Examination of *in vivo* protein-DNA interactions and CpG-methylation patterns in the EBER promoter region of latent Epstein-Barr virus genomes

Thesis of the Ph.D. dissertation

of

Ferenc Bánáti

Place of research: National Center for Epidemiology, Microbiological Research Group

Supervisor: Dr. János Minárovits M.D., D.Sc., Head of the Microbiological Research Group

2010
INTRODUCTION

Multicellular organisms consist of different cell types, carrying the same or nearly the same genetic information. The complicated processes resulting in this phenotypic diversity can be attributed to the differentially controlled regulation of gene activity within a single cell. The regulatory pathways can manifest through processes depending directly on the nucleotide sequence of the genome and through the so-called epigenetic processes that are based on the post-synthetic modifications of the genomic DNA.

Epstein-Barr virus (EBV), a human pathogenic herpesvirus, is the causative agent of infectious mononucleosis and is closely associated with several neoplasms (e.g. Burkitt’s lymphoma). After infection of human B cells and oropharyngeal epithelial cells, the virus normally persists in a latent state in B cells without virus production. In healthy individuals lytic (productive) EBV replication occurs only occasionally when the virus is activated. In latently infected cells (memory B cells, tumour cells and lymphoblastoid cell lines) there is a restricted viral gene expression pattern, corresponding to different latency types (I, II, III), which can be distinguished based on their alternative promoter usage. In the various latency forms maximum 10 out of the approximately hundred EBV genes are active. Due to the strictly limited expression of latent genes, EBV remains hidden from the host immune system during persistence.

Among the latently transcribed genes the EBER transcripts (EBER1 and 2) are the most abundant. Transcription of these protein non-coding RNAs is governed by the RNA polymerase II (Pol II) and RNA polymerase III (Pol III) elements located on the promoter. They are constitutively transcribed regardless of the latency type, suggesting an indispensable role in the life cycle of the virus. Although, their involvement in tumorigenesis is already confirmed, their exact mechanism of action is not clearly described yet.

It is well established that the methylation of cytosines at cytosine-guanine dinucleotides (CpG) and the histone modifications within promoter areas are key regulators of Pol II transcribed cellular and viral genes. The role of CpG-methylation is less clear at Pol III-transcribed promoters. There are CpG-methylation sensitive and insensitive Pol III promoters as well.

Because the EBER1 and 2 genes are always active and clearly participate in tumorigenesis, it is important to examine the protein binding and CpG-methylation patterns and the chromatin structure at their regulatory sequences, in order to broaden our knowledge about processes regulating their activities.
OBJECTIVES

1) Using the bisulphite sequencing method I wished to examine the regulatory and coding sequences of the EBER genes to create the detailed CpG-methylation maps of the EBER region in different cell lines.

2) Applying quantitative real-time PCR I planned to determine the activity of the EBER1 and the EBER2 genes in different cell lines.

3) By transfection of an in vitro CpG-methylated or unmethylated EBER-carrying vector into EBV-negative cells, I wished to analyse the effect of CpG-methylation on the activity of the EBER promoters.

4) To complement the previously published DMS in vivo footprinting results of B cell lines, I intended to examine the protein binding pattern at the regulatory sequences of EBER1 in the nasopharyngeal carcinoma cell line C666-1.

5) Using a fragment of the EBER1 regulatory sequence, by applying the method of in vitro footprinting, I wished to analyse the protein binding patterns at unmethylated and in vitro methylated EBER1 promoters.

6) To analyse, whether the binding of different nuclear proteins (ATF and CTCF), depends on the methylation status of their recognition sequence, I planed to use EMSA (electrophoretic mobility shift assay) experiments with unmethylated and methylated binding sequences.

7) Using the chromatin immunoprecipitation assay one can confirm the in vivo existance of suspected or in vitro described protein-DNA interactions. On this way, I wished to study the binding of transcription factors to the EBER1 promoter sequence in vivo.

8) Because of the relevant role of different histone modifications in the regulation of the activity of promoters, I proposed to quantify the abundance of histone modifications at the regulatory and the coding regions of the EBER locus.

MATERIALS AND METHODS

Cell lines

The cell lines used in my experiments represented latency type I and latency type III Burkitt’s lymphoma and latency type III lymphoblastoid cell lines. In addition, a
nasopharyngeal carcinoma cell line carrying latency type I EBV genomes was also used. Besides, EBV-negative cell lines of B cell and epithelial origin were examined, too.

**DNA and RNA purification**

Total genomic DNAs were isolated from cultured cells according to a conventional protocol. The RNA samples were isolated from cell cultures, using the Tri-reagent following the recommendations of the manufacturer. DNA and RNA samples were dissolved in ultrapure dH₂O.

**In vitro methylation and transfection**

An EBER harbouring pBS- plasmid was methylated using the MSsI CpG-methyltransferase enzyme. The purified methylated construct and the unmethylated control plasmid were transfected into B and epithelial cell lines using DEAD-dextran or the Fugene transfection agent. The total DNA and RNA of transfected cells were isolated 48 hrs after transfection. DNA samples were analysed for CpG-methylation, RNAs were used to quantify the activity of the control genes (GFP and β-actin) and the EBER1 and EBER2 genes.

**Bisulphite sequencing**

In single-stranded DNA the methyl-cytosines (5mC) resist to bisulphite treatment and will appear as cytosines during PCR amplification, while unmethylated cytosines will be converted to uraciles and PCR-amplified as thymines. As a consequence, by comparing the sequencing results, one can distinguish between methylated and unmethylated cytosines. The bisulphite modification was performed according to Frommer *et al.* (1992) and Clark *et al.* (1994). The resulting modified DNA samples were sequenced as described by Myöhänen *et al.* (1994).

**Reverse transcription**

cDNA was synthesized using the Superscript III reverse transcriptase enzyme and 1 µg of DNase-treated RNAs applying gene-specific oligos, according to the instruction of the manufacturers’.

**Real-time PCR**

Real-time PCR is appropriate for quantitative analysis of DNA or cDNA. I used this method for defining the level of the DNA amounts, gene activity of different genes and the
quantitation of \textit{in vivo} binding of different proteins (CTCF, acetylated and methylated histones) to the EBER region.

LightCycler FastStart DNA Master SYBR Green I kit was used for the amplification. The concentration of samples was determined with the help of a plotted standard curve obtained from a 10x dilution series. Relative concentrations of the samples were calculated using the Light Cycler Software 4.05.

\textbf{DMS \textit{in vivo} fotprinting}

Genomic footprinting was performed essentially as described previously (Niller \textit{et al.} 1995). For the footprinting reaction C666-1 cells were harvested and incubated with dimethyl-sulphate (DMS). The reaction was stopped, and genomic DNA was prepared. Footprinted DNAs were subjected to piperidine treatment to break the strands at the DMS-modified guanines. For visualization of footprints by Ligation Mediated-PCR (LM-PCR), 2 µg of sequenced (DMS-treated naked DNA) or footprinted (\textit{in vivo} DMS-treated) DNA were analyzed as described previously with slight modifications.

\textbf{\textit{In vivo} footprinting}

To find out, which transcription factors are sensitive to the methylation of their recognition sites, I used \textit{in vitro} DNase I footprinting in the case of \textit{in vitro} methylated and unmethylated EBER1 promoter sequences. Comparing the protein binding patterns one can conclude, which interactions are impaired by CpG-methylation. This method was a combination of the Core Footprinting System and the method described by Niller \textit{et al.} (1991).

\textbf{EMSA}

Electromobility shift assay and antibody supershift experiments can be used to find out if a DNA sequence can be bound to a given protein, and to verify suspected protein binding sequences. In this method a short double stranded DNA is incubated together with nuclear proteins in the presence or absence of an antibody, directed against the protein of interest. The protein complexes are separated on an acrylamide gel. If the antibody causes alteration in DNA mobility compared to the antibody negative sample, one can conclude, that the tested protein binds to the sequence of interest \textit{in vitro}. In this work I examined the binding of proteins to the ATF and CTCF binding sequences on the 5’ EBER1 promoter.
The method was performed essentially as described by Niller and Hennighausen, (1990) and Hennighausen and Lubon in Methods in Enzymology.

**Chromatin Immunoprecipitation**

Chromatin Immunoprecipitation (ChIP) is a method used to detect protein-DNA interactions *in vivo*. During this procedure, proteins binding to the DNA are covalently cross-linked via a chemical reaction, and the isolated DNA is broken into pieces between 300-1500 bps in length, thereafter. The protein of interest and the covalently cross-linked DNA are precipitated together by an antibody specific to the protein, and by Protein A beads binding the Fc fragment of the antibody. After these, we can identify the DNA fragments and determine the degree of binding of the examined protein to a given sequence *in vivo*. I examined the *in vivo* binding of the c-Myc and CTCF proteins and different histone modifications within the EBER locus.

ChIP was performed according to the method of Farnham *et al.*, published in 2002 with slight modifications.

**RESULTS**

1. I found **extensive hypomethylation in the whole EBER locus** in different cell lines and cell types.
2. In all of the examined cell lines the EBER1 and EBER2 genes were found to be **highly active** and abundantly transcribed.
3. When I analysed, whether the *in vitro* methylation of the EBER genes inhibits their activity, I found, that **CpG-methylation of the region resulted in a significant decrease of EBER transcription** in transfected cells.
4. The *in vivo* genomic footprinting assay showed, that in the nasopharyngeal carcinoma cell line C666-1, **the protein binding pattern in the EBER region was similar to** the earlier results observed in the case of **B cell lines** (Niller *et al.*, 2003).
5. According to the *in vitro* footprinting experiments with methylated and unmethylated EBER1 promoter sequences, **many important regulatory sites** (c-Myc, Sp1, ATF, Box-A and -B) **showed altered protein-DNA interactions, when the sequence was methylated**. The **TATA binding site was invariably occupied** on both the methylated and unmethylated fragments.
6. Using the electrophoretic mobility shift assay I observed that the unmethylated ATF binding sequence formed a specific complex with nuclear proteins, while such a complex could not be detected using the in vitro methylated sequence. The antibody supershift assay using anti-ATF-2 antibody did not verify, however, the binding of ATF-2 to the sequence. In similar experiments, using the putative CTCF binding site of the EBER1 promoter, no specific interaction could be observed in the presence of nuclear extracts from latency type I or type III BL cells.

7. I could demonstrate the binding of the transcription factor and oncoprotein c-Myc to the 5’ promoter region of the EBER1 in vivo, using chromatin immunoprecipitation. The putative CTCF binding to the upstream regulatory region of EBER1 (Day et al. 2007) was found to be rather week, as in the ChIP experiments only a hardly detectable enrichment could be observed in that region.

8. The ChIP experiments aiming the quantitation of activating histone modifications showed, that the entire EBER locus was enriched in acetylated H3, acetylated H4 and dimethlylated H3K4 modifications. Smaller differences, however, could be observed among the various cell lines. In addition, the complete deficiency of H3K4me2 modification in the 5’ region of EBER1 in the C666-1 cell line, was an unexpected result.

**DISCUSSION**

Fine-mapping of the CpG-methylation of the constitutively transcribed EBER genes showed, that the entire DNA region is extensively hypomethylated in all of the examined cell lines. Based on these results, we suspected a methylation-sensitive transcription of the EBER genes, which could be further supported by transfection experiments: the activity of the EBER genes was inhibited following the methylation of the region. In vitro experiments showed, that CpG-methylation silenced the transcription of the EBER genes by blocking the binding of activating transcription factors.

The in vivo protein binding pattern on the EBER1 gene was described earlier in cell lines with B cell origin. I could complement these results by applying the DMS in vivo footprinting method in case of the C666-1 nasopharyngeal carcinoma cell line. Comparing the
data I concluded, that the protein binding pattern in cell lines of B cell and epithelial origin is identical and the binding of proteins to their recognition sequences is not cell type specific.

I think, that the verification of \textit{in vivo} c-Myc binding to the 5’ sequence of EBER1 was an important discovery. This binding can provide an explanation for the development of Burkitt’s lymphoma, among others, because the elevated c-Myc protein level caused by the translocation characteristic for Burkitt’s lymphomas, can transactivate the antiapoptotic EBER genes. As a consequence, the chance for survival of B cell clones carrying c-myc translocation will increase in the germinal center, raising the probability of Burkitt’s lymphoma development. The formation and maintenance of an open chromatin conformation can also be explained by the binding of c-Myc to the EBER region, as its participation in chromatin remodelling processes is also confirmed. Besides, c-Myc may mediate the attachment of the c-Myc-bound DNA to the nuclear matrix, which can be important in the maintenance of the EBV episomes during cell division.

The open chromatin conformation was further supported by the observed enrichment of activating histone modifications in the entire EBER locus, that is indicative for active chromatin.

\textbf{This thesis is based on the following publications:}


Other publications related to the topic:


Oral presentations in connection with the results


**Ferenc Bánáti**, Anita Koroknai, Mária Takács, Dániel Salamon, Hans Helmut Niller, János Minárovits. Epigenotypes of EBER1 and 2 genes of Epstein-Barr virus in lymphoid and nasopharyngeal carcinoma cell lines. 1st Central European Forum for Microbiology, Keszthely, 2005.