The Epstein-Barr Virus Nuclear Antigens 1 & 5

Study of virus-host cellular protein interactions

UNIVERSITY OF GOTHENBURG

Alma Forsman
2009

Institute of Biomedicine
Department of Clinical Chemistry and Transfusion Medicine

Front page: Epstein-Barr virus budding from a B lymphocyte. Image provided by the Analytical Imaging Facility at the Albert Einstein College of Medicine.

ISBN 978-91-628-7948-8

till Henrik och Elton

The beginning of knowledge is the discovery of something we do not understand - Frank Herbert

ABSTRACT

The Epstein-Barr virus (EBV) is the causative agent or cofactor in the aetiology of several human malignancies such as Burkitt's lymphoma, Hodgkin's disease, nasopharyngeal carcinoma (NPC) and lymphoproliferative disorder in immunocompromised patients. EBV is a lymphotropic γ-herpes virus infecting more than 90 percent of the population worldwide. Following acute infection the virus establishes a life-long latency in resting memory B cells. The virus is remarkable for the efficiency with which it causes proliferation and immortalization of the infected B cells through expression of several latent gene products. All of the viral EBNA proteins have been proposed to play a role in the control of gene expression in the EBV infected lymphoblastoid cell.

The present thesis is mainly focused on further elucidating the molecular mechanisms of the EBNA1 and EBNA5 proteins using proteomic technologies as approach. In paper I we used an improved tandem affinity purification procedure for identification and characterization of factors in the EBNA5 interaction proteome. The majority of the 37 validated interactors could be assigned to one of three groups according to function: protein folding and degradation, pre-mRNA processing, or ribosomal proteins, implicating functional relationships with EBNA5 in these processes. We also showed that EBNA5 is part of high molecular protein complexes, supporting the notion that functional units in the cell are not single proteins but well-structured complexes composed of multiple proteins i.e. modules.

The previously reported repressor activity of EBNA5 was further investigated in paper II. The study identified the novel interactor BAG2 as a major target for the function of EBNA5 via the chaperone-mediated folding and proteasome-degradation pathways. Taken together, the results are consistent with the hypothesis that EBNA5 tune the balance between protein rescue and destruction in a way that disfavour the path of degradation.

The constituents of the large macromolecular complex that initiates transcription from the viral C promoter were investigated in paper III. Using a DNA affinity procedure we showed that the transcription factors E2F1, ARID3A/Bright and Oct-2 binds the Cp as well as EBNA1 and *oriPI*, possibly facilitating long-distance promoter-enhancer interactions.

While the study of genes and proteins continues to be important, looking at isolated components is not enough to understand most biological processes. Modularity has been proposed as a general principle for the molecular architecture of living systems. These assemblies interact with other large protein complexes, thus the proteins are part of a protein-protein interaction network inside the cell. A common feature of these interaction networks is that it contains junctions of proteins that are highly interconnected, also called hubs. Hubs have a tendency of being essential and involved in cancer development. Two central pathways in cancer biology are the Rb- and p53-pathways, which

are targets for both EBNA1 and EBNA5 action. This is consistent with the hypothesis that several viral proteins target the same hubs in the host, which ensures the takeover of the cellular machineries essential for the viral infection and persistence processes, and contribute to the robustness of the viral infectious system.

Keywords: *Epstein-Barr virus, EBNA1, EBNA5, EBNA-LP, cellular network, virus-host protein interactions*

PUBLICATIONS

This thesis is based on the following papers:

I. Identification of intracellular proteins associated with the EBV-encoded nuclear antigen 5 using an efficient TAP procedure and FT-ICR mass spectrometry.
Forsman, A., Rüetschi U., Ekholm, J. and Rymo, L.
Journal of Proteome Research (2008)

II. Epstein-Barr virus nuclear antigen 5 is a multi-functional protein with a possible role in the chaperone-mediated protein folding and ubiquitin-proteasome degradation systems.
Ekholm, J., **Forsman, A.,** Kashuba, E., Andersson, M., Rüetschi, U. and Rymo, L.
In manuscript (2009)

III. E2F1, ARID3A/Bright and Oct-2 factors bind the Epstein-Barr virus C promoter, as well as to EBNA1 and *oriPI*, possibly facilitating long-distance promoter-enhancer interactions.
Boreström, C., **Forsman, A.,** Rüetschi, U. and Rymo, L.
In manuscript (2009)

ABBREVIATIONS

ARID3A	AT-rich interaction domain DNA binding protein 3A, a.k.a. Bright
BAG2	BCL2-associated athanogene 2
BART	*Bam*HI A rightward transcripts
BCR	B cell receptor
BL	Burkitt's lymphoma
bp	base pair
CAT	chloramphenicol acetyl transferase
CCT	cytoplasmic c-terminal tail
cDNA	complementary DNA
ChIP	chromatin immuniprecipitation
CHIP	c terminus of Hsp70-interacting protein
CTAR	c-terminal activation region
co-IP	co-immunoprecipitation
CR	conserved region
Cp	the EBV C promoter
CpG	cytosine and guanine separated by a phosphate
DNA	deoxyribonucleic acid
DS	dyad symmetry, i.e. *oriPII*
E2F	a eukaryotic transcription factor family
EBER	Epstein-Barr virus encoded RNA
EBNA	Epstein-Barr virus nuclear antigen
EBV	Epstein-Barr virus
EMSA	electromobility shift assay
FR	family of repeats, i.e. *oriPI*
FT-ICR	fourier transform ion cyclotron resonance
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GC	germinal center
HD	Hodgkin's disease
HDAC	histone deacetylase
HnRNP	heterogeneous ribonucleoprotein
Hsp	heat shock protein
Ig	immunoglobulin
IM	infectious mononucleosis
IR	internal region
IP	immunoprecipitation
JNK	c-Jun N terminal kinase
kb	kilobase pair
LCL	lymphoblastoid cell line
LMP	latent membrane protein

LUC	luciferase
MAPK	mitogen-activated protein kinase
MAR	matrix association/attachment regions
MCM	minichromosome maintainance complex
MHC	major histocompatability complex
mRNA	messenger RNA
MS	mass spectrometry
MS/MS	tandem mass spectrometry
m/z	mass-to-charge ratio
NE	nuclear extract
NFκB	nuclear factor κB
N_{het}	denoting heterogenity in the number of terminal repeats in different virus isolates
NK cell	natural killer cell
NPC	nasopharyngeal carcinoma
nt	nucleotide
Oct-2	octamer binding factor 2
OriP	origin of replication
OriPI	part of *oriP* with the family of repeats
OriPII	part of *oriP* with the dyad symmetry
ORC	origin of replication complex
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PIC	preinitiation complex
PI3K	phosphatidylinositol 3-kinase pathway
PML NB	promyelocytic leukaemia nuclear body
PTLD	post transplant lymphoproliferative disease
PU.1	Ets family transcription factor
PVDF	polyvinylidene fluoride
Qp	the EBV Q promoter
Q-PCR	quantitative PCR
Rb	retinoblastoma protein
RBP-Jκ	recombination signal binding protein for immunoglobulin kappa J region
RNA	ribonucleic acid
RT-PCR	reverse transcription PCR or real time PCR
TAP	tandem affinity purification
TNF	tumor necrosis factor
TR	terminal repeats
TRAFs	tumor necrosis factor receptor–associated factors
Wp	the EBV W promoter
XLPD	X-linked lymphoproliferative disease

Table of contents

1

EPSTEIN-BARR VIRUS

1.1 The discovery of EBV

The discovery of Epstein-Barr virus was brought about by a young British missionary surgeon named Denis Burkitt. In the late 1950s Denis Burkitt practiced medicine in the University Hospital in Kampala, the capital of Uganda. During the course of his work, he noticed a child suffering from a massive swelling in the four angles of the jaw. Shortly afterwards he observed several other young patients displaying the same aggressive jaw lesions. Intrigued by his observations and the study of the hospitals medical records, together with massive correspondence and travel to almost 60 hospitals in east, central and south of Africa, Burkitt and his colleagues showed that the lymphoma was common in regions of equatorial Africa, also known as the lymphoma belt and that the distribution overlapped with severe endemic malaria (O'Conor, 1961). The strange tumours caught the attention of Anthony Epstein, a young experimental pathologist, at a lecture given by Burkitt at the Middlesex Hospital, London. Epstein postulated the idea that a novel virus could be involved in the development of these tumours and during a couple of years in the early 1960s the Epstein lab regularly received biopsy samples from Burkitt in Uganda. From one of these deliveries the first Burkitt's lymphoma (BL) derived continuous cell line was derived; EB1 (Epstein & Barr, 1964). Interestingly, this cell line was produced due to outgrowth of B cells from one of the biopsies that had been delayed in transport. By examining the cells by electron microscopy, Epstein and his co-workers Yvonne Barr and Bert Achong found virus particles in the EB1 cells and observed that the virus had the typical morphology of a herpes virus. In order to study the virus further, Epstein contacted the Henle lab in Philadelphia, one of the foremost virus diagnostic and research laboratories at that time. Through a joint effort the immunological uniqueness of the virus was demonstrated, as well as its biochemical singularity. It was also Henle & Henle who applied the designation "EBV" after the "EB"

BL-derived cell lines in which the virus was first found. Henle & Henle furthermore discovered that EBV is the causative agent for infectious mononucleosis. The Henle lab also showed after seroepidemiogical research, that the EBV infection is ubiquitously spread worldwide, infecting more that 90% of the adult population.

1.2 Primary infection and persistence

Epstein-Barr Virus (EBV) is an exclusively human, lymphotropic γ-herpes virus that infects more than 90% of the population worldwide (Kieff & Rickinson, 2001). Primary infection usually occurs during early childhood and does not result in any recognized disease. If the primary infection is delayed through adolescence or adulthood, a dramatic immune response occurs which frequently results in infectious mononucleosis, a self-limiting lymphoproliferative disorder (Niederman *et al.*, 1968). EBV is transmitted via saliva and the virus enters the host through the oropharynx. It is believed to initially infect either squamous epithelial cells or resting B cells sitting at or close to the surface of tonsillar epithelia or other lymphoid organs in Waldeyer's ring where crypt structures dip into underlying lymphoid tissue. The viral glycoprotein gp350/220 binds to the receptor CD21 and gp42/gH/gL with the major histocompatibility complex (MHC) class II on the B cell. Upon binding the Epstein-Barr virions partly enters the cell, and partly remains on the B cell surface to enable efficient transfer to CD21-negative epithelial cells (Shannon-Lowe *et al.*, 2006). Thereby, EBV has developed a way to simultaneously access both B cells and squamous epithelial cells of the oropharynx. B cells are currently considered to be the host for EBV infection, while epithelial cells may be seen as helpful enhancers for viral transfer and for establishing latency in B cells (Borza & Hutt-Fletcher, 2002; Wolf *et al.*, 1984).

EBV infection in epithelial cells generally leads to activation of the lytic viral program. However, in B cells the default viral pathway is latency. During latent infection only a limited number of latent proteins are expressed, which comprise six EBV nuclear antigens (EBNA1-6), three latent membrane proteins (LMP1, LMP2A and B) and the two non-polyadenylated RNAs (EBERs). The latent proteins have all been directly implicated in the immortalization process (Kieff & Rickinson, 2001).Upon virus infection all latency genes are expressed and the rapid proliferation of the infected B cells leads to lymphoblast formation and is therefore often referred as the growth program or latency III (Fig. 1). The virus mimics antigen activation of B cells by binding antigens to the surface of naïve B cells which causes the cell to differentiate. At this stage the transformed B cells are efficiently eliminated by the host cytotoxic T cell response (Klein, 1994).

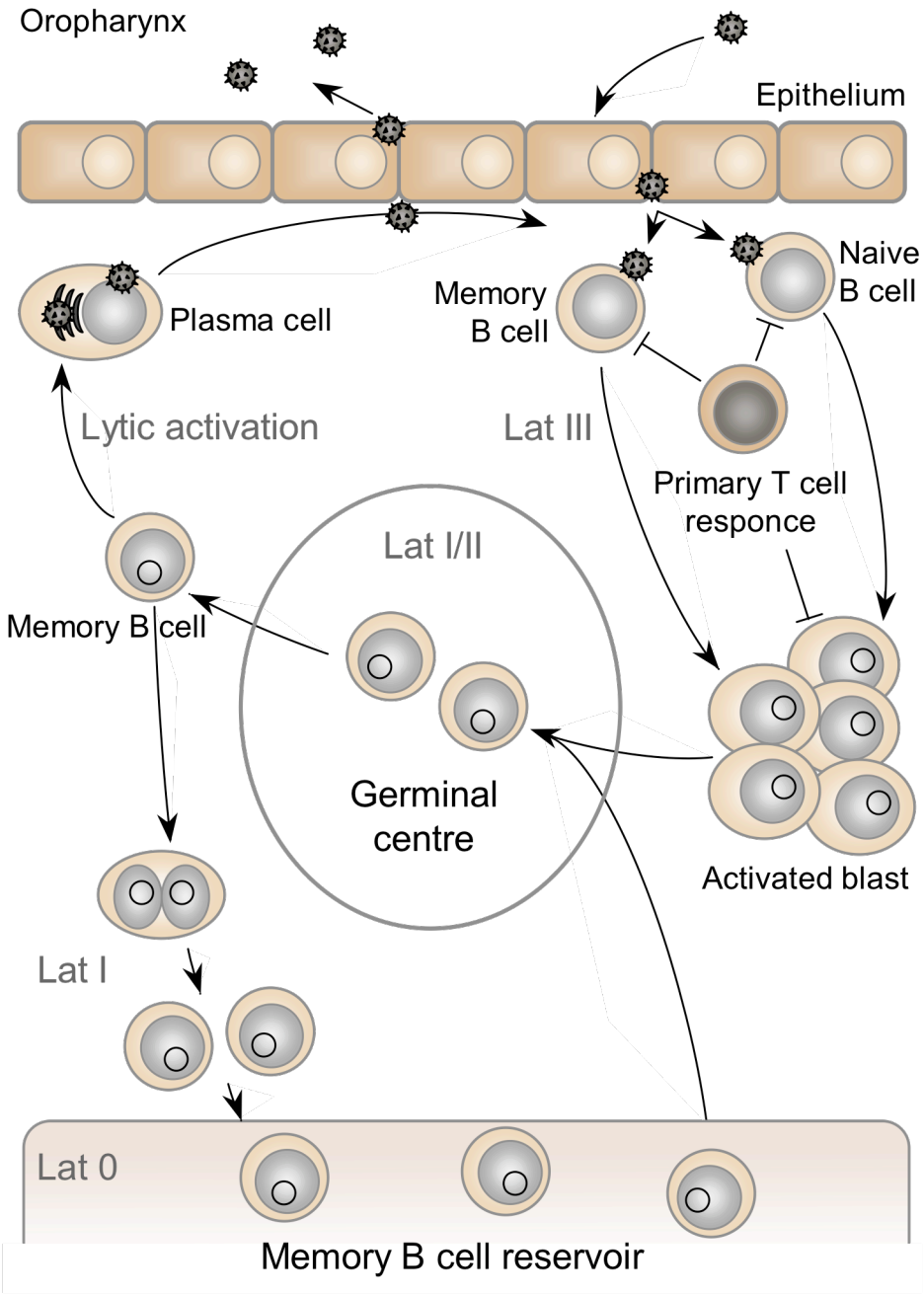

Figure 1. A model for Epstein–Barr virus (EBV) infection and persistence.
In the oropharynx, EBV infects underlying B cells and expresses a full spectrum of latent proteins (latency III, EBNAs1-6, LMPs 1 and 2A) also referred to as the growth program. The virus can thereby drive the activation and proliferation of the infected B cells. At this stage, many of these lymphoblasts are killed by a cytotoxic T cell response. However, some of the cells escape the immune response and undergo a germinal centre (GC) reaction where a more limited set of viral genes are expressed (latency I/II or the default program). The infected GC cells are subsequently rescued and develop into memory B cells, where no EBV gene expression is detected (latency 0 or latency program). Intermittent expression of EBNA1 during replication of these memory B cells allows the virus genome to be distributed to each of the daughter B cells (latency I). As B cells recirculate to the oropharynx, a switch into the EBV lytic cycle may occur, possibly triggered by maturation of B cells into plasma cells, allowing for virus replication, shedding into saliva and transmission both to new hosts and to previously uninfected B cells within the same host. Adapted from (Young & Rickinson, 2004).

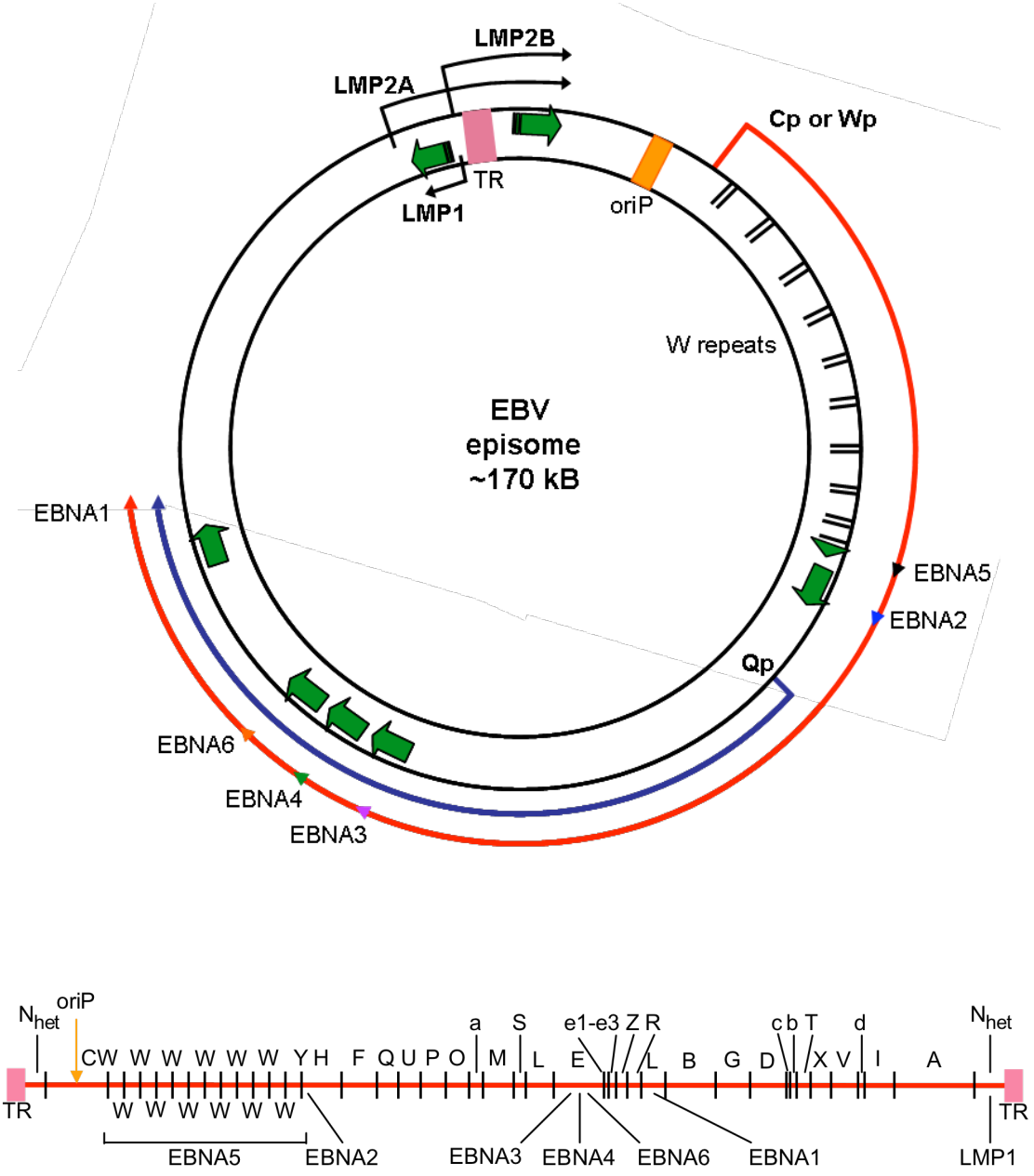

Figure 2. The Epstein-Barr virus genome
The EBV genome shown as a double-stranded DNA episome or its linear form down under. The linear viral DNA circularizes at the terminal repeats (TR-pink box) upon cell entry. The origin of plasmid replication (*oriP*) is shown in orange. The six latent antigens (EBNAs1-6) are encoded by individual mRNAs generated by differential splicing of the same long primary transcript initiated from either the W promoter (Wp) or the C promoter (Cp) in cells with a latency III expression pattern. The Q promoter (Qp) drives the expression of EBNA1 in cells with a latency I and II expression pattern. The exons of the latent membrane proteins (LMP2s) A and B are located on either side of the terminal repeat (TR) regions, therefore their expression require circularization of the viral DNA. The LMP1 protein is transcribed in a leftward direction downstream of the TR. Adapted from (Murray & Young, 2001).

The cells that escape the immune response migrate into a follicle to form a germinal centre (GC) and it is proposed that EBV-activated B cells also undergo a germinal centre reaction where a more restricted set of EBV genes are expressed which are designated latency II. The expression pattern of latency II, or default program as it is also known as, exhibits EBNA1, LMP1 and LMP2A as well as EBERs expression. The infected GC cells are subsequently rescued and develop into memory B cells, where no EBV antigens are detected due to absence of antigen binding and T cell signals (referred to as the latency program or latency 0). EBV infected memory B cells have no latent gene expression but express EBNA1 during B cell proliferation to insure that the EBV genome is replicated and passed on to the daughter cells (EBNA1 only program or latency I) (Kieff & Rickinson, 2001). The lack of immunogenic EBV antigens and immune response are the reason that in most immunocompetent individuals the virus is harboured for life within these latently infected resting memory B cells, without causing any symptoms (Klein, 1994; Thorley-Lawson *et al.*, 1996). As recirculation of B cells to the oropharynx occurs, a switch into the EBV lytic cycle may be triggered, possibly by maturation of B cells into plasma cells. Reactivation of the lytic cycle results in new virus release that may lead to shedding of virions into saliva and transmission to both new hosts and to previously uninfected B cells in the same host (Kang *et al.*, 2001). The fact that EBV can access and persist in memory B cells without causing disease is the key to its success in infecting most of the human population and is the evidence of its co-evolution with man.

1.3 Gene regulation

Significant efforts have been made to unravel the regulation of EBV gene expression since latent expression programs are central in EBV biology and pathogenesis. The virus genome is a linear double-stranded DNA that is approximately 172 kilobase pairs (kb) which encodes approximately 85 proteins (Figure 2). After infection, a repeated DNA sequence (the terminal repeats, TR) located at the ends of the linear form, mediates the formation of a circular episome. Usually each latently infected cell contains 10-20 copies of the genome (Farrell, 2005).

The EBV DNA also contains an unusually large tandem repeat, known as the major internal repeat or IR1. The IR1 is functionally significant as it contains the W promoter which is the first promoter to be activated during primary infection of B lymphocytes. The W promoter (Wp) drives the expression of all EBNA proteins. As Wp is located in the IR1 sequence, it is therefore present in multiple copies (Figure 2). During primary infection, EBNA2 and EBNA5 are the first

viral proteins to be expressed from the strong Wp and the expression of these proteins are sufficient to drive the cells into early G_1 phase of the cell cycle. Wp driven EBNA2 protein functions as an transcriptional activator which is recruited to the C promoter (Cp). Within a few days the Cp becomes the dominant latent promoter and drives the expression of the EBNAs (Woisetschlaeger *et al.*, 1991). Cp is located in the unique region just upstream of IR1. The EBNA genes belong to the same transcription unit and the different mRNAs are generated by alternative splicing from a large primary transcript originating from Wp or Cp. Each EBNA transcript contains multiple copies of the W exons, which in most cases splice to two or three small exons. Alternative splicing dictates the 3´coding exon present in mature EBNA gene transcripts (Speck, 2005).

EBNA1 is transcribed relatively early after primary infection and the binding of EBNA1 to the origin of replication (*oriP*) activates both Wp and Cp initiated transcription. An enhancer element consisting of three domains (UAS1, 2 and 3) located upstream of Wp has been shown to upreglate both Wp and Cp activity (Speck, 2005).

Cp is tightly regulated by the EBNA gene products. While EBNA1 and EBNA2 upregulates Cp, the EBNA3, 4 and 6 proteins are able to interfere with EBNA-mediated transactivation. EBNA5 augments EBNA2 activation showing the exquisite autoregulation of Cp by all latent gene products. In the absence of Cp activity, Wp can fully compensate for activation of the EBNAs (Tao *et al.*, 2006; Woisetschlaeger *et al.*, 1990).

In latency I and II, Wp and Cp are not active and the only EBNA protein expressed is EBNA1. In these restricted latency phenotypes, EBNA1 gene expression is driven by the TATA-less promoter Q promoter (Qp). Qp-initiated EBNA1 gene transcripts differ from those initiated by Wp/Cp in that they contain a short 5´exon, designated Q. The Qp is a cell cycle regulated promoter and EBNA1 is a potent autoregulator of its activity. This feature may lead to the silencing of the Qp in latency 0 B lymphocytes, where no latent gene expression is observed (Schaefer *et al.*, 1997). The regions around Qp have been found hypomethylated while the Cp and Wp are hypermethylated in tumour cell lines, indicating a mechanism where methylation plays an important role in EBV gene regulation at different stages of its life cycle (Tao *et al.*, 2006).

Contrary to the EBNA gene transcription, the organization of the LMP genes is relatively simple. There are three LMP gene products expressed in latency II and III – LMP1, LMP2A and LMP2B. The LMP1 protein is encoded by 3 closely spaced exons located at the right-hand end of the viral genome (see Figure 2), in the unique sequence near the terminal repeats (Figure 2). The ED-L1 promoter drives the expression of LMP1, and is also referred to as the LMP1 regulatory sequense (LRS) (Fahraeus *et al.*, 1990). LMP2A and LMP2B gene expression is driven by independent promoters located in the unique region and their 3´exons are located at the left-hand end of the viral genome. Thus, the

expression of LMP2 gene products requires circularization of the viral genome (Speck, 2005). The promoters regulating the LMP gene expression are highly dependent on expression of EBNA2, like Cp (reviewed in(Kieff & Rickinson, 2001)). The patterns of gene expression in different latency types are summarized in Table 1.

Table 1. EBV latent gene expression in different latency types (Cohen, 2005)

Latency	*EBNA1*	*EBNA2-6*	*LMP1*	*LMP2*
0	-	-	-	-
I	+	-	-	-
II	+	-	+	+
III	+	+	+	+

1.4 Latent protein function

All of the EBNA proteins have been proposed to play a role in the control of gene expression in the EBV infected lymphoblastoid cell.

EBNA1

The EBNA1 protein is a DNA binding nuclear phosphoprotein involved in viral replication and is essential for maintenance of the viral episome (Yates *et al.*, 1984). EBNA1 binds the origin of replication (*oriP*) as a homodimer and the main role of the EBNA1/*oriP* complex is to ensure stable replication by recruiting the cellular origin recognition complex (ORC) and the minichromosome maintenance (MCM) complex, which enables viral transfer to each daughter cell parallel with mitosis (Mackey & Sugden, 1999a). EBNA1 is a transcriptional transactivator, upregulating both the Cp and LMP1 promoter (Kieff & Rickinson, 2001). The protein also interacts with two sites downstream of Qp to negatively regulate its own transcription (Murray & Young, 2001).

The EBNA1 protein contains a repeated peptide sequence (Gly-Gly-Ala) which acts as an inhibitor of antigen processing (Levitskaya *et al.*, 1995). By inhibiting the ubiquitin-proteasome dependent degradation, EBNA1 prevents endogenous antigen presentation through the MHC I class pathway and immune detection. This could explain why EBV positive cells in which viral antigen

presentation is limited to EBNA1 can escape the host immune surveillance system. Some investigations have also suggested that EBNA1 may have a direct role in oncogenesis (Wilson *et al.*, 1996) and inhibit p53 induced apoptosis (Kennedy *et al.*, 2003).

EBNA2

EBNA2 has a crucial role in the immortalization process, functioning as a master transcription factor of all the nine viral proteins needed for transformation, and a limited number of cellular target genes (Cohen *et al.*, 1989). In primary infection of B lymphocytes, EBNA2 together with EBNA5 are the only proteins needed to advance the cells to early G_1 phase of the cell cycle (Sinclair & Farrell, 1995). EBNA2 is a nuclear phosphoprotein and acts as a transactivator of Cp, which promotes the switch from Wp to Cp observed in early B cell infection. EBNA2 is also a specific transcriptional activator of both cellular and viral genes, upregulating the B cell antigens CD21 and CD23, as well as LMP1 and LMP2, and the *c-myc* and *c-fgr* proto-oncogenes (Kieff & Rickinson, 2001). EBNA2 is recruited to promoters via its interaction with the ubiquitous DNA-binding protein RBP-Jκ as well as other DNA binding proteins such as PU.1, AUF1, DP103, the POU domain protein and members of the ATF/CREB protein family (Zetterberg, 2005). The *c-myc* oncogene is also a target of EBNA2 and this effect is likely to be important for EBV-induced B cell proliferation (Kaiser *et al.*, 1999).

EBNA3, 4 and 6

The EBNA3, 4 and 6 (also referred to as EBNA3A, 3B and 3C - the EBNA3 family) proteins are encoded by three genes adjacent to each other. EBNA3 and EBNA6 are essential for B cell transformation whilst EBNA4 appears to be dispensable (Robertson *et al.*, 1996). EBNA6 can induce upregulation of both cellular and viral genes but also acts as a repressor of Cp activity. It has also been proposed that EBNA6 interacts with the retinoblastoma protein pRb, to promote transformation (Allday & Farrell, 1994; Radkov *et al.*, 1997). Recently, the EBNA6 oncoprotein was shown to stabilize the cellular oncoprotein c-Myc and to repress the transcriptional activity of p53 which suggests that EBNA6 may change the pattern of cellular gene expression in EBV associated human cancers (Bajaj *et al.*, 2008; Yi *et al.*, 2009).

Binding of each one of these three EBNAs to the transcription factor RBP-Jκ disrupts its binding to EBNA2 thus repressing EBNA2 mediated transactivation (Robertson *et al.*, 1996). Thus, the EBNA2 and EBNA3 family work together to fine-tune the activity of RBP-Jκ, thereby regulating the expression of cellular

and viral promoters. EBNA6 has also been shown to interact with a histone deacetylase, HDAC1, which contributes to the transcriptional repression of Cp via RBP-Jκ (Radkov *et al.*, 1999).

In addition, these three proteins also have a major role in the induction of the immune rejection response that makes mononucleosis a self-limiting disease (Rickinson & Kieff, 2001).

EBNA5

EBNA5 is encoded by the 5' proximal part of the EBNA mRNAs and is also referred to as EBNA-LP (EBNA leader protein). The EBNA5 polypeptide consists of a multiple-repeat domain encoded by repeating W_1 and W_2 exons in the major internal repeat (IR1), and a unique carboxy-terminal encoded by the Y_1 and Y_2 exons, which are located downstream of the IR1 (Sample *et al.*, 1986; Speck *et al.*, 1986). Several isoforms of EBNA5 with different numbers of W_1W_2 repeats generated by alternative splicing can be detected during primary infection (Nitsche *et al.*, 1997). Still, most established lymphoblastoid cell lines (LCLs) express one major and a few minor species.

EBNA5 is not essential for EBV mediated B cell transformation but greatly enhances the efficiency of the process (Mannick *et al.*, 1991). The protein is phosphorylated at multiple sites, one of which has been shown to be essential for EBV-induced immortalization of B-lymphocytes and to be involved in the tight association between EBNA5 and the nuclear matrix (Yokoyama *et al.*, 2001). The so far best characterized function of EBNA5 is its cooperation with EBNA2 in upregulating transcriptional targets of EBNA2, including the LMP1 promoter (Harada & Kieff, 1997; Masciarelli *et al.*, 2002; McCann *et al.*, 2001). EBNA5 together with EBNA2 can also activate the expression of cyclin D2, thereby inducing the G_0 to G_1 transition (Sinclair *et al.*, 1994).

Furthermore, in addition of being a positive transcription factor, at high but biologically relevant concentrations EBNA5 has also been shown to act as a repressor of gene expression in a dose dependent manner (Dufva *et al.*, 2002). The repressive effect is partly explained as a consequence of RNA processing by inhibition of 3′-end cleavage and polyadenylation of certain cellular and viral pre-mRNAs. It was also excluded that repression occurred as a consequence of apoptosis induced by EBNA5. This is consistent with the notion that a regulatory mechanism exists in the cell that confers specificity to the selection by EBNA5 of target genes for repression.

LMP1

LMP1 is the major EBV oncogene. LMP1 is essential for transformation of B lymphocytes and LMP1-mediated signalling is critical for EBV-associated pathogenesis. LMP1 is an integral membrane protein that contains a short cytoplasmic N-terminal region, six transmembrane segments, and a cytoplasmic C-terminal tail (CCT). The transmembrane segments oligomerize within the membrane and mediate constitutive activation. The CCT contains two subdomains that have been implicated in LMP1 signalling, C-terminal activation region (CTAR) 1 and 2, and these two regions play distinct, yet overlapping roles in EBV-associated lymphoproliferation (Huen *et al.*, 1995; Xie *et al.*, 2004).

At least four signalling pathways have been implicated in the function of LMP1, namely the Nuclear Factor κB (NF-κB), c-Jun N-terminal kinase (JNK), p38 kinase and phosphatidylinositol 3-kinase (PI3K) pathways (Dawson *et al.*, 2003).

LMP1 mimics signalling events and effector functions of CD40 in B lymphocytes in many ways, but not all. This includes activation of NF-κB through tumour necrosis receptor associated factors (TRAFs). Thus, LMP1 can activate the same signalling pathways as the TNF receptor (Mosialos *et al.*, 1995) and this association leads to signalling through NF-κB and JNK pathways to activate expression of anti-apoptotic genes.

LMP2A and 2B

The LMP2 gene encodes two distinct proteins, LMP2A and LMP2B. The structure of both proteins are similar, 12 transmembrane domains and a cytoplasmic C-terminus. In addition, LMP2A also has a cytoplasmic N-terminal domain. Neither LMP2A nor LMP2B are required for B cell transformation. LMP2B lacks the kinase interaction domain but is thought to modulate the aggregating effect of LMP2A and thereby regulate LMP2A function (Longnecker, 2000; Rovedo & Longnecker, 2007).

The LMP2A protein is a substrate for the src family of tyrosine kinases, and associates with a cellular phosphotyrosine protein (Longnecker *et al.*, 1991). LMP2A mediates blocking of the signal transduction through the B cell receptor (BCR), thus inhibiting B cell development by blocking calcium mobilization and tyrosine phosphorylation (Fruehling & Longnecker, 1997; Miller *et al.*, 1993). In absence of normal BCR signalling, LMP2A can drive the proliferation and survival of B cells, supporting the role of LMP2A as a modifier of normal growth program of B cell development in order to favour maintenance of EBV latency and to prevent inappropriate activation of the lytic cycle (Caldwell *et al.*, 1998).

1.5 Associated diseases

Usually EBV is harboured for life without any symptoms. EBV pathogenesis however does arise when the immune system is compromised. The list of benign and malignant diseases that are associated with EBV is still growing. The most common ones are listed below. For a detailed review of EBV-associated diseases, see reference (Kieff & Rickinson, 2001).

Infectious Mononucleosis

Infectious mononucleosis (IM) is a self-limiting lymphoproliferative disorder that might arise if primary EBV infection is delayed through adolescence or adulthood. The usual clinical symptoms are high fever, headache, chills and sweats, fatigue and a severe soar throat. The convalescence can be very long, ranging from several days to several weeks. Patients with acute IM shed high titres of infectious virus in the throat from lytic infection in the oropharynx (Vetsika & Callan, 2004).

Burkitt's lymphoma

Burkitt's lymphoma (BL) can be divided into three main clinical variants: the endemic, the sporadic and the immunodeficiency-associated variants. The high-incidence endemic form of BL is the most common childhood cancer in equatorial Africa. Annually ~5-10 cases per 100'000 children suffer from the disease. Also children in Papua New Guinea are affected and the common denominator in both areas is that malaria is holoendemic. EBV is present in all endemic BLs and in up to 85% of cases in areas of intermediate incidence, such as Brazil and North Africa. In the sporadic variant (also called non-African BL) only 15% of the low-incidence tumours of children in the developed world carry the EBV genome. BL is also quite common among immunodeficient adults such as HIV carriers in the developed world. Some 30-40% of these tumours are EBV-associated (Kieff & Rickinson, 2001) and it has been proposed that BL can be an initial manifestation of AIDS. All BL lymphomas carry one of three characteristic chromosomal translocations that place the MYC gene under the control of the Ig heavy chain promoter or of one of the light chain loci. This is the key factor in the pathogenesis of BL (Kovalchuk *et al.*, 2000; Li *et al.*, 2003; Polack *et al.*, 1996).

Nasopharyngeal carcinoma

Nasopharyngeal carcinoma (NPC) is an epithelial tumour that is characterized by marked geographic and population differences in incidence. This is by far the most common malignant tumour of the nasopharynx. NPC is uncommon in western populations, occurring at a rate of 0.1 cases per 100'000 people per year (Buell, 1974). However, in southern China and Southeast Asia it accounts for nearly 20% of all cancers, with an incidence of 100 cases per 100'000 people per year in some regions. The tumour is also quite frequent in the Inuit population as well as the population of Mediterranean Africa (Nielsen *et al.*, 1977). Interestingly, while NPC is seen primarily in middle-aged persons in Asia, a high proportion of African cases appear in children. The EBV genome is consistently detected in all undifferentiated NPC, regardless of geographic distribution or racial background (Raab-Traub *et al.*, 1987). In spite of the strong association with EBV, the aetiology of the disease is poorly understood. The distinct geographical distribution of this malignancy suggests that environmental factors as well as disease susceptibility genes are involved in the oncogenic process. Several traditional foods such as Cantonese Chinese salted fish, and Harissa, a common ingredient in North African cooking, contain compounds that have been shown to reactivate latent EBV infected B cells in vitro (Shao *et al.*, 1988).

Hodgkin's disease

Hodgkin's disease (HD) is an unusual lymphoma of the human lymphatic system. Originally, epidemiological studies suggested a possible role for EBV in the aetiology of HD, as IM is a predisposing factor. However, EBV is only present in 30-50% of the HD cases. HD is characterized by the presence of a special type of lymphatic cells called Reed-Sternberg (RS) cells, which are of B-cell origin. The malignant Reed-Sternberg cells account for only 1-2% of the total tumour mass, which the virus is localized to and is clonal. The EBV association seems to be age-related as it is more frequent in childhood, in older adults (>45 years) and in mixed cellularity cases. The survival rate is generally 90% or higher when the disease is detected during early stages, making it one of the more curable forms of cancer.

Lymphoproliferative disorders in immunodeficiencies

The ability of EBV to cause malignant diseases is most clearly indicated by the development of lymphoproliferative disorders and immunoblastic B-cell lymphomas in immunocompromised patients. The immunodeficiency can either

be congenital or acquired including patients with the rare genetic disease X-linked lymphoproliferative disease (XLPD), Wiskott-Aldrich syndrome, recipients of organ or bone marrow transplants and AIDS-patients. All these individuals have an impaired cytotoxic T-cell response and are unable to to control the proliferation of EBV-infected cells. Most post-transplant lymphoproliferative disorders (PTLDs) arise in the first year after transplantation, when the immunosuppression is the most severe. The lesions are mostly of B cell origin and range from atypical polyclonal B cell proliferations, which often regress following withdrawal or reduction of immunosuppression, to aggressive non-Hodgkin lymphomas, which generally do not resolve following immune reconstitution (Murray & Young, 2001).

2

PROTEIN INTERACTIONS

2.1 Molecular networks

While the study of genes and proteins continues to be important, looking at isolated components is not enough to understand most biological processes. For instance, the analysis of the signalling pathway involving the p53 tumour-suppressor gene is more important than looking at the gene only (Vogelstein *et al.*, 2000). This is further supported by the notion that a combined attack on genes connected to p53 caused more severe effects than the removal of the gene itself (Franklin *et al.*, 2000).

A protein complex is defined as a group consisting of two or more proteins. Protein complexes are a form of quaternary structures and the proteins in a protein complex are linked by both covalent and non-covalent protein-protein interactions. Identifying the components of protein complexes is a vital aspect of cell biology because cellular processes are often carried out by highly dynamic and stable protein complexes and their characterization often provides insight into their function. The formation of a protein complex often serves to activate or inhibit one or more of the members of the complex and in this way protein complex formation can be similar to phosphorylation. Furthermore, individual proteins can participate in the formation of a variety of different protein complexes. Different complexes perform different functions, and the same complex can perform very different functions depending on a variety of factors, for example in which cellular compartment the complex resides and at which stage of the cell cycle.

We know that nearly all major processes in the cell are carried out by multi-protein complexes consisting of five or more proteins. These assemblies interact with other large protein complexes, thus the proteins are part of a protein-protein interaction network inside the cell (Alberts, 1998; Spirin & Mirny, 2003). The protein-protein network is composed of proteins represented as nodes and the links or edges representing their interactions (see Figure 3). Similarly, metabolic

pathways can be represented as networks formed by metabolites, reactions and enzymes connected by two types of relationships, mass flow and catalytic regulation. Furthermore, transcriptional regulation can be represented by complex networks, where nodes represent genes and directed links denote regulatory effects on the target genes. These interactions are essential to keep the molecular systems of living cells working properly. A common feature of these interaction networks is that they contain junctions of proteins that are highly interconnected, also called hubs. Investigation of 43 organisms has revealed that only a minor fraction of all proteins in the cell are hubs, meaning that few hubs concentrate a high number of protein-protein connections (Jeong *et al.*, 2000). The topology of these networks is extremely heterogeneous and described as scale-free, which is in contradiction to the classical random network theory, introduced previously by Erdös & Rényi in the 1960s (see Figure 3). The hubs tie together most of the other less connected proteins (nodes) with only one or two interaction partners. In addition, new nodes are preferentially attached to already existing nodes, a property that is also thought to characterize the evolution of biological systems (Hartwell *et al.*, 1999). The resulting network is very dense, and the distance between any two proteins is generally short. Due to the small distances between any two proteins in these networks, they are referred to as small-world networks. This general topology can also be observed in the structure of the World Wide Web, citation networks and metabolic networks for example (Jeong *et al.*, 2000). The protein-protein interaction networks are highly dynamic, allowing for rapid changes in the proteome to accommodate environmental changes (Hartwell *et al.*, 1999).

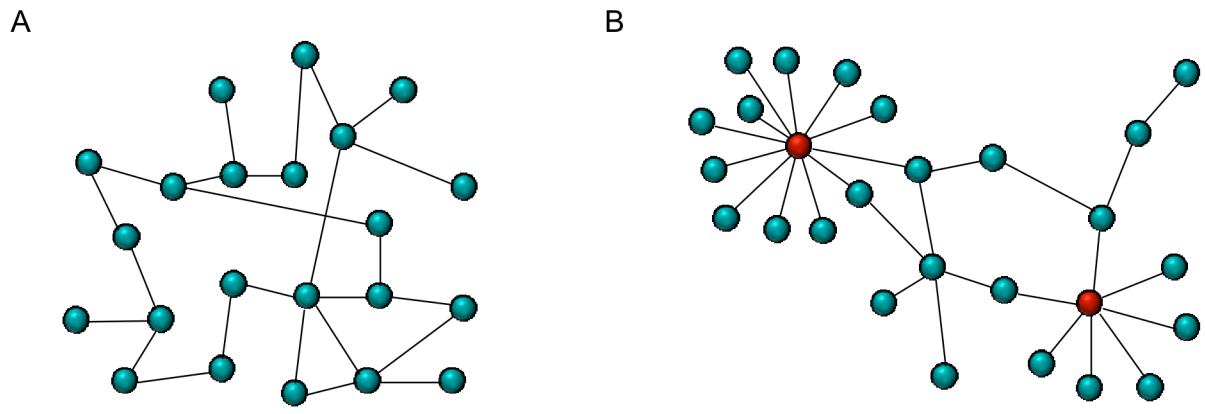

Figure 3. Network topology.
A. The random network is homogeneous: most nodes (visualized as green dots) have approximately the same number of links. B. The scale-free network is heterogeneous: the majority of the nodes have one or two links (also commonly referred as edges) but a few nodes have a large number of links i.e. hubs (visualized as red dots), guaranteeing that the system is fully connected.

The multi-protein complexes in these networks are cornerstones of most biological processes and together they form various types of molecular machineries or modules. For instance, protein-protein interactions, related to cellular communication by signal transduction, activate or repress the transcription of genes, changing the molecular composition of the cell. These separate modules are the 'building blocks' of the biological systems, working together to shape the phenotypic pattern of the cells and organisms. Spirin and co-workers have discovered that modules are densely connected within themselves but sparsely connected with other modules of the network meaning that cellular entities with similar functions interact mainly with each other (Spirin & Mirny, 2003). They also divided these modules into two subgroups using experimental data and functional annotations of the genes; (a) protein complexes or (b) functional modules. Protein complexes are groups of proteins that interact with each other at the same time and place, forming a single multimolecular machinery such as the spliceosome, transcription factor complexes, RNA splicing and polyadenylation machinery etc. In contrast, functional modules consist of proteins participating in a particular process while binding to each other at different times and places, like the MAP signalling cascade or the CDK/cyclin module responsible for cell-cycle progression.

Recently, two comprehensive studies of the *Saccharomyces Cerevisiae* protein-protein interactions were performed (Gavin *et al.*, 2006; Krogan *et al.*, 2006). The results showed that half of the proteome was to be in complex and that there was a substantial protein sharing between the complexes. These experimental results constitute a strong support for the modular organization principle which has been further strengthened in several other studies with the same approach (Calderwood *et al.*, 2007; Uetz *et al.*, 2006; Zhang *et al.*, 2009). The modules arise through evolution as it is easier to build a complex system from stable functioning subsystems and modularity also contributes to the robustness of the system against random component failure (Albert *et al.*, 2000; Jeong *et al.*, 2001). In contrast, human cellular protein networks have a high sensitivity to targeted attacks on the highly interconnected hubs. Removal of the hubs one by one, quickly results in a total collapse of the overall network (Albert *et al.*, 2000). Notably, hubs have a tendency of being essential and have also been suggested to have a higher probability of being involved in cancer development (Jonsson & Bates, 2006). With hubs seemingly having such a central role in interaction networks and the biology of the organisms, there is now an increased effort to identify the hub proteins and to discover in what sense these proteins differ from other proteins. Analysis of network structure can be useful to determine the importance of these proteins and to predict their functions.

2.2 Virus-host protein interactions

Infectious diseases caused by viral agents are still one of the major health threats all over the world. Viruses are among the simplest pathogens, and thus require the host cell machinery for replication and survival. Viruses have evolved to regulate and exploit these machineries in numerous ways. The host has likewise evolved complex defensive mechanisms to protect itself against pathogenic threats, but the host immune system includes several trade-offs that can be exploited by pathogens and induces undesirable inflammatory reactions. Molecular understanding of the interactions between the host and the virus is a key challenge in modern virology research, and can provide valuable insights into the basic mechanism that regulate normal and pathological cellular behaviour.

Cellular protein networks are not static and the robustness of the systems may change dynamically according to various factors like tissue, cell type or changes in the cellular environment such as viral infections (Kitano, 2007). The pathogenesis and survival of the viruses depend on a complex interplay between viral and host proteins, both acting in a complex virus-host protein-protein interaction network. A comprehensive analysis of the virus protein interaction networks of the Kaposi's sarcoma-associated herpes virus and the Varicella-zoster virus, shows a network topology of single, highly coupled modules, whereas cellular networks are organized into separate functional submodules (Uetz *et al.*, 2006). Due to this unusual topology, viral networks are thought to be more resistant to deliberate attacks as the network remains more stable even though the most highly connected proteins are removed. The analysis of the combined virus-host protein interaction network revealed that the viral network adopts cellular network properties to a large extent. This interplay may lead to new system properties that represent specific features of the viral pathogenesis as the host protein-protein network is altered by the invading virus proteins. The creation of novel interactions and modifications or the destruction of others results in a network topology that favours the excessive amount of virus production in a stressed host cell network.

Recently, a comprehensive mapping of the EBV virus-host protein interaction network was performed (Calderwood *et al.*, 2007). The result revealed that the human proteins targeted by EBV usually are highly interconnected proteins, hubs, and that this might be an efficient mechanism for EBV in order to reorganize the cellular processes. Last year, the first release of VirHostNet (Virus-Host network) (Navratil *et al.*, 2009) was launched, and when examining the virus-host network of EBV one can distinguish two highly interconnected viral proteins – EBNA1 and EBNA5 – both of which will be the focus of the remaining of this thesis.

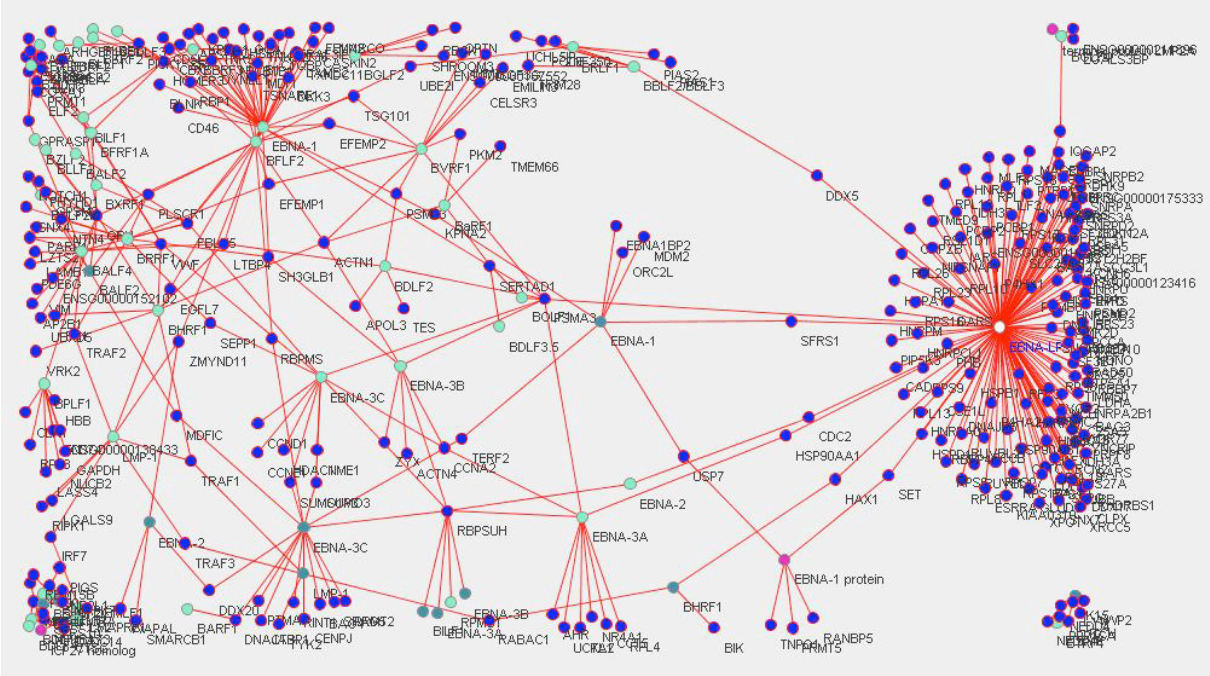

Figure 4. EBV-host interactome network.
Map of the EBV-host protein interactions derived from the VirHostNet visualizator. Cellular proteins are shown as blue nodes and virus proteins have different nodes colours according to viral taxonomy.

3

THE PRESENT INVESTIGATION

3.1 Aim of the study

Overall aim

Throughout the life cycle of EBV a multitude of protein-protein interactions take place, ensuring the maintenance of the viral genome and the production of new infectious viral particles. Gaining a better understanding of how the viral proteins interact with each other and with the host will help elucidate the details of the cellular targets in regulation of lytic and latent virus infection.

The EBV nuclear antigens 1 and 5 are two viral proteins of high interest in EBV biology. The EBNA1 protein is essential for viral episome maintenance in the infected host due to its function as the DNA binding protein that recruits the cellular complexes needed for viral replication. In addition, EBNA1 bound to the EBV *oriP* family of repeats functions as a transcriptional enhancer of neighbouring promoters, including the C promoter. The mechanism for this interaction is however not completely understood at the molecular level. The so far best characterized function of EBNA5 is its cooperation with EBNA2 in the transactivation of the LMP1 promoter. A considerable amount of evidence has accumulated indicating that EBNA5 is a multifunctional protein but its precise role in EBV biology has yet to be defined.

Paper I

- To identify and characterize factors in the EBNA5 interaction proteome using an improved tandem affinity purification procedure coupled to high resolution mass spectrometry analysis.

Paper II

- To further elucidate the cellular and molecular mechanisms underlying the repressor function of EBNA5.

Paper III

- To identify factors that bind to the Cp core and to EBNA1 and *oriPI*, possibly facilitating long-distance promoter-enhancer interactions.

3.2 Methods and Materials

The purpose of this section is to provide an overview of the methods used in the work of the thesis. Detailed descriptions are available in the papers and manuscripts at the end of the thesis.

Cell lines

The cell lines used in this study were chosen based on their phenotype, transfectability and their endogenous expression of latent EBV genes (see table 2).

Table 2. Cell lines used in this thesis

Cell line	*EBV*	*Latency*	*Comment*
DG75	No	-	Burkitt's lymphoma cell line
BJAB	No	-	Burkitt's lymphoma cell line
Rael	Yes	I	Burkitt's lymphoma cell line
P3HR1	Yes	II/III	EBNA2 deficient Burkitt's lymphoma cell line
Raji	Yes	II/III	Expresses EBNA2 but lacks EBNA6 and is defective in lytic replication
B95.8	Yes	III	Marmoset derived, infected with the B95.8 strain
IB4	Yes	III	Placenta lymphocytes infected with the B95.8 strain
Cbc-Rael	Yes	III	Cord blood cells infected with the Rael strain
B95.8-LCL	Yes	III	B lymphocytes infected with the B95.8 strain
WW1-LCL	Yes	III	B lymphocytes infected with the QIMR-Wil virus isolate
HEK293A	No	-	Human embryonic kidney monolayer epithelial cell line of neuroendothelial origin
Molt4	No	-	T cell line from acute lymphoblastic leukaemia
MCF7	No	-	Human breast adenocarcinoma epithelial cell line

Molecular cloning and sequencing

All manipulations involved in constructing expression plasmids were carried out by standard procedures (Sambrook & Russel, 2001). All plasmids made were

verified by dideoxy sequencing, utilizing the ABI Prism®Big Dye™Terminator Cycle Sequencing Ready Reaction Kits (Applied Biosystems). Sequences were compared to the published EBV-genome strain B95-8 (GenBank V01555).

Transfection and reporter gene assays

Transfection is the process of introducing DNA into a cell by non-viral methods. This technique is a routine tool for studying gene structure and function in mammalian cells. In this thesis two different transfection techniques have been used: electroporation (suspension cells) and lipid-mediated transfection (adherent cells).

Reporter genes have become an invaluable tool in studies of gene expression. The purpose of the reporter gene assay is to measure the regulatory potential of a promoter of interest. This can be done by linking the promoter sequence to an easily detectable reporter gene such as that encoding for the firefly luciferase or chloramphenicol acetyl transferace. The promoter region and the transcription factors will regulate the expression of the reporter gene. We have used the chloramphenicol acetyltransferase (CAT) and luciferase (LUC) reporter gene assays in our studies as an indirect measurement of the promoter activity. CAT is a bacterial enzyme that catalyzes the transfer of acetyl groups from acetyl coenzyme A to radioactively labelled chloramphenicol.

In the luciferase assay, light is emitted when the reporter luciferase acts on the appropriate luciferin substrate and the amount of emitted light is correlated to the presence of reporter protein. The substrate luciferin is not toxic to mammalian cells and as luciferase is not naturally present in target cells the assay is virtually background-free. Therefore, the reporter system is highly sensitive and ideal for detecting low-level gene expression.

Immunoblot analysis

Immunoblot (or western blot) is an analytical technique for detecting specific proteins in a given cell extract or tissue homogenate. Gel electrophoresis is used to separate native or denatured proteins by the size of the polypeptide (denaturing conditions) or by the 3D structure of the protein (native/non-denaturing conditions). The proteins are then transferred to a membrane (typically nitrocellulose or PVDF), where they are detected using antibodies specific to the target protein. In this thesis we have used whole cell extracts from continually growing cells or the lysates from the reporter assays of the transfected cells to study the levels of endogenously expressed proteins and to verify the expression of transiently expressed proteins, respectively. Immunoblotting is only a non-linear *semi*-quantitative method; it is therefore

necessary to ensure equal loading of all samples on the gel and to use standard curves. Equal loading can be confirmed by measuring total protein expression using Bradford reagent, Ponceau S staining of the blotted membrane and/or probing for protein levels of β-actin or housekeeping genes.

PCR, RT-PCR and Q-PCR

Polymerase chain reaction (PCR) is a technique used to amplify a specific region of a DNA strand several orders of magnitude. The method relies on thermal cycling, consisting of cycles of repeated heating and cooling to facilitate DNA melting and enzymatic replication of the DNA. In order to robustly detect and quantify gene expression from small amounts of RNA, amplification of the gene transcript is necessary. Reverse transcription PCR (RT-PCR) is a variant of PCR where RNA first is transcribed into cDNA and then amplified in the same manner as in the PCR procedure. These techniques allow both detection and amplification of a specific target sequence in real time by fluorescent dyes or probes (real time or quantitative PCR, Q-PCR). The key feature of Q-PCR is that the amplified DNA is quantified as it accumulates in the reaction in real time after each amplification cycle. The data generated can be analysed by a computer software to calculate the relative gene expression in several samples at the same time. Real-time PCR can also be applied to the detection and quantification of DNA in samples to determine the presence and abundance of a particular DNA sequence in these samples.

Immunostaining

Immunostaining is an antibody-based method to detect a specific protein in a cell or tissue. The interaction can be visualized by adding a secondary antibody with a fluorescent probe attached. Fixation is performed in order to preserve the cell morphology and tissue architecture. To gain accessibility for the antibodies, permeabilization of the cell membrane is performed usually by using detergents at low concentrations. One of the main difficulties with immunostaining is overcoming specific or non-specific backgrounds. Optimisation of all the experimental steps is usually necessary to circumvent these problems. In addition, the presence of positive and negative controls for staining is essential for determining the specificity.

Immunoprecipitation

Immunoprecipitation (IP) is a technique for precipitating a protein antigen out of solution using an antibody that specifically binds to that particular protein. Immunoprecipitation of intact protein complexes is known as co-immunoprecipitation (Co-IP). This process can be used to isolate and concentrate a specific protein or protein complex from a sample containing many thousands of different proteins and complexes. Both techniques require that the antibody is coupled to a solid matrix for isolation of the protein/complex of interest and removal of non-specific interactors. In order to exclude false-positive interactions, control experiments where the immunoprecipitations are performed with the IgG fraction of the relevant species need to be performed.

DNA affinity purification

DNA affinity purification is a powerful technique used for isolation and identification of protein binding a specific DNA sequence of interest. The DNA affinity purification procedure performed in Paper III was conducted as described by Atanasiu et al with some modifications (Atanasiu *et al.*, 2005). The biotinylated double stranded DNA (the bait) was coupled to streptavidine-coated magnetic beads and subsequently incubated with nuclear extracts or purified proteins. The resulting DNA-protein complexes were subjected to magnetic separation, wash and elution, and the highly enriched protein preparations were analyzed using mass spectrometry for identification of unknown interactors, or immunoblot to verify binding of known factors. DNA-binding proteins such as histones can give rise to background problems. The use of DNA baits no longer than 500 bases and preclearing of the nuclear extract with competitor DNA, such as sonicated salmon sperm, can reduce these problems.

Electrophoretic Mobility Shift Assay (EMSA)

The electrophoretic mobility shift assay is a common affinity method used to study protein-DNA or protein-RNA interactions. The binding reaction takes place *in vitro* by mixing nuclear extracts or *in vitro* translated proteins with a labelled double-stranded DNA probe. The protein-DNA or RNA complexes are separated by non-denaturing polyacrylamide gel electrophoresis (PAGE), based on the differences in size, charge and to a lesser extent the conformation. Binding of a nuclear protein to the labelled probe causes it to move more slowly during gel electrophoresis and results in the appearance of a shifted band. An antibody that recognizes the protein can be added to this mixture to create an even larger complex with a greater shift. This is called a supershift assay, and is

used to identify a protein present in the protein-nucleic acid complex. Supershifting is not always possible, as this requires specific antibodies to the target proteins that are able to bind the protein when it is bound to the DNA. Another way to try to deduce the identity of the protein is by competition experiments with unlabelled competitor sequences of the target-protein binding site. An excess of the unlabelled consensus sequence is added to the binding reaction, if any of the shifted bands are reduced or disappear, the band is most likely caused by binding of the target protein. Variants of the competition assay are useful for measuring the specificity of the binding and for measurement of the association and dissociation kinetics.

Chromatin immunoprecipitation assay

Chromatin immunoprecipitation (ChIP) is a powerful tool used to determine the location of DNA binding sites on transcription regulatory sequences in the chromatin structure. This technique gives a picture of the protein-DNA interactions that occur inside the nucleus of living cells or tissues *in vivo*. The DNA-binding proteins, such as histones and transcription factors, are chemically cross-linked to the DNA in order to maintain the low affinity interactors as well. The cross-linked cells are lysed and the chromatin is sheared by sonication to an average size of 200-600 bp. The protein-DNA complex can thereafter be immunoprecipitated out of the lysate by using an antibody specific to the putative DNA-binding protein. To verify if your sequence of interest is pulled down with the antibody, the chromatin is de-crosslinked, the proteins are degraded by proteinase K and the resulting "naked DNA" is analyzed by quantitative PCR (Q-PCR). The amount of DNA specifically pulled down is compared to the amount pulled down by a non-related antibody or pre-immune serum (normal IgG). It is also important to monitor the amount of DNA that is pulled down unspecifically by the antibody of interest. This is achieved by running a Q-PCR of an unrelated region of the genome.

Tandem affinity purification

Tandem affinity purification is a method used to isolate and identify multi-protein complexes associated with your protein of interest under native conditions. A fusion protein is made with the bait protein attached to a double affinity tag at the C- or N-terminus. A protease cleavage site separates the affinity tag and the protein sequence. In Paper I we used an affinity tag composed of the StrepTagII sequence and the Protein A domain from *Saccharomyces Cerevisiae*, divided by two Tobacco Etch Virus cleavage sites. The construct is transfected into mammalian cells and the resulting cell extracts

are subjected to purification by two consecutive affinity chromatography steps. There is a possibility that a tag added to a protein might obscure binding of the new protein to its interacting partners. In addition, the tag may also affect protein expression levels. On the other hand, the tag may also not be sufficiently exposed to the affinity beads, hence skewing the results. All these factors are taken into consideration when designing your purification procedure. To identify individual proteins in the resulting eluate, the enzyme trypsin is often used to generate smaller peptides that are more suitable for mass spectrometric analysis. Trypsin has a very well defined specificity, as it hydrolyzes only the peptide bonds in which the carbonyl group is contributed either by an arginine or lysine residue. The purified protein eluates were separated by gel electrophoresis and the protein-containing lanes of the gel were excised and subjected to in-gel tryptic digestion for further mass spectrometric analysis.

Mass spectrometry analysis and protein identification

Mass spectrometry (MS) is the core method of choice when trying to identify unknown compounds, to quantify known compounds, and to elucidate the structure and chemical properties of molecules. A mass spectrometer is an instrument that measures the masses of individual molecules or fragments that have been electrically charged, i.e. converted into ions. Thus, the mass spectrometer does not actually measure the molecular mass directly, but rather the mass-to-charge (m/z) ratio of the ions formed from the molecules. By coupling two mass analysis stages (MS/MS) one can create fragmented ion spectra of the precursor ions detected in the first MS mode. The precursor mass and fragment ion spectrum for each peptide analysed are used for identification of proteins in a complex mixture by comparing the patterns of the spectra against a theoretical fragment ion spectrum derived from a database of known protein sequences. In this thesis we have analysed tryptic digests of proteins by a liquid chromatography-tandem mass spectrometry using a linear ion trap - Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer. FT-ICR mass spectrometry is a high-resolution technique where mass to charge ratios can be determined with very high accuracy. Accordingly, this technique also allows the signals of two ions of similar masses to be detected as distinct ions. This contributes to improved detection range and reduction of false positive protein identification.

In situ proximity ligation assay

We have used the highly sensitive *in situ* Proximity ligation method to examine the subcellular localization of protein-protein interactions at single-molecule

resolution. Antibodies specific for immunoglobulins from different species, modified by attachment of oligonucleotides, were used as secondary proximity probes together with a pair of primary antibodies from the corresponding species, specific for the proteins of interest. The proximity probes guide the formation of circular DNA strands when bound in close proximity. The DNA circles in turn serve as templates for a localized rolling-circle amplification (RCA) procedure, allowing individual interacting pairs of protein molecules to be visualized and counted in human cell lines and clinical specimens. Combining the dual recognition of the protein-protein interaction by proximity probes offers an improved selectivity and sensitivity, as well as visualization of low abundant protein interactors.

3.3 Results and Discussion

Paper I – Isolation and identification of EBNA5 multi-protein complexes

The biological events occurring in the body are highly complex. The expression, production, secretion and interaction of proteins, peptides and small molecules often occur rapidly and at low concentrations. Methods for isolation and identification of such events must be rapid, selective, sensitive and robust. In recent years, new variations of affinity methodologies have been developed for that purpose.

EBNA5 is an unusual protein consisting of a 22 and 44-amino acid multiple-repeat domain encoded by the W_1 and W_2 exons in the major internal repeat (IR1), and a unique 45-amino acid carboxy-terminal end encoded by the Y1 and Y2 exons, located downstream of the IR1 (Sample *et al.*, 1986; Speck *et al.*, 1986). Several isoforms of EBNA5 due to alternative splicing can be detected during primary infection, but still most established lymphoblastoid cell lines expresse only one major and a few minor species (Nitsche *et al.*, 1997). EBNA5 is a multifunctional protein with several previously identified interaction partners including pRb, p53, Hsp70/Hsc70, Hsp27, DNA-PKcs, α- and β-tubulin, Prolyl-4-hydroxylase, HA95, HAX-1, hERR-1, p14ARF, S3a and Sp100. The biological implications of these interactions have however not been fully understood.

In our attempt to further elucidate the biological activity of EBNA5, we focused on the isolation and identification of EBNA5 interaction partners in processes that do not necessarily depend on the presence of EBNA2. We were also particularly interested in the mechanism underlying the previously described repressor effect of EBNA5 (Dufva *et al.*, 2002). Purification of protein complexes and identification of their protein components are important strategies to unveil the function of proteins with a known sequence but an

unknown biological activity. Several approaches have been employed for this purpose such as yeast two-hybrid methods, enzyme assays and tandem affinity purification (TAP) procedures followed by mass spectrometry analysis. We have used an improved tandem affinity purification method for purification of proteins interacting with EBNA5 in mammalian cells. Initial experiment using the original TAP method described by Rigault *et al* (Rigaut *et al.*, 1999) resulted in high background levels. In order to circumvent this problem, we modified the affinity tag by replacing the calmodulin binding peptide with one copy of the StrepTagII sequence followed by two Tobacco etch virus (TEV) protease cleavage sites and the IgG-binding Protein A domains.

Identification of proteins co-purified with EBNA5-StrepTAP or StrepTAP alone, which served as a control, was performed by LC-MS/MS analysis of tryptic digests on a Hybrid Linear Ion Trap (LTQ)-FT-ICR mass spectrometer. All acquired MS/MS spectra were searched against the complete Swiss-Prot protein sequence database. To address the issue of false positive protein identifications, a randomized version of the Swiss-Prot database was constructed for parallel data searches. With this approach together with stringent criteria during selection of positive protein identifications and subtraction of proteins identified in the control samples, we identified 147 proteins as novel EBNA5 interactors.

Validation of EBNA5 interacting factors was performed by coimmunoprecipitaions experiments and/or by split-tag experiments using two novel protein interactors as baits, Bcl2-associated Athanogene 2 (BAG2) and heterogeneous ribonucleoprotein M (hnRNP M). With these two approaches, 37 proteins were validated as EBNA5 interacting partners. The size and composition of the isolated EBNA5-StrepTAP complexes where further investigated by gel filtration experiments under native conditions. The results revealed that most of the EBNA5 protein is part of high molecular weight protein complexes of 2000 kDa or more. This supports the notion that the functional units of the cell forms interconnected structures and that the TAP procedure allows purification of delicate subcellular structures of a complex nature.

Among the novel interaction partners identified, several proteins are interesting with regard to the function of EBNA5. Our finding that the co-chaperone BAG2 and the 26S proteasome subunit 2 are members of the EBNA5 interactor family, as is the well-characterized Hsp/Hsc70, is very interesting since they are all components constituting the cytoplasmic protein quality control system. Furthermore, the identification of Hsp40 as an associated partner to EBNA5 reveals the fact that EBNA5 interacts with all the factors necessary for an active Hsp70 chaperone complex. In addition, the majority of the 37 novel interactors could be assigned to one of three groups according to function: protein folding and degradation, pre-mRNA processing, or ribosomal proteins. This is

consistent with the notion that functional units of the cell are not single proteins but well-structured macrocomplexes i.e machineries, suggesting that EBNA5 have functional relationship to these distinct groups. In addition, the cellular protein-protein networks contain highly connected proteins i.e. hubs. These hubs have a tendency of being essential and are usually targeted by viral proteins. Our results support the notion that EBNA5 targets protein factories of central positions which may have functional contributions in EBV biology.

Paper II – Regulation of the cytoplasmic quality-control protein-degradation pathway by EBNA5.

We have previously reported that EBNA5 can inhibit pre-mRNA cleavage and polyadenylation in transient transfection systems (Dufva *et al.*, 2002). Since the total level of reporter gene repression could not solely be explained by this mechanism, we continued the search for other biological functions that could be responsible for this phenomenon.

The EBNA5 and EBNA2 proteins are the first viral proteins expressed in primary EBV infection. The two proteins cooperate and upregulates the expression of cyclin D2 and thereby increase the progression of resting B cells into the G_1 phase of the cell cycle (Sinclair *et al.*, 1994). It has also been shown that EBNA5 is important for efficient immortalization of naïve B cells *in vitro* (Hammerschmidt & Sugden, 1989). Furthermore, EBNA5 functions as a coactivator of EBNA2 in the transactivation of the LMP1 promoter. This upregulation is partly mediated by the inducible form of the chaperone Hsp70 and the interferon-inducible gene Sp100 (Ling *et al.*, 2005; Peng *et al.*, 2007). EBNA5 is highly expressed during the first two days of EBV infection, then decreases to levels found in established lymphoblastiod cell lines (Finke *et al.*, 1987; Szekely *et al.*, 1995). EBNA5 is diffusely distributed in the whole nucleoplasm during the early hours after infection but by the end of the first day condenses into discrete foci, which coincide with nuclear bodies known as ND10 domains or PML-oncogenic domains (PODs). EBNA5 is translocated to the nucleolus under conditions of cellular stress, including heat shock, high cell density and proteasome inhibition (Mattsson *et al.*, 2001; Sample *et al.*, 1986; Szekely *et al.*, 1996). In addition to its role in regulating promoter activity, EBNA5 have been implicated in the regulation of degradation of specific nuclear proteins due to the similar behaviour of EBNA5 and components of the proteasome dependent degradation machinery (Kashuba *et al.*, 2003; Mattsson *et al.*, 2001).

In the present study, we have used a model system to monitor effects of EBNA5 on activity, solubility and localization of proteins that are transiently overexpressed in B lymphocytes. The intention was to mimic conditions found during acute primary EBV infection when the level of EBNA5 expression has

reached its maximum. We have used the well-characterized reporter protein luciferase, which has been widely used in studies of gene expression. The results showed that co-transfected EBNA5 decreased the amount of active and soluble luciferase protein in the cytoplasm but increased the total amount of luciferase due to accumulation of insoluble reporter protein in the nucleus. We further observed that luciferase localized at nucleoli in the presence of EBNA5 and that the majority of EBNA5 was accumulated in these nucleoli. Mutation analysis of EBNA5 identified a conserved region in the EBNA5 W_1W_2 repeat domain (CR3) as essential for EBNA5-mediated accumulation of insoluble luciferase and nucleolar localization of luciferase. Our previous study indicated functional relationships between EBNA5 and proteins in the chaperone dependent protein-folding and ubiquitin-proteasome degradation systems (Paper I). To examine the interactions between EBNA5 and the associated factors Hsp70 and BAG2, we employed the *in situ* proximity ligation assay where the interaction of interest can be monitored at the single molecule resolution. The results showed a significant interaction between wild-type EBNA5 and BAG2 in the cells the corresponding interaction could not be detected between the CR3-EBNA5 mutant and BAG2. The results strongly supported the hypothesis that BAG2 is a major target of EBNA5, which was further strengthened by the observation that the mutant CR3-EBNA5 interacted with Hsp70 but not with BAG2.

EBNA5 has long been implicated in regulation of the protein quality control system constituting two pathways: chaperone-mediated folding and ubiquitin-proteasome degradation. The choice between folding and degradation has come to be known as the "molecular triage". Molecular chaperones as Hsp70 are considered as folding and assembly factories with a role in degradation of misfolded proteins as well. The activity of Hsp70 depends on its interaction with specific co-chaperones. Labelling of chaperone substrates for degradation is mainly regulated by the protein CHIP (C terminus of Hsp70-interacting protein). CHIP is an E3 ligase, and with the ubiquitin-conjugating enzyme Ubc mediates ubiquitinylation of substrates presented by Hsp70, thereby initiating sorting of proteins to the proteasome. Co-chaperones do not only facilitate chaperone-assisted degradation, but also control and restrict degradation functions of molecular chaperones. Binding of BAG2 to the Hsp70/CHIP complex inhibits the CHIP-mediated ubiquitinylation, thus competes with the degradation-stimulating co-chaperone BAG1 in the regulation of the Hsp70/CHIP complex (Arndt *et al.*, 2005; Dai *et al.*, 2005). Inhibition of CHIP occurs only in the context of a ternary complex consisting of Hsp70, BAG2, and CHIP. The observed specificity of wild-type EBNA5 binding abilities to BAG2 support the notion that inhibition of CHIP might occur in the present ternary complex and that BAG2 is a major target of EBNA5. This is further supported by the observed additive effect of aggregation on luciferase conducted by overexpression of BAG2 and EBNA5, implicating synergism between the two

proteins. The stimulatory ability of BAG2 on client binding to Hsp70, together with inhibition of the ubiquitin ligase activity of CHIP, argues for the hypothesis that EBNA5 act on several levels in the chaperon-dependent protein-folding and ubiquitin-proteasome degradation systems. This might be of high importance, particularly during the early phase of EBV infection where EBNA5 is known to play a major role. During viral infections the creation of novel interactions and modifications or the destruction of others, results in a stressed host cell where the production of viral protein is favourable. One plausible outcome may involve the ability of EBNA5 to elicit a stress response that in the end increases expression of EBV proteins and subsequent cellular transformation. Taken together a model can be put forward that the EBNA5/Hsp70/BAG2 complex identified in the present study represent a chaperone complex that has been turned into a machine with abrogated protein folding activity and inhibited protein degradation ability by the presence of EBNA5.

Paper III – Identification of factors binding to the Cp core that may facilitate long-distant promoter-enhancer interactions

To further characterize the molecular and mechanistic details of the presumably large macromolecular complex that initiates transcription from the C promoter (Cp), we set out to find additional Cp interacting proteins by using DNA affinity purification and mass spectrometric analysis.

The Cp transcript produces all six latent EBNA-proteins from the same primary transcript modulated by alternative splicing. The activity of Cp is directed by several regulatory sequences of which the *oriPI* family of repeats (FR) region is the major regulatory site. The *oriPI* element is the EBNA1-dependent enhancer of the Cp. EBNA1 binds as a dimer to the regulatory core consensus sequence, which consists of multiple EBNA1 and octamer binding sites. The enhancement by the EBNA1/*oriPI* complex follows a complicated pattern, where at least 8 of the available 20 sites need to be occupied by EBNA1 to approach full transcriptional activation (Zetterberg *et al.*, 2004). The mechanism for the interaction between the *oriPI*-EBNA1 complex and Cp is, however, still not understood at the molecular level.

The observation that EBNA1-dependent *oriPI*-enhancement of Cp activity was far greater in cells of the B cell origin, indicated that there must be additional interacting proteins, presumably B cell specific factors needed for efficient activation. In order to isolate the transcriptional initiation complex, we employed the DNA affinity purification procedure. In DNA affinity purification it is vital not to use long baits, since this causes massive background problems due to the large pool of DNA binding proteins in the cell. We therefore decided to use baits covering the -170Cp region in the initial purification. Baits amplified from the ampicillin coding region were used as negative control. The

affinity purified eluted proteins were separated by 1D gel electrophoresis and subjected to in-gel tryptic digestion. The resulting peptide mixture was analyzed by a Hybrid Linear Ion Trap (LTQ)-FT-ICR mass spectrometer, and the required data was subjected to bioinformatic analysis. The list of proteins binding to the Amp negative control bait was subtracted from the list of proteins binding to the Cp fragment. Further data reduction was performed by filtering the data to contain only proteins annotated as transcription factors. Among the transcription factors identified as putative Cp interactors, we selected E2F1, ARID3A and Oct-2 for further characterization. Interestingly, E2F1 as well as Oct-2 have previously been shown to bind the *oriP*-region (Almqvist *et al.*, 2005; Maser *et al.*, 2001) as well as to the Cp (Chau *et al.*, 2008). We repeated the affinity purification using the *oriPI* fragment with either three or eight EBNA1 repeats (Δ*oriPI*). Immunoblot experiments on the *oriPI* eluates showed that all three factors also bound the Δ*oriPI* sequence. ARID3A-binding was highly enriched in the activation-competent Δ*oriPI* with eight EBNA1 binding repeats, in comparison to the activation-incompetent Δ*oriPI*-construct with three EBNA1 repeats, whereas the degree of binding of the other factors were unaffected by the number of repeats. Using a series of competition EMSAs and additional DNA affinity purifications with mutated Cp baits, we were able to map the binding sites of all three proteins to a short segment of the Cp core in close proximity of each other. This region (-149/-109) has previously been shown to be required for both *oriPI*-dependent and -independent transcriptional activation, indicating that these interactions are important for the activity of Cp (Nilsson *et al.*, 2001). *In vivo* binding of the proteins to both the -170Cp and the *oriPI* regions was verified using Chromatin Immunoprecipitation (ChIP).

In addition, the role of EBNA1 in the novel isolated complex was investigated by co-immunoprecipitation experiments using specific antibodies and western blot analysis. The results showed that all three factors, E2F1, ARID3A and Oct-2, bound the EBNA1 protein as well. These results were further confirmed for all three factors by *in situ* proximity ligation assays in which the interactions were investigated at the single molecule level. The interactions were located to the nuclear compartment, strengthening the notion that these factors are part of a functional complex in the nucleus.

The identification of these three factors as part of the large Cp-*oriPI*-EBNA1 complex leaves several interesting contributions in the context of EBV pathogenesis. We have in a collaboration shown that the Oct-1 and Oct-2 factors bind the *oriPI* motif *in vitro* and upregulates the transcriptional activation of Cp (Almqvist *et al.*, 2005). It has been suggested that EBNA1 and Oct-2 are responsible for the switch from Qp to Cp (Werner *et al.*, 2007). Recently Oct-2 has also been shown to play an important role in the regulation of viral reactivation in Karposi's Sarcoma-associated Herpesvirus (KSHV) infected cells (Di Bartolo *et al.*, 2009).

E2F1 is a ubiquitously expressed protein that belongs to a family of transcription factors that is essential for the regulation of cell growth and plays an important role in almost every phase of the cell cycle. Ectopic expression of E2F1 is sufficient to drive quiescent cells to S phase, suggesting its pivotal role in G_1/S transition and oncogenic potential. The transcriptional activity of E2F1 is regulated through its interaction with the retinoblastoma protein (Rb). At the early G_1 hypophosphorylated Rb binds to E2F1 and inhibits -phase, its activity. During mid-G_1, Rb is subsequently phosphorylated by cyclin-dependent kinases which leads to transcriptional activation of several G_1/S transition genes.

It is well known that the pRB/E2F1 pathway is deregulated in almost all human cancers due to frequent Rb inactivation. Chau and co-workers recently reported binding of E2F1 and Rb to a site in the distal Cp upstream of the EBNA2 responsive element (Chau *et al.*, 2008). Using synchronised EBV-positive cells they could show that the E2F1-Rb-Cp interaction vary in a cell cycle dependent way, indicating that the Cp is cell cycle regulated. It has also been shown that E2F1 is involved in regulation of genes affecting p53-dependent apoptosis (DeGregori & Johnson, 2006). The tumour suppressor protein p53 plays a pivotal role in preventing tumourigenesis by inducing cell cycle arrest or apoptosis. MDM2 is an oncogene that negatively regulates p53 activity by blocking its transcriptional activity as well as targeting it for degradation. E2F1 can induce apoptosis via the p14ARF/MDM2/p53-dependent pathway. Interestingly, both the EBNA5 and EBNA6 proteins, which are essential for transformation of B cells *in vitro*, have been shown to target the Rb and p53 proteins as well (Parker *et al.*, 1996; Szekely *et al.*, 1993). In addition, association with p14ARF is also a common feature of E2F1 and EBNA5, via the deregulatory function of EBNA5 on the p14ARF-MDM2-p53 complex (Kashuba *et al.*, 2003).

The E2F1 protein has been shown to be essential for c-Myc-induced carcinogenesis both *in vitro* and *in vivo* (Baudino *et al.*, 2003). In an EBV context the interaction with c-Myc is particularly interesting since this protein is upregulated by EBNA2 in early EBV infection and is the most prominent target essential for the ability of EBNA2 to transform B-cells (Kaiser *et al.*, 1999). In addition, it has recently been shown that E2F1 expression is deregulated and plays an oncogenic role in sporadic Burkitt's lymphoma due to its collaboration with c-Myc (Molina-Privado *et al.*, 2009).

ARID3A (aka DrilI/E2FBP1) is a novel player in the EBV field. The protein was originally isolated as an E2F1 heterodimeric partner that stimulates E2F-dependent transcription (Suzuki, 1998). ARID3A is a member of the AT-rich interaction domain (ARID) family of transcription factors that are known to exert pleiotropic roles in embryonic patterning, cell lineage gene regulation, cell cycle control, chromatin remodelling and transcriptional regulation (Wilsker *et al.*, 2005). ARID3A appears to be ubiquitously expressed in human cells, raising the possibility that the function of the protein depends on the species and

cellular context. It was recently demonstrated that ARID3A deregulated the E2F1/Rb pathway as well as being a direct target gene of p53 (Ma *et al.*, 2003; Peeper *et al.*, 2002). The functional connection of ARID3A to both the E2F1/Rb and p53 tumour suppressor pathways in addition to its pivotal role in promoting either cellular proliferation or growth arrest implies a tight and finely tuned regulation of the protein's activity and binding affinities. In addition, the EBNA1 protein interactor HAUSP/USP7 has been shown to stabilize both p53 and its regulator MDM2, indicating that EBNA1 is involved in p53 regulation/destabilisation as well (Li *et al.*, 2004).

The ARID3A mouse orthologue Bright is a B cell specific, matrix-associating region-binding protein that transactivates gene expression from the IgH intronic enhancer (Eμ). The Eμ are flanked on both sides by nuclear matrix associating regions (MARs), that serve to anchor higher order chromatin to the nuclear matrix. Transfection studies have implicated that Bright facilitates spatially separated promoter-enhancer interactions and the model is further supported by the ability of Bright to induce DNA bending upon binding within the Eμ MAR (Kaplan *et al.*, 2001). These properties are also common features of the EBNA1 protein as shown by the demonstration that the *oriPI* and *oriPII* elements interacts trough EBNA1 by a DNA-looping mechanism (Frappier & O'Donnell, 1991). Bright binds the promoter as well as the enhancer as a tetrameric complex in a manner analogous to the mechanism proposed for *oriPI*-EBNA1-induced activation of the Cp (Herrscher *et al.*, 1995; Mackey & Sugden, 1999b; Zetterberg *et al.*, 2004).

Furthermore, Bright interacts with the ubiquitous autoantigen Sp100, a component of the promyelocytic leukaemia nuclear bodies (PML NBs) (Zong *et al.*, 2000). The PML NBs are a cellular structure that appears to be involved in the pathogenesis of a variety of human diseases including acute promyelocytic leukaemia and viral infections, and are suggested to affect or regulate a wide range of cellular processes. Sp100 is also an interactor of the EBNA5 protein, functioning as an important mediator in the coactivation process with EBNA2 (Ling *et al.*, 2005). Identification and characterization of Bright's interaction with Sp100 may provide a link between nuclear subdomains, matrix attatchment and transcription.

3.4 Conclusions and Future Perspectives

Viruses are among the simplest of pathogens, thus requiring the host cell machinery for replication and survival. Since the genomes of many important viruses have been cloned and sequenced, structure and functional studies have been performed on viral genes and gene products. While the study of genes and proteins continues to be important, looking at isolated components is not enough to understand most biological processes. Recent advances in cell biology have allowed in a shift from this reductionist approach of just studying the virus itself to an approach, which explores the molecular interactions between the virus and its host. Protein-protein interactions define literally all important aspects of the viral biology, ranging from the initial attachment to the host cell, to assembly and release of newly formed virus particles.

One of the greatest challenges in the post-genomic era is the prediction of protein functions. Proteins that share connections in a protein-protein interaction network tend to have similar functions. In this way, by analyzing the neighbourhood of known proteins, it is possible to infer some of the functional roles of their direct neighbours. In the present thesis we have primarily focused on further elucidating the molecular mechanisms of the EBNA1 and EBNA5 proteins with proteomic technologies as approach by isolation and identification of multi protein complexes. We have used an improved tandem affinity purification strategy for identification of novel putative interaction partners to EBNA5. The majority of validated protein interactors could be assigned to one of three groups according to function: protein folding and degradation, pre-mRNA processing, and ribosomal proteins, implicating functional relationships with EBNA5 in these processes. We also showed that EBNA5 is part of high molecular protein complexes, supporting the notion that functional units in the cell are not single proteins but well-structured complexes composed of multiple proteins i.e. factories or modules.

The previously reported repressor activity of EBNA5 was further investigated in paper II. In this study we identified the novel interactor BAG2 as a major target for the function of EBNA5 via the chaperone-mediated folding and proteasome-degradation pathways. Taken together, the results are consistent with the hypothesis suggesting that EBNA5 tune the balance between protein rescue and destruction in a way that disfavour the path of degradation by inhibition of ubiquitinylation.

The constituent of the large macromolecular complex that initiates transcription from the viral C promoter (Cp) was further investigated in paper III. Using the DNA affinity procedure we revealed that the transcription factors E2F1, ARID3A/Bright and Oct-2 binds the Cp core, as well as the EBNA1 protein and oriPI. We have identified interactions between individual factors, involved in either of the Rb- or p53-pathways, and the EBNA1 protein as well

as the oriPI/Cp regulatory regions of the EBV genome, possibly facilitating long-distance promoter-enhancer interactions. The tight association of these factors with the p53 and Rb pathways contributes to several interesting implications according to EBV pathogenesis, as they both are central pathways in cancer biology (Klein, 2009). The p53 tumour suppressor protein is associated with approximately 50% of all human cancers and inactivation of this protein is a key mechanism for viral proteins for altering host cells and initiating tumour growth. One major interesting finding presented herein is that both the EBNA1 and EBNA5 proteins target the same highly interconnected cellular protein i.e. hubs. This is consistent with the hypothesis that several viral protein target the same hubs in the host, which ensures the takeover of the cellular machineries essential for the viral infection and persistence processes, and contribute to the robustness of the viral infectious system. These hubs are mainly part of the central mechanisms or pathways of the cell such as cell cycle regulation, proliferation and cell survival. The general isolation and identification procedures of protein complexes presented herein are powerful tools to extract important clues about virus-host interaction dynamics. Elucidation of the pathogenic mechanisms of viruses and specific host-pathogen interactions are a key question in EBV biology.

In addition of helping our understanding of the basic biology of herpesviruses, these large-scale protein-protein interaction studies can be beneficial from a clinical point of view. Cancer is a complex set of diseases in which genetic and epigenetic alterations of cells have an important part in its progression. To combat cancer it is important to understand in molecular detail the effector mechanisms that lead to activation and inhibition of tumour progression because it is these interactions that can be potentially exploited for therapy. The virus proteins alter the host interaction networks by competing with host proteins for binding in the host protein-protein interaction network. By identification of these functionally important interactions may prove valuable in the search for new forms of anti viral therapy.

4

ACKNOWLEDGEMENTS

I wish to express my sincerest gratitude to all former and present colleagues and friends who in one way or another have helped, supported and inspired me throughout this PhD period. This thesis would not have existed without them. In particular I would like to thank:

Lars Rymo, my supervisor, I am deeply grateful for you seeing the researcher in me and inspiring me to join your group. It has been a long and hard journey but it has ceratinly been a very educational one. I would also like to thank you for letting me research with freedom and for givening me guidance when I have needed it. You have also tought me to strive to do good science and not to take the easy way out.

Ulla Rüetschi, for being my co-supervisor. I don't know how I would have managed without you. Thank you for your support and for all fruitful talks a chemist to another during all these projects technical problems.

Cecilia Boreström, for being a great colleague and friend. Thank you for all the helpful comments and pep-talks durig the last pieces of this thesis.

To all former and present collegues at Molecular biology and Clinical Chemistry, especially **Pegah Johansson, Susanne Nyström, Annica Sjölander, Ann Jansson, Åsa Nilsson, Josefine Ekholm, Aida Muslimovic, Malin Berggren, Sara Landgren, Malin von Otter, Petra Bergström, Henrik Zetterberg, Lars Palmquist, Ismail Hassan Ismail** and **Maria Olsson.** You have all in one way or the other been a fantastic part of this PhD period. Thank you very much for all advices, giving discussions and moral support through good times and bad. You have been both collegaues and friends contributing to a great time and a giving research environment.

Genanalys for always being helpful and friendly, and for being an excellent technical resource for us students. I really miss our joint coffee breaks.

To all my **friends** and **relatives** that has supported me in everything, both during this time and as long as you have been part of my life. You enrich my life with so much and for that I am truly grateful.

Marie and **Bror**, thank you for being wonderful parents in-law. Its not easy to have a daughter in-law with such a strong will and I am greatful for all your support and help. Your house in Gotland is a safe heaven for our little family every summer.

My wonderful parents **Lisbeth** and **Björn,** for always encouraging and believing in me. You have taught me to believe in myself and never to doubt my abilities. Your support is invaluable and my gratitude endless.

Stina and **Elin**, my sisters. Its not easy to be the youngest of three sisters but you are a big part of who I am today. Thank you for all your love and encouragement.

Finally, my husband **Henrik**, for being my other half. You came into my life more than 11 years ago, we have shared so much and I love you for every bit of it. Thank you for always being in my corner and for giving me: **Elton** - the sunshine of my life - thank you for showing mom a whole new wonderful side of life.

This study was supported by grants from the Swedish Medical research Council, The Swedish Cancer society, the Sahlgrenska University Hospital, the Assar Gabrielsson Foundation for Clinical Research and the King Gustav V Jubilee Clinic Cancer Research Foundation.

5
REFERENCES

ALBERT, R., JEONG, H. & BARABASI, A. L. (2000). ERROR AND ATTACK TOLERANCE OF COMPLEX NETWORKS. *NATURE* **406**, 378-382.

ALBERTS, B. (1998). THE CELL AS A COLLECTION OF PROTEIN MACHINES: PREPARING THE NEXT GENERATION OF MOLECULAR BIOLOGISTS. *CELL* **92**, 291-294.

ALLDAY, M. J. & FARRELL, P. J. (1994). EPSTEIN-BARR VIRUS NUCLEAR ANTIGEN EBNA3C/6 EXPRESSION MAINTAINS THE LEVEL OF LATENT MEMBRANE PROTEIN 1 IN G1-ARRESTED CELLS. *J VIROL* **68**, 3491-3498.

ALMQVIST, J., ZOU, J., LINDERSON, Y., BORESTROM, C., ALTIOK, E., ZETTERBERG, H., RYMO, L., PETTERSSON, S. & ERNBERG, I. (2005). FUNCTIONAL INTERACTION OF OCT TRANSCRIPTION FACTORS WITH THE FAMILY OF REPEATS IN EPSTEIN-BARR VIRUS ORIP. *J GEN VIROL* **86**, 1261-1267.

ARNDT, V., DANIEL, C., NASTAINCZYK, W., ALBERTI, S. & HOHFELD, J. (2005). BAG-2 ACTS AS AN INHIBITOR OF THE CHAPERONE-ASSOCIATED UBIQUITIN LIGASE CHIP. *MOL BIOL CELL* **16**, 5891-5900.

ATANASIU, C., LEZINA, L. & LIEBERMAN, P. M. (2005). DNA AFFINITY PURIFICATION OF EPSTEIN-BARR VIRUS ORIP-BINDING PROTEINS. *METHODS MOL BIOL* **292**, 267-276.

BAJAJ, B. G., MURAKAMI, M., CAI, Q., VERMA, S. C., LAN, K. & ROBERTSON, E. S. (2008). EPSTEIN-BARR VIRUS NUCLEAR ANTIGEN 3C INTERACTS WITH AND ENHANCES THE STABILITY OF THE C-MYC ONCOPROTEIN. *J VIROL* **82**, 4082-4090.

BAUDINO, T. A., MACLEAN, K. H., BRENNAN, J., PARGANAS, E., YANG, C., ASLANIAN, A., LEES, J. A., SHERR, C. J., ROUSSEL, M. F. & CLEVELAND, J. L. (2003). MYC-MEDIATED PROLIFERATION AND LYMPHOMAGENESIS, BUT NOT APOPTOSIS, ARE COMPROMISED BY E2F1 LOSS. *MOLECULAR CELL* **11**, 905-914.

BORZA, C. M. & HUTT-FLETCHER, L. M. (2002). ALTERNATE REPLICATION IN B CELLS AND EPITHELIAL CELLS SWITCHES TROPISM OF EPSTEIN-BARR VIRUS. *NAT MED* **8**, 594-599.

BUELL, P. (1974). THE EFFECT OF MIGRATION ON THE RISK OF NASOPHARYNGEAL CANCER AMONG CHINESE. *CANCER RES* **34**, 1189-1191.

CALDERWOOD, M. A., VENKATESAN, K., XING, L., CHASE, M. R., VAZQUEZ, A., HOLTHAUS, A. M., EWENCE, A. E., LI, N., HIROZANE-KISHIKAWA, T., HILL, D. E., VIDAL, M., KIEFF, E. & JOHANNSEN, E. (2007). EPSTEIN-BARR VIRUS AND

VIRUS HUMAN PROTEIN INTERACTION MAPS. *PROC NATL ACAD SCI U S A* **104**, 7606-7611.

CALDWELL, R. G., WILSON, J. B., ANDERSON, S. J. & LONGNECKER, R. (1998). EPSTEIN-BARR VIRUS LMP2A DRIVES B CELL DEVELOPMENT AND SURVIVAL IN THE ABSENCE OF NORMAL B CELL RECEPTOR SIGNALS. *IMMUNITY* **9**, 405-411.

CHAU, C. M., DENG, Z., KANG, H. & LIEBERMAN, P. M. (2008). CELL CYCLE ASSOCIATION OF THE RETINOBLASTOMA PROTEIN RB AND THE HISTONE DEMETHYLASE LSD1 WITH THE EPSTEIN-BARR VIRUS LATENCY PROMOTER CP. *J VIROL* **82**, 3428-3437.

COHEN, J. I., (ED) (2005). *CLINICAL ASPECTS OF EPSTEIN-BARR VIRUS INFECTION*. WYMONDHAN: CALISTER ACADEMIC PRESS.

COHEN, J. I., WANG, F., MANNICK, J. & KIEFF, E. (1989). EPSTEIN-BARR VIRUS NUCLEAR PROTEIN 2 IS A KEY DETERMINANT OF LYMPHOCYTE TRANSFORMATION. *PROC NATL ACAD SCI U S A* **86**, 9558-9562.

DAI, Q., QIAN, S. B., LI, H. H., MCDONOUGH, H., BORCHERS, C., HUANG, D., TAKAYAMA, S., YOUNGER, J. M., REN, H. Y., CYR, D. M. & PATTERSON, C. (2005). REGULATION OF THE CYTOPLASMIC QUALITY CONTROL PROTEIN DEGRADATION PATHWAY BY BAG2. *J BIOL CHEM* **280**, 38673-38681.

DAWSON, C. W., TRAMOUNTANIS, G., ELIOPOULOS, A. G. & YOUNG, L. S. (2003). EPSTEIN-BARR VIRUS LATENT MEMBRANE PROTEIN 1 (LMP1) ACTIVATES THE PHOSPHATIDYLINOSITOL 3-KINASE/AKT PATHWAY TO PROMOTE CELL SURVIVAL AND INDUCE ACTIN FILAMENT REMODELING. *J BIOL CHEM* **278**, 3694-3704.

DEGREGORI, J. & JOHNSON, D. G. (2006). DISTINCT AND OVERLAPPING ROLES FOR E2F FAMILY MEMBERS IN TRANSCRIPTION, PROLIFERATION AND APOPTOSIS. *CURR MOL MED* **6**, 739-748.

DI BARTOLO, D. L., HYJEK, E., KELLER, S., GUASPARRI, I., DENG, H., SUN, R., CHADBURN, A., KNOWLES, D. M. & CESARMAN, E. (2009). ROLE OF DEFECTIVE OCT-2 AND OCA-B EXPRESSION IN IMMUNOGLOBULIN PRODUCTION AND KAPOSI'S SARCOMA-ASSOCIATED HERPESVIRUS LYTIC REACTIVATION IN PRIMARY EFFUSION LYMPHOMA. *J VIROL* **83**, 4308-4315.

DUFVA, M., FLODIN, J., NERSTEDT, A., RUETSCHI, U. & RYMO, L. (2002). EPSTEIN-BARR VIRUS NUCLEAR ANTIGEN 5 INHIBITS PRE-MRNA CLEAVAGE AND POLYADENYLATION. *NUCLEIC ACIDS RES* **30**, 2131-2143.

EPSTEIN, M. A. & BARR, Y. M. (1964). CULTIVATION IN VITRO OF HUMAN LYMPHOBLASTS FROM BURKITT'S MALIGNANT LYMPHOMA. *LANCET* **1**, 252-253.

FAHRAEUS, R., JANSSON, A., RICKSTEN, A., SJOBLOM, A. & RYMO, L. (1990). EPSTEIN-BARR VIRUS-ENCODED NUCLEAR ANTIGEN 2 ACTIVATES THE VIRAL LATENT MEMBRANE PROTEIN PROMOTER BY MODULATING THE ACTIVITY OF A NEGATIVE REGULATORY ELEMENT. *PROC NATL ACAD SCI U S A* **87**, 7390-7394.

FARRELL, P. J., (ED) (2005). *EPSTEIN-BARR VIRUS GENOME*. WYMONDHAM: CAISTER ACADEMIC PRESS.

FINKE, J., ROWE, M., KALLIN, B., ERNBERG, I., ROSEN, A., DILLNER, J. & KLEIN, G. (1987). MONOCLONAL AND POLYCLONAL ANTIBODIES AGAINST EPSTEIN-BARR VIRUS NUCLEAR ANTIGEN 5 (EBNA-5) DETECT MULTIPLE PROTEIN SPECIES IN BURKITT'S LYMPHOMA AND LYMPHOBLASTOID CELL LINES. *J VIROL* **61**, 3870-3878.

FRANKLIN, D. S., GODFREY, V. L., O'BRIEN, D. A., DENG, C. & XIONG, Y. (2000). FUNCTIONAL COLLABORATION BETWEEN DIFFERENT CYCLIN-DEPENDENT KINASE INHIBITORS SUPPRESSES TUMOR GROWTH WITH DISTINCT TISSUE SPECIFICITY. *MOL CELL BIOL* **20**, 6147-6158.

FRAPPIER, L. & O'DONNELL, M. (1991). EPSTEIN-BARR NUCLEAR ANTIGEN 1 MEDIATES A DNA LOOP WITHIN THE LATENT REPLICATION ORIGIN OF EPSTEIN-BARR VIRUS. *PROC NATL ACAD SCI U S A* **88**, 10875-10879.

FRUEHLING, S. & LONGNECKER, R. (1997). THE IMMUNORECEPTOR TYROSINE-BASED ACTIVATION MOTIF OF EPSTEIN-BARR VIRUS LMP2A IS ESSENTIAL FOR BLOCKING BCR-MEDIATED SIGNAL TRANSDUCTION. *VIROLOGY* **235**, 241-251.

GAVIN, A. C., ALOY, P., GRANDI, P., KRAUSE, R., BOESCHE, M., MARZIOCH, M., RAU, C., JENSEN, L. J., BASTUCK, S., DUMPELFELD, B., EDELMANN, A., HEURTIER, M. A., HOFFMAN, V., HOEFERT, C., KLEIN, K., HUDAK, M., MICHON, A. M., SCHELDER, M., SCHIRLE, M., REMOR, M., RUDI, T., HOOPER, S., BAUER, A., BOUWMEESTER, T., CASARI, G., DREWES, G., NEUBAUER, G., RICK, J. M., KUSTER, B., BORK, P., RUSSELL, R. B. & SUPERTI-FURGA, G. (2006). PROTEOME SURVEY REVEALS MODULARITY OF THE YEAST CELL MACHINERY. *NATURE* **440**, 631-636.

HAMMERSCHMIDT, W. & SUGDEN, B. (1989). GENETIC ANALYSIS OF IMMORTALIZING FUNCTIONS OF EPSTEIN-BARR VIRUS IN HUMAN B LYMPHOCYTES. *NATURE* **340**, 393-397.

HARADA, S. & KIEFF, E. (1997). EPSTEIN-BARR VIRUS NUCLEAR PROTEIN LP STIMULATES EBNA-2 ACIDIC DOMAIN-MEDIATED TRANSCRIPTIONAL ACTIVATION. *J VIROL* **71**, 6611-6618.

HARTWELL, L. H., HOPFIELD, J. J., LEIBLER, S. & MURRAY, A. W. (1999). FROM MOLECULAR TO MODULAR CELL BIOLOGY. *NATURE* **402**, C47-52.

HERRSCHER, R. F., KAPLAN, M. H., LELSZ, D. L., DAS, C., SCHEUERMANN, R. & TUCKER, P. W. (1995). THE IMMUNOGLOBULIN HEAVY-CHAIN MATRIX-ASSOCIATING REGIONS ARE BOUND BY BRIGHT: A B CELL-SPECIFIC TRANS-ACTIVATOR THAT DESCRIBES A NEW DNA-BINDING PROTEIN FAMILY. *GENES DEV* **9**, 3067-3082.

HUEN, D. S., HENDERSON, S. A., CROOM-CARTER, D. & ROWE, M. (1995). THE EPSTEIN-BARR VIRUS LATENT MEMBRANE PROTEIN-1 (LMP1) MEDIATES ACTIVATION OF NF-KAPPA B AND CELL SURFACE PHENOTYPE VIA TWO EFFECTOR REGIONS IN ITS CARBOXY-TERMINAL CYTOPLASMIC DOMAIN. *ONCOGENE* **10**, 549-560.

JEONG, H., MASON, S. P., BARABASI, A. L. & OLTVAI, Z. N. (2001). LETHALITY AND CENTRALITY IN PROTEIN NETWORKS. *NATURE* **411**, 41-42.

JEONG, H., TOMBOR, B., ALBERT, R., OLTVAI, Z. N. & BARABASI, A. L. (2000). THE LARGE-SCALE ORGANIZATION OF METABOLIC NETWORKS. *NATURE* **407**, 651-654.

JONSSON, P. F. & BATES, P. A. (2006). GLOBAL TOPOLOGICAL FEATURES OF CANCER PROTEINS IN THE HUMAN INTERACTOME. *BIOINFORMATICS* **22**, 2291-2297.

KAISER, C., LAUX, G., EICK, D., JOCHNER, N., BORNKAMM, G. W. & KEMPKES, B. (1999). THE PROTO-ONCOGENE C-MYC IS A DIRECT TARGET GENE OF EPSTEIN-BARR VIRUS NUCLEAR ANTIGEN 2. *J VIROL* **73**, 4481-4484.

KANG, M. S., HUNG, S. C. & KIEFF, E. (2001). EPSTEIN-BARR VIRUS NUCLEAR ANTIGEN 1 ACTIVATES TRANSCRIPTION FROM EPISOMAL BUT NOT INTEGRATED DNA AND DOES NOT ALTER LYMPHOCYTE GROWTH. *PROC NATL ACAD SCI U S A*.

KAPLAN, M. H., ZONG, R. T., HERRSCHER, R. F., SCHEUERMANN, R. H. & TUCKER, P. W. (2001). TRANSCRIPTIONAL ACTIVATION BY A MATRIX ASSOCIATING REGION-BINDING PROTEIN. CONTEXTUAL REQUIREMENTS FOR THE FUNCTION OF BRIGHT. *J BIOL CHEM* **276**, 21325-21330.

KASHUBA, E., MATTSSON, K., POKROVSKAJA, K., KISS, C., PROTOPOPOVA, M., EHLIN-HENRIKSSON, B., KLEIN, G. & SZEKELY, L. (2003). EBV-ENCODED EBNA-5 ASSOCIATES WITH P14ARF IN EXTRANUCLEOLAR INCLUSIONS AND PROLONGS THE SURVIVAL OF P14ARF-EXPRESSING CELLS. *INT J CANCER* **105**, 644-653.

KENNEDY, G., KOMANO, J. & SUGDEN, B. (2003). EPSTEIN-BARR VIRUS PROVIDES A SURVIVAL FACTOR TO BURKITT'S LYMPHOMAS. *PROC NATL ACAD SCI U S A* **100**, 14269-14274.

KIEFF, E. & RICKINSON, A. B. (2001). EPSTEIN-BARR VIRUS AND ITS REPLICATION. IN *FUNDAMENTAL VIROLOGY*, 4 EDN, PP. 1185-1248. EDITED BY D. M. KNIPE & H. P.M. PHILADELPHIA: LIPPINCOTT, WILLIAMS AND WILKINS.

KITANO, H. (2007). BIOLOGICAL ROBUSTNESS IN COMPLEX HOST-PATHOGEN SYSTEMS. *PROG DRUG RES* **64**, 239, 241-263.

KLEIN, G. (1994). EPSTEIN-BARR VIRUS STRATEGY IN NORMAL AND NEOPLASTIC B CELLS. *CELL* **77**, 791-793.

KLEIN, G. (2009). BURKITT LYMPHOMA-A STALKING HORSE FOR CANCER RESEARCH? *SEMIN CANCER BIOL*.

KOVALCHUK, A. L., QI, C. F., TORREY, T. A., TADDESSE-HEATH, L., FEIGENBAUM, L., PARK, S. S., GERBITZ, A., KLOBECK, G., HOERTNAGEL, K., POLACK, A., BORNKAMM, G. W., JANZ, S. & MORSE, H. C., 3RD (2000). BURKITT LYMPHOMA IN THE MOUSE. *J EXP MED* **192**, 1183-1190.

KROGAN, N. J., CAGNEY, G., YU, H., ZHONG, G., GUO, X., IGNATCHENKO, A., LI, J., PU, S., DATTA, N., TIKUISIS, A. P., PUNNA, T., PEREGRIN-ALVAREZ, J. M., SHALES, M., ZHANG, X., DAVEY, M., ROBINSON, M. D., PACCANARO, A., BRAY, J. E., SHEUNG, A., BEATTIE, B., RICHARDS, D. P., CANADIEN, V., LALEV, A., MENA, F., WONG, P., STAROSTINE, A., CANETE, M. M., VLASBLOM, J., WU, S., ORSI, C., COLLINS, S. R., CHANDRAN, S., HAW, R., RILSTONE, J. J., GANDI, K., THOMPSON, N. J., MUSSO, G., ST ONGE, P., GHANNY, S., LAM, M. H., BUTLAND, G., ALTAF-UL, A. M., KANAYA, S., SHILATIFARD, A., O'SHEA, E., WEISSMAN, J. S., INGLES, C. J., HUGHES, T. R., PARKINSON, J., GERSTEIN, M., WODAK, S. J., EMILI, A. & GREENBLATT, J. F. (2006). GLOBAL LANDSCAPE OF PROTEIN COMPLEXES IN THE YEAST SACCHAROMYCES CEREVISIAE. *NATURE* **440**, 637-643.

LEVITSKAYA, J., CORAM, M., LEVITSKY, V., IMREH, S., STEIGERWALD-MULLEN, P. M., KLEIN, G., KURILLA, M. G. & MASUCCI, M. G. (1995). INHIBITION OF ANTIGEN PROCESSING BY THE INTERNAL REPEAT REGION OF THE EPSTEIN-BARR VIRUS NUCLEAR ANTIGEN-1. *NATURE* **375**, 685-688.

LI, M., BROOKS, C. L., KON, N. & GU, W. (2004). A DYNAMIC ROLE OF HAUSP IN THE P53-MDM2 PATHWAY. *MOL CELL* **13**, 879-886.

LI, Z., VAN CALCAR, S., QU, C., CAVENEE, W. K., ZHANG, M. Q. & REN, B. (2003). A GLOBAL TRANSCRIPTIONAL REGULATORY ROLE FOR C-MYC IN BURKITT'S LYMPHOMA CELLS. *PROC NATL ACAD SCI U S A* **100**, 8164-8169.

LING, P. D., PENG, R. S., NAKAJIMA, A., YU, J. H., TAN, J., MOSES, S. M., YANG, W. H., ZHAO, B., KIEFF, E., BLOCH, K. D. & BLOCH, D. B. (2005). MEDIATION OF EPSTEIN-BARR VIRUS EBNA-LP TRANSCRIPTIONAL COACTIVATION BY SP100. *EMBO J* **24**, 3565-3575.

LONGNECKER, R. (2000). EPSTEIN-BARR VIRUS LATENCY: LMP2, A REGULATOR OR MEANS FOR EPSTEIN-BARR VIRUS PERSISTENCE? *ADV CANCER RES* **79**, 175-200.

LONGNECKER, R., DRUKER, B., ROBERTS, T. M. & KIEFF, E. (1991). AN EPSTEIN-BARR VIRUS PROTEIN ASSOCIATED WITH CELL GROWTH TRANSFORMATION INTERACTS WITH A TYROSINE KINASE. *J VIROL* **65**, 3681-3692.

MA, K., ARAKI, K., ICHWAN, S. J., SUGANUMA, T., TAMAMORI-ADACHI, M. & IKEDA, M. A. (2003). E2FBP1/DRIL1, AN AT-RICH INTERACTION DOMAIN-FAMILY TRANSCRIPTION FACTOR, IS REGULATED BY P53. *MOL CANCER RES* **1**, 438-444.

MACKEY, D. & SUGDEN, B. (1999A). APPLICATIONS OF ORIP PLASMIDS AND THEIR MODE OF REPLICATION. *METHODS ENZYMOL* **306**, 308-328.

MACKEY, D. & SUGDEN, B. (1999B). THE LINKING REGIONS OF EBNA1 ARE ESSENTIAL FOR ITS SUPPORT OF REPLICATION AND TRANSCRIPTION. *MOL CELL BIOL* **19**, 3349-3359.

MANNICK, J. B., COHEN, J. I., BIRKENBACH, M., MARCHINI, A. & KIEFF, E. (1991). THE EPSTEIN-BARR VIRUS NUCLEAR PROTEIN ENCODED BY THE LEADER OF THE EBNA RNAS IS IMPORTANT IN B-LYMPHOCYTE TRANSFORMATION. *J VIROL* **65**, 6826-6837.

MASCIARELLI, S., MATTIOLI, B., GALLETTI, R., SAMOGGIA, P., CHICHIARELLI, S., MEARINI, G. & MATTIA, E. (2002). ANTISENSE TO EPSTEIN BARR VIRUS-ENCODED LMP1 DOES NOT AFFECT THE TRANSCRIPTION OF VIRAL AND CELLULAR PROLIFERATION-RELATED GENES, BUT INDUCES PHENOTYPIC EFFECTS ON EBV-TRANSFORMED B LYMPHOCYTES. *ONCOGENE* **21**, 4166-4170.

MASER, R. S., MIRZOEVA, O. K., WELLS, J., OLIVARES, H., WILLIAMS, B. R., ZINKEL, R. A., FARNHAM, P. J. & PETRINI, J. H. (2001). MRE11 COMPLEX AND DNA REPLICATION: LINKAGE TO E2F AND SITES OF DNA SYNTHESIS. *MOL CELL BIOL* **21**, 6006-6016.

MATTSSON, K., POKROVSKAJA, K., KISS, C., KLEIN, G. & SZEKELY, L. (2001). PROTEINS ASSOCIATED WITH THE PROMYELOCYTIC LEUKEMIA GENE PRODUCT (PML)-CONTAINING NUCLEAR BODY MOVE TO THE NUCLEOLUS UPON INHIBITION OF PROTEASOME-DEPENDENT PROTEIN DEGRADATION. *PROC NATL ACAD SCI U S A* **98**, 1012-1017.

MCCANN, E. M., KELLY, G. L., RICKINSON, A. B. & BELL, A. I. (2001). GENETIC ANALYSIS OF THE EPSTEIN-BARR VIRUS-CODED LEADER PROTEIN EBNA-LP AS A CO-ACTIVATOR OF EBNA2 FUNCTION. *J GEN VIROL* **82**, 3067-3079.

MILLER, C. L., LONGNECKER, R. & KIEFF, E. (1993). EPSTEIN-BARR VIRUS LATENT MEMBRANE PROTEIN 2A BLOCKS CALCIUM MOBILIZATION IN B LYMPHOCYTES. *J VIROL* **67**, 3087-3094.

MOLINA-PRIVADO, I., RODRIGUEZ-MARTINEZ, M., REBOLLO, P., MARTIN-PEREZ, D., ARTIGA, M. J., MENARGUEZ, J., FLEMINGTON, E. K., PIRIS, M. A. & CAMPANERO, M. R. (2009). E2F1 EXPRESSION IS DEREGULATED AND PLAYS AN ONCOGENIC ROLE IN SPORADIC BURKITT'S LYMPHOMA. *CANCER RES* **69**, 4052-4058.

MOSIALOS, G., BIRKENBACH, M., YALAMANCHILI, R., VANARSDALE, T., WARE, C. & KIEFF, E. (1995). THE EPSTEIN-BARR VIRUS TRANSFORMING PROTEIN LMP1 ENGAGES SIGNALING PROTEINS FOR THE TUMOR NECROSIS FACTOR RECEPTOR FAMILY. *CELL* **80**, 389-399.

MURRAY, P. G. & YOUNG, L. S. (2001). EPSTEIN-BARR VIRUS INFECTION: BASIS OF MALIGNANCY AND POTENTIAL FOR THERAPY. *EXPERT REV MOL MED* **3**, 1-20.

NAVRATIL, V., DE CHASSEY, B., MEYNIEL, L., DELMOTTE, S., GAUTIER, C., ANDRE, P., LOTTEAU, V. & RABOURDIN-COMBE, C. (2009). VIRHOSTNET: A KNOWLEDGE BASE FOR THE MANAGEMENT AND THE ANALYSIS OF PROTEOME-WIDE VIRUS-HOST INTERACTION NETWORKS. *NUCLEIC ACIDS RES* **37**, D661-668.

NIEDERMAN, J. C., MCCOLLUM, R. W., HENLE, G. & HENLE, W. (1968). INFECTIOUS MONONUCLEOSIS. CLINICAL MANIFESTATIONS IN RELATION TO EB VIRUS ANTIBODIES. *JAMA* **203**, 205-209.

NIELSEN, N. H., MIKKELSEN, F. & HANSEN, J. P. (1977). NASOPHARYNGEAL CANCER IN GREENLAND. THE INCIDENCE IN AN ARCTIC ESKIMO POPULATION. *ACTA PATHOL MICROBIOL SCAND [A]* **85**, 850-858.

NILSSON, T., ZETTERBERG, H., WANG, Y. C. & RYMO, L. (2001). PROMOTER-PROXIMAL REGULATORY ELEMENTS INVOLVED IN ORIP-EBNA1-INDEPENDENT AND -DEPENDENT ACTIVATION OF THE EPSTEIN-BARR VIRUS C PROMOTER IN B-LYMPHOID CELL LINES. *J VIROL* **75**, 5796-5811.

NITSCHE, F., BELL, A. & RICKINSON, A. (1997). EPSTEIN-BARR VIRUS LEADER PROTEIN ENHANCES EBNA-2-MEDIATED TRANSACTIVATION OF LATENT MEMBRANE PROTEIN 1 EXPRESSION: A ROLE FOR THE W1W2 REPEAT DOMAIN. *J VIROL* **71**, 6619-6628.

O'CONOR, G. T. (1961). MALIGNANT LYMPHOMA IN AFRICAN CHILDREN. II. A PATHOLOGICAL ENTITY. *CANCER* **14**, 270-283.

PARKER, G. A., CROOK, T., BAIN, M., SARA, E. A., FARRELL, P. J. & ALLDAY, M. J. (1996). EPSTEIN-BARR VIRUS NUCLEAR ANTIGEN (EBNA)3C IS AN

IMMORTALIZING ONCOPROTEIN WITH SIMILAR PROPERTIES TO ADENOVIRUS E1A AND PAPILLOMAVIRUS E7. *ONCOGENE* **13**, 2541-2549.

PEEPER, D. S., SHVARTS, A., BRUMMELKAMP, T., DOUMA, S., KOH, E. Y., DALEY, G. Q. & BERNARDS, R. (2002). A FUNCTIONAL SCREEN IDENTIFIES HDRIL1 AS AN ONCOGENE THAT RESCUES RAS-INDUCED SENESCENCE. *NAT CELL BIOL* **4**, 148-153.

PENG, C. W., ZHAO, B., CHEN, H. C., CHOU, M. L., LAI, C. Y., LIN, S. Z., HSU, H. Y. & KIEFF, E. (2007). HSP72 UP REGULATES EPSTEIN-BARR VIRUS EBNALP CO-ACTIVATION WITH EBNA2. *BLOOD*.

POLACK, A., HORTNAGEL, K., PAJIC, A., CHRISTOPH, B., BAIER, B., FALK, M., MAUTNER, J., GELTINGER, C., BORNKAMM, G. W. & KEMPKES, B. (1996). C-MYC ACTIVATION RENDERS PROLIFERATION OF EPSTEIN-BARR VIRUS (EBV)-TRANSFORMED CELLS INDEPENDENT OF EBV NUCLEAR ANTIGEN 2 AND LATENT MEMBRANE PROTEIN 1. *PROC NATL ACAD SCI U S A* **93**, 10411-10416.

RAAB-TRAUB, N., FLYNN, K., PEARSON, G., HUANG, A., LEVINE, P., LANIER, A. & PAGANO, J. (1987). THE DIFFERENTIATED FORM OF NASOPHARYNGEAL CARCINOMA CONTAINS EPSTEIN-BARR VIRUS DNA. *INT J CANCER* **39**, 25-29.

RADKOV, S. A., BAIN, M., FARRELL, P. J., WEST, M., ROWE, M. & ALLDAY, M. J. (1997). EPSTEIN-BARR VIRUS EBNA3C REPRESSES CP, THE MAJOR PROMOTER FOR EBNA EXPRESSION, BUT HAS NO EFFECT ON THE PROMOTER OF THE CELL GENE CD21. *J VIROL* **71**, 8552-8562.

RADKOV, S. A., TOUITOU, R., BREHM, A., ROWE, M., WEST, M., KOUZARIDES, T. & ALLDAY, M. J. (1999). EPSTEIN-BARR VIRUS NUCLEAR ANTIGEN 3C INTERACTS WITH HISTONE DEACETYLASE TO REPRESS TRANSCRIPTION. *J VIROL* **73**, 5688-5697.

RICKINSON, A. B. & KIEFF, E. (2001). EPSTEIN-BARR VIRUS. IN *FIELDS VIROLOGY*, 4TH EDN, PP. 2575-2627. EDITED BY D. KNIPE, P. HOWLEY, D. GRIFFIN, M. MARTIN, R. LAMB, B. ROIZMAN & S. STRAUS. PHILADELPHIA, PA: LIPINCOTT-RAVEN PUBLISHERS.

RIGAUT, G., SHEVCHENKO, A., RUTZ, B., WILM, M., MANN, M. & SERAPHIN, B. (1999). A GENERIC PROTEIN PURIFICATION METHOD FOR PROTEIN COMPLEX CHARACTERIZATION AND PROTEOME EXPLORATION. *NAT BIOTECHNOL* **17**, 1030-1032.

ROBERTSON, E. S., LIN, J. & KIEFF, E. (1996). THE AMINO-TERMINAL DOMAINS OF EPSTEIN-BARR VIRUS NUCLEAR PROTEINS 3A, 3B, AND 3C INTERACT WITH RBPJ(KAPPA). *J VIROL* **70**, 3068-3074.

ROVEDO, M. & LONGNECKER, R. (2007). EPSTEIN-BARR VIRUS LATENT MEMBRANE PROTEIN 2B (LMP2B) MODULATES LMP2A ACTIVITY. *J VIROL* **81**, 84-94.

SAMBROOK, J. & RUSSEL, D. W. (2001). *MOLECULAR CLONING: A LABORATORY MANUAL*. COLD SPRING HARBOR, NEW YORK: COLD SPRING HARBOR LABORATORY PRESS.

SAMPLE, J., HUMMEL, M., BRAUN, D., BIRKENBACH, M. & KIEFF, E. (1986). NUCLEOTIDE SEQUENCES OF MRNAS ENCODING EPSTEIN-BARR VIRUS NUCLEAR PROTEINS: A PROBABLE TRANSCRIPTIONAL INITIATION SITE. *PROC NATL ACAD SCI U S A* **83**, 5096-5100.

SCHAEFER, B. C., PAULSON, E., STROMINGER, J. L. & SPECK, S. H. (1997). CONSTITUTIVE ACTIVATION OF EPSTEIN-BARR VIRUS (EBV) NUCLEAR ANTIGEN 1 GENE TRANSCRIPTION BY IRF1 AND IRF2 DURING RESTRICTED EBV LATENCY. *MOL CELL BIOL* **17**, 873-886.

SHANNON-LOWE, C. D., NEUHIERL, B., BALDWIN, G., RICKINSON, A. B. & DELECLUSE, H. J. (2006). RESTING B CELLS AS A TRANSFER VEHICLE FOR EPSTEIN-BARR VIRUS INFECTION OF EPITHELIAL CELLS. *PROC NATL ACAD SCI U S A* **103**, 7065-7070.

SHAO, Y. M., POIRIER, A., OHSHIMA, H., MALAVEILLE, C., ZENG, Y., THE, G. D. & BARTSCH, H. (1988). EPSTEIN-BARR VIRUS ACTIVATION IN RAJI CELLS BY

EXTRACTS OF PRESERVED FOOD FROM HIGH RISK AREAS FOR NASOPHARYNGEAL CARCINOMA. *CARCINOGENESIS* **9**, 1455-1457.

SINCLAIR, A. & FARRELL, P. (1995). HOST CELL REQUIREMENTS FOR EFFICIENT INFECTION OF QUIESCENT PRIMARY B LYMPHOCYTES BY EPSTEIN-BARR VIRUS. *J VIROL* **69**, 5461-5468.

SINCLAIR, A. J., PALMERO, I., PETERS, G. & FARRELL, P. J. (1994). EBNA-2 AND EBNA-LP COOPERATE TO CAUSE G0 TO G1 TRANSITION DURING IMMORTALIZATION OF RESTING HUMAN B LYMPHOCYTES BY EPSTEIN-BARR VIRUS. *EMBO J* **13**, 3321-3328.

SPECK, S. H., (ED) (2005). *REGULATION OF EBV LATENCY-ASSOCIATED GENE EXPRESSION*. WYMONDHAM: CAISTER ACADEMIC PRESS.

SPECK, S. H., PFITZNER, A. & STROMINGER, J. L. (1986). AN EPSTEIN-BARR VIRUS TRANSCRIPT FROM A LATENTLY INFECTED, GROWTH-TRANSFORMED B-CELL LINE ENCODES A HIGHLY REPETITIVE POLYPEPTIDE. *PROC NATL ACAD SCI U S A* **83**, 9298-9302.

SPIRIN, V. & MIRNY, L. A. (2003). PROTEIN COMPLEXES AND FUNCTIONAL MODULES IN MOLECULAR NETWORKS. *PROC NATL ACAD SCI U S A* **100**, 12123-12128.

SUZUKI, M. (1998). A NOVEL E2F BINDING PROTEIN WITH MYC-TYPE HLH MOTIF STIMULATES E2F DEPENDENT TRANSCRIPTION BY FORMING A HETERODIMER. *ONCOGENE* **17**, 853-865.

SZEKELY, L., POKROVSKAJA, K., JIANG, W. Q., DE THE, H., RINGERTZ, N. & KLEIN, G. (1996). THE EPSTEIN-BARR VIRUS-ENCODED NUCLEAR ANTIGEN EBNA-5 ACCUMULATES IN PML-CONTAINING BODIES. *J VIROL* **70**, 2562-2568.

SZEKELY, L., POKROVSKAJA, K., JIANG, W. Q., SELIVANOVA, G., LOWBEER, M., RINGERTZ, N., WIMAN, K. G. & KLEIN, G. (1995). RESTING B-CELLS, EBV-INFECTED B-BLASTS AND ESTABLISHED LYMPHOBLASTOID CELL LINES DIFFER IN THEIR RB, P53 AND EBNA-5 EXPRESSION PATTERNS. *ONCOGENE* **10**, 1869-1874.

SZEKELY, L., SELIVANOVA, G., MAGNUSSON, K. P., KLEIN, G. & WIMAN, K. G. (1993). EBNA-5, AN EPSTEIN-BARR VIRUS-ENCODED NUCLEAR ANTIGEN, BINDS TO THE RETINOBLASTOMA AND P53 PROTEINS. *PROC NATL ACAD SCI U S A* **90**, 5455-5459.

TAO, Q., YOUNG, L. S., WOODMAN, C. B. & MURRAY, P. G. (2006). EPSTEIN-BARR VIRUS (EBV) AND ITS ASSOCIATED HUMAN CANCERS--GENETICS, EPIGENETICS, PATHOBIOLOGY AND NOVEL THERAPEUTICS. *FRONT BIOSCI* **11**, 2672-2713.

THORLEY-LAWSON, D. A., MIYASHITA, E. M. & KHAN, G. (1996). EPSTEIN-BARR VIRUS AND THE B CELL: THAT'S ALL IT TAKES. *TRENDS MICROBIOL* **4**, 204-208.

UETZ, P., DONG, Y. A., ZERETZKE, C., ATZLER, C., BAIKER, A., BERGER, B., RAJAGOPALA, S. V., ROUPELIEVA, M., ROSE, D., FOSSUM, E. & HAAS, J. (2006). HERPESVIRAL PROTEIN NETWORKS AND THEIR INTERACTION WITH THE HUMAN PROTEOME. *SCIENCE* **311**, 239-242.

WERNER, M., ERNBERG, I., ZOU, J., ALMQVIST, J. & AURELL, E. (2007). EPSTEIN-BARR VIRUS LATENCY SWITCH IN HUMAN B-CELLS: A PHYSICO-CHEMICAL MODEL. *BMC SYST BIOL* **1**, 40.

VETSIKA, E. K. & CALLAN, M. (2004). INFECTIOUS MONONUCLEOSIS AND EPSTEIN-BARR VIRUS. *EXPERT REV MOL MED* **6**, 1-16.

WILSKER, D., PROBST, L., WAIN, H. M., MALTAIS, L., TUCKER, P. W. & MORAN, E. (2005). NOMENCLATURE OF THE ARID FAMILY OF DNA-BINDING PROTEINS. *GENOMICS* **86**, 242-251.

WILSON, J. B., BELL, J. L. & LEVINE, A. J. (1996). EXPRESSION OF EPSTEIN-BARR VIRUS NUCLEAR ANTIGEN-1 INDUCES B CELL NEOPLASIA IN TRANSGENIC MICE. *EMBO J* **15**, 3117-3126.

VOGELSTEIN, B., LANE, D. & LEVINE, A. J. (2000). SURFING THE P53 NETWORK. *NATURE* **408**, 307-310.

WOISETSCHLAEGER, M., JIN, X. W., YANDAVA, C. N., FURMANSKI, L. A., STROMINGER, J. L. & SPECK, S. H. (1991). ROLE FOR THE EPSTEIN-BARR VIRUS NUCLEAR ANTIGEN 2 IN VIRAL PROMOTER SWITCHING DURING INITIAL STAGES OF INFECTION. *PROC NATL ACAD SCI U S A* **88**, 3942-3946.

WOISETSCHLAEGER, M., YANDAVA, C. N., FURMANSKI, L. A., STROMINGER, J. L. & SPECK, S. H. (1990). PROMOTER SWITCHING IN EPSTEIN-BARR VIRUS DURING THE INITIAL STAGES OF INFECTION OF B LYMPHOCYTES. *PROC NATL ACAD SCI U S A* **87**, 1725-1729.

WOLF, H., HAUS, M. & WILMES, E. (1984). PERSISTENCE OF EPSTEIN-BARR VIRUS IN THE PAROTID GLAND. *J VIROL* **51**, 795-798.

XIE, P., HOSTAGER, B. S. & BISHOP, G. A. (2004). REQUIREMENT FOR TRAF3 IN SIGNALING BY LMP1 BUT NOT CD40 IN B LYMPHOCYTES. *J EXP MED* **199**, 661-671.

YATES, J., WARREN, N., REISMAN, D. & SUGDEN, B. (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**, 3806-3810.

YI, F., SAHA, A., MURAKAMI, M., KUMAR, P., KNIGHT, J. S., CAI, Q., CHOUDHURI, T. & ROBERTSON, E. S. (2009). EPSTEIN-BARR VIRUS NUCLEAR ANTIGEN 3C TARGETS P53 AND MODULATES ITS TRANSCRIPTIONAL AND APOPTOTIC ACTIVITIES. *VIROLOGY* **388**, 236-247.

YOKOYAMA, A., TANAKA, M., MATSUDA, G., KATO, K., KANAMORI, M., KAWASAKI, H., HIRANO, H., KITABAYASHI, I., OHKI, M., HIRAI, K. & KAWAGUCHI, Y. (2001). IDENTIFICATION OF MAJOR PHOSPHORYLATION SITES OF EPSTEIN-BARR VIRUS NUCLEAR ANTIGEN LEADER PROTEIN (EBNA-LP): ABILITY OF EBNA-LP TO INDUCE LATENT MEMBRANE PROTEIN 1 COOPERATIVELY WITH EBNA-2 IS REGULATED BY PHOSPHORYLATION. *J VIROL* **75**, 5119-5128.

YOUNG, L. S. & RICKINSON, A. B. (2004). EPSTEIN-BARR VIRUS: 40 YEARS ON. *NAT REV CANCER* **4**, 757-768.

ZETTERBERG, H., BORESTROM, C., NILSSON, T. & RYMO, L. (2004). MULTIPLE EBNA1-BINDING SITES WITHIN ORIPI ARE REQUIRED FOR EBNA1-DEPENDENT TRANSACTIVATION OF THE EPSTEIN-BARR VIRUS C PROMOTER. *INT J ONCOL* **25**, 693-696.

ZETTERBERG, H., RYMO, L., (ED) (2005). *EBNA2 TRANSCRIPTION REGULATION IN EBV LATENCY*. WYMONDHAM: CALISTER ACADEMIC PRESS.

ZHANG, L., VILLA, N. Y., RAHMAN, M. M., SMALLWOOD, S., SHATTUCK, D., NEFF, C., DUFFORD, M., LANCHBURY, J. S., LABAER, J. & MCFADDEN, G. (2009). ANALYSIS OF VACCINIA VIRUS-HOST PROTEIN-PROTEIN INTERACTIONS: VALIDATIONS OF YEAST TWO-HYBRID SCREENINGS. *J PROTEOME RES* **8**, 4311-4318.

ZONG, R. T., DAS, C. & TUCKER, P. W. (2000). REGULATION OF MATRIX ATTACHMENT REGION-DEPENDENT, LYMPHOCYTE-RESTRICTED TRANSCRIPTION THROUGH DIFFERENTIAL LOCALIZATION WITHIN PROMYELOCYTIC LEUKEMIA NUCLEAR BODIES. *EMBO J* **19**, 4123-4133.