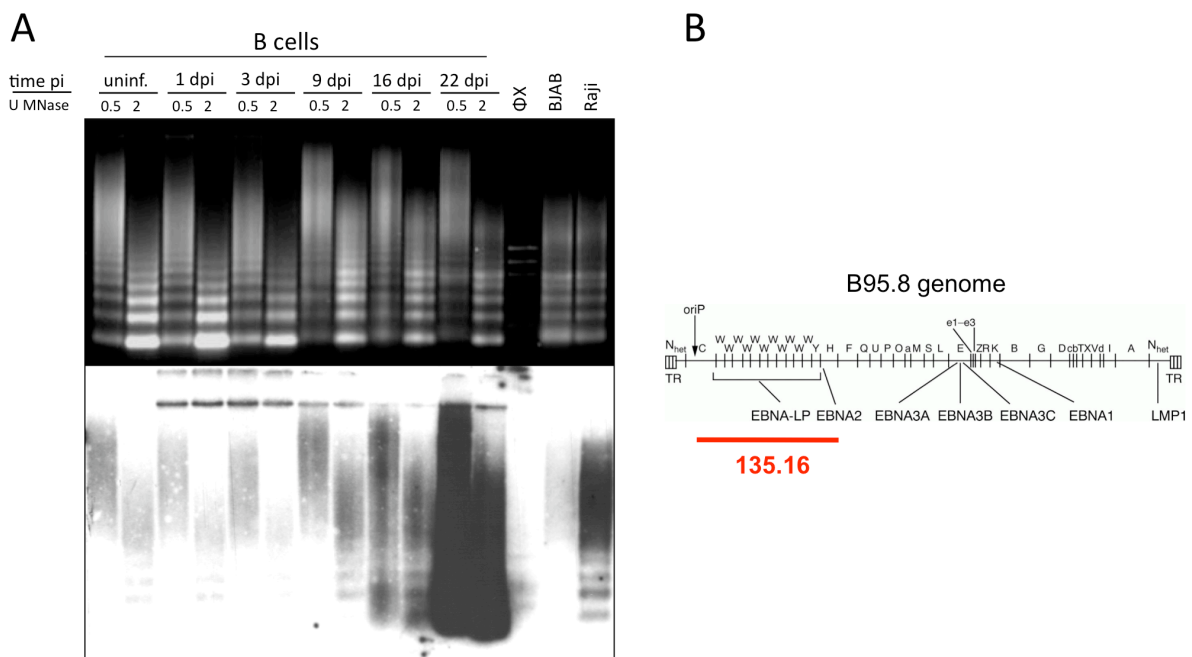
Fig. 4.5 Southern Blot hybridization estimates the temporal formation of nucleosomal occupancy of viral DNA

(A) Chromatin of uninfected and freshly infected B cells was prepared at various time points pi as indicated. Chromatin was treated with different amounts of MNase (0.5 U and 2 U per sample) to achieve a partial digestion. After Proteinase K treatment and DNA purification, equal amounts of DNA were separated on a 1% TAE agarose gel with ethidium bromide (upper part). Separated DNA was blotted to a nylon membrane and hybridized over night with a radioactively labeled probe (135.16). The probe will detect DNA only, which is already part of nucleosomal structures, because MNase will degrade nucleosome-free DNA. Raji cells served as a positive control compared to the EBV-negative cell line BJAB. As early as day nine pi, viral DNA appeared partially protected from MNase-mediated degradation as can be seen from the occurrence of a typical MNase ladder in the autoradiogram (lower part). The signal intensity of the MNase ladder increased over time and reached its maximum 22 days pi. The strong signal increase between day 16 and day 22 probably reflects a raise in the viral genome load per cell.
(B) The diagram provides a general overview of the EBV genome. The probe used in (A) was a recombinant plasmid DNA covering a part of EBV's genome ranging from nucleotide coordinates 7315 to 56081, with two BamW repeats, only.

4.3.3   *OriP* nucleosomes are detected as early as three days pi in a qPCR approach

The detection of DNA in Southern blot experiments has one limitation, namely sensitivity. To obtain a signal, DNA amounts of >10pg are necessary. This is particularly difficult in bulk cell cultures of primary B cells early after infection, because EBV will not infect every cell. Thus, nucleosomal positioning could take place earlier than its first detection in a Southern blot experiment. A more sensitive method is quantitative real time PCR (qPCR), because in theory, one molecule of DNA is enough to obtain a significant signal after PCR amplification.

The positions of two nucleosomes flanking the dyad symmetry element (DS) of *oriP* are known (Zhou et al., 2005). With this information, I designed primer pairs for the amplification of the respective nucleosomal DNA (NucI and NucII, details in Fig. 4.6). A

third primer pair, termed NucII/III, encompassed the internucleosomal space between nucleosome II and III and served as an internal negative control. Chromatin was treated with MNase, and the mononucleosomal DNA was purified and analyzed by qPCR. The negative control PCR product of NucII/III was always set to one to normalize for different genome loads in a time course experiment or in different cell lines.
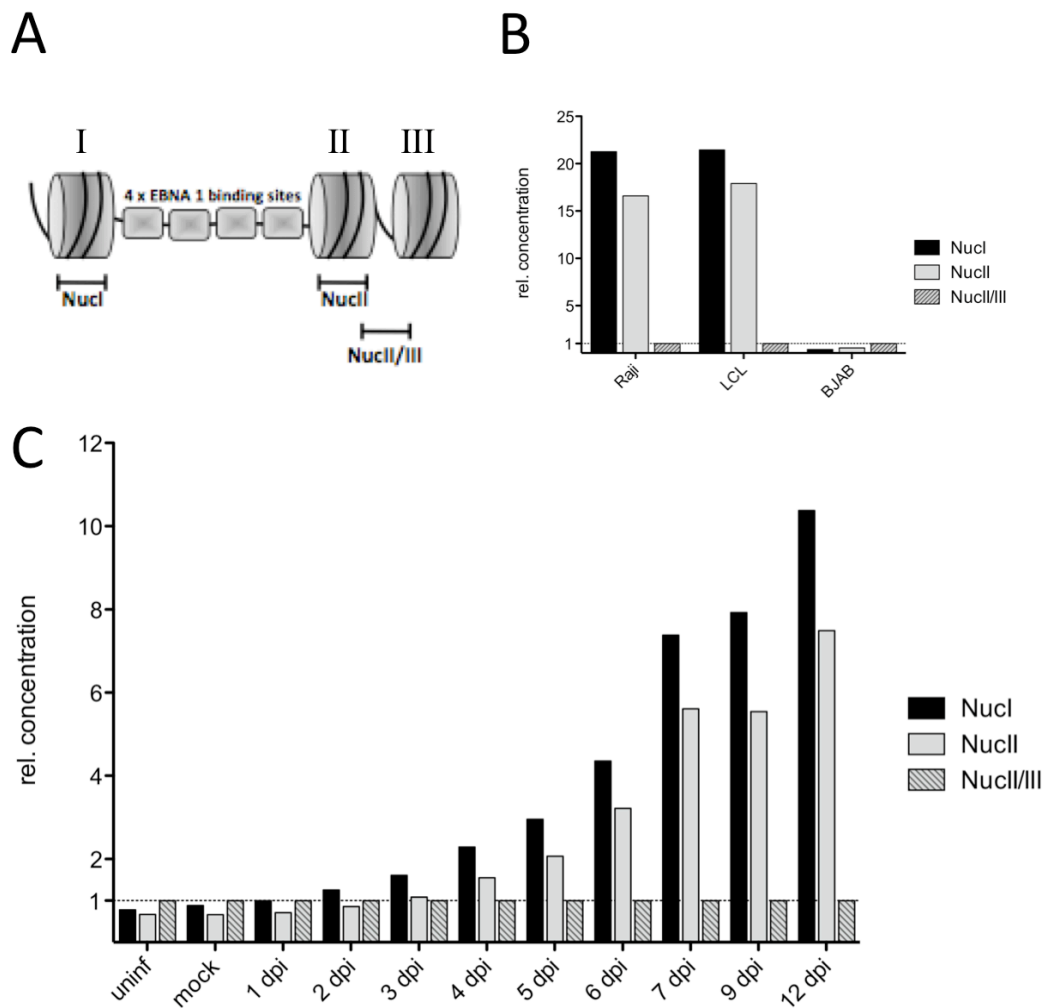
Fig. 4.6 EBV DNA packed in nucleosomes at *oriP* can be detected in a qPCR approach
(A) The exact positions of three nucleosomes around the dyad symmetry element (DS) of OriP is known (Zhou et al., 2005). With this information at hand, primers were designed to assess the nucleosomal positioning in the course of an EBV infection. Two primer pairs amplify the nucleosomal DNA (NucI and NucII, as indicated in A), whereas a third primer pair (termed NucII/III) spans the internucleosomal space between nucleosome II and III. The negative control PCR product of NucII/III was always set to one to calculate the relative enrichment of the PCR products of NucI and NucII.
(B) Chromatin of different cell lines was prepared and digested with MNase to determine the relative occupancy and positioning of nucleosomes I and II. After gel purification of mononucleosomal DNA, quantitative real time PCR analysis was carried out with the primer pairs described in (A). Mononucleosomal DNA from Raji cells and a long term LCL showed a clear enrichment of the products of NucI and NucII as compared to the product of NucII/III. The cell line BJAB served as a negative control.
(C) Primary B cells infected with EBV (strain B95.8) revealed an enrichment of the products of NucI and NucII as early as three days pi, which was increasing subsequently. Uninfected and mock-infected cells did not show enrichment of any PCR product. The result of one out of four biological replicates is shown. This analysis clearly indicated that the positioning of nucleosomes on EBV DNA flanking DS is a fast but gradual process.

Long-term LCL and Raji cells were analyzed for the positioning of the two DS-nucleosomes. Both cell lines were highly enriched in the PCR products for NucI and NucII at comparable levels (Fig. 4.6 B), indicating that the nucleosomes are likewise positioned. The EBV-negative cell line BJAB served as a negative control and did not show any enrichment of the PCR products. The first detectable enrichment of the NucI PCR product was visible in infected primary B cells on day three pi. The signal increased gradually, demonstrating a rapid and fast incorporation of EBV's DNA into nucleosomes (Fig. 4.6 C). The NucII PCR product showed a similar pattern, but was delayed by one day. Uninfected and mock-infected cells did not display any enrichment of either PCR product, as expected.

The qPCR approach identified EBV DNA to become part of nucleosomal structures very early after infection at *oriP*. In terms of timing, qPCR was more sensitive in comparison with the Southern blot experiment. But both experiments demonstrated a gradual increase of nucleosome occupancy of EBV DNA. As indicated above, this could reflect different genome loads in the cells, or a steady increase of EBV nucleosomes over time.

5. RESULTS II

EPIGENETIC MODIFICATIONS ON EBV'S DNA DEFINE LATENCY − AND A CHANGEOVER CAUSES THE LYTIC SWITCH

DNA methylation at CpG dinucleotides is important for the regulation of EBV's life cycle. The viral transcription factor BZLF1 has the unique characteristic to activate genes, even if their promoters contain BZLF1 responsive motifs with methylated CpG dinucleotides (Bhende et al., 2004; Bergbauer et al., 2010; Kalla et al., 2010). We previously identified numerous viral promoters and the cognate sequence motifs in a genome-wide ChIP sequencing approach. So-called meZREs[*] are preferentially located in promoters of early lytic genes. BZLF1's ability to bind and activate promoters, which are repressed by extensive DNA methylation, is a viral strategy to escape from latency and induce the lytic phase. DNA methylation is one epigenetic mechanism but the overall chromatin configuration at lytic promoters was unclear when I started my PhD work. It was unknown whether promoters, which encompass meZREs, are methylated *in vivo.* This aspect is fundamental because our hypothesis relies on this concept to mediate escape from viral latency. Similarly, the occupancy of EBV DNA with nucleosomes and its role during latency was unclear. Certain post-translational modifications of histone tails have been studied at latent gene promoters, but the contribution of the histone code to establish and maintain the repressed state of lytic promoters during latency has not been analyzed thoroughly. It was the aim of my work to

---

[*] methylated BZLF1 responsive elements

address and clarify these open issues. I wanted to reveal the epigenetic status of the lytic promoters during latency and identify the changes that accompany induction of EBV's lytic phase, which overrides repressive epigenetic events at lytic promoters.

5.1   EBV's DNA methylation at the nucleotide level

The *BRLF1* promoter and the *BBLF4* promoter are two prominent examples of promoters that contain meZREs. Both promoters were strongly bound by purified BZLF1 protein in a pull down assay with *in vitro* methylated EBV DNA (Fig. 5.1, red line, lower y-axis). Unmethylated DNA was not or only weakly bound (Fig. 5.1, black line, lower y-axis). The *in vivo* binding of BZLF1 protein to DNA in viral chromatin in Raji cells was assessed by Chromatin Immunoprecipitation (ChIP) experiments (Fig. 5.1, grey line, upper y-axis). The grey line is a nearly perfect mirror image of the red line indicating that the meZREs identified *in vitro* were important binding sites for BZLF1 *in vivo*. Nevertheless, the DNA methylation state of the meZREs *in vivo* was uncertain.
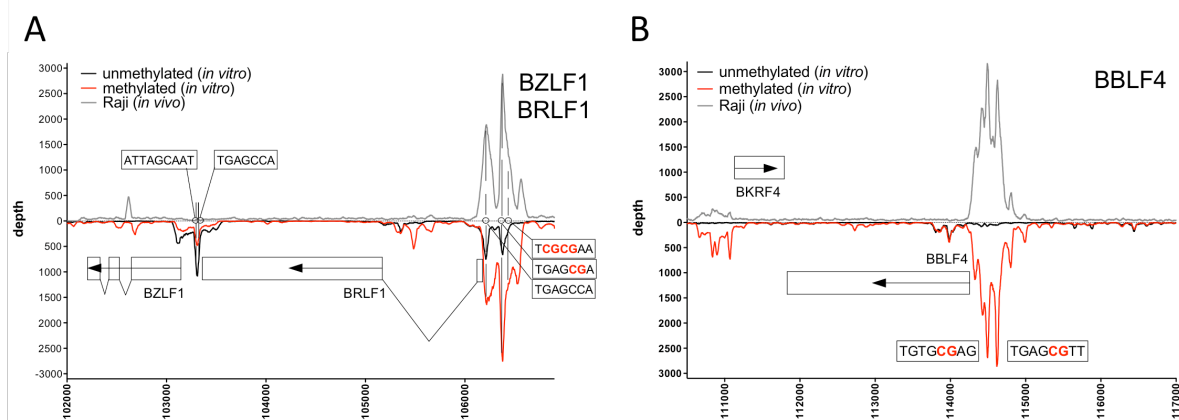
Fig. 5.1 BZLF1 can bind to its cognate sequences even if they carry methylated CpG dinucleotides
As described previously by us, a genome-wide screen identified BZLF1-responsive elements (ZREs) with a ChIP sequencing approach (Bergbauer et al., 2010). The grey line, mapped on the upper part of the y-axis, represents the result of a conventional Chromatin Immunoprecipitation (ChIP) experiment in Raji cells, while the red and the black lines (represented on the lower part of the y-axis) depict the results of *in vitro* pulldown assays, using either completely methylated (red) or unmethylated (black) EBV DNA as a substrate for binding to purified BZLF1 protein. Plotted is the sequencing depth, i.e. the number of sequence reads per base pair (y-axis), *versus* the genome coordinates of EBV strain B95.8 (x-axis).
(A) The two promoters of *BZLF1* and *BRLF1* encompass ZREs. The promoter of *BRLF1* contains ZREs, which are exclusively or preferentially bound by BZLF1 protein when CpG-methylated. These motifs are termed meZREs. The grey line (top part of the y-axis) depicts the situation in Raji as *in vivo*; the red line (bottom part of the y-axis) is almost a mirror image at the *BRLF1* promoter. The *BZLF1* promoter is different, because it does not appear to be accessible to BZLF1 protein *in vivo*.
(B) The *BBLF4* promoter contains several binding sites for BZLF1, which are exclusively bound if the DNA is methylated *in vitro*. The meZREs perfectly match the binding sites of *in vivo* ChIP experiments.

I analyzed selected regions of the EBV genome with deep bisulfite sequencing to assess their state of methylation including meZRE motifs of early lytic promoters. The methylation pattern of certain latent and late lytic promoters was also determined to understand the overall concept of viral gene regulation by DNA methylation. I could prove that all meZREs are highly methylated in latent Raji cells and that induction of the lytic phase does not cause a change in the methylation profile of EBV's DNA.

5.1.1 Deep bisulfite sequencing assesses EBV DNA methylation at the nucleotide level

The chemical treatment of DNA with bisulfite allows the discrimination between methylated and unmethylated cytosines. The addition of a sulfur trioxide group to the C5 position of the pyrimidine ring causes the hydrolytic desamination of the cytosine, i.e. the amino group is substituted by a ketone. As a consequence, cytosines are converted into uracils. Methylated cytosines are protected from the conversion, because the methyl-group is located at the C5 position. Fig. 5.2 provides an overview about the working steps I performed to obtain the DNA methylation pattern of certain viral genomic sequences on the nucleotide level.

Genomic DNA from Raji cells was bisulfite modified. 60 partly overlapping PCR products encompassing a total of 26 EBV regions formed contigs, which contained a selection of latent, early lytic, and late lytic gene promoters. PCR products had an average length of 461bp and covered 27869bp of EBV's genome. They were purified and pooled in equimolar amounts and sequenced on a Roche Genome Sequencer FLX system in cooperation with the Microsynth AG (http://www.microsynth.ch).

I used the CLC genomic workbench software for sequence alignment. Sequence reads were aligned to a modified version of the B95.8 wildtype sequence, in which I converted all cytosines in a non-CpG context to thymines. Cytosines of CpG dinucleotides were kept unconverted. The parameters of the sequence alignment were set to tolerate RY-mismatches in the context of CpG dinucleotides, allowing mismatches of purines and pyrimidines, which are expected if cytosines are unmethylated. A total number of ~88000 reads with an average length of 346bp was successfully mapped to the EBV genome. Aligned sequences were extracted and the percentage of methylated cytosines was assessed with the BiQ analyzer HT software/beta-test version (Bock et al., 2005).
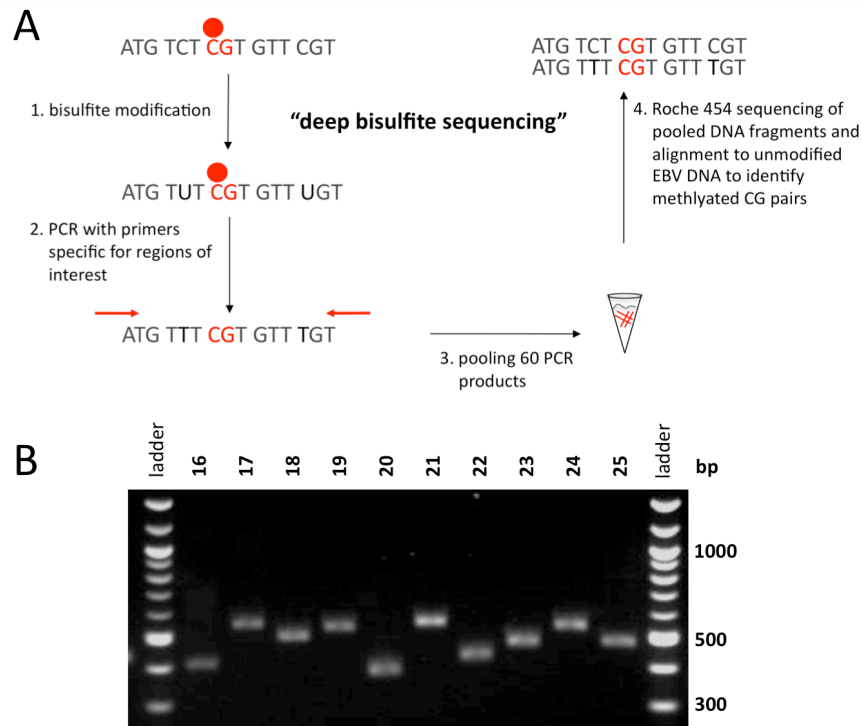
Fig. 5.2 High-resolution analysis of viral DNA methylation at the nucleotide level by deep bisulfite sequencing

(A) DNA of Raji cells was isolated and treated with bisulfite, which leads to the hydrolytic desamination of cytosines to uracils. After PCR amplification and purification of 60 regions of the viral genome, samples were pooled in equimolar amounts and sequenced with a Roche Genome Sequencer FLX system ("454 sequencing"). Sequences were aligned to the EBV genome to identify methylated CpG dinucleotides.

(B) Example of ten PCR products after purification and equimolar adjustment. PCR products had an average length of 461bp and covered 27869bp of the viral genome, representing 20% of EBV's unique sequences. The sequenced regions include latent, immediate early, early, and late promoters of EBV.

The deep sequencing results of four contigs are graphically represented in Fig. 5.3. Each bar represents a CpG dinucleotide in the EBV B95.8 sequence and is plotted against its genome location on the x-axis. The percentage of methylation of each CpG dinucleotide is depicted in red on the left y-axis. The sequence coverage is indicated as grey area and is pictured on the right y-axis in logarithmic scale. The coverage of each CpG dinucleotide was 840 reads/bp on average and >10 reads/bp were considered as representative. Selected features of the EBV genome are shown above the graphs. Deep bisulfite sequencing detected certain sequence variations between Raji DNA and B95.8 wildtype DNA. CpG dinucleotides missing in Raji DNA are illustrated as black bars with a star on top, additional CpG dinucleotides have a circle on top of the red or grey bar.

The *F/Q* promoter was completely unmethylated in Raji cells between genome coordinates 62200 and 62600, in sharp contrast to the highly methylated region upstream (Fig. 5.3 A). A previous study compared the methylation state of 20 CpG dinucleotides in a region between genome coordinates 62200 and 62500 in several cell lines including Raji cells (Tao et al.,

1998). The 20 CpG dinucleotides were hypomethylated in most cell lines, but a minor fraction showed variable methylation in the first eight CpG sites that encompass the *F* promoter. My deep sequencing results demonstrated that Raji cells completely lack CpG-methylation in this region. Two CTCF binding sites flank the eight CpG sites that are methylated to some degree in certain cell lines (Tempera et al., 2010) and presumably shield the *Q* promoter from DNA methylation and other repressive modification in its close vicinity.

The *BZLF1* promoter does not encompass meZREs and CpG sites are underrepresented in this promoter in comparison with the average CpG density in EBV (Fig. 5.3 B). The six CpG sites downstream of the *BZLF1*-TATA box were hypomethylated, although *BZLF1* is not expressed in latent Raji cells. This finding indicated that repressive mechanisms distinct from DNA methylation are responsible to repress the BZLF1 promoter. For example, ZEB-proteins[*] were shown to bind to the transcriptional initiation site of *BZLF1* and to suppress gene activation during latency (Kraus et al., 2003; Yu et al., 2011).

The CpG-rich *BBLF4* promoter was heavily methylated, reaching nearly 100% methylation at most CpG dinucleotides (Fig. 5.3 C). The high percentage of methylation especially in meZREs indicated a homogenous pattern of CpG-methylation of all genome copies of EBV in Raji cells. The high degree of CpG-methylation at the *BBLF4* promoter also indicated that all cells within the population of Raji cells have a repressed *BBLF4* gene during latency. Other early lytic promoters showed an equally high and uniform methylation of meZREs (Tab. A4, appendix). Methylated meZREs offer a platform for BZLF1 binding, once the protein is expressed. The consistently high degree of CpG-methylation of meZREs is presumably a mean to target the BZLF1 protein to most motifs and induce a high expression of early lytic genes upon the onset of the lytic phase. Other CpG sites in the *BBLF4* promoter, like the five CpGs between genome coordinate 114525 and 114575, were variably methylated. This suggested that the methylation state of the respective sites was less important for epigenetic regulation of EBV.

The *BDLF4/BDRF1* promoter showed a high degree of CpG-methylation (Fig. 5.3 D), indicating a tight repression of these late lytic genes during latency. Remarkably and in contrast to most early lytic genes, meZREs are absent in the *BDLF4/BDRF1* promoter, as are conventional ZREs (Bergbauer et al., 2010). It is clear from this observation that BZLF1 does not reverse the repression of this late lytic promoter. The deep bisulfite sequencing results of the remaining 22 contigs are presented as a table in the appendix (Tab A4).

---

[*] zink finger E-box binding proteins
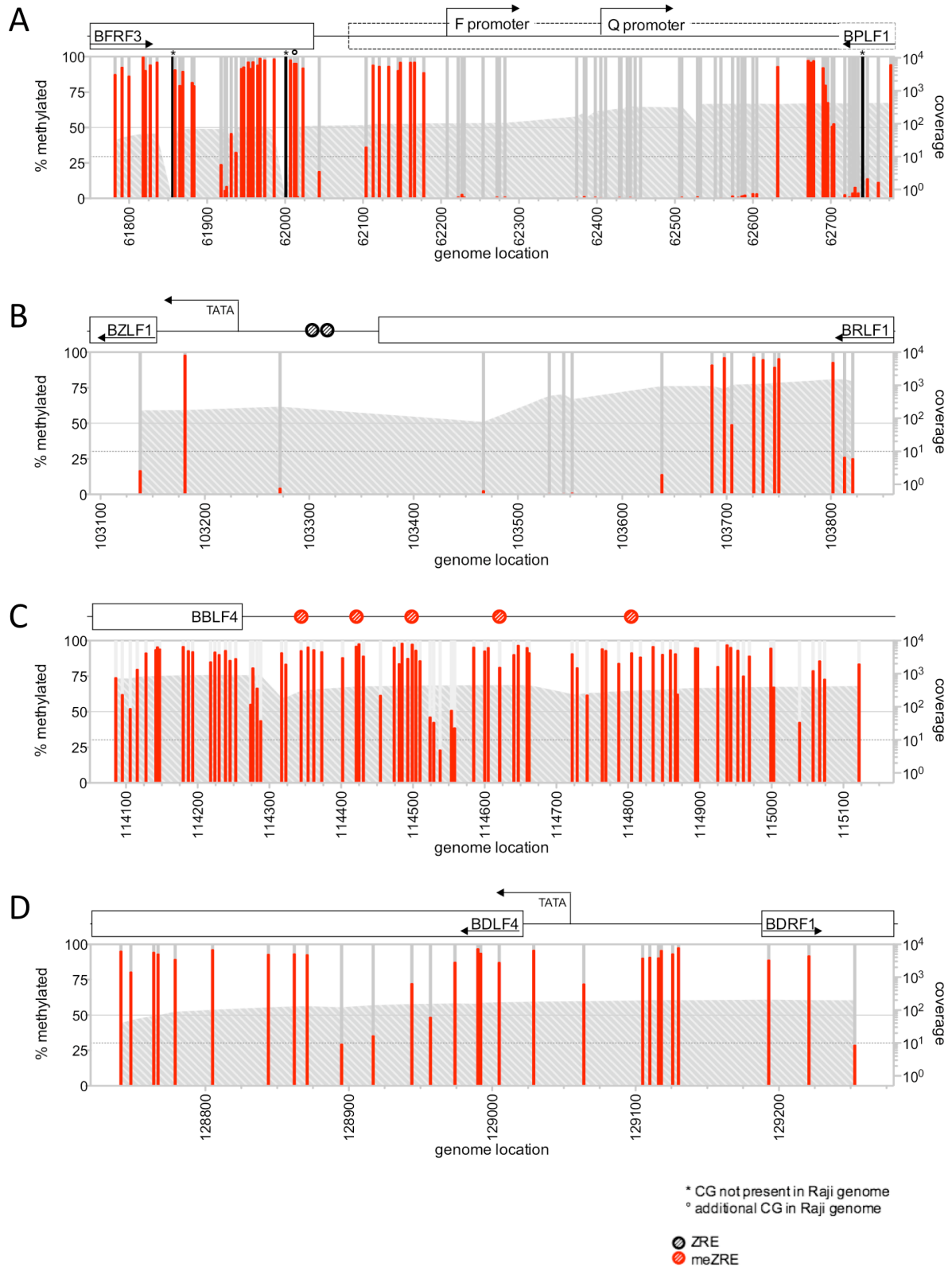
Fig. 5.3 Graphical representation of four selected regions of the viral genome in Raji cells with results from deep bisulfite sequencing

Each bar represents a CpG dinucleotide and is plotted *versus* its genome coordinate on the x-axis. The percentage of methylation is displayed on the left y-axis and is indicated in red. The right y-axis provides the sequencing coverage, which is depicted as a grey area in the diagrams in logarithmic scale. Coverage above ten reads/bp was considered to be representative for the analysis. Deep bisulfite sequencing detected single sequence variations between Raji DNA and B95.8 wildtype sequence. CpG dinucleotides missing in the Raji genome but present in B95.8 are indicated as black bars with a star on top, while additional CpG dinucleotides in Raji cells

are indicated with a circle on top of the red or grey bar. Selected annotation of the EBV genome can be found above the graph, with ZREs and meZREs indicated as black or red circles, respectively.

(A) Methylation pattern of the region with the latent *Q* promoter. The *Q* promoter was clearly spared from CpG-methylation, in sharp contrast to the upstream and downstream regions, which were heavily methylated. These findings indicate that the *Q* promoter is protected from methylation and is active in latency, which is in line with a recently identified CTCF binding site, situated exactly at the left border between methylated and unmethylated regions (Tempera et al., 2010).

(B) Methylation profile of the immediate early promoter of *BZLF1*. The *BZLF1* promoter contains only six CpG dinucleotides, which were not methylated in Raji cells, pointing to a repressive mechanism distinct from DNA methylation during latency.

(C) Methylation pattern at the early lytic promoter of *BBLF4*. The *BBLF4* promoter is highly enriched in CpG dinucleotides, most of them showed a very high percentage of methylation. All CpG dinucleotides, which are part of known meZREs and bound by *BZLF1* in a methylation dependent manner, were methylated to nearly 100%.

(D) The ZRE-free late lytic promoters *BDLF4* and *BDRF1* were methylated indicating their tight repression during latency.

5.1.2   Lytic induction does not change the DNA methylation of EBV promoters

I established a conditional system for the induced expression of *BZLF1* to analyze possible changes of epigenetic modifications on viral DNA in the lytic phase of EBV. The plasmid pRTS-2 comprises the constitutively active, bicistronic coding sequence of the Tet-repressor-KRAB fusion gene and the tetracycline controlled transactivator rtTA2s-M2 (Fig. 5.4 A). The repressor binds tightly to the bidirectional *tet* promoter and inhibits activation of the *tet*-controlled genes in the absence of tetracycline or its analogue, doxycycline. Upon addition of doxycycline, the activator replaces the repressor, and the *tet*-controlled genes are expressed (Bornkamm et al., 2005). The coding sequences of *BZLF1* and the green fluorescent protein (GFP) were placed under the control of the *tet* promoter to obtain an inducible *BZLF1* expression that can be monitored by GFP positive cells with, for example, flow cytometry (Fig. 5.4, A and B).
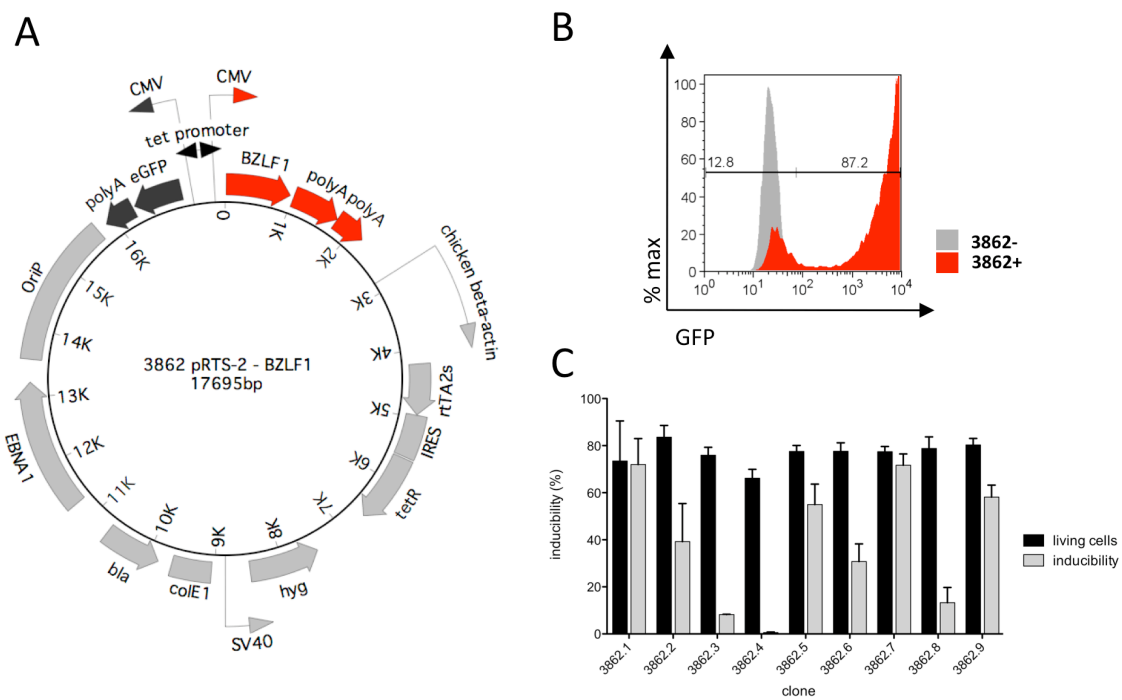
Fig. 5.4 Conditional expression of *BZLF1* in Raji cells
(A) The plasmid 3862 pRTS-2-BZLF1 contains the bicistronic coding sequence of the Tet-repressor-KRAB fusion gene and the tetracycline controlled transactivator rtTA2s-M2 under the control of the constitutively active chicken beta-actin promoter. The two genes are separated by an internal ribosomal entry site (IRES). In absence of tetracycline or its analogue, doxycycline, the repressor binds tightly to the so-called *tet* promoter, and inhibits its activation. After addition of doxycycline to the cells, the repressor is released from the promoter and replaced by the activator, which induces transcription. The promoter in the plasmid is bidirectional, allowing the simultaneous expression of the two transgenes *GFP* and *BZLF1* (Bornkamm et al., 2005).
(B) GFP expression in single Raji cells is a measure of induced transgene expression of *BZLF1*. The addition of 100 ng/ml doxycycline to the cells for 15 h led to 87% of GFP positive cells, as measured by flow cytometry.
(C) Nine single cell clones of Raji cells stably transfected with the plasmid 3862 pRTS-2-BZLF1 were analyzed for cell viability and inducibility of transgene expression by flow cytometry. Cells were cultivated and induced in three independent experiments to validate a robust and reproducible transgene expression. Clone 3862.1 was chosen for further analysis.

Raji cells were stably transfected with the plasmid 3862 pRTS-2-BZLF1. Nine single cell clones of Raji 3862 cells were analyzed for viability and inducibility of the transgenes using flow cytometry. The inducibility varied between the clones, but a robust and high expression of the GFP gene was obtained in Raji 3862 cells clone 1. This clone was chosen for further experiments (Fig. 5.4 C). An important characteristic of Raji cells and its derivates is the lack of the *BALF2* locus. The BALF2 protein is essential for the lytic DNA replication of the viral genome. As a consequence, downstream gene expression of all late viral genes is blocked and Raji 3862 cells can be induced to express early lytic genes, only, but the late lytic phase does not commence and remains incomplete. This aspect is important for the analysis of epigenetic modifications on EBV's DNA upon lytic induction, because this Raji cell model does not give rise to the synthesis of viral progeny DNA, which would be a major problem as it is free of CpG-methylation and other epigenetic modifications.

The methylation state of four selected regions was assessed by bisulfite sequencing after lytic induction. Raji 3862 cells were incubated with 100ng/ml doxycycline over night and sorted for GFP positive, i.e. lytically induced cells, with a FACS Aria II instrument. DNA of lytically induced and parental Raji cells was isolated, bisulfite treated and amplified with primers spanning the *F/Q* promoter, the *BZLF1* promoter, the *BBLF4* promoter, and the *BDLF4/ BDRF1* promoter. Resulting PCR fragments were purified and directly sequenced by the Sanger method and a commercial service provider (http://www.sequiserve.de). The ratio of the height of the cytosine peak to the height of the thymine peak in the sequencing chromatogram provides the conversion rate, i.e. the methylation state of the promoters.

Fig. 5.5 represents the results graphically in a way comparable to Fig. 5.3. The methylation state of parental Raji cells is depicted in red; the percentage of methylation in induced Raji 3862 cells is shown in green. There was no remarkable difference between the DNA methylation pattern of parental Raji cells and induced Raji 3862 cells in any locus. Slight variations were visible only in regions, which are variably methylated in parental Raji cells, like the five CpG dinucleotides around position 61925 in the *F/Q* promoter. The region downstream of the *F* promoter remained hypomethylated (Fig. 5.5 A), similar to the *BZLF1* promoter (Fig. 5.5B). The meZREs in the *BBLF4* promoter remained highly methylated after lytic induction (Fig. 5.5 C). Also the late lytic *BDLF4/BDRF1* promoter was heavily methylated in both cell lines. My findings indicated that lytic induction does not alter the methylation state of EBV's DNA.
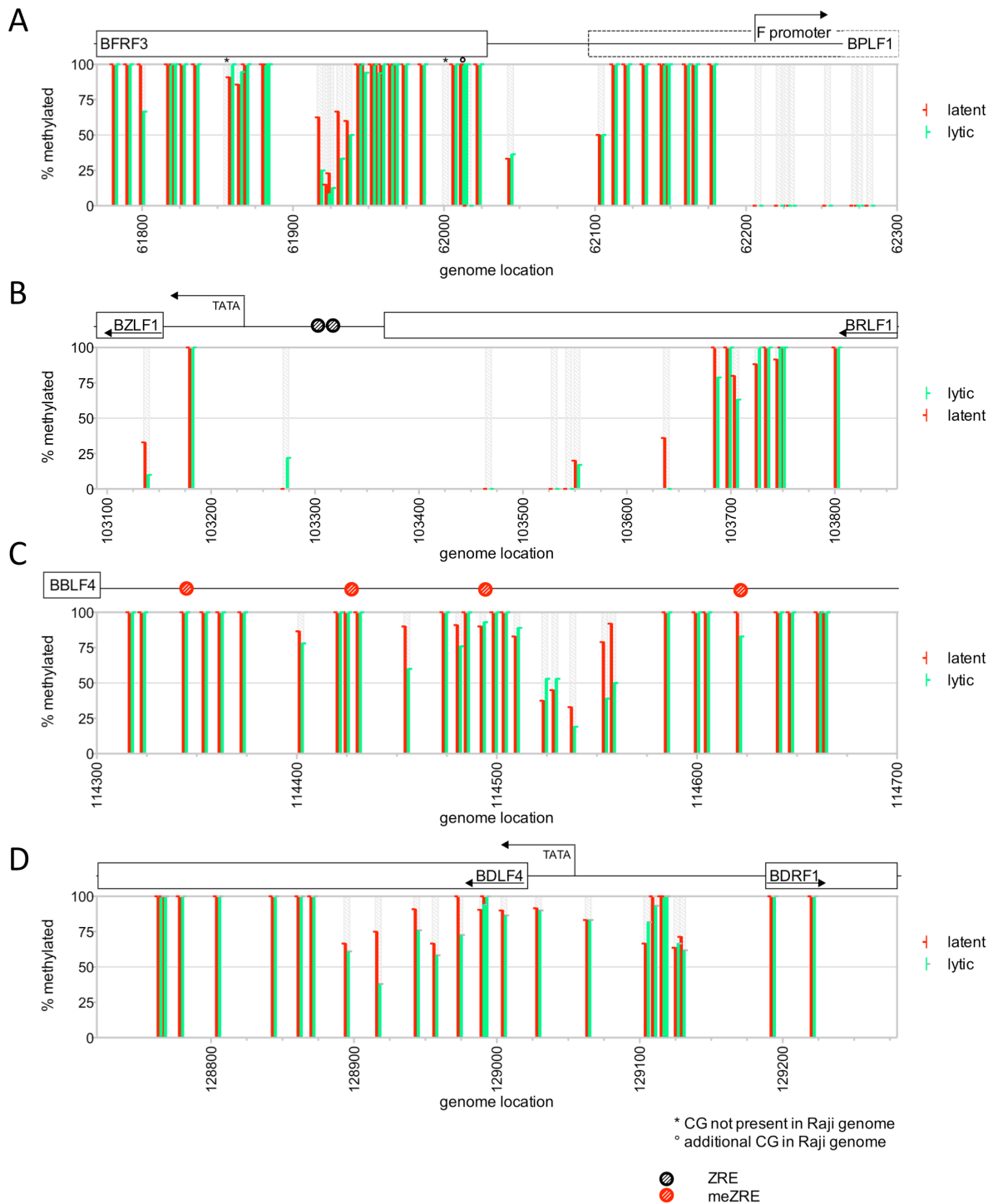
Fig. 5.5 The DNA methylation profile of Raji EBV DNA does not change upon induction of EBV's lytic phase

DNA of Raji cells or induced Raji 3862 cells, which had been sorted for expression of GFP, was isolated and treated with bisulfite. Bisulfite modified DNA was amplified with primers specific for **(A)** the latent *Q/F* promoter, **(B)** the immediate early *BZLF1* promoter, **(C)** the early lytic *BBLF4* promoter, and **(D)** the late lytic *BDLF4/BDRF1* promoter. PCR fragments were directly sequenced by the Sanger method and the resulting chromatograms were analyzed for the ratio of the height of the cytosine peak to the thymine peak, which reflects the rate of conversion after chemical bisulfite modification. Similar to Fig. 5.3, the percentage of methylation in Raji cells (latent) is depicted in red, while the percentage of methylation in lytically induced Raji 3862 cells (lytic) is indicated as green bars.

5.2   Nucleosome occupancies in EBV

Nucleosomes are barriers to transcription. The tightly packed complex of histone proteins and DNA constricts the accessibility of the DNA for proteins that are important for gene regulation and transcription (Bai et al., 2010). Most of the knowledge about the role of nucleosomes during transcriptional regulation originates from research in yeast. *S. cerevisiae* holds several inducible promoters, which serve as model systems for nucleosomal rearrangements upon transcriptional induction (i.e. the *PHO* promoter, Svaren et al., 1997; Adkins et al., 2004), but similar studies are rare in humans.

The nucleosome occupancy in EBV's DNA and its role during latency and lytic reactivation has not been studied so far. It was part of my work to analyze the occupation of EBV's DNA by nucleosomes with a focus on promoters that become activated upon lytic induction through the binding of BZLF1. I could show that certain BZLF1-regulated early lytic promoters were densely packed in nucleosomes during latency. Induction of the lytic phase caused major rearrangements in those promoters as compared to the rest of the genome. My studies identified different promoter subgroups, which have alternative modes of regulation by nucleosome occupancy and eviction, indicating different classes of genes and a hierarchy in the stringency of repression during latency.

5.2.1   Mononucleosomal DNA on Chip (MND-on-Chip) experiments assess the nucleosomal occupancy of viral DNA

MND-on-Chip experiments rely on the different accessibility of DNA to MNase cleavage in the context of chromatin (compare chapter 4.3.1 and Fig. 4.4). A digestion of chromatin with MNase leads to the degradation of free DNA, whereas DNA that is part of nucleosomes is protected and can be identified in a microarray hybridization.

Chromatin of different Raji cell derivates was treated with MNase. The DNA was purified and size-selected for mononucleosomal DNA on an agarose gel. Mononucleosomal DNA and sonicated input DNA with an average length of 150bp were labeled with the fluorochromes Cy5 and Cy3, respectively, and hybridized to a high-resolution EBV microarray (Fig. 5.6). The microarray contained the whole EBV B95.8 sequence in tiling oligonucleotides of 50 nucleotides in a step size of ten and an offset of five nucleotides between upper and lower DNA strand. In theory it provided a final resolution of five nucleotides.
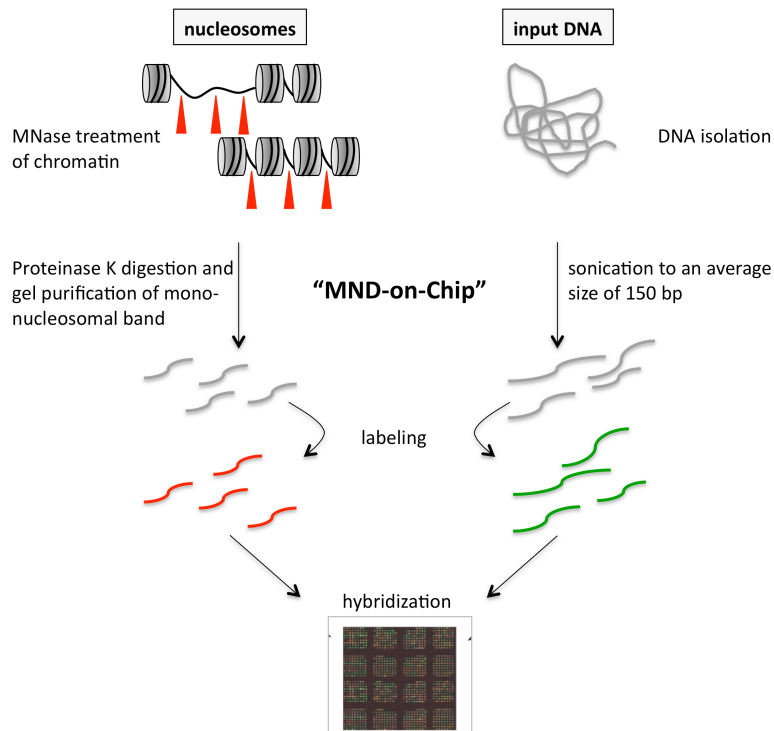
Fig. 5.6 Mononucleosomal DNA on Chip (MND-on-Chip) experiments analyze the nucleosome occupancy of DNA
Chromatin was treated with 50U/sample MNase and digested with Proteinase K. After DNA purification, mononucleosomal DNA was size selected on agarose gels and purified. Nucleosomal DNA was labeled with the fluorochrome Cy5 and hybridized together with sheared, Cy3-labeled DNA from the same cells ("input") to a custom Nimblegen microarray. The microarray encompasses the whole EBV B95.8 genome in tiling oligomers with a length of 50, a spacing of ten and an offset of five nucleotides between upper and lower DNA strand, providing a final resolution of five nucleotides.

5.2.2　Nucleosome occupancy in EBV's BZLF1-responsive promoters

MND-on-Chip experiments were conducted with three different sets of cell lines. Parental Raji cells were analyzed as a model of the latent state. Doxycycline-induced Raji 3862 cells express full-length *BZLF1* and *GFP* as transgenes upon induction (compare chapter 5.1.2 and Fig. 5.4). Chromatin of lytically induced and GFP-sorted cells was prepared to analyze the nucleosome occupancy in EBV's lytic phase. Raji 3927 cells constitutively express a truncated version of *BZLF1* that lacks its N-terminal activation domain. Chromatin of this cell line was analyzed to study the influence of BZLF1's DNA binding domain (DBD), only, on nucleosomal occupancies at BZLF1-regulated promoters. Experiments were performed with two independent preparations of mononucleosomal DNA each. The ZREs that had been identified in the EBV genome (Bergbauer et al., 2010) were reviewed for the following criteria for a subsequent bioinformatical analysis: Clusters of ZREs and meZREs or single BZLF1 binding sites had to be at least 1500bp apart from each other in the EBV genome. In regions that encompass two or more closely spaced ZREs in a cluster, the position with the strongest binding of BZLF1 was selected. 32 ZREs and meZREs fulfilled the criteria and

were selected for my analysis. For simplicity, they are further denoted as "ZREs", because the discrimination between methylation dependent and methylation independent ZREs is not important for this analysis. A list of the 32 ZREs can be found in the appendix (Tab. A5). The "enriched *versus* input" $\log_2$-ratios ($\log_2$-ratios) indicate the occupancy of DNA with nucleosomes.

The $\log_2$-ratios of the 32 ZREs were averaged and the average nucleosome occupancy profiles of the three Raji derivates were overlaid in a window ±2000bp, centered at the start of the ZRE (Fig. 5.7 A) The average nucleosome occupancy profile displayed an increased nucleosome occupancy at ZREs during latency (parental Raji cells, black line), indicated by a high $\log_2$-ratio. The elevated $\log_2$-ratio dropped after lytic induction (Raji 3862 cells, red line) and after binding of the DBD of BZLF1 (Raji 3927 cells, green line). This observation indicated that full length BZLF1, but also binding of the truncated version of BZLF1, which encompassed the C-terminal half of BZLF1 with the DBD, only, caused the loss of nucleosomes at ZREs.

Subtraction of the $\log_2$-ratios of the individual ZREs of lytically induced Raji 3862 cells and parental Raji cells are shown in a heat map in Fig. 5.7 B (left panel). The rows depict the subtracted values shown in Fig. 5.7 A at each individual ZRE in a window size of ±2000bp, centered at the start of the ZRE. A cluster analysis was performed to probe for functional groups among all ZREs. The cluster analysis of all ZREs was done on the same heat map matrix with a window size of ±350bp, centered at the ZRE start. The resulting dendrogram shown on the left side of the heat map identified three functional groups of ZREs. We concentrated the following analysis on the two most divergent groups, situated at the upper and lower part of the dendrogram, i.e. group 1 and group 2, respectively. Group 1 of ZREs was comprised of 15 ZREs as indicated in the heat map. The average nucleosome occupancy profile of group 1 ZREs during latency (right panel, parental Raji cells, black line) did not display elevated $\log_2$-ratios at the ZRE. The average $\log_2$-ratios did not change in lytically induced cells expressing full-length BZLF1 (Raji 3862 cells, red line) or in cells containing the binding domain of BZLF1, only (Raji 3927 cells, green line), indicating that the average nucleosome occupancy was similar in all Raji derivates. In contrast, group 2 ZREs, which contained the lower cluster of ZREs as indicated in the heat map, had elevated average nucleosome occupancies at the ZRE during latency (parental Raji cells, black line). The average $\log_2$-ratios dropped after induction of the lytic phase (Raji 3862 cells, red line). Binding of BZLF1 to the ZREs caused a similar reduction of average nucleosome occupancies at the ZREs (Raji 3927 cells, green line).
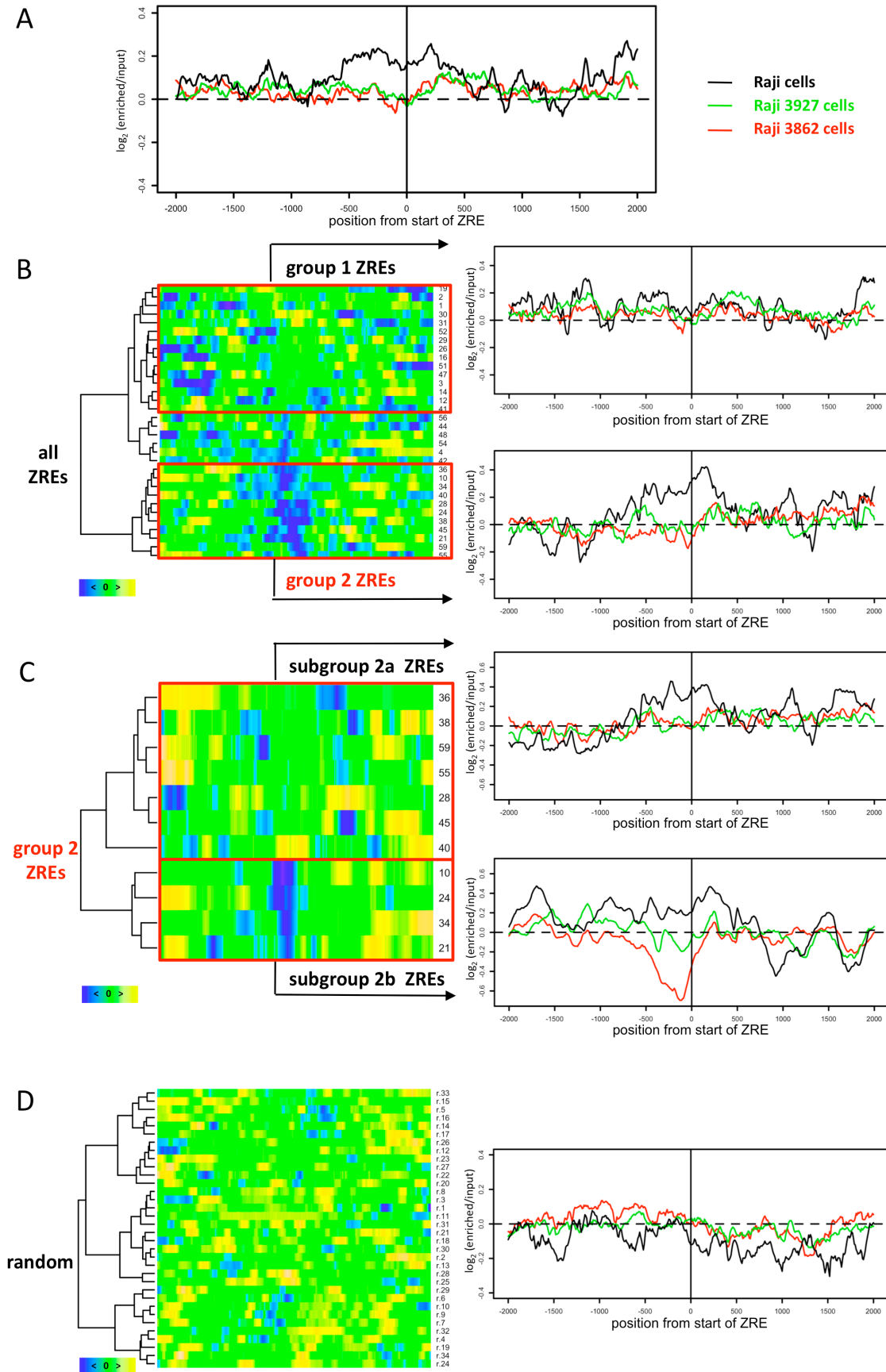
Fig. 5.7 Nucleosome occupancy in BZLF1-regulated promoters
MND-on-Chip experiments were carried out as described in Fig. 5.6. Parental Raji cells (black line), Raji 3927 cells expressing the DNA binding domain (DBD) of BZLF1 constitutively (green line), and induced Raji 3862

cells, which express full length BZLF1 protein upon addition of doxycycline were analyzed. The $\log_2$-ratios of enriched *versus* input data, cluster analyses, and heat maps were obtained as described in the methods section.

(A) Overlay of the average nucleosome occupancy profiles of 32 ZREs in Raji cells (black line), Raji 3927 cells (green line), and lytically induced Raji 3862 cells (red line) in a window of ±2000bp, centered at the start position of each ZRE.

The average nucleosome occupancy appeared elevated at the positions of ZREs in latency as indicated by the increased $\log_2$-ratio in parental Raji cells. Upon induction of the lytic phase (red line), the ratio dropped, indicating a loss of nucleosomes in those cells. Raji 3927 cells, which express the truncated BZLF1, showed a reduced average nucleosome occupancy in comparison to Raji cells, similar to induced Raji 3862 cells.

(B) Left panel: Subtractions of the $\log_2$-ratios of lytically induced Raji 3862 cells and latent parental Raji cells for the 32 single ZREs were visualized in a heat map in a window of ±2000bp, centered at the start position of each ZRE. The rows of the heat map represent individual ZREs. Negative values of the subtraction are in blue, positive values are shown in yellow, and values close to zero are marked in green. A dendrogram was obtained after performing a cluster analysis on the same heat map matrix with a window of ±350bp, centered at the start position of each ZRE, and is shown on the left side of the heat map. Based on this dendrogram, different categories of functional groups among the ZREs were uncovered. We concentrated our analysis of the two most divergent groups of the upper and the lower part of the dendrogram. Group 1 comprised 15 ZREs, which are in the upper part of the dendrogram, as indicated. Group 2 contained all 11 ZREs of the lower cluster, as indicated. Right panel: The overlay of the average nucleosomal occupancy for group 1 ZREs and group 2 ZREs were obtained as in (A). The average nucleosomal occupancy of group 1 ZREs was low and did not reveal noticeable differences between the three Raji cell derivates, as the $\log_2$-ratios indicated. Group 2 ZREs showed a high nucleosome occupancy during latency (black line, right panel), which disappeared in lytically induced Raji 3862 cells (red line). Again, Raji cells, which constitutively express the DBD of BZLF1 only, displayed a pattern comparable to Raji 3862 cells, which express full length BZLF1.

(C) Subtractions of the $\log_2$-ratios of lytically induced Raji 3862 cells (full length BZLF1) and Raji 3927 cells (truncated BZLF1) of the eleven single ZREs of group 2 were analyzed in a heat map (left panel) and in an overlay of the average nucleosome occupancy of the three datasets (right panel) similar to (B). The dendrogram shown on the left side of the heat map clearly illustrated two subgroups. Subgroup 2a ZREs comprised seven ZREs of the upper cluster of the dendrogram. The average nucleosome occupancy of this group was reduced as compared to parental Raji cells but nearly identical between lytically induced Raji 3862 cells and Raji 3927 cells. Subgroup 2b clustered on the lower part of the dendrogram and encompassed four ZREs. An accentuated loss of average nucleosome occupancy was observed in this group, when full length (wild-type) BZLF1 bound to the ZREs of this subgroup as compared to Raji 3927 cells, which expressed only the DBD of BZLF1.

(D) Bioinformatical evaluation of 32 randomly chosen regions that lack ZREs. Left panel: Subtractions of the $\log_2$-ratios of lytically induced Raji 3862 cells and latent parental Raji cells for the 32 positions were visualized in a heat map similar to (B). The subtraction of the two datasets mostly resulted in values close to zero, indicating that there was no significant difference between the datasets. The dendrogram clustered the positions into three groups, but the datasets did not display a noticeable difference as compared to (B) and (C) (*right* panel). Average nucleosome occupancy profiles of parental Raji cells (black line), Raji 3927 cells (green line) and lytically induced Raji 3862 cells (red line) of the 32 positions were displayed as in (A). The profiles did not reveal obvious differences between the analyzed datasets of three different Raji derivates.

The previous analysis looked at differences between latent parental Raji cells and lytically induced Raji 3862 cells. Next, I concentrated on the differences between Raji 3927 cells that express the truncated BZLF1 protein constitutively and lytically induced Raji cells. This analysis addressed the role of BZLF1's N-terminal transactivation domain on the displacement of nucleosomes. Subtractions of the $\log_2$-ratios of the eleven ZREs of group 2 in lytically induced Raji 3862 cells and Raji 3927 cells are visualized in a second heat map (Fig. 5.7 C). A second cluster analysis was performed with this heat map matrix, similar to the first analysis. The dendrogram distinguished two subgroups of ZREs. Subgroup 2a of ZREs showed similar average nucleosome occupancies between the two datasets, indicating a similar displacement of nucleosomes, as shown in the heat map. The average nucleosome

occupancy profile of subgroup 2a certified the previous observation (Fig. 5.7 C, upper right panel): The $\log_2$-ratios of both MND-on-Chip experiments with Raji 3862 cells and Raji 3927 cells dipped likewise as compared to the high average nucleosome occupancy in parental Raji cells. In addition, full length BZLF1 including its transactivation domain induced the precipitous loss of nucleosomes at ZREs of subgroup 2b (red matrix in Fig. 5.7 C, lower right panel). The overlay of the average nucleosome occupancy profiles of subgroup 2b showed a high nucleosome occupancy in latent parental Raji cells (black line), as expected. Binding of truncated BZLF1 (Raji 3927 cells, green line) caused a drop in the $\log_2$-ratios, indicating a lower average nucleosome occupancy. The induction of the lytic phase in Raji 3862 cells resulted in a collapse of the $\log_2$-ratios (red line) as compared to Raji 3927 cells and parental Raji cells, indicating the formation of hypersensitive sites.

The subtraction of the $\log_2$-ratios of MND-on-Chip data with lytically induced Raji 3862 cells and parental Raji cells at 32 randomly chosen positions of the EBV genome without ZREs did not reveal differences between the two datasets (Fig. 5.7 D, left panel). A cluster analysis identified three different groups, as indicated in the dendrogram on the left side of the heat map, but noticeable differences were not obtained as compared to the analysis of the 32 ZRE sites. The overlay of the average nucleosome occupancy profiles confirmed, that there was no remarkable difference between the three Raji cell derivates (Fig. 5.7 D, right panel).

In conclusion, the MND-on-Chip experiments suggested different functional groups of ZREs according to nucleosomal changes. Group 1 ZREs had low average nucleosome occupancies during latency, indicating that the promoters belonging to this group were not tightly occupied with nucleosomes. Binding of full length or truncated BZLF1 did not rearrange nucleosomes. Subgroup 2a of ZREs displayed a high average nucleosome occupancy in latency, which disappeared after binding of truncated BZLF1 or after induction of the lytic phase with full length BZLF1. The initial finding indicated a tight repression of the respective promoters during latency, which BZLF1 could overcome by simply binding to the ZREs. ZREs of subgroup 2b revealed a high average nucleosome occupancy in latency, similar to subgroup 2a, but binding of truncated BZLF1 was insufficient to remodel the nucleosomes completely. The transactivation domain of BZLF1 induced the complete loss of nucleosomes. This finding implied, that certain ZRE-regulated promoters are tightly controlled and repressed during latency. As a consequence, their activation requires energy to remodel the nucleosomes upon lytic induction, which could be provided by cellular chromatin remodeling complexes that rely on ATP as the energy source. Fig. 5.8 shows the single profiles of two representative ZREs for each group.
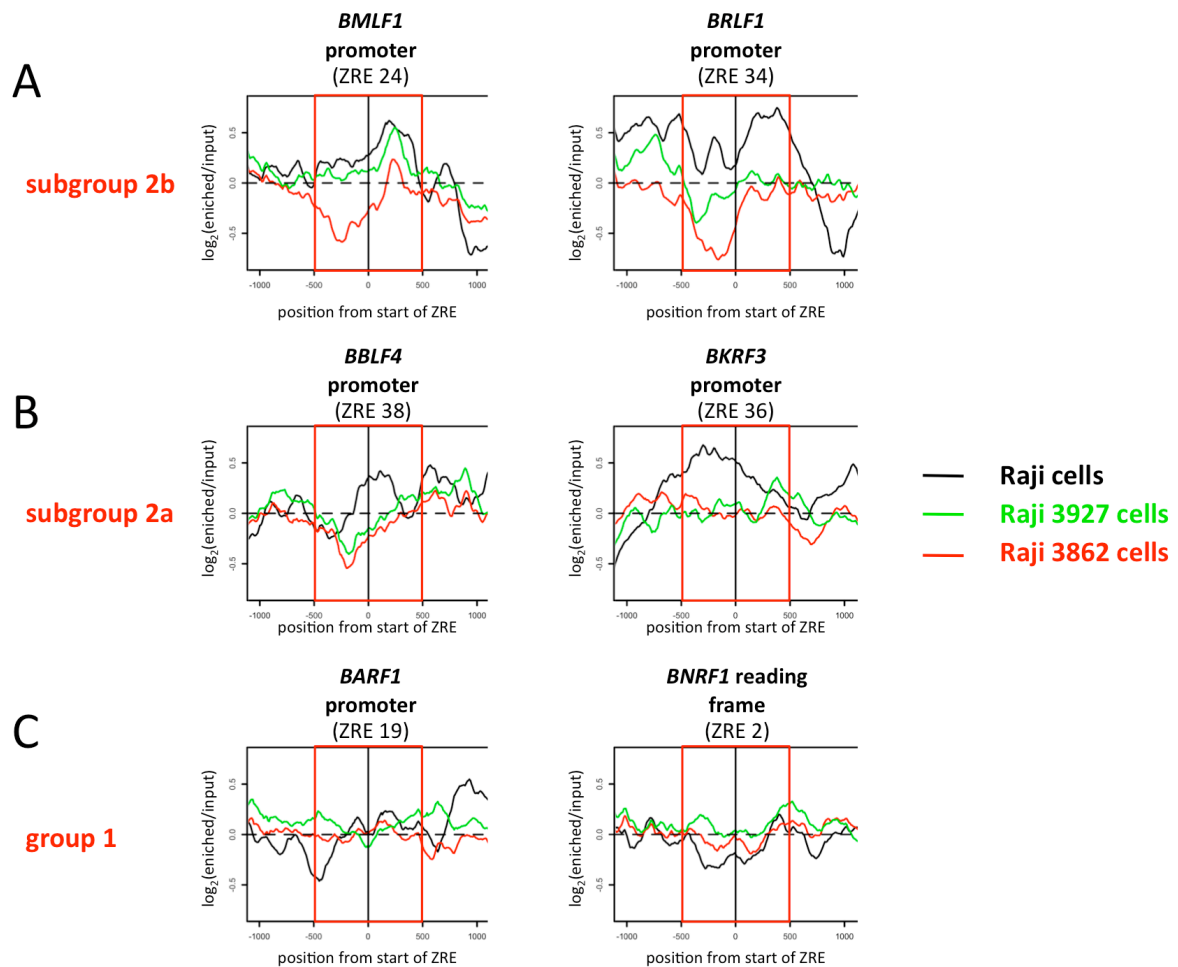
Fig. 5.8 Close-up of six single ZREs of different ZRE groups
MND-on-Chip experiments in conjunction with bioinformatical analysis identified ZREs, which belong to different groups. The nucleosome occupancy profiles of two members of each group is shown in a window of ±1000bp, centered at the start of the ZRE, similar to Fig. 5.7.
(A) Subgroup 2b of ZREs had a high average nucleosome occupancy during latency (parental Raji cells, black line), which dramatically dropped after induction of the lytic phase (induced Raji 3862 cells, red line). This observation was confirmed by looking at the individual nucleosome occupancy profiles of the promoters of *BMLF1* and *BRLF1*. The binding of the DBD of BZLF1, only, (Raji 3927 cells, green line) induced modest changes at the ZRE of the *BRLF1* promoter, but did not modify the nucleosome occupancy at the *BMLF1* promoter.
(B) ZREs of the subgroup 2a displayed high average nucleosome occupancy in latency in parental Raji cells, similar to subgroup 2b. Unlike ZREs of the subgroup 2b, BZLF1 DBD alone was sufficient to rearrange and diminish the average nucleosome occupancy (green line). Full length BZLF1 did not alter the average nucleosomal pattern any further (red line). The *BBLF4* promoter and the *BKRF3* promoter exemplify the nucleosome pattern of members of this subgroup.
(C) ZREs of group 1 comprised moderate average nucleosome occupancy during latency. Upon induction of the lytic phase, this nucleosomal pattern remained stable, as it was observed in the *BARF1* promoter and the ZRE in the *BNRF1* reading frame.

The individual nucleosome occupancy profiles of both promoters in a window of ±1000bp, centered at the start position of the ZRE, showed elevated $\log_2$-ratios during latency in parental Raji cells (black line). The $\log_2$-ratios dropped in the *BRLF1* promoter, but not in the *BMLF1* promoter in Raji 3927 cells, expressing the DBD of BZLF1, only. At both promoters, the $\log_2$-ratios collapsed in Raji 3862 cells, presumably forming a hypersensitive site

downstream of the ZRE (Fig. 5.8 A). The ZRE of the *BBLF4* promoter and the *BKRF3* promoter are examples of members of subgroup 2a. The individual nucleosome occupancy was high in latent parental Raji cells in both promoters (black line). Both Raji derivates exhibited a loss of the nucleosomal signal, indicating that the binding of truncated BZLF1 lacking its transactivation domain was sufficient to rearrange nucleosomes. At the *BBLF4* promoter, a hypersensitive site formed upon binding of BZLF1. The nucleosome occupancy in the *BKRF3* promoter dropped to average levels (Fig. 5.8 B). Examples of members of group 1 are the *BARF1* promoter and the *BNRF1* reading frame. Both ZREs did not show elevated levels of nucleosome occupancies in parental Raji cells, and a substantial difference between all Raji derivates was not visible (Fig 5.8 C). Table 5.2 lists the ZREs hierarchically and according to their functional group.

Tab. 5.2 Catalog of functional groups of ZREs
Cluster analysis identified three groups of ZREs as indicated. The table gives an overview of the members of each group and the function of the corresponding gene (DBD: DNA binding domain, AD: activation domain, n.o.s. not otherwise specified).

group	ID	location	function of respective gene
Subgroup 2b - high occupancy in latency - DBD and AD are necessary for remodelling of nucleosomes	10	*oriLyt*	origin of ***lytic replication***
	24	*BMLF1* promoter	RNA export factor, ***lytic replication***
	34	*BRLF1* promoter	transcription factor, ***lytic replication***
	21	*BMRF1* promoter	polymerase associated factor, ***lytic replication***
Subgroup 2a - high occupancy in latency - DBD is sufficient for remodelling of nucleosomes	40	*BBLF2/3* promoter	primase associated factor, ***lytic replication***
	38	*BBLF4* promoter	helicase, ***lytic replication***
	55	*BALF5* reading frame	DNA polymerase, ***lytic replication***
	36	*BKRF3* promoter	uracil DNA glycosidase
	28	*BLLF1* reading frame	gp350/220, glycoprotein
	45	*BDLF3* promoter	gp150, glycoprotein
	59	*BNLF2a*/b promoter	immune evasion
Group 1 - low occupancy in latency - no remodelling of nucleosomes	19	*Barf1* promoter	glycoprotein
	2	*BNRF1* reading frame	tegument protein
	1	*BNRF1* reading frame	tegument protein
	30	*EBNA3B* reading frame	latency
	31	*BZLF1* promoter	transcription factor
	52	*BILF2* promoter	glycoprotein
	29	n.o.s.	n.o.s.
	26	*BSRF1* promoter	tegument protein
	16	*BPLF1* reading frame	tegument protein
	51	*BdRF1* promoter	structural protein
	47	*BcLF1* reading frame	major capsid antigen, structural protein
	3	*BCRF1* reading frame	viral IL10, immune evasion
	14	*BFLF1* promoter	glycoprotein
	12	*BFLF2* promoter	viral maturation at the membrane
	41	*BGLF5* promoter	alkaline exonuclease

The individual ZRE clusters indicated a functional relationship between the group and the function of the regulated gene. Genes that are absolutely essential for EBV's DNA replication upon lytic induction are *BMLF1*, *BRLF1*, *BMRF1*, *BBLF2*/3, *BBLF4*, *BALF2,* and *BALF5* (Fixman et al., 1992; Feederle et al., 2000). Five out of seven promoters of genes responsible for lytic DNA replication belonged to subgroup 2b or subgroup 2a, indicating that theses promoters are repressed very tightly during latency. The promoters of *BMLF1*, *BRLF1* and *BMRF1* probably require active chromatin remodeling for nucleosome rearrangements upon lytic induction. The promoters of *BBLF2/3* and *BBLF4* also appeared tightly repressed during latency. The *BALF2* gene is missing in Raji cells and is thus not included in this analysis. The seventh gene important for lytic replication is the DNA polymerase *BALF5*. A ZRE located in the gene body of *BALF5* behaved similar to the ZREs of the five other promoters of genes important for DNA replication, which are present in Raji cells. Its role is enigmatic but could contribute to the transcriptional regulation of *BALF5*.

5.2.3   Indirect   and   Chromatin   Immunoprecipitation experiments confirm the loss of nucleosomes at ZREs

MND-on-Chip experiments indicated a loss of nucleosomes at ZREs, which was validated in two additional, independent experiments (Fig. 5.9) with parental Raji cells (latent) and doxycycline induced Raji 3862 cells (lytic). Chromatin Immunoprecipitation (ChIP) is a method to analyze interactions of proteins with DNA. Chromatin of parental Raji cells and induced Raji 3862 cells, that express full length BZLF1, was isolated and crosslinked with formaldehyde to covalently fix the interaction between proteins and DNA. Chromatin was sonicated to an average DNA size of 400bp.

Incubation with an antibody specific for histone H3 and precipitation with sepharose beads resulted in an enrichment of DNA that was wrapped around nucleosomes containing histone H3. DNA was purified and analyzed by qPCR for the enrichment of two different promoters that contained ZREs (*BRLF1* and *BMRF1*) in comparison with the ZRE-free *W* promoter (*Wp*). The three promoters were enriched in histone H3 during latency in parental Raji cells (2-3% of input DNA). The ZRE-containing promoters of *BRLF1* and *BMRF1* showed a clear reduction of the histone H3 signal (1% of input DNA) in lytically induced Raji 3862 cells. The ZRE-free *Wp* was not affected by lytic induction (Fig. 5.9 A).
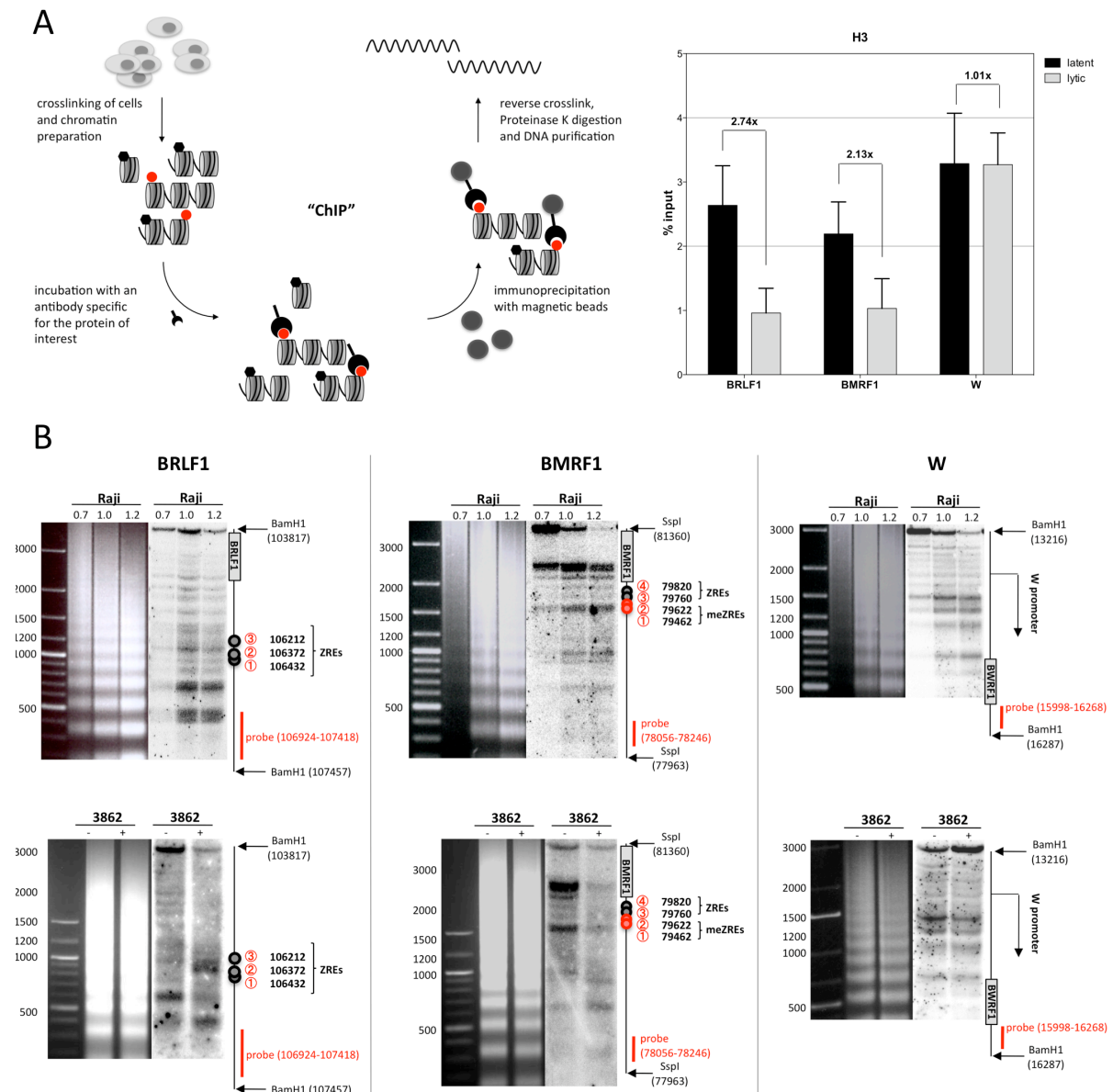
Fig. 5.9 Chromatin Immunoprecipitation and indirect endlabeling experiments confirm the loss of nucleosomes at ZREs

(A) Left panel: Chromatin Immunoprecipitation (ChIP) experiments were performed to confirm the drop in nucleosomal occupancy around ZREs in the presence of BZLF1. Chromatin was crosslinked with formaldehyde and sheared to an average DNA size of 400bp. After incubation with an antibody specifically recognizing histone 3 (H3), the complexes of antibody, protein, and DNA were precipitated with sepharose beads. After three washing steps with increasing stringencies and subsequent Proteinase K digestion, DNA was purified and analyzed by qPCR. The experiment was repeated three times with chromatin preparations from parental Raji cells and uninduced Raji 3862 cells (latent) and compared with three chromatin preparations of Raji 3862 cells induced with 100ng/ml doxycycline over night. Right panel: Quantitative PCRs of histone H3 ChIPs confirmed the loss of nucleosomes in two promoters containing ZREs (*BRLF1* and *BMRF1*) in lytically induced cells. In contrast, the ZRE-free *W* promoter was not affected.

(B) Indirect endlabeling experiments confirmed the microarray hybridization and ChIP results. Chromatin was prepared and treated with limited amounts of MNase. Subsequently, the chromatin was cleaved with a suitable restriction endonuclease to obtain a defined starting point for the analysis. After separation on an agarose gel and Southern blotting, nucleosomal patterns were visualized by hybridizing with a probe specific for a region close to the restriction site as indicated. Bands in the autoradiogram after hybridization reveal the boundaries of nucleosomes and hypersensitive sites. Upper row: Indirect endlabeling experiments in Raji cells covering three different promoters as indicated. The annotations on the right side of the autoradiogram show the cleavage site of the restriction endonuclease, the binding site of the probe, the locations of ZREs, the start of the three transcripts and the coding region of *BRLF1*, *BMRF1,* and the *W* promoter and adjacent genes. Lower row: The three regions were re-analyzed in Raji 3862 cells, either uninduced (-) or induced with 100ng/ml doxycycline over night (+).

In general, the pattern in uninduced cells was similar to the one seen in parental Raji cells. Induction of the lytic phase led to rearrangements in the nucleosomal pattern at ZREs. The *BRLF1* promoter gained a hypersensitive site situated in the ZRE regions after induction of the lytic phase, reflected by a strong additional band of 1000bp in the induced cells. Concomitantly, the signal of faster and more slowly migrating bands are reduced. The *BMRF1* locus also lost nucleosomes at the ZRE locations. A strong signal beyond the cluster of ZRE sites disappeared. In line with the results of panel (A), the nucleosomal pattern of the *W* promoter, which does not contain ZREs, was not affected by the induction of the lytic phase, as expected.

The same promoter sites were also analyzed in indirect endlabeling experiments (Fig. 5.9 B). Chromatin was treated with limited amounts of MNase and cleaved with an appropriate restriction endonuclease at a site close to the region of interest. The DNA was purified, separated on an ethidium bromide agarose gel and transferred to a nylon membrane by Southern blotting. The membrane was hybridized to a radioactively labeled probe that was complementary to a region downstream but close to the cleavage site of the restriction endonuclease. The bands that are visible in the autoradiogram show the boundaries of nucleosomes and hypersensitive sites with fragment sizes that indicate the distance of the boundaries from the restriction endonuclease cleavage site.

The upper panel of Fig. 5.9 B shows the results of indirect endlabeling experiments in parental Raji cells. Raji chromatin was partially digested with increasing amounts of MNase as indicated and analyzed for nucleosome occupancies in the promoter regions of *BRLF1*, *BMRF1,* and *Wp*. The restriction endonuclease cleavage sites, the probe location, and selected features of EBV's genome are depicted on the right side of the autoradiogram. Bands in the autoradiogram displayed nucleosome boundaries and hypersensitive sites. The evaluation of exact positions is difficult and only possible in very regularly spaced arrays of nucleosomes, as it can be seen in *Wp* upstream of the *BWRF1* reading frame. The bands were 200bp apart from each other, which corresponds to the average size of DNA encompassing one nucleosome plus its linker region. One band seemed to be missing at a height of 1000bp in the array of nucleosomes indicating that this site was not protected from cleavage by a nucleosome. Probably a different protein occupied the DNA in this region. A pattern indicative of a hypersensitive site was present in the *BMRF1* promoter, approximately 2200bp from the restriction endonuclease cleavage site in the *BMRF1* locus. The autoradiogram displayed a very strong band and signals from slower migrating fragments were missing. This observation indicated that MNase preferentially and entirely cleaved this site in the Raji DNA. The lower panel (Fig. 5.9 B) shows the same promoter sites in Raji 3862 cells, without or with addition of doxycycline and induction of the lytic phase of EBV. The uninduced cells displayed a pattern similar to the latent parental Raji cells, as expected. The induction of EBV's lytic phase after the doxycycline-dependent expression of full length BZLF1 caused

major rearrangements in the promoters of *BRLF1* and *BMRF1*, but not in the ZRE-free *W* promoter. The *BRLF1* promoter acquired a pattern that was typical for a hypersensitive site approximately 800bp upstream of the restriction endonuclease cleavage site. The hypersensitive site was situated in the same location as the *BRLF1* ZRE sites, confirming the microarray and ChIP experiments. The *BMRF1* promoter showed overall reduced signal intensities in bands that were more than 1000bp apart from the restriction endonuclease cleavage site after lytic induction, indicating a loss of nucleosomes in this BZLF1-regulated promoter as well. Residual signal originated probably from uninduced cells, as the cells were not sorted for GFP expression after doxycycline induction and induced to 80%, only.

My experiments suggested that different lytic promoters are occupied and presumably repressed by nucleosomes. A fraction of early lytic genes, including genes important for EBV's lytic DNA replication, are tightly repressed during latency, because their ZREs are densely occupied by nucleosomes. Microarray experiments, Chromatin Immunoprecipitations, and indirect endlabeling experiments confirmed that the induction of the lytic phase caused a loss of nucleosomes at these promoters. Furthermore, microarray experiments revealed that the DNA binding domain of BZLF1 was sufficient to remodel chromatin at certain promoters. Other promoters required the transactivation domain of BZLF1 for chromatin remodeling and eviction of nucleosomes. A different subgroup of early lytic gene promoters that contain ZREs did not display high levels of nucleosomes during latency (Fig. 5.7 and 5.8). This observation implied that the corresponding genes were repressed less stringently during the latent state. Induction of the lytic phase did not lead to rearrangements of nucleosomes indicating that nucleosomes in these promoters were no functional barriers to transcriptional activation. The analysis revealed that the positioning of nucleosomes is an important regulatory mechanism for EBV, repressing lytic genes during viral latency.

5.3   Histone modifications and chromatin modifying enzymes in EBV

Post-translational modifications of histone tails influence the functional state of a promoter. Histone modifications can either have repressive or activating effects on transcription, for example by adjusting the contact strength between DNA and histone proteins by adding negative or positive charges or by recruiting non-histone proteins to the site of modification (Kouzarides, 2007). Methylation of lysine residues can occur at different lysines in the

histone tail to different degrees, as lysines can be mono-, di-, or trimethylated. The diversity of lysine methylation can result in many different functions. Two prototypical lysine modifications that repress gene expression are trimethylation of histone H3 at lysine 9 or lysine 27 (H3K9me3 and H3K27me3). H3K9me3 is implicated in heterochromatin formation and H3K27me3 leads to a stable and inherited repression of genes through interaction with Polycomb group proteins (Hansen et al., 2008). A characteristic lysine methylation mark that has a positive effect on transcription is trimethylation of histone H3 at lysine 4 (H3K4me3). H3K4me3 is mainly enriched in promoters and 5' ends of the coding region of active genes at transcriptional start sites. H3K4me3 is connected to actively transcribing RNA Polymerase II (Barski et al., 2007). In certain gene promoters of silent genes, H3K27me3 colocalizes with H3K4me3; these regions also encompass inactive RNA polymerase II. This so-called "bivalent chromatin" is indicative of genes that are "poised" to become active (Bernstein et al., 2006). The epigenetic mode of repressed viral lytic genes has never been assessed before. I performed Chromatin Immunoprecipitation (ChIP) experiments to explore the pattern of histone modifications at latent, early lytic, and late lytic genes in parental Raji cells and in lytically induced Raji 3862 cells that express full length BZLF1 upon addition of doxycycline. I could show that lytic genes typically encompass H3K27me3 in the latent state, presumably because Polycomb group proteins (PcGs) sustain this repressive histone code at these promoters. Induction of the lytic phase caused major changes in the histone modification pattern. The repressive mark H3K27me3 entirely disappeared in the EBV genome, as did PcG-proteins. H3K4me3 was established selectively at promoters of early lytic genes upon lytic induction, where it was absent during latency.

5.3.1 Histone H3 and its post-translational modifications at EBV promoters

ChIP experiments were performed as described before (chapter 5.2.3 and Fig. 5.9) with parental Raji cells to assess the histone modification pattern in the latent state and with doxycycline induced Raji 3862 cells to analyze the changes in the histone modification pattern after the lytic switch. The promoters of latent, early lytic, and late lytic genes were analyzed by qPCR for footprints of histone H3, H3K4me3, H3K9me3, and H3K27me3. Fig. 5.10 shows the results of this survey. The amount of ChIP DNA was assessed as percentage of input DNA (% input) and is depicted on the y-axis; the loci tested are indicated on the x-axis. Black bars show results for EBV's latent state in parental Raji cells (latent) and light grey bars depict the results after the lytic switch in induced Raji 3862 cells (lytic).

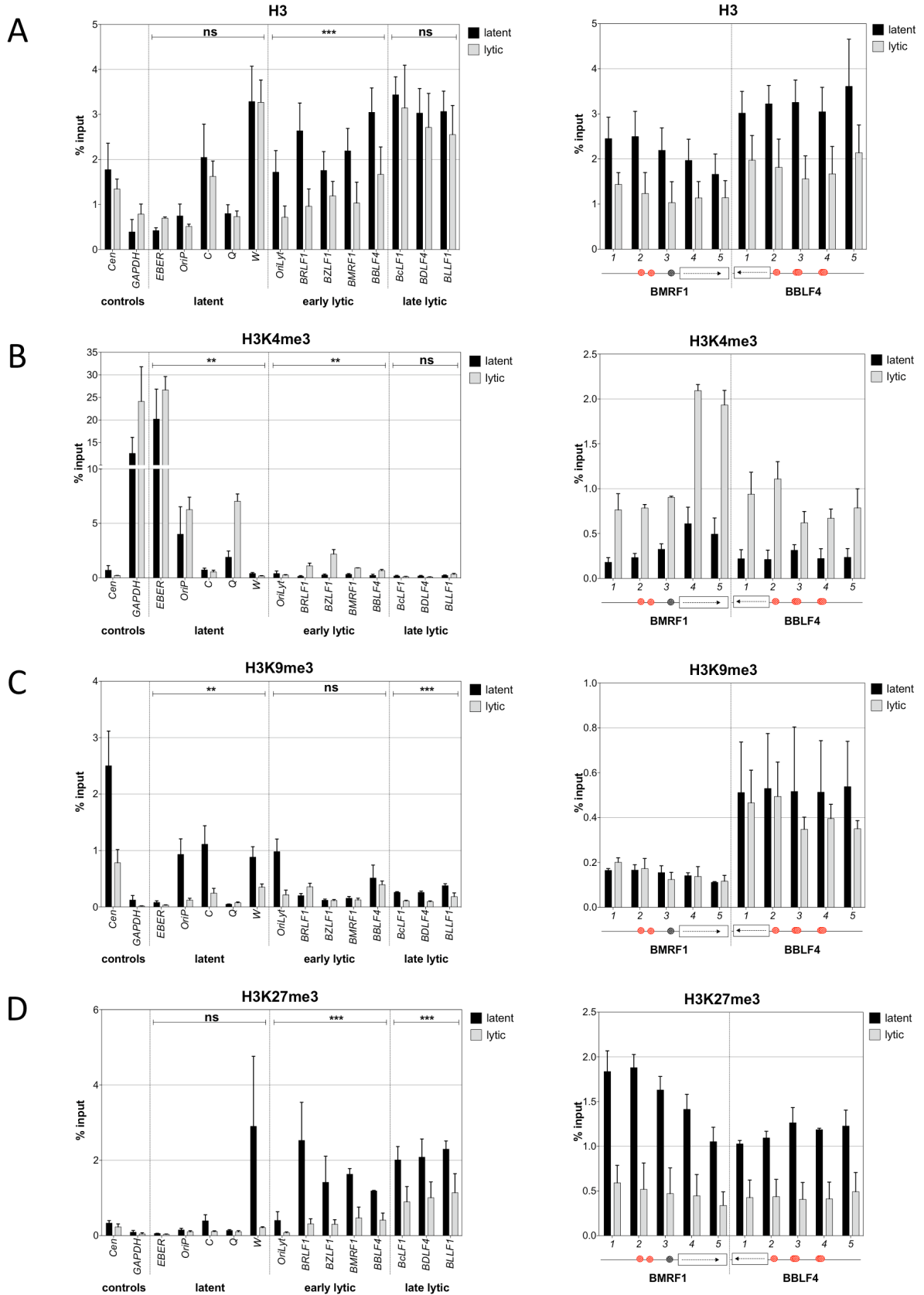**Fig. 5.10 Chromatin immunoprecipitation of histone H3 and its N-terminally modified versions**
ChIPs were performed as described in Fig 5.9 with antibodies directed against histone H3 and its posttranslational modifications. DNA was analyzed by qPCR for enrichment at different loci in the human or the EBV genome. Amplicons of Centromer1 (Cen) served as a control of an epigenetically repressed locus, while

the promoter of the housekeeping gene *GAPDH* is indicative of chromatin with active epigenetic marks. Enrichment of EBV DNA was evaluated with primer pairs covering latent, early lytic, and late lytic promoters. The left panel provides the results of the promoter regions analyzed, while the right panel gives a close-up of two early lytic promoters (*BMRF1* and *BBLF4*). Note the different scales of the y-axis.

(A) Lytic induction caused the loss of H3 at early lytic gene promoters (p-value: $6x10^{-9}$), which is in line with the results depicted in Fig. 5.7-5.9. Late lytic promoters were slightly affected, while latent promoters were completely unchanged.

(B) The activation mark H3K4me3 is only found at the *EBERs*, the *Q* promoter, and at *oriP* in latency. After induction of the lytic phase, H3K4me3 levels were found elevated at early lytic promoters (p-value: $1.4x10^{-3}$) and at the *Q* promoter. Late lytic genes do not show any sign of this modification. H3K4me3 was more prevalent at the transcribed gene body than at the promoter because this modification is linked to RNA Pol II active transcription (right panel).

(C) The latent *C* and *W* promoters contained H3K9me3, and also *oriP* displayed this repressive mark at low levels in latently infected cells, which was lost upon induction of the lytic phase (p-value: $1x10^{-3}$). In two early lytic promoters, *BMRF1* and *BBLF4*, induction of EBV's lytic phase did not alter the levels of H3K9me3 (right panel), indicating that this modification is not important for the repression of early lytic genes.

(D) The repressive modification H3K27me3 correlated with repressed early lytic and late lytic gene promoters, indicated by the strong enrichment in latently infected cells. The repressive mark was completely lost in early lytic promoters and at the *W* promoter after induction of the lytic phase (p-value: $6x10^{-5}$). Late lytic promoters showed a significant reduction in H3K27me3 levels, but the enrichment at theses promoters was still slightly above background level.

The left panel provides two cellular control loci: Centromer 1 (Cen), a silent locus and GAPDH, an active locus, together with promoters and transcripts of latent genes (*EBER*, *oriP*, *Cp*, *Qp,* and *Wp*), five early lytic promoters, that encompass ZREs and meZREs (*oriLyt*, *BRLF1*, *BZLF1*, *BMRF1,* and *BBLF4*), and three late lytic promoters, that are not regulated by BZLF1 (*BcLF1*, *BDLF4,* and *BLLF1*). The right panel is a close-up of two early lytic promoters (*BMRF1* and *BBLF4*).

Histone H3 levels were high in *Cp* and *Wp* during latency, but not in the *EBER* locus, *oriP* and the *Q* promoter (Fig. 5.10 A). The lytic switch did not cause any change in the tested latent promoters. Early lytic promoters had elevated levels of histone H3 during latency, but the signal diminished significantly after the induction of the lytic phase. The close-up of the *BMRF1* promoter and the *BBLF4* promoter confirmed this observation. Late lytic promoters were also occupied with histone H3. Similar to latent promoters, the lytic switch did not induce a change in the histone H3 pattern. The ChIP results of histone H3 were in line with the results of the MND-on-Chip analysis. Histone H3 is part of the protein core of nucleosomes; a reduction in the amount of this protein indicates a loss of the nucleosome. ChIP experiments confirmed, that the nucleosomal loss is specific for promoters of early lytic genes that encompass ZREs. The ChIP results of latent promoters indicated that *Cp* and *Wp* are occupied and presumably repressed by nucleosomes during latency. The *Q* promoter was the only latent promoter with low levels of histone H3, implying that latent gene expression in Raji cells is initiated at this promoter (Fig. 5.10 A).

H3K4me3 levels were dramatically high at the *EBER* locus, indicating that the EBERs are

actively transcribed in latent Raji cells (Fig. 5.10 B). Enrichment of H3K4me3 was also detected at *oriP* and *Qp* but to a lesser extent. Early and late lytic promoters did not show any sign of H3K4me3 modification during latency. Lytic induction caused a slight but significant increase in H3K4me3 levels at early lytic promoters, which increased towards the 5' end of the gene body, as indicated in the close-up of the *BMRF1* promoter and the *BBLF4* promoter.

Also *EBERs*, *oriP*, and *Qp* showed elevated levels of H3K4me3 indicating an upregulation of latent gene expression upon lytic induction as well. Trimethylation of H3K9 was elevated in *oriP*, *Cp* and *Wp* during latency (Fig. 5.10 C) but to a lesser extent than the positive control (Cen). Lytic induction caused a loss of this modification in all latent H3K9me3-positive loci. The early lytic promoters of *BRLF1*, *BZLF1*, and *BMRF1* were not enriched in H3K9me3. Only *BBLF4* appeared to encompass little amounts of this epigenetic modification. Lytic induction did not induce any change in the H3K9me3 pattern in early lytic promoters, which was confirmed in the close-up of the promoters of *BMRF1* and *BBLF4*. *OriLyt* showed a similar pattern as *oriP*; both origins of replication carry H3K9me3 in latency, but lose the modification upon lytic induction. Late lytic promoters showed only background levels of H3K9me3. The results indicated that H3K9me3 did not contribute to repression of lytic gene promoters.

H3K27me3 levels were dramatically increased in early lytic and late lytic promoters during latency (Fig 5.10 D). This modification was also prevalent at *Wp*. Induction of the lytic phase caused a major drop of the H3K27me3 signal, indicating a complete loss of this repressive mark. The close-up of the *BMRF1* promoter and the *BBLF4* promoter showed high levels of H3K27me3 at binding sites of BZLF1 (ZREs and meZREs, indicated as black and red circles, respectively) during latency. Lytic induction reduced the levels of H3K27me3 in the tested regions, signals dropped virtually to background levels.

In summary, levels of histone H3 and its N-terminally modified versions suggested the mode of repression during latency. Latent gene promoters that are not used in latent Raji cells, i.e. *Wp* and *Cp*, were enriched in histone H3 and carried H3K9me3 and H3K27me3 marks, indicating a strong repression of these promoters. H3K4me3 modification, an active mark, was absent. The active latent regions, i.e. the *EBER* locus, *oriP*, and *Qp*, showed an inverse picture with elevated levels of H3K4me3 and missing repressive modifications. Surprisingly, *oriP* encompassed H3K9me3 in addition to activation marks. A colocalization of H3K9me3 and H3K4me3 is so far not described, but resembles bivalent chromatin, that comprises H3K4me3 in conjunction with H3K27me3. But it is not clear whether H3K9me3 and

H3K4me3 colocalize on the same EBV episome. Sequential ChIP assays that make use of two antibodies in a row could address this point. At early lytic gene promoters, H3K27me3 marks dominated during latency, together with an unusually high occupancy with histone H3. Induction of the lytic phase caused a loss of histone H3 and H3K27me3 epigenetic marks at early lytic promoters and an increase in H3K4me3 levels, indicating an activation of the BZLF1-regulated early lytic genes. H3K9me3 appeared to be unimportant for early lytic gene regulation. Late lytic gene promoters exhibited a similar pattern as early lytic promoters during latency. Induction of EBV's lytic phase erased the H3K27me3 marks but histone H3 levels were unchanged and elevated, and trimethylation of H3K4 was undetectable.

Together my findings demonstrated that the loss of the repressive modification H3K27me3 upon lytic induction affects the entire EBV genome but depletion of nucleosomes and the addition of activating marks are specific for early lytic genes, only.

5.3.2   Polycomb proteins set up the repressed state in EBV's lytic promoters

H3K27me3 is a modification that is associated with repression by PcG-proteins. ChIP experiments were performed with an EZH2[*]-specific antibody. EZH2 is a methyltransferase responsible for the trimethylation of H3K27 and is part of one out of two Polycomb repressive complexes (PRCs), termed PRC2.
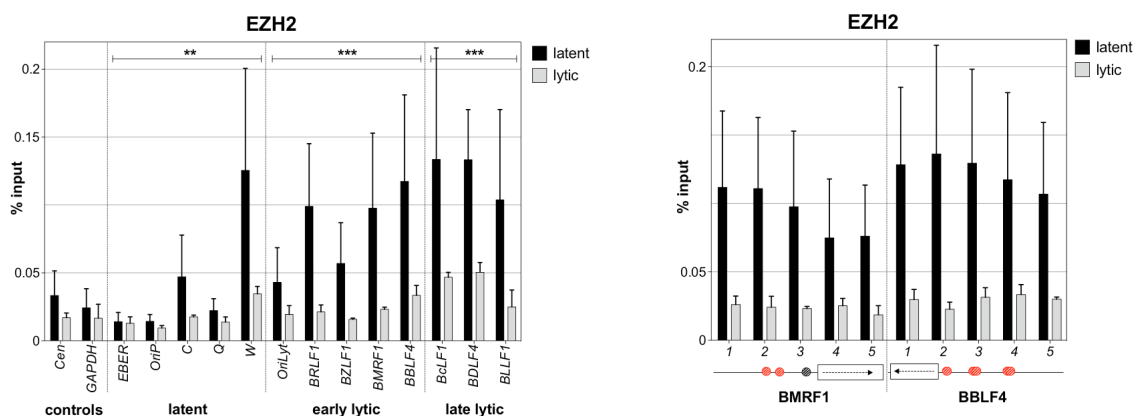
Fig. 5.11 Chromatin immunoprecipitation of the H3K27-methyltransferase EZH2
Chromatin Immunoprecipitations and qPCR were carried out as described in Fig. 5.10. Enhancer of zeste 2 (EZH2) is a histone methyltransferase responsible for trimethylation of histone H3 at lysine 27 (H3K27me3). It is part of the Polycomb Repressive Complex 2 (PRC2), which together with PRC1 leads to an inherited and stable repression of local chromatin. In line with the results of H3K27me3 ChIPs, this protein was present at promoters of early and late lytic genes, and also at the W promoter during latency. After induction of the lytic phase, the signal is lost (p-value: 7×10^{-4} for early lytic genes; 4×10^{-4} for late lytic genes).

---

[*] human enhancer of zeste 2

ChIP experiments were performed with the same conditions and with the same readout as in 5.3.1 to test whether PRC2 colocalizes with H3K27me3. Fig. 5.11 shows the results of EZH2 ChIPs in parental Raji cells (latent, black bars) and in induced Raji 3862 cells (lytic, light grey bars). Similar to H3K27me3, *Cp* and *Wp* showed elevated levels of EZH2 during latency. Lytic induction resulted in a loss of the EZH2 signal to background levels. Early lytic and late lytic genes likewise showed a high enrichment of EZH2 protein in the promoter region. The induction of EBV's lytic phase caused the loss of the protein at the early lytic promoters. EZH2 ChIPs perfectly mirrored the results of H3K27me3 ChIPs indicating that this methyltransferase is responsible for the trimethylation of H3K27 in repressed EBV promoters. The results demonstrated that Polycomb repression is an important mechanism to maintain EBV's lytic genes in a silent state during latency.

5.4 Occupation of EBV promoters with RNA polymerase II

Transcription of all protein-coding genes is carried out by RNA polymerase II (RNA Pol II). The largest subunit of RNA Pol II, i.e. RPB1, contains a long carboxy-terminal domain (CTD) with 52 repeats of the consensus heptapeptide sequence YSPTSPS (Corden et al., 1985). The heptapeptide is target of a number of post-translational modifications that regulate the function of RNA Pol II. Non-consensus repeats are more frequent towards the C-terminus and might provide special functions to specific regions within the CTD (Fong et al., 2001). Unphosphorylated RNA Pol II is termed Pol IIA; RNA Pol II encompassing phosphorylated serine residues in the CTD is described as Pol II0 form. The two forms of RNA Pol II can be discriminated in a ChIP experiment by the use of CTD-specific antibodies. RNA Pol II is detected in promoter regions of active genes, paused genes, and so-called poised genes; and the CTD-modification pattern of Pol II is indicative for the promoter type (reviewed in Brookes et al., 2009).

RNA Pol II is recruited in a hypophosphorylated form to promoters of active genes. The initiation of transcription, the so-called promoter escape, requires the phosphorylation of serine 5 of the CTD. For transcript elongation, RNA Pol II is additionally phosphorylated at serine 2, and termination of transcription involves the erasure of phosphorylation marks, which again allows the recruitment of RNA Pol II to active promoters. This active cycle of RNA Pol II is characterized by a specific pattern of CTD- and histone modifications at the promoter of the active gene. ChIP experiments detect a peak of serine 2-unphosphorylated

RNA Pol II at the transcriptional start site (TSS), elevated levels of serine 5-phosphorylated RNA Pol II at the TSS and the gene body, and enrichment of serine 2-phosphorylated RNA Pol II in the gene body. Additionally, H3K4me3 can be found upstream of the CDS.

In paused genes, initiation of RNA Pol II takes place similar to active genes, but a lack of serine 2-phosphorylation and the involvement of negative elongation factors stall the polymerase at a promoter-proximal position. Low levels of serine 2-unphosphorylated and serine 5-phosphorylated RNA Pol II, and a minor enrichment of H3K4me3 at the TSS characterize paused genes.

A *poised polymerase* is phosphorylated at serine 5 and lacks serine 2-phosphorylation. Transcriptional elongation of these genes is possible at very low levels without the phosphorylation of serine 2. Poised polymerases are associated with "bivalent chromatin" that is repressed by PcG-proteins through trimethylation of H3K27 but is additionally trimethylated at H3K4. In ChIP experiments, the polymerase can only be detected with an antibody directed against the serine 5-phosphorylated form of RNA Pol II. An antibody recognizing unphosphorylated serine 2 at the CTD is not able to detect the polymerase indicating that the CTD owes a not characterized modification or is in an unusual conformation (reviewed in Brookes et al., 2009).

In my experiments, epigenetic modifications in early lytic promoters indicated a tight repression during latency and an activation of transcription after EBV's switch to the lytic phase. Occupation of the promoters by RNA Pol II has not been analyzed previously. ChIP experiments using antibodies specific for different RNA Pol II CTD modifications were performed to clarify whether lytic promoters were paused, poised, or silent during latency and whether an active polymerase could be detected after the switch to the lytic phase. I could show that lytic promoters did not encompass any form of the polymerases during latency indicating that the genes were completely silent. After the switch to the lytic phase, RNA Pol II was detected in promoter regions of early lytic genes with an antibody specific for the amino-terminal domain of the polymerase, but all tested CTD-specific antibodies failed to detect any form of the polymerase. Cleavage of the CTD was excluded by western blot immunodetection, indicating that the RNA Pol II CTD is intact and present in lytically induced cells but might be in an atypical conformation. The unusually modified CTD of RNA Pol II clearly supported transcription of early lytic genes.

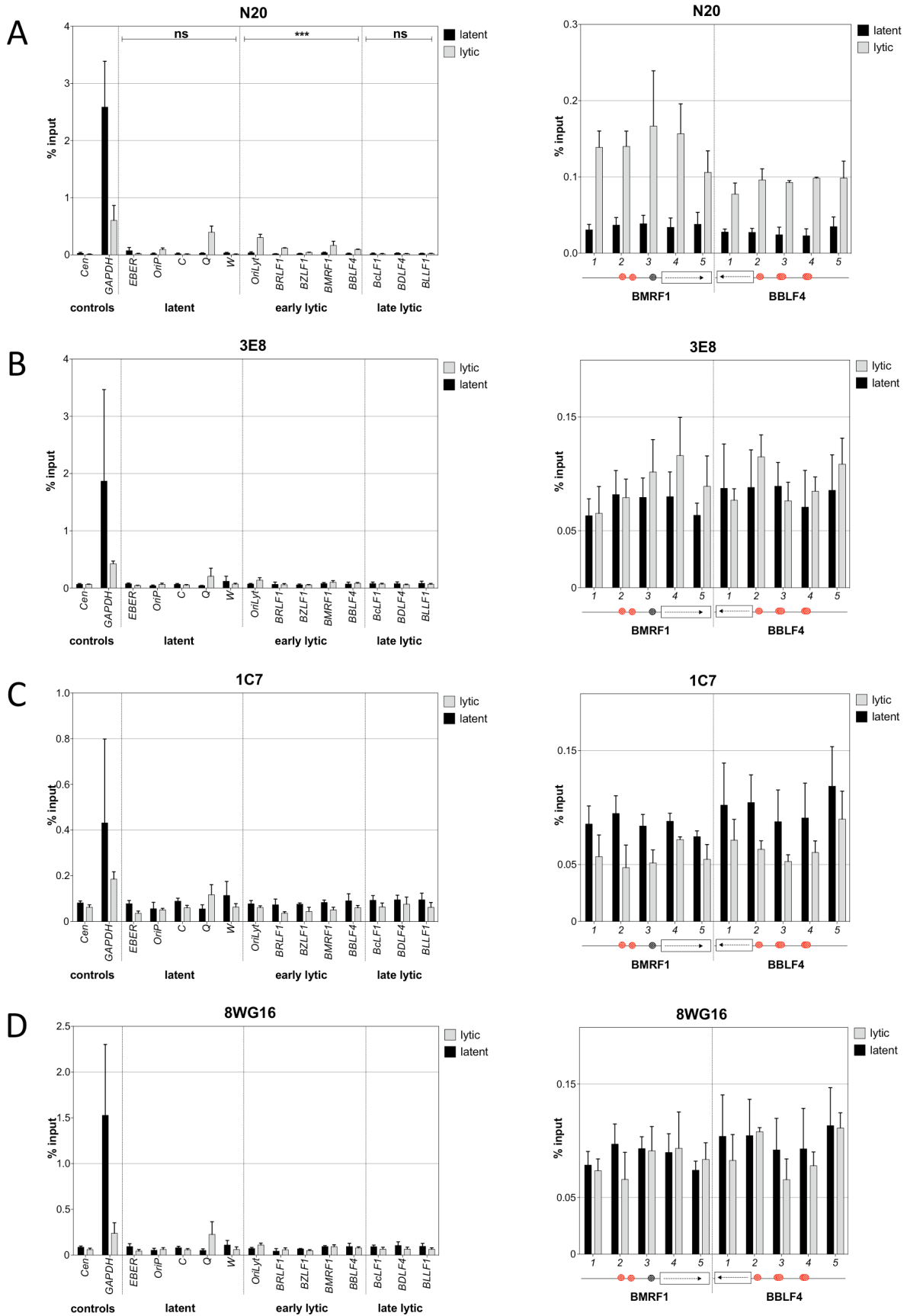**Fig. 5.12 Chromatin Immunoprecipitations of RNA Pol II and its CTD-modified versions**
Chromatin Immunoprecipitations and qPCR were carried out as described in Fig. 5.10 with antibodies recognizing the inactive (Pol IIA) and the initiated form (Pol II0) of the polymerase. N20 is an antibody specific for the N-terminal part of RNA Pol II, which detects both forms, 3E8 is directed against the serine5-

phosphorylated C-terminal domain (CTD) of the polymerase and recognizes the Pol II0 form; 1C7 and 8WG16 are mainly specific for the Pol IIA form.

(A) ChIP experiments using an antibody specific for the N-terminal part of RNA Pol II (N20) showed an increase at early lytic promoters (p-value: $4x10^{-4}$), at *oriP* and the *Q* promoter after induction of the lytic phase. During latency, the polymerase was hardly detectable at any EBV locus. Only the EBERs showed a tiny amount of RNA Pol II.

(B-D) Antibodies against different modifications of the C-terminal domain (CTD) of RNA Pol II failed to immunoprecipitate the polymerase at EBV promoters. Note the different scales of the y-axes of the left and the right panel.

5.4.1   ChIP experiments of RNA Pol II and its CTD-modified versions

ChIP experiments were carried out as described before with parental Raji cells (latent) and lytically induced Raji 3862 cells (lytic). Different antibodies were used to detect the different forms of RNA Pol II. The antibody N20 is specific for both forms of the polymerase, Pol IIA and Pol II0, because it recognizes an epitope in the amino-terminal domain of the polymerase. The Ser5-P (3E8) antibody is specific for the serine 5-phosphorylated form of RNA Pol II and recognizes the Pol II0 form. Anti-RNAPII (1C7) antibody was generated with a serine 7-phosphorylated peptide. It recognizes a variety of phosphorylated CTD forms in ELISA, but in a western blot immunodetection it is specific for the Pol IIA form. The antibody CTD (8WG16) recognizes a serine 2-unphosphorylated form of the CTD, and is mainly specific for the Pol IIA form. Fig. 5.12 illustrates the result of the ChIP experiments, which are similarly presented as the experiments in Fig. 5.10 and Fig. 5.11.

ChIP experiments with the N20 antibody detected RNA Pol II at the *EBER* locus during latency at a low level compared to the positive control, i.e. *GAPDH* (Fig. 5.12 A). Early lytic and late lytic promoters did not give rise to any enrichment of the polymerase in ChIP experiments in Raji cells. This indicated, that the promoters are completely silent during latency. In lytically induced cells, *oriP* and the latent *Q* promoter showed increased presence of RNA Pol II. Early lytic promoters encompassed significantly higher levels of the polymerase after lytic induction although the standard was low compared to *GAPDH* enrichment in latent Raji cells. RNA Pol II was not detectable at late lytic promoters after lytic induction. CTD-specific antibodies did not detect the polymerase at any EBV locus after lytic induction, except very minor enrichments at the *Q* promoter (Fig. 5.12 B-D). The effective enrichment of the positive control indicated, that the ChIP experiments were technically successful. This finding was surprising because RNA Pol II is clearly recruited to promoters of early lytic genes as revealed with an antibody (N20), which binds to the N-terminal domain of RNA Pol II. Lytic induction resulted also in a loss of the polymerase at *GAPDH*, which is a housekeeping gene.

The results prompted the hypothesis, that the onset of the lytic phase induced a peculiar change of RNA Pol II, which might include a proteolytical cleavage of the CTD. The CTD-less form of the polymerase is described as Pol IIB form. The Pol IIB form was originally isolated from calf thymus DNA (Corden et al., 1985) but described to be an artifact from sample preparation later (Kim et al., 1986). Pol IIB is not able to initiate and support transcription on promoter sites but mediates nonselective transcription on calf thymus DNA at nicks and gaps in the DNA, as demonstrated *in vitro* (Dahmus et al., 1983). Later, RNA Pol II mutants lacking the last non-consensus repeat (CTD52) and mutants lacking the repeats one to three (CTD1-3) of CTD were shown to transform into the Pol IIB form, suggesting a role for the removal of the CTD *in vivo*. It was hypothesized that dipping the CTD could be a response to stress, particularly after UV irradiation (Chapman et al., 2004; Chapman et al., 2005), but the mechanism of regulation is still unclear. A cleavage of CTD upon lytic induction could be a viral strategy to govern the transcription machinery and to concentrate the activity of RNA polymerase II entirely to the expression of viral genes.

5.4.2 The Pol IIB form cannot be detected in western blot immunodetection after induction of the lytic phase

The Pol IIB form can be detected in a western blot immunodetection with an antibody specific for the amino-terminal part of RNA Pol II (Pol 3-3). This antibody is able to recognize all forms of the polymerase. Raji 3862 cells were induced for 6, 12, and 24 h with 100ng/ml doxycycline. Uninduced cells, parental Raji cells with and without doxycycline, and B95.8 cells served as controls. Fig. 5.13 shows the result of the western blot immunodetection.

Raji 3862 cells contained two forms of the polymerase, Pol II0 and Pol IIA. The CTD-less polymerase (Pol IIB) was not detected in the samples, neither during latency nor after lytic induction with doxycycline. A minor fraction of B95.8 cells spontaneously supports the lytic phase of EBV. This cell line contained all three forms of the polymerase. Whether the Pol IIB signal originated from cells undergoing the lytic switch remained an open question, as the induced lytic phase in Raji 3862 cells did not yield the truncated Pol IIB form.

The western blot immunodetection did not detect the presumed proteolytic cleavage of the CTD, which might support lytic gene transcription upon induction of the lytic phase. The failure of CTD-specific antibodies to detect the polymerase in ChIP experiments could be a hint for unknown CTD modifications that are not recognized by the antibodies used in this study, but likely favors viral as opposed to cellular transcription. Another explanation for the

result could be an atypical folding of the CTD that precludes the antibodies from binding to their epitopes.
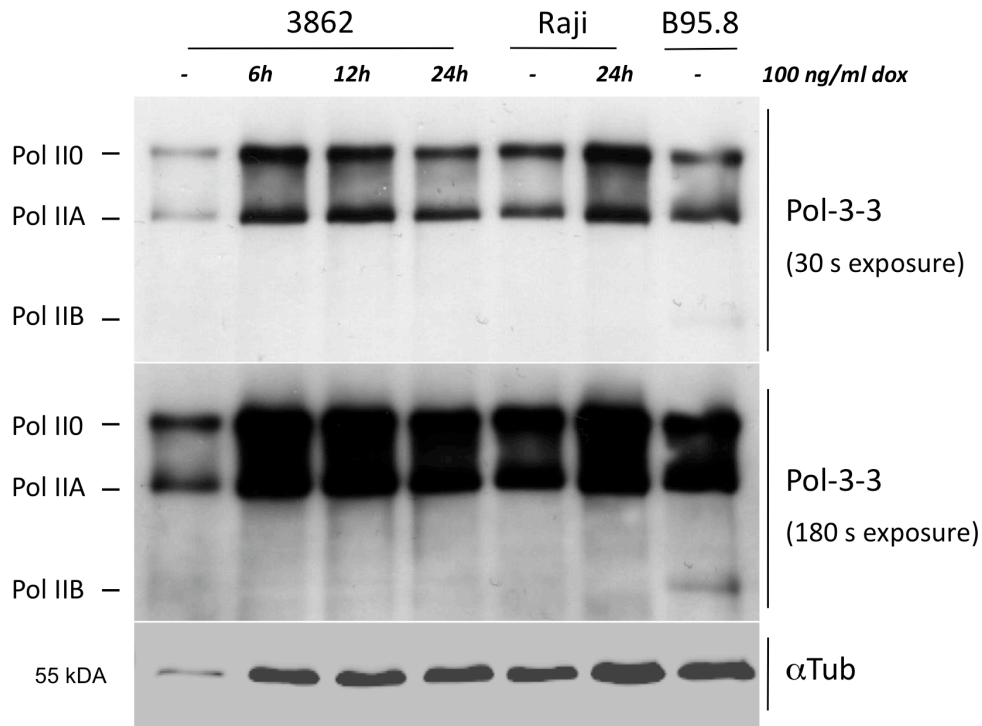
Fig. 5.13 Western blot immunodetection of RNA Pol II in lytically induced cells
Protein lysates of Raji 3862 cells after different time points *post* induction with doxycycline, parental Raji cells, and B95.8 cells were analyzed in a western blot immunodetection with an antibody recognizing the N-terminal part of Pol II (Pol-3-3). This antibody recognizes the initiated form of Pol II (Pol II0), the unphosphorylated form of Pol II (Pol IIA), and the CTD-less form of Pol II (Pol IIB). Tubulin was used as a loading control (α-Tub, bottom panel). Signal detection was performed either after 30s (top panel) or after 180s (middle panel). All cells contained the Pol IIA and II0 forms of the polymerase. Only B95.8 cells showed a faint signal for the CTD-less Pol IIB.

5.4.3   The binding of RNA Pol II at promoters of early lytic genes strongly increased gene transcription

Quantitative RT-PCR experiments were performed in Raji 3862 cells in the time course of lytic induction to analyze the potential of the uncommonly modified RNA Pol II to activate gene expression (Fig. 5.14 A). Raji 3862 cells were induced with 100ng/ml doxycycline for 46 h. RNA was prepared every four hours and treated with DNase I to remove contaminating DNA from the preparation. The success of DNase I treatment was controlled by PCR. RNA was reversely transcribed into cDNA and absolute transcript levels of selected mRNAs were quantified by qPCR. Expression of a set of early and late lytic genes was tested together with the latent gene *EBNA1* and the housekeeping gene *cytochrome C (cytC)*. Transcript levels were calculated on the basis of a single cell.

Fig. 5.14 Kinetics of viral gene expression after induction of EBV's lytic phase

(A) Flowchart of the experimental procedure. Raji 3862 cells were induced with 100ng/ml doxycycline for 46 h. A fraction of the cells was harvested every four hours, RNA was prepared and DNase I treated to el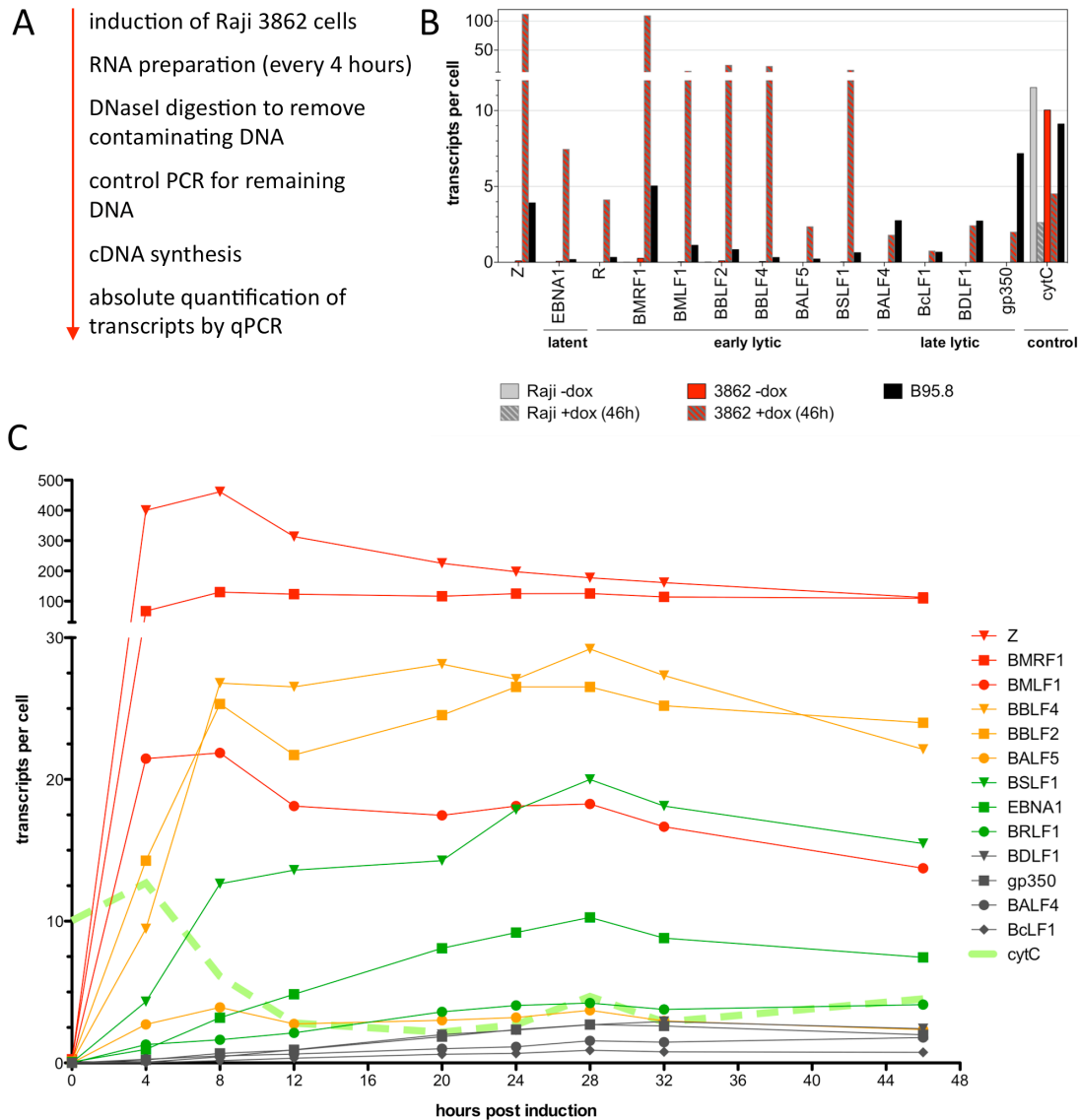iminate contaminating DNA. DNase treatment was controlled by PCR. RNA was transcribed into cDNA and analyzed in a Roche LightCycler 480 real-time PCR system. Expression levels of selected transcripts were determined and calculated at the level of single cells.

(B) In a control experiment, different latent and lytic cells (parental Raji cells, Raji 3862 cells ± doxycycline induction, B95.8 cells) were analyzed for the expression of latent, early lytic, and late lytic EBV genes as well as *cytochrome C* as a housekeeping gene. Results are displayed as transcript levels per cell on the y-axis. Parental Raji cells and uninduced Raji 3862 cells did not show expression of lytic genes except *BMRF1* in uninduced Raji 3862 cells. Expression of *EBNA1*, which is a paradigmatic latent gene, is extremely low. After induction of the lytic phase with doxycycline for 46 h, Raji 3862 cells show a dramatic increase in the expression of early lytic genes, and a slight increase in the expression of late lytic genes. In the B95.8 cell line, a minor fraction of cells spontaneously undergoes the lytic phase of EBV. B95.8 cells served as a positive control. *Cytochrome C* is affected slightly by the addition of doxycycline, as can be seen in parental Raji cells that had been incubated in medium with doxycycline.

(C) The temporal induction of different EBV genes suggests grouping into four different functional groups of viral genes. Group 1 (red) responds within four hours to the addition of doxycycline and encompasses *BZLF1 (Z)*, *BMRF1,* and *BMLF1*. Group 2 (orange) consists of *BBLF4, BBLF2*, *BSLF1,* and *BALF5*. They reached the expression peak eight hours post induction. *EBNA1* and *BRLF1* levels increased slowly over time (group 3, green) peaking at 28 hours after induction, while the late lytic genes (group 4, grey) showed a very low expression upon lytic induction with kinetics comparable to group 3.

In a control experiment, transcript levels of different latent Raji cell derivates (Raji cells with or without doxycycline treatment, Raji 3862 cells without doxycycline treatment) and lytically induced Raji cells (Raji 3862 cells with doxycycline treatment for 46 h), as well as B95.8 cells, which contain a fraction of virus-producing cells, were determined (Fig. 5.14 B). As expected, lytic genes were not detectably expressed in latent cells. Only uninduced Raji 3862 cells displayed very low levels of *BMRF1* expression, which was probably caused by a certain leakiness of the *tet*-promoter. Lytic induction increased the levels of certain early lytic genes per cell, but late lytic genes were expressed at low levels, only. A minor fraction of B95.8 cells spontaneously undergoes lytic reactivation, which is reflected in the expression of early and late lytic genes. The expression of the housekeeping gene *cytC* varied under doxycycline treatment in Raji cells and in Raji 3862 cells.

The induction of lytic gene expression was studied in a time course experiment in Raji 3862 cells after induction with doxycycline (Fig. 5.14 C). The kinetics of induction differed among the early lytic genes. The expression levels of the transgene *BZLF1* peaked after four hours of doxycycline induction. Expression of *BMRF1* and *BMLF1* was equally fast. The *BBLF4*, *BBLF2,* and *BALF5* genes were maximally expressed eight hours post induction. *BSLF1*, *EBNA1,* and *BRLF1* levels continuously increased for a time period of 28 h of induction. Late lytic genes were expressed at very low levels, with a kinetic similar to *BSLF1*, *EBNA1,* and *BRLF1*. The results of the analysis indicated that induction of early lytic gene expression was possible with the unconventionally modified RNA polymerase. The time course revealed different kinetics of activation indicating that some genes are direct targets of BZLF1, while other genes are probably induced by a combination of transcription factors or are secondary targets of BZLF1. Expression of late lytic genes was detectable but at very low levels. This observation indicated, that the expression of late lytic genes is in principle possible without lytic DNA replication. Probably the very low expression levels are not adequate for the synthesis of sufficient amounts of viral structural proteins to form capsids and to release progeny virus.

6. DISCUSSION

Textbooks and most of temporary literature state that EBV's life cycle is biphasic, consisting of an initial latent and a second lytic phase. Accordingly, EBV-encoded genes are classified either as latent or lytic genes. Latent genes and transcripts comprise the EBV-encoded nuclear antigens (*EBNAs*), the latent membrane proteins (*LMPs*), the two non-coding RNAs *EBER1* and *EBER2*, and 44 EBV-encoded microRNAs (miRNAs). Lytic genes are further subdivided into immediate early, early, and late lytic genes. Immediate early genes are defined as genes that are expressed upon lytic induction prior to the translation of new proteins. *BZLF1* and *BRLF1* are classified as immediate early genes. They encode EBV's own transcription factors. Early genes are direct targets of BZLF1, BRLF1, or both. They code mainly for proteins that are essential for lytic viral DNA replication. Late lytic genes are classically defined as genes that are expressed after the onset of lytic DNA replication. Agents that inhibit lytic replication also inhibit late gene expression, indicating that DNA replication in *cis* might be required for their synthesis. This model describes EBV's life cycle to follow the general concepts of herpes simplex virus replication (Roizman et al., 2007).

In the list of previous results of others and our group, the model of a biphasic life cycle appears to be oversimplified in the case of EBV. The most apparent conflict with the commonly accepted concept is the expression of a set of lytic genes immediately after infection of primary B cells (Zeidler et al., 1997; Altmann et al., 2005; Wen et al., 2007; Kalla et al., 2010). But lytic gene expression during this so-called "abortive lytic phase" fails to give rise to viral progeny. Another controversy is the immediate early status of *BRLF1* because its expression is induced by the transcription factor BZLF1 demonstrating that *de novo* protein

synthesis of BZLF1 is mandatory for *BRLF1* expression (Flemington et al., 1991; Amon et al., 2004). Late lytic genes are likewise a matter of debate. Previous reports demonstrated that BRLF1 is able to induce the expression of *gp350* without viral replication (Feederle et al., 2000). Other reports claim that late gene expression of *BcLF1* and *BFRF3* does not require DNA replication in *cis* (Serio et al., 1997). Together, these findings indicate that the conventional classification of herpesviral genes and life phases is insufficient to describe EBV's infectious cycle.

I propose here another concept of EBV's life cycle, in which genes and life phases are classified according to the epigenetic status of the respective promoter and the dependence on viral cofactors. Characteristics of each promoter class and their temporal activation during the infectious cycle of EBV are summarized in Tab. 6.1.

Tab. 6.1 classification of EBV promoters
Promoters are clustered according to the epigenetic state of their template and the requirement for cofactors. The table shows examples for each promoter class, the favored chromatin template, the requirement of a viral cofactor, and the activity during EBV's life cycle.

examples	class	template	viral factor	active during
BZLF1, BRLF1, BALF1, BHRF1, BNLF2a, BCRF1	*default-on* (former: latent, early lytic & late lytic promoters)	naïve chromatin	no	abortive lytic & lytic phase after DNA replication
Cp, Qp, gp350/220	*poised-on* (former: a fraction of latent & late lytic promoters)	naïve chromatin	yes	latent & lytic phase after DNA replication
BBLF4, BMRF1, BALF5 (meZRE promoters)	*default-off* (former: a fraction of latent promoters)	CpG-methylated	yes	lytic phase prior to DNA replication
no evidence	-	CpG-methylated	no	-

Promoters that are operative during the early phase of infection are termed *default-on* promoters. *Default-on* promoters are free of repressive chromatin modifications because incoming virus is epigenetically "naïve". *Default-on* promoters do not need any additional viral factor for their activation, but encounter a cellular environment, in which the basal cellular transcription machinery binds and activates their promoter elements. *Default-on* promoters include the *W* promoter that initiates the early expression of latent genes, the promoters of the two EBV-encoded transcription factors *BZLF1* and *BRLF1*, and the promoters of *BALF1, BHRF1, BNLF2a,* and *BCRF1*. It is unclear whether the promoters need any additional component such as transcription and enhancer factors provided by the cell for their activation. The efficiency of transcription probably varies. I hypothesize that the multiplicity of the *W* promoter (*Wp*), which is part of a repeat array in EBV's genome, could reflect a rather inefficient binding of the transcription machinery. The establishment of several *Wp* copies would compensate a weak affinity of RNA Pol II. This initial phase, in

which *default-on* promoters are active, comes to an end when the viral DNA acquires components of cellular chromatin. Reactivation of silenced *default-on* promoters during lytic replication probably requires DNA replication in *cis* to obtain an epigenetically unmodified template again.

The second class of EBV promoters persists in a poised-to-be-on modus and the promoters are accordingly termed *poised-on* promoters. The chromatin state of them tolerates expression of genes in principle, but a viral factor that is necessary for their activation is missing early after infection. *Poised-on* promoters encompass for example the *C* promoter and the *Q* promoter. It is well established, that the *C* promoter is activated by EBNA1, a gene product of the *default-on* promoter *Wp*. But also the promoters of genes encoding for EBV's structural proteins belong to this class. They rely on the expression of a so far unknown, probably viral factor, which is provided upon lytic DNA replication. *Poised-on* promoters that are not actively hold in an open configuration acquire repressive chromatin marks to be stalled in an inactive state later during infection similar to *default-on* promoters. Silenced *poised-on* promoters need DNA replication in *cis* for efficient gene activation in order to provide an epigenetically unmodified template.

The third and most interesting class of EBV promoters belongs to the *default-off* class of promoters. They encompass BZLF1-regulated, methylation-dependent promoters, like the *BBLF4* promoter. The regulation of *default-off* promoters has long been enigmatic, because the early and *default-on* expression of *BZLF1* was insufficient to induce gene transcription. The observation that meZREs[*] are only bound and transactivated by BZLF1 if they are methylated in their central CpG dinucleotide solved this paradox: *Default-off* promoters are simply not in the correct epigenetic state at the time of viral infection. Epigenetic modifications are not that early and thus prevent an early expression of the meZRE-regulated genes. Moreover, BZLF1's *default-on* promoter is shut down upon chromatinization, preventing the expression of this molecular switch gene when latency is established. Once the signaling cascades that lead to *BZLF1*`s reactivation are induced, BZLF1 protein rescues and activates *default-off* promoters from their repressed state. Until now, there is no evidence for a fourth promoter class, which depends on CpG-methylation but does not require a viral factor to become active.

Fig. 6.1 summarizes the characteristics of each promoter class and their temporal activation during EBV's infectious cycle graphically.

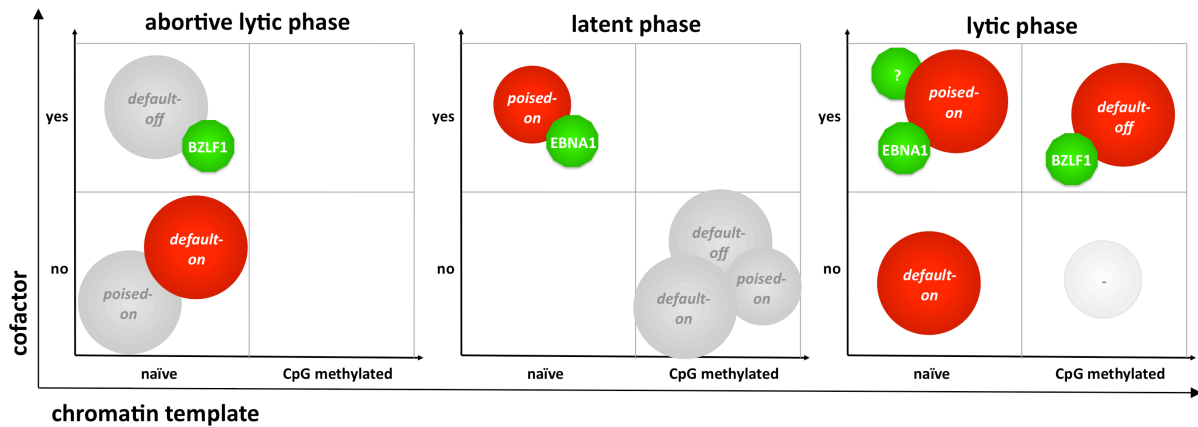[*] methylated BZLF1 responsive elements

Fig. 6.1 EBV promoters can be classified into different subgroups dependent on the epigenetic state of the template and the presence of cofactors

Upon infection of primary B cells, the viral chromatin is in a naïve configuration. The transcription of *default-on* promoters is possible, because they do not depend on the presence of a cofactor. BZLF1 as a *default-on* gene is expressed during the abortive lytic phase but fails to induce *default-off* promoters because they are not CpG-methylated (left panel). Upon chromatinization, *default-on* promoters become silenced. A fraction of *poised-on* promoters is provided with the necessary cofactor, which is the EBNA1 protein, and is kept in an epigenetically unmodified and active state during latency. *Poised-on* promoters that are not provided with the essential cofactor, *default-on* promoters, and *default-off* promoters are epigenetically silenced and hold inactive during latency (middle panel). CpG-methylated DNA is the correct chromatin template for the activation of *default-off* promoters. Upon BZLF1 expression, they become active and express genes that are essential for lytic viral DNA replication. In the onset of the lytic phase, viral DNA replication in *cis* gives rise to epigenetically unmodified, naïve templates again, which allows the re-initiation of transcription at *default-on* promoters. Presumably, a so far unknown viral cofactor is also provided, which permits the expression of the full set of *poised-on* genes (right panel).

In the next chapter I want to integrate established concepts of epigenetic regulation in EBV into my proposed new concept of viral regulation. I want to recapitulate the results of my thesis in light of this model. Last, I want to discuss open questions and possible means to address them.

6.1   State of the art

Promoter selection during the establishment of viral latency is one key process in EBV's life cycle that depends on epigenetic regulation. Upon infection of primary B cells, expression of *EBNA1*, *EBNA2*, and *EBNA3A*-C is initiated at *Wp* (Woisetschlaeger et al., 1990). *Wp* is a *default-on* promoter and free of repressive epigenetic modifications at that stage of infection. Therefore, *Wp* is accessible for the basal transcription machinery. As mentioned above, *Wp* is part of a repetitive unit and present in multiple copies, which could indicate its rather weak affinity for RNA polymerase II. How *Wp* is repressed about four days *post* infection is not completely understood and discussed controversially. DNA methylation was proposed to be responsible for the promoter shutdown (Tierney et al., 2000), but studies with low-passage

lymphoblastoid cell lines (LCL) showed a variable methylation of *Wp*, only (Elliott et al., 2004). Methylation increased with higher passage number. Therefore, it was proposed that alternate repressive mechanisms proceed and DNA methylation of *Wp* is a secondary event (Park et al., 2007). Initially, *EBNA1* (among other *EBNA* genes) is expressed from the *Wp* promoter. The *EBNA1* gene product binds to and activates the second promoter of *EBNAs*, *Cp* (Altmann et al., 2006). *Cp* is a *poised-on* promoter and, once its viral cofactor is present, induces in cell culture expression of all *EBNA* genes, which represents the so-called "latency III" gene expression profile. *In vivo*, the latency III program triggers an immune response in healthy individuals, which in turn requires downregulation of *Cp* by DNA methylation and a second promoter switch to *Qp* (Schaefer et al., 1995; Paulson et al., 1999). This scenario demonstrates that the epigenetically open condition of *poised-on* genes is vulnerable for repressive epigenetic changes as well. Transcription initiation from *Qp* results in the expression of *EBNA1,* only, and is described as the "latency I" program. *Qp* is the only promoter that maintains an active chromatin configuration *in vivo,* which is probably supported by the binding of the chromatin insulator CTCF that prevents the spread of repressive epigenetic modifications into the promoter region of *Qp* (Tempera et al., 2010). This observation suggests that *poised-on* promoters do not only depend on additional factors in *trans* but also have to be actively protected from repression by epigenetic modifications.

Another key checkpoint during EBV's life cycle depends on epigenetic mechanisms as well: the switch to the lytic phase is controlled by DNA methylation. BZLF1 is the master regulator of EBV's lytic switch in latently infected B cells (Countryman et al., 1987). The viral transcription factor has the ability to bind to methylated DNA (Bhende et al., 2004; Karlsson et al., 2008; Dickerson et al., 2009). We could previously show that CpG-methylation of promoters of early lytic genes is a prerequisite to enable the switch to the lytic phase. Binding of EBV's transcription factor BZLF1 to the methylated promoter elements of the class of early lytic genes abrogates repression, induces the expression of lytic genes, and promotes the production of progeny virus (Bergbauer et al., 2010; Kalla et al., 2010). This scenario illustrates the characteristics of the *default-off* state: Virion DNA with *default-off* promoters reaches the cell in an epigenetically unsuitable configuration because *default-off* promoters are initially unmethylated. They first have to be core DNA-methylated before the viral factor, i.e. the transcription factor BZLF1, is able to bind to and activate this class of promoters.

6.2 Scope and aim of my thesis work

Epigenetic mechanisms play a remarkable role regulating EBV's life cycle. But the picture is still incomplete because we have only snapshots of this viral life cycle. In my thesis work I wanted to understand the underlying principles of epigenetic modifications that are necessary to guarantee an efficient infection, successful transformation of B cells, long-term persistence in the host cell, and escape from latency to spread the progeny virus. The two major questions at the beginning of my thesis work were:

1. How does the virus establish and maintain latency?

2. How does the virus escape from latency?

We know now that the viral life cycle consists of three stages: an initial lytic but abortive phase, a latent phase, and lastly, a complete lytic phase, which supports viral synthesis. In the perspective of this model of EBV's life cycle, I wanted to unravel the epigenetic mechanism that lead to the repression of *default-on* promoters and the concomitant termination of the initial abortive lytic phase. As mentioned above, DNA methylation and its role in the switch from *Wp* to *Cp* is controversially discussed. I wanted to address this uncertainty. The epigenetic state of latent promoters, which are according to my model a fraction of *poised-on* promoters, is well understood. The methylation profile and the histone modification pattern of certain promoters are known (Minarovits, 2006). But it is unclear how *default-on* promoters, *default-off* promoters, and *poised-on* promoters except *Cp* and *Qp* are kept in a silent modus. The virus has to cope with two problems: (i) the promoters require tight repression to evade an immune response during latency, but (ii) activation has to proceed rapidly and synchronously after induction of the lytic phase. It was uncertain, how *default-off* promoters can become activated. BZLF1 can bind to and activate methylated *default-off* promoters, but their methylation state was unclear. Whether these promoters are equipped with other repressive epigenetic modifications that need to be overcome upon lytic induction was unknown.

I wanted to elucidate these open issues. I analyzed the early phase of infection and focused on the establishment of epigenetic marks on EBV's DNA. Furthermore, I studied the epigenetic state during latency to understand the repression of lytic genes in latently infected B cells. At last, I investigated epigenetic changes on EBV's DNA upon lytic induction to understand how

EBV is able to escape epigenetic repression. I used different model systems to portray epigenetic events in EBV: Primary infected B cells from adenoids were used for experiments in the early phase of an EBV infection. The cell line Raji is our model system for a latency III infected cell line. I analyzed two different derivates of Raji cells in order to reproduce epigenetic events that occur upon lytic induction: one Raji cell line constitutively expressed a truncated form of BZLF1, which lacks the N-terminal activation domain. A second Raji cell line was engineered to carry stably a doxycycline-inducible plasmid encoding full length (wildtype) BZLF1. My analysis of the different states of an EBV infection together with published data provided information that led to novel insights and improved our model about EBV's life cycle.

6.3   How to establish and maintain latency?

6.3.1   The temporal establishment of an epigenetic pattern on EBV DNA

The initial abortive lytic phase lasts approximately four days, until the transcription from *default-on* promoters is downregulated (Altmann et al., 2006; Kalla et al., 2010). The mechanisms of promoter shutdown were unclear but hardly discussed (Tierney et al., 2000; Elliott et al., 2004; Park et al., 2007). I performed experiments with freshly infected B cells from adenoids to obtain information about early epigenetic events during an EBV infection. I concentrated on the kinetics of DNA methylation and the kinetics of nucleosome acquisition during an EBV infection. I wanted to clarify, whether DNA methylation could play a role in the termination of the abortive lytic phase and the downregulation of *default-on* promoters.

CpG-methylation of EBV's DNA was assessed by MeDIP-on-Chip analysis on primary infected B cells at different time points *post* infection (pi). The DNA methylation profile of EBV showed weak DNA methylation at one week pi, with the most substantial increase between week two and three pi. The methylation profile appeared to be stable thereafter and MeDIP experiments with LCLs originating from different donors had similar profiles. These observations indicated the existence of a characteristic LCL methylation pattern that evolves slowly after infection of primary B cells. Cellular methyltransferases were found upregulated one week pi, which is in line with the time point of increase in EBV's DNA methylation over time (Fig. 4.3). Similar experiments with EBV's close relative KSHV[*] also indicated a slow increase in the viral-specific methylation pattern upon infection (Gunther et al., 2010). These

---

[*] Kaposi Sarcoma-associated herpesvirus

observations exclude the possibility that EBV's DNA methylation could be a mechanism for the promoter shutdown of *default-on* promoters, because the temporal course of both events does not coincide and CpG-methylation traits behind.

The packaging of EBV's DNA into nucleosomes could be an alternate mean to establish an epigenetically repressed gene expression profile in LCLs and to downregulate *default-on* gene expression. I analyzed the occupancy of EBV's DNA with *oriP* nucleosomes in the time course of an EBV infection using a qPCR approach. The occupancy of EBV's DNA with nucleosomes appeared to be a fast and gradual process, which starts immediately after infection (Fig. 4.6). The first discovery of nucleosome-incorporated EBV DNA matched the timepoint at which *default-on* promoters are shut down approximately three to four days pi. *Default-on* promoters include the lytic gene promoters of *BALF1, BHRF1, BZLF1, BRLF1, BNLF2a, BCRF1,* and *Wp* (Zeidler et al., 1997; Altmann et al., 2005; Kalla et al., 2010). *BHRF1* and *BALF1* are two homologs of the cellular anti-apoptotic gene *BCL-2*. BZLF1 has a pro-proliferative effect on resting primary B cells, similar to its cousin AP-1. BNLF2a is targeted to the TAP-transporter[*], which impairs HLA class I-restricted antigen presentation (Horst et al., 2009). *BNLF2a* is a classic "late lytic" gene and is expressed already eight hours pi. In addition, its mRNA is incorporated into virions and transferred to newly infected cells (Simon Jochum, personal communication), which reflects its important role during the abortive lytic phase. BCRF1 is a homolog to cellular IL-10 and also involved in the suppression of innate host immune responses. The anti-apoptotic, pro-proliferative, and immunosuppressive effects that the viral proteins provide to the cells initially are pivotal for the success of viral infection and the establishment of the virus in the host cell. The function of *default-on* genes is substantial in the very early phase of infection, but a rapid downregulation is crucial for a persistent infection. The prevalence of BZLF1 and BRLF1 specific CD8[+] T-cells very early after infection illustrates the need to downregulate *default-on* promoters in order to escape an early immune response prior to the establishment of latency (Iskra, 2010). The rapid acquisition of nucleosomes establishes a defined chromatin structure on EBV's DNA, which terminates the activity of *default-on* promoters. Nucleosomes are not only barriers to transcription: Histone proteins are also important for a multitude of other cellular tasks. For example, packaging of DNA into nucleosomes is essential during the cell cycle for a proper DNA replication during S-phase and successful chromosome segregation during mitosis (Lipford et al., 2001). It is essential for EBV to adapt to these cellular mechanisms to guarantee the heredity transmission of its genomes in dividing cells. The fast

---

[*] transporter associated with antigen processing

occupancy of EBV's DNA with nucleosomes seemingly reflects the importance of an early establishment of the viral episome in infected cells, too.

6.3.2   Epigenetic modifications help to maintain the latent state

Herpesviral latency is associated with genomic persistence and a highly restricted viral gene expression in the host cell (Speck et al., 2010). The initiation of latent gene expression from different *poised-on* promoter types has been studied extensively (Minarovits, 2006). As pointed out before, EBV acquires chromatin modifications during its infectious cycle. Nucleosomes are positioned very early while DNA methylation is a slow process. However, according to my results, the chromatin state that allows the induction of *default-off* genes is set after two to three weeks pi, which matches the timepoint of the first detectable virus in LCLs (Kalla et al., 2010). I could also show that latent Raji genomes are highly methylated (Fig. 4.2), but *default-off* genes are not expressed in latent cells until *BZLF1* is expressed. The mode of repression of *default-off* genes during latency has not been studied so far. Therefore, I wanted to determine the state of latent genomes in order to understand the epigenetic quality of the repression of *default-off* genes during latency in our model cell system Raji. I assessed the DNA methylation pattern, the nucleosome occupancy profile, and a set of important histone modifications to understand the epigenetic nature of latency. The results I obtained delivered the following picture:

Qp was the only *poised-on* EBNA promoter that allowed latent gene expression in Raji cells. The promoter region was entirely free of methylation and the nucleosome occupancy was low. The repressive histone modifications H3K9me3 and H3K27me3 were absent, but the promoter encompassed significant levels of the activation mark H3K4me3. *EBERs* and *oriP* comprised a similar epigenetic pattern indicating an active state during latency. Furthermore, my experiments demonstrated a strict repression of *default-on*, *default-off*, and *poised-on* promoters except *Qp* during latency. All three promoter classes were highly methylated during latency. Importantly, *default-off* promoters encompassed a strong invariable CpG-methylation of meZREs. The strong methylation of meZREs provides the *default-off* promoters with an epigenetic background that is permissive for transcriptional activation by BZLF1. But unless BZLF1 is expressed, DNA methylation represses transcription according to the "textbook" knowledge, which links silencing and DNA methylation.

I could show that DNA methylation is not the only mean to establish a tight repression during latency. The promoters encompassed densely positioned nucleosomes, a finding that was typical for BZLF1-regulated promoters (*default-off* promoters). More interestingly, MND-on-

Chip experiments indicated a hierarchy of gene repression during latency. Promoters of genes, which were shown to be essential for lytic DNA replication (Fixman et al., 1992; Feederle et al., 2000), comprised the highest nucleosome occupancy indicating the strictest repression. A hierarchy of repression appears plausible: If proteins indispensable for lytic replication are absent due to a strong gene repression during latency, the whole signaling cascade of lytic reactivation cannot come to pass. Infrequent and accidental expression of genes that are not required for lytic DNA replication can probably be tolerated during latency because protein products will not trigger the activation of further promoters that could cause an immune response. *Default-off*, *default-on*, and *poised-on* promoters have additional means of repression during a stable latency: I could demonstrate that PcG[*]-proteins bind and seal repressed promoters with the repressive mark H3K27me3 to stably maintain repression during cell cycle progression. During latency, activation marks were absent among *default-off*, *default-on*, and repressed *poised-on* promoters. RNA Pol II was oblivious, but it is questionable whether a polymerase would ever be able to bind to promoter elements that are padded tightly with nucleosomes, as it is the case for *default-off* promoters (Lorch et al., 1987). My observations argue against the existence of a so-called bivalent chromatin structure in the EBV genome. Bivalent marks were described to be important for lytic reactivation of KSHV, a close relative of EBV (Gunther et al., 2010; Toth et al., 2010). Sequential ChIP assays indicated the co-existence of H3K4me3 and H3K27me3 at the immediate early promoters of ORF50 and ORF48. The concomitant presence of these modifications is a hallmark of bivalent chromatin (Bernstein et al., 2006). ORF50 encodes the BRLF1 homolog RTA. RTA is sufficient to induce the complete cascade of KSHV's lytic replication. Furthermore, H3K27me3 and the histone H3K27 methyltransferase EZH2[†] were detected in the entire KSHV genome. The third characteristic for a bivalent chromatin structure, i.e. a poised polymerase, was not analyzed in the KSHV background. I analyzed all features of bivalent chromatin in EBV but I could not detect any colocalization of H3K4me3 and H3K27me3. A polymerase was not present at *default-off* promoters during latency. My experiments apparently demonstrated that bivalent chromatin is not present in EBV and pointed to a different mode of gene regulation in EBV as compared to KSHV.

---

[*] proteins of the Polycomb group
[†] human enhancer of zeste 2

6.4   How to escape the latent state?

Exogenous signals such as antigen contact or unknown stochastic events induce the lytic phase, which leads to the expression of the full set of viral genes, lytic reproduction of the virus, and the release of progeny to infect other cells. In EBV, signaling cascades downstream of the B cell receptor most likely trigger induction of the lytic phase (Campbell, 1999). The foremost event of a productive lytic phase is the activation of *default-off* genes. *Default-off* genes encode proteins for lytic DNA replication. DNA replication yields epigenetically unmodified EBV genomes, so that the initial epigenetic "default setting" of the virus is restored. This is a prerequisite for the activation of silent *poised-on* genes, which are necessary for the synthesis of structural proteins and packaging of newly replicated viral DNA into capsids. Activation of *default-on* genes, which were silenced after the establishment of latency by epigenetic repression, follows the same rules. I wanted to gain deeper insights into the mechanism of lytic reactivation with special focus on the activation of *default-off* genes.

6.4.1   Primary infected B cells are ready for lytic reactivation two weeks *pi*

DNA methylation is indispensable for the switch to the lytic phase in EBV. Early expression of BZLF1 fails to give rise to progeny virus, indicating that lytic induction is only possible in latently infected B cells (Kalla et al., 2010). *Default-off* promoters encompass meZREs, which are exclusively bound by BZLF1 when they carry methylated cytosines (Bergbauer et al., 2010). As a consequence, EBV's DNA requires DNA methylation marks to switch to the lytic and productive phase. As I pointed out before, I could show that DNA methylation is predominantly established between week two and three pi in primary infected B cells (Fig. 4.3). In agreement with my observation, progeny virus is not released earlier than two weeks pi (Kalla et al., 2010). These findings demonstrated that the timing of EBV's DNA methylation governs its capacity to switch to the lytic phase.

6.4.2   Chromatin changes at *default-off* promoters after lytic induction

My experiments demonstrated the solid repression of *default-off* promoters during latency. The promoters were not only intensively methylated, but also occupied by nucleosomes (Fig. 5.3 and 5.7). They encompassed repressive histone tail modifications and binding of PcG-proteins promoted long-term silencing (Fig. 5.10 and 5.11). It was previously shown that BZLF1 is able to bind to methylated DNA (Bergbauer et al., 2010), but the role of the

alternate repressive epigenetic marks was unclear. A major question remained. Does activation of the lytic phase include a remodeling of EBV's chromatin or is the binding of BZLF1 to methylated DNA sufficient to recruit the transcription machinery to *default-off* promoters?

The methylation state of *default-off* promoters remained unchanged upon lytic reactivation demonstrating that DNA methylation does not prevent transcriptional activation *per se* (Fig. 5.5). In contrast, other repressive epigenetic marks were strongly affected by the induction of the lytic phase. I could show that binding of BZLF1 to promoters of a subset of *default-off* genes resulted in major rearrangements of nucleosomes. The high nucleosome occupancy that was present in latency was lost, indicating a local opening of the promoter to permit access for the transcription machinery. Remodeling was mainly independent of the N-terminal activation domain of BZLF1, as a truncated form that lacks the N-terminus was also able to remodel EBV's chromatin at most BZLF1-responsive promoters. The N-terminal transactivation domain of BZLF1 constituted to the loss of nucleosomes at certain promoters like *BMRF1*, only. I hypothesize that BZLF1 interacts with different chromatin remodelers in order to cause a local release of DNA in promoters of *default-off* genes. Previous studies already identified the histone acetylase CBP[*] as an interaction partner of BZLF1. CBP induces chromatin rearrangements and mediates RNA Pol II recruitment to promoter sites of DNA. Interaction of BZLF1 and CBP required both the C-terminal domain of BZLF1 and the N-terminal activation domain of BZLF1 (Adamson et al., 1999; Zerby et al., 1999). However, I already obtained chromatin remodeling at certain BZLF1-regulated promoters with a BZLF1-mutant that lacked the N-terminal activation domain, which was reported to be detrimental for CBP interaction (Adamson et al., 1999). A closer look at the protein structure of BZLF1 might be helpful to understand why the mutant-BZLF1 is still able to remodel chromatin. BZLF1 is a bZIP (basic leucine zipper) transcription factor. The protein has a length of 245 amino acids and consist of several protein domains: The N-terminal activation domain, a regulatory domain, a central DNA binding domain, a coiled-coil dimerization domain, and a C-terminal tail of 23 amino acids, whose function is unclear and debated. The small C-terminal tail was proposed to be important for stabilizing the coiled-coil dimer. The proximal part of the C-terminal domain was also assumed to be required for promoter transactivation (Hicks et al., 2003; Schelcher et al., 2007). A more recent study denies the impact of the C-terminal domain for transactivation, and instead supposes a role for BZLF1-mediated DNA replication (McDonald et al., 2009). The truncated version of BZLF1 I used in

---

[*] CREB binding protein

my experiments comprised this C-terminal tail of unknown function, which might be important for BZLF1-dependent chromatin remodeling upon lytic reactivation of the virus. Probably the domain provides interaction surfaces for so-far unidentified chromatin remodelers.

The differences in the two modes of chromatin remodeling at the two subsets of promoters matched the hierarchy, i.e. the extend of repression I mentioned in the last chapter: Genes that are central for lytic viral DNA amplification displayed the tightest repression by nucleosomes during latency. The N-terminal domain of BZLF1 comprising the C-terminal tail was not sufficient and the activation domain was essential to remodel nucleosomal occupancy in the majority of promoters of this subset indicating that they require the most energy to be activated for gene expression. Similar to my findings, a previous study on genes activated by Toll-like receptors in macrophages suggested a hierarchical regulation of nucleosome remodeling that depended on the strength of repression in the inactivated state. The authors found a correlation between the nucleosome density, the need for chromatin remodelers, and the function of the corresponding gene: Promoters of genes with highly specialized tasks assembled into stable nucleosomal structures and required remodeling upon induction. A minimal basal transcription of the corresponding genes would be detrimental to the cell. In contrast, promoters of genes that are activated by common transcription factors encompassed only unstable nucleosomes and did not rely on the energy consuming remodeling of their promoter nucleosomes (Ramirez-Carrozzi et al., 2006; Ramirez-Carrozzi et al., 2009). Another hallmark of *default-off* promoters in EBV during latency was trimethylation of histone H3K27 and the local prevalence of the PcG-protein EZH2. Lytic induction caused the rapid loss of H3K27me3 and EZH2. Concomitantly, the H3K4me3 activation mark was established at *default-off* promoters. The epigenetic changes at *default-off* promoters created a chromatin environment that allowed binding of RNA Pol II and an efficient initiation of transcription.

My experiments revealed that *default-off* promoters are accessible for BZLF1-binding during latency but that remodeling of the repressive chromatin structure is necessary for the switch to the lytic phase presumably to provide admission for the basal cellular transcription machinery. These observations clearly demonstrated that binding of BZLF1 rescues *default-off* promoters from an epigenetically repressed state. Interestingly, KSHV uses a different mechanism to escape from latency. As mentioned before, the KSHV protein RTA is sufficient to induce the lytic phase and the promoter region of ORF50, the gene that encodes RTA, encompasses bivalent chromatin structures during latency. In KSHV, inhibition of the

methylase EZH2 or overexpression of the H3K27me3 demethylase was sufficient to reactivate the RTA promoter, but not the BZLF1 promoter in cell lines that carry both viruses (Gunther et al., 2010; Toth et al., 2010). This observation indicates, that epigenetic repression and reactivation of KSHV and EBV has evolved independently. While KSHV makes use of bivalent chromatin for its rapid lytic reactivation, EBV has evolved a crafty mechanism to overcome tight epigenetic repression by its transcription factor BZLF1.

6.4.3   Two *default-off* promoters in a close-up: Epigenetic regulation at *BBLF4* and *BMRF1*

Two examples of *default-off* promoters are the promoter of the EBV-encoded helicase *BBLF4* and the promoter of the DNA-polymerase associated factor *BMRF1*. Both gene products belong to the seven proteins that are indispensable for lytic DNA replication of EBV's genome. Fig. 6.2 and 6.3 graphically portray the results of all epigenetic information that were collected during my thesis, in order to demonstrate the complexity of epigenetic regulation of the promoters during the different phases of EBV's life cycle.

The *BBLF4* promoter encompasses several meZREs (Fig. 6.2, lower y-axis, red and green line). All meZREs were highly methylated during latency (Fig. 6.2, upper panel, lower y-axis, each CpG dinucleotides is represented as a bar, the degree of methylation is indicated in red). The promoter site of *BBLF4* was highly enriched in nucleosomes in contrast to the *BBLF4* gene body (Fig. 6.2, grey area). The modifications of histone tails and the binding of chromatin-associated proteins are depicted as differently colored areas on the upper y-axis. The whole promoter encompassed high levels of the repressive mark H3K27me3 (Fig. 6.2, dark grey area) and the H3K27 methyltransferase EZH2 (Fig. 6.2, green area). Repression by H3K9me3 was present, but on a low level, only (Fig. 6.2, blue area). The activation mark H3K4me3 was absent, similar to RNA Pol II. Induction of the lytic phase caused dramatic changes at the promoter site (Fig. 6.2, lower panel): Nucleosomes in the promoter region were lost; a hypersensitive site was established at the 5' end of the gene. H3K27me3 and EZH2 levels dropped and became indistinguishable to background levels. The activation mark H3K4me3 was established at the promoter site (Fig. 6.2, red area), and was even more prominent towards the gene body of *BBLF4*. RNA Pol II followed the same pattern (Fig. 6.2, orange area). Histone H3K9me3 levels remained unchanged indicating that this modification is not important for the maintenance and for reversion of gene repression.
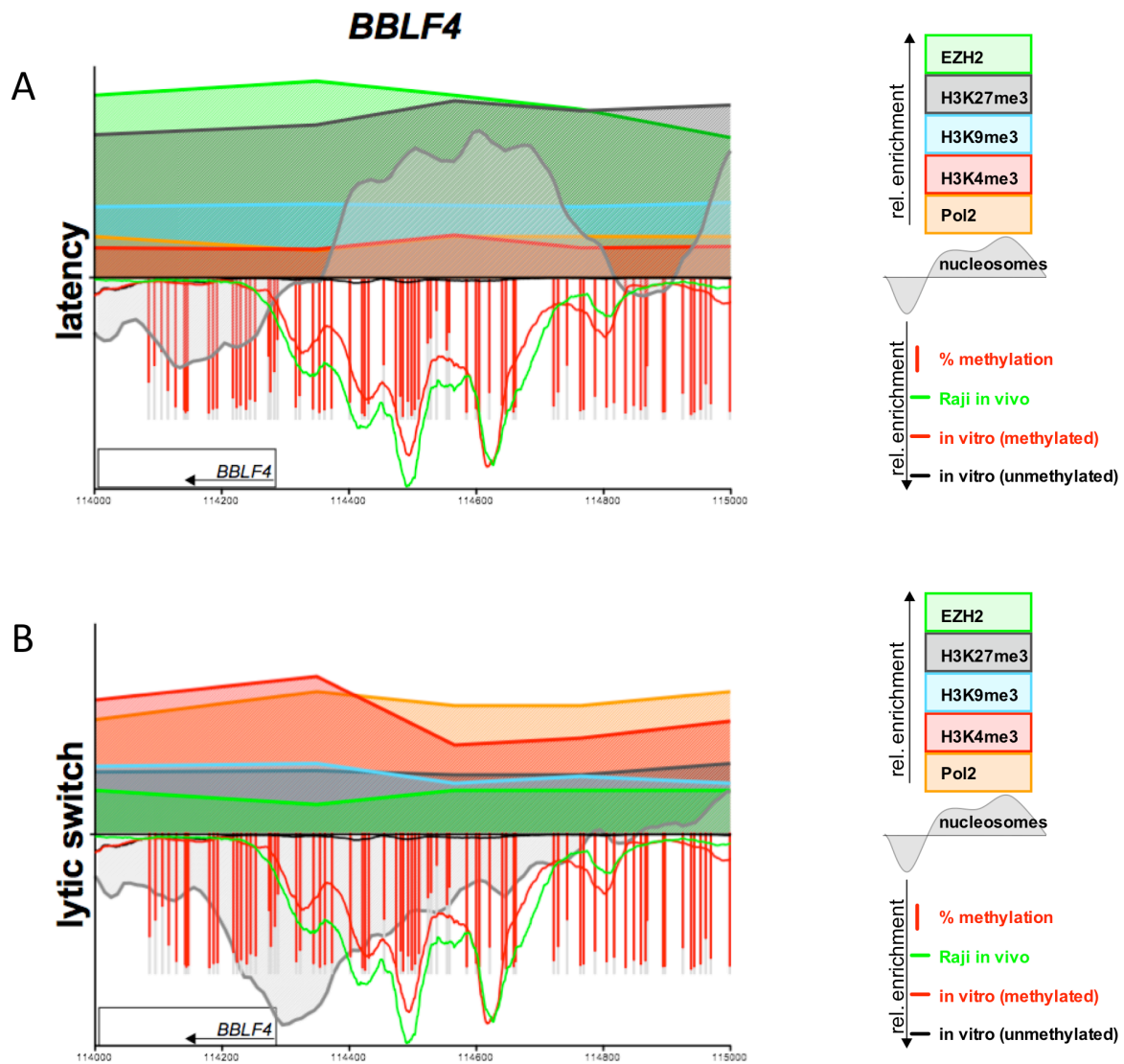
Fig. 6.2 Epigenetic regulation of EBV: Close-up of the *default-off BBLF4* promoter
Summary of all information that I collected during my PhD work to assess the epigenetic state of the early gene promoter of *BBLF4* during latency **(A)** and after lytic induction **(B)**. Displayed are binding sites of BZLF1 (plotted on the lower y-axis, please refer to Fig. 5.1 for details), the methylation state of the promoters evaluated by bisulfite sequencing on the lower y-axis (experimental details are given in Fig. 5.3), the nucleosomal occupancy analyzed by MND-on-Chip analysis (grey area, compare Fig. 5.8), and the results of ChIP experiments of EZH2 (light green area), H3K27me3 (dark green area), H3K9 (blue area), H3K24me3 (red area) and RNA Pol II (orange area) plotted on the upper y-axis (compare Fig. 5.10-12). Results are provided as relative, dimensionless numbers, but the scaling of each experiment is equal in all the graphs shown.
Repressive modifications characterize the *BBLF4* promoter during latency. They are erased after induction of the lytic phase with the exception of CpG-methylation. The ZREs are densely occupied by nucleosomes, which presumably serve to repress this promoter in the latent phase. Repression by Polycomb group proteins maintain the repressed state, as can be seen by the enrichment of the PRC2 protein EZH2 and the histone modification H3K27me3. Nucleosomes and repressive marks are removed after lytic induction. A hypersensitive site becomes established upstream of the ZREs in front of the coding sequence of *BBLF4*. H3K9me3 levels are low but visible during latency and after lytic induction. The activation mark H3K4me3 is absent in latency, but is found at the promoter and at higher levels in the gene body after lytic induction. RNA Pol II parallels the H3K4me3 pattern.
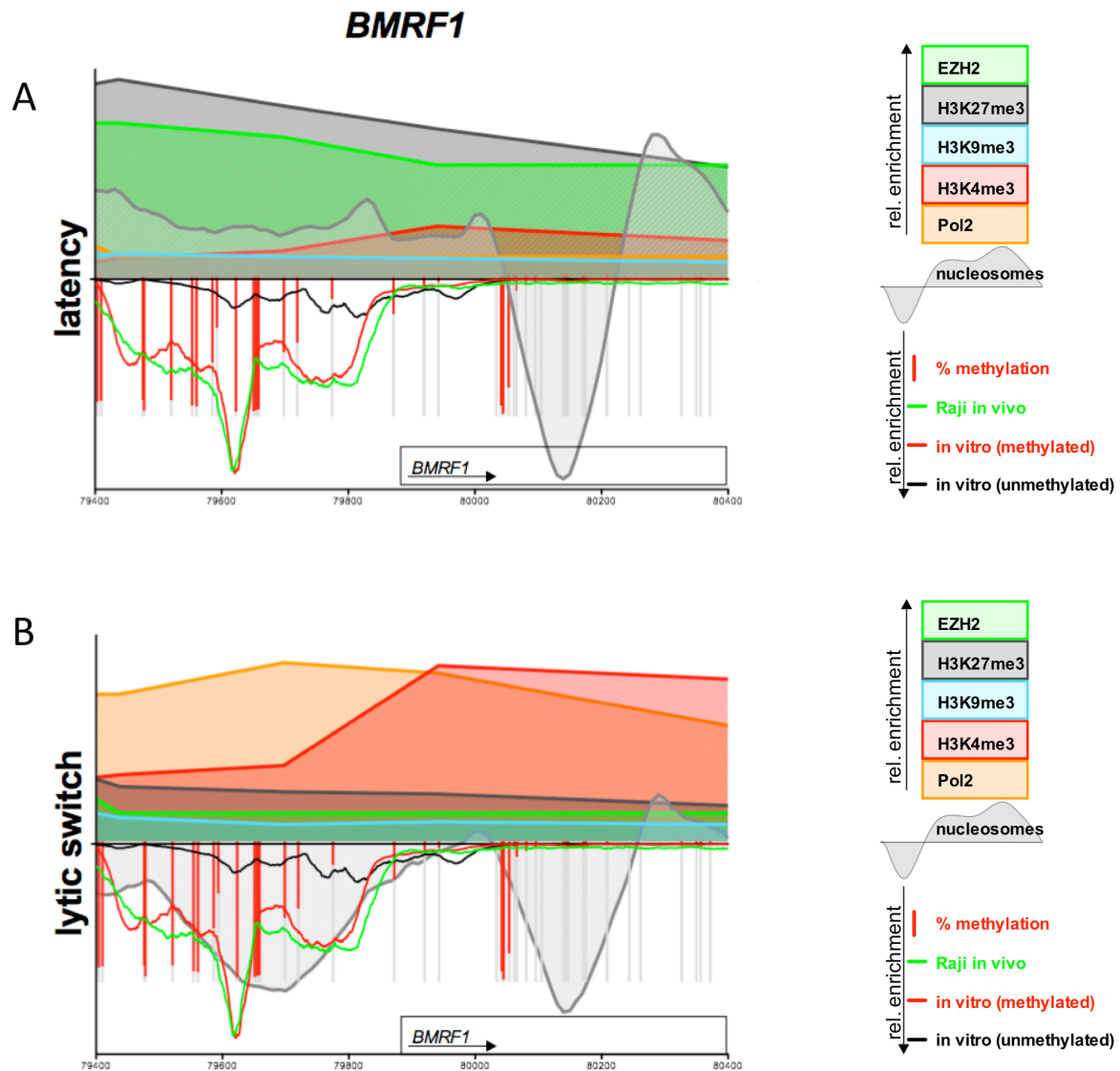
Fig. 6.3 Epigenetic regulation of EBV: Close-up of the *default-off BMRF1* promoter
Summary of all information that I collected during my PhD work to assess the epigenetic state of the early gene promoter of *BMRF1* during latency **(A)** and after lytic induction **(B)**. Figure details can be found in Fig. 6.2.
In principle, the *BMRF1* promoter shows the same epigenetic pattern seen in the *BBLF4* promoter, although repression by nucleosomal occupancy and methylation seems to be less tight, as not all CpG sites are completely methylated and nucleosomes are slightly reduced during latency. However, after the switch to the lytic phase, the mechanisms found in the *BBLF4* promoter also apply here: Repressive marks are removed and activation marks and RNA Pol II are enriched in the lytic phase.

The *BMRF1* promoter exhibited a similar epigenetic regulation. The promoter comprises meZREs (red and green line, lower y-axis), which were highly methylated (red bars, lower y-axis). The nucleosome occupancy was relatively high at the promoter site. The *BMRF1* gene body contained a hypersensitive site, which sharply separated the methylated promoter site from the hypomethylated region downstream within the gene body. The complete region displayed high levels of H3K27me3 (dark grey area) and EZH2 (green area), but only low amounts of H3K9me3 (blue area). Activation marks and RNA Pol II binding were absent during latency. The picture changed dramatically upon lytic induction: The promoter was

completely remodeled; nucleosomes and repressive modifications were evicted and erased, respectively. Instead, the H3K4me3 levels and RNA Pol II occupancy rose (red and orange area, respectively), inducing the transcription of *BMRF1*.

Summing up, the pictures illustrate the dramatic changes that are necessary to induce transcription of *BBLF4* and *BMRF1*, two representatives of *default-off* genes. Clearly, the promoter sites are tightly repressed by epigenetic mechanisms during latency. The chromatin configuration allows binding of BZLF1, but the epigenetic state of the promoter has to be reverted to make it available for RNA Pol II binding and gene expression. This demonstrates that BZLF1 is able to bind to and access a repressed promoter, but its binding is not sufficient to establish RNA Pol II at the promoter site. Hence, after BZLF1 binding, the chromatin has to be remodeled in a "classical" way.

6.4.4 Chromatin changes at *default-on* and *poised-on* promoters upon lytic induction

The epigenetic pattern of *default-on* and *poised-on* promoters during latency pretty much resembles the mode of repression of *default-off* promoters. However, both promoter classes do not encompass binding sites for BZLF1. I wanted to figure out the epigenetic consequences of lytic induction on *default-on* and *poised-on* promoters to dissect the influence of BZLF1 binding on the chromatin state.

Interestingly, PcG-mediated repression was erased at *default-on* and repressed *poised-on* promoters similar to *default-off* promoters: H3K27me3 and EZH2 levels dropped upon lytic reactivation. Repressive modifications appeared to be completely removed from the EBV genome. This finding indicated that BZLF1 binding in *cis* is not a prerequisite for the removal of repressive modifications. However, certain chromatin changes at *default-off* promoters occur locus-specifically: *Default-on* and *poised-on* promoters did not acquire activation marks, and RNA Pol II did not bind to the promoters. The nucleosome levels were still elevated and appeared to inhibit the binding of the transcription machinery. The kinetics of transcriptional activation in lytically induced cells pictured a low-level induction of some *poised-on* genes. Perhaps, erasure of repressive marks led to a certain leakiness of these promoters. This state might be an artifact of the Raji cell line because lytic viral DNA replication does not proceed as the viral *BALF2* gene is deleted in the viral genome. Therefore, activation of *poised-on* promoters cannot be studied in the model cell line Raji.

But comparison of *default-off* and *poised-on* promoters provides important insights about the

contribution of BZLF1 to epigenetic changes: nucleosome remodeling is a direct effect of BZLF1 binding but Polycomb repression disappears independently of BZLF1 binding.

6.5   Open questions and outlook

The life cycle of EBV clearly depends on epigenetic modifications. My results contributed to a clearer understanding of the dynamic epigenetic changes that govern the regulation of EBV's life cycle. Still, many questions remain unanswered and need to be addressed in the future.

6.5.1   Why are certain parts of the EBV genome kept in an open configuration?

This is also a key question in embryonic development and its epigenetic regulation. What defines the trajectory the marble takes on its way down the hill in Waddington's epigenetic landscape (Fig. 1.1)? One theory is the importance of early transcription: In a previous study it was found that transcription initiation and H3K4me3 are hallmarks of promoters of most protein-coding genes in human ES cells. This situation could serve to create a chromatin structure at transcriptional start sites that is more accessible to transcription than the rest of the genome (Guenther et al., 2007). H3K4me3 modification inhibits the binding of DNMT3L protecting these sites from *de novo* CpG-methylation. Thus, active RNA polymerase probably protects CpG islands and promoter sites from DNA methylation (Cedar et al., 2008). The cooperation of early transcription and chromatin insulators like CTCF could be a mechanism to prevent establishing repressive epigenetic modifications on parts of EBV's genome that are active during latency. The solution of this enigma will explain how a stable latency is established and maintained in EBV-infected cells. In addition, EBV could serve as a model system to understand the establishment of a restricted gene expression pattern in metazoan cells.

6.5.2   What is the mechanism of chromatin remodeling upon lytic induction?

Binding of BZLF1 to promoters of *default-off* genes induces major nucleosome rearrangements and causes a local opening of the tightly packed chromatin. My results suggested at least two different means to remodel chromatin upon BZLF1 binding: One mechanism relied on the function of the N-terminal activation domain of BZLF1. The second depended on the C-terminal half of BZLF1, only. I speculated that the C-terminal domain

could contribute to chromatin remodeling, but it is just as well possible that binding of BZLF1 alone is sufficient to induce a local chromatin rearrangement. The mechanisms of chromatin remodeling can be addressed using additional BZLF1 mutant forms that lack the C-terminal domain. Interaction partners of BZLF1 can be identified by immunoprecipitation experiments, probably in conjunction with mass spectrometry. Knowledge about cellular factors that are necessary for promoter activation upon lytic induction will enhance our understanding about epigenetic regulation in EBV.

6.5.3   What is the mechanism of *poised-on* and *default-on* promoter activation upon lytic induction?

Raji cells are deficient in the gene locus *BALF2*, which encodes for the single strand DNA binding protein of EBV's DNA replication machinery (Hung, 1999, Nakayama 2009). *BALF2* is one of the genes that is absolutely essential for lytic DNA replication and the completion of the full lytic cycle because repressed *poised-on* and *default-on* genes probably require the newly replicated, naïve DNA templates for their transcription (Summers et al., 1976; Yajima et al., 1976; Datta et al., 1981). However, there is at the moment no proper model system to elucidate the mechanistic background of *poised-on* and *default-on* promoter activation upon lytic induction. Late events in EBV's lytic phase still remain a black box. As a consequence, indications that newly replicated transcripts are necessary for the transcription of "late lytic genes", i.e. *poised-on* genes, is only based on indirect hints.

The complete lytic cycle can be rescued by the introduction of the *BALF2* open reading frame into Raji cells (Decaussin et al., 1995). A Raji cell derivate, in which the lytic phase can be reproduced, would be a perfect means to analyze events that are necessary for the expression of *poised-on* genes and epigenetically silenced *default-on* genes. Ideally, the engineered Raji cell line should encompass an inducible version of *BZLF1* and *BALF2* in order to produce progeny virus. Whether the *poised-on* transcripts originate from the established EBV episome or from newly synthesized EBV genomes is a still unanswered key question regarding late events of the lytic phase. A way to examine the nature of the template is to measure its methylation state. Chromatin immunoprecipitation of the RNA Pol II-associated template DNA and a subsequent methylation analysis by bisulfite sequencing could be a means to answer this open question.

6.5.4 What is the nature of RNA Polymerase II during lytic replication of EBV?

I performed Chromatin Immunoprecipitations of RNA Pol II with different CTD[*]-modification-specific antibodies but I could only enrich RNA Pol II at *default-off* promoters by using an antibody directed against the N-terminal part of the protein. All antibodies recognizing epitopes of the regulatory CTD of RNA Pol II were unable to precipitate DNA upon lytic induction, but their functionality was demonstrated in positive controls in the same samples. This is a surprising finding. First, we hypothesized an EBV-induced clip-off of the CTD to redirect the cellular transcription machinery to EBV gene transcription, but this hypothesis could not be confirmed in western blot immunodetections: The CTD-less polymerase was not detected in lytically induced Raji cells. However, the virus apparently uses a special and so far unknown mechanism for the transcription of its *default-off* genes because standard CTD-modifications appear to be absent during lytic transcription. The usage of CTD-specific antibodies that are specific for uncommon modifications would be a means to shed light on this issue.

6.5.5 Which concepts of EBV can be transferred to cellular epigenetic mechanisms?

As I pointed out above, EBV could serve as a model system to study the establishment of an epigenetic pattern in metazoan cells. Furthermore, epigenetic regulation in EBV comprises concepts, that are so far not known in the epigenetic regulation of cellular genes: It is questionable whether EBV has invented or adopted a transcription factor that is able to convert an epigenetic repressed state, as it is the case for BZLF1. The same holds true for the unidentified mechanism of RNA Pol II regulation: Peculiar modifications of RNA Pol II's CTD could be a virus specific mechanism to redirect cellular transcription to the benefits of the virus, but it could just as well be a crucial mechanism of transcriptional regulation in other scenarios of stress response, for example in gene regulation induced by UV-irradiation or inflammatory cytokines. Conceivably, the examination of epigenetic events in EBV may contribute to extend our knowledge about epigenetic regulation of human DNA during development and disease.

---

[*] C-terminal domain

7. SUMMARY

Cellular gene regulation depends on fundamental epigenetic mechanisms, but epigenetic modifications also govern the regulation of the life cycle of Epstein-Barr virus (EBV). Promoter usage during latency depends on DNA methylation of the viral genome and CpG-methylation of certain promoters with meZREs[*] is an indispensable prerequisite to switch from the latent to the lytic phase. In my thesis, I wanted to assess the underlying epigenetic principles of EBV's gene regulation during the establishment of latency and upon lytic reactivation. My results suggested a new classification of viral promoters and phases of gene regulation that depend on the epigenetic state of the viral chromatin. According to this new model, EBV's infectious cycle consists of an initial abortive lytic, a latent, and a productive lytic phase, and the viral promoters can be classified into *default-on*, *poised-on*, and *poised-off* promoters.

Default-on promoters are immediately active upon infection of primary B cells leading to the so-called abortive lytic phase of an EBV infection. *Default-on* promoters encounter the cell in an environment that supports binding of the basal transcription machinery to activate viral gene transcription. *Default-on* promoters include the promoters of *BALF1*, *BHRF1*, *BZLF1*, *BRLF1*, *BNLF2a, BCRF1,* and *Wp*. The protein products are indispensible for the permanent establishment of EBV's genome in latently infected cells supporting growth transformation, immune evasion, and anti-apoptotic cellular pathways. *Default-on* promoters are epigenetically silenced upon occupancy of EBV's DNA with nucleosomes very early after infection and a switch to *poised-on* promoters is initiated.

---

[*] methylated BZLF1 responsive element

Poised-on promoters embody an epigenetically active state upon infection, but their activation requires an additional, virus-encoded factor to allow initiation of transcription. *Wp*-induced expression of EBNA1 promotes the switch to the *poised-on* promoter *Cp* to sustain long-term EBNA expression. Other *poised-on* genes including the viral structural proteins are not provided with their cofactor initially. During latency, these promoters are repressed through compaction of chromatin by high nucleosome occupancy, trimethylation of H3K27, and a stable transmission of repressive modifications by Polycomb-mediated long-term silencing. The establishment of a defined DNA methylation pattern on EBV's DNA further represses *poised-on* promoters.

DNA methylation is a prerequisite for the activation of a third promoter class, termed *default-off* promoters. *Default-off* promoters are bound and transactivated by BZLF1 in a methylation-dependent manner. Upon infection of primary B cells, EBV's DNA is completely unmethylated, impeding an early expression of *default-off* genes. Only two to three weeks *post* infection the viral genome has acquired a proper epigenetic configuration that supports transcription of *default-off* genes. Binding of BZLF1 alone does not suffice to recruit the cellular transcription machinery including RNA polymerase II, but the chromatin requires remodeling, including a loss of nucleosomes and repressive modifications at *default-off* promoter sites. *Default-off* genes encode the viral lytic DNA replication machinery. The newly synthesized DNA templates lack epigenetic modifications because lytic DNA amplification is uncoupled from cellular DNA replication, eliminating the epigenetic maintenance mechanisms during the synthesis of viral progeny. As a consequence, *default-on* promoters and silenced *poised-on* promoters, which rely on unmethylated, epigenetically naïve templates, become also activated in the onset of the lytic phase. Silenced *poised-on* promoters require additionally a viral cofactor for their activation. This so-far unknown factor is probably provided upon lytic DNA amplification allowing the transcription of genes encoding for structural proteins that are necessary for the packaging of viral progeny. The released EBV progeny is epigenetically unmodified and ready to infect other cells.

In essence, the regulation of EBV's life cycle by epigenetic mechanisms is a paradigm for viral coevolution with its host. Repressive epigenetic modifications are common cellular defense mechanism to fight invading pathogens. EBV has hijacked this system for the regulation of promoter usage during its own life cycle, which has become a key principle of EBV's success in infecting and persisting in its host.

8. ABBREVIATIONS

5mC	5-methylcytosine
A	adenosine
AA	amino acid
AD	activation domain
AP-1	activator protein 1
ATP	adenosine-5'-triphosphate
BCR	B cell receptor
BL	Burkitt's lymphoma
bp	base pair
BSA	bovine serum albumine
bZIP	basic leucine zipper
C	cytosine
C-terminal	carboxy-terminal
CD	cluster of differentiation
cDNA	complementary DNA
CENPA	centromer protein A
ChIP	Chromatin Immunoprecipitation
ChIP seq	Chromatin Immunoprecipitation with subsequent genome-wide sequencing
CMV	cytomegalovirus
CpG	cytosine-phosphatidyl-guanosine
CRE	cAMP responsive element
CTCF	CCCTC-binding factor
CTD	carboxy-terminal domain
cytC	cytochrome C
DBD	DNA binding domain
DD	DNA dimerization domain
DMSO	dimethylsulfoxid
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
DNMT	DNA methyltransferase
dsDNA	double stranded DNA
DTT	dithiothreitol
E	early
E.coli	Escherichia coli
e.g.	exempli gratia

EBERs	EBV-encoded RNAs
EBNA	EBV-encoded nuclear antigen
EBV	Epstein-Barr virus
ECL	enhanced chemiluminescence
EDTA	ethyldiamintetraacetic acid
ES cells	embryonic stem cells
et al.	et alii
EZH	human enhancer of zeste
f.e.	for example
Fig.	figure
G	guanosine
GAL	galactose metabolism
GFP	green fluorescent protein
gp	glycoprotein
h	hour
H_2O	water
H3	histone H3
H3K27	histone H3 lysine 27
H3K27me3	trimethylated histone H3 at lysine 27
H3K4	histone H3 lysine 4
H3K4me3	trimethylated histone H3 at lysine 4
H3K9	histone H3 lysine 9
H3K9me3	trimethylated histone H3 at lysine 9
HAT	histone acetyltransferase
HDAC	histone deacetylase
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HKMT	histone lysine methyl transferase
HLA	human leukocyte antigen
hmC	5-hydroxymethylcytosine
HOX	homeobox
HP1	heterochromatin protein 1
HRP	horseradish peroxidase
i.e.	id est
IE	immediate early
IFN	interferon
Ig	immune globuline
IL	interleukin
IM	infectious mononucleosis
INF	interferon
INO	inositol
ISWI	imitation swi
kDa	kilo dalton
KSHV	Kaposi Sarcoma associated herpesvirus
L	late
LCL	lymphoblastoid cell line
LMP	latent membrane protein
M.HhaI	DNA cytosine C5 methyltransferase from Haemophilus hemolyticus
MAP	mitogen-activated protein
MBD	methyl-CpG binding domain
MBP	methyl binding protein
MeCP	methyl cytosine binding protein
meZRE	methylated BZLF1 responsive element
MHC	major histocompatibility complex
min	minute

miRNA	microRNA
mRNA	messenger RNA
N-terminal	amino-terminal
ncRNA	noncoding RNA
NDR	nucleosome depleted region
NFkB	nuclear factor "kappa-light-chain-enhancer" of activated B cells
NURD	nucleosome remodeling factor
ORC	origin recognition complex
oriLyt	origin of lytic replication
oriP	origin of plasmid replication
PAA	polyacrylamide
PBMCs	peripheral blood mononuclear cells
PBS	phosphate buffered saline
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
PHO	phosphate metabolism
PcG	Polycomb group
PRC	Polycomb repressive complex
PRE	polycomb responsive element
PRMT1	Protein arginine N-methyltransferase 1
RNA	ribonucleic acid
RNA Pol II	RNA polymerase II
rpm	rounds per minute
RPMI	Roswell Park Memorial Institute
RSC	remodels the structure of chromatin
RT	room temperature
RT-PCR	reverse transcriptase PCR
RTA	replication and transcription activator
s	second
S. cerevisiae	Sacharomyces cerevisiae
S. pombe	Schizosaccharomyces pombe
SDS	sodium dodecyl sulfate
SDS page	SDS polyacrylamide gel electrophoresis
SUMO	small ubiquitin-like modifier
SWI/SNF	switch/sucrose nonfermentable
T	thymine
TAP	transporter associated with antigen processing
TBP	TATA box binding protein
TEMED	tetramethylendiamine
TET	ten-eleven translocation proteins
TF	transcription factor
TRD	transcriptional repressing domain
Tris	2-amino-2-hydroxymethyl-propane-1,3-diol
tRNA	transfer RNA
TSS	transcriptonal start site
U	untis
U	uracil
UV	ultra violett
V	volt
w/v	weight per volume
WHO	world health organization
ZRE	BZLF1-responsive element

9. LITERATUR

Adams, A. (1987). "Replication of latent Epstein-Barr virus genomes in Raji cells." J Virol **61**(5): 1743-6.

Adamson, A. L. and S. Kenney (1999). "The Epstein-Barr virus BZLF1 protein interacts physically and functionally with the histone acetylase CREB-binding protein." J Virol **73**(8): 6551-8.

Adkins, M. W., S. R. Howar, et al. (2004). "Chromatin disassembly mediated by the histone chaperone Asf1 is essential for transcriptional activation of the yeast PHO5 and PHO8 genes." Mol Cell **14**(5): 657-66.

Agalioti, T., S. Lomvardas, et al. (2000). "Ordered recruitment of chromatin modifying and general transcription factors to the IFN-beta promoter." Cell **103**(4): 667-78.

Altmann, M. and W. Hammerschmidt (2005). "Epstein-Barr virus provides a new paradigm: a requirement for the immediate inhibition of apoptosis." PLoS Biol **3**(12): e404.

Altmann, M., D. Pich, et al. (2006). "Transcriptional activation by EBV nuclear antigen 1 is essential for the expression of EBV's transforming genes." Proc Natl Acad Sci U S A **103**(38): 14188-93.

Amon, W., U. K. Binne, et al. (2004). "Lytic cycle gene regulation of Epstein-Barr virus." J Virol **78**(24): 13460-9.

Atanassov, B. S., E. Koutelou, et al. (2010). "The role of deubiquitinating enzymes in chromatin regulation." FEBS Lett.

Bai, L. and A. V. Morozov (2010). "Gene regulation by nucleosome positioning." Trends Genet **26**(11): 476-83.

Barski, A., S. Cuddapah, et al. (2007). "High-resolution profiling of histone methylations in the human genome." Cell **129**(4): 823-37.

Bell, A. C., A. G. West, et al. (1999). "The protein CTCF is required for the enhancer blocking activity of vertebrate insulators." Cell **98**(3): 387-96.

Bergbauer, M., M. Kalla, et al. (2010). "CpG-methylation regulates a class of Epstein-Barr virus promoters." PLoS Pathog **6**(9).

Bernstein, B. E., T. S. Mikkelsen, et al. (2006). "A bivalent chromatin structure marks key developmental genes in embryonic stem cells." Cell **125**(2): 315-26.

Bestor, T. H. and D. Bourc'his (2004). "Transposon silencing and imprint establishment in mammalian germ cells." Cold Spring Harb Symp Quant Biol **69**: 381-7.

Bhende, P. M., W. T. Seaman, et al. (2004). "The EBV lytic switch protein, Z, preferentially binds to and activates the methylated viral genome." Nat Genet **36**(10): 1099-104.

Bird, A. P. and A. P. Wolffe (1999). "Methylation-induced repression--belts, braces, and chromatin." Cell **99**(5): 451-4.

Bock, C., S. Reither, et al. (2005). "BiQ Analyzer: visualization and quality control for DNA methylation data from bisulfite sequencing." Bioinformatics **21**: 4067-8.

Bornkamm, G. W., C. Berens, et al. (2005). "Stringent doxycycline-dependent control of gene activities using an episomal one-vector system." Nucleic Acids Res **33**(16): e137.

Brookes, E. and A. Pombo (2009). "Modifications of RNA polymerase II are pivotal in regulating gene expression states." EMBO Rep **10**(11): 1213-9.

Bryant, G. O., V. Prabhu, et al. (2008). "Activator control of nucleosome occupancy in activation and repression of transcription." PLoS Biol **6**(12): 2928-39.

Burkitt, D. (1958). "A sarcoma involving the jaws in African children." Br J Surg **46**(197): 218-23.

Burkitt, D. (1962). "A children's cancer dependent on climatic factors." Nature **194**: 232-4.

Campbell, K. S. (1999). "Signal transduction from the B cell antigen-receptor." Curr Opin Immunol **11**(3): 256-64.

Cao, R., Y. Tsukada, et al. (2005). "Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing." Mol Cell **20**(6): 845-54.

Carrozza, M. J., B. Li, et al. (2005). "Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription." Cell **123**(4): 581-92.

Cedar, H. and Y. Bergman (2008). "Epigenetic silencing during early lineage commitment."

Chapman, R. D., M. Conrad, et al. (2005). "Role of the mammalian RNA polymerase II C-terminal domain (CTD) nonconsensus repeats in CTD stability and cell proliferation." Mol Cell Biol **25**(17): 7665-74.

Chapman, R. D., B. Palancade, et al. (2004). "The last CTD repeat of the mammalian RNA polymerase II large subunit is important for its stability." Nucleic Acids Res **32**(1): 35-44.

Chuang, L. S., H. I. Ian, et al. (1997). "Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1." Science **277**(5334): 1996-2000.

Corden, J. L., D. L. Cadena, et al. (1985). "A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II." Proc Natl Acad Sci U S A **82**(23): 7934-8.

Countryman, J., H. Jenson, et al. (1987). "Polymorphic proteins encoded within BZLF1 of defective and standard Epstein-Barr viruses disrupt latency." J Virol **61**(12): 3672-9.

Dahmus, M. E. and C. Kedinger (1983). "Transcription of adenovirus-2 major late promoter inhibited by monoclonal antibody directed against RNA polymerases IIO and IIA." J Biol Chem **258**(4): 2303-7.

Datta, A. K. and R. E. Hood (1981). "Mechanism of inhibition of Epstein-Barr virus replication by phosphonoformic acid." Virology **114**(1): 52-9.

Decaussin, G., V. Leclerc, et al. (1995). "The lytic cycle of Epstein-Barr virus in the nonproducer Raji line can be rescued by the expression of a 135-kilodalton protein encoded by the BALF2 open reading frame." J Virol **69**(11): 7309-14.

Delecluse, H. J., T. Hilsendegen, et al. (1998). "Propagation and recovery of intact, infectious Epstein-Barr virus from prokaryotic to human cells." Proc Natl Acad Sci U S A **95**(14): 8245-50.

Dickerson, S. J., Y. Xing, et al. (2009). "Methylation-dependent binding of the epstein-barr virus BZLF1 protein to viral promoters." PLoS Pathog **5**(3): e1000356.

Dirmeier, U., B. Neuhierl, et al. (2003). "Latent membrane protein 1 is critical for efficient growth transformation of human B cells by epstein-barr virus." Cancer Res **63**(11): 2982-9.

Dong, A., J. A. Yoder, et al. (2001). "Structure of human DNMT2, an enigmatic DNA methyltransferase homolog that displays denaturant-resistant binding to DNA." Nucleic Acids Res **29**(2): 439-48.

Dyson, P. J. and P. J. Farrell (1985). "Chromatin structure of Epstein-Barr virus." J Gen Virol **66 (Pt 9)**: 1931-40.

Ehrlich, M. (2003). "The ICF syndrome, a DNA methyltransferase 3B deficiency and immunodeficiency disease." Clin Immunol **109**(1): 17-28.

Elliott, J., E. B. Goodhew, et al. (2004). "Variable methylation of the Epstein-Barr virus Wp EBNA gene promoter in B-lymphoblastoid cell lines." J Virol **78**(24): 14062-5.

Epstein, M. A., B. G. Achong, et al. (1964). "Virus Particles in Cultured Lymphoblasts from Burkitt's Lymphoma." Lancet **1**(7335): 702-3.

Ernst, J. and M. Kellis (2010). "Discovery and characterization of chromatin states for systematic annotation of the human genome." Nat Biotechnol **28**(8): 817-25.

Feederle, R., M. Kost, et al. (2000). "The Epstein-Barr virus lytic program is controlled by the co-operative functions of two transactivators." EMBO J **19**(12): 3080-9.

Fernandez, A. F., C. Rosales, et al. (2009). "The dynamic DNA methylomes of double-stranded DNA viruses associated with human cancer." Genome Res **19**(3): 438-51.

Fixman, E. D., G. S. Hayward, et al. (1992). "trans-acting requirements for replication of Epstein-Barr virus ori-Lyt." J Virol **66**(8): 5030-9.

Fixman, E. D., G. S. Hayward, et al. (1995). "Replication of Epstein-Barr virus oriLyt: lack of a dedicated virally encoded origin-binding protein and dependence on Zta in cotransfection assays." J Virol **69**(5): 2998-3006.

Flemington, E. K., A. E. Goldfeld, et al. (1991). "Efficient transcription of the Epstein-Barr virus immediate-early BZLF1 and BRLF1 genes requires protein synthesis." J Virol **65**(12): 7073-7.

Floer, M., X. Wang, et al. (2010). "A RSC/nucleosome complex determines chromatin architecture and facilitates activator binding." Cell **141**(3): 407-18.

Fong, N. and D. L. Bentley (2001). "Capping, splicing, and 3' processing are independently stimulated by RNA polymerase II: different functions for different segments of the CTD." Genes Dev **15**(14): 1783-95.

Francis, N. J., R. E. Kingston, et al. (2004). "Chromatin compaction by a polycomb group protein complex." Science **306**(5701): 1574-7.

Gires, O., U. Zimber-Strobl, et al. (1997). "Latent membrane protein 1 of Epstein-Barr virus mimics a constitutively active receptor molecule." EMBO J **16**(20): 6131-40.

Goldberg, A. D., C. D. Allis, et al. (2007). "Epigenetics: a landscape takes shape." Cell **128**(4): 635-8.

Goll, M. G. and T. H. Bestor (2005). "Eukaryotic cytosine methyltransferases." Annu Rev Biochem **74**: 481-514.

Goll, M. G., F. Kirpekar, et al. (2006). "Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2." Science **311**(5759): 395-8.

Gowher, H. and A. Jeltsch (2001). "Enzymatic properties of recombinant Dnmt3a DNA methyltransferase from mouse: the enzyme modifies DNA in a non-processive manner and also methylates non-CpG [correction of non-CpA] sites." J Mol Biol **309**(5): 1201-8.

Grant, P. A., D. Schieltz, et al. (1998). "A subset of TAF(II)s are integral components of the SAGA complex required for nucleosome acetylation and transcriptional stimulation." Cell **94**(1): 45-53.

Grune, T., J. Brzeski, et al. (2003). "Crystal structure and functional analysis of a nucleosome recognition module of the remodeling factor ISWI." Mol Cell **12**(2): 449-60.

Guenther, M. G., S. S. Levine, et al. (2007). "A chromatin landmark and transcription initiation at most promoters in human cells." Cell **130**(1): 77-88.

Gunther, T. and A. Grundhoff (2010). "The epigenetic landscape of latent Kaposi sarcoma-associated herpesvirus genomes." PLoS Pathog **6**(6): e1000935.

Gutierrez, A. and R. J. Sommer (2004). "Evolution of dnmt-2 and mbd-2-like genes in the free-living nematodes Pristionchus pacificus, Caenorhabditis elegans and Caenorhabditis briggsae." Nucleic Acids Res **32**(21): 6388-96.

Hansen, K. H., A. P. Bracken, et al. (2008). "A model for transmission of the H3K27me3 epigenetic mark." Nat Cell Biol **10**(11): 1291-300.

Hansen, R. S., C. Wijmenga, et al. (1999). "The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome." Proc Natl Acad Sci U S A **96**(25): 14412-7.

Hassan, A. H., P. Prochasson, et al. (2002). "Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes." Cell **111**(3): 369-79.

Henle, G. and W. Henle (1966). "Studies on cell lines derived from Burkitt's lymphoma." Trans N Y Acad Sci **29**(1): 71-9.

Henle, G., W. Henle, et al. (1968). "Relation of Burkitt's tumor-associated herpes-ytpe virus to infectious mononucleosis." Proc Natl Acad Sci U S A **59**(1): 94-101.

Hermann, A., H. Gowher, et al. (2004a). "Biochemistry and biology of mammalian DNA methyltransferases." Cell Mol Life Sci **61**(19-20): 2571-87.

Hermann, A., R. Goyal, et al. (2004b). "The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites." J Biol Chem **279**(46): 48350-9.

Hicks, M. R., S. S. Al-Mehairi, et al. (2003). "The zipper region of Epstein-Barr virus bZIP transcription factor Zta is necessary but not sufficient to direct DNA binding." J Virol **77**(14): 8173-7.

Higuchi, R., C. Fockler, et al. (1993). "Kinetic PCR analysis: real-time monitoring of DNA amplification reactions." Biotechnology (N Y) **11**(9): 1026-30.

Holliday, R. (2006). "Epigenetics: a historical overview." Epigenetics **1**(2): 76-80.

Horst, D., D. van Leeuwen, et al. (2009). "Specific targeting of the EBV lytic phase protein BNLF2a to the transporter associated with antigen processing results in impairment of HLA class I-restricted antigen presentation." J Immunol **182**(4): 2313-24.

IARC (1997). Epstein-Barr virus and Kaposi's sarcoma herpesvirus/human herpesvirus 8. Lyon, IARC.

Iskra, S. (2010). Angeborene und adaptive Immunreaktionen in der Frühphase der Epstein-Barr-Virus-Infektion. HelmholtzZentrum München, Department of gene vectors. Munich, LMU.

Ito, S., A. C. D'Alessio, et al. (2010). "Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification." Nature **466**(7310): 1129-33.

Jeltsch, A., W. Nellen, et al. (2006). "Two substrates are better than one: dual specificities for Dnmt2 methyltransferases." Trends Biochem Sci **31**(6): 306-8.

Jochum, S. (2011). Munich, HelmholtzZentrum München, Department of gene vectors.

Jones, P. A. and B. S. Archer TK, Beck S, Berger S, Bernstein BE, Carpten JD, Clark SJ, Costello JF, Doerge RW, Esteller M, Feinberg AP, Gingeras TR, Greally JM, Henikoff S, Herman JG, Jackson-Grusby L, Jenuwein T, Jirtle RL, Kim YJ, Laird PW, Lim B, Martienssen R, Polyak K, Stunnenberg H, Tlsty TD, Tycko B, Ushijima T, Zhu J, Pirrotta V, Allis CD, Elgin SC, Jones PA, Martienssen R, Rine J, Wu C. (2008). "Moving AHEAD with an international human epigenome project." Nature **454**(7205): 711-5.

Kalla, M. (2007). Epigenetik von Epstein-Barr Virus: Regulation viraler Gene in der frühen Phase der Infektion. HelmholtzZentrum München, Department of gene vectors. Munich, LMU.

Kalla, M., A. Schmeinck, et al. (2010). "AP-1 homolog BZLF1 of Epstein-Barr virus has two essential functions dependent on the epigenetic state of the viral genome." Proc Natl Acad Sci U S A **107**(2): 850-5.

Kaneda, M., M. Okano, et al. (2004). "Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting." Nature **429**(6994): 900-3.

Kaplan, N., I. K. Moore, et al. (2009). "The DNA-encoded nucleosome organization of a eukaryotic genome." Nature **458**(7236): 362-6.

Karlsson, Q. H., C. Schelcher, et al. (2008). "Methylated DNA recognition during the reversal of epigenetic

silencing is regulated by cysteine and serine residues in the Epstein-Barr virus lytic switch protein." PLoS Pathog **4**(3): e1000005.

Katsumura, K. R., S. Maruo, et al. (2009). "Quantitative evaluation of the role of Epstein-Barr virus immediate-early protein BZLF1 in B-cell transformation." J Gen Virol **90**(Pt 10): 2331-41.

Keogh, M. C., S. K. Kurdistani, et al. (2005). "Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex." Cell **123**(4): 593-605.

Kieff, E. and A. Rickinson (2007). Epstein-Barr Virus and its replication. Philadelphia, Lippincott-Williams Wilkins.

Kim, W. Y. and M. E. Dahmus (1986). "Immunochemical analysis of mammalian RNA polymerase II subspecies. Stability and relative in vivo concentration." J Biol Chem **261**(30): 14219-25.

Kirchmaier, A. L. and B. Sugden (1995). "Plasmid maintenance of derivatives of oriP of Epstein-Barr virus." J Virol **69**(2): 1280-3.

Klose, R. J., S. A. Sarraf, et al. (2005). "DNA binding selectivity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG." Mol Cell **19**(5): 667-78.

Kornberg, R. D. (1974). "Chromatin structure: a repeating unit of histones and DNA." Science **184**(139): 868-71.

Kouzarides, T. (2007). "Chromatin modifications and their function." Cell **128**(4): 693-705.

Kraus, R. J., J. G. Perrigoue, et al. (2003). "ZEB negatively regulates the lytic-switch BZLF1 gene promoter of Epstein-Barr virus." J Virol **77**(1): 199-207.

Kriaucionis, S. and N. Heintz (2009). "The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain." Science **324**(5929): 929-30.

Ku, M., R. P. Koche, et al. (2008). "Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains." PLoS Genet **4**(10): e1000242.

Lander, E. S., L. M. Linton, et al. (2001). "Initial sequencing and analysis of the human genome." Nature **409**(6822): 860-921.

Langst, G. and P. B. Becker (2001). "Nucleosome mobilization and positioning by ISWI-containing chromatin-remodeling factors." J Cell Sci **114**(Pt 14): 2561-8.

Lantermann, A. B., T. Straub, et al. (2010). "Schizosaccharomyces pombe genome-wide nucleosome mapping reveals positioning mechanisms distinct from those of Saccharomyces cerevisiae." Nat Struct Mol Biol **17**(2): 251-7.

Leonhardt, H., A. W. Page, et al. (1992). "A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei." Cell **71**(5): 865-73.

Lewis, E. B. (1978). "A gene complex controlling segmentation in Drosophila." Nature **276**(5688): 565-70.

Lipford, J. R. and S. P. Bell (2001). "Nucleosomes positioned by ORC facilitate the initiation of DNA replication." Mol Cell **7**(1): 21-30.

Lister, R., M. Pelizzola, et al. (2009). "Human DNA methylomes at base resolution show widespread epigenomic differences." Nature **462**(7271): 315-22.

Lomvardas, S. and D. Thanos (2002). "Modifying gene expression programs by altering core promoter chromatin architecture." Cell **110**(2): 261-71.

Lorch, Y., J. W. LaPointe, et al. (1987). "Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones." Cell **49**(2): 203-10.

Loyola, A. and G. Almouzni (2004). "Histone chaperones, a supporting role in the limelight." Biochim Biophys Acta **1677**(1-3): 3-11.

Lu, X., M. D. Simon, et al. (2008). "The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure." Nat Struct Mol Biol **15**(10): 1122-4.

Luger, K., A. W. Mader, et al. (1997). "Crystal structure of the nucleosome core particle at 2.8 A resolution." Nature **389**(6648): 251-60.

Maga, G. and U. Hubscher (2003). "Proliferating cell nuclear antigen (PCNA): a dancer with many partners." J Cell Sci **116**(Pt 15): 3051-60.

Mancao, C., M. Altmann, et al. (2005). "Rescue of "crippled" germinal center B cells from apoptosis by Epstein-Barr virus." Blood **106**(13): 4339-44.

Mancao, C. and W. Hammerschmidt (2007). "Epstein-Barr virus latent membrane protein 2A is a B-cell receptor mimic and essential for B-cell survival." Blood **110**(10): 3715-21.

Marechal, V., A. Dehee, et al. (1999). "Mapping EBNA-1 domains involved in binding to metaphase chromosomes." J Virol **73**(5): 4385-92.

Martens, J. A. and F. Winston (2003). "Recent advances in understanding chromatin remodeling by Swi/Snf complexes." Curr Opin Genet Dev **13**(2): 136-42.

Martin, C. and Y. Zhang (2005). "The diverse functions of histone lysine methylation." Nat Rev Mol Cell Biol **6**(11): 838-49.

Marzluff, W. F., P. Gongidi, et al. (2002). "The human and mouse replication-dependent histone genes." Genomics **80**(5): 487-98.

Mavrich, T. N., C. Jiang, et al. (2008). "Nucleosome organization in the Drosophila genome." Nature **453**(7193):

358-62.

McDonald, C. M., C. Petosa, et al. (2009). "Interaction of Epstein-Barr virus BZLF1 C-terminal tail structure and core zipper is required for DNA replication but not for promoter transactivation." J Virol **83**(7): 3397-401.

Meilinger, D., K. Fellinger, et al. (2009). "Np95 interacts with de novo DNA methyltransferases, Dnmt3a and Dnmt3b, and mediates epigenetic silencing of the viral CMV promoter in embryonic stem cells." EMBO Rep **10**(11): 1259-64.

Metivier, R., G. Penot, et al. (2003). "Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter." Cell **115**(6): 751-63.

Metzger, E., A. Imhof, et al. (2010). "Phosphorylation of histone H3T6 by PKCbeta(I) controls demethylation at histone H3K4." Nature **464**(7289): 792-6.

Miller, G., T. Shope, et al. (1972). "Epstein-Barr virus: transformation, cytopathic changes, and viral antigens in squirrel monkey and marmoset leukocytes." Proc Natl Acad Sci U S A **69**(2): 383-7.

Minarovits, J. (2006). "Epigenotypes of latent herpesvirus genomes." Curr Top Microbiol Immunol **310**: 61-80.

Minarovits, J., L. F. Hu, et al. (1992). "RNA polymerase III-transcribed EBER 1 and 2 transcription units are expressed and hypomethylated in the major Epstein-Barr virus-carrying cell types." J Gen Virol **73 (Pt 7)**: 1687-92.

Mullis, K., F. Faloona, et al. (1992). "Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. 1986." Biotechnology **24**: 17-27.

Nakayama, J., J. C. Rice, et al. (2001). "Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly." Science **292**(5514): 110-3.

Neefjes, J. J., F. Momburg, et al. (1993). "Selective and ATP-dependent translocation of peptides by the MHC-encoded transporter." Science **261**(5122): 769-71.

Nemerow, G. R., C. Mold, et al. (1987). "Identification of gp350 as the viral glycoprotein mediating attachment of Epstein-Barr virus (EBV) to the EBV/C3d receptor of B cells: sequence homology of gp350 and C3 complement fragment C3d." J Virol **61**(5): 1416-20.

NOE. (2004-2008). "The Epigenome Network of Excellence." from http://www.epigenome-noe.net/.

Okano, M., D. W. Bell, et al. (1999). "DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development." Cell **99**(3): 247-57.

Okano, M., S. Xie, et al. (1998). "Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases." Nat Genet **19**(3): 219-20.

Panning, B. and R. Jaenisch (1996). "DNA hypomethylation can activate Xist expression and silence X-linked genes." Genes Dev **10**(16): 1991-2002.

Park, J. H., J. P. Jeon, et al. (2007). "Wp specific methylation of highly proliferated LCLs." Biochem Biophys Res Commun **358**(2): 513-20.

Paulsen, M. and A. C. Ferguson-Smith (2001). "DNA methylation in genomic imprinting, development, and disease." J Pathol **195**(1): 97-110.

Paulson, E. J. and S. H. Speck (1999). "Differential methylation of Epstein-Barr virus latency promoters facilitates viral persistence in healthy seropositive individuals." J Virol **73**(12): 9959-68.

Pulvertaft, J. V. (1964). "Cytology of Burkitt's Tumour (African Lymphoma)." Lancet **1**(7327): 238-40.

Ramirez-Carrozzi, V. R., D. Braas, et al. (2009). "A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling." Cell **138**(1): 114-28.

Ramirez-Carrozzi, V. R., A. A. Nazarian, et al. (2006). "Selective and antagonistic functions of SWI/SNF and Mi-2beta nucleosome remodeling complexes during an inflammatory response." Genes Dev **20**(3): 282-96.

Rickinson, A. and E. Kieff (2007). Epstein-Barr virus. Philadelphia, Lippincott-Williams Wilkins.

Ringrose, L. and R. Paro (2007). "Polycomb/Trithorax response elements and epigenetic memory of cell identity." Development **134**(2): 223-32.

Roizman, B. and A. Sears (2007). Herpes simplex viruses and their replication. Philadelphia, Lippincott-Williams Wilkins.

Rottach, A., C. Frauer, et al. (2009). "The multi-domain protein Np95 connects DNA methylation and histone modification." Nucleic Acids Res **38**(6): 1796-804.

Saha, A., J. Wittmeyer, et al. (2005). "Chromatin remodeling through directional DNA translocation from an internal nucleosomal site." Nat Struct Mol Biol **12**(9): 747-55.

Salamon, D., M. Takacs, et al. (2001). "Protein-DNA binding and CpG methylation at nucleotide resolution of latency-associated promoters Qp, Cp, and LMP1p of Epstein-Barr virus." J Virol **75**(6): 2584-96.

Sambrook and Russell (2001). Molecular Cloning: A Laboratory Manual, CSHL press.

Schaefer, B. C., J. L. Strominger, et al. (1995). "Redefining the Epstein-Barr virus-encoded nuclear antigen EBNA-1 gene promoter and transcription initiation site in group I Burkitt lymphoma cell lines." Proc Natl Acad Sci U S A **92**(23): 10565-9.

Schelcher, C., S. Al Mehairi, et al. (2007). "Atypical bZIP domain of viral transcription factor contributes to

stability of dimer formation and transcriptional function." J Virol **81**(13): 7149-55.

Schepers, A., M. Ritzi, et al. (2001). "Human origin recognition complex binds to the region of the latent origin of DNA replication of Epstein-Barr virus." EMBO J **20**(16): 4588-602.

Schermelleh, L., A. Haemmer, et al. (2007). "Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation." Nucleic Acids Res **35**(13): 4301-12.

Schoeftner, S., A. K. Sengupta, et al. (2006). "Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing." EMBO J **25**(13): 3110-22.

Schreiber, S. L. and B. E. Bernstein (2002). "Signaling network model of chromatin." Cell **111**(6): 771-8.

Sears, J., M. Ujihara, et al. (2004). "The amino terminus of Epstein-Barr Virus (EBV) nuclear antigen 1 contains AT hooks that facilitate the replication and partitioning of latent EBV genomes by tethering them to cellular chromosomes." J Virol **78**(21): 11487-505.

Serio, T. R., J. L. Kolman, et al. (1997). "Late gene expression from the Epstein-Barr virus BcLF1 and BFRF3 promoters does not require DNA replication in cis." J Virol **71**(11): 8726-34.

Seto, E., A. Moosmann, et al. (2010). "Micro RNAs of Epstein-Barr virus promote cell cycle progression and prevent apoptosis of primary human B cells." PLoS Pathog **6**(8).

Shaw, J. E., L. F. Levinger, et al. (1979). "Nucleosomal structure of Epstein-Barr virus DNA in transformed cell lines." J Virol **29**(2): 657-65.

Shimizu, N., S. Sakuma, et al. (1989). "Identification of an enhancer-type sequence that is responsive to Z and R trans-activators of Epstein-Barr virus." Virology **172**(2): 655-8.

Shivaswamy, S., A. Bhinge, et al. (2008). "Dynamic remodeling of individual nucleosomes across a eukaryotic genome in response to transcriptional perturbation." PLoS Biol **6**(3): e65.

Silva, A. J. and R. White (1988). "Inheritance of allelic blueprints for methylation patterns." Cell **54**(2): 145-52.

Smith, A. E., C. Chronis, et al. (2009). "Epigenetics of human T cells during the G0-->G1 transition." Genome Res **19**(8): 1325-37.

Song, C. X., K. E. Szulwach, et al. (2010). "Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine." Nat Biotechnol.

Speck, S. H. and D. Ganem (2010). "Viral latency and its regulation: lessons from the gamma-herpesviruses." Cell Host Microbe **8**(1): 100-15.

Strahl, B. D. and C. D. Allis (2000). "The language of covalent histone modifications." Nature **403**(6765): 41-5.

Summers, W. C. and G. Klein (1976). "Inhibition of Epstein-Barr virus DNA synthesis and late gene expression by phosphonoacetic acid." J Virol **18**(1): 151-5.

Svaren, J. and W. Horz (1997). "Transcription factors vs nucleosomes: regulation of the PHO5 promoter in yeast." Trends Biochem Sci **22**(3): 93-7.

Tahiliani, M., K. P. Koh, et al. (2009). "Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1." Science **324**(5929): 930-5.

Tao, Q., K. D. Robertson, et al. (1998). "The Epstein-Barr virus major latent promoter Qp is constitutively active, hypomethylated, and methylation sensitive." J Virol **72**(9): 7075-83.

Tempera, I., A. Wiedmer, et al. (2010). "CTCF prevents the epigenetic drift of EBV latency promoter Qp." PLoS Pathog **6**(8).

Thomae, A. W., D. Pich, et al. (2008). "Interaction between HMGA1a and the origin recognition complex creates site-specific replication origins." Proc Natl Acad Sci U S A **105**(5): 1692-7.

Tierney, R. J., H. E. Kirby, et al. (2000). "Methylation of transcription factor binding sites in the Epstein-Barr virus latent cycle promoter Wp coincides with promoter down-regulation during virus-induced B-cell transformation." J Virol **74**(22): 10468-79.

Toth, Z., D. T. Maglinte, et al. (2010). "Epigenetic analysis of KSHV latent and lytic genomes." PLoS Pathog **6**(7): e1001013.

Tsai, C. L., H. P. Li, et al. (2006). "Activation of DNA methyltransferase 1 by EBV LMP1 Involves c-Jun NH(2)-terminal kinase signaling." Cancer Res **66**(24): 11668-76.

Tsai, C. N., C. L. Tsai, et al. (2002). "The Epstein-Barr virus oncogene product, latent membrane protein 1, induces the downregulation of E-cadherin gene expression via activation of DNA methyltransferases." Proc Natl Acad Sci U S A **99**(15): 10084-9.

Tsurumi, T., M. Fujita, et al. (2005). "Latent and lytic Epstein-Barr virus replication strategies." Rev Med Virol **15**(1): 3-15.

Valouev, A., J. Ichikawa, et al. (2008). "A high-resolution, nucleosome position map of C. elegans reveals a lack of universal sequence-dictated positioning." Genome Res **18**(7): 1051-63.

Vignali, M., A. H. Hassan, et al. (2000). "ATP-dependent chromatin-remodeling complexes." Mol Cell Biol **20**(6): 1899-910.

Waddington, C. H. (1957). The strategy of the Genes; a Discussion of Some Aspects of Theoretical Biology, George Allen & Unwin.

Watson, J. D. and F. H. Crick (1953). "Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid." Nature **171**(4356): 737-8.

Watt, F. and P. L. Molloy (1988). "Cytosine methylation prevents binding to DNA of a HeLa cell transcription

factor required for optimal expression of the adenovirus major late promoter." Genes Dev **2**(9): 1136-43.

Weber, M., J. J. Davies, et al. (2005). "Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells." Nat Genet **37**(8): 853-62.

Wen, W., D. Iwakiri, et al. (2007). "Epstein-Barr virus BZLF1 gene, a switch from latency to lytic infection, is expressed as an immediate-early gene after primary infection of B lymphocytes." J Virol **81**(2): 1037-42.

Woisetschlaeger, M., C. N. Yandava, et al. (1990). "Promoter switching in Epstein-Barr virus during the initial stages of infection of B lymphocytes." Proc Natl Acad Sci U S A **87**(5): 1725-9.

Wood, A., J. Schneider, et al. (2003). "The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p." J Biol Chem **278**(37): 34739-42.

Wysocka, J., T. Swigut, et al. (2006). "A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling." Nature **442**(7098): 86-90.

Xu, D., J. Bai, et al. (2009). "Covalent modifications of histones during mitosis and meiosis." Cell Cycle **8**(22): 3688-94.

Yajima, Y., A. Tanaka, et al. (1976). "Inhibition of productive replication of Epstein-Barr virus DNA by phosphonoacetic acid." Virology **71**(1): 352-4.

Yang, Y. H., S. Dudoit, et al. (2002). "Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation." Nucleic Acids Res **30**(4): e15.

Yates, J., N. Warren, et al. (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 U S A **81**(12): 3806-10.

Yates, J. L., N. Warren, et al. (1985). "Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells." Nature **313**(6005): 812-5.

Yoder, J. A., C. P. Walsh, et al. (1997). "Cytosine methylation and the ecology of intragenomic parasites." Trends Genet **13**(8): 335-40.

Young, L. S. and P. G. Murray (2003). "Epstein-Barr virus and oncogenesis: from latent genes to tumours." Oncogene **22**(33): 5108-21.

Yu, X., P. J. McCarthy, et al. (2011). "ZIIR Element of BZLF1 Promoter of Epstein-Barr Virus Plays Central Role in Establishment and Maintenance of Viral Latency." J Virol.

Yuan, G. C., Y. J. Liu, et al. (2005). "Genome-scale identification of nucleosome positions in S. cerevisiae." Science **309**(5734): 626-30.

Zeidler, R., G. Eissner, et al. (1997). "Downregulation of TAP1 in B lymphocytes by cellular and Epstein-Barr virus-encoded interleukin-10." Blood **90**(6): 2390-7.

Zerby, D., C. J. Chen, et al. (1999). "The amino-terminal C/H1 domain of CREB binding protein mediates zta transcriptional activation of latent Epstein-Barr virus." Mol Cell Biol **19**(3): 1617-26.

Zhou, J., C. M. Chau, et al. (2005). "Cell cycle regulation of chromatin at an origin of DNA replication." EMBO J **24**(7): 1406-17.

zur Hausen, H., H. Schulte-Holthausen, et al. (1970). "EBV DNA in biopsies of Burkitt tumours and anaplastic carcinomas of the nasopharynx." Nature **228**(5276): 1056-8.

10. APPENDIX

10.1   Oligonucleotides

10.1.1   RT-PCR Primer

Tab. A1 PCR primer pairs for reverse transcript PCR (RT-PCR)

locus	forward	backward
BZLF1	GGTTTCCGTGTGCGTCGTG	AGCCTGCTCCTGAGAATGCTT
BBLF4	CTGGGCAAGGTGACAAATGTAATC	GAAGCAGGCGAGGCAAGAAC
BALF4	CTGGGGGGTGAGGAAGTCG	CAACACAACCGTGGGCATAGAG
BcLF1	CCTCCCTGACCGTTCCCAG	GCAGTTTGAGACCGCCACATC
BDLF1	GCACCTCCTCTGCTATGGGC	TGATACTCACCAAGATTGTTCCAGG
gp350	ACCGAGCATTTCTGTTTTTACGC	GATGTCTACTTTCAAGATGTGTTTGGAAC
cytC	CAATGCTCCGTTGTTGGCAG	CCTGGTGGGCGTGTGCTAC
DNMT1	AAGGAGCCCGTGGATGAGG	CTCGCTGGAGTGGACTTGTGG
DNMT3a	CTTGGGCATTCAGGTGGACC	GCGAGCAGGGTTGACGATG
DNMT3b	TACTGCCCCGCACCCAAG	GAGCCCCCCCTCAAAGAGAG
NP95	TGCTGCTCCTCCTCCTCCC	GGCAACCGCTACGATGGC
DNMT3LF1B3	TGCGGAAGTCTCCAGGTTCAC	CACAGGTAGCACACCCAGTTGC
gusB	CGCCCTGCCTATCTGTATTC	TCCCCACAGGGAGTGTGTAG
NP95neu	AGATCCAGGAGCTGTTCCAC	AAGAGGGTATGGCCGTCCT
PCNA	TGGAGAACTTGGAAATGGAAA	GAACTGGTTCATTCATCTCTATGG

10.1.2 qPCR Primer

Tab. A2 Primer pairs for qPCR analysis

	forward	backward
Cp	ACAAGGGGACAAGTGTGGCA	TGACTGGTGGGGGGCATC
EBER	CGCTACATCAAACAGGACAGC	AGCCGAATACCCTTCTCCCAG
Lmp1p	GAAATGGAAAGGCAGTGCGG	TCACCTGAACCCCCCTAAAGC
OriP	TCAAACACTTGCCCACAAAC	GCCCACCGTGCTCTCAGC
Qp	TGTCACCACCTCCCTGATAATGTC	CATACACCGTGCGAAAAGAAGC
Wp1	GCCTAAAACCCCCAGGAAGC	GACCCCCTCTTACATTTGTGT
Wp2	CAGACAGGGGAGTGGGCTTG	GATTCCTTGGAGGGGGCG
CEN	AAGGTCAATGGCAGAAAAGGA	CAACGAAGGCCACAAGATGTC
GAPDH	TACTAGCGGTTTTACGGGCG	TCGAACAGGAGGAGCAGAGAGCGA
OriLyt1	GGTCTCTGTGTAATACTTTAAGGTTTGCTC	AGCCCCTCCTCCTCTCGTTATC
BARF1	AACGCATTGTGAGGGATGGG	CGGGAGGCAGGAGACACG
BBLF3	CTGCTGTCCACCCTGGTCAAG	CCCGAGGTCAGAAAGAGGCTC
BBLF4	GTGCTACACAGCCGCTCCG	CGATACTCCTGATGGTCCTCTCG
BFRF1	TCCAGGGCACAAAGTCCTCC	AACCCAGAAGACCTCACCTTGC
BKRF4	ACTCAACACCATCCTGACCGTG	CTTCTGCTGGGCACTCTTTCG
BMLF1	GCTGACCCAGGCGACGAG	TTCAAAACCTCTTACATCACTCACTGC
BMRF1	CACACCACCCCCCAAGGAC	GCAGCAGCAGAAGCCAACG
BRLF1	CCGGCTGACATGGATTACTGG	AGGAACCAAAATAACCGAGCCTC
BSRF1	CCAAAAATAGTAAGCAGCCGTGAC	GAAACAGCCACAGGGGGATG
BVRF2	GTAAATAAACTCATCGCACGGGG	TGCCTTGTCCACTGGGGTG
BZLF1	GGTGCAATGTTTAGTGAGTTACCTGTC	TGACACCAGCTTATTTTAGACACTTCTG
BcLF1	AATCAAATGGTTGGACACGGC	TCAGGGTGGGCAGAGGACC
BDLF4	GTGTCCGTAATGGATGGGGG	CAGCCAGCGACTTGGAGGG
BLLF1	TACAAGGGGGGTGCGGTG	CAGGTGGGCATCTTCTGCTTC
BMRF1 1	TCCACGGGGAGACTTATGCTAAC	GCCACCTTGTCACGGAGCC
BMRF1 2	GCCAATAACTACATAAGTAGGGATGAGC	TCTCCACGAGCCTCGATGAAG
BMRF1 3	CACACCACCCCCCAAGGAC	GCAGCAGCAGAAGCCAACG
BMRF1 4	CCGTCCTGTCCAAGTGCTATGAC	TCTGGGCTCTGGTGATTCTGC
BMRF1 5	CGCACCTGTGTCCTGGGC	CAATCATCTGCTCGTTCCTCAGC
BBLF4 1	TCCCGTGACCAGGCAGTCC	TCCCCCGAGTTTGACCCG
BBLF4 2	TCACGGAGGCGTCTGAGGTC	TTGGCGTAGATGGTCTGTCCC
BBLF4 3	GAGGCAGGTGTTTACCCATTTG	CAGGTCACGCACGGTCAGC
BBLF4 4	GTGCTACACAGCCGCTCCG	CGATACTCCTGATGGTCCTCTCG
BBLF4 5	TCCCCCAGTTTACGGAGGTG	CATGAAAATGTTCTGGTAGAGGCG

10.1.3 Deep bisulfite sequencing primer

Tab. A3 List of Primers for deep bisulfite sequencing
The table includes the contig (Cg.), the ID, the promoter site, the start and stopp position of the PCR product, and the sequences for the forward and reverse primer.

Cg.	ID	prom.	start	stopp	forward (5' 3')	reverse (5' 3')
1	33	EBER	6303	6741	AAGGTTAGTTTGTAAGGTGGATG	TAAAAAAACAACCACAAACACC
	33_3	EBER	6722	7245	TGTTTGTGGTTGTTTTTTTAGA	TATCCAAAAAACCCTATCACAA
	33_2	EBER	6726	7107	TGTGGTTGTTTTTTTAGATTT	TACCCTTCTCCCAAAAAAATTA
	57	OriP	7125	7441	TTTTTTGTGGTTAGTTTTGTATTT	TACAATCAACATATAATACCCAATAA
2	31	OriP	8357	8930	TTGTTTTTGTTTATGGGTTTTATT	AACCATTTTAATCACAAAAACAA
	32	OriP	8909	9437	TGTTTTTGTGATTAAAATGGTTT	CCAAACACAAAAAAATTCTCTAA
3	50	BCRF1	9457	9946	TGGAATAGGGTTTAGTTGGTAA	ACCTATAACATAACTTCCTCCAAA
4	26	Cp	11172	11522	GGGTTGGGTAAAGGGGTTT	ATTAAAACCTCCTAAAAAAACCTCAA
	25	Cp	11499	11982	GAGGTTTTTTTAGGAGGTTTTAA	CAATAAAAAACATCTAAAACCCA
5	34	Wp	14033	14379	AGTTTTGGAGGATTTTAAAATTT	CCAAAAATAACTACAAAACCACT
	35	Wp	14411	14918	GTAAGAGGGGGTTTTTTATTTT	ACTAACCCCCAAATTCCTATA
6	46	BHLF1	49563	50128	GGTAGGGGTAAGGGTAAGTTTA	AAAACACCAATTTTAACAAACC
	47	BHLF1	50106	50560	AGGTTTGTTAAAATTGGTGTTT	TACCCATAAAATACTCAAACCC
7	14	OriLyt	52565	52947	GTTTAGGGGTTTAGGGGGTAG	CAAATACCACCCACCTAATAACA
	15	OriLyt	52925	53355	TGTTATTAGGTGGGTGGTATTT	TAAAATTCCCTTCCTCTAAATTCT
	16	OriLyt	53183	53590	TTTAGTTAGGTTTTTATTGGGG	TCAAAATAATACATCCCAAAAA
	17	BHRF1	53567	54118	TGTTTTTGGGATGTATTATTTT	CTAAAAAAAAAAATAACCCACC
8	62	BFLF2 BFLF1	57055	57509	TTTTGGTTATTTGTGTGGATTT	CCTCCTAACCCCCATATTTAA
	63	BFLF2 BFLF1	57490	57881	TAAATATGGGGGTTAGGAGGTA	CTTCAAATCCATATAACCCAAAA
	64	BFLF2 BFLF1	57856	58547	TGGTTTTGGGTTATATGGATTT	CACCTTACCACCCAAATATATCATA
	65	BFLF2 BFLF1	58516	58964	TGTGAGTTATGATATATTTGGGTG	AACCCAAAATCACACAAAATA
9	27	Fp/Qp	61755	62339	TGGGGTTGGTGTTTATGTAT	AATCACCCAATTTCTATCTATTAAAA
	28	Fp/Qp	62334	62829	GTGATTATTGAGGGAGTGTTTT	AAACACAAAAACCCCAAAATA
10	51	BORF1	74893	75254	AGGGTYGTTTTTAGATGGT	AACCCCTAAACCTTCATAAC
11	7	BMRF1	79351	79823	TGTTGATTGAAGGTATTTTTTTT	AATAACCACCCTAACTCAAAAA
	8	BMRF1	79813	80406	GGGTGGTTATTTGTTTAGGTTTT	TAAACATCAACAACACCTACCC
12	52	BMRF2	80706	81167	TTTGGTATTATAGTTGTGGTGGT	AAACAAAAAACACAAACCCC
13	18	BSLF2	83317	83820	AGTTATGTGTTTTTGTTGGTTGAT	TTACACCAAAAACCACAAAAAA
	19	BSLF2	83809	84342	TTTTTGGTGTAAGAGGAGTTGTT	AAACCAATCAAAACAACACTACA
	20	BSLF2	84266	84649	GAGAGTTTTTGAGAAGGAATTATTT	CTCATAAACACCCTAACAACAA
14	21	BSLF2	85120	85678	TAATAGTATGGTTTTTAGGTAGGGG	CAATAAAATCCCAAAAAACAA
	22	BSLF2	85250	85678	TGGGTGGGTTTTGGATAT	CCAATAAAATCCCAAAAAACA
	23	BSLF1 /BSRF1	86389	86862	GGGAGTTAAAGGTGTTAATGAATT	TCAAAACCCTTATAACCTCAAAC
	24	BSLF1 /BSRF1	86695	87237	GTTTTTTGGGTAGGTGAAGTTGT	TTAACCACCTAAACCTCCACC
15	53	BLLF1	91966	92291	AGATTAAGTTGATGTTTTTTGTTT	CTTCCATATTATCATCCAAAAC
16	42	BZLF1	103093	103493	AGGTATTTGGTATGGGTTAGGT	TAAATCTAAACTCCCCCCTAAC
	41	BZLF1	103501	103856	TGTTATGGATTTTAGTGTTGTGG	ATCAACCAAAAAAAATCAAACC
17	1	BRLF1	105121	105560	TTGTTTTTTGATTTTAGGAGTTAGTTT	TCCTAATCAAAACCCAAAATCT
	2	BRLF1	105541	105946	ATTTTGGGTTTTGATTAGGAAAT	CCATCCTAAAATATTATCCAACTACA
	3	BRLF1	105925	106503	TTGGATAATATTTTAGGATGGTATG	TCCATAAAACAAACAAAAAAAAA
	4	BRLF1	106487	107035	TTTGTTTGTTTTATGGAATGTT	ATCCAATAACCACTAAAAACCT
18	43	BBLF4	114054	114315	ATATAGGTTGAAAAGGGTAGGAA	ACTTTCCCTACAAAATCTCAAA
	44	BBLF4	114286	114714	TTYGATTTTTTGAGATTTTGTA	CCCTCAAACTCTTTATCTTAAAA
	45	BBLF4	114701	115165	AAAGAGTTTGAGGGATAGGTG	AACTTAACCAACAAATTCAACAA

Cg.	ID	prom.	start	stopp	forward (5' 3')	reverse (5' 3')
19	11	BBLF2/3	118531	119068	GAGTATGGGGGTATGTTTGTTTT	AAAAATCCAATCCTCCTCAAAT
	12	BBLF2/3	119047	119623	ATTTGAGGAGGATTGGATTTT	CCAAAAAATACAAAAAATACACAAAA
	13	BBLF2/3	119758	120356	TTGTATTTGGGGTGTTTTGT	CATTAAAAAAAAATCTCCTCCTC
20	54	BGRF1	124496	124945	GTGGGTTAATATTTGTAAGGTTT	TACRAAACATACAACATATTCAA
21	55	BDRF1	128721	129279	GTTGGTTTTGTGTGTTTTG	ACCTCCTTCTTAATAAAATTAACC
22	60	BcLF1	137428	137762	GGGGAAGGGTTTGTTTTTTATAT	ACTCCAAACTCTAAACTCCAAA
	61	BcLF2	137674	137999	AATGAATGATTGTTAGGAGTTGT	AACCAACATCTTTACAAAATTTC
23	9	BALF5	156458	156978	AGATATGGAAGTTTAGAGGGTTTT	AAAAATCCAAACATTAATCCCA
	10	BALF5	156956	157422	AATGGGATTAATGTTTGGATTT	AAAACATTAACTTTACCTCCCTAAA
24	5	BALF2	164666	165223	TTTTTTAAGAGGGTAGGTGGT	CTCATCAAAACCTCCCTAAAA
	6	BALF2	165230	165736	GAGGATTAGGGTTGGTAAAGGT	AAAAAACCCCTAAAAAACCACT
25	37	LMP2a	166736	167245	TTTATGAGGATTTATATTGGGGT	CCTTCAAAAAAATATCCAAAAAA
	38	LMP2a	167222	167695	ATTTTTTGGATATTTTTTTGAAG	ACCTAACTCATACCAACAAAAA
26	39	LMP1	169017	169496	ATTTTTAGGGAATGTTAGATTTTATT	CTCCTAACACACTACCCTAAAA
	40	LMP1	169480	169772	GGGTAGTGTGTTAGGAGTAAGGT	CCTTTCCATTCCTATTACACTTAA

10.2 Deep bisulfite sequencing analysis

10.2.1 Matlab script

```matlab
1  inputfile='6303-7441.xls';
2  outputfile=[inputfile(1:end-4) 'out.xls'];
3  data=xlsread(inputfile);
4  startpos=data(1,1);
5  data=data(3:end,2:end);
6  dim=size(data);
7  numlines=dim(1);
8  numcol=dim(2);
9  distances=[];
10 c=1;
11 while c<=numcol
12   distances=[distances data(1,c)];
13   c=c+2;
14 end
15 positions=distances;
16 positions(1)=distances(1)+startpos+1;
17 for c=2:length(distances)
18   positions(c)=positions(c-1)+distances(c)+2;
19 end
20 positions=positions(1:end-1);
21 meth=[];
22 notmeth=[];
23 c=2;
24 while c<=numcol
25   states=data(:,c);
26   meth=[meth length(find(states==1))];
27   notmeth=[notmeth length(find(states==0))];
28   c=c+2;
29 end
30 percentmeth=[];
31 for i=1:length(meth)
32   percentmeth(i)=meth(i)/(meth(i)+notmeth(i))*100;
33 end
34 both=meth+notmeth;
35 outdata=[[1:1:length(meth)]' positions' percentmeth' both'];
36 out=cell(0,0);
37 out{1,1}='CpG';
38 out{1,2}='Position';
39 out{1,3}='Percent methylated';
40 out{1,4}='Total';
41 for i=1:length(meth)
42   out{i+1,1}=outdata(i,1);
43   out{i+1,2}=outdata(i,2);
44   out{i+1,3}=outdata(i,3);
45   out{i+1,4}=outdata(i,4);
46 end
47 xlswrite(outputfile,out);
```

10.2.2 Complete list of deep bisulfite sequencing results

Tab. A4 Methylation state of viral CpG dinucleotides
ID, genome location (pos.), percentage of methylation (meth.), and coverage (cov.) of all CpG positions analyzed. Results, which originate from LCLs are colored in yellow. A coverage below ten reads/bp is indicated in red. CpG dinucleotides that are not present (n.p.) in the Raji EBV genome are indicated.

ID	pos.	meth.	cov.	ID	pos.	meth.	cov.	ID	pos.	meth.	cov.	ID	pos.	meth.	cov.
1	6327	0.00	73	50	7366	0.00	1	99	14115	100.00	38	148	50381	96.87	1659
2	6348	0.00	63	51	7384	0.00	1	100	14143	91.89	37	149	50388	96.29	1670
3	6353	0.00	90	52	8049	45.27	634	101	14161	97.44	39	150	50410	97.57	1688
4	6389	0.00	89	53	8060	15.50	684	102	14259	97.62	42	151	50413	97.63	1690
5	6402	1.10	91	54	8066	4.63	713	103	14261	97.62	42	152	50438	97.66	1710
6	6408	1.92	52	55	8099	9.16	928	104	14288	95.45	44	153	50474	95.30	1744
7	6413	1.41	71	56	8107	5.35	954	105	14290	97.73	44	154	50514	95.00	1741
8	6427	0.00	89	57	8136	94.85	1010	106	14296	97.78	45	155	50518	97.88	1748
9	6481	0.00	96	58	8145	86.99	1007	107	14381	n.p.	/	156	50537	93.91	1642
10	6485	1.02	98	59	8147	75.37	1023	108	14391	n.p.	/	157	52587	97.64	127
11	6487	0.00	98	60	8163	60.52	1031	109	14445	81.25	64	158	52589	100.00	133
12	6504	1.03	97	61	8190	72.15	1034	110	14462	95.68	139	159	52597	100.00	144
13	6538	0.00	92	62	8229	96.64	1131	111	14497	85.90	624	160	52599	98.61	144
14	6565	0.00	99	63	8241	96.16	1147	112	14509	94.17	617	161	52604	96.58	146
15	6570	0.00	101	64	8291	95.62	1255	113	14513	94.91	609	162	52610	98.69	153
16	6575	0.99	101	65	8298	96.94	1274	114	14531	92.49	799	163	52621	95.21	167
17	6581	0.00	102	66	8301	96.14	1271	115	14534	94.68	808	164	52623	97.60	167
18	6637	0.97	103	67	8315	97.05	1286	116	14543	94.41	823	165	52627	92.86	168
19	6668	0.00	100	68	8317	94.70	1284	117	14566	93.90	869	166	52643	95.86	169
20	6688	0.00	98	69	9517	93.51	185	118	14590	83.97	923	167	52646	88.82	161
21	6693	0.00	97	70	9519	96.30	189	119	14595	88.25	919	168	52661	0.00	2
22	6706	1.06	94	71	9528	97.04	203	120	14604	95.82	934	169	52672	65.54	177
23	6719	2.27	88	72	9545	95.63	206	121	14610	96.16	911	170	52698	94.12	187
24	6757	0.68	1612	73	9580	86.88	221	122	14693	66.01	1018	171	52700	95.21	188
25	6763	0.90	777	74	9681	98.02	252	123	14709	93.54	1021	172	52707	92.43	185
26	6786	0.26	2286	75	9801	18.85	244	124	14731	94.78	1034	173	52727	94.23	208
27	6836	0.49	2637	76	9830	16.48	267	125	14735	70.52	1038	174	52730	95.22	209
28	6852	0.15	2729	77	11193	93.46	474	126	14770	79.31	1068	175	52738	97.60	208
29	6862	0.18	2765	78	11197	59.10	489	127	14776	93.18	1071	176	52741	97.61	209
30	6883	0.28	2827	79	11209	95.85	530	128	14790	92.76	1078	177	52746	95.24	210
31	6893	0.18	2799	80	11220	97.23	541	129	14817	95.32	1089	178	52768	97.64	212
32	6903	0.24	2870	81	11237	96.22	556	130	49596	97.37	38	179	52772	98.12	213
33	6908	0.24	2883	82	11243	94.78	556	131	49604	97.44	39	180	52780	96.26	214
34	6938	0.38	2915	83	11275	96.27	590	132	49670	35.19	54	181	52783	97.67	215
35	6943	0.21	2916	84	11364	95.38	692	133	49727	88.61	79	182	52789	97.22	216
36	6964	0.58	2940	85	11386	97.32	709	134	49733	62.96	81	183	52794	97.27	220
37	6981	0.24	2951	86	11404	96.08	714	135	49753	94.57	92	184	52799	97.26	219
38	6990	0.50	2974	87	11439	95.31	725	136	49766	97.14	105	185	52804	96.40	139
39	6996	0.33	2989	88	11479	94.52	693	137	49786	95.80	119	186	52905	36.55	290
40	7005	0.54	2989	89	11649	21.05	133	138	49826	93.22	118	187	52969	94.45	1711
41	7014	0.17	2975	90	11766	92.03	138	139	49856	94.41	143	188	52993	93.47	1777
42	7019	0.27	2936	91	11772	94.16	137	140	49868	91.03	145	189	52995	92.72	1785
43	7032	0.21	1903	92	11778	97.04	135	141	49950	43.05	151	190	53011	85.50	1841
44	7056	0.42	1887	93	11852	11.19	143	142	50039	26.09	161	191	53069	96.18	2040
45	7110	1.82	220	94	14077	97.06	34	143	50094	6.33	158	192	53074	95.72	2057
46	7119	0.46	217	95	14085	94.44	36	144	50177	86.93	1163	193	53147	5.12	2109
47	7165	0.00	224	96	14101	92.11	38	145	50182	94.05	1211	194	53230	53.22	3322
48	7331	0.00	1	97	14103	94.74	38	146	50223	95.28	1376	195	53248	96.28	2714
49	7356	0.00	1	98	14105	94.74	38	147	50356	89.13	1619	196	53394	27.63	2092

Tab. A4 (continued) Methylation state of viral CpG dinucleotides
ID, genome location (pos.), percentage of methylation (meth.), and coverage (cov.) of all CpG positions analyzed. Results, which originate from LCLs are colored in yellow. A coverage below ten reads/bp is indicated in red. CpG dinucleotides that are not present (n.p.) in the Raji EBV genome are indicated.

ID	pos.	meth.	cov.	ID	pos.	meth.	cov.	ID	pos.	meth.	cov.	ID	pos.	meth.	cov.
197	53428	70.52	2225	246	57566	64.57	1682	295	58718	94.38	427	344	62145	90.20	102
198	53436	92.20	2256	247	57577	93.71	1733	296	58720	93.91	427	345	62148	96.08	102
199	53446	95.47	2275	248	57580	93.52	1729	297	58759	89.38	433	346	62161	96.12	103
200	53448	94.37	2275	249	57590	93.89	1750	298	58790	84.74	439	347	62166	96.15	104
201	53492	96.02	2359	250	57592	94.25	1775	299	58803	94.39	446	348	62178	88.68	106
202	53499	96.61	2359	251	57601	83.01	1783	300	58851	73.33	450	349	62208	0.00	106
203	53538	17.09	1182	252	57625	51.06	1516	301	58865	95.19	457	350	62222	0.95	105
204	53558	42.14	2404	253	57635	65.78	1812	302	58895	91.36	463	351	62227	2.83	106
205	53604	20.56	107	254	57647	61.83	1842	303	58901	94.86	467	352	62230	0.93	108
206	53611	16.39	183	255	57673	63.68	1875	304	58920	62.58	465	353	62254	0.00	108
207	53663	93.60	297	256	57712	99.05	1690	305	58932	65.24	466	354	62272	0.91	110
208	53681	96.01	326	257	57738	49.34	1753	306	61782	87.50	32	355	62275	0.00	109
209	53686	97.98	346	258	57769	91.84	1778	307	61791	92.50	40	356	62282	0.91	110
210	53692	95.94	345	259	57787	92.22	1760	308	61800	86.36	44	357	62374	0.59	170
211	53751	22.11	502	260	57790	75.57	1764	309	61818	100.00	52	358	62384	1.30	231
212	53777	88.81	581	261	57893	54.00	50	310	61821	90.20	51	359	62386	0.00	155
213	53804	89.83	639	262	57904	0.00	2	311	61827	94.23	52	360	62398	0.80	251
214	53840	92.38	682	263	57943	87.72	57	312	61836	96.15	52	361	62405	0.40	253
215	53883	94.89	705	264	57979	81.48	54	313	61856	n.p.	/	362	62411	0.00	260
216	53893	71.97	710	265	57993	93.22	59	314	61859	90.63	64	363	62429	0.66	303
217	53953	30.22	771	266	58066	86.96	69	315	61865	79.69	64	364	62431	0.00	280
218	53958	41.12	783	267	58137	86.24	109	316	61869	89.55	67	365	62439	0.00	345
219	53981	94.58	793	268	58192	47.66	107	317	61881	81.82	77	366	62443	0.58	346
220	53998	91.13	812	269	58222	21.49	121	318	61883	79.45	73	367	62449	0.29	343
221	54008	87.73	815	270	58239	31.71	123	319	61918	23.94	71	368	62456	0.00	339
222	54054	82.84	816	271	58267	55.91	127	320	61923	5.48	73	369	62506	0.31	325
223	54096	32.74	733	272	58279	49.61	129	321	61925	8.33	72	370	62509	0.84	359
224	57092	22.39	804	273	58307	97.64	127	322	61931	45.71	35	371	62529	1.00	100
225	57101	12.27	823	274	58309	94.57	129	323	61937	32.47	77	372	62533	0.00	409
226	57133	73.88	896	275	58345	93.65	126	324	61944	91.67	84	373	62552	0.00	417
227	57143	75.22	900	276	58354	93.50	123	325	61947	92.77	83	374	62559	0.24	420
228	57193	71.20	986	277	58378	60.00	125	326	61953	96.30	81	375	62574	1.67	420
229	57197	83.25	985	278	58418	92.74	124	327	61956	91.67	84	376	62581	0.71	420
230	57203	60.46	994	279	58420	93.39	121	328	61959	96.43	84	377	62586	1.68	417
231	57235	83.68	1017	280	58441	69.49	118	329	61965	94.05	84	378	62588	1.90	421
232	57263	96.18	1022	281	58451	47.01	117	330	61968	98.81	84	379	62590	2.38	421
233	57317	30.75	1031	282	58480	96.43	112	331	61974	97.70	87	380	62600	3.37	415
234	57343	87.34	1027	283	58498	85.71	126	332	61986	98.41	63	381	62605	3.31	423
235	57373	93.50	1030	284	58504	98.41	126	333	62001	n.p.	/	382	62632	93.15	409
236	57377	0.68	1026	285	58508	88.89	126	334	62007	97.73	88	383	62671	97.44	430
237	57402	72.16	1020	286	58570	60.33	305	335	62012	95.45	88	384	62673	96.06	431
238	57421	57.32	1024	287	58576	93.23	310	336	62014	95.45	88	385	62675	96.05	430
239	57425	88.09	1016	288	58597	94.94	336	337	62016	1.28	78	386	62678	97.19	427
240	57436	79.53	1031	289	58629	45.72	339	338	62023	92.05	88	387	62690	92.24	438
241	57445	95.08	1037	290	58637	0.00	353	339	62044	18.89	90	388	62693	79.86	437
242	57465	72.94	1031	291	58646	92.54	268	340	62104	36.17	94	389	62696	67.65	439
243	57521	94.68	1411	292	58669	89.34	366	341	62113	94.12	102	390	62702	50.92	436
244	57529	94.57	1528	293	58691	92.29	402	342	62121	93.20	103	391	62704	52.74	438
245	57539	95.67	1549	294	58705	97.62	421	343	62133	93.20	103	392	62718	2.59	425

Tab. A4 (continued) Methylation state of viral CpG dinucleotides
ID, genome location (pos.), percentage of methylation (meth.), and coverage (cov.) of all CpG positions analyzed. Results, which originate from LCLs are colored in yellow. A coverage below ten reads/bp is indicated in red. CpG dinucleotides that are not present (n.p.) in the Raji EBV genome are indicated.

ID	pos.	meth.	cov.	ID	pos.	meth.	cov.	ID	pos.	meth.	cov.	ID	pos.	meth.	cov.
393	62724	0.90	443	442	79561	93.02	673	491	80999	95.10	4369	540	84240	83.47	242
394	62728	3.83	444	443	79586	61.73	682	492	81029	73.44	4375	541	84263	93.90	246
395	62731	7.74	439	444	79593	37.04	683	493	81031	71.98	4389	542	84303	93.66	584
396	62735	3.84	443	445	79623	96.74	705	494	81046	93.79	4431	543	84306	93.59	577
397	62741	0.00	7	446	79651	96.35	713	495	81049	95.41	4424	544	84367	85.56	381
398	62747	13.86	440	447	79655	95.51	712	496	81125	14.02	4030	545	84373	94.55	404
399	62761	11.29	443	448	79659	95.10	715	497	83356	88.04	443	546	84391	94.71	416
400	62777	94.43	449	449	79699	54.38	730	498	83377	94.91	609	547	84395	93.79	419
401	74898	83.73	1518	450	79720	47.76	735	499	83384	80.34	590	548	84414	96.24	425
402	74912	73.59	1882	451	79775	16.47	771	500	83429	94.48	870	549	84418	95.77	426
403	74918	87.16	1901	452	79872	27.27	11	501	83436	91.92	928	550	84421	94.81	424
404	74926	94.14	1912	453	79920	7.69	13	502	83445	88.54	951	551	84446	93.04	431
405	74943	94.13	1960	454	79943	4.55	22	503	83457	63.22	949	552	84527	91.40	442
406	74947	95.11	1962	455	80034	n.p.	/	504	83556	65.29	1190	553	84543	46.88	433
407	74963	80.98	1972	456	80042	92.45	53	505	83565	9.09	11	554	84574	91.87	443
408	74971	96.65	1970	457	80045	98.11	53	506	83584	94.30	1229	555	84587	74.89	462
409	74979	96.20	1976	458	80054	79.63	54	507	83595	90.48	1240	556	85153	83.92	3115
410	74982	95.70	1975	459	80062	3.64	55	508	83618	92.82	1282	557	85155	94.26	3120
411	74985	94.87	1967	460	80066	10.71	56	509	83626	0.00	7	558	85172	66.75	2271
412	74993	94.33	1974	461	80081	6.52	46	510	83638	95.40	1304	559	85211	92.12	3593
413	74999	93.46	1971	462	80096	4.76	63	511	83665	56.90	1327	560	85216	95.01	3666
414	75029	0.00	15	463	80105	0.00	63	512	83685	32.42	1351	561	85219	92.96	3665
415	75031	96.61	2003	464	80140	1.37	73	513	83724	0.00	19	562	85278	94.43	4994
416	75042	96.30	1998	465	80144	2.78	72	514	83736	75.25	1305	563	85281	96.69	5022
417	75093	97.19	2029	466	80147	1.35	74	515	83749	96.80	1470	564	85290	97.25	5135
418	75095	97.34	2029	467	80171	5.13	78	516	83752	90.57	1474	565	85302	95.60	5176
419	75097	98.32	2028	468	80175	1.25	80	517	83771	97.22	1475	566	85305	95.90	5152
420	75103	96.94	2023	469	80209	1.19	84	518	83792	95.98	1442	567	85309	90.05	5154
421	75107	95.77	2035	470	80244	0.00	88	519	83794	95.03	1469	568	85314	95.19	5202
422	75110	96.52	2009	471	80262	0.00	82	520	83797	92.30	1468	569	85340	39.33	5151
423	75116	79.41	2045	472	80327	1.77	113	521	83846	57.89	19	570	85394	82.72	5520
424	75139	95.71	2050	473	80350	1.89	106	522	83852	63.64	22	571	85401	86.87	5512
425	75141	97.18	2055	474	80357	3.45	116	523	83872	51.67	60	572	85422	95.24	5520
426	75146	93.47	2051	475	80372	0.86	116	524	83928	69.37	111	573	85441	94.75	5528
427	75148	96.63	2046	476	80729	28.24	2737	525	83934	83.05	118	574	85499	73.89	5585
428	75152	84.31	2052	477	80740	39.40	2855	526	83954	30.30	132	575	85512	93.50	5526
429	75161	86.53	1960	478	80746	59.23	2813	527	83965	0.69	145	576	85523	58.35	5479
430	75188	53.51	2052	479	80758	87.07	3025	528	83978	78.81	151	577	85534	40.40	5520
431	75200	98.69	1992	480	80789	95.55	3239	529	83985	86.39	147	578	85541	78.89	5406
432	75206	95.33	2057	481	80795	96.40	3248	530	84047	99.48	194	579	85567	95.43	5468
433	75209	94.65	2055	482	80803	97.08	3286	531	84071	95.98	224	580	85616	96.48	5402
434	75230	95.02	2009	483	80807	95.72	3438	532	84096	88.11	227	581	85623	96.23	5404
435	75234	93.88	1992	484	80815	96.44	3483	533	84101	80.70	228	582	85635	96.86	5155
436	79404	89.73	292	485	80844	85.86	3691	534	84177	85.36	239	583	85642	94.69	5030
437	79410	88.82	170	486	80877	17.04	4060	535	84187	60.70	229	584	86420	91.72	640
438	79476	88.44	545	487	80890	32.17	4140	536	84195	92.47	239	585	86431	97.45	667
439	79478	96.08	561	488	80914	87.15	3899	537	84201	92.50	240	586	86433	97.46	669
440	79521	88.50	626	489	80947	93.41	4202	538	84212	93.85	244	587	86442	94.92	649
441	79554	92.41	659	490	80968	83.42	4211	539	84231	95.10	245	588	86445	95.13	678

Tab. A4 (continued) Methylation state of viral CpG dinucleotides

ID, genome location (pos.), percentage of methylation (meth.), and coverage (cov.) of all CpG positions analyzed. Results, which originate from LCLs are colored in yellow. A coverage below ten reads/bp is indicated in red. CpG dinucleotides that are not present (n.p.) in the Raji EBV genome are indicated.

ID	pos.	meth.	cov.	ID	pos.	meth.	cov.	ID	pos.	meth.	cov.	ID	pos.	meth.	cov.
589	86458	93.27	683	638	87035	93.69	206	687	105513	24.93	726	736	106836	53.23	124
590	86469	96.64	715	639	87041	94.69	207	688	105518	44.49	717	737	106859	96.15	130
591	86477	91.03	736	640	87046	93.30	209	689	105535	45.40	718	738	106862	83.97	131
592	86482	94.99	739	641	87108	91.51	212	690	105566	96.53	1614	739	106866	89.31	131
593	86484	95.02	743	642	87139	91.43	210	691	105613	96.10	1767	740	106876	94.57	129
594	86518	95.67	785	643	87162	91.24	217	692	105625	94.82	1797	741	106889	0.79	126
595	86539	92.33	795	644	87176	97.63	211	693	105630	96.91	1815	742	106904	93.70	127
596	86546	97.00	800	645	87184	89.05	210	694	105643	96.76	1823	743	106913	98.45	129
597	86552	97.76	802	646	87187	96.65	209	695	105646	95.83	1823	744	106925	95.35	129
598	86556	97.26	803	647	87191	97.14	210	696	105676	96.42	1873	745	106929	91.54	130
599	86567	94.15	803	648	87193	95.24	210	697	105678	95.57	1873	746	114086	73.97	726
600	86569	95.91	806	649	87209	88.83	206	698	105689	96.51	1893	747	114095	62.03	748
601	86571	96.79	810	650	91991	97.08	342	699	105720	63.98	1921	748	114106	52.16	740
602	86592	92.18	818	651	91997	90.31	351	700	105726	90.98	1928	749	114116	79.74	844
603	86597	96.93	815	652	92025	86.11	396	701	105756	83.65	1945	750	114128	91.29	873
604	86599	96.82	817	653	92061	21.29	404	702	105760	95.07	1948	751	114142	93.66	883
605	86606	98.16	817	654	92097	15.99	444	703	105777	88.93	1970	752	114144	95.48	885
606	86609	95.72	818	655	92159	96.69	484	704	105781	95.69	1971	753	114147	94.14	888
607	86617	96.95	819	656	92165	95.47	486	705	105788	95.55	1978	754	114180	95.77	923
608	86619	96.46	819	657	92208	97.23	541	706	105790	97.36	1972	755	114187	92.96	923
609	86642	96.45	789	658	92259	18.35	545	707	105839	94.17	2041	756	114193	92.22	925
610	86645	95.06	809	659	92268	10.73	410	708	105904	7.54	2016	757	114218	84.99	946
611	86657	91.40	814	660	103138	16.67	180	709	105950	72.40	250	758	114224	91.97	946
612	86671	95.93	811	661	103181	97.84	185	710	105994	95.48	310	759	114230	90.27	946
613	86676	93.12	814	662	103272	4.29	233	711	106000	96.96	329	760	114239	93.11	944
614	86694	93.75	816	663	103467	2.50	80	712	106017	95.82	359	761	114245	86.03	945
615	86725	93.39	938	664	103530	0.40	494	713	106042	94.38	480	762	114253	87.37	942
616	86735	95.12	942	665	103544	0.36	558	714	106129	82.52	715	763	114274	55.28	937
617	86740	76.43	946	666	103552	0.78	387	715	106185	87.89	743	764	114277	80.75	935
618	86753	83.39	963	667	103638	13.97	981	716	106229	87.14	824	765	114283	66.49	934
619	86758	94.34	971	668	103686	91.01	990	717	106238	95.25	842	766	114288	43.76	962
620	86781	95.70	976	669	103698	96.16	808	718	106276	81.52	844	767	114317	91.41	198
621	86791	93.24	976	670	103705	49.03	1077	719	106293	95.87	848	768	114323	83.42	199
622	86799	96.02	980	671	103726	96.55	1132	720	106324	71.62	821	769	114344	92.95	312
623	86801	94.09	981	672	103735	94.82	1140	721	106376	89.36	846	770	114354	95.45	330
624	86815	83.11	983	673	103746	89.50	1219	722	106434	95.35	860	771	114362	93.55	341
625	86828	95.89	974	674	103750	95.50	1244	723	106436	96.61	855	772	114373	92.18	358
626	86866	91.14	158	675	103802	92.89	1576	724	106457	68.46	764	773	114402	88.01	392
627	86869	92.54	134	676	103813	26.06	1558	725	106541	87.50	32	774	114421	95.98	423
628	86875	96.77	155	677	103821	25.16	1391	726	106556	93.18	44	775	114425	97.64	424
629	86892	78.62	159	678	105209	89.54	478	727	106559	88.89	45	776	114431	89.12	432
630	86905	84.05	163	679	105256	25.00	552	728	106576	64.15	53	777	114455	61.52	434
631	86908	48.97	145	680	105292	14.29	602	729	106634	28.95	76	778	114474	95.31	448
632	86922	88.55	166	681	105298	9.56	607	730	106647	88.61	79	779	114481	83.74	449
633	86949	83.67	196	682	105304	10.08	615	731	106655	87.18	78	780	114485	98.02	454
634	86956	94.85	194	683	105406	89.44	682	732	106660	58.23	79	781	114493	87.42	453
635	86975	97.51	201	684	105479	32.13	722	733	106687	77.91	86	782	114499	97.37	457
636	87012	94.09	203	685	105487	41.24	725	734	106691	95.40	87	783	114504	93.22	457
637	87030	92.20	205	686	105490	31.22	724	735	106742	87.50	96	784	114510	85.81	458

Tab. A4 (continued) Methylation state of viral CpG dinucleotides
ID, genome location (pos.), percentage of methylation (meth.), and coverage (cov.) of all CpG positions analyzed. Results, which originate from LCLs are colored in yellow. A coverage below ten reads/bp is indicated in red. CpG dinucleotides that are not present (n.p.) in the Raji EBV genome are indicated.

ID	pos.	meth.	cov.	ID	pos.	meth.	cov.	ID	pos.	meth.	cov.	ID	pos.	meth.	cov.
785	114524	46.29	458	834	118747	96.77	31	883	119438	81.25	16	932	124738	87.42	946
786	114529	42.58	458	835	118750	90.91	33	884	119452	100.00	16	933	124746	93.90	951
787	114538	23.14	458	836	118756	100.00	33	885	119465	87.50	16	934	124753	69.77	956
788	114554	50.98	459	837	118759	69.70	33	886	119483	50.00	16	935	124792	16.15	997
789	114558	38.66	463	838	118780	47.06	34	887	119495	76.47	17	936	124814	16.87	1002
790	114585	95.27	465	839	118785	88.57	35	888	119514	73.33	15	937	124837	26.85	1002
791	114600	92.74	468	840	118797	88.57	35	889	119583	77.78	18	938	124841	18.17	1007
792	114605	95.07	467	841	118806	97.06	34	890	119594	100.00	11	939	124890	95.48	1039
793	114621	81.20	468	842	118822	91.18	34	891	119786	100.00	9	940	124894	78.52	1043
794	114641	90.27	473	843	118829	88.57	35	892	119798	100.00	12	941	124901	95.56	1036
795	114647	96.60	470	844	118835	88.89	36	893	119805	83.33	12	942	124909	94.91	1041
796	114659	95.10	469	845	118837	88.89	36	894	119865	88.89	27	943	124918	95.97	1043
797	114662	91.45	468	846	118846	56.67	30	895	119909	84.00	25	944	124928	95.77	1041
798	114722	90.76	249	847	118871	94.59	37	896	119935	67.65	34	945	124930	95.56	1037
799	114729	80.83	240	848	118874	94.74	38	897	119958	95.45	44	946	128741	95.12	41
800	114743	61.76	272	849	118894	74.36	39	898	119970	96.30	54	947	128748	80.39	51
801	114764	94.18	275	850	118909	72.50	40	899	120006	93.44	61	948	128764	94.29	70
802	114769	93.13	291	851	118930	100.00	40	900	120012	93.85	65	949	128767	93.06	72
803	114787	83.95	299	852	118939	87.50	40	901	120024	95.71	70	950	128779	89.25	93
804	114805	91.48	317	853	118949	85.37	41	902	120080	53.52	71	951	128805	96.19	105
805	114817	88.47	321	854	118959	95.12	41	903	120106	n.p.	/	952	128844	92.86	126
806	114835	95.72	327	855	118966	90.00	40	904	120124	94.59	74	953	128862	93.23	133
807	114848	90.38	343	856	118994	80.49	41	905	120152	93.42	76	954	128871	92.59	135
808	114859	93.47	352	857	119002	n.p.	/	906	120158	96.10	77	955	128895	29.37	126
809	114866	90.83	360	858	119004	50.00	40	907	120185	80.52	77	956	128917	35.42	144
810	114869	62.36	364	859	119014	97.50	40	908	120200	89.61	77	957	128944	72.15	158
811	114894	95.03	382	860	119021	92.31	39	909	120202	93.42	76	958	128957	48.43	159
812	114897	94.76	382	861	119023	92.50	40	910	120206	92.00	75	959	128974	87.27	165
813	114925	81.77	384	862	119078	100.00	3	911	120216	90.79	76	960	128990	96.86	159
814	114938	96.95	393	863	119084	33.33	3	912	120220	98.68	76	961	128992	93.64	173
815	114943	95.18	394	864	119130	100.00	4	913	120226	98.72	78	962	129005	87.15	179
816	114953	93.18	396	865	119159	57.14	7	914	120229	100.00	78	963	129029	95.72	187
817	114961	75.13	398	866	119163	100.00	7	915	120232	98.72	78	964	129064	71.88	192
818	114969	89.11	404	867	119174	100.00	6	916	120242	75.32	77	965	129105	90.10	202
819	114999	94.65	411	868	119186	100.00	8	917	120282	61.84	76	966	129110	90.82	207
820	115003	67.31	413	869	119219	100.00	10	918	120300	91.03	78	967	129116	90.34	207
821	115039	42.51	414	870	119222	80.00	10	919	120309	96.20	79	968	129118	95.65	207
822	115058	78.88	412	871	119238	100.00	10	920	120317	100.00	73	969	129126	93.24	207
823	115067	85.71	385	872	119277	92.31	13	921	120323	97.22	72	970	129130	97.56	205
824	115074	72.90	428	873	119292	100.00	12	922	120326	93.94	66	971	129193	88.84	215
825	115122	83.49	436	874	119342	93.75	16	923	120333	95.45	66	972	129221	91.83	208
826	118564	57.14	7	875	119357	93.75	16	924	124550	94.81	636	973	129253	28.64	206
827	118570	90.91	11	876	119360	93.75	16	925	124557	95.52	647	974	137500	20.32	1378
828	118602	64.29	14	877	119394	100.00	16	926	124626	86.43	781	975	137533	81.45	1547
829	118616	87.50	16	878	119396	87.50	16	927	124631	57.97	790	976	137549	93.32	1588
830	118690	100.00	25	879	119403	100.00	16	928	124671	13.36	876	977	137569	64.98	1622
831	118693	100.00	26	880	119421	100.00	16	929	124706	28.32	925	978	137604	13.13	1676
832	118707	96.00	25	881	119427	93.75	16	930	124726	80.79	937	979	137632	94.47	1718
833	118713	68.00	25	882	119430	93.75	16	931	124734	90.58	945	980	137634	90.79	1716

Tab. A4 (continued) Methylation state of viral CpG dinucleotides

ID, genome location (pos.), percentage of methylation (meth.), and coverage (cov.) of all CpG positions analyzed. Results, which originate from LCLs are colored in yellow. A coverage below ten reads/bp is indicated in red. CpG dinucleotides that are not present (n.p.) in the Raji EBV genome are indicated.

ID	pos.	meth.	cov.	ID	pos.	meth.	cov.	ID	pos.	meth.	cov.	ID	pos.	meth.	cov.
981	137639	96.67	1710	1030	157009	98.39	62	1079	165109	9.08	804	1128	165424	11.63	3372
982	137648	95.71	1724	1031	157015	90.74	54	1080	165111	11.93	805	1129	165453	9.23	4451
983	137651	92.87	856	1032	157034	95.24	63	1081	165118	12.30	805	1130	165473	20.05	4503
984	137662	79.02	1735	1033	157043	95.24	63	1082	165132	13.88	814	1131	165475	13.56	4500
985	137665	94.50	1744	1034	157053	92.31	65	1083	165136	11.84	819	1132	165518	20.63	4950
986	137672	95.49	1753	1035	157055	96.92	65	1084	165140	7.33	818	1133	165551	21.66	5119
987	137701	59.87	5774	1036	157059	93.85	65	1085	165154	7.25	745	1134	165554	14.50	5173
988	137721	95.71	6148	1037	157064	96.92	65	1086	165160	6.83	805	1135	165569	32.09	4132
989	137728	92.85	6248	1038	157094	92.75	69	1087	165182	15.49	768	1136	165585	20.73	5441
990	137773	69.90	4322	1039	157120	92.86	70	1088	165184	14.70	796	1137	165623	9.94	5613
991	137803	16.82	4381	1040	157123	95.77	71	1089	165225	75.00	4	1138	165687	19.58	5915
992	137817	77.73	3804	1041	157184	33.77	77	1090	165267	15.55	2874	1139	165698	8.82	5919
993	137842	86.21	4548	1042	157216	75.32	77	1091	165278	17.97	5020	1140	166764	0.00	3
994	137850	65.34	4536	1043	157249	86.84	76	1092	165280	20.78	5134	1141	166768	0.00	3
995	137865	55.35	4600	1044	157255	84.21	76	1093	165325	13.50	6103	1142	166775	0.00	3
996	137879	96.81	4696	1045	157270	89.33	75	1094	165329	19.59	6170	1143	166830	0.00	16
997	137891	97.67	4720	1046	157299	93.06	72	1095	165332	15.96	6290	1144	166833	0.00	18
998	137918	96.89	4893	1047	157306	94.81	77	1096	165335	16.23	6363	1145	166842	0.00	19
999	137926	96.21	4992	1048	157309	93.51	77	1097	165341	22.95	6518	1146	166870	3.03	33
1000	137939	97.34	5009	1049	157319	97.37	76	1098	165343	16.88	6551	1147	166896	12.82	39
1001	137941	97.48	5009	1050	157338	94.44	72	1099	165350	21.42	6644	1148	166904	0.00	38
1002	137949	95.97	4985	1051	157363	55.41	74	1100	165355	17.70	6361	1149	166979	0.00	47
1003	137951	96.69	4983	1052	157374	60.81	74	1101	165359	19.48	6324	1150	166989	0.00	47
1004	137954	97.14	70	1053	157381	82.43	74	1102	165409	21.56	7866	1151	167064	3.92	51
1005	137976	0.04	4887	1054	157386	55.71	70	1103	165424	11.18	6787	1152	167086	0.00	54
1006	156482	87.50	56	1055	164710	11.00	200	1104	165453	9.02	8915	1153	167103	1.89	53
1007	156498	94.52	73	1056	164714	8.17	208	1105	165473	19.71	9046	1154	167108	1.96	51
1008	156507	92.59	81	1057	164720	7.69	247	1106	165475	13.69	9042	1155	167113	0.00	55
1009	156523	91.26	103	1058	164749	18.45	206	1107	165518	20.90	9973	1156	167117	0.00	55
1010	156529	89.57	115	1059	164785	12.59	405	1108	165551	21.43	10346	1157	167130	0.00	55
1011	156554	71.01	69	1060	164790	14.98	327	1109	165554	14.39	10425	1158	167132	0.00	55
1012	156570	64.84	128	1061	164832	9.77	532	1110	165569	31.86	8350	1159	167140	7.02	57
1013	156613	89.24	158	1062	164897	2.94	613	1111	165585	20.23	10990	1160	167149	1.75	57
1014	156632	95.51	156	1063	164902	6.13	620	1112	165623	9.34	11298	1161	167166	89.29	56
1015	156635	93.29	164	1064	164913	8.95	637	1113	165687	19.35	11908	1162	167172	87.50	56
1016	156667	92.51	187	1065	164921	4.59	653	1114	165698	8.76	11921	1163	167179	94.44	54
1017	156740	94.17	206	1066	164935	13.56	671	1115	165267	14.99	1421	1164	167189	100.00	52
1018	156781	93.49	215	1067	164945	14.04	698	1116	165278	17.54	2492	1165	167197	90.74	54
1019	156784	98.62	218	1068	164961	13.13	731	1117	165280	20.40	2549	1166	167248	95.42	502
1020	156787	96.80	219	1069	164983	11.32	751	1118	165325	13.15	3058	1167	167251	96.53	518
1021	156799	67.27	220	1070	164998	5.97	770	1119	165329	19.75	3083	1168	167261	95.19	561
1022	156807	94.71	227	1071	165004	6.89	769	1120	165332	15.99	3139	1169	167286	14.73	577
1023	156816	96.51	229	1072	165007	8.09	766	1121	165335	16.39	3172	1170	167290	22.54	639
1024	156831	20.09	234	1073	165033	11.04	761	1122	165341	23.06	3252	1171	167375	95.55	741
1025	156874	92.74	248	1074	165038	6.96	776	1123	165343	16.57	3271	1172	167387	96.14	725
1026	156878	92.34	248	1075	165075	5.93	792	1124	165350	21.30	3314	1173	167415	96.88	769
1027	156923	95.31	256	1076	165077	4.40	795	1125	165355	17.83	3186	1174	167426	95.76	779
1028	157000	96.77	62	1077	165088	11.78	798	1126	165359	19.27	3145	1175	167435	95.01	781
1029	157006	87.10	62	1078	165097	10.17	806	1127	165409	22.27	3933	1176	167467	89.00	791

Tab. A4 (continued) Methylation state of viral CpG dinucleotides

ID, genome location (pos.), percentage of methylation (meth.), and coverage (cov.) of all CpG positions analyzed. Results, which originate from LCLs are colored in yellow. A coverage below ten reads/bp is indicated in red. CpG dinucleotides that are not present (n.p.) in the Raji EBV genome are indicated.

ID	pos.	meth.	cov.	ID	pos.	meth.	cov.	ID	pos.	meth.	cov.	ID	pos.	meth.	cov.
1177	167477	6.94	792	1186	169150	97.14	70	1195	169550	8.22	657	1204	169645	59.15	541
1178	167489	0.76	791	1187	169348	1.39	72	1196	169556	6.96	661	1205	169648	46.37	537
1179	167512	8.36	718	1188	169370	5.88	68	1197	169567	25.42	657	1206	169669	13.21	477
1180	167571	90.71	807	1189	169425	2.17	46	1198	169569	16.95	655	1207	169691	4.26	517
1181	167642	95.77	804	1190	169434	2.00	50	1199	169591	3.05	459	1208	169712	1.99	453
1182	167659	64.17	801	1191	169437	4.08	49	1200	169603	19.02	552	1209	169720	1.64	61
1183	169054	86.15	65	1192	169439	0.00	45	1201	169613	21.29	559	1210	169742	0.23	440
1184	169093	90.00	70	1193	169445	8.33	48	1202	169624	33.94	548				
1185	169111	94.29	70	1194	169465	n.p.	/	1203	169639	82.53	538				

10.3 Nucleosome occupancies at BZLF1 responsive promoters

Tab. A5 List of ZREs that were part of the MND-on-ChIP analysis
The ZREs and meZREs had to be at least 1500bp apart from each other in the EBV genome. In promoters that encompass more than one ZRE, the position with the strongest binding of BZLF1 was chosen.

ID	ZRE type	start of ZRE	end of ZRE	EBV genome feature
1	meZRE	2531	2538	*BNRF1* reading frame
2	meZRE	4275	4282	*BNRF1* reading frame
3	ZRE	9159	9166	*BCRF1* reading frame
4	ZRE	49288	49295	n.o.s.
10	ZRE	53630	53636	*oriLyt*
12	meZRE	56777	56784	*BFLF2* promoter
14	ZRE	58518	58524	*BFLF1* promoter
16	meZRE	64602	64609	*BPLF1* reading frame
19	meZRE	77941	77948	*Barf1* promoter
21	ZRE	79760	79766	*BMRF1* promoter
24	meZRE	84646	84653	*BMLF1* promoter
26	meZRE	86548	86555	*BSRF1* promoter
28	meZRE	91000	91007	*BLLF1* reading frame
29	ZRE	93977	93984	n.o.s.
30	ZRE	97961	97968	n.o.s.
31	ZRE	103303	103311	*BZLF1* promoter
34	meZRE	106372	106378	*BRLF1* promoter
36	ZRE	110108	110115	*BKRF3* promoter
38	meZRE	114495	114502	*BBLF4* promoter
40	meZRE	119796	119803	*BBLF2* promoter
41	ZRE	122476	122483	*BGLF5* promoter
42	meZRE	126796	126803	*BGLF2* promoter
44	ZRE	129593	129600	*BDLF4* promoter
45	ZRE	131379	131386	*BDLF3* promoter
47	ZRE	137155	137162	n.o.s.
48	ZRE	145154	145131	*BVRF1* promoter
51	ZRE	148198	148205	*BdRF1* promoter
52	ZRE	150517	150524	*BILF2* promoter
54	ZRE	152264	152271	n.o.s.
55	meZRE	155182	155189	*BALF5* reading frame
56	meZRE	156993	157000	*BALF5* promoter
59	meZRE	167567	167574	*BNLF2a*/b promoter

PUBLICATIONS

Parts of this work were published in scientific literature.

Kalla, M., **Schmeinck, A**., Bergbauer, M., Pich, D. & Hammerschmidt, W. (2010). "AP-1 homolog BZLF1 of Epstein-Barr virus has two essential functions dependent on the epigenetic state of the viral genome." Proc Natl Acad Sci U S A 107(2): 850-5.

Bergbauer, M., Kalla, M., **Schmeinck, A**., Göbel, C., Rothbauer, U., Eck, S., Benet- Pagès, A., Strom, T., and Hammerschmidt, W. (2010). "CpG-methylation regulates a class of Epstein-Barr virus' promoters. ", PloS Pathogens 6(9)

CURRICULUM VITAE

Personal Information

Name	Anne Schmeinck
Date of birth	11.01.1982
Place of birth	Bocholt
Nationality	German

Education

1988-1992	Primary school ("Pestalozzi-Schule") Bocholt
1992-2001	Elementary school ("Euregio-Gymnasium") Bocholt
	A-levels and qualification for university: May, 2001
2001-2006	Biotechnology, University of Münster
2005-2006	Diploma thesis at the institute for general botany and microbiology
	Department of biology and biotechnology of fungi
	Prof. Dr. B. Tudzynski/ Prof. Dr. P. Tudzynski
	"The nitrogen regulation in *Fusarium fujikuroi*: molecular analysis of important elements and their protein-protein interactions"

PhD work

2007-2011	PhD thesis at the Helmholtz Center Munich
	Department of gene vectors
	Prof. Dr. Wolfgang Hammerschmidt
	"Acquisition and loss of chromatin modifications during an Epstein-Barr virus infection"

VIELEN DANK