To Elucidate the Epstein-Barr Virus Replisome

By Christopher Traylen

A Thesis submitted for the degree of Doctor of Philosophy

School of Life Sciences

University of Sussex

September 2015

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another University for the award of any other degree.

Signature:.....

Ackowledgements

I would first like to thank my supervisor Professor Alison Sinclair for all the guidance and supervision since the first day of my PhD. I appreciate all the support and time Alison was given to me to develop myself the best way possible. All of the members of the Sinclair lab have made my experience very enjoyable and provided a great work environment. Past members Dr. Nicolae Balan and Dr Sharada Ramasubramanyan have helped me as colleagues and as friends. Dr. Kay Osborn was always available to ask for help and I also thank everyone else who helped me throughout my time in the lab. I'd like to thank Professor Michelle West for her support and advice, and also members of the West lab who have helped me. I appreciate both Dr. David Wood and Dr. Michael McClellan for being good friends as well as work colleagues. My parents have given me the best support possible in able for myself to complete my studies. I would not have been able to finish without their help.

UNIVERSITY OF SUSSEX

CHRISTOPHER TRAYLEN

PhD BIOCHEMISTRY

To Elucidate the Epstein - Barr Virus Replisome

SUMMARY

Epstein - Barr virus (EBV) is a member of the γ-herpesvirus subfamily of Herpesviridae. EBV is a double stranded DNA virus infecting humans causing a variety of disease from asymptomatic infection to association with certain tumours including Burkitts lymphoma, Hodgkin's disease and nasopharyngeal carcinoma. EBV encodes an immediate-early protein called Zta (BZLF1, EB1, ZEBRA), which is an important transcription factor and replication factor direct in disrupting latency. EBV encodes viral proteins that assemble as a replisome at the viral lytic origin recognition site (Ori-Lyt). Zta binds Ori-Lyt and it is unclear how Zta interacts and recruits the complex to the site of DNA replication, while coordinating and recruiting host factors. After a mutation to three alanines (ZtaAAA) data implicates that the extreme C-terminus of Zta is essential for replication.

The question posed is how does Zta assemble the replisome? Identification of the lytic changes that contribute to lytic replication, including cellular components that may contribute to EBV replication is attempted.

Transfected control, Zta and ZtaAAA in HEK293-BZLF1-KO cells was compared. Size exclusion chromatography identified a higher molecular weight complex containing Zta during viral replication. SILAC (Stable isotope labelling by amino acids in cell culture) coupled to proteomics analysis identified the elution fraction composition. An interpretation of these cellular components in the context of lytic replication is explored. Identification of interactions of Zta with cellular proteins was attempted by SILAC histidine tagged Zta with pull down assay. Quantitative data was returned and a confirmation of interactions was attempted. A global proteomics approach was also performed. An enrichment method to isolate SILAC labeled Burkitts Lymphoma cells undergoing EBV lytic replication was coupled to mass spectrometry analysis to identify changes in host and viral proteins.

Overall, cellular targets that may interact with Zta are to be confirmed. The global proteomics study recognized for the first time by proteomic analysis the identification of three EBV lytic replication cycle protein

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AP-1	Activator Protein 1
BART	BamH1 A transcripts
BCR	B Cell Receptor
BL	Burkitt's Lymphoma
BP	Base Pair
bZIP	basic-leucine zipper
C/EBPα	CCAAT/enhancer binding protein alpha
SBP	Creb Binding Protein
ChIP	Chromatin Immunoprecipitation
Cp	latency promoter within BamH1 C digestion fragment
CpG	Cytosine-phosphate-Guanine
CREB	cAMP Response Element binding
CR2	Complement receptor type 2
CT	Carboxy-terminus
DDR	DNA Damage Response
DMEM	Dulbecco/Vogt Modified Eagle's Minimal Essential Medium
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DPBS	Dulbecco’s Phosphate-Buffered Saline
DS	Dyad Symmetry
EBER	Epstein Barr encoded RNA
EBNA	Epstein Barr Nuclear Antigen
EBNA-LP	Epstein Barr Nuclear Antigen Leader Protein
EBV	Epstein – Barr Virus
ECL	Enhanced Chemiluminescence
FCS	Foetal Calf Serum
FR	Family of Repeats
GFP	Green Fluorescent Protein
Gp	Glycoprotein
HCMV	Human Cytomegaolvirus
HD	Hodgkins Disease
HDAC	Histone deacetylase
HEK293	Human embryonic kidney 293 cell line
HHV-1	Human Herpes Virus 1, or Herpes Simplex Virus 1 (HSV1)
HHV-2	Human Herpes Virus 2, or Herpes Simplex Virus 2 (HSV2)
HHV-3	Human Herpes Virus 3, or Varicella Zoster Virus (VZV)
HHV-4	Human Herpes Virus 4, or Epstein Barr Virus (EBV)
HHV-5	Human Herpes Virus 5, or Human Cytomeglovirus (HCMV)
HHV-8	Human Herpes Virus 8, or Kaposi's Sarcoma-associated Herpesvirus (KSHV)
HIV	Human Immunodeficiency Virus
HL	Hodgkin's Lymphoma
HLA	Human Leukocyte Antigen
HRP	Horse Radish Peroxidase
IgG	Immunoglobulin G
IgM	Immunoglobulin M
IM	Infectious Mononucleosis
KO	Knock-out

KSHV	Kaposi's Sarcoma-associated Herpesvirus
LCL	Lymphoblastoid Cell Lines
LMP	Latent Membrane Protein
MHC	Major Histocompatibility Complex
MOPS	3-(N-morpholino)propanesulfonic acid
MS	Multiple Sclerosis
mRNA	Messenger Ribonucleic Acid
NaCl	Sodium Chloride
NCBI	National Center for Biotechnology Information
NHDL	Non-Hodgkin's Lymphoma
NFκB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NGFR	Neuronal Growth Factor Receptor
NPC	Nasopharyngeal Carcinoma
OHL	Oral Hairy Leokplakia
ORF	Open Reading Frame
oriLyt	Origin of lytic replication
oriP	Origin of plasmid replication
P	Pellet
PI	Protein Interaction
PKC	Protein Kinase C
PML	Promyelocytic leukaemia
PSG	Penicillin, Streptomycin, L-Glutamine
PTLD	Post Transplantation Lymphoproliferative Disease
Qp	Latency promoter within BamH1 Q digestion fragment
QPCR	Quantitative polymerase chain reaction
RA	Rheumatoid Arthiritis
RNA	Ribonucleic Acid
Rp	Promoter of *BRLF1*
Rpm	Rotation per minute
RPMI	Roswell Park Memorial Institute medium
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SLE	Systemic lupus erythematous
SM	Protein product of mRNA spanning genes BSLF1 BMRF1
SN	Supernatant
SS	Sjögren's syndrome
TBE	Tris-Borate-EDTA
TPA	12-0-tetradecanoyl phorbol-13-acetate
TMV	Tobacco Mosaic Virus
TR	Terminal Repeats
v/v	Volume / volume
VZV	Varicella Zoster Virus
w/v	Weight / volume
WB	Western Blot
Wp	Latency promoter within BamH1 W digestion fragment
ZRE	Zta Response Element

1. Introduction

1.1. Historical Background of Virology

The scientific study of microbiological diseases began in the late 1800s. The ability to separate some microorganisms using filters initiated the understanding that microbes were causative of disease and this was well documented. In 1840, Henle reported that these infectious agents could not be seen by the light microscope and by the late 19th century Robert Koch had introduced criteria designed to establish the relationship between the microorganism and disease termed Koch's postulates. With the development of molecular biology and our understanding of pathogens that cause disease, these criteria became limited. The contribution of infectious agents towards the development of cancer meant that the criteria could not be fulfilled. Also, subclinical infections that include the majority of herpesviruses present themselves as an asymptomatic infection for an indefinite period, without causing cancer in the majority of the world population. Therefore, revised versions of the criteria proposed to include viruses, as they do not fit into the criteria originally suggested (Fredricks & Relman 1996).

Virology research began with the observation of a disease first described as lesions of tobacco plant leaves by Adolf Mayer in 1879 (Lustig & Levine 1992). Transmission of the virus was performed by inoculating healthy plants with the liquid from viral infected leaves. Beijerinck displayed that this agent could replicate itself in living tissues and referred to this as a contagious living liquid. In addition to plant viruses, bacteriophages can infect bacteria. These viruses that are parasitic to bacteria were discovered in the early 1900s (Duckworth 1976). It also became apparent viruses can infect animal cells and viruses infect humans causing disease including cancer.

The pathology of diseases was described but the causative agents were not determined until the ability for these to be visualized and improved molecular biology techniques were established. It wasn't until the invention of the electron

microscope that viruses were visualized and in 1939 the first images of the tobacco mosaic virus (TMV) were developed (Levine & Enquist 2007).

The concept of a transmissible agent that could be the foundation of oncogenesis in animals began when the Rous sarcoma virus was found to have oncogenic properties, the cell free transmission of filtrate could convey a tumour sarcoma in chickens (Rous 1910; Rous 1911). However, the idea that an agent that could cause cancer in humans was not accepted until the discovery of Epstein - Barr virus (EBV) by Epstein (Epstein et al. 1964). When EBV was discovered it was established as the first tumour associated virus identified in humans. Since the middle of the 20th century, tumour virology has become one major focus of cancer research. It is now known that viruses cause at least 10% of human cancers. Research into the understanding of viral pathogenesis will reveal therapeutic targets against which drugs can be developed that will contribute to the prevention and treatment of human cancers (Levine & Enquist 2007).

1.2. Herpesviruses

The herpesvirus family consists of many DNA viruses that undergo a true latency period in a wide range of different cells. They are large enveloped, double stranded DNA viruses and are extremely widespread among humans, and considered to be one of the most successful virus family that infects the human population. Herpesvirus virions are spherical with a core, capsid, tegument and envelope to enclose the viral DNA. They have been widely studied and their genomes sequenced (Whitley, 1996). At least eight herpesviruses have been identified that infect humans. It is estimated that the eight human herpesviruses share a significant homology in about 40 genes (McGeoch 1989). Between herpesviruses, the size of their DNA varies in size from 124kbp to 235kbp containing about 70 to 200 genes (Roizmann et al. 1992). There are three major subfamilies of herpesviruses that include alpha (α), beta (β, and gamma (γ) classified by the International Committee on Taxonomy of Viruses. Alpha herpesviruses can infect a wide range of species, whereas beta and gamma herpesviruses have restricted infectivity (Davison 2007).

Alpha herpesviruses include Human herpesvirus 1 (HHV-1) known as Herpes simplex virus 1 (HSV1), Human herpesvirus 2 (HHV-2) known as Herpes simplex

virus 2 (HSV2), and Human herpesvirus 3 (HHV-3) known as Varicella Zoster Virus (VZV) and these all infect neurons. Beta-herpesviruses include Human herpesvirus 5 (HHV-5) known as human cytomegalovirus (HCMV), human herpesvirus 6 (HHV6) and human herpesvirus 7 (HHV7) that infect monocytes. Gamma-herpesviruses include Human herpesvirus 4 (HHV4) known as Epstein-Barr virus (EBV) and Human herpesvirus 8 (HHV-8) known as Kaposi's Sarcoma Herpesvirus (KSHV), these infect lymphocytes and either epithelial or endothelial cells. Both of these viruses have been shown to cause cancer in humans. EBV and KSHV are also known as the lymphocryptoviruses (Longnecker & Neipel 2007) or gamma-1 herpes virus (for EBV), gamma-2 herpes virus (for KSHV) (Crawford et al. 2014).

Following infection of cells by herpesviruses, the linear viral DNA locates to the nucleus. Depending on the cellular environment either viral replication proceeds or viral latency is established. During latency, a small subset of latent genes are expressed to maintain the viral episome in cells until a cellular event initiates the reactivation into lytic cycle. A cascade of lytic genes can then express proteins to promote replication of the viral genome within the nucleus. During lytic cycle around 70 genes are expressed and these contribute to genome viral replication and virion production.

1.3. Epstein - Barr virus

Epstein - Barr virus (EBV) infects 90% of the world population (Young & Rickinson 2004). EBV research began after the discovery and categorisation of tumours that were present in people living in regions of Africa by Denis Burkitt (Burkitt 1958). The distinct tumours were first described in 1958 and named as Burkitt's lymphoma (BL). From the publications of this data and attending various lectures and discussions, tumour virologist Epstein, with Barr and Achong studied biopsies taken from these tumours. A virus was isolated from one of the BL cell line from Africa and this virus was identified under an electron microscope in 1964, (Epstein et al. 1964). The virus was named Epstein-Barr virus after the research team of Epstein and Barr. Since the breakthrough of the identification of this virus, many other viruses have been discovered that contribute to the

pathogenesis of cancer. These include human papillomavirus, Merkel cell polyomavirus and Kaposi's sarcoma herpesvirus (KSHV) (Moore & Chang 2010)

EBV has been extensively studied since its discovery in 1964, although we still do not fully understand how it causes disease. As with other herpesviruses, survival in the body is associated with the latent state, with only a sporadic switch into its lytic cycle to spread infectious virions. Both latency and lytic cycles have been studied in the contribution of the viral mechanisms towards disease.

1.4. EBV Disease

Although EBV infects the majority of the world population asymptomatically, the virus can have oncogenic properties (Young & Rickinson 2004). The beginning of the association with EBV and disease was demonstrated using immunofluorescence studies by the Henle laboratory in the 1960s. This initially indicated that antibodies against the virus now known as EBV are found in patients with very high levels when presenting with infectious mononucleosis (Henle & Henle 1966). Most diseases that arise are associated with the latent form of EBV, which is unlike other human herpes viruses, where disease is related to the lytic phase of the virus. EBV exhibits a distinct tropism for both B cells and epithelial cells and EBV infection of B cells results in 'one step' immortalization in vitro, but within the body this is countered by immunosurveillance. When this fails, the EBV-associated diseases progress. The mechanisms of EBV infection and contribution to disease in both B cells and epithelial cells is beginning to be more understood.

1.5. EBV associated with Cancer

1.5.1. Burkitt's lymphoma

Denis Burkitt first described the disease now known as Burkitt 's lymphoma (BL) as outlined previously. This disease is common in children was first seen as facial swellings and later classified as a lymphoma (Burkitt 1983). As described earlier, these samples led to the first identification of EBV.

There are three variants of BL that differ in their biology, presentation and association with EBV. These are endemic, sporadic and immunodeficiency

variants (Kutok & Wang 2006). EBV is present in 95% of all endemic BL cases in Africa. BL is prevalent in equatorial Africa and Papua New Guinea, where EBV plays a central role, causing an incidence rate over 50-100 per 1,000,000 individuals (Kutok & Wang 2006). A sporadic form of BL is evident, where 10–15% tumours are EBV-positive. There is a low base-line rate of BL developing without EBV in all populations worldwide, and that any increases above this baseline are due to the influence of EBV (Magrath 2012).

The majority of BL in patients is characterized by a chromosomal translocation between chromosome 8 and 14 (Zech et al. 1976), involving a c-myc proto-oncogene and an immunoglobulin gene. The activated immunoglobulin gene leads to the overexpression of myc, a transcription factor involved in many cellular targets that ultimately leads to tumorigenesis (Baumforth et al. 1999). The mark of this lymphoma is the deregulation of myc contributes to BL by clonal expansion, mutagenesis, and escape from immune surveillance (Orem et al. 2007). EBV also contributes to the pathogenesis of BL by providing anti-apoptotic signals. These override c-myc-induced cell death (Rowe, Fitzsimmons, et al. 2014). Altered myc expression may also replace EBV-driven cell proliferation and allow cells to survive and proliferate with downregulation of the Epstein-Barr Nuclear Antigen (EBNAs) and latent membrane proteins (LMPs), which may in turn enable the infected cells to evade immunosurveillance (Rowe et al. 1987).

BL cells display an EBV latency I expression, of EBNA1 with Epstein–Barr virus-encoded small RNAs (EBERs) (Kutok & Wang 2006). EBNA1 and EBERs have an effect on cell growth in some experimental systems. EBV may act by increasing the frequency of genome instability within the B-lymphocytes the virus infects.

1.5.2. Hodgkin’s lymphoma

Hodgkin’s disease (HD) or Hodgkin’s lymphoma is classified by the presence of Reed-Sternberg cells, which are derived from B-lymphocytes. The evidence that EBV may play a role in the development of HD was the observation of raised antibody titres against EBV antigens when compared to other lymphoma patients (Levine et al. 1971). It was first demonstrated that there was an increased

occurrence of EBV DNA in HD tissue specimens (Weiss et al. 1991). EBV was then further suggested to contribute to the development of the disease (Murray et al. 1992).

HD is a tumour of the lymph nodes and can be diverse in morphology. HD has been classified into four histological subtypes: lymphocyte predominance (LP), nodular sclerosis (NS), mixed cellularity (MC), and lymphocyte depletion (LD), based on the morphology of Reed-Sternberg cells (Harris et al. 1994). EBV is mostly associated with the MC subtype of HD (Pallesen et al. 1991). HD of patients greater than 55 years of age, and children less than 10 years of age are most likely to develop HD that is EBV associated (Armstrong et al. 1998).

LMP1 has a high level of expression in Reed-Sternberg cells and there is a correlation between LMP1 expression and EBV positive HD (Durkop et al. 1999; Murray et al. 2001). LMP1 can prevent entry into the lytic cycle via NFκB and downregulating BLIMP1α, required for plasma cell differentiation and induction of the lytic cycle (Prince et al. 2003; Vrzalikova et al. 2011). Whereas LMP2 can induce the lytic cycle acting as a B-cell receptor homologue (Schaadt et al. 2005). The loss of the B-cell receptor contributes to the progression of classical HD, which may suggest that cells without a BCR or BCR signalling are positively selected towards EBV-associated HL, with latent EBV proteins playing a central role (Vockerodt et al. 2013) presenting as a Latency II disease (Cesarman & Mesri 1999).

1.5.3. Nasopharyngeal Carcinoma

Nasopharyngeal carcinoma (NPC) (Type 3) has the most consistent association with EBV in the world population. This undifferentiated form of nasopharyngeal carcinoma is characterized by the presence of undifferentiated carcinoma cells. These are joined with prominent infiltrating lymphocytes; this is believed to be important for the growth of the tumour cells (Young & Murray 2003). NPC is common in areas of China and South-East Asia, with environmental cofactors adding to genetic factors (Yu et al. 1986).

EBV and the expression of EBNA proteins were first shown to be present in NPC tumour cells (zur Hausen et al. 1970). Serological screening of elevated antibody titres have aided diagnosis and monitoring therapy of NPC (Zeng 1985) as the association is well established between EBV and NPC. It has been proposed that NPC may arise from the expansion of a single EBV infected nasopharyngeal epithelial cell (Pathmanathan et al. 1995). LMP1 may be central to the pathogenesis of the disease. LMP1 inactivation of cellular pathways contributes to the proliferation and transformation of epithelial cells (Lo et al. 2013). NPC can present as a Latency II disease (Cesarman & Mesri 1999), again with latent proteins playing a role in the pathogenesis of disease.

1.5.4. Post Transplant Lymphoproliferative Disorders (PTLD)

Nearly all Post Transplant Lymphoproliferative Disorders (PTLDs) are associated with EBV. It is the most common disease associated with transplantation, arising in 10% of all transplant recipients (Penn 1994). Immunosuppressive agents that are used to prevent rejection of transplants result in a lack of immune control against EBV, this leads to unrestricted proliferation of B-lymphocytes and possible tumour formation. The infection of B-lymphocytes with EBV can lead to a hyper-proliferative state that can vary in severity (Loren et al. 2003).

The origin of EBV in this disease varies. Bone marrow transplantation can lead to the reactivation of EBV from donor cells containing EBV (Zutter et al. 1988). Patients who are seronegative for EBV have a higher risk of developing PTLD after transplantation, as they have no T cell response to EBV. A 10-75-fold incidence of PTLD occurs in recipients that are seronegative over donor seropositive transplant patients. The main immune control of EBV primary infection is a CD8+ T cell response. Tumours arising from PTLD include polymorphic PTLD or monomorphic PTLD, arising from B-lymphocytes.

1.5.5. NK and T Cell lymphomas

EBV has been found in many T cell and natural killer (NK) lymphoproliferations but this is a rare event. EBV can enter B cells through the CD21 receptor (Fingeroth et al. 1984), but T cells including NK cells lack CD21 expression (Fox et al. 2011). The EBV role in pathogenesis of T cell lymphomas has been investigated.

Five T-cell tumours carry EBV DNA: aggressive NK cell leukaemia, extranodal NK/ T-cell lymphoma nasal type, enteropathy-type T-cell lymphoma, angioblastic T-cell lymphoma, and peripheral T-cell lymphoma, unspecified (Jaffe & Ralfkiaer 2001). The model for B-cell infection and EBV life cycle cannot explain how these diseases may arise, raising questions such as the mechanism of infection *in vivo* and the heterogeneity of infection within the same tissue sample (Fox et al. 2011).

1.6. Other diseases associated with Epstein - Barr virus

1.6.1. Infectious Mononucleosis

The link between EBV and infectious mononucleosis (IM) was discovered by the Henle laboratory, as described previously (Henle et al. 1968). EBV was also shown to replicate in oropharyngeal epithelial cells during infectious mononucleosis (Lemon et al. 1977).

IM is a symptomatic infection. IM generally lasts a few weeks but some symptoms will remain such as fatigue for months after the initial primary infection (Rea et al. 2001). After infection, an incubation period of four to six weeks is evident (Hoagland 1955), where a high viral load can be detected in the oral cavity and blood (Balfour et al. 2015). IM is characterized by a global expansion of CD8+ T cells with activated T cell response to EBV infected B cells (Hislop 2015). Once the infection is controlled by the immune system, a delayed antibody response against EBNA1 develops (Henle et al. 1987) IM may be relevant to cancer as a history of IM leads to a higher risk of developing EBV-positive HD (Hjalgrim et al. 2010).

1.6.2. Oral Hairy Leukoplakia

Oral hairy leukoplakia (OHL) was first described in 1984 and is a mucosal disease associated with EBV, with all cases occurring in immunosuppressed patients. The disease is present on the tongue of these patients, and is common with immunosuppressed individuals infected with HIV (Reichart et al. 1989). EBV can replicate in oral epithelial cells (Greenspan et al. 1985). B95-8 EBV from HEK293 cells shows a preferential to infect differentiated primary epithelial cells (Feederle

et al. 2007), displaying EBV can preferentially infect epithelial cells at a high infection rate, possibly due to the differentiation of epithelial cells.

The disease can be treated with acyclovir, a viral DNA replication inhibitor, although acyclovir-resistant EBV has been detected in patients treated with acyclovir for oral hairy leukoplakia (Walling et al. 2003). It is unknown whether treatment failures for oral hairy leukoplakia occur because of sub-therapeutic levels of acyclovir or an EBV resistance to acyclovir.

1.6.3. Autoimmune diseases

EBV is linked with systemic autoimmune diseases such as rheumatoid arthritis (RA), Sjögren's syndrome (SS), systemic lupus erythematous (SLE) and multiple sclerosis (MS). These diseases may have genetic and environmental factors contributing to their development and EBV is suggested to play a role, however this remains controversial (Fust 2011).

Rheumatoid arthritis (RA) involves a persistent synovial inflammation of the joints, B and T cells with macrophages and cytokines infiltrate and cause tissue damage and cartilage destruction (Scott et al. 2010). Infection with EBV has been suggested to contribute to RA, the frequency of EBV being higher in patients with RA (Balandraud et al. 2003)

Sjögren's syndrome (SS) can present as disorders with the exocrine glands and lymphocyte infiltration (Jonsson et al. 2011). EBV antigens and DNA were detected from biopsies of patients with severe SS, indicating the virus may play a role in the pathogenesis (Fox et al. 1986).

Systemic lupus erythematous SLE presents as a rare autoimmune disease that may include the failed clearance of the early phases of apoptosis, leading to an autoimmune reaction (Gaipl et al. 2007). SLE patients have an abnormally high frequency of EBV-infected cells in the host, independent of immunosuppressive therapy (Gross et al. 2005). Also EBV mRNA expression was significantly increased in SLE (Poole et al. 2009).

Multiple sclerosis (MS) has also been linked to EBV infection. MS is a chronic disease of the nervous system, evident with demyelination of the myelin sheaths surrounding axons. In vitro cells from patients with MS had an increased tendency for spontaneous EBV induced B-lymphocyte transformation (Fraser et al. 1979). Molecular mimicry may play a role in disease as CD4+ T lymphocytes specific for myelin and the EBV DNA polymerase protein have been isolated (Wucherpfennig & Strominger 1995).

1.7. EBV genome

EBV is a double-stranded DNA virus of about 172kb in length and can express around 86 latent or lytic cycle genes. Latent promoters are clustered near the terminal repeats of the fused genome, where the viral DNA circularizes maintaining as a viral episome (Speck & Ganem 2010). The genome contains the Wp promoter which is the initial promoter expressed after infection. The genome is also composed of multiple exons for latent genes. EBNAs are transcribed from the Cp or Wp promoter during latency III, and different EBNA products arise from differential splicing of the primary transcript (Figure 1.1) (Young & Murray 2003). A switch from Wp to Cp during the establishment of latency represents a shift from control by cellular transcription factors to the control by viral transcription factors (Speck & Ganem 2010). The Qp region is active in Latency I and II.

Currently there are two major subtypes of EBV that have been identified worldwide, EBV-1 and EBV-2 (Zimber et al. 1986). These subtypes differ in their geographical distribution, EBV-1 is established in most Western populations whereas both EBV-1 and EBV-2 are widespread in equatorial Africa (Sixbey et al. 1989).

Prior to recent advances in DNA sequencing, DNA polymorphisms were used to determine EBV type in tissue samples. Some of the nucleotide sequence from the B95-8 EBV genome was first determined from BamHI fragments of the virus (Cheung & Kieff 1982; Dambaugh & Kieff 1982). A BamHI digest created fragments of the viral DNA and a cloned library was created. The viral DNA was cloned as restriction fragments and into vectors, sequenced by the Sanger

method. The restriction fragments were named by size and open reading frames determined by their orientation from left (L) to right (R) (Baer et al. 1984). The B95-8 EBV-1 virus was the first herpesvirus to be completely sequenced (Baer et al. 1984).

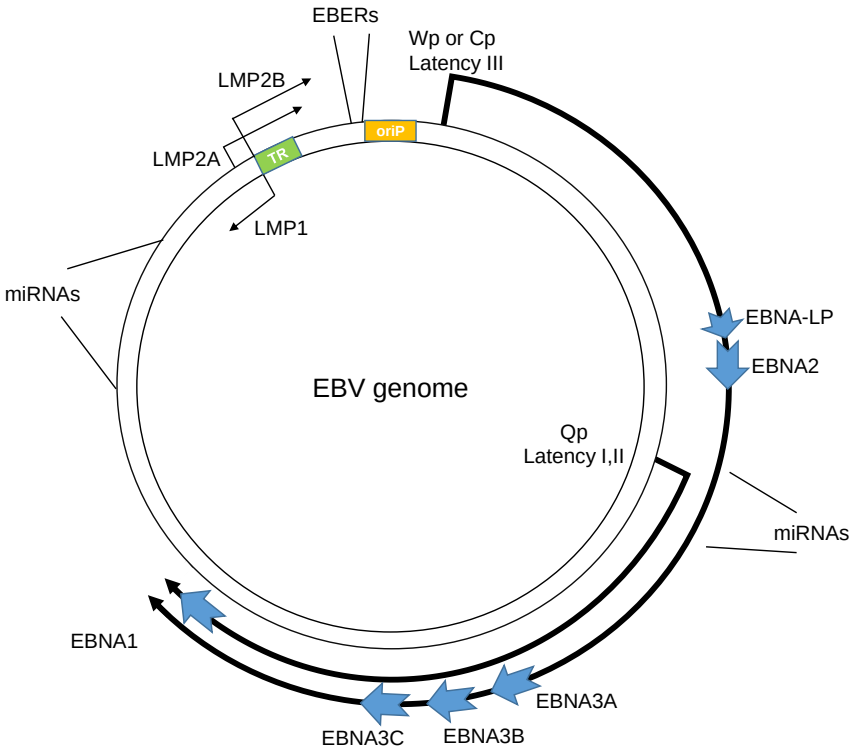


Figure 1.1 Diagram showing the location and transcription of the EBV latent genes on the double-stranded viral DNA episome. The linear viral DNA ligates together at the Terminal Repeats (TR) shown in green. The origin of plasmid replication (oriP) is shown in orange. The arrows indicate the direction in which they are transcribed; the latent proteins include the six nuclear antigens (EBNAs 1, 2, 3A, 3B and 3C, and EBNA-LP) and the three latent membrane proteins (LMPs 1, 2A, 2B). The EBNAs are transcribed from either Wp or Cp promoter in latency III. EBNA1 is transcribed in Latency I or II from Qp. EBERs are also highly transcribed. Diagram of genome from (Young & Murray 2003). MicroRNAs are also expressed from regions of the genome (Zhu et al. 2009; Lin et al. 2013; Pfeffer et al. 2004).

The EBV GD1 (Guangdong strain 1) sequence was assembled from a lymphoblastoid cell line obtained from a patient with nasopharyngeal carcinoma (NPC) in China (Zeng et al. 2005) and a Type 2 sequence was annotated from an African BL cell line AG876 (Dolan et al. 2006). EBV within the Raji cell line has also been sequenced, this genome has a 12kb deletion compared to the B95-8 virus sequence (Parker et al. 1990). This deleted sequence has been shown to include the BALF2 gene, which is essential for lytic replication. Therefore within Raji cells EBV is unable to replicate in viral lytic cycle due to the absence of the lytic replisome protein.

An updated wild type EBV sequence was drawn together from the B95-8 sequence and the Raji sequence. This aided the assembly of the current standard reference sequence of EBV, which is the RefSeq HHV4 (EBV) sequence located on the NCBI Reference Sequence database (RefSeq Accession NC_007605). To date, 84 EBV genomes have been compared which is currently the most comprehensive analysis so far (Palser et al. 2015).

There are variations between the genome sequences for EBV genes. EBV-1 and EBV-2 have a distinct variation in the EBNA2 gene with a 70% identity at the genome sequence level and 54% identity in protein sequences (Tzellos & Farrell 2012). EBV-1 and EBV-2 also differ in the organization of genes that code for the Epstein-Barr Nuclear Antigen (EBNA) proteins (Sample et al. 1990). These include EBNA2, EBNA3A, EBNA3B and EBNA3C.

1.8. EBV Life Cycle

1.8.1. Virion Structure

Mass spectrometry analysis has identified components of the EBV virion (Johannsen et al. 2004). The capsid has an icosahedral shape that is surrounded by tegument proteins and a lipid envelope covered in glycoproteins. Electron microscopy revealed a structure similar to other herpesviruses. On the viral capsid, the viral glycoprotein gp350/220 attaches to the CD21 (Cr2) molecule located on the surface of B-lymphocytes (Fingeroth et al. 1984; Nemerow et al. 1985). The viral protein is abundantly expressed and has two forms named gp350 and gp220 (Hummel et al. 1984). Glycoprotein gp350 is highly abundant, followed by glycoprotein gH, gB, gM, gp42, gL, gp78 and gp150 (Johannsen et al, 2004). These viral factors are essential for viral entry to cells. The viral glycoprotein gp42 binds to HLA class II cell surface protein (CD74), which acts as a cofactor (Li et al. 1997). Also gH has a role in attachment and penetration into epithelial cells (Molesworth et al. 2000), along with gp85 (Oda et al. 2000). gB is also important to mediate viral entry into cells (McShane & Longnecker 2004).

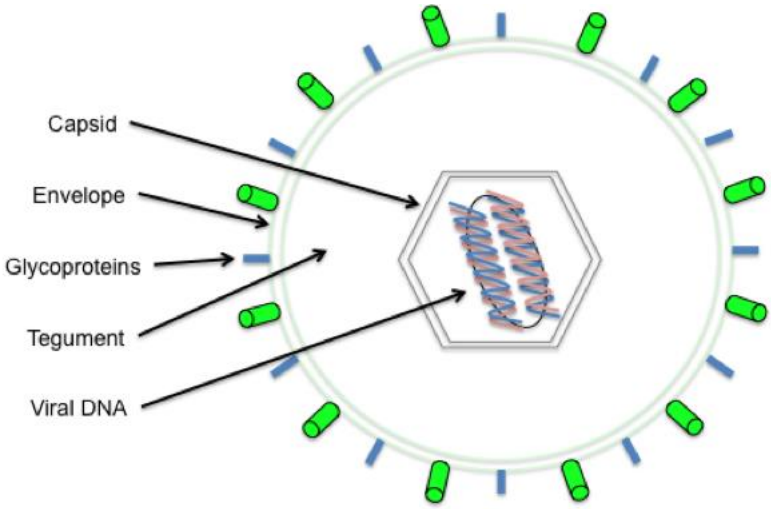


Figure 1.2 Schematic diagram illustrating the multilayer organization of human herpesviruses. The virion contains a capsid, envelope, glycoproteins and tegument that surround viral DNA. Adapted from (Liu & Zhou 2007)

The main cells for EBV infection include B-lymphocytes and epithelial cells. The virus is spread by oral transmission (Cohen 2000). Virus is passed to B-lymphocytes that then become infected. The method of entry between cell types becomes complex as the presence of gp350, essential for B cell entry, is inhibitory for the entry into epithelial cells (Shannon-Lowe et al. 2006). The fusion of virion envelope with the cellular membrane mediates entry into B-lymphocytes following endocytosis, and entry into epithelial cells may be different (Miller & Hutt-Fletcher 1992). EBV can enter epithelial cells through CD21 independent pathways (Tugizov et al. 2003). Transfection of epithelial cell lines with the receptor CD21 shows an increased efficiency of infection, expressing only EBERs and EBNA1 (Li et al. 1992). Transfer infection of EBV to efficiently infect epithelial cells involves gH and gL glycoproteins but not gp42 (Shannon-Lowe & Rowe 2011). Therefore the method of viral entry into B cells and epithelial cells are different. Infection of epithelial cells can be from direct and B-cell mediated transmission, and this suggests that EBV infects epithelial cells using both pathways (Feederle et al. 2007).

Once the virus enters the cell then cellular pathways are activated. The signal transduction pathways involved in EBV infection may include tyrosine and phosphoinotiside-3 kinases, both of which contribute to viral gene expression after infection (Sinclair & Farrell 1995). This exploitation of the host signal transduction pathway enables efficient infection by the virus. After successful entry into cells, the viral genome is transported to the nucleus. The linear genome

circularises to form a closed circular episome, through its terminal repeats (TR) sequences (Kintner & Sugden 1979). These TRs are approximately 500bp in size and circularise by homologous recombination. The expression of genes LMP2a and LMP2b (also termed Tp1 and Tp2) genes require this circularization of the linear EBV genome (Laux et al. 1989). This produces fused termini of unique length for each independent circularization. EBV in the virion is nucleosome free (Shaw et al. 1979) and the virus exists as an extrachromosomal element in the cell nucleus (Nonoyama & Pagano 1972). The latent phase enables the viral episome to be maintained within the nucleus while being associated with histones in a similar manner to cellular chromatin (Tempera & Lieberman 2010).

EBV persists in a circulating subset of resting memory B cells at a frequency of ~1 in 1×10^5 to 1×10^6 cells (Thompson & Kurzrock 2004). The viral DNA is replicated once per cell cycle during S phase in synchrony with the cellular DNA (Adams 1987). There are many different initiation sites of latent viral replication across the viral genome (Norio & Schildkraut 2001). A cellular recognition complex binds to the viral latent origin of replication with EBNA1 to initiate copying of the viral genome (Schepers et al. 2001). Pre-replication complexes are important in latent replication, assembling onto the viral genome (Ritzi et al. 2003). Only a few of these pre-recognition complexes are activated per viral genome, suggesting this is a highly organized and precise mechanism to keep maintain a viral episome in each dividing cell (Papior et al. 2012).

1.8.2. Latency Cycle

The reversibility of EBV to enter into a lytic cycle after a silent phase within cells is a key characteristic of true latency by herpesviruses. This reversibility, along with persistence allows EBV to be a very successful virus where both latency and lytic cycles are key to viral survival (Speck & Ganem 2010). Viral gene expression within the latency phase produces proteins and RNA for viral genome maintenance and host modification.

Upon primary infection of B lymphocytes, the EBV latency-associated proteins are expressed including EBNAs, their mRNAs initiate from a common and B-cell-specific promoter, Wp (Alfieri et al. 1991). This is shown in Figure 1.1. The origin

of plasmid replication (OriP) was found to be a region essential for stable replication of latent EBV (Yates et al. 1984). This region is the site of replication for the virus, once per cell cycle and cellular factors including viral protein EBNA1 are important (Kirchmaier & Sugden 1995). The origin of plasmid replication (OriP) on the EBV genome allows for episome maintenance through the action of EBNA1, for the latent replication of the viral DNA. The C-terminus of EBNA1 binds to the FR and DS tandem repeats within OriP (Rawlins et al. 1985) and regions of the N-terminus of EBNA1 facilitates anchoring of the episome to cellular chromatin (Marechal et al. 1999; Hung et al. 2001).

EBV expresses eight latency-associated genes, with different patterns of expression in different cells resulting in four latency programs. EBNA1, EBNA2, EBNA3A, EBNA3B, EBNA3C, EBNA-LP, LMP1, LMP2A/LMP2B are the latent proteins expressed. A low level of EBNA1, EBNA3A, EBNA3B, EBNA3C and LMP1 and LMP2 expression are initiated by EBNA2 and EBNA-LP. The Wp promoter is used before a switch to the Cp promoter and EBNA2 mediates this promoter switch (Jin & Speck 1992). There are 4 distinct patterns of EBV latent gene expression observed in infected lymphocytes, with a subset of the genes expressed in different cells and malignancies.

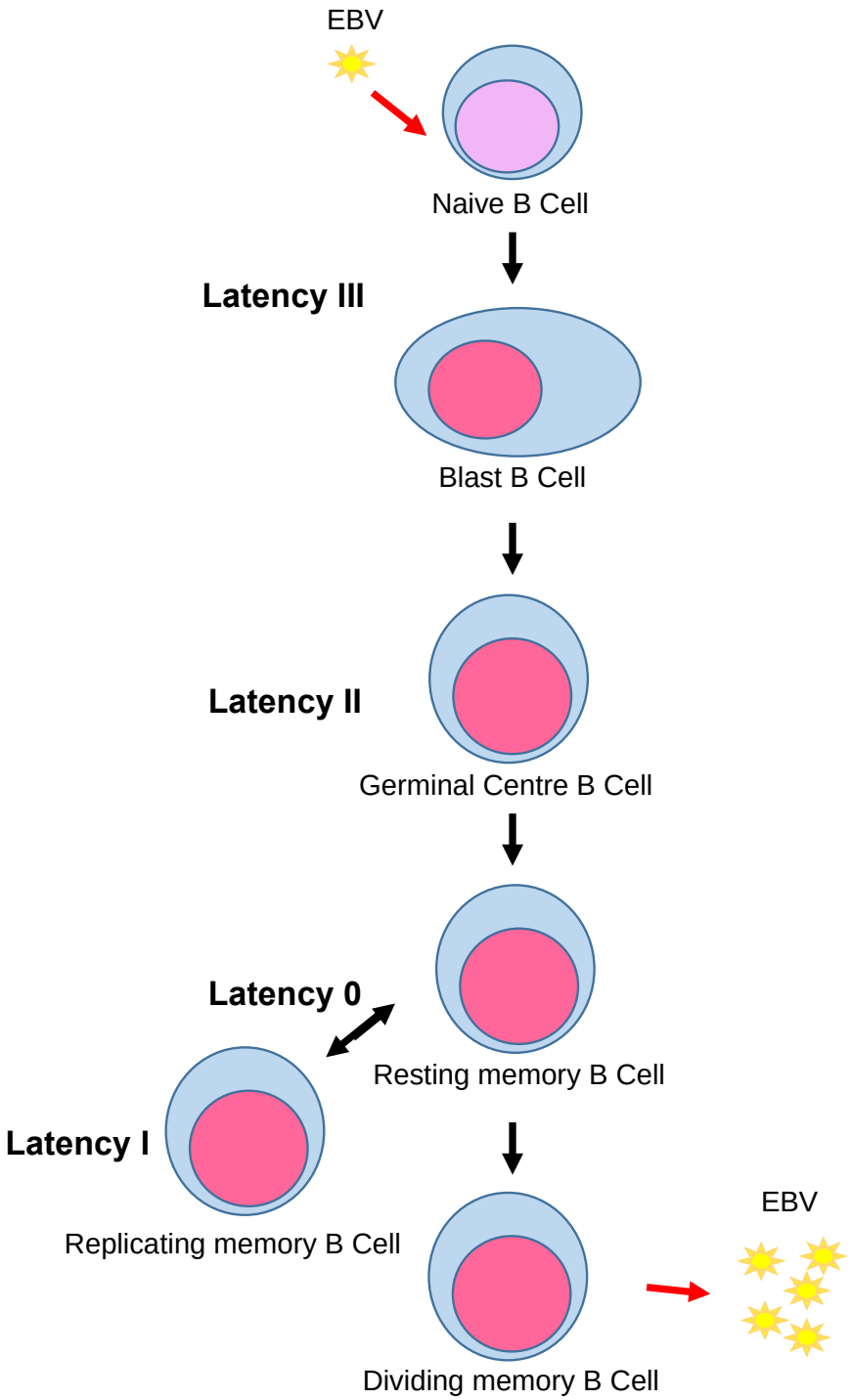


Figure 1.3 EBV life cycle in B Cells (Adapted from (Vockerodt et al. 2015) EBV is transmitted in saliva. Following infection of naive B cells in the lymph node all latent genes are expressed in Latency Type III and the viral latent proteins drive proliferation in the absence of antigen. B-cell differentiation into the memory compartments occurs in germinal centres driven by latency type II proteins. Infected memory B cells exiting the germinal centre down-regulate viral proteins and are invisible to the immune response as Latency Type 0. EBNA1 is expressed during homeostatic proliferation to maintain the latent viral episome in Latency I. An establishment of a life-long infection of memory B cells are detected in the peripheral circulation. Differentiation to plasma cells results in reactivation of the virus to the lytic cycle, expression of lytic proteins and production of infectious virus.

Memory B-cells can express a latency 0 profile (Shaknovich et al. 2006). EBERs and miRNAs are only expressed in latency 0. The other expression profiles include latency I, II and III, these are outlined in different states of infected cells in Figure 1.3 (Vockerodt et al, 2015).

The latency gene expression within epithelial cells is more restricted compared to B-cells and the cellular phenotype of epithelial cells. Also latent viral replication in epithelial cells requires expression of lytic cycle genes, suggesting that the infection and life cycle of EBV in epithelial cells is dependent on the cellular environment (Shannon-Lowe et al. 2009). Chromatin changes that include histone modifications are decreased and redistributed heterochromatin marks are associated with growth transformation (Hernando et al. 2013).

The latency patterns in Figure 1.4 are outlined below:

Latency III expresses all six EBNA (EBNA1, 2, 3A, 3B, 3C and LP) proteins, LMP1 and LMP2 and EBER1 and EBER2. Also BamA rightward transcripts (BARTs) with microRNAs are expressed. This form of latency is characteristic of lymphoblastoid cell lines (LCLs) that have been transformed from resting B-lymphocytes in vitro with EBV. These cells are routinely used in cell culture to study latency III (Thorley-Lawson 2001).

Latency II is characteristic of NPC tumours. Only EBNA1, two latent infection integral membrane proteins (LMP1 and LMP2a/LMP2b) and two EBV encoded smalls RNAs (EBER1 and EBER2) are expressed with microRNAs (Brooks et al. 1992).

In vivo, latency I is characteristic of BL cells. Only EBNA1 and EBERs are expressed and is similar to the latency pattern seen *in vivo* for memory cells (Rowe et al. 1987).

Latency 0 has rarely been identified where only EBERs and BARTs are expressed in memory B-lymphocytes and plasma cells (Shaknovich et al. 2006).

Several microRNAs (miRNAs) are expressed throughout the EBV latency cycles. These perform many cellular functions, including RNA silencing guided by these small RNAs derived from double-stranded RNA, acting as viral regulators of host and viral gene expression (Zhu et al. 2009; Lin et al. 2013; Pfeffer et al. 2004).

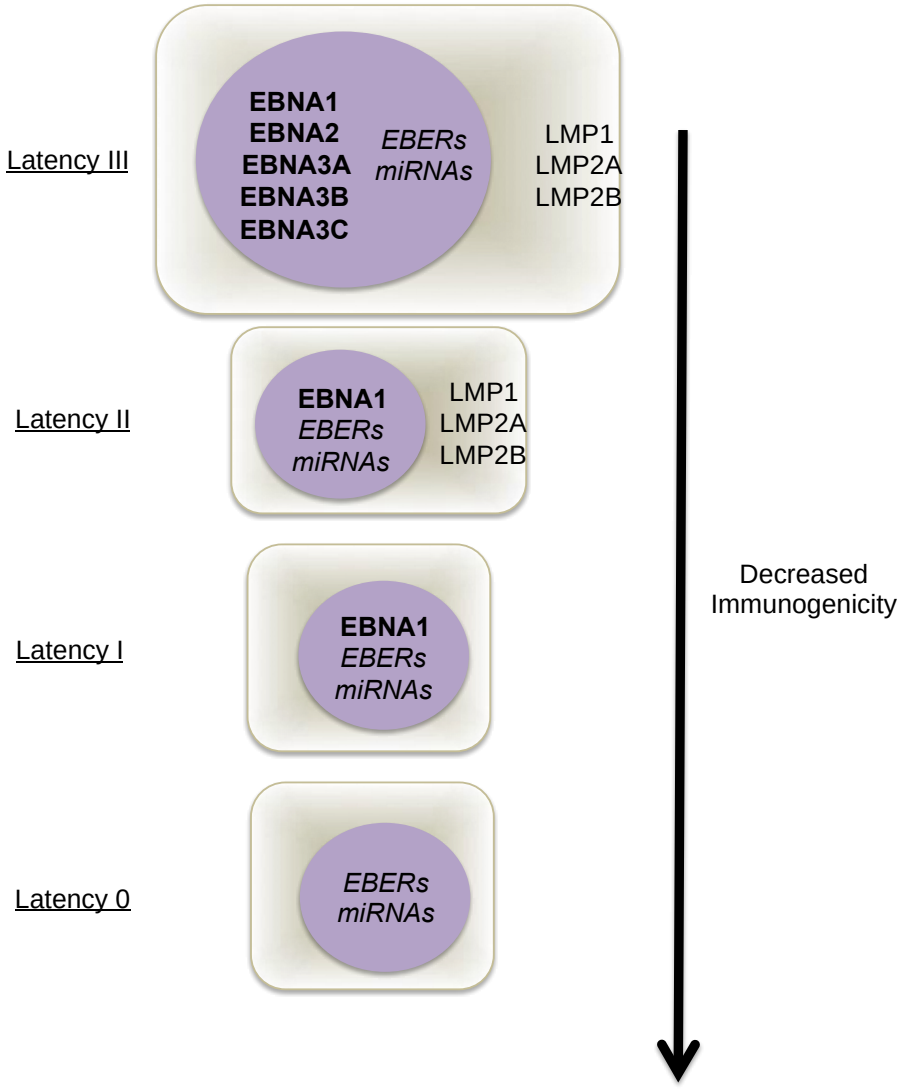


Figure 1.4 The latency pattern of viral genes expression in EBV-associated tumours. (Fox et al. 2011). Latency 0 can be observed in non-dividing circulating infected memory B cells of healthy carriers; it is possible that the majority of these cells express no viral genes at all, but that a minority may express non-coding RNAs. Latency I expresses EBNA1 with EBERs and microRNAs, Latency II expresses EBNA1, EBERs, microRNAs and LMP1, 2A and 2B. Latency III expresses all the EBNA proteins, EBERs, microRNAs and LMP1, 2A and 2B.

1.9. Lytic Cycle

1.9.1. Reactivation

EBV remains latent within the cells that it infects until a reactivation event occurs. The lytic switch for gamma herpesviruses has been significantly studied to research the causality leading to the activation of the lytic cycle. The lytic cycle is the reactivation from latency that leads to the production of infectious virions. This is essential for the virus to ensure the virus survival and spread to other hosts. There is interplay between cell signalling during the latency stage that suppresses lytic cycle. The Notch pathway is highly conserved cell signal system which contains four Notch receptors, NOTCH1, NOTCH2, NOTCH3 and NOTCH4 (Hori et al. 2013). EBNA2-dependent gene expression can be modulated by Notch and also NOTCH2 inhibited entry into lytic cycle by upregulating Zeb2, to repress BZLF1 transcription (Rowe, et al 2014). Therefore, the establishment of latency in memory B cells is being modulated through varuous cellular mechanisms to inhibit lytic cycle and to maintain the viral episome in latency. Despite this, the switch between latency and lytic cycle can be induced a number of ways. Current models suggest that B cell differentiation from resting cells to plasma cells stimulated through the antigen receptor stimulates the lytic cycle (Laichalk & Thorley-Lawson 2005). The activation of EBV to produce lytic proteins and enter the lytic cycle by exposure to anti-IgM was documented early (Tovey et al. 1978) and the crosslinking of surface immunoglobulin lead to EBV replication in Akata cells (Takada 1984). This receptor activation allows signalling pathways to be activated that ultimately lead to the induction of lytic cycle.

Within lytic cycle there are epigenetic mechanisms that also control the mechanism of the EBV life cycle. The lytic cycle genes within the EBV genome are silenced in latency. Control of reactivation includes viral and cellular factors that may activate or repress many elements in the pathway to lytic cycle. This switch can be manipulated in cell culture systems using many different stimuli. This repression involves epigenetic mechanisms such as DNA methylation and chromatinisation (Fernandez et al. 2009). The Zp promoter is highly methylated during latency (Li et al. 2012). Transcriptional repression is initiated by

methylation of DNA. This repressive mark attracts proteins with methyl binding sites, such as histone deacetylases and methyl transferases. These impose a repressive chromatin structure upon the DNA (Jones & Baylin 2007). Methylation of DNA in humans occurs primarily on carbon 5 of the pyrimidine ring of cytosine residues in the context of a cytosine guanine (CpG) dinucleotide site (Schones & Zhao 2008). Numerous cellular processes are regulated by DNA methylation, used to transcriptionally repress genes (Robertson 2005). This is evident for the regulation of lytic viral gene expression

1.9.2. Zta

Zta is a 245 amino acid residue protein and is a bZIP protein with four defined regions: an N-terminal transactivation domain, a basic DNA contact region and a dimerization region (zipper) and a c-terminal region. The basic and zipper regions are characterised as a basic leucine zipper motif (Sinclair, 2006), which shares homology to AP-1 transcription factors c-fos and c-jun. AP1 transcription factors contain basic region leucine zipper (bZIP) and this region can dimerise with related bZIP factors. In contrast, Zta only forms homodimers. Zta can bind to DNA with consensus AP-1 DNA elements in addition to other sites (Farrell et al. 1989; Flower et al. 2011).

Zta (BZLF1, EB1, Z, and ZEBRA) is the first protein expressed in the lytic cycle, from the *BZLF1* gene (Countryman & Miller 1985). Zta plays a role as both a transcription factor and as a replication factor and is expressed within 30 minutes after lytic activation (Sinclair 2003). Zta binds to and interacts with AP-1 binding sites (Urier et al. 1989). Zta can bind to Zta response elements (ZREs), which are 7bp DNA motif that closely resemble AP-1 binding sites. There are three different classes of ZREs: Class I, Class II and Class III. Class I ZREs do not contain a CpG motif, whereas Class II and III do. Zta may bind to these methylated CpG motifs, bypassing the inhibitory effect of DNA methylation (Bhende et al. 2004).

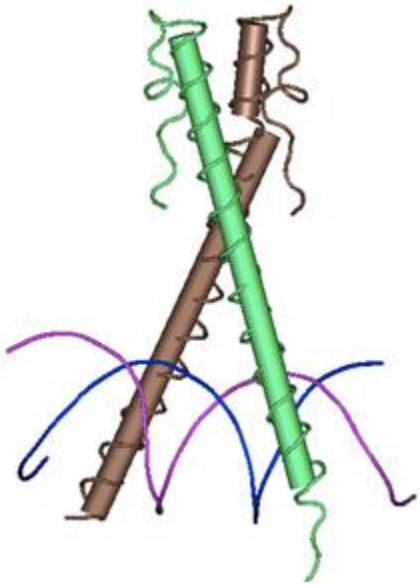


Figure 1.5 Zta dimerization and DNA binding domain structure (Petosa et al 2006) using the Cn3D macromolecular structure viewer NCBI. Two strands of Zta are present (green and brown). Zta can form a dimer that is bound to the DNA double helix (blue and purple).

Zta can interact with key ZREs within methylated DNA sequences, as well as human promoters containing these methylated ZREs (Flower et al. 2010). The strategy Zta uses to enhance lytic cycle is to bind viral promoters that are methylated (Dickerson et al. 2009). Zta can interact with many parts of the viral genome and this has been mapped in detail, noting many novel sites of interaction (Ramasubramanyan et al. 2012; Woellmer & Hammerschmidt 2013).

The amino terminal of Zta encompasses amino acids 1 to 167 and includes the transactivation domain. The transactivation domain was characterized and mapped, with residues 28 to 78 critical for the function of the domain (Flemington et al. 1992). . Transcription preinitiation complexes interact with this region of Zta (Chi & Carey 1993). TATA box binding protein associated factors (TAF) interact with the Zta transactivation domain to form stable complexes on promoter DNA. This leads to the activation of transcription (Lieberman & Berk 1994).

The bZIP domain encompasses the basic DNA binding region and the leucine zipper region. The DNA binding domain of Zta is located in the 175aa - 196aa region. This basic region of amino acids are important for the transcriptional function of Zta and activating gene expression of lytic genes, such as BRLF1 (Heston et al. 2006). Mutations of this domain also alters the sub-nuclear

localization of Zta to replication compartments (Park et al. 2008). Also residues Y180A, Y180E, K188A mutations are deficient in activating EBV DNA replication (Heston et al. 2006). Serine 186 (S186) is important for the DNA binding (Baumann et al. 1998). It mediates the recognition of the Rp promoter (Francis et al. 1999), and mediates methylated activation of the Rp promoter (Bhende et al. 2005). A redox sensitive cysteine at amino acid 189 (C189) is essential for the lytic replication function of the protein (Wang et al. 2005; Karlsson et al. 2008). Transcription of early gene promoters is not affected but transcription of late genes are with this mutation. This residue is conserved among the bZIP family of proteins.

The bZIP domain bound to a ZRE sequence of DNA has been crystalized and shows unique residues that contribute to the dimer interface (Petosa et al. 2006). The dimer formed at the two asparagine 211 (Asn211) residues share a hydrogen bond conserved in most bZIP structures. But unique to Zta, there is an extended hydrogen bond network that links these residues to the C-terminal tail at the top of the coiled coil, two Cys222 residues position their thiol groups 3.9Å apart, which is too far to form a disulphide bond. This may be functionally significant. Evidence suggests that Cys222 partly contributes to the redox sensitivity of ZEBRA's DNA binding activity (Wang et al. 2005).

The leucine zipper and dimerization domain is located between residues 197aa-221aa. A coiled-coil structure is structurally related to leucine zippers while the Zta dimerization domain shares amino acids in common with C/EBPα leucine zipper (Kouzarides et al. 1991). Zta has a unique dimer region that lacks the heptad leucine zipper motifs. Instead this region has non-leucine coiled-coil domain, the leucine coiled-coil is characteristic of bZIP factors (Chang et al. 1990; Flemington & Speck 1990b). Zta can form a homodimer with and without DNA binding. The amino acids 196 to 227 can direct the folding of the coiled-coil domain and this interaction is not as strong as other bZIP proteins (Hicks et al. 2001). The contribution of the c-terminal region of Zta has been shown to have high importance for the function of the coiled-coil domain (Hicks et al. 2003).
The c-terminal region is unique to EBV and is required for EBV genome replication. An unexpected discovery was that a region of the c-terminus of Zta

intercalates with the ZIP region. The zipper of Zta is located adjacent to the c-terminus of the protein. A fold back structure of the zipper domain is supported by the c-terminal region where the proximal region essential (Schelcher et al. 2007). These residues fold back against the zipper region and maintain the protein structure. Therefore the c-terminus Is essential for viral DNA replication (McDonald et al. 2009). The final 9 amino acids from the crystal structure have not been crystallised so the structure of the c-terminus is yet to be elucidated (Petosa et al, 2006). The formation of replication compartments rely on Zta and crucial amino acids for nuclear localization and induction of lytic proteins (Park et al. 2008).

In vitro activation of EBV lytic cycle can be initiated by chemical agents for example by sodium-butyrate activating as a histone deacetylase inhibitor (Luka et al. 1979). The resulting histone acetylation changes at the promoter of Zta can lead to induction of gene expression and lytic cycle (Jenkins et al. 2000). Histone acetylation allows a permissive or repressive state of chromatin. The acetylation of histones weakens the affinity of histones with DNA, making the DNA more accessible for transcription factors. This inhibits the deacetylation of histones, leading to a hyper acetylation state where the agent acts.

12-0-tetradecanoyl phorbol-13-acetate (TPA) differentiates B-lymphocytes into plasma cells and can induce EBV replication (zur Hausen et al. 1978; Anisimova et al. 1984). Calcium ions and protein kinase C can mediate lytic expression of the viral genome (Castagna et al. 1982; Faggioni et al. 1986; Gradoville et al. 2002). TPA phorbol ester binds the C1 domain of protein kinase C (PKC), mimicking antigen receptor activation. 5-aza-cytidine is another lytic inducing agent. This is incorporated into DNA but can not be methylated and so prevents cytidine methylation in DNA, through an unknown mechanism this reactivates viral gene expression (Ben-Sasson & Klein 1981)

Although these agents initiate lytic cycle, the mechanism of reactivation is complex and there are variations between cell lines. For example, phorbol esters do not activate a HH514-16 BL cell line, whereas activation is evident in B95-8 cells (Gradoville et al. 2002). Raji cells require both sodium butyrate and TPA to

initiate the expression of early lytic genes, although Raji cells cannot produce virions (Nutter et al. 1987). The EBV genome in Raji cells contain a deletion for the BALF2 gene, which is an essential lytic replication protein (Zhang et al. 1988). This suggests additional mechanisms are involved between cell lines for activation.

Reactivation involves a temporally regulated cascade of gene expression. The signal transduction pathways involved can be complex. This leads to the lytic cascade of viral genes in lytic cycle. The lytic cycle can be separated into three stages of gene expression; these are termed immediate-early, early and late. Cellular transcription factors are also involved in stimulating replication (Baumann et al. 1999). Sp1 and Sp3 are cellular transcription factors that bind to the Zta promoter, contributing to activity (Liu et al. 1997).

1.9.3. Immediate Early genes

Immediate early (IE) genes are the first lytic genes expressed after the reactivation of EBV from latency. These include two transcription factors with the ability to activate lytic genes (Chevallier-Greco et al. 1986). Immediate early genes initiate lytic cycle and expression of early and late genes. Central to the activation of the lytic cycle are BZLF1 as mentioned previously, and BRLF1. The importance of the BZLF1 gene in induction of lytic cycle was demonstrated by the transfection of B95-8 BamHI BZLF1 fragments into EBV positive Raji cells (Takada et al. 1986). The BZLF1 fragment induced the expression of transcription factor Zta. The expression of Zta leads to the productive infection of active virions, capable of the immortalization of B-lymphocytes (Grogan et al. 1987).

BZLF1 expresses Zta and BRLF1 expresses Rta. Both of these proteins are transcription factors, and both transactivate other lytic genes and coordinate the cascade together (Countryman & Miller 1985; Hardwick et al. 1988). Zta can also auto-regulate its own promoter (Flemington & Speck 1990a). Both are essential for the induction of lytic cycle, both a BZLF1 knockout virus and a BRLF1 knockout virus are incapable of reactivating EBV from latency (Feederle et al. 2000).

When the lytic cycle is initiated, both BZLF1 and BRLF1 are expressed immediately, Zta first activating the Rta promoter, Rp (Sinclair et al. 1991). Zta can then co-activate genes with Rta (Ragoczy & Miller 1999). Rta can also interact with CREB-binding protein CBP to induce the activation of early lytic genes (Swenson et al. 2001). Rta with Zta plays a key role in lytic replication, Rta interacting with the origin of lytic replication (OriLyt) on the viral genome and this interaction may stabilize the replication proteins that locate there (El-Guindy et al. 2013).

Another immediate-early protein that facilitates the stability of lytic gene mRNA is indispensible for EBV lytic cycle is BMLF1, a nuclear RNA binding protein. BMLF1 has been demonstrated to be essential for lytic replication (Lieberman et al. 1986).Most lytic genes are intronless and BMLF1 enhances expression of these viral intronless genes by binding their mRNA and facilitating its transport (Ruvolo et al. 1998). With the stability of gene expression, lytic cycle can progress and the production of infectious virus (Gruffat et al. 2002).

1.9.4. Early genes

Viral early (E) genes after IE lytic genes are next expressed in the cascade of lytic cycle gene expression, in order to faciltate the production of newly replicated viral DNA and to manipulate the cellular architecture for viral synthesis. These accessory proteins aid virion synthesis and modify the cellular environment to enable the execution of EBV lytic replication. Some key genes are outlined below:

BMRF1 is essential for lytic virus replication, recruited to viral replication compartments as a DNA polymerase processivity factor (Neuhierl & Delecluse 2006). This oligomeric ring shaped structure acts as a sliding clamp (Makhov et al. 2004) and forms a tetrameric ring from head to head dimers (Murayama et al. 2009). BMRF1 can also act a transcription factor, activating the BHLF1 promoter (Zhang et al. 1996; Zhang et al. 1997). The BHLF1 promoter is activated by Zta or BMRF1 or both together, this region located in the EBV oriLyt region is crucial for lytic replication (Zhang et al. 1997). Another lytic protein expressed is BHRF1, which shows homology to cellular Bcl-2. This protein inhibits apoptosis, which is the cellular response to the onset lytic cycle (Henderson et al. 1993).

EBV encodes a viral kinase protein expressed named BGLF4. This protein is a serine/threonine protein kinase and is the only protein kinase expressed from the EBV genome (Chee et al. 1989). The BGLF4 protein kinase phosphorylates viral genes including Zta. BGLF4 phosphorylates BMRF1 (Chen et al. 2000) as well as EBNA-LP (Kato et al. 2003) and may play other roles in the phosphorylation of cellular proteins. A stable complex between BGLF4 and Zta is formed with the amino acid serine 209 of Zta being essential for this interaction (Asai et al. 2009) colocalising with BGLF4 in replication compartments, implying a direct regulation of EBV lytic replication (Asai et al. 2009).

BNLF2a is an early lytic gene that blocks CD8+ lymphocytes through HLA Class II interference, blocking TAP processing (Hislop et al. 2007). This early protein creates a window of opportunity for the virus to remain undetected from T cell responses to lytic cycle (Horst et al. 2009).

An early gene immune response modulator BARF1, interacts with the colony stimulation factor 1 (CSF1). The lytic protein acts as a decoy receptor for the cytokine CSF1 to limit the differentiation of macrophages (Strockbine et al. 1998; Hoebe et al. 2012).

BRRF1 early gene encodes a transcription factor Na, and this protein enhances the induction of the lytic cycle with Rta. Here, Na activates the Zp promoter through c-jun phosphorylation (Hong et al. 2004).

BGLF5 is an alkaline exonuclease that contributes to immune evasion. BGLF5 promotes mRNA degradation and shut off of HLA class I (Rowe et al. 2007). The recognition by T cells therefore becomes impaired by the action of BGLF5 (Zuo et al. 2008). This protein is also essential for viral replication, as viral DNA replication in cells lacking the BGLF5 gene generated abnormal linear genomes (Feederle et al. 2009).

1.10. EBV Lytic Replication

One essential role of Zta is as a replication factor. Zta coordinates the transition from latency to lytic cycle involving the recruitment of essential viral factors for viral DNA replication to the origin of lytic replication (OriLyt). Replication units are also referred to as the replisome. Cellular replisome protein machinery consists of DNA polymerase, single strand DNA binding protein complex, DNA primase and ligase, DNA topoisomerases, clamp loading complex and a processivity factor (Chagin et al. 2010). These proteins conduct DNA synthesis in an organised manner in the 5' to 3' direction, creating a replication fork with a leading and lagging strand (Perumal et al. 2010).

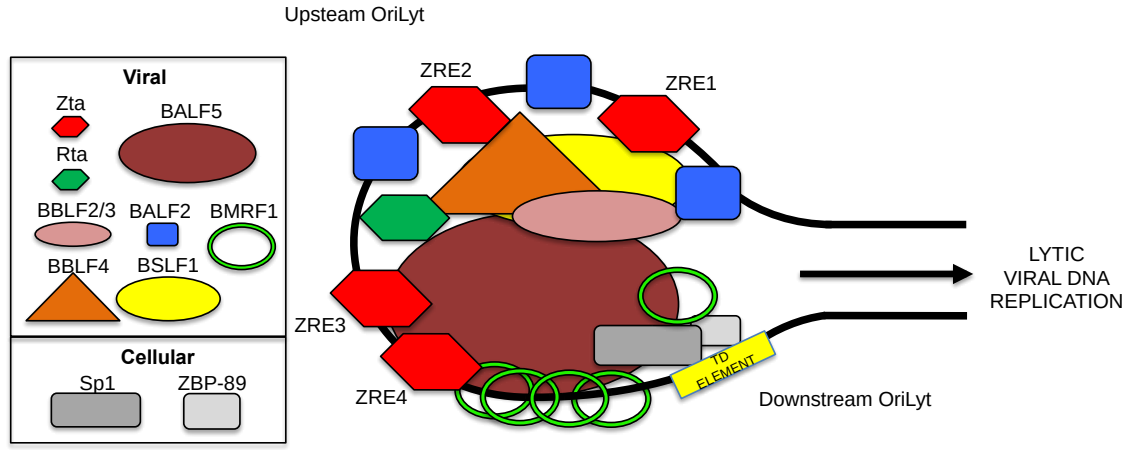


Figure 1.6 A model of the EBV lytic replisome adapted from both (Baumann et al. 1999) and (El-Guindy et al. 2013). Cellular Sp1 and ZBP89 presumably bind to the TD element during the latent as well as lytic phase of EBV's life cycle. Zta activates oriLyt directly by binding to ZREs, located in the upstream component (Lieberman & Berk 1990; Schepers et al. 1993; Schepers et al. 1996). Protein–protein interactions exist between the Zta and the DNA polymerase accessory factor (BMRF1) (Zhang et al. 1996), and between the transactivation domain of Zta and both the viral helicase (BBLF4) and the primase complex BSLF1 and BBLF2/3 (Gao et al. 1998). BALF2 may contact the BSLF1 and BBLF2/3 subcomplex (Gao et al. 1998). BMRF1 interacts with BALF5 (Murayama et al. 2009). Rta interacts with oriLyt. Rta is required for synthesis of the BHLF1 transcript, which facilitates strand separation and recruitment of replication proteins to oriLyt during replication (El-Guindy et al. 2013). BALF5 and BMRF1 interact (Kiehl & Dorsky 1995) with the cellular transcription factors Sp1 and ZB89.

The identification of the origin of lytic replication (OriLyt) (Hammerschmidt & Sugden 1988) led research to be focused on this point of newly synthesized viral DNA. Zta interacts with and contacts ZREs at OriLyt in order to initiate lytic replication (Schepers et al. 1996) (Figure 1.6). The majority of EBV genomes contain two copies of OriLyt, OriLyt L and R. Some EBV genomes contain just

one copy, but no genome has been found that lacks both (Xue & Griffin 2007). The KSHV genome contains two copies of OriLyt, but there is not an absolute requirement of two origins of lytic replication for viral genomes (AuCoin et al. 2002).

Initiation of lytic replication requires a stable RNA-DNA hybrid to form at OriLyt, composed of a BHLF1 RNA transcript (Rennekamp & Lieberman 2011). This guanine-cytosine rich RNA template is essential for the DNA strand separation and loading of the core replication proteins. The replication machinery then initiates and synthesizes leading and lagging viral DNA strands. The viral genes and proteins essential for EBV replication include BALF5 DNA polymerase, BMRF1 DNA polymerase processivity factor, BALF2 single-stranded DNA binding-protein, BSLF1 Primase, BBLF4 Helicase, BBLF2/3 associated component of the helicase-primase complex (Fixman et al. 1992).

The BALF5 DNA polymerase is a 113kDa protein that extends viral DNA and can extend RNA primers on template DNA (Tsurumi 1991). Through the interaction with BMRF1 processivity factor, the polymerase increases in replication activity (Kiehl & Dorsky 1995). BMRF1 interacts with BALF5 as a monomer (Murayama et al. 2009). The BALF2 single stranded (ss) DNA binding protein (SSB) is about 130kDa in size, and binds ss-DNA covering about 30 nucleotides to one protein monomer, thus destabilising double stranded DNA (Tsurumi et al. 1998). The helicase-primase interaction (BBLF4-BSLF1) stabilises the formation of the replisome, localising within the nucleus with BBLF2/3, the primase-associated protein (Gao et al. 1998) (Figure 1.5)

The majority of this replication machinery is well conserved with Herpesviridae. Most replication proteins can be interchanged with other herpes virus replication proteins, but the correct origin-binding protein is essential for each complex to initiate replication (Rennekamp & Lieberman 2010).

Nuclear architecture is important for viral DNA replication. It is reprogrammed during the transition into lytic phase. Several hours before lytic viral DNA replication, host chromatin recesses to the nuclear boundaries, not associated

with EBV amplification factories (Chiu et al. 2013). The synthesis of viral DNA takes place in these amplification compartments formed within the nucleus, Zta and BMRF1 colocalise here (Takagi et al. 1991). The viral replication proteins localize to the replication compartments and BMRF1 readily distributes on newly formed DNA as well as acting as an accessory factor for the viral DNA polymerase (Daikoku et al. 2005). The architecture changes to allow the viral replication compartments to be surrounded externally by cellular DNA and histones. A redistribution of cellular proteins is also involved in DNA replication. The EBV replication compartments associate with promyelocytic leukaemia protein PML bodies and disrupt nuclear domain 10 (ND10). (Bell et al. 2000; Amon et al. 2006). Zta may compete with PML bodies for SUMOylation, as Zta itself can be SUMO-1 modified (Adamson & Kenney 2001). This redistribution of cellular factors and modification of the immediate nuclear environment allows EBV to take control its own genome replication. The viral replication factories contain PCNA (Chiu et al. 2013). PCNA is loaded onto newly synthesized viral DNA with mismatch repair factors assembling in these compartments (Daikoku et al. 2006). Also recombination repair factors such as ATM and MRN complexes localize to the replication compartments (Kudoh et al. 2009).These factors may aid replication and repair damage to the viral DNA. Two essential cellular factors involved with EBV lytic replication are Sp1 and ZBP-89. These are transcription factors that interact with the viral DNA polymerase BALF5 and processivity factor BMRF1, Sp1 and ZBP-89 tether these viral proteins to OriLyt aiding assembly of the viral replisome (Baumann et al. 1999).

BMRF1 plays a transcriptional role as well as a replicative role, further activating the OriLyt BHLF1 promoter when together with Zta. An interaction between BMRF1 DNA polymerase processivity factor and Zta takes place between the first 45 amino acids of BMRF1 and the bZIP domain of Zta (Zhang et al. 1996). BBLF4, BBLF2/3 and BSLF1 form a tripartite complex that localizes to the nucleus (Gao et al. 1998). BBLF4 and BBLF2/3 plus BSLF1 bind to the transactivation domain of Zta at 1aa to 133aa, stabilize the replication complex. The binding of BBLF4 to Zta has been mapped to 22aa to 86aa of Zta transactivation domain (Liao et al. 2001).

Some cellular proteins contribute and aid lytic replication. The protein Sp1 forms a weak complex with BALF5 in vivo, and Zta can interact with BALF5 through its transactivation domain between 26aa and 88aa (Baumann et al. 1999). Also ZBP89 Zta interacts with BMRF1, with Sp1 and ZBP89 interacting the downstream component of OriLyt, the TD element 88aa (Baumann et al. 1999) Zta interacts with nearly all viral replication components of the replisome except for BALF2, the single-stranded DNA binding protein (El-Guindy et al. 2010) (Figure 1.6). Overall, EBV DNA is newly synthesised from OriLyt into concatamer DNA that is then processed and packaged into virions.

1.11. Zta-Host Interactions

There are many stages where viruses can interact with the host cell through viral protein interactions. Host proteins are manipulated in order to facilitate survival of the virus. Herpesviruses have been extensively investigated for viral-host interactions. Ranging from viral entry to the cell, DNA damage, apoptosis, immune evasion and viral replication.

Zta has been shown to interact with a variety of different viral and cellular proteins for a number of different functions in the cell. These range from transcription of host and viral genes, manipulating the cellular environment to facilitate the lytic cycle and viral lytic replication. The switch between latency and lytic cycle can be regulated by the control of cellular proteins and pathways.

p53 is a key tumour suppressor protein and is central to the formation of some cancers. Other viruses have manipulated and disrupted p53. Zta interacts with p53 through its bZIP domain and the C-terminus of p53 (Zhang et al. 1994). Knockdown of p53 in LCLs leads to interruption of EBV reactivation (Chua et al. 2012), suggesting that a direct interaction between Zta and p53 may play a role in regulating reactivation into lytic cycle.

CREB-binding protein (CBP) is a histone acetylase protein. Histone acetylation weakens the interaction between DNA and nucleosomes, changing the chromatin conformation (Norton et al. 1989). Zta interacts with CBP through the transactivation domain and zipper domain of Zta and the amino-terminus of CBP.

Zta may target cellular histone acetylases to early viral promoters to assist in transcription (Adamson & Kenney 1999).

Cellular CCAAT/enhancer binding protein α (C/EBPα) is a bZIP transcription factor that shows similarity to myc. The leucine zipper shares homology with Fos and Jun (Landschulz et al. 1988). C/EBPα may also play a role in cell cycle arrest. This protein interacts with Zta through the bZIP domain, inducing p21 expression and also mediating cell cycle arrest (Wu et al. 2003).
NFkB is a central protein involved with controlling transcription and cellular signalling pathways. Zta can interact with the p65 subunit of NFkB through the Zta dimerization domain, inhibiting Zta transactivation function (Gutsch et al. 1994).

Zta and Rta can form a complex through an intermediary protein MCAF1 on ZREs. The proteins work in synergy for transcription using MCAF1 at ZREs to promote gene expression (Chang et al. 2010).

Zta interacts with a member of the host DNA damage pathway. 53BP1 is connected to the ataxia telangiectasia mutated (ATM) signal transduction pathway. The c-terminal region of Zta is responsible for the interaction with 53BP1 (Bailey et al. 2009). Upon the induction of lytic cycle, ATM pathway is activated and Zta may reduce this effect by interacting with 53BP1, facilitating lytic replication

Ku80 is part of the DNA-dependent protein kinase family and involved in the DNA repair pathway. The n-terminal 425aa of Ku80 interacts with 168aa and 245aa of Zta, enhancing Zta-activated transcription (Chen et al. 2011).

Zta can associate with the mitochondrial single-stranded DNA binding protein (mtSSB). This protein is involved in mitochondrial DNA replication, and this interaction with Zta may play a role in viral replication (Wiedmer et al. 2008). The Zta interaction with mtSSB facilitates viral replication and blocks mitochondrial DNA replication.

The basic transcription factor TFIID interacts with 25aa to 86aa of Zta and may form a stable promoter complex, to stimulate transcription (Lieberman & Berk 1991). The myb protein, a member of myeloblastosis transcription factors, can activate the BMRF1 promoter in synergy with Zta (Kenney et al. 1992). Ubinuclein, another nuclear transcription factor can interact with Zta through the bZIP domain (Aho et al. 2000). The functional significance of these interactions is currently unclear.

Zta also plays a role in negatively regulating genes. The retinoic acid receptor (RAR) and retinoid X receptor (RxR) play a role in hormone receptor and transcription. Zta interacts here through its dimerization domain, together with RAR to negatively regulate the promoter for BMRF1 (Sista et al. 1995).

As outlined previously, cellular factors are also recruited to replication compartments. Topoisomerases acts by unwinding supercoiled DNA. RecQL1 maintains chromosome stability as a DNA helicase enzyme. RecQL1 associates with Zta at OriLyt, while topoisomerase I is required for this recruitment. Topoisomerase I also associates with Zta at OriLyt in EBV-positive cells, promoting replication assembly at this viral origin of lytic replication (Wang et al. 2009).

1.12. Aims of project

Cellular factors are recruited to the origin of lytic replication for EBV and for the closely related gamma herpesvirus, KSHV. These include Topoisomerase I and II, MSH2/6, RecQL, poly(ADP-ribose) polymerase I (PARP-1), DNA-PK, Ku86/70 autoantigens, and scaffold attachment factor A (SAF-A) (Wang et al. 2008). There is evidence for a wide range of host factors to be recruited that may contribute to herpesvirus lytic replication. A proteomics direction will be taken to investigate questions of EBV lytic replication and cellular proteins.

The aims are:

- Develop a method to identify cellular proteins that interact with Zta and investigate these protein-protein interactions. Transfection of histidine tagged Zta into EBV negative cells (U2OS) will allow an establishment of

protein extraction conditions. Through the poly-histidine tag, binding conditions to immobilized metal ion affinity chromatography (IMAC) will be optimised. EBV negative SILAC labelled extracts transfected with Zta and then applied to this pull down assay will be coupled to mass spectrometry analysis. Identified proteins of interest will be attempted to be assessed for interaction with Zta

- The developed method for protein-protein interactions would be taken further by applying cross-linking and denaturing conditions to identify novel interactions with Zta. EBV positive SILAC labelled extracts transfected with Zta and applied to this optimised pull down assay will be coupled to mass spectrometry analysis. Proteins of interest will be assessed for interaction with Zta.

- A difference in protein complexes created during EBV lytic replication will be dissected through size exclusion chromatography. Size exclusion chromatography of EBV negative cells transfected with Zta will be optimised, before EBV positive cells transfected with Zta are applied to the chromatography column. Differences in the elution profile between control cell extract, full lytic cycle extract and extract unable to undergo lytic cycle will be assessed. Specific elutions from the column will be investigated by SILAC labelled mass spectrometry analysis. Protein complexes will be dissected and evaluated.

- Whole cell proteomics of BL cells undergoing lytic cycle will be assessed for protein expression in control and Zta expressing cells. Using an enriched SILAC labelled population of cells undergoing latent and lytic cycle coupled with mass spectrometry, a global proteomics view will determine any differences between latent and lytic protein expression of host and viral proteins.

2. Materials and Methods

2.1. Materials

2.1.1. Plasmids

Vector	Method	Source
pBABE	Transfection of cells	(Morgenstern & Land 1990)
pCDNA3	Transfection of cells	Invitrogen
pBABE - Zta	Transfection of cells	(Hicks et al, 2001)
pCDNA3 - hisZta	Transfection of cells	(Bailey et al 2009)
pCDNA3 - hisZtaAAA	Transfection of cells	Cloned by Q. Karlsson
pcDNA3-flag-BMRF1	Transfection of cells	Cloned by Q. Karlsson
pcDNA3-flag-NFκB p65	Transfection of cells	Cloned by Q. Karlsson
pcDNA3-flag-p53	Transfection of cells	Addgene
pcDNA3-flag-BGLF4	Transfection of cells	(Tarakanova et al. 2007)
pCpGL-BHLF1	Luciferase assay	(Bergbauer et al. 2010)
pCpGL-BHLF1 mutant	Luciferase assay	(Ramasubramanyan et al. 2015)

Table 2.1 List of plasmids used in the experiments

2.1.2. Antibodies

Antibody (order)	Species, clonality	Source
BZ1 mouse monoclocal (Primary)	Mouse, monoclonal	(Young et al, 1991)
ScZ α-Zta(Primary)	Goat, polyclonal	Santa Cruz Biotechnology
Monoclonal α-FLAG (Primary)	Rabbit, polyclonal	Sigma
Actin (Primary)	Rabbit, polyclonal	Sigma
FANCA (Primary)	Rabbit, polyclonal	AbCam
BRD4 (Primary)	Rabbit, monoclonal	AbCam
ELP3 (Primary)	Rabbit, polyclonal	AbCam
ELK4 (Primary)	Rabbit, polyclonal	AbCam
HSP90 (Primary)	Mouse, monoclonal	AbCam
BMRF1 (Primary)	Mouse, monoclonal	AbCam
BALF5 (Primary)	Rat, monoclonal	(Barth et al. 2008)
Poly(A) RNA Pol (Mitochondrial) (Primary)	Rabbit, polyclonal	AbCam
Casein Kinase II alpha (Primary)	Mouse, monoclonal	AbCam

53BP1 (Primary)	Rabbit, polyclonal	Sigma
ECL α-rabbit HRP-linked (Secondary)	Goat	GE Healthcare
ECL α-mouse HRP-linked (Secondary)	Horse	GE Healthcare
IR Dye α-rabbit 680CW (Secondary)	Goat	LICOR
IR Dye α-mouse 800CW (Secondary)	Goat	LICOR
IR Dye α-mouse 680CW (Secondary)	Goat	LICOR
IR Dye α-rabbit 800CW (Secondary)	Donkey	LICOR

Table 2.2 List of antibodies used for western blotting

2.1.3. Purchased reagents and materials

Reagent	Use	Supplier
DMEM	Cell Culture	GIBCO
RPMI	Cell Culture	GIBCO
DMEM R0K0 (SILAC)	Cell Culture	Dundee Cell
DMEM R6K4 (SILAC)	Cell Culture	Dundee Cell
RPMI R0K0 (SILAC)	Cell Culture	Dundee Cell
RPMI R6K4 (SILAC)	Cell Culture	Dundee Cell
RPMI R10K8 (SILAC)	Cell Culture	Dundee Cell
Fetal Calf Serum	Cell Culture	GIBCO
Fetal Calf Serum (SILAC)	Cell Culture	Dundee Cell
PBS	Cell Culture	GIBCO
Antimycotic Antibiotic (AA)	Cell Culture	GIBCO
Penicillin/Strep/Glutamine	Cell Culture	GIBCO
Trypsin	Cell Culture	GIBCO
Hygromycin B	Cell Culture	Invitrogen
Dimethyl Sulfoxide (DMSO)	Cell Culture	Sigma
6-well plate (3.5cm diameter)	Cell Culture	NUNC
$25cm^3$ small flask	Cell Culture	NUNC
$75cm^3$ medium flask	Cell Culture	NUNC
$175cm^3$ large flask	Cell Culture	NUNC
Anti IgG	Induce EBV lytic cycle in cell culture	DAKO
Doxycycline	Induce doxycycline promoter	Sigma
Formaldehyde	Cross-linking cells	Sigma
Sample Buffer, Laemmli 2X concentration	Protein Gel	Sigma
3-8% Precast Tris-acetate gel	Protein Gel	Invitrogen
10% Precast BIS-tris gel	Protein Gel	Invitrogen

12% Precast BIS-tris gel	Protein Gel	Invitrogen
SimplyBlue Safestain	Protein Gel	Invitrogen
Flag control peptide	Protein Gel	Sigma
BSA	Protein Gel	Sigma
HIS-Select Nickel Affinity Gel	Pull Down assay	Sigma
Imidazole	Pull Down assay	Sigma
CellLytic MT Cell lysis reagent	Protein Extraction	Sigma
Benzonase Nuclease	Protein Extraction	Sigma
Protease inhibitor EDTA-free	Protein Extraction	Roche
Sodium orthovanadate (Na_3VO_4) Phosphatase Inhibitor	Protein Extraction	Sigma
Sodium Fluoride (NaF) Phosphatase inhibitor	Protein Extraction	Sigma
β-Glycerolphosphate Phosphatase inhibitor	Protein Extraction	Sigma
Trypsin (Mass Spectrometry grade)	Mass spectrometry	Promega
Superose 6 10/300GL Column	SEC	GE Healthcare
Acetone	Acetone precipitation	Fisher
Odyssey Blocking buffer	Western Blot	Licor
PBS (100 tablets)	Western Blot	Oxoid
Tween 20	Western Blot	Fisher
Nitrocellulose Membrane	Western Blot	Licor
MOPS SDS running buffer (20x)	Western Blot	Invitrogen
Tris-acetate SDS running buffer (20x)	Western Blot	Invitrogen
Luminol	Western Blot	Sigma
P-Coumaric acid	Western Blot	Sigma
Hydrogen Peroxide (30%)	Western Blot	Sigma

Table 2.3 List of reagents and materials purchased from various suppliers.

2.1.4. Solutions

Solutions	Use	Composition
Solution A	Nuclear Extract	10mM Hepes pH 7.9 10mM KCl 1.5mM MgCl2 0.1mM EGTA (chelating agent) 0.5mM DTT (reducing agent) 0.5mM PMSF (phenyl methane sulfoyl fluoride) (serine protease inhibitor)
Solution C	Nuclear Extract	10mM Hepes pH 7.9 400mM NaCl 1.5mM MgCl2 0.1mM EGTA (chelating agent) 0.5mM PMSF (phenyl methane sulfoyl fluoride) (serine protease inhibitor) 5% Glycerol
HIS-Select Equilibration Buffer	Pull down	50 mM sodium phosphate, pH 8.0, with 0.3 M sodium chloride and 10 mM imidazole
HIS-Select Wash Buffer	Pull down	50 mM sodium phosphate, pH 8.0, with 0.3 M sodium chloride and 10 mM imidazole
HIS-Select Equilibration Buffer with 8M urea	Pull down	0.1M Sodium Phosphate at pH 8.0 with 8M urea
HIS-Select Wash Buffer with 8M urea	Pull down	0.1M Sodium Phosphate at pH 6.2 with 8M urea
HIS-Select Elution Buffer with 8M urea	Pull down	0.1M Sodium Phosphate at pH 4.5 with 8M urea
WB Blocking Buffer	Western Blot	5% dried milk (w/v), 10mM Tri-Base pH 7.4, 150mM NaCl, 1mM EDTA, 0.1% Tween 20 (v/v).
Transfer Buffer	Western Blot	0.025M Tris, 0.192M Glycine, 15% methanol (v/v)

PBS Tween	Western Blot	10mM Tris-Base, 150mM NaCl, 1mM EDTA, 0.1% Tween 20 (v/v), pH 7.4
Ponceau-S concentrate (10X)	Western Blot	2% (w/v) Ponceau-S, 30% (v/v) Trichloroacetic acid, 30% (v/v) Sulphosaliylic acid (Sigma)
Laemmli Protein Sample buffer	Western Blot	2% SDS, 10% glycerol, 5% 2-mercaptoethanol, 0.002% bromphenol blue, 0.0625 Tris HCl (Sigma)
ECL solution I	Western Blot	10mM Tris HCl pH8.5, 0.4mM p-Coumaric acid, 2.5mM Luminol
ECL solution II	Western Blot	10mM Tris HCl pH8.5, 0.02% Hydrogen Peroxide
Size Exclusion buffer	SEC	20mM Tris 100mM NaCl pH 8.0 buffer

Table 2.4 List of Solutions made in the laboratory

2.1.5. Kits used and their suppliers.

Kit	Manufacturer	Use
QIAquick PCR Purification Kit	Qiagen	DNA purification
QIAprep Maxi-Prep	Qiagen	Maxi-Prep of expression vectors
Effectene Transfection Kit	Qiagen	Transfection
Luciferase Assay System	Promega	Luciferase Assay

Table 2.5 Kits used and their suppliers

2.1.6. QPCR primers

Name	Sequence (5' → 3')	Reference
BALF5	(a) agtccttcttggctagtctgttga (b) ctttggcgcggatcctc	(Gallagher et al. 1999)
β-Globin	(a) ggcaaccctaaggtgaaggc (b) ggtgagccaggccatcacta	(Gallagher et al. 1999)

Table 2.6 QPCR primers

2.1.7. Cell Lines

Cell Lines	Reference
Akata containing pRTS-CD2-BZLF1	(Ramasubranyan et al, 2015)
Akata	(Takada, 1984)
HEK293-*BZLF1*-KO	(Feederle et al, 2000)
U2OS	ATCC Catalog No. HTB-96

Table 2.7 Cell lines used

2.2. Methods

2.2.1. Transformation/Maxi Prep

100ng of plasmid DNA was added to Top10 chemically competent *E.Coli* cells. The cells were incubated on ice for 5 minutes before adding preheated LB media (42°C), heat shocked at 42°C for 30 seconds and incubated on ice for 2 minutes. Cells were plated onto pre-warm LB agar plates containing 100µg/ml of ampicillin and incubated at 37°C overnight. Single colonies were chosen and added to 5ml of LB media containing ampicillin and incubated at 37°C shaking at 225rpm in a ThermoScientific MaxQ4000 incubated shaker for 8 hours. The media was then transferred to flasks containing 150ml LB media with 100µg/ml of ampicillin and incubated at 37°C shaking overnight. Qiagen maxi preps were performed as recommended by the manufacturers protocol. The plasmid stocks were then quantified for dsDNA concentration with an Eppendorf Biophotometer.

2.2.2. Luciferase Assay

U2OS cells were grown in 6-well plates and transfected with BHLF1 wild type promoter sequence or BHLF1 mutated promoter sequence, with control or hisZta expression vectors. Cells were harvested after 48 hours and washed in 1ml PBS before being split into 2 x 500µl. Half of the harvested cells were used for Western Blot protein expression and the other half of cells used for luciferase assay. 125µl of passive cell lysis buffer (Promega) was added to the cell pellet and incubated for 15 minutes. The cells were centrifuged and the supernatant was pipetted into clean tubes. For the luciferase activity analysis, 10µl of each lysate sample was pipetted into a white 96-well luminescence plate. 50µl of the luciferase kit reagents (Substrate and Buffer) were added to the samples and the

bioluminescence was measured by a Glomax multidetection machine. The samples were repeated in triplicate.

2.2.3. Cell Culture

Cells were grown in an incubator at 37°C with 5% CO_2. EBV positive Burkitt 's lymphoma (BL) cell line (B Cells) are maintained in RPMI (Invitrogen).

HEK293-*BZLF1*-KO cells (Feederle et al, 2000) are adherent cells maintained in RPMI (Invitrogen) supplemented with hygromycin (100µg/ml). These cells contain the B-958 EBV strain but the *BZLF1* gene has been removed. The cells containing the recombinant virus are maintained with the addition of hygromycin. SILAC labelled HEK293-*BZLF1*-KO cells (Feederle et al, 2000) maintained in light (R0K0), medium (R6K4) or heavy (R10K8) RPMI (Dundee Cell) supplemented with hygromycin (100µg/ml).

U2OS cells maintained in DMEM (Invitrogen). U2OS are an adherent osteocarcinoma cell line. SILAC labelled U2OS cells maintained in light (R0K0), medium (R6K4) or heavy (R10K8) DMEM (Dundee Cell).

R0K0 media containing metabolically unlabelled arginine and lysine amino acids (R0K0). R6K4 media containing 13C labelled arginine and 2D labelled lysine amino acids (R6K4). R10K8 media containing 13C and 15N labelled arginine and 13C and 15N labelled lysine (R10K8)

All media for cell growth is supplemented with 100U/ml of penicillin, 100µg/ml of streptomycin and 2mM of L-glutamine solution and AA. 10% (v/v) final volume Foetal Calf Serum (FCS) (Invitrogen) was also added or 10% (v/v) final volume Foetal Calf Serum (FCS) (Dundee Cell) for SILAC labelling of cells.

For long term storage of cell lines, cells were washed and then resuspended in 90% FCS supplemented with 10% DMSO at $1x10^7$ cells/ml. Cells were frozen at -80°C overnight before long term storage in liquid nitrogen. Frozen cell stocks were recovered by rapid thawing, followed by a subsequent wash in 10ml of chilled media before culturing.

2.2.4. Determination of cell count

10µl of cells were taken from the cell populations and placed on a haemocytometer with a coverslip. Four quadrants of cells were counted and the number averaged before being multiplied by 1 x 10^4. This was the total number of cells per millilitre.

2.2.5. Small Scale Transfection

All transfections were performed in a class II microbiology cabinet. Cells were washed and resuspended in PBS before being counted and plated.

U2OS cells were plated at 2.5 x 10^5 cells in 6-well plates in 3mls of media. Transfection was done using a Qiagen Effectene Transfection Kit. 1µg of DNA was added to 100µl of EC buffer with 3.2µl enhancer. After vortexing the DNA mixture was incubated at room temperature for 5 minutes. 2.5µl of effectene was added and vortexed before being incubated for 10 minutes at room temperature. The media in the wells were replaced with 2ml DMEM. 1ml of DMEM media was added to the DNA mixture before being added to the well. Cells were incubated for 72 hours before being washed in PBS and then harvested.

2.2.6. Large Scale Transfection

U2OS cells were plated at 3.75 x 10^6 cells in large flasks in 30mls of media. Transfection was done using a Qiagen Effectene Transfection Kit. The transfection ratio was scaled up to 17.5x. 17.5µg of DNA was added to 1.75ml of EC buffer with 56µl enhancer. After vortexing the DNA mixture was incubated at room temperature for 5 minutes. 44µl of effectene was added and vortexed before being incubated for 10 minutes at room temperature. The media in the wells were replaced with 27ml DMEM. 3ml of DMEM media was added to the DNA mixture before being added to the well. Cells were incubated for 72 hours before being washed in PBS and harvested.

HEK293-*BZLF1*-KO cells were plated into large flasks in 30mls of media and grown to 60% confluency. Transfection was done using a Qiagen Effectene Transfection Kit. 17.5µg of DNA was added to 1.75ml of EC buffer with 56µl enhancer. After vortexing the DNA mixture was incubated at room temperature for 5 minutes. 44µl of effectene was added and vortexed before being incubated for 10 minutes at room temperature. The media in the wells were replaced with

27ml DMEM. 3ml of DMEM media was added to the DNA mixture before being added to the well. Cells were incubated for 72 hours before being washed in PBS and harvested.

2.2.7. Akata cell induction

EBV positive Akata cell lines have been established using the expression plasmid pRTS-CD2-BZLF1 which as well as carrying endogenous EBV BZLF1 inducible by IgG (0.125%) also has a bidirectional DOX inducible promoter B-Tet, which drives expression of BZLF1 together with a non-functional neuronal growth factor receptor (NGFR) and green fluorescent protein (GFP) as markers of DOX-induced expression. The same cell line with the same construct but carrying the reverse BZLF1 sequence was used to control for any dox induced effects. Doxycycline was added to cells to give a 500ng/ml final concentration for a 24-hour induction.

For Akata cells induction, IgG (0.125% of final volume) was added to cells for 48-hour induction.

2.2.8. DNA purification

DNA was extracted from cells using Promega Wizard Genomic DNA Extraction kit. This was performed following the manufacturers protocol for the kit (Promega). DNA was stored at -20°c.

2.2.9. QPCR

Quantitative polymerase chain reaction method was used to detect EBV genome replication. SYBR green (Promega) intercalates with double strand DNA and the SYBR will fluoresce, detected by the qPCR machine (StepOnePlus, Applied Biosystems) controlled by StepOne software version 2.3 (Applied Biosystems). Viral lytic gene BALF5 is used to detect the amplified viral genome and β-globin used as an internal control. 1 x 10^6 HEK293-*BZLF1*-KO cells. A standard curve was constructed by using serial dilutions of a 100ng sample with one primer set.

A master mix of reagents was used and distributed onto the 96-well sample plate for each sample. A master mix prepared with 12.5µl SYBR, 8.5µl dH2O, 1µl

forward primer 10μM, and 1μl reverse primer 10μM. 2μl of each 10ng sample were plated in triplicate and the master mix added.

The QPCR reaction conditions included:

Temperature	Time	
50°c	2 Minutes	
95°c	10 Minutes	
95°c	15 Seconds	⎫
60°c	1 minute	⎬ 40 cycles
95°c	15 Seconds	⎭
60°c	1 Minute	
95°c	15 Seconds	

The qPCR results were analysed by StepOne software 2.3 (Applied Biosystems). The melting curve was analysed for each reaction to ensure results were acceptable, with no contamination, mispriming or artefacts affecting results

2.2.10. Cell Extraction

CellLytic MT Cell lysis reagent (Sigma) was used for cell lysis and extraction of proteins. 1mM of phosphatase inhibitors: Sodium Fluoride, Sodium Orthovanadate, β-glycerophosphate were added to the reagent. Protease inhibitor 1X (Roche) and the determined benzonase at 125/Units were also supplemented to the reagent. Cells were harvested and washed in PBS before being resuspended in this cell lysis reagent for 25 minutes at room temperature. Centrifugation for 10 mins at 13000rpm in a Fisher accuSpin micro R centrifuge allowed the supernatant kept as the protein extract.

Nuclear extract protocol was adapted from Lee et al 1998. Cells were harvested and washed in PBS. Before spun at 2000rpm in a Fisher accuSpin micro R centrifuge for 5 minutes. The cell pellet was resuspended in one packed cell volume of Solution A and left on ice for 15 minutes. NP40 was added to final concentration of 0.6% and cells were then vortexed briefly before centrifuging. The supernatant was discarded and the remaining pellet resuspended in one

nuclear volume of Solution C rotating at 4°C for 30 minutes. The cell suspension was centrifuged and the supernatant was kept as the nuclear fraction.
The cell lysis reagent was prepared as described but made up to the required volume with 8M urea.

2.2.11. Native Affinity Tag Pull Down

HIS-Select Nickel Affinity Gel (Sigma) used for small-scale purification of histidine-tagged recombinant Zta. The affinity gel was aliquoted into 50µl 50% suspension and spun at 8000rpm in Fisher accuSpin micro R centrifuge for 5 mins. The ethanol was removed and the affinity gel washed with dH2O twice. The gel was equilibrated with equilibration buffer for 5 minutes before centrifuging and removing the buffer. Cell extract that was prepared with cell lysis buffer was applied to the gel and incubated for 60 minutes rotating at 4°C. The supernatant was removed and kept; the gel was washed with Wash Buffer twice for 15 minutes each rotating at 4°C, with or without imidazole. The gel was either eluted with elution buffer containing imidazole or finally boiled in an equal volume protein sample buffer before being loaded onto SDS-PAGE.

2.2.12. Denatured Affinity Tag Pull Down

A denatured pull down was performed as described but with minor adjustments to the buffers. Equilibration buffer containing 8M urea, wash buffer containing 8M urea were used. An elution buffer containing 8M urea was also used before the elutions were collected.

2.2.13. FPLC

An AKTA purifier system was used for size exclusion chromatography. The AKTA purifier (GE Healthcare) was controlled using Unicorn software (v.5.20). A Superose 6 10/300GL (GE Healthcare) is a prepacked Tricorn column with a protein fractionation range between 5,000da and 5,000,000da. All water and buffers used were filtered using a 0.22µm filter before being applied to the Akta system and column. This column was connected to the AKTA purifier with appropriate tubing and washed with two column volumes of distilled water after a pump wash of the system with pump A and B. The column was then equilibrated with two column volumes of 20mM Tris buffer with 100mM NaCl at pH 8.0. 500µl of protein extract was centrifuged at 13000rpm in a Fisher accuSpin micro R

centrifuge for 5 minutes before being loaded into a 500µl loading loop connected to the AKTA purifier. The protein was fractionated at 0.2ml/min for two column volumes, either collecting into 2ml fractions or 500µl fractions. The fractions were then kept and labeled according to the collection plate used. Samples were mixed with a minimal 4X acetone before being placed in -20°C until further analysis.

The UV detector measures protein by milli absorbance unit (mAU). This is one-thousandth of an absorbance unit, where an absorbance unit is defined as the absorbance at 280nm. This determines protein concentration of the sample being detected.

2.2.14. Acetone precipitation

Proteins in solution were precipitated with 4X volume prechilled (-20°C) acetone in acetone compatible tubing before being placed at -20°C overnight. The tubes were spun at 15000G for 15 minutes in a Sorvall RC-5B Centrifuge in an SS-34 rotor at 4°C before carefully removing the buffer. The proteins that were precipitated into a white pellet were resuspended in 2x protein sample buffer and then heated at -95°C for 5 minutes before being analysed by western blotting.

2.2.15. Protein electrophoresis SDS-PAGE

Cells, protein extracts or agarose gel were mixed with an equal volume of 2x protein sample buffer (Sigma). Samples were heated at 95°C for 5 minutes before being spun at 13000rpm in a Fisher accuSpin micro R centrifuge for 5 minutes. Samples were loaded onto precast 10% or 12% Bis-Tris NuPage (Invitrogen), or 3-8% Tris-Acetate gel (Invitrogen). SeeBlue marker (Invitrogen) was loaded as a molecular weight marker. The gels were run at 200 volts for 50 minutes in MOPS SDS running buffer for 10 or 12% gels, 150 volts for 60 minutes in Tris-acetate SDS running buffer for 3-8% gels (Invitrogen).

2.2.16. Western Blotting

After samples are separated by SDS-PAGE, proteins transferred to nitrocellulose membrane in a Bio-Rad Transfer tank at 75 volts for 90 minutes containing transfer buffer. The addition of Ponceau Red stain to the membrane checked that proteins were transferred correctly. The stained transferred membrane was washed with water to display stained protein bands. The membrane was washed

three times in PBS-tween for 5 minutes each before blocking the membrane. The membranes were blocked for 45 minutes using either 5% milk in PBS-tween or Odyssey Blocking Buffer (Licor). The membrane was incubated with primary antibody overnight at 4°C. The membrane was washed three times with PBS-Tween before the addition of secondary antibody. Two methods of secondary antibody were performed. A secondary horseradish peroxidase linked antibody was added selected against the animal from which the primary antibody was generated. Or the addition of a secondary fluorescent linked antibody (Licor) was added selected against the animal from which the primary antibody was generated. The secondary antibodies were incubated with the membrane moving at room temperature for one hour. The membranes were washed three times with PBS-tween after incubation. The HRP antibody complexes were detected using enzymatic chemiluminescence. ECL buffer A was mixed with ECL buffer B for one minute in the dark. Excess buffer was removed before the membrane sealed in saran wrap and placed in a cassette with photographic film. This exposed the antibody specific protein bands. If the secondary fluorescent antibodies (700nm or 800nm) (Licor) were used then the membrane was exposed directly to the Odyssey Fc imager. After detection at either 700nm or 800nm wavelength by the Odyssey Fc system, the antibody specific bands were detected.

2.2.17. Mass Spectrometry

Mass spectrometry analysis was either performed myself at the University of Sussex or sent to the University of Bristol Proteomics Facility for analysis.

The protocol that was used at the University of Sussex Mass Spectrometry:
Samples in protein sample buffer (Sigma) were mixed together equally and separated by SDS-PAGE before stained with SimplyBlue safestain (Invitrogen) for one hour. The gel was washed with distilled water for one hour to remove residual stain. The gel was placed on a light box and bands were cut using a clean scalpel. The gel pieces were cut into finer pieces and the residual buffer removed. The gel was washed three times with 200µl of 50% acetonitrile in 25mM NH_4HCO_3 on a rotator. The gel pieces were subjected to 5 minutes in a SpeedVac concentrator with no heat. The gel pieces were reduced by addition of 100µl 10mM DTT at 20°C for 45 minutes. Alkylation performed by the addition of + 200µl

50mM iodoacetamide (IAA) in 25mM NH_4HCO_3. This was added to the tubes which were placed in a lightproof box for 45 minutes. The gel was washed three times with 200µl of 50% acetonitrile in 25mM NH_4HCO_3 on a rotator with 5 minutes speed vac on no heat at the end. 12.5ng/µl trypsin was added to each gel piece and left on ice for 10 minutes. Excess trypsin was removed before the addition of 25mM NH_4HCO_3 at 37°C overnight. Formic acid was added to a final concentration of 5% and the peptides were moved to clean tubes. The gel pieces had 1X volume of acetonitrile vortexed and rotated at room temperature for 15 minutes. The supernatant added to the tubes containing peptides. The peptides were applied to speed vac until the supernatant was removed. The peptides were resuspended in 8µl of 0.1% trifluoroacetic acid (TFA) and placed into -20°C until further analysis.

An LTQ-OrbitrapXL mass spectrometer was used to obtain the mass spectrometry spectra and created an .RAW output file of the peptides identified. Mascot Daemon was first used to investigate the dataset. Within the Task Editor, the parameters and dataset were set for the search. The parameter editor to format parameters included the UniProt/SwissProt database, Carbamidomethyl and Oxidation as modifications. Trypsin was the enzyme at a peptide tolerance set to 7Da and a peptide charge 2+ and 3+. SILAC quantitation was also inputted here with K+4, K+8, R+6 and R+10. The mascot result report for a peptide mass fingerprint search enabled a list of proteins, each of which the matches some of the experimental peptide masses. The result was kept as a screenshot.

MaxQuant (v1.5.0.25) was also used to obtain protein lists from the .RAW output files from the LTQ-OrbitrapXL mass spectrometer, using a FASTA file containing the human proteome alongside the EBV proteome.

The parameters used in Maxquant included:

Parameter	**Value**
Version	1.5.0.25
User name	ct245
Fixed modifications	Carbamidomethyl (C)
Decoy mode	revert
Special AAs	KR

Include contaminants	True
MS/MS tol. (FTMS)	20 ppm
Top MS/MS peaks per 100 Da. (FTMS)	10
MS/MS deisotoping (FTMS)	True
MS/MS tol. (ITMS)	0.8 Da
Top MS/MS peaks per 100 Da. (ITMS)	6
MS/MS deisotoping (ITMS)	False
MS/MS tol. (TOF)	0.1 Da
Top MS/MS peaks per 100 Da. (TOF)	10
MS/MS deisotoping (TOF)	True
MS/MS tol. (Unknown)	0.5 Da
Top MS/MS peaks per 100 Da. (Unknown)	6
MS/MS deisotoping (Unknown)	False
PSM FDR	0.01
Protein FDR	0.01
Site FDR	0.01
Use Normalized Ratios For Occupancy	True
Min. peptide Length	7
Min. score for unmodified peptides	0
Min. score for modified peptides	40
Min. delta score for unmodified peptides	0
Min. delta score for modified peptides	6
Min. unique peptides	0
Min. razor peptides	1
Min. peptides	1
Use only unmodified peptides and	True
Modifications included in protein quantification	Acetyl (Protein N-term);Oxidation (M)
Peptides used for protein quantification	Razor
Discard unmodified counterpart peptides	True
Min. ratio count	2
Re-quantify	True
Use delta score	False
iBAQ	True
iBAQ log fit	True
Match between runs	False
Find dependent peptides	False
Site tables	True
Decoy mode	Oxidation (M)Sites
Special AAs	revert
Include contaminants	KR
RT shift	True
Advanced ratios	False
AIF correlation	True
First pass AIF correlation	0.47
AIF topx	0.8
AIF min mass	20
AIF SIL weight	0

AIF ISO weight	4
AIF iterative	2
AIF threshold FDR	True 0.01

Table 2.8 Maxquant parameters used for mass spectrometry anaylysis.

The protocol that was used at the University of Bristol Proteomics Facility:
The gel lane was cut into 3 slices and each slice subjected to in-gel tryptic digestion using a ProGest automated digestion unit (Digilab UK). The resulting peptides were fractionated using a Dionex Ultimate 3000 nanoHPLC system in line with an LTQ-Orbitrap Velos mass spectrometer (Thermo Scientific). In brief, peptides in 1% (vol/vol) formic acid were injected onto an Acclaim PepMap C18 nano-trap column (Dionex). After washing with 0.5% (vol/vol) acetonitrile 0.1% (vol/vol) formic acid peptides were resolved on a 250 mm × 75 μm Acclaim PepMap C18 reverse phase analytical column (Dionex) over a 150 min organic gradient, using 7 gradient segments (1-6% solvent B over 1min., 6-15% B over 58min., 15-32%B over 58min., 32-40%B over 3min., 40-90%B over 1min., held at 90%B for 6min and then reduced to 1%B over 1min.) with a flow rate of 300 nl min^{-1}. Solvent A was 0.1% formic acid and Solvent B was aqueous 80% acetonitrile in 0.1% formic acid. Peptides were ionized by nano-electrospray ionization at 2.1 kV using a stainless steel emitter with an internal diameter of 30 μm (Thermo Scientific) and a capillary temperature of 250°C. Tandem mass spectra were acquired using an LTQ- Orbitrap Velos mass spectrometer controlled by Xcalibur 2.1 software (Thermo Scientific) and operated in data-dependent acquisition mode. The Orbitrap was set to analyze the survey scans at 60,000 resolution (at m/z 400) in the mass range m/z 300 to 2000 and the top six multiply charged ions in each duty cycle selected for MS/MS in the LTQ linear ion trap. Charge state filtering, where unassigned precursor ions were not selected for fragmentation, and dynamic exclusion (repeat count, 1; repeat duration, 30s; exclusion list size, 500) were used. Fragmentation conditions in the LTQ were as follows: normalized collision energy, 40%; activation q, 0.25; activation time 10ms; and minimum ion selection intensity, 500 counts. The raw data files were processed and quantified using Proteome Discoverer software v1.2 (Thermo Scientific) and searched against the UniProt/SwissProt Human database release version 57.3 (20326 entries) plus the HisZta sequence using the SEQUEST (Ver. 28 Rev. 13) algorithm. Peptide precursor mass tolerance

was set at 10ppm, and MS/MS tolerance was set at 0.8Da. Search criteria included carbamidomethylation of cysteine (+57.0214) as a fixed modification and oxidation of methionine (+15.9949) and appropriate SILAC labels ($^{2}H_{4}$-Lys, $^{13}C_{6}$-Arg) as variable modifications. Searches were performed with full tryptic digestion and a maximum of 1 missed cleavage was allowed. The reverse database search option was enabled and all peptide data was filtered to satisfy false discovery rate (FDR) of 5%. The Proteome Discoverer software generates a reverse "decoy" database from the same protein database and any peptides passing the initial filtering parameters that were derived from this decoy database are defined as false positive identifications. The minimum cross-correlation factor (Xcorr) filter was readjusted for each individual charge state separately to optimally meet the predetermined target FDR of 5% based on the number of random false positive matches from the reverse decoy database. Thus each data set has its own passing parameters. Quantitation was done using a mass precision of 2ppm. After extracting each ion chromatogram the Proteome Discoverer software runs several filters to check for, among other things,

2.2.18. Antibody search using BLAST

For each antibody, the epitope amino acid sequence that was recognised by the antibody was retrieved from the manufacturers datasheet. This sequence was entered into the http://web.expasy.org/blast/ tool and a BLAST search performed against the Homo sapiens UniProtKB taxonomic subset. This returned a graphical overview of the alignments and a list of the matches with an expectation value. The proteins that were aligned with a 100% match were investigated for their molecular weight and amino acid sequence length and collated into a table. Therefore the proteins identified from this BLASTP search contain the amino acid sequence the selected antibody is reported to recognize.

3. Identify novel interacting partners of Zta

3.1. Introduction

A method to identify cellular proteins that interact with Zta and investigate these protein-protein interactions was developed. The extraction of Zta from cells is essential to enable the identification and study of any associated proteins. Protein extracts that maximise the release of transfected Zta into the supernatant were optimised. Existing proteins that have been demonstrated to interact with Zta were also assessed for optimal extraction from cells, in an attempt to extract Zta in the context of possibly binding to these interaction partners.

A polyhistidine tagged version of Zta was used so that nickel affinity chromatography could attempt to purify tagged Zta with any associating proteins. Immobilized metal ion affinity chromatography (IMAC) is based on the specific covalent bonds of amino acids to metal ions. Poly-histidine (6x His-tag residues) has an affinity for nickel ions (Zhao et al. 2010).

The identification of proteins by mass spectrometry is a regular proteomics method. Functional proteomics assays have integrated quantitative routes to answer questions in more detail. Quantitative proteomics has traditionally been performed by two-dimensional gel electrophoresis. Advancements of quantitative proteomics included the combination of stable isotope labelling and mass spectrometry. This has been achieved through stable isotope labelling of amino acids in cell culture (SILAC). SILAC was first described and implemented and has since been applied to many studies (Ong et al. 2002). The in vivo incorporation of labelled amino acids in cell culture can be identified and quantitated using mass spectrometry. Using software designed to identify peptides with a higher molecular mass these SILAC peptides can be differentiated from unlabelled peptides. This technique has furthered the field of proteomics and allowed the identification and quantitation of proteins from complex biological mixtures. SILAC can be used as a quantitative proteomic approach in any cell culture system and applied to affinity purification (Blagoev et al. 2003). Protein populations from control and experimental samples can be mixed directly after harvesting or samples can be mixed prior to mass spectrometry analysis.

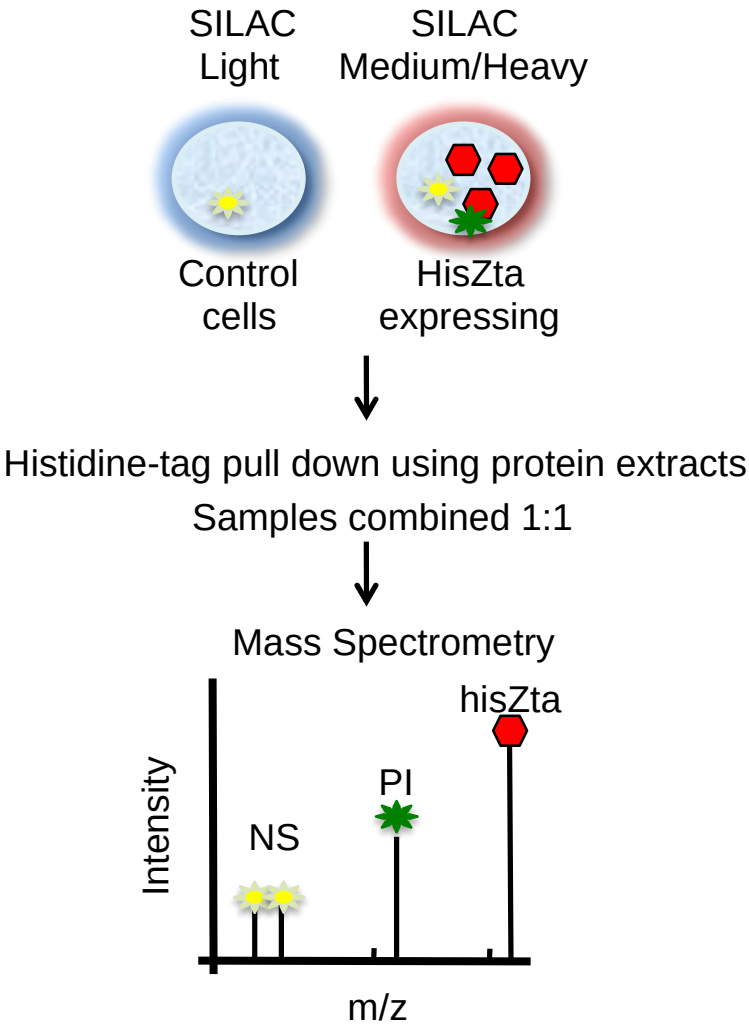


Figure 3.1 Diagram of SILAC-immunoprecipitation principle Control cells are labelled with unlabeled arginine and lysine containing normal isotopes of carbon and hydrogen. Cells are labelled medium or heavy with arginine and lysine containing heavier isotopes for carbon and hydrogen amino acids. These cells can be transfected with a control or the histidine tagged expression vector of interest. An immobilized metal ion affinity chromatography (IMAC) pull down can pull the histidine tagged protein from the extracts. This can be combined with a control pull down sample and the difference in peptide mass for detected peptides can result in a quantitated list of proteins identified. NS is Non Specific proteins, equally binding nonspecifically to both pull down affinity gels. PI is Protein interaction, more proteins that interact or associate with Zta have a higher abundance in the hisZta pull down.

Cell populations are SILAC labelled with either normal (light) isotope labelled amino acids or heavy isotope amino acids. The arginine amino acid in the cell growth media contains ^{13}C compared to ^{12}C and the cells incorporate the amino acids into new proteins. A control sample contains normal naturally occurring isotopes, versus the treated sample with heavier isotopes. Leucine, arginine and lysine can be altered to contain heavier isotopes of carbon. Arginine is not an essential amino acid but is essential in many cell lines (Scott et al. 2000). Commercially available amino acids contain ^{13}C, ^{15}N or ^{2}H, or medium/heavy isotopes in cell culture media.

Using an EBV negative cell line (U2OS) cells, mock transfection of the normal cell population are labelled with ‘light’ media and transfection with histidine tagged Zta of the heavy cell population are labelled with ‘medium’ media was performed.

SILAC labelled cell extracts were established and a quantifiable pull down method coupled to mass spectrometry was executed in order to identify any possible protein interactions with Zta. Using control ‘light labelled’ cell extracts and Zta expressing ‘medium labelled’ cell extracts a nickel pull down experiment can bind Zta via the histidine tag to nickel affinity gel (Figure 3.1). Identification of protein interaction partners in mammalian cells using SILAC-immunoprecipitation quantitative proteomics is possible using this defined technique (Emmott & Goodfellow 2014).

Combining samples of both gels and separating the proteins using SDS-PAGE, the samples can be prepared for mass spectrometry analysis. To identify which proteins are present in both the light and heavy samples, mass spectrometry was performed on the proteins bound to the nickel affinity gel.

Software analysis on the raw data obtained may reveal what proteins are present in the samples, as well as protein abundance due to a ‘mass shift’ of identical peptides. A mass shift of ‘heavier’ peptides due to the metabolic uptake of amino acids with heavier isotopes into proteins, sees a mass shift by mass spectrometry analysis as the retention time in the mass spectrometer differs for these peptides.

Peptides identified with an identical amino acid sequence will have a longer retention time due to being ‘heavier’ in the mass spectrometer and can be compared to normal peptides. Any proteins that have been enriched in the Zta containing sample would indicate that the corresponding protein may be interacting with Zta, therefore having an increased protein ratio higher than 1.0. A higher ratio would suggest that Zta is interacting with this protein of interest.

If protein interactions with Zta were established then the cellular role of the proteins would be further investigated in the context of EBV replication. Is the protein-protein interaction essential and required for EBV replication?

3.2. Results

3.2.1. Extraction of Zta and associated proteins from the nucleus

Zta has been demonstrated to translocate to the nucleus and accumulate there within 24 hours of transfection (Mahot et al., 2005). In order to be able to study this protein an efficient extraction protocol was required.

I compared a gentle cell lysis and a nuclear extraction protocol to determine efficient extraction conditions. U2OS cells were transiently transfected with a number of plasmids designed to express Zta and proteins known to co-associate with it and the protein expression and extraction efficiency levels were determined by western blot.

First, U2OS cells were transfected with plasmids expressing wild type Zta and also Zta binding partners: FLAG-tagged p53, NFκB, BMRF1 and BGLF4.
Both cell lysis and nuclear extract protocols resulted in a final supernatant of extract containing soluble proteins and a cell pellet after extraction (Figure 3.2A). The supernatant and pellet were resuspended in protein sample buffer and separated by SDS-PAGE and western blots were performed and probed for the presence of these proteins.

The first experiment confirmed that the transfection of Zta into U2OS cells was successful. The presence of wild type Zta, was confirmed by western blot (Figure 3.2B). Comparing the cell lytic reagent extraction and nuclear extract methods, Zta remained in the pellet for both methods. Only a very small amount of the transfected Zta protein is extracted and solubilised into the supernatant.

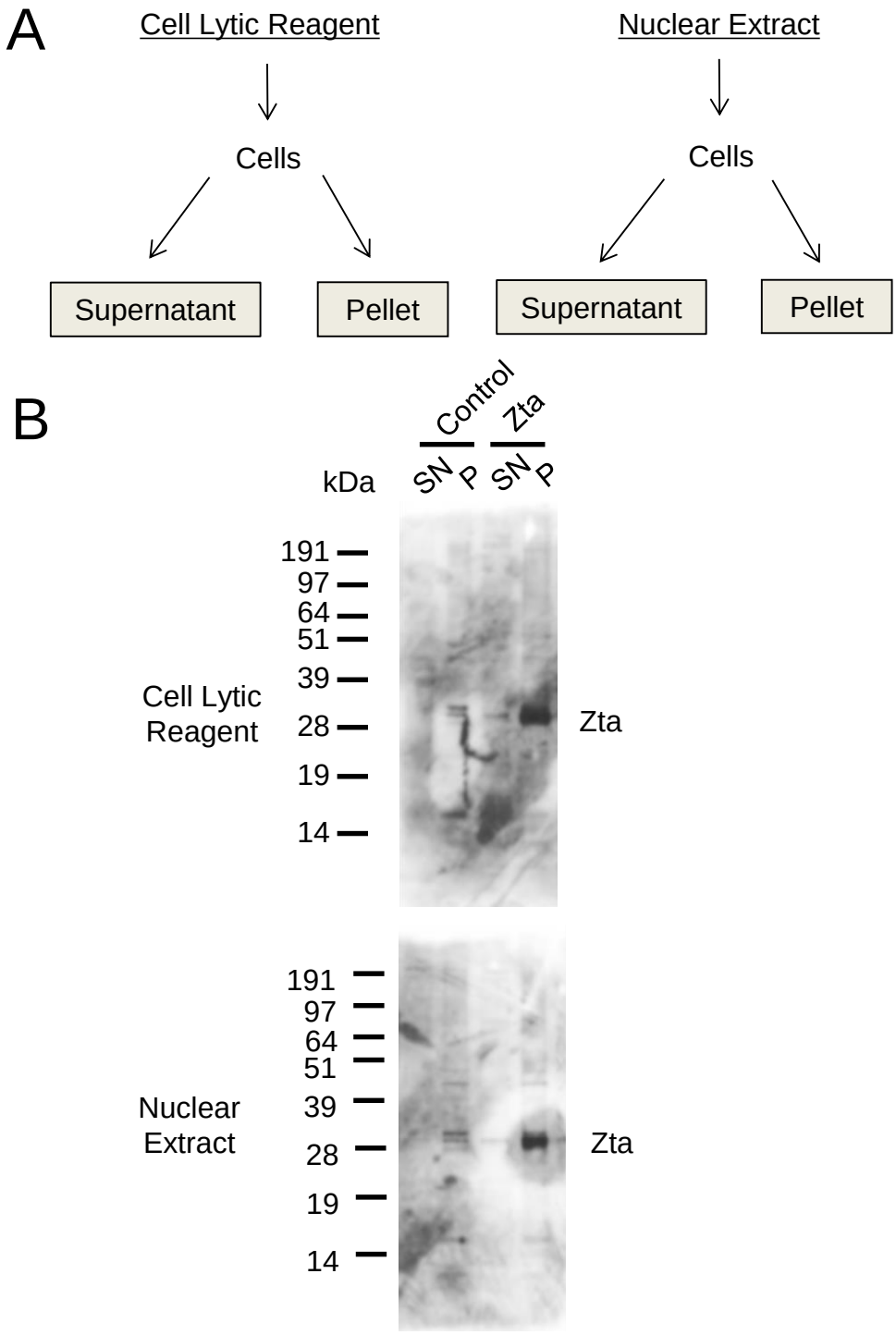


Figure 3.2 Cell lytic reagent and nuclear extract methods to determine extraction efficiency of Zta A Flow diagram of the two methods (Cell lytic reagent and nuclear extract) used for protein extraction. Both methods extracted proteins into a supernatant and a pellet remained. **B** U2OS cells were transfected with control, and Zta expression vectors and harvested after 72 hours. Proteins were extracted using CellLytic reagent or nuclear extract. The supernatant was mixed with an equal volume of protein sample buffer and the pellet boiled in the same volume of protein sample buffer. The proteins were separated by SDS-PAGE and a western blot for Zta is shown. The antibody used here was BZ1. SN refers to supernatant, P refers to Pellet

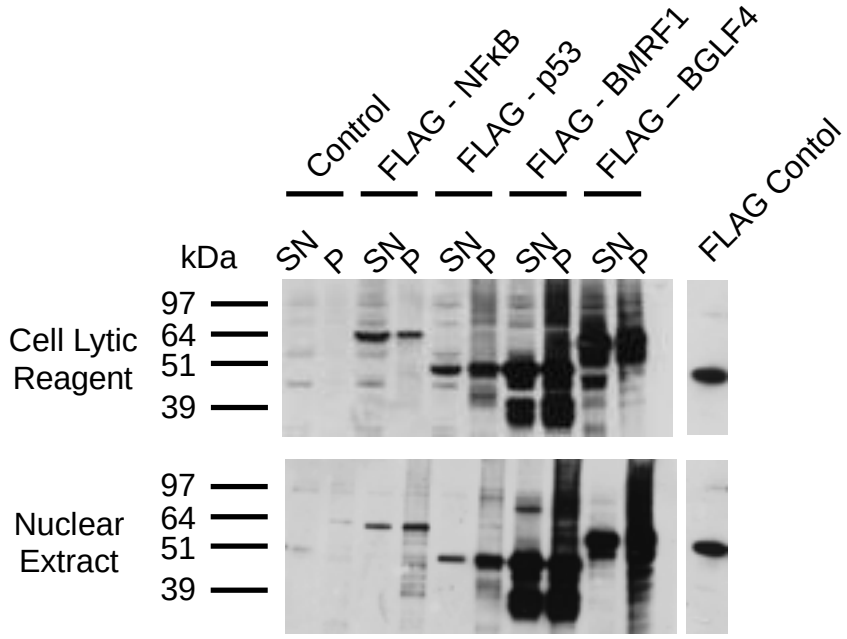


Figure 3.3 Cell lytic reagent and nuclear extract methods to determine extraction efficiency of FLAG-tagged cellular proteins. U2OS cells were transfected with control, and FLAG-tagged cellular expression vectors and harvested after 72 hours. Cellular proteins were extracted using Cell Lytic reagent and nuclear extract. The supernatant was mixed equally with protein sample buffer and the pellet boiled in the same volume of protein sample buffer. Supernatant and protein pellet were compared between each transfection. The proteins were separated by SDS-PAGE and a western blot for NFκB, p53, BMRF1 and BGLF4 shown. The antibody used here was anti-FLAG. SN refers to supernatant, P refers to Pellet

The extraction of the transfected host and viral proteins into U2OS cells was also successful. The cell lytic reagent and nuclear extract methods of protein extraction of these Flag-tagged transfected proteins was compared (Figure 3.3). For NFκB, p53, BMRF1 and BGLF4 both extraction methods allowed some soluble protein to be extracted from the cells. Not much difference was observed between the two methods apart from NFκB being exception here, where more protein was extracted into the supernatant from the pellet using the cell lytic reagent compared to the nuclear extract protocol.

From the standard cell lytic reagent and nuclear extraction methods, the majority of transfected Zta and Flag-tagged proteins remain in the nuclear pellet after a first round of extraction. The issue with proteins remaining in the nuclear pellet was addressed with the addition of a benzonase nuclease (Biedermann et al. 1989). Benzonase can degrade all forms of DNA and RNA with no proteolytic activity and releases nuclear proteins intimately associated with DNA such as histones. Benzonase was applied to further rounds of the two protein extractions applied to the pellet that was obtained from a first round of extraction. For the nuclear extract protocol, benzonase was applied to the pellet generated after the addition of solution A. A schematic diagram illustrates each of the three rounds of extraction for each method (Figure 3.4). The samples highlighted in yellow were separated by SDS-PAGE and investigated for protein extraction by western blot.

Some Zta protein is solubilised into the supernatant with the addition of benzonase to the reagents (Figure 3.5A) although some remains in the pellet. The same result is found for the nuclear extract protocol, some Zta is solubilised into the supernatants (Figure 3.5B).

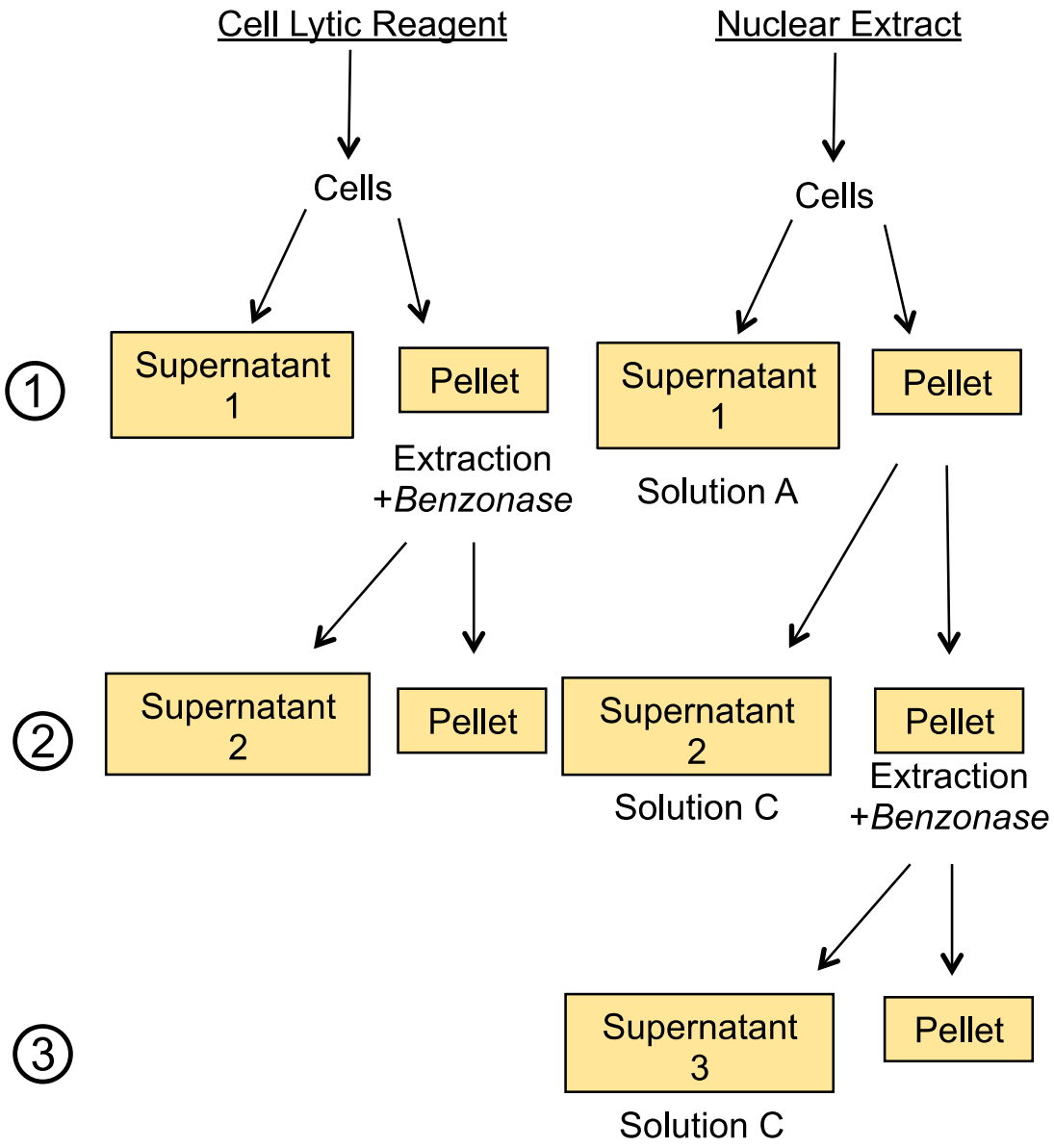


Figure 3.4 Diagram of cell lytic reagent and nuclear extract method applied to transfected U2OS cells. After transfection, U2OS cells were harvested and spun into a cell pellets before protein extraction with either cell lytic reagent or nuclear extract. The extraction methods were repeated on the cell pellet from the first extraction and 25 Units of benzonase were added to the lysis buffers after a first round of protein extraction. Highlighted boxes represent the samples explored by western blot analysis.

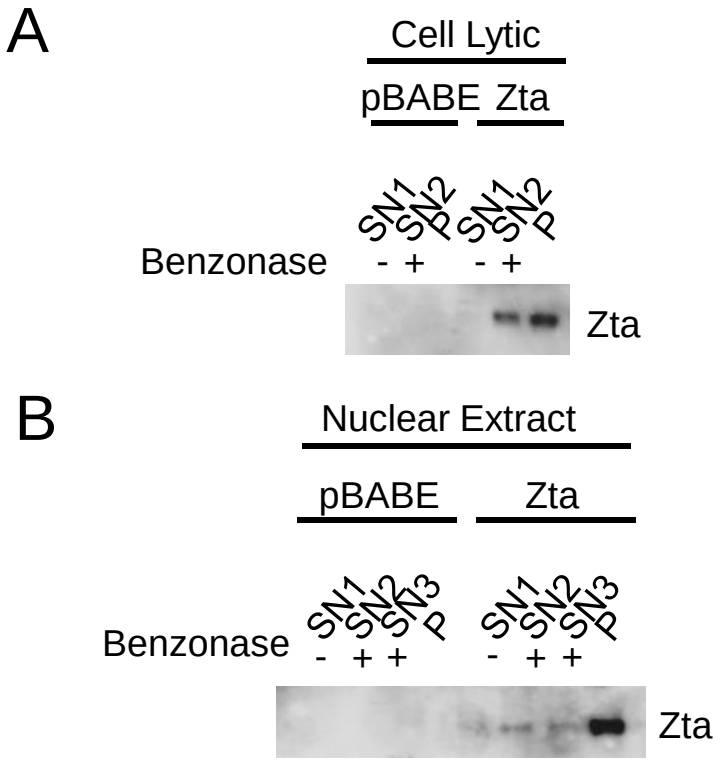


Figure 3.5 Comparison of extraction efficiency of Zta with the addition of benzonase to the extraction protocols. Cell lytic reagent or nuclear extract protocol with the addition of benzonase were used to extract transfected proteins. The supernatant was mixed with an equal volume of protein sample buffer and the pellet boiled in the same volume of protein sample buffer. The proteins were separated by SDS-PAGE and a western blot for Zta was performed. The antibody used here was BZ1. SN refers to supernatant, P refers to Pellet. **A** Cell lytic reagent with the addition of benzonase to the second round of extract attempt of transfected Zta **B** Nuclear extract with the addition of benzonase to the second and third round of extract attempt of transfected Zta

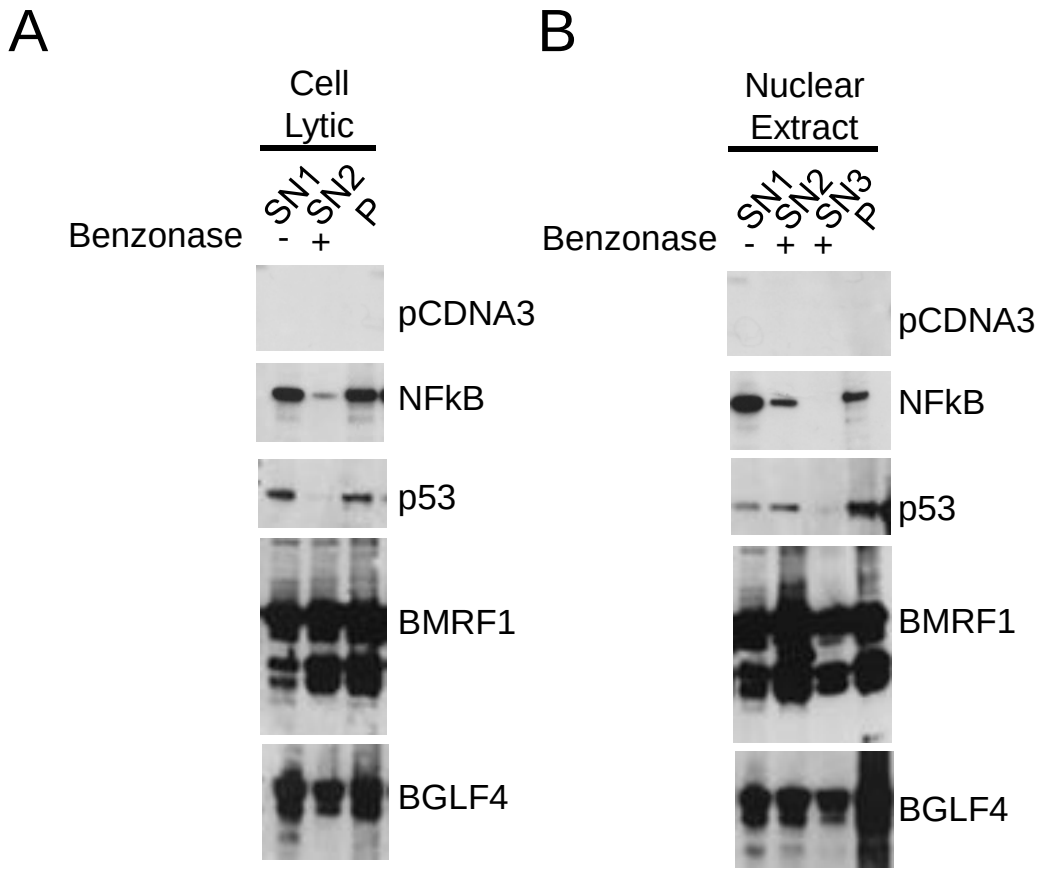


Figure 3.6 Comparison of extraction efficiency of FLAG-tagged proteins with the addition of benzonase to the extraction protocols. Cell lytic reagent or nuclear extract protocol with the addition of benzonase were used to extract transfected proteins. The supernatant was mixed with an equal volume of protein sample buffer and the pellet boiled in the same volume of protein sample buffer. The proteins were separated by SDS-PAGE and a western blot for Zta is shown. The antibody used here was anti-FLAG. **A** Cell lytic reagent with the addition of benzonase to the second round of extract attempt for extraction of transfected cellular FLAG-tagged proteins **B** Nuclear extract with the addition of benzonase to the second and third round of extract attempt for extraction of transfected cellular FLAG-tagged proteins. SN refers to supernatant, P refers to Pellet

The transfected host and FLAG-tagged viral proteins displayed a similar profile between the two extraction methods with the addition of benzonase on repeat extractions (Figure 3.6). NFκB is extracted into the first supernatant using the cell lytic reagent and the nuclear extract. The remaining proteins stay within pellet, some being solubilised into the supernatant. Some p53 is extracted into the first supernatant, with the equivalent protein amount remaining in the pellet. The viral proteins (BMRF1 and BGLF4) are again very well expressed and extracted into the first supernatant and subsequent supernatants, with some protein still remaining in the pellet, BMRF1 and BGLF4 are extracted into the supernatant well with some protein remaining in the cell pellet.

As the addition of benzonase releases some of Zta from the pellet and may be aiding the other transfected proteins here to be extracted, the concentration of benzonase and the temperature of extraction were altered in order to determine a more optimal environment for the extraction methods. It was also decided to continue with only the cell lytic reagent as the method of extraction, as currently the extraction results were comparable to the nuclear extract protocol.

The benzonase stock is a 250U/µl (Units/µl) solution and this was diluted accordingly with cell lytic reagent 1:10 in order to obtain a 25U/µl benzonase stock solution. Cell lytic reagent extraction was performed on ice, at room temperature or at 37°C on six sets of U2OS cells transfected with Zta wild type only. Six different conditions were investigated, varying the benzonase concentration and temperature of the extraction method of cell lytic reagent. The Zta transfected cell pellets were lysed with cell lytic reagent before a second extraction attempt with benzonase (Figure 3.7). Some Zta was extracted into the first supernatant at 4°C for the addition of 25U, 125U and 250U of benzonase to the cell lytic reagent. Benzonase was added to the second round of extraction for these conditions but mostly Zta remained in the cell pellet. The first round of extraction was repeated at 4°C before a second round of extraction using 1250U of benzonase. This higher amount of benzonase completely released the Zta from the cell pellet into the second supernatant at this benzonase concentration and temperature.

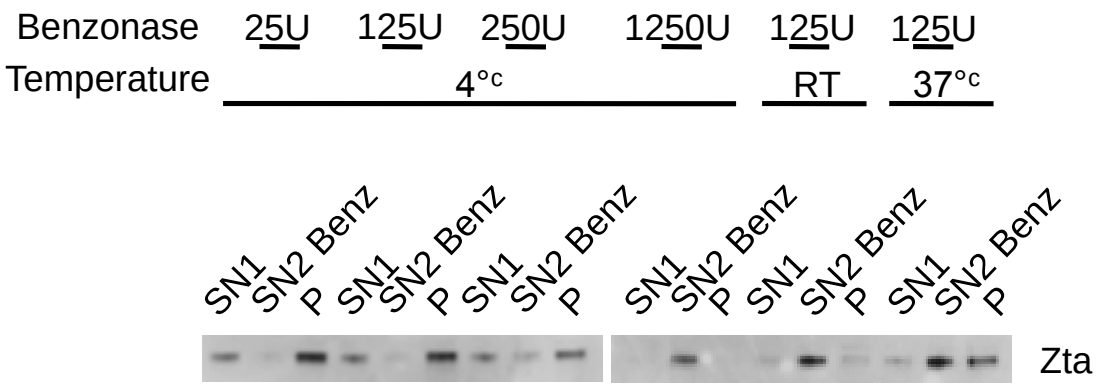


Figure 3.7 Changing benzonase conditions for cell lytic reagent extraction. U2OS cells were transfected with Zta. Cells were harvested after 72 hours and lysed with the cell lytic reagent with the addition of benzonase and lysate and pellet was run on a protein gel. The benzonase concentration was altered and shown per Unit and the temperature altered between 4°C, room temperature and 37°C. The proteins were separated by SDS-PAGE and a western blot for Zta performed. The antibody used here was BZ1.

The addition of 125U of benzonase during a second cell lysis extraction, incubating the lysis at room temperature, efficiently extracts Zta. Performing the second extraction with 125U benzonase at 37°C allowed some Zta to be extracted into the second supernatant but some Zta remains in the pellet, perhaps due to precipitation into an insoluble pellet. The addition of 125U of benzonase condition added to 100µl of cell lytic reagent at room temperature is the most optimal condition here to efficiently extract Zta proteins from the nucleus into the supernatant. The addition of 10 times more benzonase at 4°C produced a similar outcome for extraction, therefore using the lower concentration of 125U benzonase at room temperature was chosen as the optimal approach of extracting cellular and transfected proteins using the cell lytic reagent.

3.2.2. Establishing histidine-tag Zta pull down conditions

The ability of histidine-tagged Zta to interact with nickel affinity gel was assessed. Histidine tagged Zta was transfected into U2OS cells, these were lysed using the cell lytic reagent protocol with the optimal benzonase digestion conditions described. This extract was applied to and incubated with equilibrated nickel agarose beads, and the unbound extract was kept for analysis. The beads were washed three times with wash buffer before being mixed with an equal volume of protein sample buffer. A western blot was performed to detect hisZta from the extract, unbound sample and protein bound to the gel (Figure 3.8). It can be seen here that hisZta was extracted into the supernatant input). This supernatant was applied to the gel, where some hisZta bound and some did not. Therefore, the equilibrated gel, and wash conditions allowed the extracted hisZta to bind to the gel.

This was repeated to determine if there are any non-specific interactions with the nickel affinity gel (Figure 3.9). Actin was considered as an indicator of a readily detectable protein that may bind to the nickel agarose beads non-specifically. By probing the western blot for actin, presence in the gel sample would indicate if this protein binds to the nickel without a histidine tag. If this protein binds non-specifically then other cellular proteins may bind as well. The harvested transfected cells were processed into cell protein extracts and the pull down was

repeated as previously. The input, unbound and gel samples were mixed equally with protein sample buffer and separated by SDS-PAGE.
In this experiment, hisZta binds to the nickel agarose beads as demonstrated previously (Figure 3.9). Actin does not bind to the nickel agarose beads although the actin signal in the input is quite minimal compared to the hisZta signal.

To minimize any non-specific binding to the gel by cellular proteins that have an affinity for nickel or agarose, it is recommended to use imidazole in washes and even for elution of histidine-tagged proteins. Imidazole has an affinity for the nickel agarose gel and is a metal ion ligand therefore this may help minimize non-specific binding. The imidazole concentration investigated to minimise non-specific binding was either a concentration of 0mM or 10mM within the wash buffer (Figure 3.10).

Two volumes of cell extract were also compared to assess whether the binding sites on the nickel affinity gel are saturated. 250µl and 50µl of the cell extract were used. Using the larger volume of cell extract resulted in more hisZta protein binding, which shows that the gel capacity for binding is not saturated by 250µl (Figure 3.10). Actin is present in the gel sample without imidazole present in the wash buffer. Using 10mM imidazole in the wash buffer decrease the non-specific binding of actin to the gel.

In order to further minimise non-specific binding, the NaCl concentration was altered in the wash buffer and up to 1000mM can tolerated. As imidazole minimises the non-specific binding of actin to the gel, the difference between 10mM and 20mM was further investigated in the wash buffer. Therefore, either 300mM or 1000mM of NaCl was included in the wash buffer, with 10mM or 20mM imidazole. The western blot and quantitation from the pull down is shown relative to the input sample (Figure 3.11).

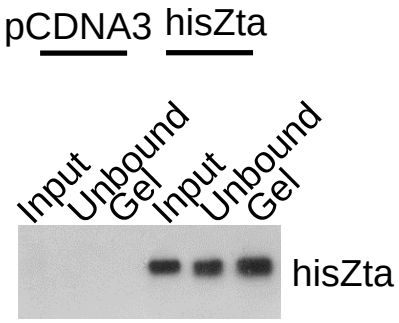


Figure 3.8 Preliminary histidine-tag Zta pull down to demostrate hisZta can bind to nickel affinity gel. U2OS cells transfected with pCDNA3 control and hisZta. After 72 hours, cell lytic extract was prepared and applied to nickel affinity gel (Sigma) and bound at room temperature. The unbound extract was retained and after washing agarose. The supernatant was mixed with an equal volume of protein sample buffer and the pellet boiled in the same volume of protein sample buffer. The proteins were separated by SDS-PAGE and a western blot for Zta performed. The antibody used here was BZ1.

Figure 3.9 Histidine-tag Zta pull down and assessment of actin binding to the nickel affinity gel. U2OS cells transfected with hisZta. After 72 hours, cell lytic extract was prepared and applied to nickel affinity gel (Sigma) and bound at room temperature. The unbound extract was retained and after washing agarose. All samples were mixed with an equal volume of protein sample buffer. The proteins were separated by SDS-PAGE and a western blot for Zta performed Antibodies used here BZ1 and α-actin.

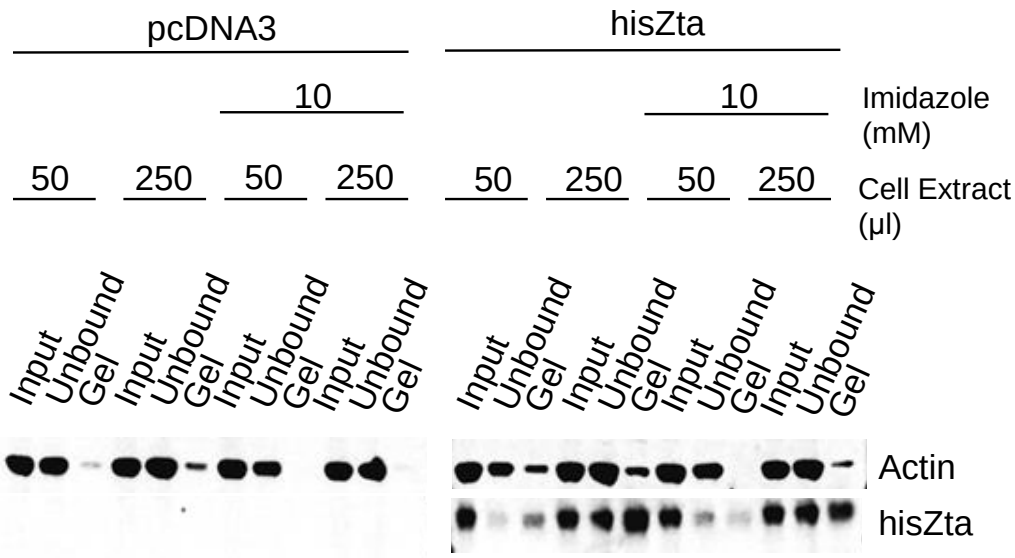


Figure 3.10 Histidine-tag Zta pull down with varying imidazole concentration and cell extract volumes. U2OS cells transfected with pCDNA3 control and hisZta. After 72 hours, cell lytic extract was prepared and applied to nickel affinity gel (Sigma) and bound at room temperature. 50µl or 250µl of extract was applied and the agarose gel washed with or without 10mM imidazole. The unbound extract was retained and after washing agarose. All samples were mixed with an equal volume of protein sample buffer. The proteins were separated by SDS-PAGE and a western blot for Zta performed Antibodies used here BZ1 and α-actin.

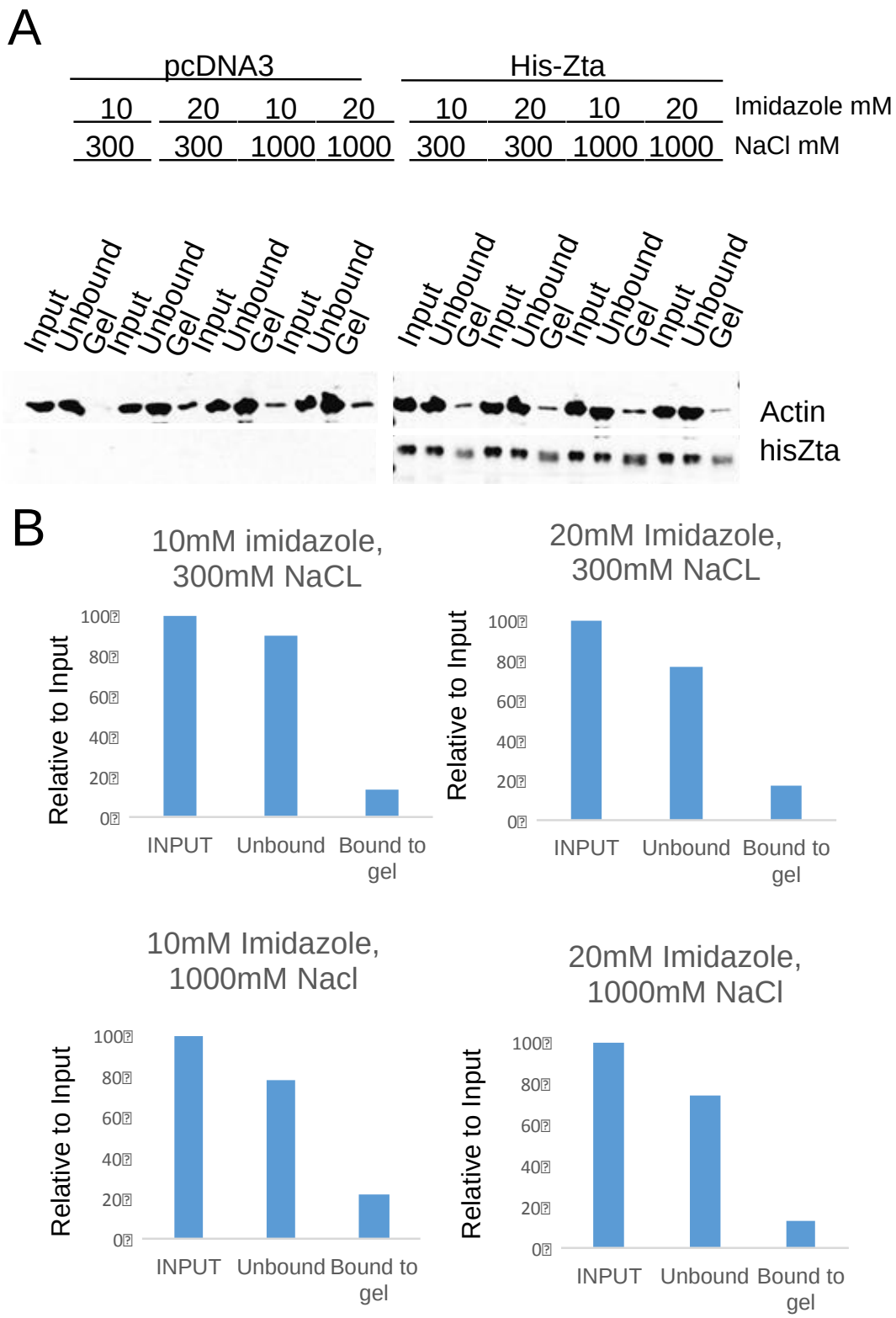


Figure 3.11 Histidine-tag Zta pull down with varying imidazole concentration and NaCl concentration in the wash buffer. A U2OS cells transfected with pCDNA3 control and hisZta. After 72 hours, cell lytic extract was prepared and 100µl applied to nickel affinity gel (Sigma) and bound at room temperature. The agarose gel washed with 10mM or 20mM imidazole and 300mM or 1000mM NaCl. The unbound extract was retained and after washing agarose. All samples were mixed with an equal volume of protein sample buffer. The proteins were separated by SDS-PAGE and a western blot for Zta performed Antibodies used here BZ1 and α-actin..**B** Quantitation of Zta bound to the gel

There was no significant difference in the binding of Zta or actin by increasing NaCl concentration from 300-1000mM or varying the imidazole from 10-20mM. Therefore 20mM of imidazole and 300mM NaCl in the wash buffer were used routinely here after.

3.2.3. SILAC labelled histidine-tagged Zta pull down and mass spectrometry

A common method to investigate protein-protein interactions includes coupling a pull down assay with mass spectrometry. This approach for interacting proteins can be used to initially direct an investigation into proteins that interact with the tagged protein of interest. It is possible to pull entire protein complexes out of the solution sample and identify interacting proteins through mass spectrometry. Progress within this field of proteomics has included the use of metabolically labelling the proteins with amino acids of varying molecular mass. This has enabled the quantitation of proteins in experiments when they are differentially labelled.

Stable isotope labelling of amino acids in cell culture (SILAC) was utilized for the next step in the experiments to investigate potential interacting partners with Zta. SILAC coupled to mass spectrometry analysis allows a quantitative method of detecting protein abundance from samples.

The U2OS cells were labelled as demonstrated by the schematic diagram (Figure 3.12). A minimum of 5 passages of cells in cell culture ensures that at least 97% the cells are metabolically labelled with the chosen isotope labels. Through cell doublings and new protein synthesis allows an efficient incorporation of these labelled amino acids (Ong & Mann 2006).

The most naturally occurring normal isotopes for carbon and hydrogen were incorporated into arginine and lysine in the control media. These were chosen as the control, named pCDNA3 R0K0. Media containing arginine as L-Arginine-13C6 Hydrochloride and lysine as L-Lysine-4,4,5,5-D4 Hydrochloride were used for the transfected hisZta cells. These were named hisZta R6K4. A total of three large T175 flasks each were transfected and harvested into cell pellets after 72

hours (Figure 3.12).

Fractions of these SILAC labelled cells were investigated for of actin and hisZta (Figure 3.13). The cells were lysed as previously described using cell lytic buffer and benzonase and a fraction of the extract was resuspended in an equal volume of protein sample buffer. The proteins were separated by SDS-PAGE. The actin levels for all six samples were equivalent. Expression of hisZta was demonstrated in all three samples that were transfected with hisZta. These SILAC labelled cell fractions would be used for future histidine-tag pull down experiments.

A consecutive histidine-tagged Zta pull down using the same SILAC labelled extracts was attempted using the optimised pull down protocol established. A schematic diagram of a consecutive histidine-tagged Zta pull down is illustrated (Figure 3.14). Three consecutive pull downs were then performed as described to attempt and maximise binding of hisZta from cell extracts.

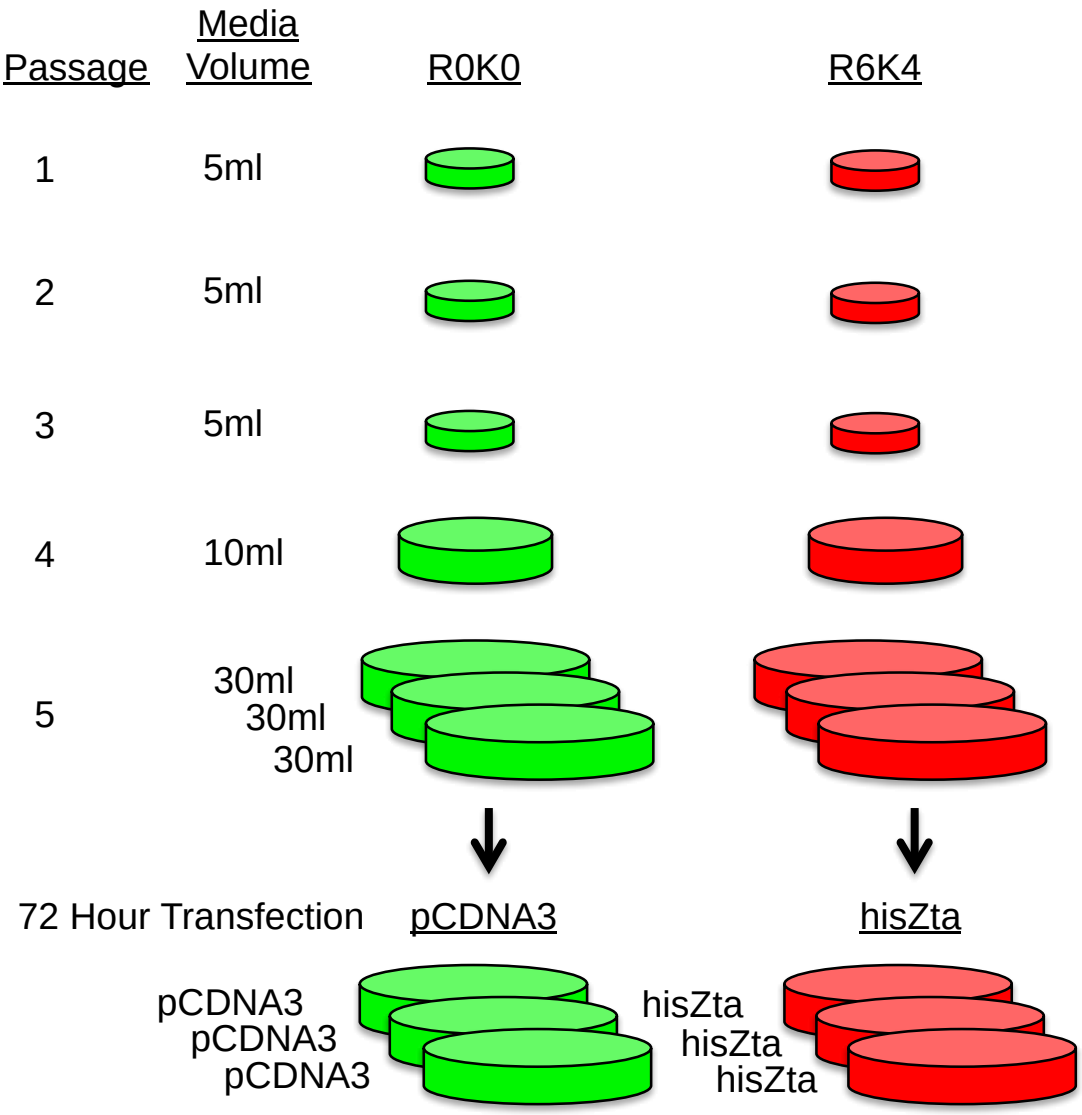


Figure 3.12 SILAC labelling of U2OS cells schematic diagram. Using light (R0K0) and medium (R6K4) media (Dundee Cell) U2OS cells are passaged 1:3 using the media volume shown. A minimum of 5 passages to ensure all of the cells are metabolically labelled. The cells are transfected with control or hisZta and harvested after 72 hours for further study

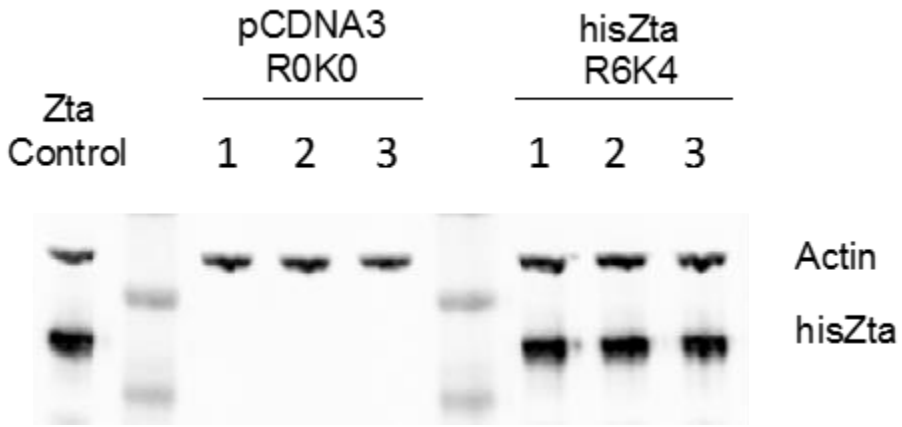


Figure 3.13 Expression of transfected control and hisZta in SILAC labeled U2OS cells. Three flasks (1,2,3) transfected with either pCDNA3 or hisZta for 72 hours. The cells were harvested and a fraction were lysed in cell lytic reagent and cell extract. A fraction of the supernatants were mixed with an equal volume of protein sample buffer. The proteins were separated by SDS-PAGE and a western blot for Zta performed Antibodies used here BZ1 and α-actin.

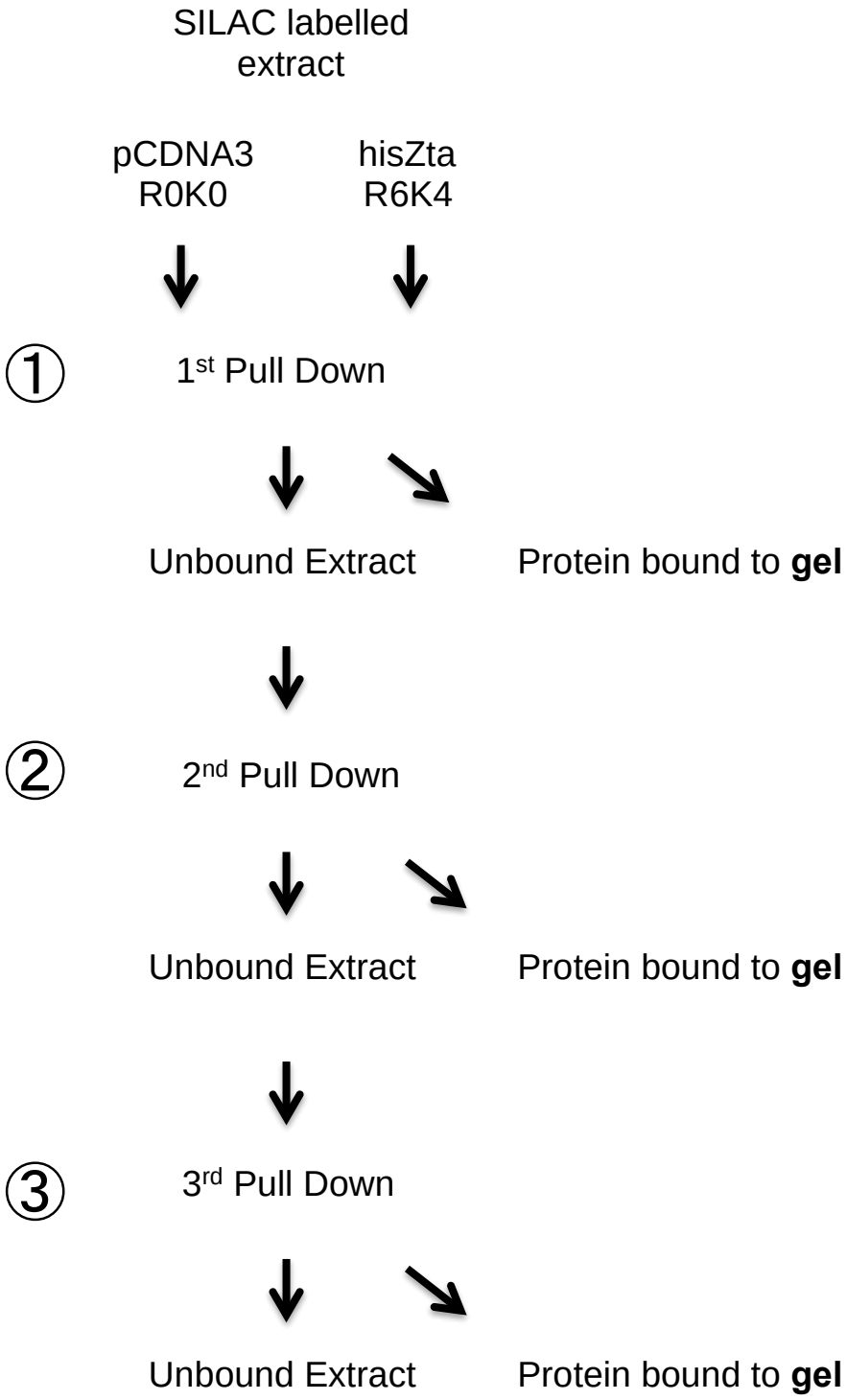


Figure 3.14 Schematic of consecutive histidine-tagged Zta pull down using SILAC labelled extracts. After attempting a histidine-tag pull down in 1, the unbound extract is applied to another gel in 2, and this unbound extract is then applied to the final gel in 3. Upon western blot of all samples, identification of the amount of hisZta binding to the gel from the same extract

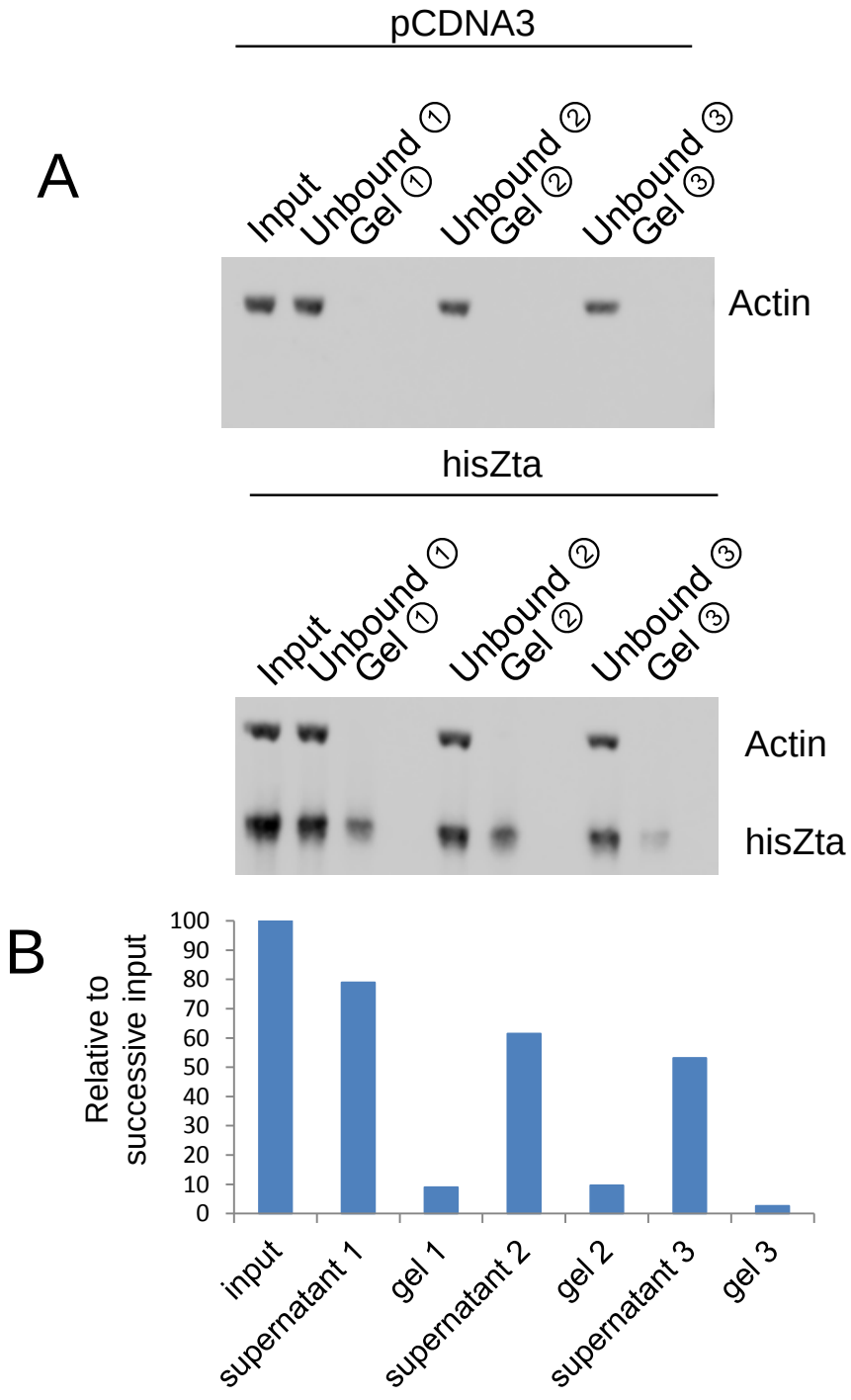


Figure 3.15 Consecutive histidine-tag Zta pull down using the same extract throughout the assay. A Histidine-tag Zta pull down using cell lytic extract in 1. Unbound extract from pull down 1 was applied to the gel in 2. This unbound extract was applied to the pull down in 3. All samples were mixed with an equal volume of protein sample buffer. The proteins were separated by SDS-PAGE and a western blot for Zta Antibodies – BZ1, α-actin, **B** Quantitation of consecutive histidine-tag Zta pull down using the same extract throughout the assay.

The control and hisZta transfected samples were applied to the gel as outlined by the schematic diagram (Figure 3.14). The samples were separated by SDS-PAGE after being boiled in protein sample buffer. The protein hisZta binds to the nickel affinity gel throughout the assay (Figure 3.15A). The unbound supernatant was kept for a second and third consecutive pull downs and the bands were quantitated relative to the input (Figure 3.15B). HisZta bound to the nickel affinity gel after incubation and washing in the first round of incubation. The western blot was detected for actin and actin did not bind non-specifically to the gel. Therefore, these SILAC labelled samples were taken forward for further investigation.

Both control and hisZta pull down samples were visualized by SimplyBlue stain (Figure 3.16A). It would be expected that if any non-specific proteins bind to the gel, the same non-specific proteins would bind to both the control gel and hisZta gel. Within the hisZta pull down, hisZta will bind through its polyhistidine tag interaction with the nickel, as well as any additional proteins through the binding of cellular proteins to hisZta. There are proteins present in both control and hisZta treated cells (Figure 3.16A). A perfect experiment would see no proteins present in the control sample, and hisZta bound to the nickel gel in the hisZta sample, other proteins present in this sample would be due to hisZta binding to any interaction partners. A western blot of the samples considered for mass spectrometry were included (Figure 3.16B)

The control and hisZta transfected samples (Figure 3.16B) were sent to the University of Bristol Proteomics facility for mass spectrometry analysis. The two samples were combined equally and separated by SDS-PAGE. The gel was cut into three gel slices before being processed into peptides for mass spectrometry. Trypsin was used to cleave the proteins into peptides. The use of trypsin to digest proteins into peptide chains is a common proteolytic enzyme. It is very specific for SILAC studies as the protein can cleave at the carboxyl-termini of lysine and arginine residues. Trypsin is the standard enzyme to use to digest protein mixtures that have been SILAC labelled (Olsen et al. 2004).

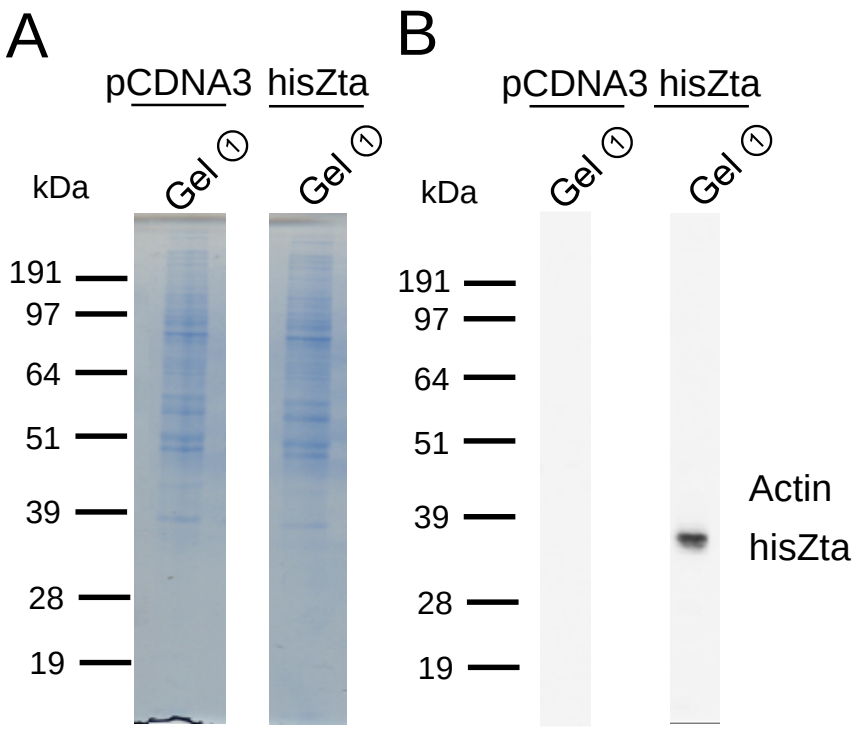


Figure 3.16 Sample confirmation before mass spectrometry analysis at University of Bristol. A The control and hisZta agarose gel boiled 1:1 in sample buffer and 2µl loaded onto a protein gel. Evidence of protein bands after SimplyBlue safestain (Invitrogen). **B** Histidine-tag Zta pull down using cell lytic extract in 1 (Figure 3.14A). The control and hisZta agarose gel boiled 1:1 in sample buffer and 2µl separated onto a protein gel. The proteins were separated by SDS-PAGE and a western blot for Zta performed. The antibody used here was BZ1.

The peptides were analysed using a LTQ-Orbitrap Velos mass spectrometer and the files were analysed using a Proteome Discoverer software (Thermo Scientific).

With the introduction here of the quantitative SILAC method to label the samples, the non-specific proteins that are pulled down and detected by mass spectrometry in both samples will effectively cancel each other out. Therefore, proteins that are in equal abundance between the two samples will return with a SILAC ratio of 1.0 after mass spectrometry analysis and quantitation. Interacting proteins with Zta will then have a SILAC ratio of greater than 1.0. A minimum threshold of speculative interactions can be applied to peptide hits and the number of SILAC peptides identified and the SILAC ratio score, in order to be more confident of any possible direct interactions.

The data from the University of Bristol Proteomics facility was returned as a Microsoft Excel spreadsheet containing proteins that were identified using peptides that the Proteome Discoverer program had recognised. Some peptides could not have a SILAC ratio attributed to them and these were not considered for analysis. A total of 3414 proteins were identified. 2546 proteins were identified with a SILAC ratio. These proteins were detected in either control or hisZta transfected sample pull downs, or just one sample. For a more illustrative analysis of these ratios, GraphPad Prism (V6.0) was utilised. A representation of the SILAC U2OS proteins returned from mass spectrometry identification was illustrated using a log2 value of the SILAC ratios. This allows a calculation of all the data points to be between -6.64 (Log2(0.01)) and 6.64 (Log2(100.0)).

Non-specific, experimental contaminants, cluster in a Gaussian (normal) distribution centred at the log2 ratio of ~0 (which corresponds to a SILAC ratio of ~1.0). A normal distribution should be centred on a log2 value of exactly 0. Any environmental contaminants will have a log2 ratio value of less than -1.0, and possible interaction partners of hisZta will have a log2 ratio value greater than 1.0.

These SILAC ratios were then converted by myself into log2 ratios and entered into GraphPad Prism software (v6.0). A histogram was created to enable a representation all of the SILAC proteins (Figure 3.17). Most of the data returned centred on a log2 value of zero. This meant that all of the non-specific proteins that were identified with a SILAC ratio of ~1.0 have a log2 value of zero. The proteins with a log2 value below -1.0 may be environmental contaminants, many were detected with a log2 value of -6.64 (log2(0.01)). The proteins with a log2 ratio of greater than 1.0 up to 6.64 (log2(100.0)) represent proteins that may interact with Zta. Anything with a ratio greater than 1.2 would suggest that the increase in protein abundance may be due to the presence of Zta. There are 41 proteins that were only identified from the hisZta sample with a SILAC that was set to 100.0 by the analysis software.

After the visualisation of the complete dataset, the proteins were individually examined in greater detail. Proteins identified in the data were ranked by peptide spectrum matches (PSM) from 100 PSMs and above (Table 3.1). This allowed an understanding of the most abundant peptides identified. Many tubulins were identified such as tubulin beta chain with the highest PSM value of 259. Heat shock proteins, structural proteins and other proteins were also identified.

To determine any possible non-specific interactions that may have bound to the nickel affinity gel, data from a user-submitted database was accessed and applied to the proteins detected here. False positives can be removed in protein-protein interaction studies (Mann 2006). The Protein Frequency Library (PFL) created by the Lamond laboratory at the University of Dundee (Boulon et al. 2010) was used. This user submitted database allows data to be collated about what non-specific proteins were detected from affinity purifications; various cell lines and matrices are presented as parameter options. This database can be accessed and a list of submitted proteins can be cross-referenced against the proteins already present, for the parameters selected. This enables a way to filter any non-specific interactions with the parameters established.

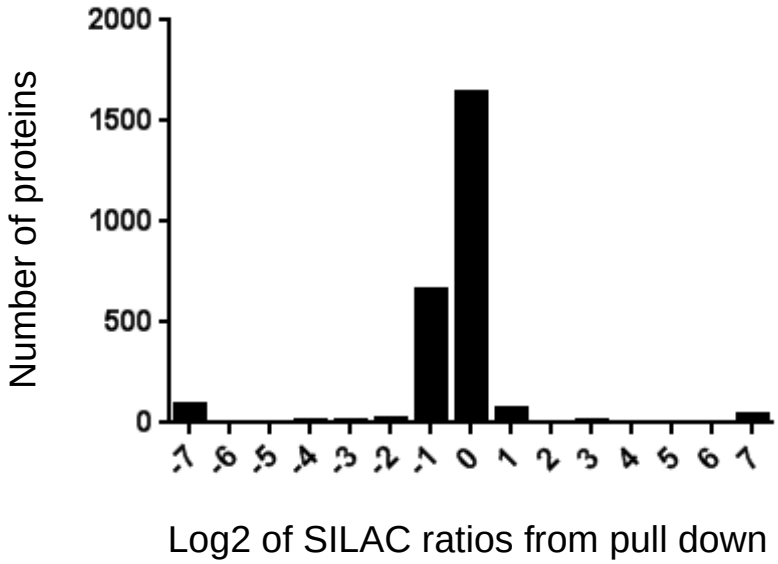


Figure 3.17 Histogram representation of log2 SILAC ratio from data returned from University of Bristol mass spectrometry analysis. Histogram of all proteins identified by mass spectrometry. Proteins pulled down with a ratio above 1.000 may be enriched by Zta.

Protein ID	Description	# PSMs
P07437	Tubulin beta chain	259
P68371	Tubulin beta-2C chain	249
Q13885	Tubulin beta-2A chain	229
Q9BVA1	Tubulin beta-2B chain	225
P68363	Tubulin alpha-1B chain	194
P04350	Tubulin beta-4 chain	188
Q9BQE3	Tubulin alpha-1C chain	188
Q71U36	Tubulin alpha-1A chain	183
P21333	Filamin-A	182
Q15233	Non-PU domain-containing octamer-binding protein	176
E7EUY0	Uncharacterized protein	168
Q13509	Tubulin beta-3 chain	167
P08670	Vimentin	158
Q13748	Tubulin alpha-3C/D chain	155
Q9BUF5	Tubulin beta-6 chain	149
P07355	Annexin A2	149
P04406	Glyceraldehyde-3-phosphate dehydrogenase	149
P02545	Prelamin-A/C	147
P14866	Heterogeneous nuclear ribonucleoprotein L	146
A8MUB1	Tubulin, alpha 1 (Testis specific), isoform CRA_a	138
P68104	Elongation factor 1-alpha 1	138
P08107	Heat shock 70 kDa protein 1A/1B	137
O43707	Alpha-actinin-4	131
Q14315	Filamin-C	127
Q3BDU5	Rhabdomyosarcoma antigen MU-RMS-40.12	120
P11142	Heat shock cognate 71 kDa protein	118
P12814	Alpha-actinin-1	116
P23246	Splicing factor, proline- and glutamine-rich	113
Q5TCI8	Lamin A/C	112
Q9Y490	Talin-1	111
Q92616	Translational activator GCN1	110
P08238	Heat shock protein HSP 90-beta	106
P14618	Pyruvate kinase isozymes M1/M2	106
B3KPW9	Tubulin, alpha 8, isoform CRA_b	104
P04264	Keratin, type II cytoskeletal 1	103
A6NMY6	Putative annexin A2-like protein	101

Table 3.1 Most abundant proteins identified by mass spectrometry analysis of U2OS SILAC hisZta pull down. Proteins ordered by peptide spectrum matches. Heat shock proteins and tubulins were most richly identified.

The protein list above SILAC ratio 1.0 was cross-referenced against the PFL for any matches of non-specific proteins. The parameters included U2OS cells and agarose gel matrix. Proteins that matched were excluded as non-specific interactions. The proteins that remained and were not excluded were considered for further analysis (Table 3.2 and Table 3.3).

Some SILAC proteins were only identified from the heavy labelled Zta extract and these were attributed a SILAC ratio of 100.0. The proteins that were only identified in the heavy labelled hisZta pull down were then ranked by PSM value (Table 3.2). Transfected hisZta was only identified in this sample and not in the control sample, which is to be expected. The PSM value for hisZta was 24 and was the most abundant protein only identified in the hisZta transfected sample (Table 3.2).

The protein list was investigated for each protein and each accession number searched within the UniProt database. Proteins with transcriptional function were initially targeted for potential further analysis, as the cell line studied here is EBV negative. Zta may be interacting with other transcription factors and forming transcriptional complexes, in addition to interactions contributing to replication function.

The majority of the proteins that were identified only had one peptide attributed to them. These proteins were identified to only be in the hisZta pull down sample. The program that processes the raw mass spectrometry data has based the identification of these proteins on one peptide identified.

Accession	Description	# PSM	SILAC Ratio
hisZta	sp|hisZta|HisZta	24	100.0
Q9P2P6	StAR-related lipid transfer protein 9	2	100.0
Q99797	Mitochondrial intermediate peptidase	2	100.0
Q96NJ3	Zinc finger protein 285	2	100.0
Q9BWV3	Cytidine and dCMP deaminase domain-containing protein 1	1	100.0
Q8TCF0	LBP protein	1	100.0
Q8NI99	Angiopoietin-related protein 6	1	100.0
Q8IYK2	Coiled-coil domain-containing protein 105	1	100.0
Q8IXL1	ELK4, ETS-domain protein (SRF accessory protein 1) (Fragment)	1	100.0
Q8IWJ2	GRIP and coiled-coil domain-containing protein 2	1	100.0
Q7L0J3	Synaptic vesicle glycoprotein 2A	1	100.0
Q6P4Q7	Metal transporter CNNM4	1	100.0
Q5T2Q4	Cyclin-Y-like protein 2	1	100.0
Q587J7	Tudor domain-containing protein 12	1	100.0
Q14573	Inositol 1,4,5-trisphosphate receptor type 3	1	100.0
P36575	Arrestin-C	1	100.0
B7ZLE7	DAPK1 protein	1	100.0
B4DS81	cDNA FLJ57186, highly similar to ARF GTPase-activating protein GIT1	1	100.0

Table 3.2 SILAC proteins only identified from the heavy labelled Zta pull down sample with an attributed ratio of 100.0. These proteins including hisZta, were identified only from the pull down sample containing hisZta. This indicates that these cellular proteins may have an interaction with hisZta.

Two proteins only identified from the hisZta pull down sample were initially considered for further analysis after their functions were assessed. The proteins chosen for further analysis were ZNF285 and ELK4.

ZNF285 is a nuclear zinc finger (ZNF) transcription factor. The ZNF family is a diverse group of proteins that have many cellular functions (Laity et al. 2001). ZNF285 may play a role in transcriptional regulation and bind DNA. Zinc finger proteins have been shown to be essential for EBV lytic replication as previously stated in Chapter 1, ZBP89 is also known as ZNF148, which is essential for lytic replication (Baumann et al. 1999).

ELK4 is also known as ETS-domain protein (E26 transformation-specific) or SRF accessory protein 1. ETS factors can act as transcriptional activators or repressors (Sharrocks 2001). ETS factors do interact with other proteins to facilitate DNA binding (Verger & Duterque-Coquillaud 2002). ELK4 has also been demonstrated to interact with BRCA1 (Chai et al. 2001). These two proteins were considered with other potential proteins outlined below.

Proteins identified with a SILAC ratio of above 1.0 could be considered to be making an interaction with Zta, a minimum threshold SILAC ratio of 1.2 was chosen as a baseline to start investigation of cellular proteins that may interact with hisZta (Table 3.3). Proteins that returned a SILAC ratio greater than 1.2 were ordered here by SILAC ratio.

Accession	Description	Zta/Control	Zta/Control Count
Q9P2X0	Dolichol-phosphate mannosyltransferase subunit 3	2.68	1
O15360	Fanconi anemia group A protein	2.59	1
Q9NZ45	CDGSH iron-sulfur domain-containing protein 1	1.85	1
Q8IYB3	Serine/arginine repetitive matrix protein 1	1.84	2
P50895	Basal cell adhesion molecule	1.82	1
Q8NBT2	Kinetochore protein Spc24	1.52	2
O60885	Bromodomain-containing protein 4	1.52	6
A2RRP1	Neuroblastoma-amplified sequence	1.48	1
Q86SK9	Stearoyl-CoA desaturase 5	1.47	2
Q96FZ7	Charged multivesicular body protein 6	1.45	1
Q07954	Prolow-density lipoprotein receptor-related protein 1	1.44	1
Q9H9T3	Elongator complex protein 3	1.42	2
P52948	Nuclear pore complex protein Nup98-Nup96	1.42	5
P10586	Receptor-type tyrosine-protein phosphatase F	1.40	1
P98172	Ephrin-B1	1.34	2

Table 3.3 SILAC proteins identified with an attributed ratio greater than 1.2. An increase in SILAC ratio indicates a possible interaction with Zta.

The protein list was investigated and each accession number searched within the UniProt database. Again, proteins with transcriptional function were initially targeted for potential further analysis, as the cell line studied here is EBV negative. Zta may be interacting with other transcription factors and forming transcriptional complexes, in addition to interactions contributing to replication function. Also possible replication factors were considered. Three more proteins with a SILAC ratio between greater than 1.2 were therefore considered for further analysis.

Fanconi anaemia complementation group A protein (FANCA) is a DNA repair protein. This protein had a SILAC ratio of 2.59. The FANC group of proteins are recognised to play a role in post-replication repair of DNA. The FANCA protein interacts with members of the FANC family and also BRCA1 (Folias et al. 2002).

Bromodomain-containing protein 4 (BRD4) is a member of the bromodomain and extra terminal domain family (Zeng & Zhou 2002). This protein had a SILAC ratio of 1.52. BRD4 can act to remodel chromatin has been shown to promote gene transcription through protein interactions aiding RNA polymerase II (Itzen et al. 2014), as well as being essential for some DNA virus replication (Wang et al. 2012).

Elongator complex protein 3 (ELP3) is a histone acetyltransferase subunit of RNA Polymerase II and may be able to remodel chromatin (Hawkes et al. 2002). This protein had a SILAC ratio of 1.42. ELP3 also regulates the transcription of the HSP70 gene (Han et al. 2007; Li et al. 2011).

Antibodies were purchased from AbCam and western blots of pull downs between control and hisZta transfected samples were performed. This attempted to try to assess whether the proteins interact with Zta.

3.2.4 Attempt to identify host protein interactions with Zta.

The investigation of the proteins chosen for follow up analysis was repeated several times using multiple transfections for both EBV negative U2OS cells and EBV positive HEK293-*BZLF1*-KO cells. The HEK293 EBV positive cells contain the EBV genome that does not contain the BZLF1 gene, therefore lytic cycle is only possible with transfection with Zta (Feederle et al. 2000). Pull down assays were repeated using cell extracts from these cells after confirmation that Zta was expressed by western blot and Zta initiated the lytic cycle by qPCR.

The replication activity of hisZta in HEK293-*BZLF1*-KO cells was assessed after transfection using a qPCR assay, before being used in a pull down for EBV positive cells. HEK293-*BZLF1*-KO cells were harvested after 96 hours after transfection of empty vector or hisZta. Zta initiates lytic replication of the EBV genome. A sample of cells was used for a western blot to confirm the expression of transfected hisZta and to check actin expression levels (Figure 3.18A). Genomic DNA was extracted from harvested cells and this was processed for qPCR analysis.

A qPCR assay was performed for new viral copy numbers after DNA replication. This was measured using primers for the viral DNA polymerase gene BALF5. Primers for β-globin were used to standardise the samples. A control transfection detected the baseline EBV genome present. A transfection with hisZta led to lytic replication, an increase in the level of EBV genome (Figure 3.18B).

The qPCR results confirmed that lytic replication was initiated upon hisZta transfection. These cells were used for subsequent pull down assay after lysis with cell lytic buffer with benzonase.

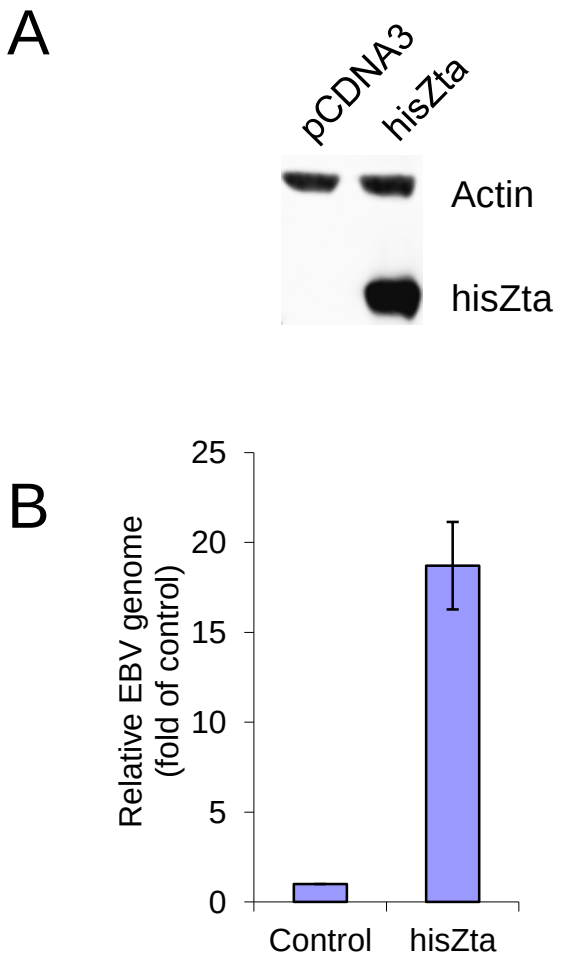


Figure 3.18 HEK293-*BZLF1*-KO transfection and qPCR of EBV genome load. HEK293-*BZLF1*-KO were transfected with empty vector or hisZta. After 96 hours cells were lysed in cell lytic reagent or genomic DNA extracted. **A** Western blot of actin and hisZta expression. The proteins were separated by SDS-PAGE and a western blot for Zta performed. The antibody used here was BZ1. **B** Quantitative Real Time PCR (QPCR) was used to detect the presence of the EBV genome and beta-globin. Results were standardised by the amount of betaglobin present。 qPCR repeated in triplicate and the standard error displayed between the experiments.

A hisZta pull down with elution was performed to investigate if there is a FANCA interaction with hisZta first using transfected U2OS extracts. Fanconi Anemia Complementation Group A (FANCA) protein had a SILAC ratio of 2.59. The predicted band size of FANCA is 130kDa. In U2OS cells, the antibody recognizes 3 prominent bands in both control and hisZta input, none of which are in this range, with multiple bands at various molecular weights (Figure 3.19A). Two of these protein bands were detected in both the control and Zta pull down sample between 39kDa and 64kDa using the FANCA antibody (Figure 3.19A). A repeat of the pull downs were performed using HEK293-*BZLF1*-KO cells that were also transfected with hisZta (Figure 3.19B). Four distinct bands were identified in both control input and pull down sample input between 64Kda and 28kDa, none of these were at the expected molecular weight.

In addition, one band was seen in the control and Zta pulldown that corresponds with a FANCA detected band (Figure 3.19B). This unique protein band was detected in the Zta pull down sample between 64kDa and 51kDa. This band is not present in the control pull down. Many bands recognised by the antibody not at the predicted protein molecular weights, the epitope for the antibody was investigated.

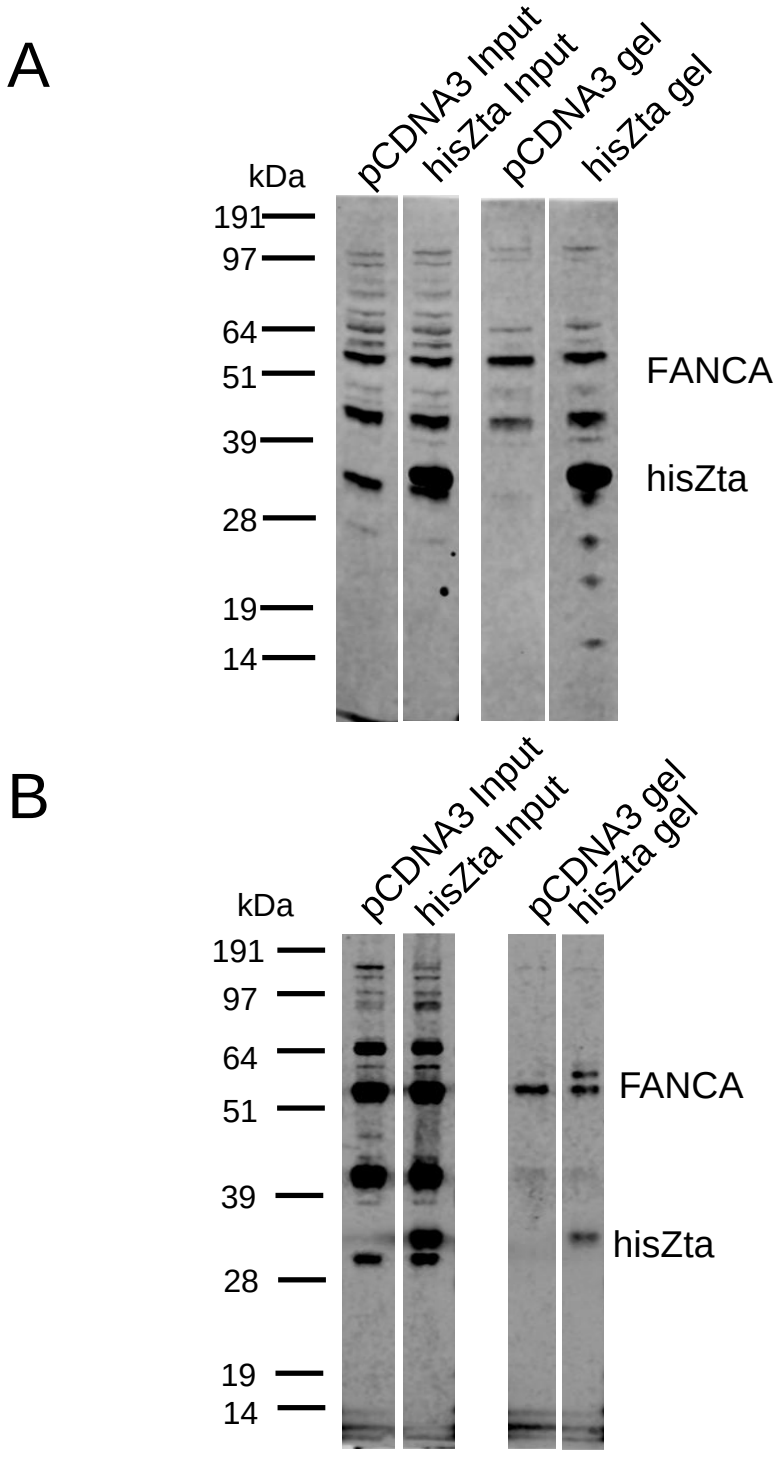


Figure 3.19 Histidine-tag Zta pull down for FANCA using transfected EBV negative and EBV positive cell extracts. Cells were transfected control or Zta vector. After 72/96 hours cells were lysed in cell lytic reagent. 100µl of supernatant was added to 25µl of HIS-select gel (Sigma) and bound at room temperature (RT). After washing, agarose was mixed 1:1 with Sample buffer and run on a protein gel for a western blot. Antibodies (ab5063) for FANCA and BZ1 for Zta. **A** U2OS cells **B** HEK293-BZLF1-KO cells

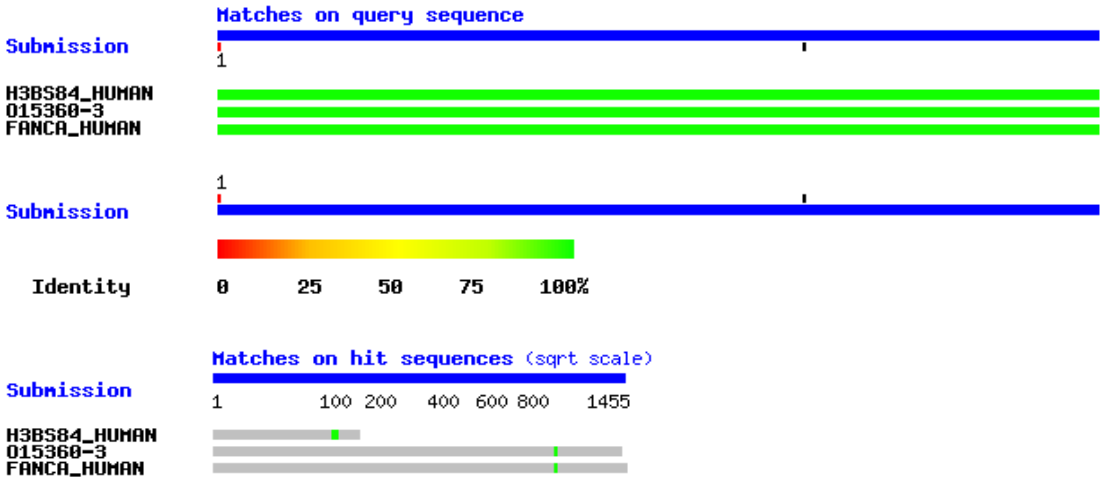


Figure 3.20 BLASTP search using UniProtKB database against the FANCA epitope recognised by the antibody. The amino acid sequence corresponding to amino acids 995 - 1009 of Human FANCA was used in a BLASTP search against the homo sapien UniProtKB database (release 2015_08 of 22-Jul-2015). Three proteins were returned that had a match on the amino acid epitope sequence recognised by the antibody. Graphical overview of the **matches** on query sequence and location of the epitope match on the protein sequence

Accession	Description	kDa	Length (aa)
H3BS84	Fanconi anemia group A protein (Fragment)	20.6	182
O15360-3	Isoform 3 of Fanconi anemia group A	159	1424
O15360	Fanconi anemia group A protein	163	1455

Table 3.4 Proteins identified from the BLASTP search of FANCA amino acid sequence (995aa – 1009aa) epitope.

A BLASTP search was performed to compare a protein query sequence against a protein sequence database. The epitope amino acid sequence recognized by the chosen antibodies was searched against the UniProtKB database (Figure 3.20). The amino acids 995-1009 of FANCA contain the epitope of where the antibody recognizes. The BLASTP search returned three proteins with an exact match for the query amino acid sequence recognised by the antibody (Figure 3.20). One of these proteins (FANCA_HUMAN) was the reported protein recognized by the antibody, O15360 at 163kDa. The location of the epitope match on the protein sequence is also displayed. The three proteins were ordered with their amino acid sequence length from smallest to largest, as found in Figure 3.20 using the UniProt database for the proteins identified from the BLASTP search (Table 3.4). The other two proteins returned include a fragment and also an isoform of FANCA. Isoform O15360-3 molecular weight is 159kDa; H3BS84 molecular weight is 21kDa.

By using these molecular weights of the proteins as a potential guide, cross-referencing against the western blot for the U2OS and HEK293-BZLF1-KO extract pull downs would indicate if the antibody may be binding to other epitopes located within proteins. The bands on both western blots do not match with these reported molecular weights.

A hisZta pull down with elution was performed to investigate if there is a BRD4 interaction with hisZta first using transfected U2OS extracts. BRD4 had a SILAC ratio of 1.53. The predicted band size for BRD4 is 152kDa. In U2OS cells, two distinct bands were identified using the BRD4 antibody both pull down samples (Figure 3.21). Being the highest most dominant band this band between 191kDa and 97kDa this was taken as BRD4.

Two distinct bands were identified in both control and pull down sample using the BRD4 antibody. There is more BRD4 present in the hisZta gel sample than the control pull down sample (Figure 3.21A). The upper band is located between 191kDa and 97kDa and this was quantitated. The quantitative value given by ImageStudio (Li-cor) for the upper band in the control gel sample was 0.429 and in the Zta pull down sample was 1.02. This was a fold change of 2.4. The SILAC ratio for BRD4 was 1.53.

Investigating BRD4 in EBV positive cells was also performed. Only one distinct band was identified in both control and pull down gel sample between 191kDa and 97kDa (Figure 3.21B). There is more BRD4 present in the control gel sample than the Zta pull down sample in this pull down.

As a lower molecular weight band was also observed (Figure 3.21A), a BLASTP search was performed to compare a protein query sequence against a protein sequence database as outlined before for FANCA antibody. The amino acids 150-250 of BRD4 contain the epitope that the antibody recognizes. The epitope amino acid sequence recognized by the BRD4 antibody was searched against the UniProtKB database. The BLASTP search returned nine proteins with an exact match for the amino acid sequence (Figure 3.22). One of these proteins was the reported protein recognized by the antibody (BRD4_HUMAN). The other proteins returned include fragments and also isoforms of varying sequence length and molecular weights, which are displayed with the amino acid sequence length (Table 3.6).

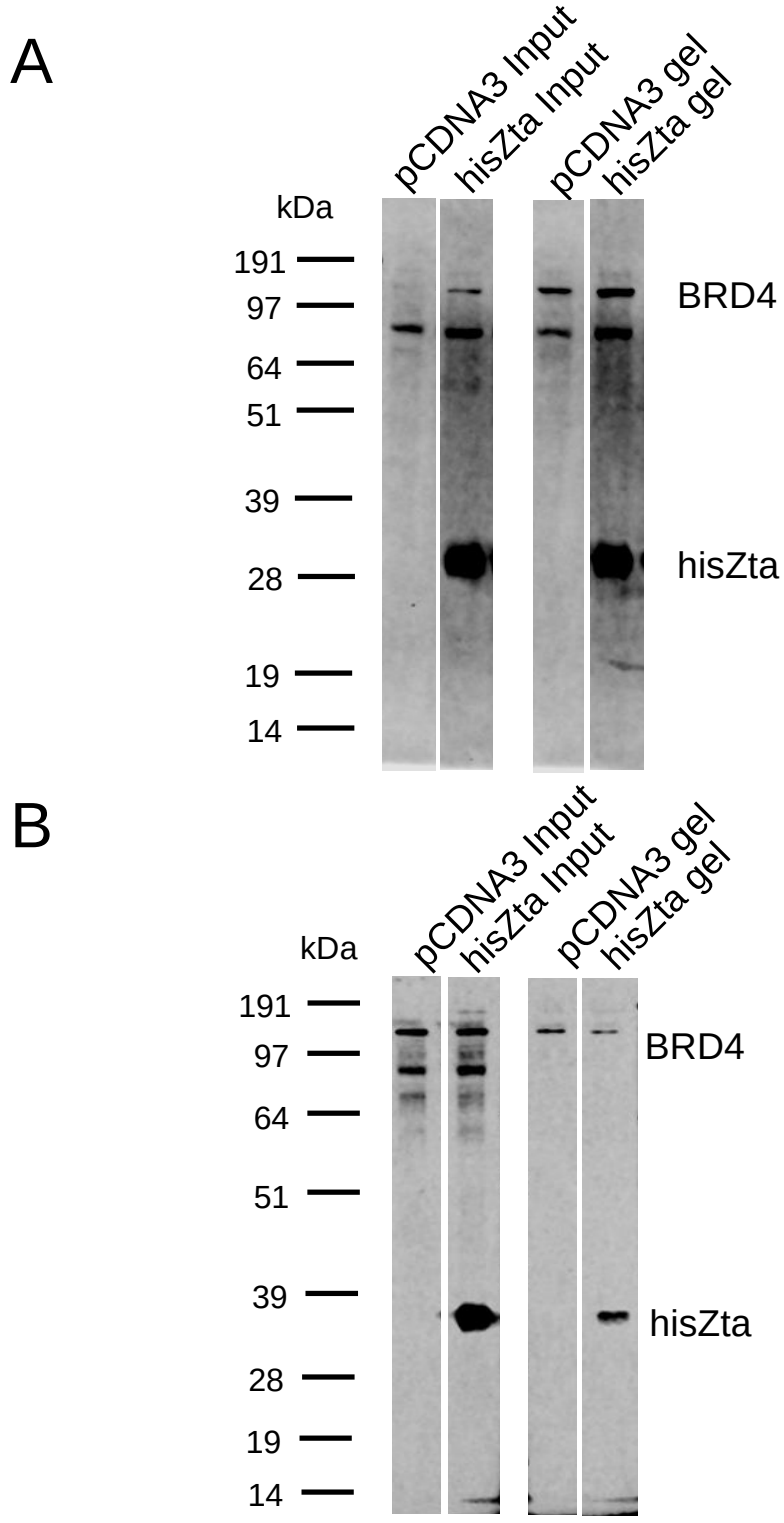


Figure 3.21 Histidine-tag Zta pull down for BRD4 using transfected EBV negative and EBV positive cell extracts. Cells were transfected control or Zta vector. After 72/96 hours cells were lysed in cell lytic reagent. 100µl of supernatant was added to 25µl of HIS-select gel (Sigma) and bound at room temperature (RT). After washing, all samples were then mixed 1:1 with protein sample buffer and run on a protein gel. This was western blotted with BZ1 for Zta and BRD4 antibodies. **A** U2OS cells **B** HEK293-BZLF1-KO cells

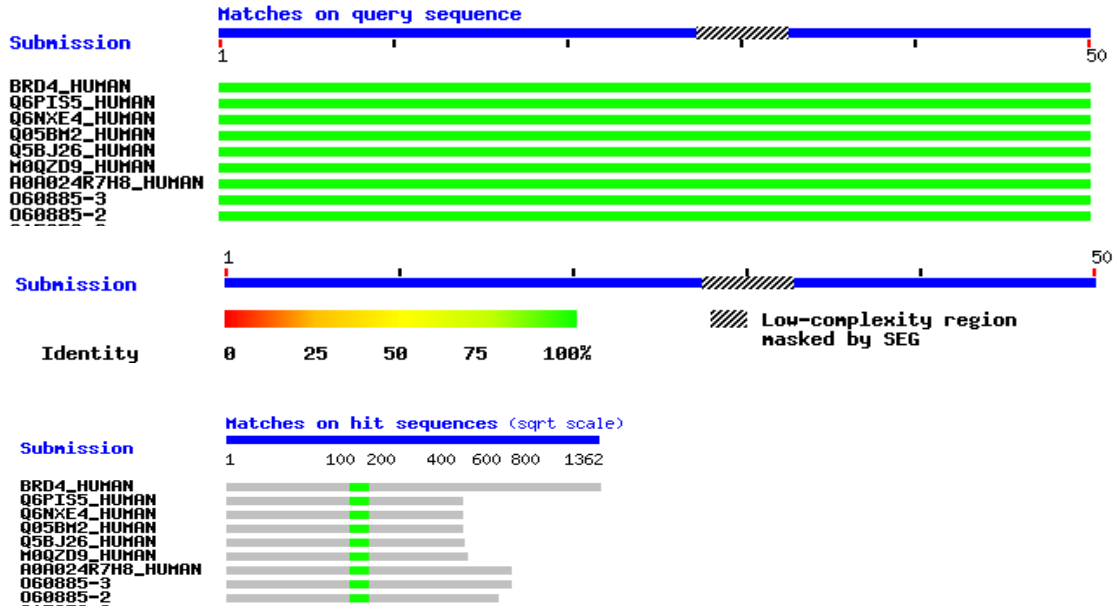


Figure 3.22 BLASTP search using UniProtKB database against the BRD4 epitope recognised by the antibody. The amino acid sequence corresponding to Human BRD4 aa 150-250 was used in a BLASTP search against the homo sapien UniProtKB database (release 2015_08 of 22-Jul-2015). Nine proteins were returned that had a match on the amino acid epitope sequence recognised by the antibody.

Accession	Description	kDa	Length (aa)
O60885	Bromodomain-containing protein 4	152	1362
Q6PIS5	BRD4 protein (Fragment)	61	548
Q6NXE4	BRD4 protein (Fragment)	61	548
Q05BM2	BRD4 protein (Fragment)	61	549
Q5BJ26	BRD4 protein (Fragment)	61	550
M0QZD9	Bromodomain-containing protein 4	63	572
A0A024R7H8	Bromodomain containing 4, isoform CRA_b	88	794
O60885-3	Isoform B of Bromodomain-containing protein 4	88	794
O60885-2	Isoform C of Bromodomain-containing protein 4	80	722

Table 3.5 Proteins identified from the BLASTP search of BRD4 amino acid sequence (150aa-250aa) epitope

By using these molecular weights of the proteins as a potential guide, cross-referencing against the western blot in Figure 3.21 would indicate if the antibody binds to other proteins containing the epitope recognised by the antibody.

The upper band on the western blot is between 191kDa and 97Kda for both U2OS cells and HEK293-*BZLF1*-KO cells. The only protein within this molecular range is the reported BRD4 protein at 152kDa; therefore this protein band may be most likely to be BRD4 detected by the antibody. This may indicate an interaction, but BRD4 binds to the control affinity gel.

The lower band on the western blot is between 97kDa and 64kDa for both U2OS cells and HEK293-*BZLF1*-KO cells. Three proteins that match this molecular weight are shown (Table 3.6). These include isoform CRA_b of BRD4 (A0A024R7H8) at 88kDa, Isoform B of BRD4 (O60885-3) at 88kDa and Isoform C of BRD4 (O60885-2) at 80kDa. Therefore, these may be variants of BRD4 observed by the antibody (Table 3.6).

A hisZta pull down with elution was performed to investigate if there is an ELK4 interaction with hisZta first using transfected U2OS extracts. ELK4 had a SILAC ratio of 100.0 therefore the protein should be only in the Zta pull down sample. The predicted band size is 40kDa for ELK4. In U2OS cells, one distinct band was identified using the ELK4 antibody that was not the predicted molecular weight. The band recognised by the antibody was above 97kDa (Figure 3.23A). The protein hisZta bound to the gel.

Investigating ELK4 in EBV positive cells was performed and only one distinct band was identified in both control and pull down sample as seen in U2OS cells. No bands were detected in the pull down here (Figure 3.23B). Protein bands corresponding to ELK4 were not readily detected in both the control and Zta pull down sample. The protein hisZta bound to the gel.

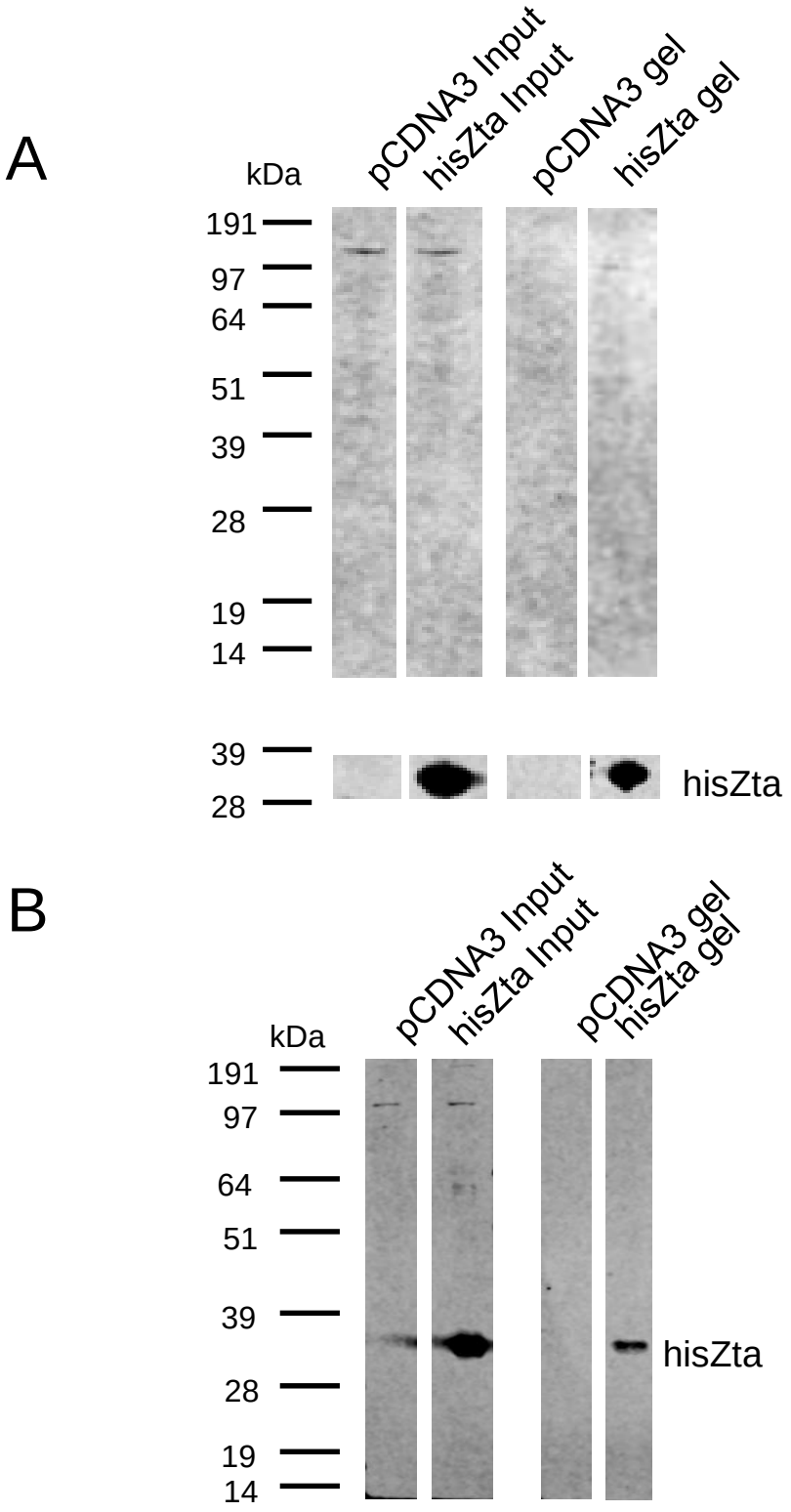


Figure 3.23 Histidine-tag Zta pull down for ELK4 using transfected EBV negative and EBV positive cell extracts. Cells were transfected control or Zta vector. After 72/96 hours cells were lysed in cell lytic reagent. 100µl of supernatant was added to 25µl of HIS-select gel (Sigma) and bound at room temperature (RT). After washing, all samples were then mixed 1:1 with protein sample buffer and run on a protein gel. This was western blotted with ELK4 antibodies and BZ1 for Zta. **A** U2OS cells **B** HEK293-BZLF1-KO cells

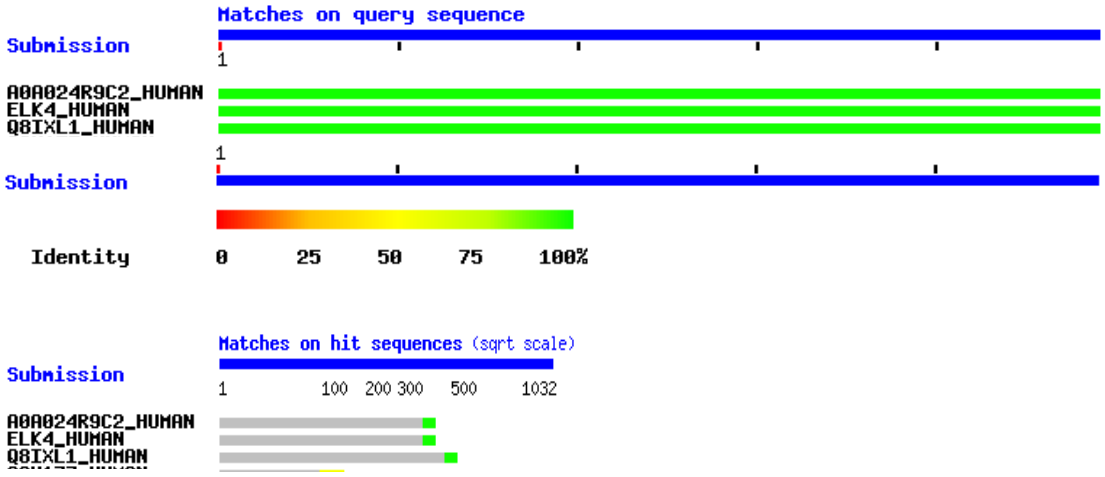


Figure 3.24 BLASTP using UniProtKB database against the ELK4 epitope recognised by the antibody. The region corresponding to internal sequence amino acids 382-431 of Human ELK4 (NP_689567) was used in a BLASTP search against the Homo sapiens UniProtKB database (release 2015_08 of 22-Jul-2015). Three proteins were returned that had a match on the amino acid epitope sequence recognised by the antibody.

Accession	Description	kDa	Length (aa)
A0A024R9C2	ELK4, ETS-domain protein	47	431
P28324	ETS domain-containing protein	47	431
Q8IXL1	ELK4, ETS-domain protein	56	521

Table 3.6 Proteins identified from the BLASTP search of ELK4 amino acid sequence epitope

The amino acids 382-431 of ELK4 are the site of epitope of where the antibody recognizes. The BLASTP search returned three proteins with an exact match for the amino acid sequence (Figure 3.24). One of these proteins was the reported protein recognized by the antibody (ELK4_HUMAN). The other proteins returned include isoforms that were 47kDa and 56kDa that also do not correspond to the band seen between 191kDa and 97kDa.

A hisZta pull down was performed to investigate if there is an interaction between hisZta for ELP3 first using transfected U2OS extracts. ELP3 had a SILAC ratio of 1.42. The predicted band size is 62kDa for ELP3. Protein bands corresponding to the molecular weight of ELP3 were detected in both the control and Zta pull down sample input (Figure 3.25). In addition, multiple bands were identified here in the sample

In U2OS cells, protein bands corresponding a band at about 62kDa (between 64kDa and 51kDa) were detected in both the control and Zta pull down sample input. The protein hisZta did bind to the gel and a band between 64kDa and 51kDa was demonstrated in the hisZta pull down only using the ELP3 antibody (Figure 3.25A)

Investigating ELP3 in EBV positive cells was also performed. Three distinct bands were identified in both control and pull down sample. No bands were detected in the pull down here. The protein hisZta bound to the gel (Figure 3.25B).

The amino acids 240-445 of ELP3 contain the epitope of where the antibody recognizes. The BLASTP search returned seven proteins with an exact match for the amino acid sequence (Figure 3.26). One of these proteins was the reported protein recognized by the antibody (ELP3_HUMAN). The other proteins returned include isoforms of ELP3. By using these molecular weights of the proteins as a potential guide, these were cross-referenced against the western blot in Figure 3.25.

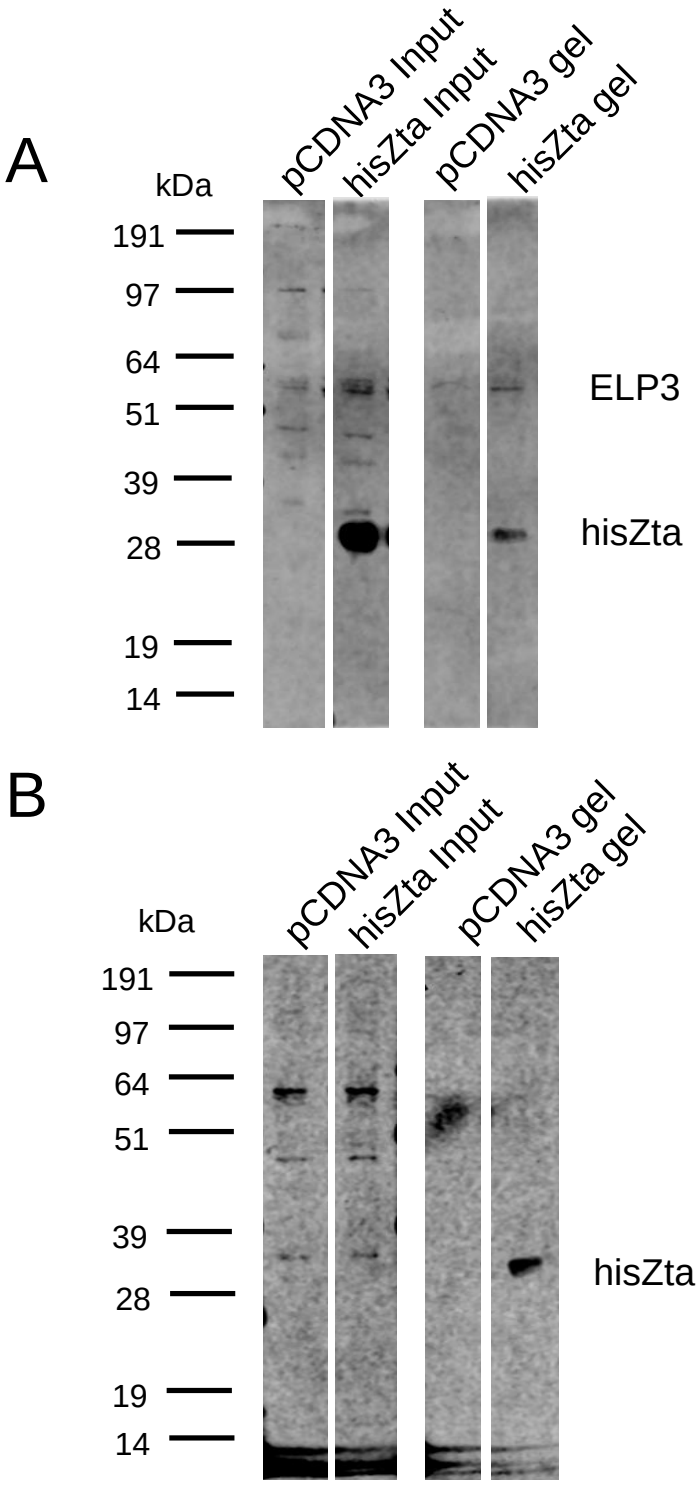


Figure 3.25 Histidine-tag Zta pull down for ELP3 using transfected EBV negative and EBV positive cell extracts. Cells were transfected control or Zta vector. After 72/96 hours cells were lysed in cell lytic reagent. 100µl of supernatant was added to 25µl of HIS-select gel (Sigma) and bound at room temperature (RT). After washing, all samples were then mixed 1:1 with protein sample buffer and run on a protein gel. This was western blotted with BZ1 for Zta and ELP3 antibodies. **A** U2OS cells **B** HEK293-BZLF1-KO cells

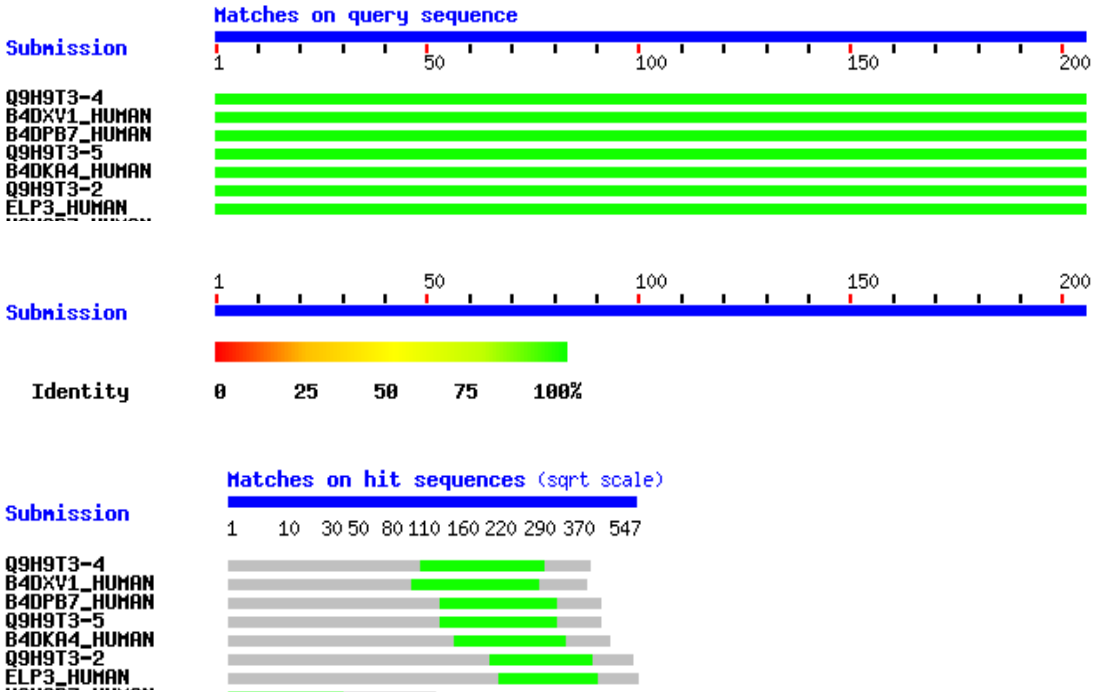


Figure 3.26 BLASTP seacrh using UniProtKB database against the ELP3 epitope recognised by the antibody. The amino acid sequence corresponding to a region within amino acids 240 - 445 of Human Elp3 (NP_060561) was used in a BLASTP search against the homo sapien UniProtKB database (release 2015_08 of 22-Jul-2015). Seven proteins were returned that had a match on the amino acid epitope sequence recognised by the antibody.

Accession	Description	kDa	Length (aa)
Q9H9T3-4	Isoform 3 of Elongator complex protein 3	49	428
B4DXV1	cDNA FLJ58642, highly similar to Homo sapien ELP3	48	418
B4DPB7	cDNA FLJ58601, highly similar to Homo sapien ELP3	52	455
Q9H9T3-5	Isoform 4 of Elongator complex protein 3	52	455
B4DKA4	Elongator complex protein 3	54	475
Q9H9T3-2	Isoform 2 of Elongator complex protein 3	61	533
Q9H9T3	Elongator complex protein 3	62	547

Table 3.7 Proteins identified from the BLASTP search of ELP3 amino acid sequence (240aa – 445aa) epitope

The band identified between 64kDa and 51kDa (Figure 3.25A and B) may be one of 5 proteins with the identical epitope recognised by the antibody (Table 3.7). This is most likely to be ELP3 (Q9H9T3), the protein identified by the antibody at 62kDa. This band is only present in the Zta pull down in U2OS cells (Figure 3.25A), this may indicate an interaction with Zta but this is not present in HEK293-*BZLF1*-KO cells. The band between 51kDa and 39kDa (Figure 3.25A, Figure 3.25B) may correspond to the proteins with 49KDa or 48KDa (Table 3.7).

A hisZta pull down was performed to investigate if there is an interaction between hisZta for ZNF285 using transfected U2OS extracts. ZNF285 had a SILAC ratio of 100.0. Therefore there should be only in the Zta pull down sample. The predicted band size is 68kDa for ZNF285.

Three distinct bands were identified using the ZNF285 antibody, as well as other non-specific bands in the control and hisZta input (Figure 3.27) Although hisZta binds to the nickel affinity gel, there are no bands recognised by the ZNF285 antibody (Figure 3.27).

The amino acids 288-337 of ZNF285 contain the epitope of where the antibody recognizes. The BLASTP search returned five proteins with an exact match for the amino acid sequence (Figure 3.28). One of these proteins was the reported protein recognized by the antibody (ZNF285_HUMAN). The other proteins returned include fragments and also isoforms of varying sequence length and molecular weights. Again, these proteins were collated with their amino acid sequence length, as found in Figure 3.27 using the UniProt database for the proteins identified from the BLASTP search (Table 3.8).

A ZNF285 isoform Q96NJ3-2 has a molecular weight of 50kDa. Another ZNF285 isoform Q6B0D6 has a molecular weight of 51kDa. One of these proteins may be the protein band observed by the antibody between 51kDa and 39kDa (Figure 3.27).

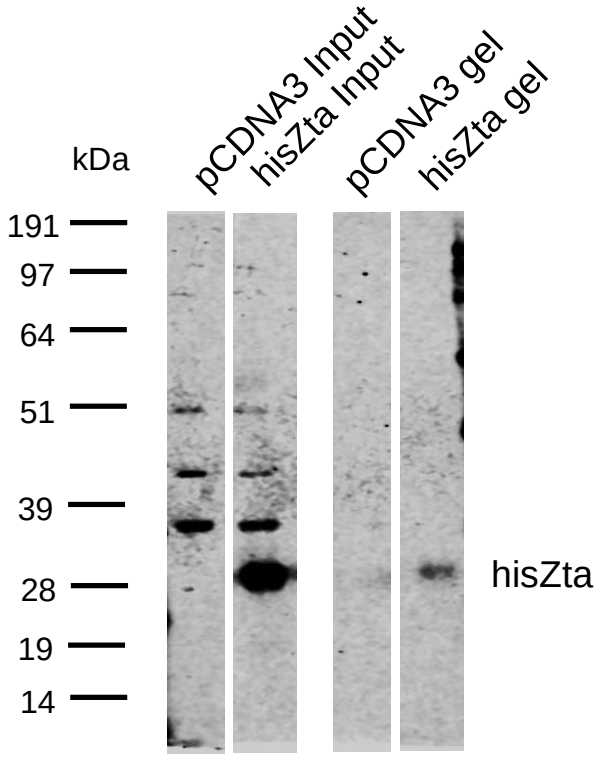


Figure 3.27 Histidine-tag Zta pull down for ZNF285 using transfected EBV negative cell extracts. U2OS cells were transfected control or Zta vector. After 72 hours cells were lysed in cell lytic reagent. 100µl of supernatant was added to 25µl of HIS-select gel (Sigma) and bound at room temperature (RT). After washing, all samples were then mixed 1:1 with protein sample buffer and run on a protein gel. This was western blotted with BZ1 for Zta and ZNF285 antibodies.

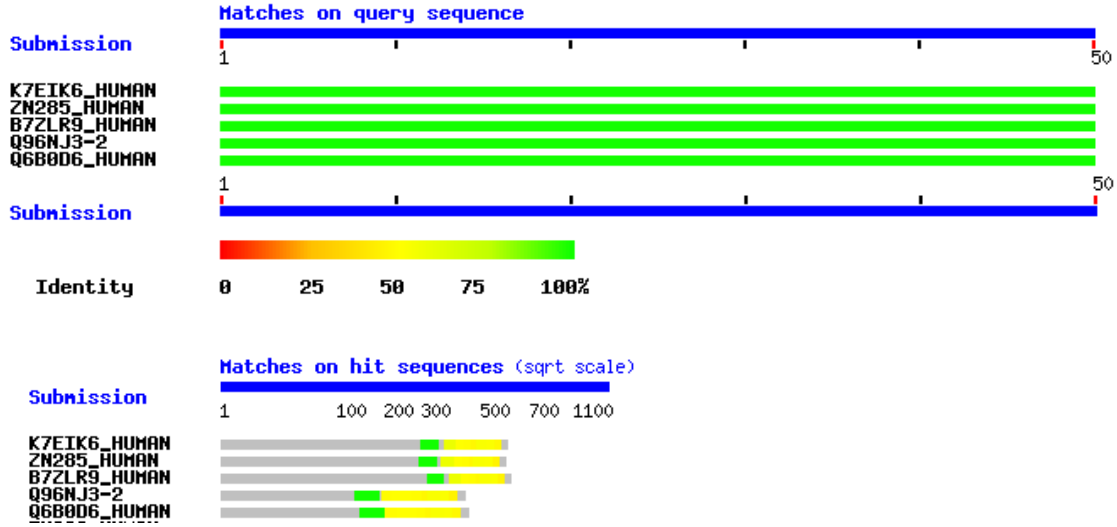


Figure 3.28 BLASTP search using UniProtKB database against the ZNF285 epitope recognised by the antibody. The region corresponding to internal sequence amino acids 288-337 of Human ZNF285 (NP_689567) was used in a BLASTP search against the homo sapien UniProtKB database (release 2015_08 of 22-Jul-2015). Five proteins were returned that had a match on the amino acid epitope sequence recognised by the antibody.

Accession	Description	kDa	Length (aa)
K7EIK6	Zinc finger protein 285	69	597
Q96NJ3	Zinc finger protein 285	68	590
B7ZLR9	ZNF285A protein (Fragment)	71	614
Q96NJ3-2	Isoform 2 of Zinc finger protein 285	50	435
Q6B0D6	ZNF285A protein (Fragment)	51	446

Table 3.8 Proteins identified from the BLASTP search of ZNF285 amino acid sequence (288aa - 337aa) epitope

3.3. Discussion

The first obstacle was to extract hisZta from transfected cells successfully for further application and investigation. Initially a cell lytic reagent was compared against a nuclear extract protocol. There was difficulty in extracting Zta from cells while known cellular partners of Zta could be extracted into the supernatant. The addition and optimization of benzonase to the cell lysis reagent allowed the extraction of Zta to be released into the supernatant. Therefore the ability to extract Zta into a suitable buffer for downstream applications was identified. The cell lytic reagent supernatant that was created was compatible with the nickel affinity gel, and this allowed the establishment of binding of the histidine tagged protein from the cell lysate

After the successful extraction of Zta from transfected cells it was determined whether hisZta could bind to the nickel affinity gel (Figure 3.8). This attempt to bind hisZta to an affinity gel in its native form was accomplished. Histidine has a high affinity for immobilized metal ions including nickel. Poly-histidine tagged Zta was demonstrated to bind to the gel using the incubation period and wash buffers, and the investigation of non-specific binding was important for the determination of interacting partners (Figure 3.10). The addition of imidazole minimized actin binding to the gel as imidazole also prevents non-specific binding of proteins to the affinity gel (Porath 1992).

The NaCl concentration was investigated as salt concentration affects the ionic strength and the binding of non-specific proteins to the affinity gel by minimizing hydrophobic interactions with the gel matrix (Bornhorst & Falke 2000). The conditions to find the optimal binding for Zta was established with the use of 300mM of NaCl and 20mM of imidazole in the wash buffer (Figure 3.11). These conditions were taken forward for hisZta to bind to the nickel affinity gel and non-specific binding of proteins minimized.

An attempt to identify novel interacting partners of Zta was performed using SILAC affinity purification. EBV negative cells were labelled with SILAC light media (R0K0) and SILAC medium media (R6K4) before being transfected with

control and hisZta expression vectors respectively. Protein extracts of these cells were produced from the extraction method outlined and the extracts were applied to the nickel affinity gel under the binding conditions established. The gel bound proteins were boiled from the gel and sent for mass spectrometry at the University of Bristol to allow a proteomics analysis of possible protein interactions. Comparing the control extract pull down to the Zta extract pull down allowed the quantitation of proteins identified from the pull down. The data returned from the University of Bristol Proteomics facility indicated an abundance of proteins with potential interactions with hisZta. These proteins were filtered against the user submitted database Protein Frequency Library (PFL) (Boulon et al. 2010). This filtered non-specific proteins that have been shown to bind to affinity matrices and remove contaminants. This allows SILAC pull down assays to identify lower abundant or less tightly bound protein complexes that may be overlooked in the background of non-specific binding.

This allowed a quantification of host proteins that had a high abundance in the Zta pull down sample. The large dataset returned from the mass spectrometry analysis did display many proteins of interest that may have potential interactions with Zta. These proteins had an increased SILAC ratio in the Zta pull down; Zta may be interacting with these proteins. 41 proteins were only identified in the Zta pull down sample which included Zta itself. 197 proteins had a SILAC ratio above 1.2, which indicates that Zta is forming an interaction or is part of a larger complex. Both sets of these higher quantitated proteins were considered for further investigation, taking into account their SILAC ratio, PSM number, and medium to light peptide count. Proteins with interesting functional properties were also considered within these categories.

Zta interacts with a number of cellular proteins, and these interactions have been published. The dataset was searched for these published interactions and only a few of these proteins were identified within the dataset. They were not abundant in the Zta pull down. DNA topoisomerase I and RecQL1 can associate with Zta (Wang 2009) but DNA topoisomerase I and RecQL1 had SILAC ratios of 0.924 and 0.660 respectively. The proteomic pull down assay here could not indicate this interaction.

The proteins of interest that remained in the data were investigated in further pull down experiments using specific antibodies. The proteins chosen for further study were FANCA, BRD4, ELP3, ELK4 and ZNF285. Repeated pull downs and western blots were attempted to confirm these possible protein interactions from Zta transfections of either EBV negative cells (U2OS) or EBV positive cells (HEK293-*BZLF1*-KO).

FANCA is a DNA repair protein. There have not been studies on FANCA and the contribution to EBV lytic cycle. FANCA also is known to interact directly with RNF8. RNF8 is an ubiquitin-binding protein, and FAAP20 links RNF8-mediated ubiquitination to the Fanconi anemia DNA repair network (Yan et al. 2012). A recent publication indicates RNF8 colocalises to EBV replication compartments (Yang et al. 2015). It may be possible that a protein complex containing FANCA is present in EBV lytic replication compartments. The Fanconi anemia pathway is linked to the activation of the ATM pathway (Yamamoto et al. 2008). Fanconi anemia (FA) pathway maintains genomic stability, and ATM and FA genes function in parallel for genome stability (Kennedy et al. 2007). ATM has also been induced by Zta alone, prior to EBV replication (Wang'ondu et al. 2015). DNA damage response (DDR) activation markers were induced even in the absence of EBV lytic replication compartments and γH2AX induction was necessary for optimal expression of early EBV genes. This suggests that markers of DDR are activated in this microenvironment of viral gene expression (Wang'ondu et al. 2015). It may be possible that Zta interacts with FANCA and recruits this factor for viral genome replication or viral gene expression. A hisZta pull down for the interaction of FANCA returned an observed band in the hisZta sample pull down, not seen in the control pull down (Figure 3.19B). The molecular weight was between 64kDa and 51kDa. This band was not at the expected molecular weight of FANCA, therefore the isoforms and variants of the target proteins were investigated.

BRD4 is a chromatin reader that can recognise acetylated histones. BRD4 has been investigated as being important for DNA virus replication. BRD4 plays a key role in Merkel cell polyomavirus DNA replication (Wang et al. 2012). Recruitment

of BRD4 to the human papillomavirus type 16 DNA replication complex is essential for replication of viral DNA (Wang et al. 2013). It has not been shown if BRD4 plays a role in EBV lytic cycle. EBV can already interact and utilize BRD4 in latency. The EBNA1 protein of Epstein-Barr virus functionally interacts with BRD4; BRD4 localises to the FR enhancer regulated by EBNA1 and aids the transcription of other EBV latency genes (Lin et al. 2008). Also KSHV latency-associated nuclear antigen 1 (LANA-1) plays a role in G1 cell cycle arrest and interacts directly with BRD4 (Ottinger et al. 2006) (Ottinger 2006). BRD4 is also recruited to latent Cp, and is required for EBNA2 –activated transcription (Palermo et al. 2011)

BRD4 reproducibly observed two higher molecular weight bands and the upper band was between 191kDa and 97kDa, the molecular weight range for BRD4 (152kDa). There were bands in the control pull down sample; therefore this protein could not be indicated as an interaction with Zta by pull down analysis.
As seen in Figure 3.19 for FANCA and Figure 3.21 for BRD4, the protein bands observed by the antibody were in high abundance in the control pull down gel. The upper molecular weight band observed using the BRD4 antibody in Figure 3.21A was in higher abundance in the hisZta pull down sample than the control sample. This may point towards Zta being the factor for a greater abundance of BRD4 in this sample.

ELP3 is a catalytic histone acetyltransferase subunit elongator protein and this subunit was purified and characterized as part of the human elongator complex (Hawkes et al. 2002). The subunit aids transcription by directly interacting with RNA polymerase II. ELP3 displays an observed band between 64kDa and 51kDa (Figure 3.25A and B), ELP3 molecular weight is 62kDa. There was a band observed only in the EBV negative Zta pull down sample (Figure 3.25A), but this could not be confirmed in the EBV positive Zta pull down sample (Figure 3.25B).

ZNF285 may play a role in transcriptional activation. A zinc finger protein ZBP89 is essential for lytic replication (Baumann et al. 1999). ZNF285 could not be confirmed in this study. ELK4 is a member of the Ets family of transcription factors. This accessory protein has been shown to bind to c-fos serum response

elements (SREs), recruited by serum response factor (Dalton 1992 Cell). ELK4 could not be confirmed in this study. The pull downs and western blot investigation of ELK4 returned inconclusive results. The protein observed by the antibody was not at the expected molecular weight, and the isoforms that contained the same epitope region recognized by the antibody did not match this molecular weight either.

The antibodies utilized to observe the proteins of interest detected multiple bands within the input samples. For FANCA, ELK4 and ZNF285 the protein bands observed were not at the molecular weight expected. The epitope of the antibodies were explored.

BLASTP searches of the epitope that the antibody recognizes allowed a comparison of proteins sharing this same epitope. Most of these proteins were isoforms and variants at varying molecular weights. Some of the protein bands that the antibody observed were at the molecular weights of some isoforms. From the UniProt BLAST search these included BRD4 isoforms (Table 3.6), may be recognized between 97kDa and 64kDa in Figure 3.21. For ELP3 bands observed, the band between 51kDa and 39kDa (Figure 3.25A, Figure 3.25B) may correspond to the proteins with 49KDa or 48KDa (Table 3.5).

A major issue that was apparent through many pull down attempts was the protein of interest would have a non-specific attraction for the nickel affinity gel. Although BRD4 and ELP3 did include bands observed by the antibody that were of the expected molecular weights.

The limitations of this study included the observation of multiple bands and non-specific binding of the proteins of interest to the nickel affinity gel. Although some protein interactions show promise, these could not be shown in confidence. The majority of western blots repeats were inconclusive to assess if the interactions were true. Zta would also bind to the nickel affinity gel but was inconsistent between assays. The particular antibodies would detect bands at varying molecular weights for FANCA, BRD4, ELP3, ELK4 and ZNF285.

It was decided that to further investigate novel protein interactions then another method would be attempted in Chapter 4. In order to indicate an interaction between Zta and the cellular proteins the issue of non-specific protein interactions with the nickel agarose gel needs to be addressed. The variability of the histidine tagged Zta binding to the nickel affinity column and sensitivity of detection was reflected upon. An adaptation of the pull down method was also considered.

Established ChIP conditions allow cross-linking of cellular proteins and DNA before being precipitated. This approach was explored and adapted under denaturing conditions for further investigation of novel proteins that may interact with Zta.

4. Identify novel interacting partners of Zta using denaturing conditions

4.1. Introduction

The question addressed here was whether novel interaction partners could be identified in EBV positive cells using an alternative approach to Chapter 3. Histidine tagged Zta was demonstrated to bind to the nickel affinity gel in Chapter 3, and although some potential interactions were investigated the overall binding efficiency of Zta to the gel and non-specific binding was inconsistent. Repeated attempts at assessing these protein interactions with Zta were unsuccessful although there were indications of these interactions from the proteins examined.

Therefore an alternative investigation into protein-protein interactions was applied. Conventional chromatin immunoprecipitation protocols include cross-linking of proteins to proteins and DNA. Sonication of cross-linked extracts allows DNA to be fragmented and sheared into pieces. The proteins in solution remain bound to DNA and interacting proteins (Collas 2010).

As the histidine tag located on the N-terminus of Zta may be buried within the structure and so impair binding of hisZta to the nickel affinity gel. By denaturing cell extracts we propose that the histidine tag would become freely exposed and so would be able to bind to the nickel affinity gel (Figure 4.1).

A denaturing method combined with cross-linking of proteins to preserve immediate interactions was optimized. The use of formaldehyde, urea, and sonication was investigated to extract of proteins from transfected cells. These extracts were optimized for binding to the nickel affinity gel in a denatured form.

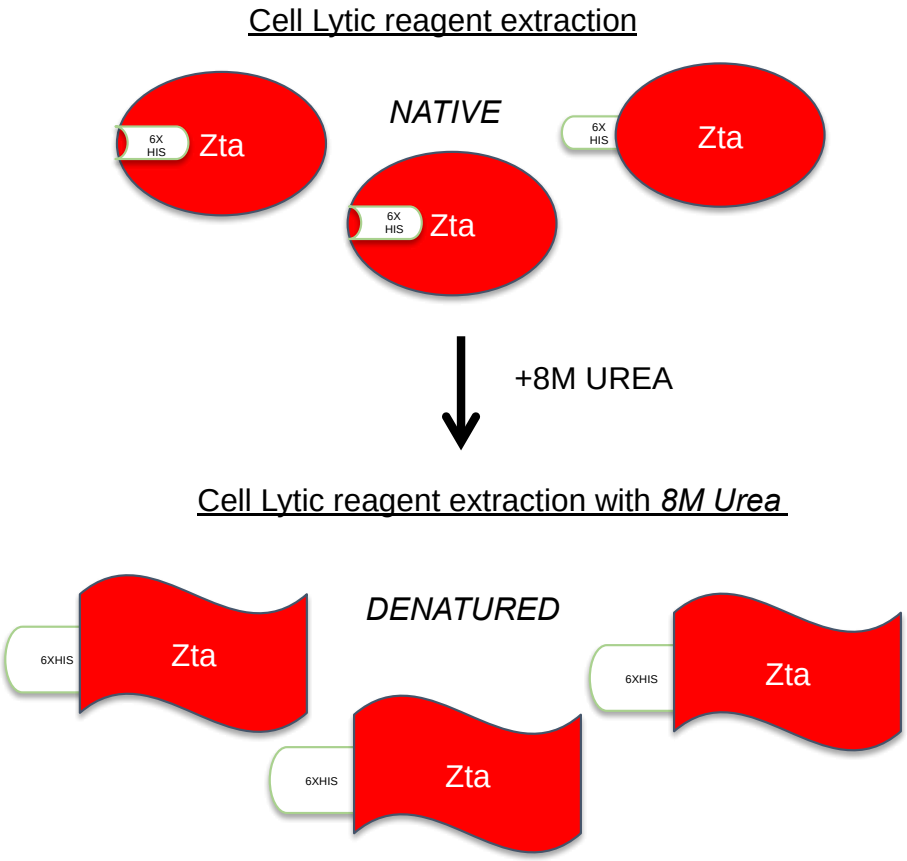


Figure 4.1 hisZta Histidine-tag buried. It is possible to extract hisZta from transfected cells with cell lytic reagent (Sigma). The binding efficiency to nickel affinity agarose beads may not be efficient between pull downs as the 6-histidine tag may be buried or not exposed. This may impair binding of hisZta to the agarose gel. With the addition of 8M urea to the cell lytic reagent, the protein extract may be denatured and all proteins including hisZta will be unfolded, exposing the histidine-tag.

4.2. Results

4.2.1. Establishing denaturing conditions to extract Zta and interacting partners

The protein hisZta transfected and expressed in HEK293-*BZLF1*-KO cells will initiate the lytic cycle from latency and form transcription and replication complexes. It has been determined that transfected hisZta and other cellular proteins can be extracted in native form from cells using the cell lytic reagent supplemented with benzonase. However, the binding efficiency to nickel affinity gel may not be efficient between pull downs. The histidine tag may be buried or not be as exposed each time between extractions and this may impair any binding of hisZta to the agarose gel. By denaturing cell extracts then the histidine tag may become freely exposed and be able to bind to the nickel affinity gel.

A schematic diagram represents how the histidine tag may be masked by Zta itself or by other proteins while in a native complex (Figure 4.1). Denaturing the protein extract leads to the loss of the tertiary structure thus exposing the poly-histidine tag to bind to nickel affinity gel.

An attempt to denature HEK293-*BZLF1*-KO cell extracts transfected with hisZta was performed. The first aim was to identify if the histidine tagged protein could bind to the nickel affinity gel in a denatured form. The cell extracts were lysed in the cell lytic reagent with benzonase as previously described but the reagent was supplemented with 8M urea. The denatured extract was applied to the nickel affinity gel and incubated together. Each gel was then washed with the wash buffer, containing different concentrations of imidazole between 0 and 50mM and the bound proteins were separated by SDS-PAGE (Figure 4.2). The input was investigated alongside the gel samples that included the transfected control gel and transfected hisZta gel.

HisZta was successfully extracted using the cell lytic reagent containing 8M urea, as seen in the input. HisZta bound efficiently. We note that actin binds to the gel in both the control and hisZta samples here. Previously, it was determined that the use of imidazole reduces non-specific binding of proteins to the nickel gel. Using 10mM imidazole in the wash buffer removes the non-specific actin binding and the other band recognised by the BZ1 antibody in the hisZta sample. 20mM imidazole in the wash buffer sees a slightly reduced amount of hisZta bound to the gel compared to 10mM imidazole wash. Using 50mM imidazole in the washes minimised hisZta binding to the gel, this imidazole concentration may compete and elute the protein from the gel. It was determined that hisZta protein binds to the nickel affinity gel in a denatured form, while the use of a low concentration of imidazole (10mM) in the gel washes minimises non-specific binding of proteins as seen by the absence of actin.

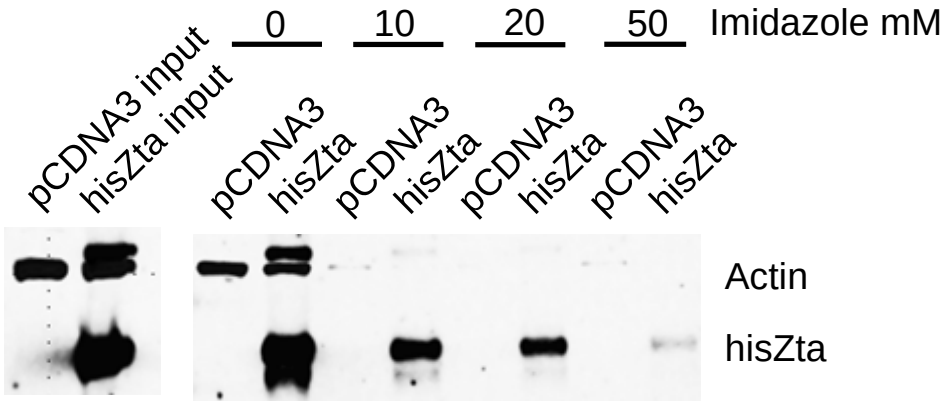


Figure 4.2 Histidine-tag Zta pull down with denatured extract. HEK293-BZLF1-KO cells were transfected with hisZta. After 96 hours cells were harvested and denatured extract produced from cell lytic reagent containing 8M urea. 100µl of extract was applied to 20µl of agarose per imidazole condition and incubated at room temperature. After varying imidazole concentration in the washing, agarose was mixed 1:1 with protein sample buffer and run on a protein gel. The proteins were separated by SDS-PAGE and a western blot for Zta and actin performed. The antibodies used were BZ1 and anti-actin.

4.2.2. Establishing crosslinking and pull down conditions for denatured Zta

By denaturing cell extracts, all of the protein in the sample will lose its tertiary structure and therefore lose the function of protein domains. Therefore protein complexes will fall apart as proteins will lose any possible interactions with each other. Taking from established ChIP assays where proteins and DNA are cross-linked, we considered crosslinking proteins within cells before creating denatured extracts. This enables the preservation of any interactions between proteins through formaldehyde cross-linking.

A schematic diagram of a cross-linked denatured cell extract pull down is illustrated (Figure 4.3). HEK293-*BZLF1*-KO would be transfected with control or hisZta and then cross-linked. Cell extracts would be prepared and denatured. The extract applied to the nickel affinity gel to bind hisZta and cross-linked interacting proteins. Then hisZta and associating proteins would then be eluted from the beads for further analysis

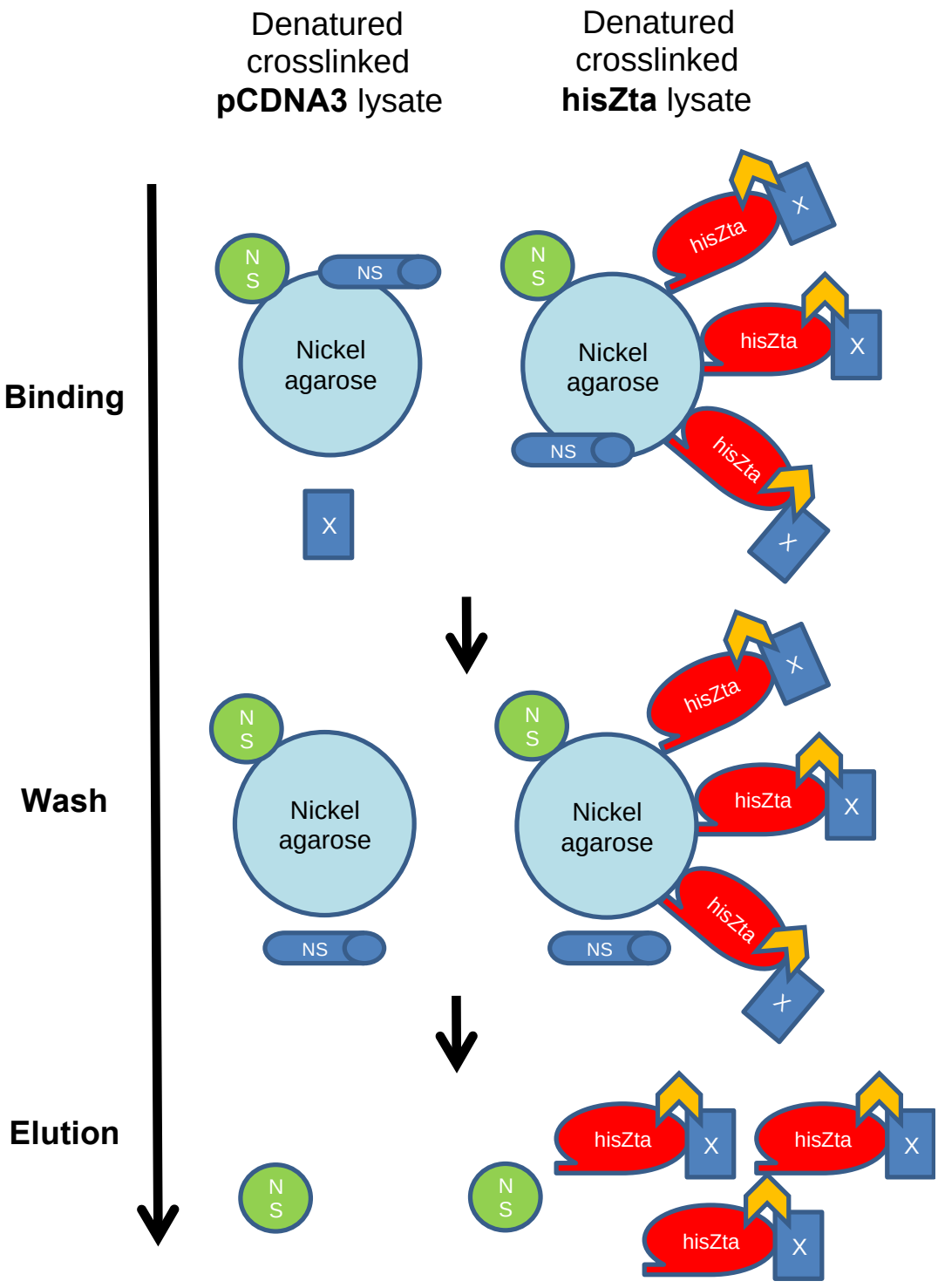


Figure 4.3 Schematic of denatured cross-linked extract pull down. HEK293-BZLF1-KO cells would be transfected with empty vector or hisZta, cross-linked, harvested after 96 hours and then denatured extract prepared. The histidine-tagged Zta would be cross-linked to interacting proteins and able to freely bind the nickel agarose gel. The non-specific proteins can be washed away from the agarose. Elution of the nickel bound hisZta will pull down cross-linked proteins for further analysis. NS nonspecific. X interacting protein

The cross-linking conditions were adjusted to determine what percentage of the cross-linking reagent would be optimal. HEK293-*BZLF1*-KO cells were transfected with hisZta and harvested after 96 hours to initiate EBV lytic replication. The cells were cross-linked with 0%, 0.1% or 1% formaldehyde before being harvested. The cells were then lysed using the native extract, or denatured with cell lytic reagent and 8M urea before protein extract collected. The cross-linked extracts were applied to nickel affinity gel and attempted to bind hisZta in native form or denatured form as a comparison. The input and gel samples were separated by SDS-PAGE (Figure 4.4).

Without cross-linking the cells, hisZta can bind to the nickel affinity gel as previously demonstrated in Chapter 3, displaying a similar binding profile. The native cross-linked extracts were not extracted efficiently here, implying that that the crosslinking of proteins impairs extraction. Therefore the amount of hisZta that was free to bind the gel was minimal, demonstrated by a lack of binding to the nickel affinity gel, possibly by being buried in a complex.

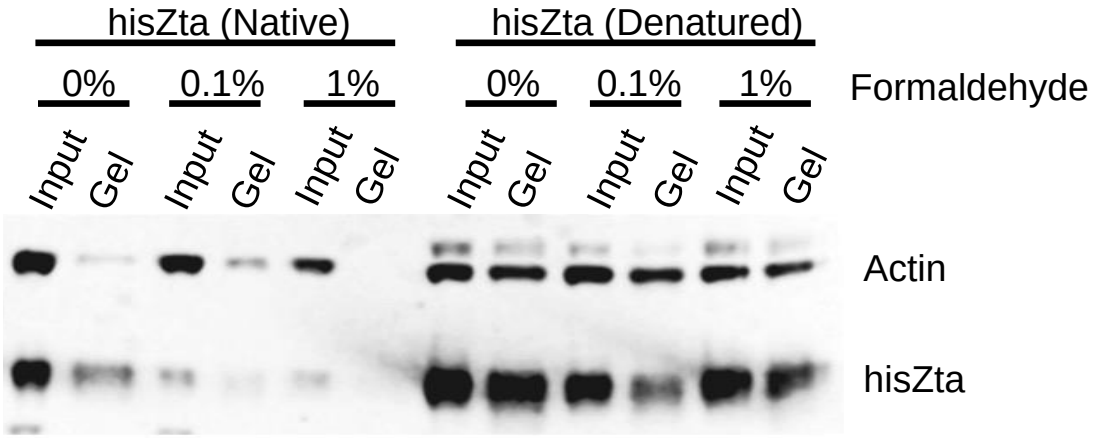


Figure 4.4 Native vs denatured pull down with varying formaldehyde percentage crosslinking. HEK293-BZLF1-KO cells were transfected with hisZta. After 96 hours, cells were treated with varying formaldehyde percentages of 0%, 0.1% and 1%. The cells were harvested and extract produced either by normal cell lytic reagent conditions or cell lytic reagent containing 8M urea, to produce native or denatured conditions. Xµl of extract was applied to 25µl of agarose and incubated at room temperature. After washing, agarose was mixed 1:1 with protein sample buffer and run on a protein gel. The proteins were separated by SDS-PAGE and a western blot for Zta and actin performed. The antibody used here was BZ1 and anti-actin

By denaturing the extracts, hisZta is readily extracted into the supernatant, as seen in the inputs. It is clear that denaturing the extracts allows the release of more Zta into supernatant. The protein binds to the nickel affinity gel without any crosslinking, and hisZta binds to the gel more readily at 1% cross-linking than 0.1% cross-linking. It was decided that 1% cross-linking conditions would be used for denaturing pull down assays.

Cross-linking cells using 1% formaldehyde and denaturing the extracts were taken forward and an optimal elution of hisZta began to be attempted. The cross-linked denatured cell extract of HEK293-*BZLF1*-KO transfected with hisZta was applied to nickel affinity gel and an elution was attempted (Figure 4.5A). The previous conditions for the pull down were repeated and the bound protein on the gel was then washed with 10mM imidazole in the wash buffer. After washing the gel, the gel was split into two. One half was mixed 1:1 with 2X sample buffer; the other half was subjected to elution with elution buffer containing 250mM imidazole. This was performed to try to elute hisZta and interacting proteins from the gel using the elution buffer containing a differing pH supplemented with this concentration of imidazole. The elution fraction could then be concentrated for further analysis, including comparing bound protein against eluted protein

HisZta was extracted from cells as shown previously. This extract bound to the nickel affinity gel as shown by analysis of half of the gel (Figure 4.5B). After an elution attempt by washing the gel with elution buffer containing 250mM imidazole, some of the bound hisZta protein was removed from the gel compared to the boiled gel fraction. The eluted hisZta was not present here in the elution fraction.

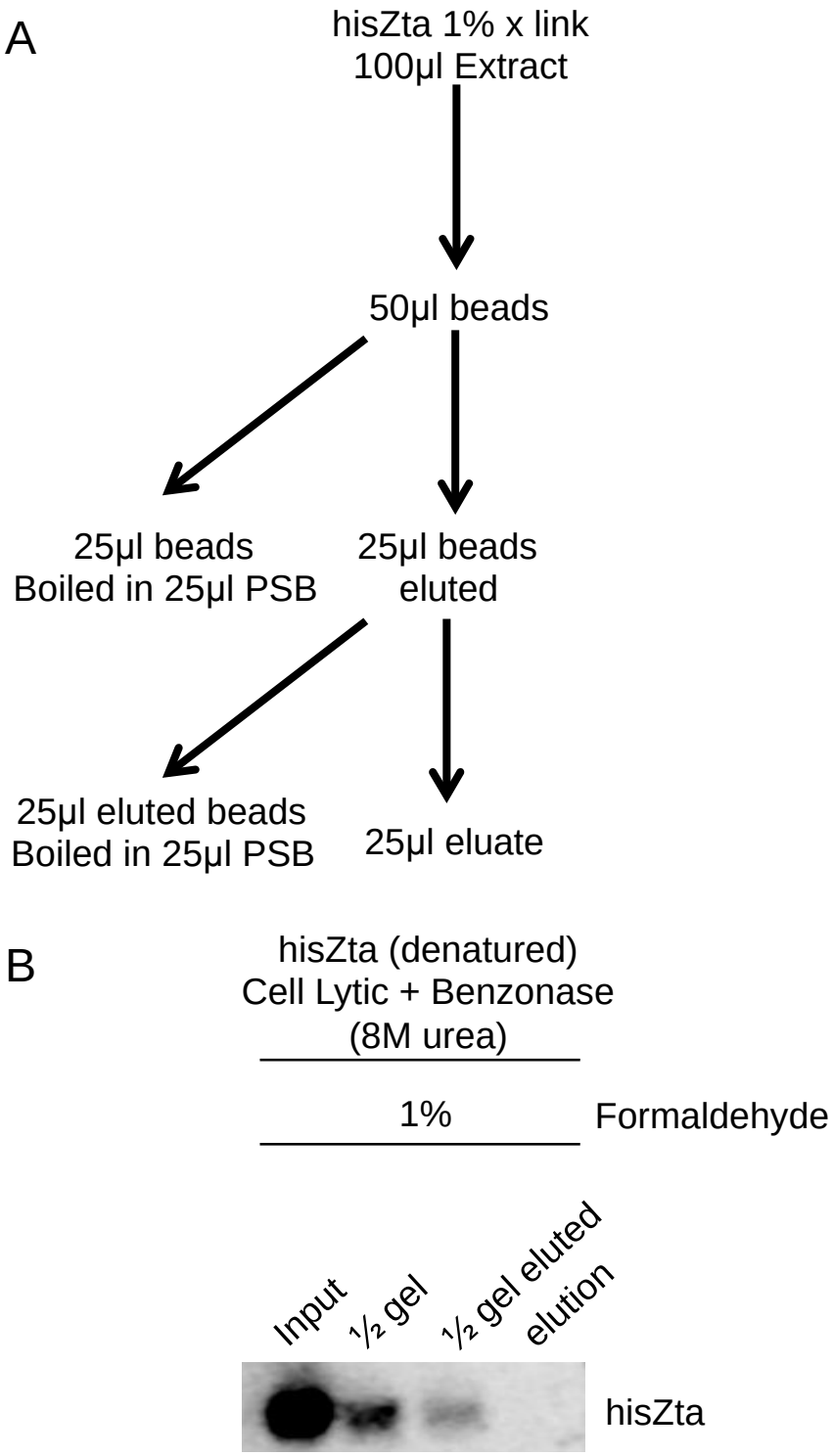
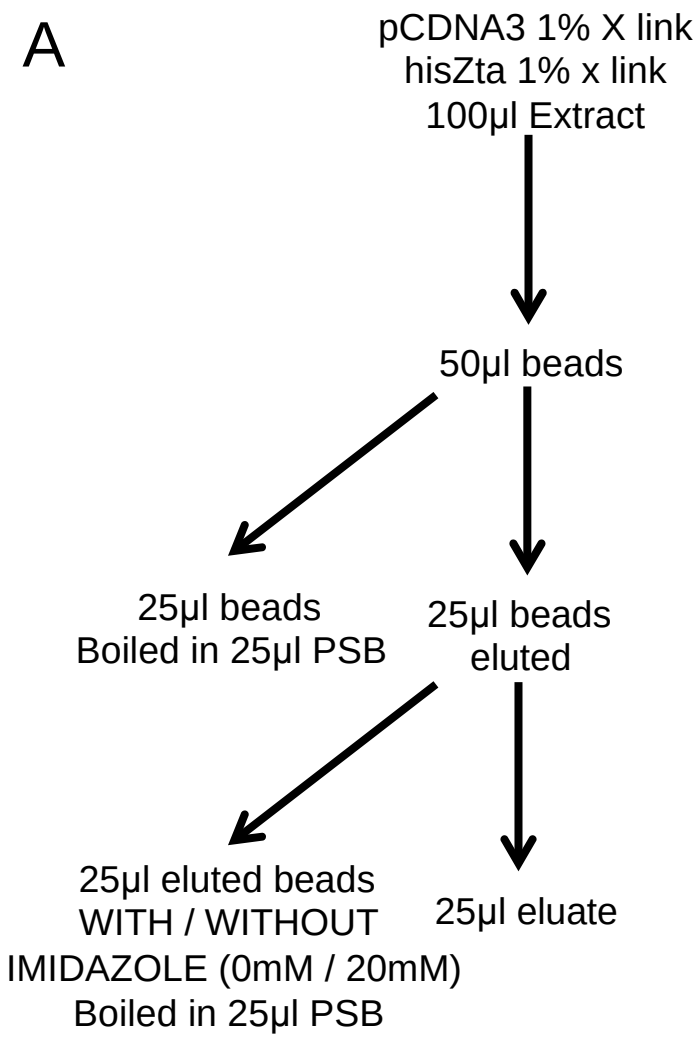


Figure 4.5 Cross-linked denatured pull down of histidine-tag Zta. HEK293-BZLF1-KO cells were transfected with hisZta. After 96 hours, cells were treated with 1% formaldehyde. **A** The cells were harvested and extract produced with cell lytic reagent containing 8M urea, to produce extract under denatured conditions. Extract was applied to 25µl of agarose and incubated at room temperature. Affinity gel was split into 2 fractions, one for gel analysis the other elution with one round with 250mM imidazole elution. **B** After washing, elution and gel fractions was mixed 1:1 with protein sample buffer and run on a protein gel. The proteins were separated by SDS-PAGE and a western blot for Zta performed. The antibody used here was BZ1.

After multiple attempts to repeat this method and elute hisZta from the nickel affinity gel, the protocol developed was reviewed. Throughout the binding procedure of histidine-tagged proteins to nickel affinity gels, imidazole is recommended throughout in order to minimize any non-specific binding to the gels and to enable the histidine-tagged protein to bind efficiently. Imidazole has been used in the lysate, wash buffers and elution buffers. It may be possible that imidazole in the elution buffer in a denatured pull down may be detrimental and interfere with the elution of hisZta from the nickel affinity gel in a denaturing assay.

It was investigated to see if imidazole has a disadvantageous effect when included in the wash and elution buffers when an elution from the gel is performed. A denatured pull down assay was performed using cross-linked control or hisZta transfected HEK293-*BZLF1*-KO cells, with or without 10mM in the elution buffer. One gel set was to be eluted without imidazole in the elution buffer, versus the other set being eluted with imidazole in the elution buffer. This is illustrated and explained as a flow diagram (Figure 4.6A). The pull down was performed using control or hisZta denatured extracts without imidazole in the elution buffer. The affinity gel was split into two groups for analysis and elution.


A
pCDNA3 1% X link
hisZta 1% x link
100µl Extract
50µl beads
25µl beads
Boiled in 25µl PSB
25µl beads
eluted
25µl eluted beads
WITH / WITHOUT
IMIDAZOLE (0mM / 20mM)
Boiled in 25µl PSB
25µl eluate

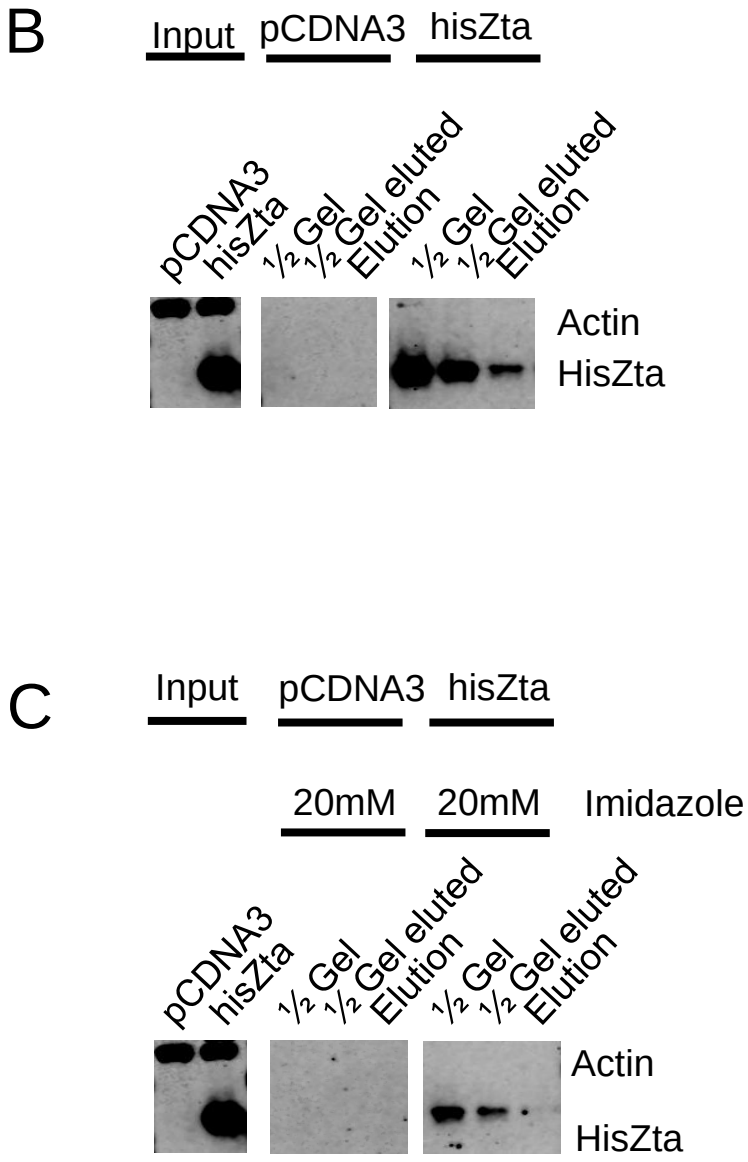


Figure 4.6 Schematic of pull down with and without imidazole (20mM) in the elution buffer. **A** Crosslinked cell extracts to be applied to the gel. The gel is then split in half where one half is mixed equally with sample buffer to assess binding. The other half of the gel is eluted with elution buffer with or without 20mM imidazole in the elution buffer. **B** Pull down with 0mM imidazole in the elution buffer **C** Pull down with 20mM imidazole in the elution buffer. Both the proteins were separated by SDS-PAGE and a western blot for Zta performed. The antibody used here was BZ1.

The pull down performed using control or hisZta denatured extracts with imidazole in the elution buffer was repeated the same way. The affinity gel was split into two groups for analysis and elution (Figure 4.6B, Figure 4.6C). The hisZta protein binds to the gel here clearly as demonstrated by the half gel sample that did not have any elution buffer applied (Figure 4.6B). The other half of the affinity gel was subjected to an elution using elution buffer with imidazole (Figure 4.6B). It can be seen in the elution lane that some hisZta was removed from the affinity gel. The gel that had the elution applied to it shows most hisZta is still bound.

The gel washed with 20mM imidazole bound in Figure 4.6C. The other half of the affinity gel was subjected to an elution using elution buffer with imidazole. It can be seen in the elution lane that there is no hisZta removed from the affinity gel. The gel that had the elution applied to it shows some hisZta is still bound. The elution buffer elutes proteins based on pH. It was decided to continue to attempt optimization of the elution from the gels using the elution buffer without imidazole due to the minimal binding of Zta with imidazole present. The elution were to be optimized without imidazole present.

Another pull down attempt was made to confirm the observation made In Figure 4.6. The denatured, cross-linked, hisZta transfected HEK293-*BZLF1*-KO cells were applied to the nickel affinity gel. An elution attempt on the whole affinity gel was performed (Figure 4.7A). The protein extract was incubated with the gel, the gel was washed with wash buffer and the gel was then eluted in elution buffer for a longer incubation period. The input, gel and elution were mixed 1:1 with 2X sample buffer and separated by SDS-PAGE for western blot analysis.

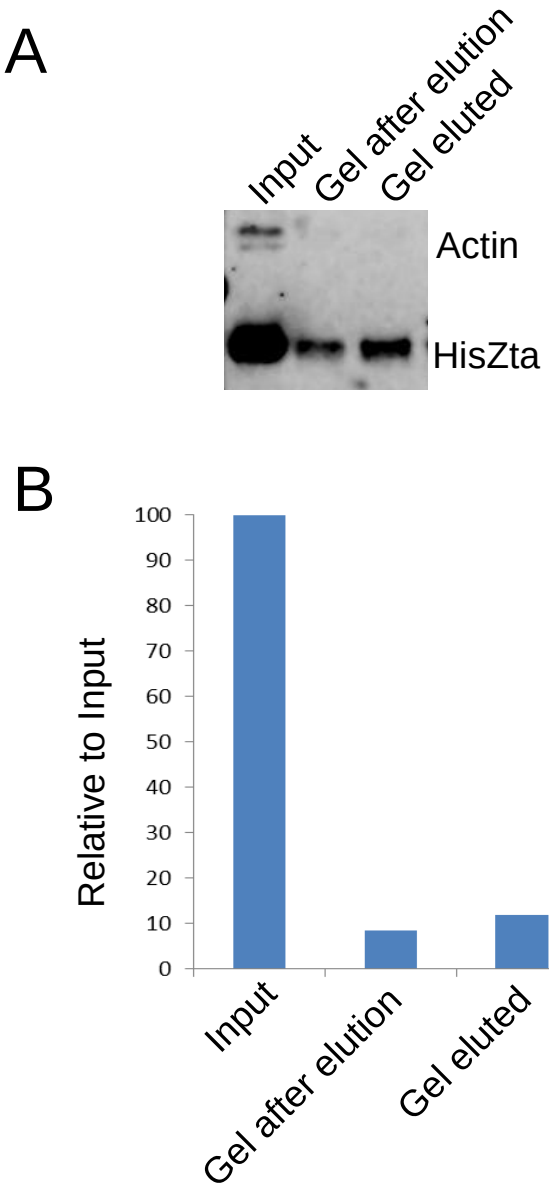


Figure 4.7 Pull down attempts without imidazole in the elution buffer. Cross-linked hisZta extract was applied to Nickel agarose gel after being denatured. The extract was applied to 50µl of beads. This was washed with wash buffer and eluted using elution buffer without imidazole. All samples were mixed with an equal amount protein sample buffer and run on a protein gel for: **A** Western blot analysis using BZ1 antibody for Zta and Anti-actin for actin. **B** Quantitation of the protein bands detected by the BZ1 antibody

The protein hisZta did bind to the nickel affinity gel with very minimal actin binding to the gel. After the single round of elution with elution buffer for a longer incubation period, more hisZta is evident in the elution, implying that the hisZta is closer to all of the protein being eluted and into the elution buffer. This was quantitated and more Zta is evident in the elution when related to input (Figure 4.7B). As this was only one round of elution then it may be possible to apply multiple elutions to the single gel and collect all of the elution fractions together. This should elute the majority of bound hisZta from the gel, in multiple stages.

A pull down attempt was repeated as previous in Figure 4.7, but with five elutions applied to the gel instead of just one. This would investigate if the gel to which denatured hisZta is bound to can be eluted from the gel in succession. By eluting the protein five times this would determine if all of the protein could be removed from the affinity gel. A flow diagram of the protocol performed is illustrated (Figure 4.8A). This demonstrates the multiple elutions steps performed. The western blot after the assay was completed shows that with multiple elutions of the gel allowed some but not all of hisZta to be eluted from the gel (Figure 4.8B). The protein hisZta did elute into the first, second and third elution fractions while very minimal in the fourth and fifth elution attempts. Some hisZta did remain on the affinity gel.

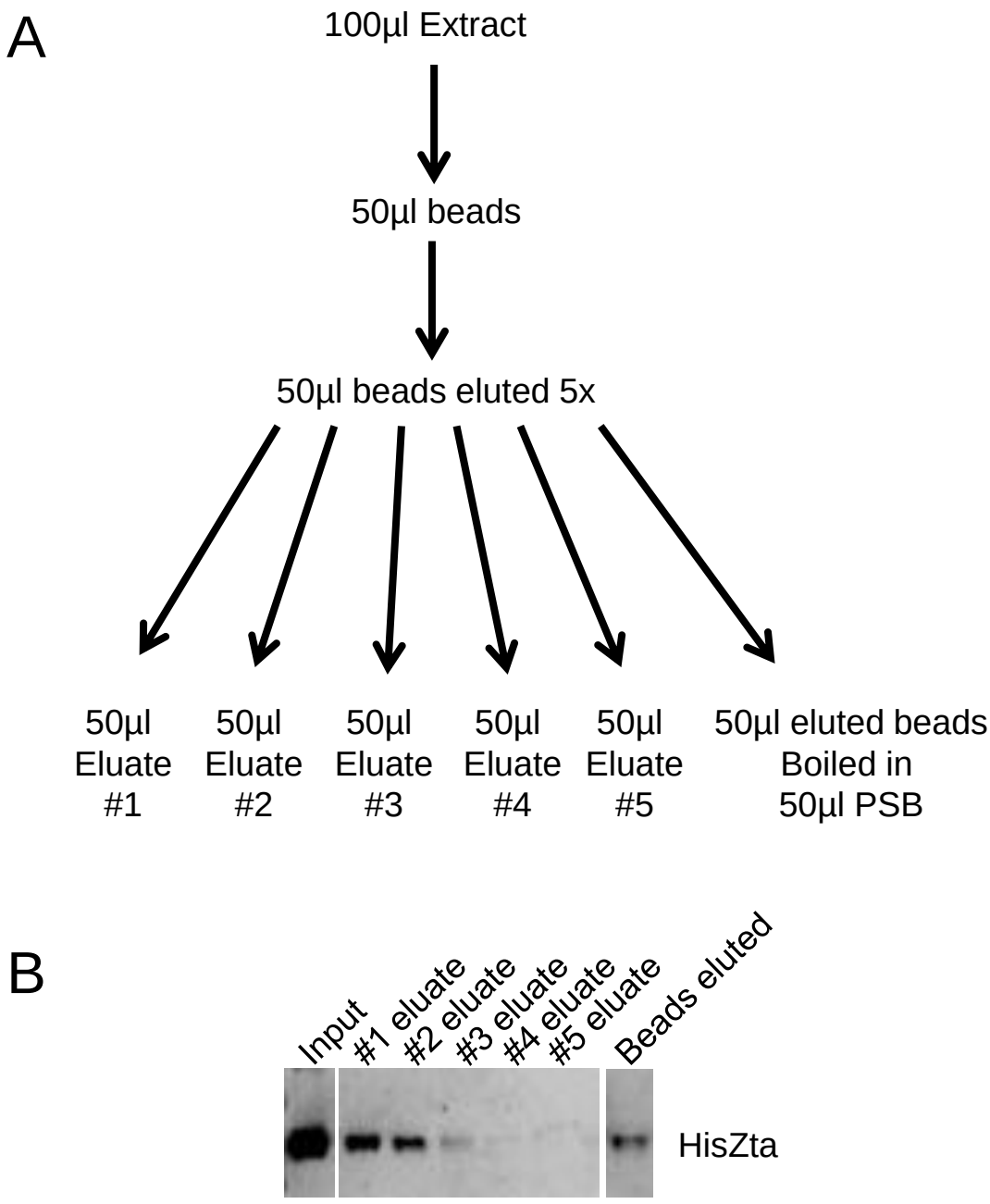


Figure 4.8 Pull down attempt with five elution stages of the gel. **A** Cross-linked hisZta extract was applied to Nickel agarose gel after being denatured. The extract was applied to 50µl of beads. This was washed with wash buffer and eluted using elution buffer without imidazole for five elution steps. This was done to try to elute of Zta from the beads. **B** All samples were mixed equally with protein sample buffer and run on a protein gel for western blot analysis. The proteins were separated by SDS-PAGE and a western blot for Zta performed. The antibody used here was BZ1.

The cumulative protocol includes cross-linking cells before cell lysis. The extracts are denatured in cell lytic reagent including 8M urea. The extract is applied to equilibrated nickel affinity gel, before being washed with wash buffer containing 10mM imidazole. Successive multiple elutions with elution buffer without imidazole demonstrates most of hisZta is eluted from the beads. It was attempted to repeat the pull down assay and combine the multiple elution fractions, this would combine the eluted hisZta into one sample and protein precipitation can be performed to concentrate this eluted protein sample.

As the cross-linking procedure applied to cells before harvesting links most of the proteins and DNA in close proximity together, the protein-protein interactions would be strengthened as well as protein-DNA interactions. The DNA would remain cross-linked to the proteins that recognise these specific DNA sequences. Benzonase will have no impact as it will be denatured in the presence of 8M urea. Although hisZta was shown to bind to the nickel affinity gel in the presence of intact DNA (Figure 4.2), it was considered that the presence of intact cross-linked DNA might interfere with the binding capacity of the nickel affinity gel. Therefore, sonication of the sample was introduced and a pull down assay performed. The cross-linked HEK293-*BZLF1*-KO transfected cells were subjected to 10 pulses of 10 seconds sonication at 30% amplitude during being denatured in cell lysis buffer containing 8M urea. The sonicated denatured extracts were applied to nickel affinity gel (Figure 4.9), to identify if hisZta binds after sonication of the sample.

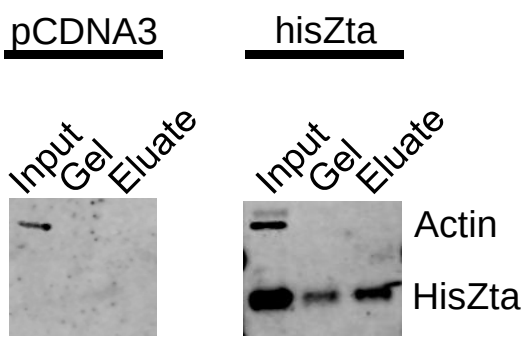


Figure 4.9 Pull down attempt with elution of the gel with sonication of protein extract.Cross-linked hisZta extract was applied to Nickel agarose gel after being denatured and sonicated. The extract was applied to 50µl of beads. This was washed with wash buffer and eluted using elution buffer without imidazole three times. The elutions were pooled and subject to acetone precipitation before being resuspended in protein sample buffer. All samples were mixed equally with protein sample buffer and the proteins were separated by SDS-PAGE and a western blot for Zta and actin performed. The antibody used here was BZ1 and anti-actin.

The binding profile of hisZta to the nickel affinity gel displays a similar binding to Figure 4.8, where some hisZta is left bound to the gel after elution (Figure 4.9). After elution with elution buffer, some hisZta is eluted from the nickel affinity gel. It was chosen to include the sonication step in the protocol for further analysis.

A pull down was attempted with extracts applied to nickel affinity gel. A transfected hisZtaAAA sample was also included here in addition to the control and hisZta cell extracts. SILAC multiplex experiments can also be performed using multiple SILAC labels, light, medium and heavy. Three treated cell populations can be analysed together in a single experiment, therefore it was chosen to include another level of analysis here with hisZtaAAA transfection

The cross-linked denatured cell extracts were sonicated and applied to the nickel affinity gel. The gel was eluted five times with elution buffer and then the elutions were combined together. This elution mixture was subjected to acetone precipitation to concentrate the proteins present. The protein pellet that formed from protein precipitation was resuspended into protein sample buffer and all of the samples were separated by SDS-PAGE for western blot analysis or coomassie staining of the gel.

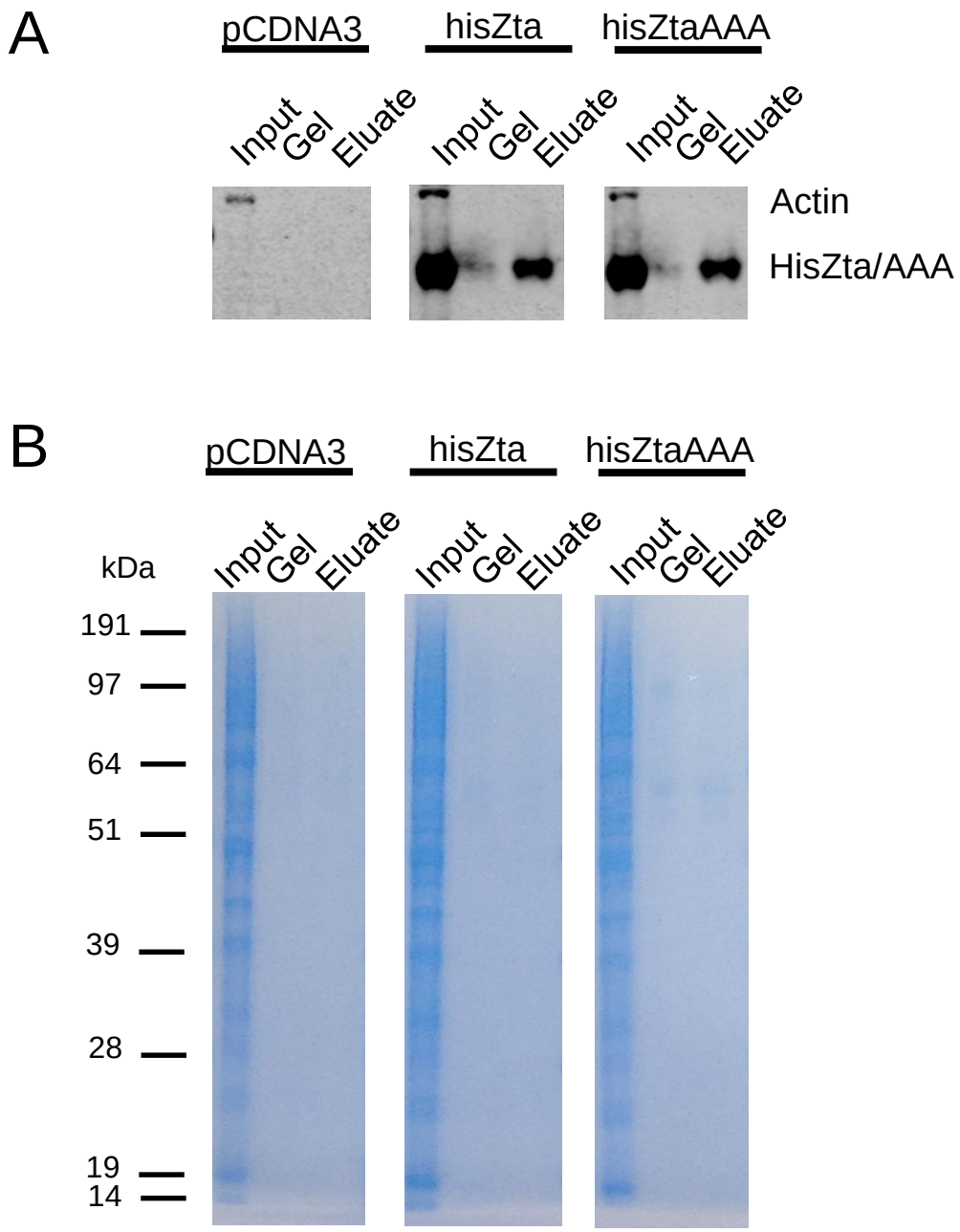


Figure 4.10 Pull down and elution attempt with sonication of protein extract. pCDNA3, hisZta, hisZtaAAA. Cross-linked hisZta extract was applied to Nickel agarose gel after being denatured and sonicated. The extract was applied to 50µl of beads. This was washed with wash buffer and eluted using elution buffer without imidazole. The samples were eluted five times and elutions combined, before precipitation with acetone and resuspended in protein sample buffer. All samples were mixed equally with protein sample buffer and run on a protein gel for: **A** The proteins were separated by SDS-PAGE and a western blot for Zta performed. The antibody used here was BZ1 **B** Coomassie staining.

The input, gel and eluate were analysed by western blot (Figure 4.10A). The hisZta protein applied to the gel displays minimal binding after elution. The consecutive elution of hisZta and hisZtaAAA from the gel was successful, after being processed and concentrated by acetone precipitation. The protein band for hisZta in the elution is more prominent compared to the gel in Figure 4.9 suggesting that the majority of protein bound to the affinity gel here is eluted from the gel. A small amount of the samples were separated independently by SDS-PAGE and this gel was analysed by coomassie staining. This demonstrated the overall protein abundance of the proteins present in each sample. The input for each of the cell extracts showed a high abundance of proteins, which is to be expected. However the proteins in the elution and gel samples were extremely minimal, including hisZta and hisZtaAAA. Although these proteins display a strong band when analysed by western blot these proteins are barely visible on coomassie staining.

To ask if protein or the nickel affinity gel were a limiting factor, a varied amount of protein was assessed. When scaling up this experiment it is possible to use a greater volume of protein extract against the affinity gel. By scaling up a pull down experiment to maximize the amount of protein to be subsequently analysed, a pull down assay was compared using the same volume of beads but a small and large extract volume. This was performed to maximize the amount of hisZta that could bind to the gel. An extract volume of 100µl or 900µl was applied to 50µl of nickel affinity gel. The gel was washed with wash buffer and then the gel was boiled in 2X sample buffer before all of the samples were analysed by western blot after being separated by SDS-PAGE (Figure 4.11). It was evident that using a greater volume of cell extract onto the same volume of affinity gel gives more protein binding and so the amount of nickel affinity gel was not a limiting factor.

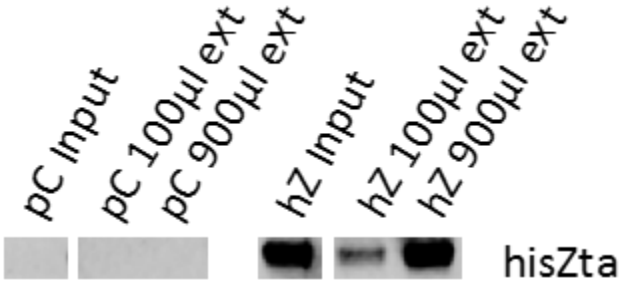


Figure 4.11 Pull down attempt with varying extract volume to the gel with protein extract. pCDNA3 and hisZta. Cross-linked hisZta extract was applied to Nickel agarose gel after being denatured and sonicated. The either 100µl or 900µl of extract was applied to 50µl of beads. This was washed with wash buffer boiled in protein sample buffer. All samples were mixed equally with protein sample buffer and the roteins were separated by SDS-PAGE and a western blot for Zta performed. The antibody used here was BZ1.

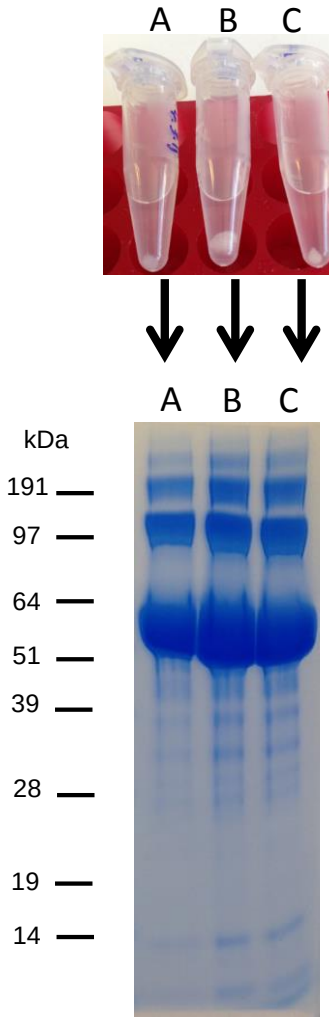


Figure 4.12 Denatured BSA precipitation using acetone and BSA solubility in urea. 200ng of BSA was denatured from elution buffer containing 0M to 8M urea and subjected to acetone precipitation. Three conditions were compared: 'A' Elution buffer without urea (0M) 'B' Elution buffer with 8M urea 'C' Elution buffer diluted to 4M urea. The precipitated samples were resuspended in protein sample buffer and separated by SDS-PAGE before staining the gel with SimplyBlue safestain

To determine if the presence of urea would have a detrimental effect on the precipitation of proteins from the elution buffer, BSA was placed into varying elution buffers with different buffer conditions. 200ng of BSA was added to elution buffers containing 0M Urea, 8M urea or elution buffer diluted to 4M urea. All three samples were precipitated using acetone and the protein pellet was resuspended in 2X protein sample buffer. The BSA present in each condition from the acetone precipitation was separated by SDS-PAGE and analysed by coomassie staining (Figure 4.12).

It can be seen that a protein pellet was present throughout the conditions here. After resuspending and boiling the protein pellet in 2X sample buffer and separation by SDS-PAGE, the stained gel shows that BSA was present in each sample. Other bands were detected by the staining. It was decided that the conditions developed and implemented here would be taken forward and executed towards a SILAC labeled pull down experiment, with the proteins eluted from the affinity gel to be analysed by mass spectrometry proteomics.

4.2.3. SILAC labeled histidine-tagged Zta pull down and elution

In order to perform an informative and successful pull down experiment to analyse hisZta protein interactions SILAC was utilized. Using SILAC labeled cells enables a quantitative level of analysis during mass spectrometry evaluation. A labeling schematic diagram illustrates how HEK293-*BZLF1*-KO would be labeled, transfected, harvested and cross-linked before being denatured and applied to nickel affinity gel (Figure 4.13).

The HEK293-*BZLF1*-KO cells were labeled with SILAC media. Flasks of cells were prepared for each transfection. The ‘light’ SILAC labeled cells were transfected with pCDNA3 (control, latent). The ‘medium’ SILAC labeled cells were transfected with hisZta (lytic). The ‘heavy’ SILAC labeled cells were transfected with hisZtaAAA (Unable to initiate lytic replication). After 96 hours the cells were cross-linked with 1% formaldehyde and harvested. A fraction of these cells were taken for analysis to determine if the transfection was successful and the early phase of lytic replication was initiated. The cells were lysed in cell lytic

reagent and sonicated before a fraction of the cell extract supernatant was mixed 1:1 with 2X protein sample buffer, and separated by SDS-PAGE for western blot analysis. A western blot for the transfected flasks demonstrates hisZta and hisZtaAAA were both readily expressed in the transfected cells, and lytic protein BMRF1 was also expressed by transfection of both hisZta and hisZtaAAA (Figure 4.14A). A qPCR was performed on a fraction of the transfected HEK293-*BZLF1*-KO cells for the presence of the EBV genome. Only hisZta initiated the lytic cycle and replication of the virus (Figure 4.14B). This demonstrated that the transfection was successful and these cells were put forward for further analysis, the cells extracts were taken forward for a larger scale pull down assay.

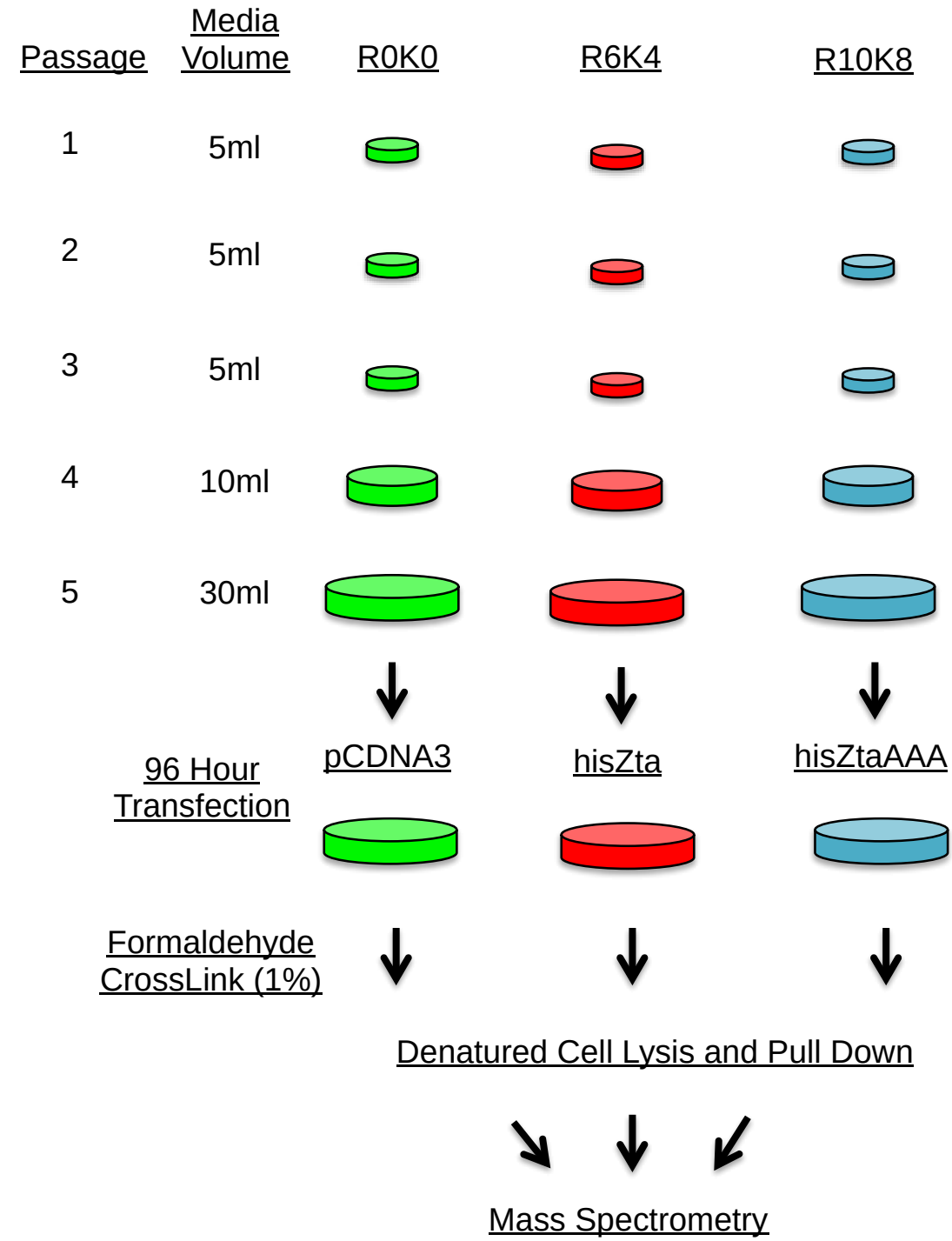


Figure 4.13 SILAC labelling schematic of HEK293-BZLF-KO cells for transfection Using light (R0K0), medium (R6K4), heavy (R10K8) media (Dundee Cell) HEK293-BZLF-KO cellsare passaged 1:3 using the media volume shown. Minimum of 5 passages to ensure all of the cells are metabolically labelled. The cells are transfected with control, hisZta or hisZtaAAA for 96 hours. Cells are crosslinked with 1% formaldehyde and then harvested for further study.

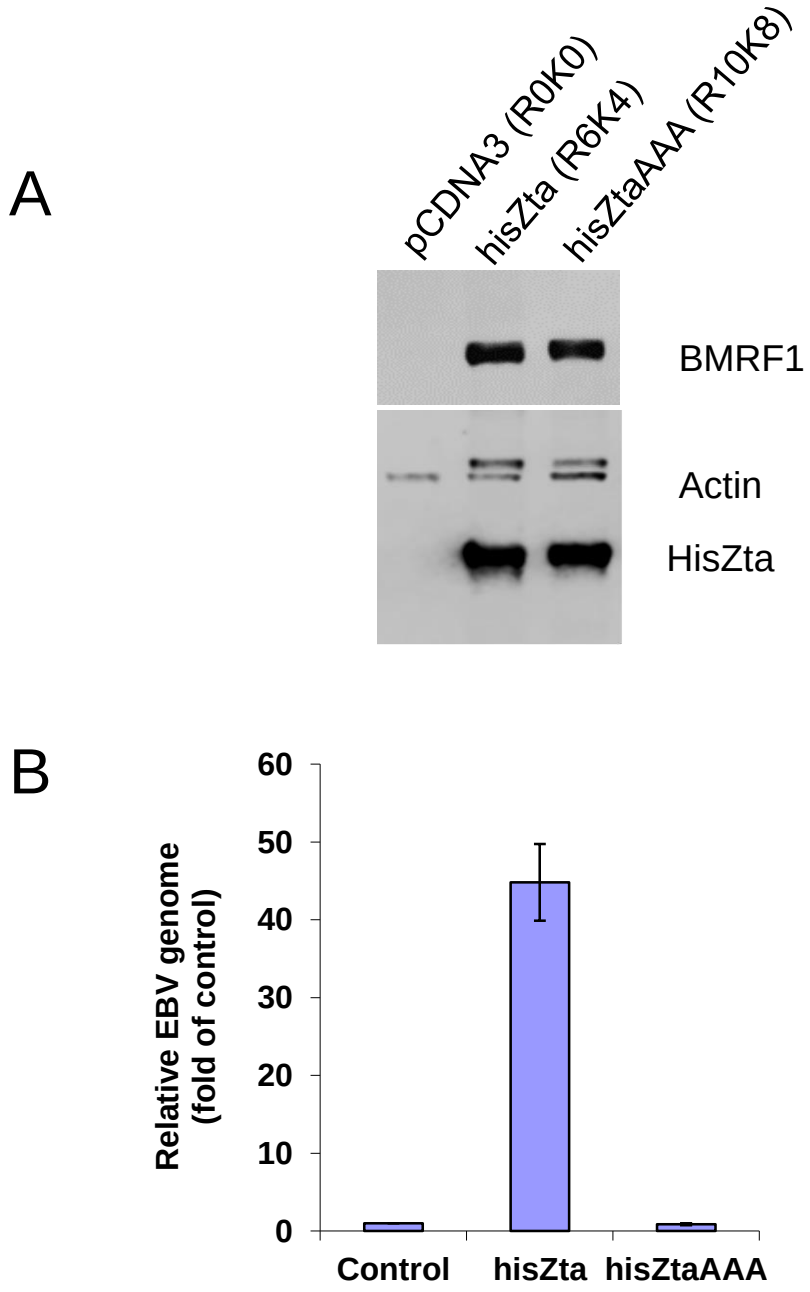


Figure 4.14 Western blot and qPCR to confirm the expression of transfected proteins and induced viral replication. **A** HEK293-BZLF-KO were transfected with pCDNA3 (R0K0 media), hisZta (R6K4 media) and hisZtaAAA (R10K8 media) and crosslinked after 96 hours. Proteins were run on a protein gel for western blot analysis. The antibodies used were BZ1, anti-BMRF1 and anti-actin **B** Quantitative Real Time PCR (qPCR) was used to detect the presence of the EBV genome and beta-globin. Results were standardised by the amount of beta-globin present. qPCR repeated in triplicate and the standard error displayed between the experiment.

A schematic diagram of the larger scale SILAC labeled pull down assay is illustrated (Figure 4.15). The HEK293-*BZLF1*-KO cells were grown in SILAC media, transfected and then cross-linked after 96 hours incubation. The cells were lysed in cell lytic reagent containing 8M urea and then sonicated, and protein expression demonstrated from a fraction of the cells (Figure 4.14A). These cell lysates were applied to a large volume of equilibrated nickel affinity gel and incubated. The gel was washed and then eluted in five successions before the multiple elutions were precipitated together. The pelleted protein was resuspended in protein sample buffer. The proteins would then be analysed by mass spectrometry proteomics analysis.

The pull down was performed to the method illustrated (Figure 4.15). The hisZta bound gel was washed and eluted and the protein buffer precipitated with acetone. The pellet that was formed from the acetone precipitation formed a large aggregate. The samples in the tube were resuspended in 10µl of 2X protein sample buffer but the pellet precipitate was too large to be resuspended in that volume of buffer. 40µl of sample buffer was added to make a total volume of 50µl (Figure 4.16). This was added to the pellet and the mixture transferred to a 1.5ml tube for western blot and coomassie analysis.

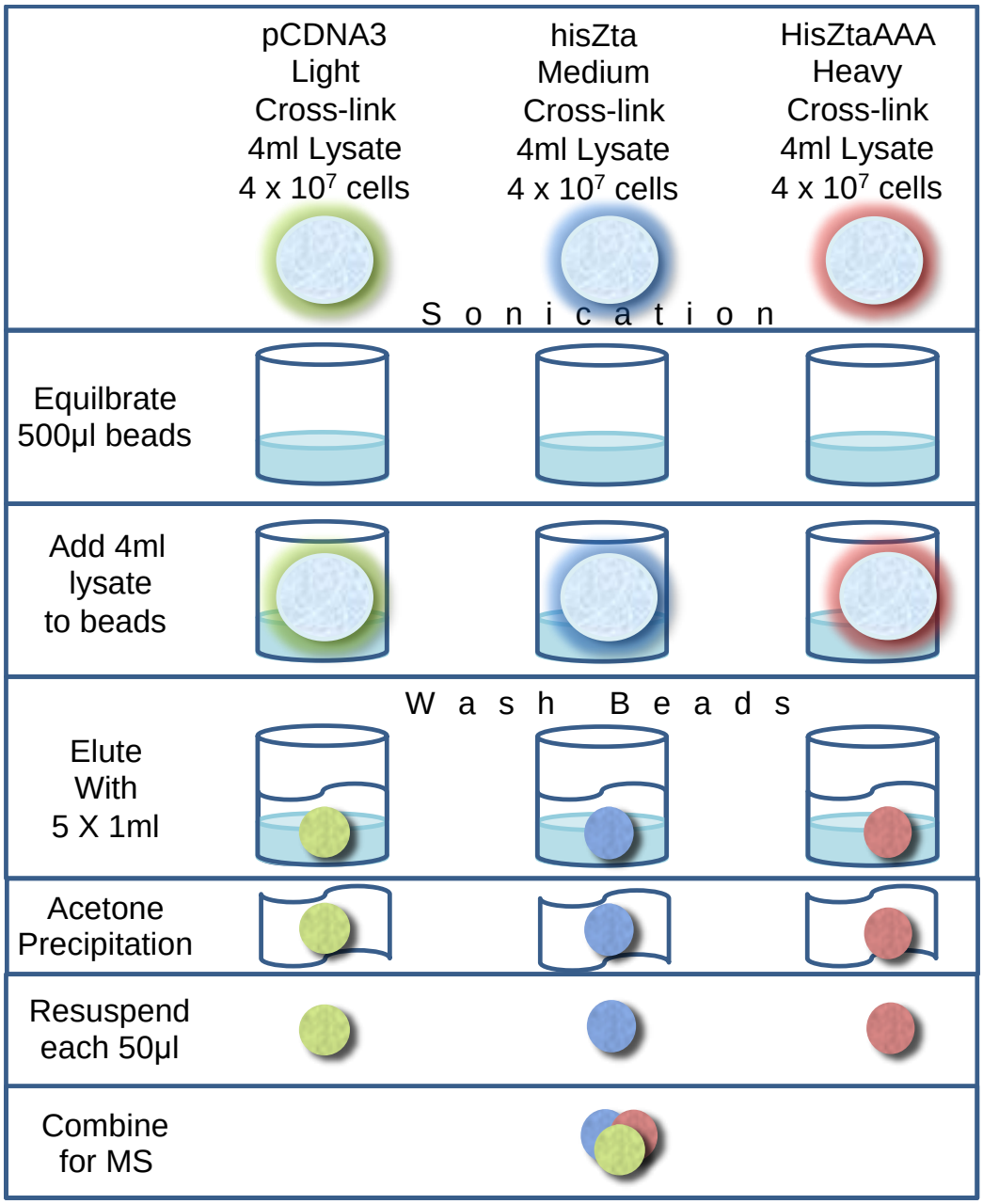


Figure 4.15 Diagram of large scale denatured crosslinked SILAC labelled pull down of cell extracts.

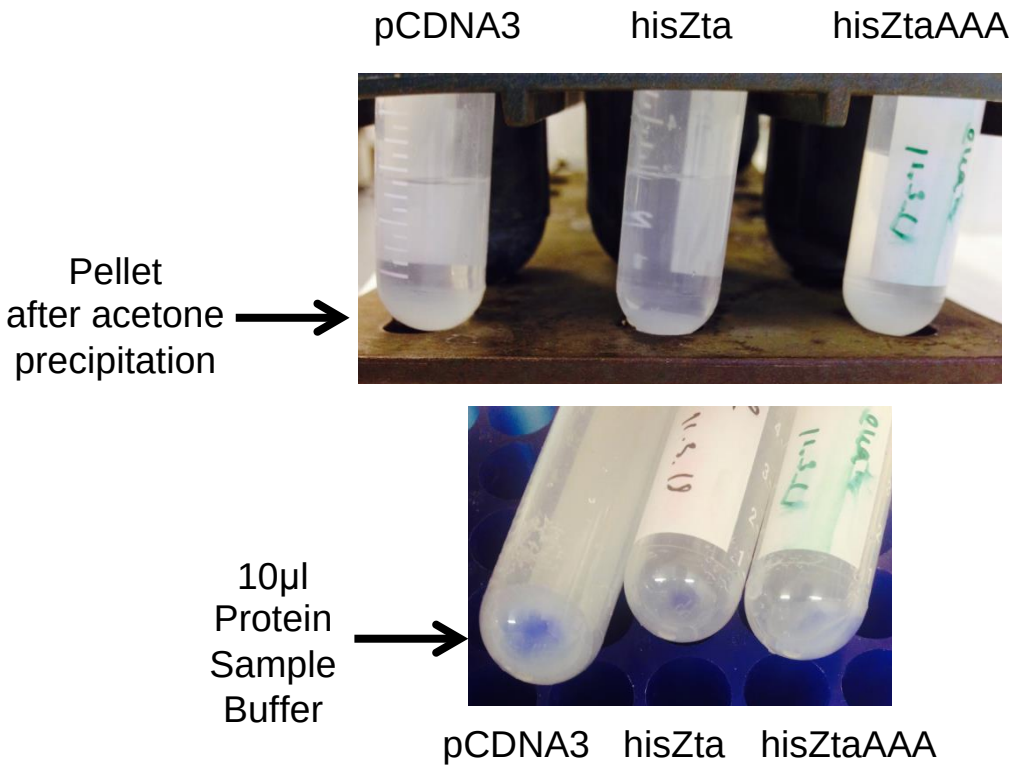


Figure 4.16 Attempt of resuspending of the protein pellet after acetone precipitation. The pull down was performed as illustrated by Figure 17. The hisZta bound gel was washed and eluted and the protein buffer precipitated with acetone. The pellet that was formed from the acetone precipitation formed a large aggreagate. The samples in the tube were to be resuspended in 10µl of 2X protein sample buffer but the pellet precipitate was too large to be resuspended in that volume of buffer. 40µl of sample buffer was added to make a total volume of 50µl. This was added to the pellet and the mixture transferred to a 1.5ml tube.

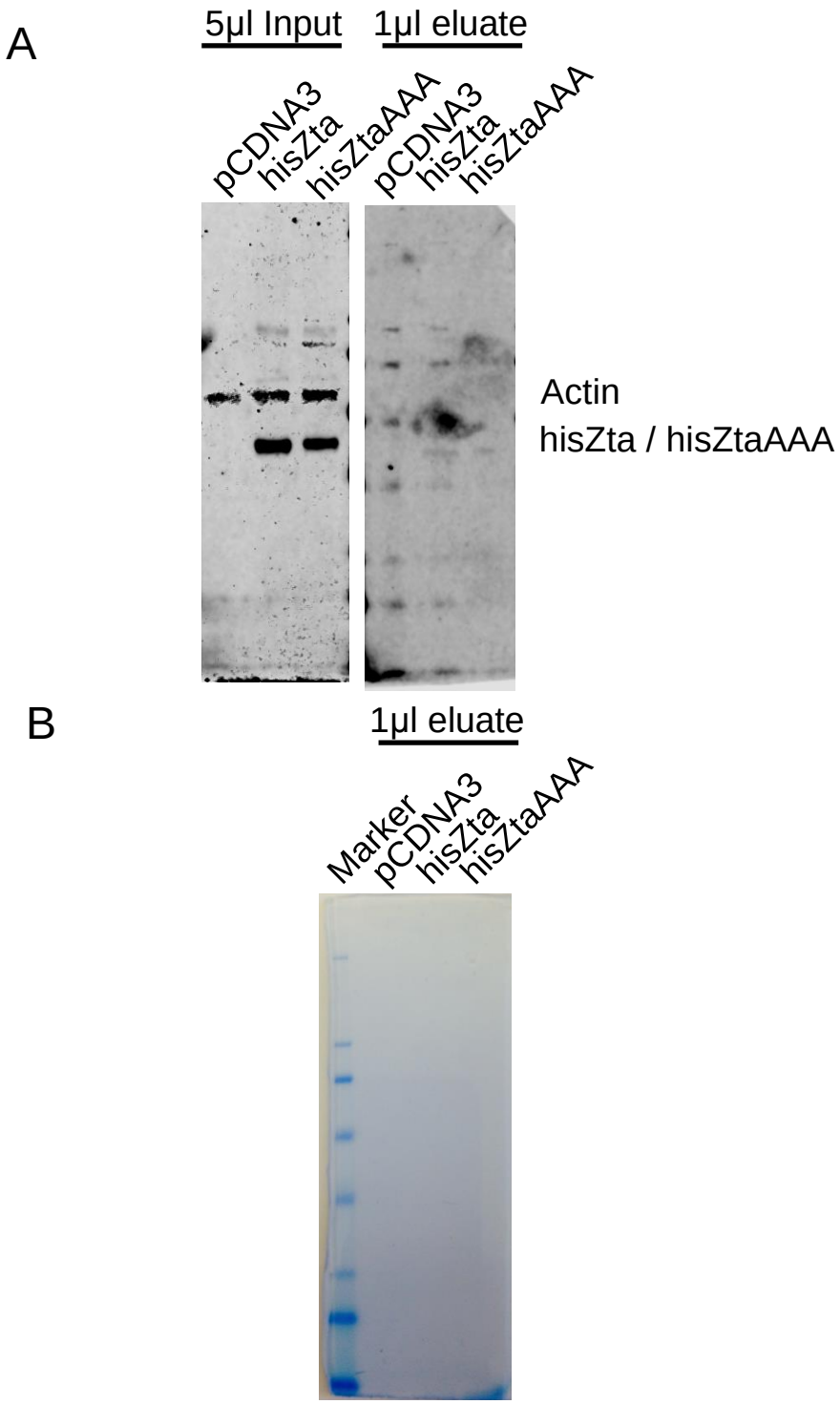


Figure 4.17 Large scale pull down western blot and coomassie staining results. The pull down was performed as previously described. The samples were boiled in protein sample buffer and a fraction of each sample separated by SDS-PAGE and analysed. **A** Western blot was performed on a small fraction of the the control, hisZta and hisZtaAAA pull down samples. The antibodies used were BZ1 for hisZta and hisZtaAAA detection and anti-actin. **B** A Coomassie staining was performed on a gel of the same samples analysed by western blot.

The results of the attempted pull down experiment were displayed in a western blot and coomassie stained gel (Figure 4.17). The input that was applied to the nickel affinity gel was reconfirmed again (Figure 4.17A). The protein samples were separated by SDS-PAGE here but there were no specific protein bands detected by the antibody. A small amount of sample should display hisZta or hisZtaAAA if they are bound, eluted and precipitated from the gel.

The same samples were also analysed by coomassie staining of a separate gel after separation by SDS-PAGE (Figure 4.17B). Again, the samples did not display evidence of protein here without any protein bands being stained, which suggest that the pull down, elution, precipitation or resuspension of the sample was not successful.

4.3. Discussion

The question asked was to identify interaction partners in EBV positive cells with an alternative approach. This was attempted using a denaturing method combined with cross-linking of proteins to preserve immediate interactions. Histidine tagged Zta was demonstrated to bind to the nickel affinity gel in Chapter 3, and although some potential interactions were investigated the overall binding efficiency of Zta to the gel was inconsistent.

The first obstacle included denaturing cell extracts. This was performed by developing the cell lytic reagent to contain a concentration of 8M of urea. After applying the denaturing cell lysis buffer to cells the supernatant was taken for analysis. The transfected protein hisZta could bind to the nickel affinity gel in denatured form. The addition of imidazole to the wash buffers minimized non-specific binding observed by actin binding, and increasing imidazole in the wash buffer reduced the binding of Zta. As these extracts were denatured, the proteins would lose their tertiary shape and function of protein domains, ultimately losing all protein interactions. Therefore the cells and proteins were cross-linked to preserve these interactions

The cross-linking of transfected cells was optimized between 0.1% and 1% formaldehyde, while native and denatured conditions were compared. It was

established that hisZta could bind to the nickel affinity gel under denatured conditions while cross-linked with 1% formaldehyde (Figure 4.4). Cross-linking of cells up to 1% is routinely used in ChIP assays. Here the application of formaldehyde allows the freezing of the cell in an intact but non-functional state (Vasilescu 2004).

After the establishment of cross-linked, denatured hisZta binding to the affinity gel, it was decided to attempt to elute the complexes from the beads using 250mM imidazole in the elution buffer. The addition of 250mM imidazole in the elution buffer minimised hisZta to bind to the nickel affinity gel. It was detrimental to Zta binding to the affinity gel (Figure 4.5); therefore it was decided to attempt to elute the complexes from the beads using only a low pH buffer, without any imidazole. The low pH alters the affinity of the histidine tagged proteins with the affinity gel (Gordon et al, 2000), therefore eluting his-tag proteins

A multiple elution stage was combined and precipitated using acetone. An optimal pull down assay was performed, where the proteins were resuspended in protein sample buffer and shown to be eluted from the affinity gel.

The maximal amount of protein in EBV positive cells was to be applied to the nickel affinity gel, and after an elution the proteins would be referred for mass spectrometry analysis. The optimal conditions for the pull down were determined while labelling HEK293-*BZLF1*-KO cells with SILAC media. The final scaled up assay including SILAC labelled transfected extracts was unsuccessful. There were no proteins bands observed using the antibody and a coomassie stain of the separated proteins did not observe any bands. Therefore it was clear that far more starting material is required to be optimal to the bead volume.

5. Interpretation of cellular components associated with the EBV lytic replisome through SILAC gel filtration

5.1. Introduction

Protein complexes control most cellular processes. Many proteins have been identified and their functions determined which include transcription, translation, protein folding, degradation and replication. Through protein network analyses of single cell organisms, identification of stable protein complexes allowed some of these functions to be determined with protein-protein interactions (Gavin et al. 2006; Krogan et al. 2006).

Analysis of human protein-protein interactions allowed a database of protein features including post-translational modifications to be created. Isoforms of proteins are also included (Mishra et al. 2006). The genomic, transcriptomic and proteomic data is being more finely characterized each year. Proteins of all sizes interact with many others of varying size and larger proteins may preferentially interact with much smaller proteins (Wong et al. 2008).

Understanding EBV proteins interacting with the host has been previously attempted. EBV-human interactome network from yeast two-hybrid screening displayed many interactions with the host in viral latency and lytic cycles (Calderwood et al. 2007). It has been previously outlined in Chapter 1 that viral proteins interact with the host proteome in order to facilitate viral replication. As these protein interactions contribute to maintaining the viral genome within the host environment, and then facilitating lytic viral replication, the question of what molecular protein complexes are formed during lytic replication? The protein subunits that may be recruited to EBV viral replication compartments have not been fully identified but studies have begun to identify components and the architecture with viral proteins ((Daikoku et al. 2005; Amon et al. 2006; Kudoh et al. 2009; Sugimoto et al. 2011).

To analyse a complex mixture of proteins, fast protein liquid chromatography can be utilized. Size exclusion chromatography (SEC) separates proteins and

complexes by molecular weight and size. A chromatography column can separate these protein samples into eluates that can be further analysed.

The AKTA purifier system (GE Healthcare) can contain a pump to move the chosen buffer through the system. The sample can be injected into the injection loop, before being passed into the SEC column. This buffer transfers a protein sample through the column. A protein concentration detector measures the absorption of ultraviolet (UV) light by the complexes at 280nm, to obtain absorbance data of the components of the elution. The protein fractions can be collected in equal volumes as an eluate, for further analysis. A schematic diagram represents the size exclusion chromatography system (Figure 5.1)

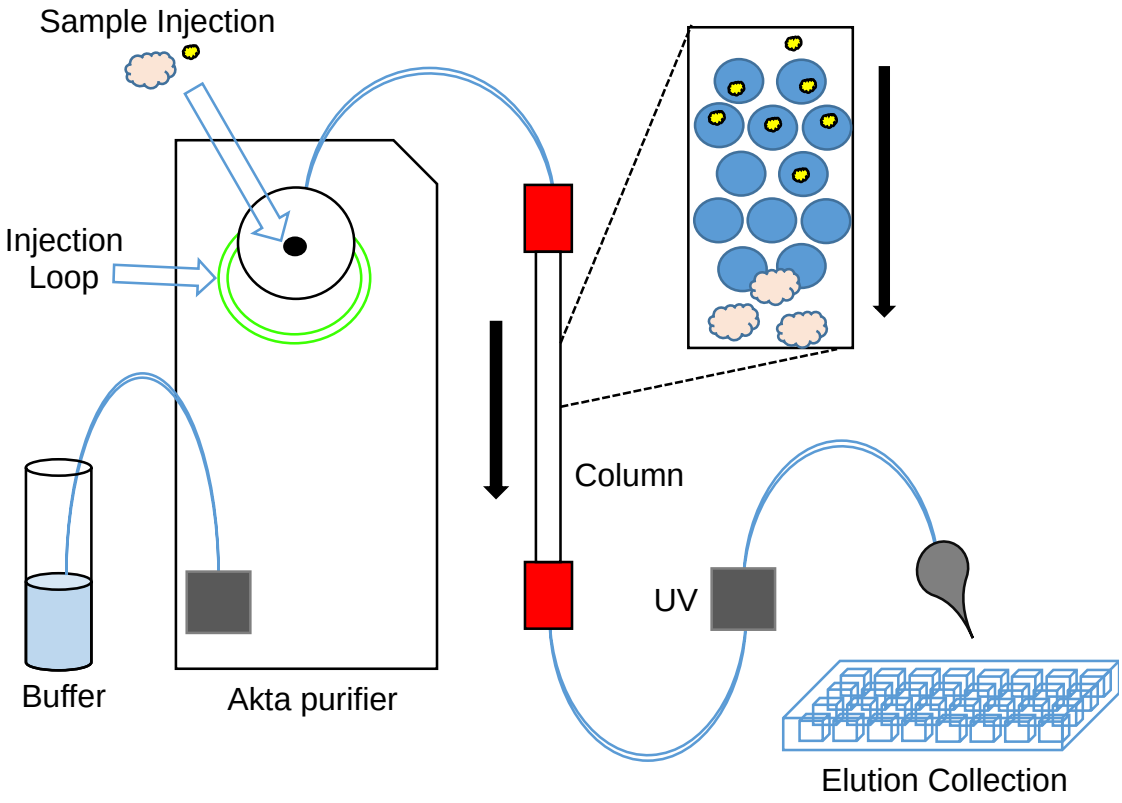


Figure 5.1 Diagram of the Akta purifier system connected to a size exclusion chromatography column. Samples are injected into the infection loop, before a buffer is applied to the equilibrated column and the sample can pass through. Proteins in the sample are separated by size as larger proteins or complexes can pass through the matrix, smaller proteins are delayed by pores in the matrix. Larger complexes elute first from the column, smaller complexes elute later.

Size exclusion chromatography of protein complexes coupled to mass spectrometry have been employed previously to identify interactions (Kirkwood et al. 2013). Their study included soluble protein complexes that included elongator factors, isolated from U2OS cells and characterised across multiple fractions.

A native size-exclusion chromatography experiment with proteomic analysis was designed and aimed to characterize soluble protein complexes isolated from transfected EBV negative and EBV positive cells. Targeting differences between a control cell population, cells undergoing full lytic replication and cells that begin the lytic cycle but not viral replication using mass spectrometry may indicate towards proteins that contribute to EBV lytic replication.

5.2. Results

5.2.1. Structure, transcription and replication function of Zta and Zta mutants

Zta is a protein composed of 245 amino acids. The protein contains a transactivation domain located within the N-terminal half of the protein Zta contains a DNA binding domain and a bZIP domain located before the C-terminus (Figure 5.2). The C-terminus is essential for lytic replication. Mutations in the C-terminal region of Zta is defective for EBV replication (Bailey et al, 2009). Making alterations to coding sequence of the wild type protein has enabled mutant Zta proteins to be produced. A six-histidine tag was engineered onto the N-terminus of all Zta proteins (Bailey et al 2009, Q Karlsson). The wild type and mutant proteins have been studied here and their transactivation and replication properties assessed. One mutant protein has the last three amino acids of the C-terminal domain mutated to alanine to assess its role in replication (Q. Karlsson). Another mutant starts at amino acid 134 that shortens the protein by removing the transactivation domain. The DNA binding domain, bZIP domain and C-terminus remain unaffected but this protein cannot initiate lytic replication because it is unable to activate transcription of lytic genes. Zta wild type and the AAA mutant's abilities to perform their functions were investigated further.

The transcriptional activity of Zta and the Zta mutant were assessed in a transactivation assay. The ability of wild type protein and mutant to activate the BHLF1 promoter sequence placed before a luciferase gene. The BHLF1 promoter contains four Zta response elements (ZREs) that Zta is known to bind strongly to. The mutation promoter sequence has these sequences mutated to CCCCTT, a sequence Zta cannot recognize to bind. (Figure 5.3A).

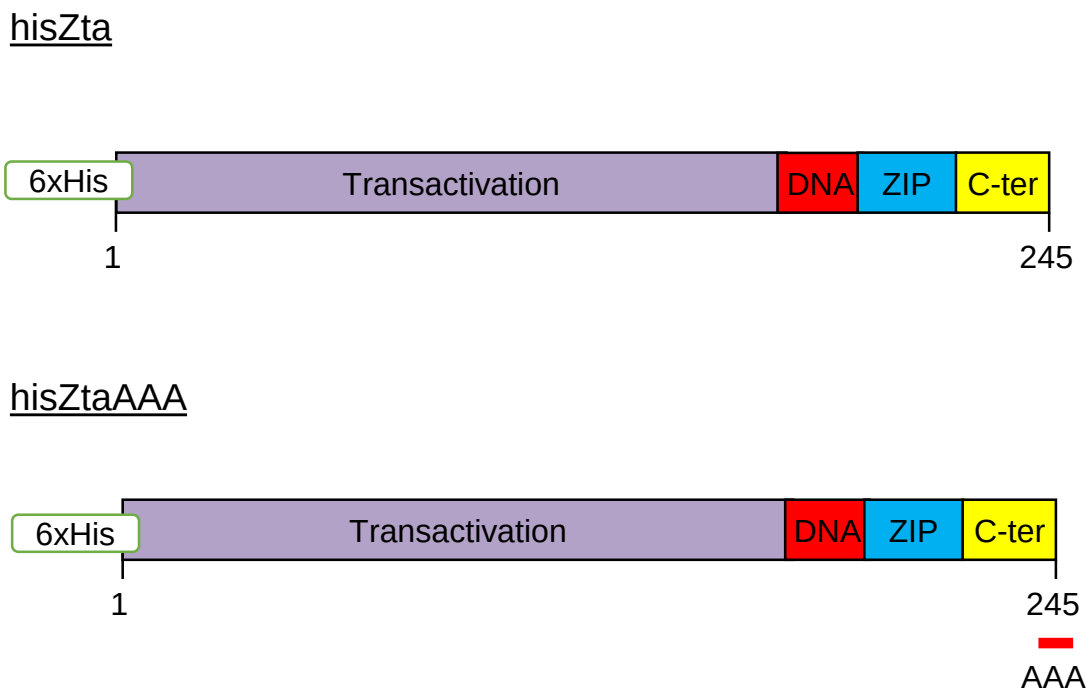


Figure 5.2 Schematic of Zta demonstrating transactivation and replication domains. A six-histidine tag is located on the N-terminus of hisZta and mutants. Zta contains a transactivation domain (purple), DNA binding domain (red), bZIP domain (blue) and C-terminus (yellow). The hisZtaAAA mutant has the last three amino acids mutated to alanine.

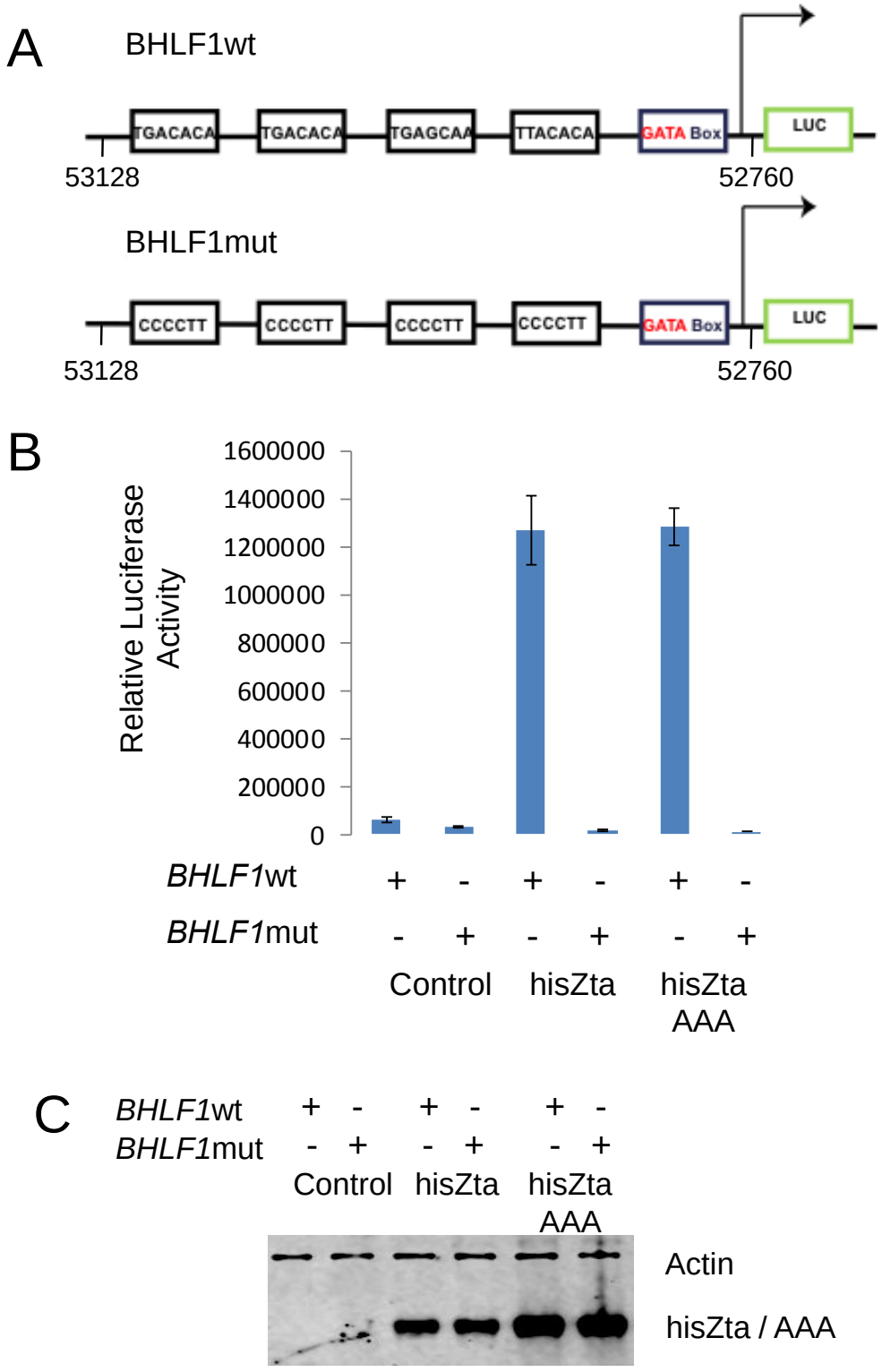


Figure 5.3 Transcription activity of hisZta and hisZta mutants. Transactivation assay to assess the transcriptional activity of Zta mutants. U2OS cells were transfected and harvested 24 hours later before luciferase levels were detected **A** The BHLF1 promoter ZRE sequence were used in a luciferase assay with these sequences mutated for BHLF1mut sequence. **B** Both hisZta wild type and hisZtaAAA are transcriptionally competent. The wild type protein and hisZtaAAA mutant can activate the BHLF1 promoter in a luciferase assay system. The standard error is shown here after the experiment performed in triplicate. **C** Western blot for hisZta and mutant expression. The proteins were separated by SDS-PAGE and a western blot for Zta and actin performed. The antibody used here was BZ1 and anti-actin. Actin was detected as loading control.

U2OS cells were co-transfected with expression vectors for hisZta and hisZtaAAA and luciferase vectors and harvested 24 hours later. Half of the cells were processed for luciferase activity using a luciferase assay system. The remaining cells were processed for western blot protein expression. The luciferase levels were detected in triplicate and plotted together.

Both hisZta and hisZtaAAA activate the wild type BHLF1 promoter (Figure 5.3B). None of the proteins activated the BHLF1 mutant.
The western blot shows the abundance the expression of hisZta and mutant proteins (Figure 5.2C). There was more expression of hisZtaAAA than hisZta but the luciferase levels were similar. This demonstrates the activity of transactivation by hisZta and hisZtaAAA.

Zta can also act as a replication factor to reactivate EBV from latency into lytic cycle. We asked whether ZtaAAA mutation impacts on this using HEK293-*BZLF1*-KO cells (Feederle et al. 2000). These cells contain a stable transfection of the EBV genome that does not contain a functional *BZLF1* gene. As Zta is essential for the activation of lytic cycle, EBV within these cells always remain in latency. Transfection of Zta is required to initiate the EBV lytic cycle and replication of the EBV genome.

Using this cell system, the replication activity of hisZta and the hisZtaAAA. hisZtaAAA and were assessed after transfection using a qPCR assay. Cells were harvested after 96 hours after transfection to allow the initiation of lytic replication of the EBV genome. A sample of cells was used for western blot to confirm the expression of the transfected proteins. The remaining cells were processed for qPCR (Figure 5.4).

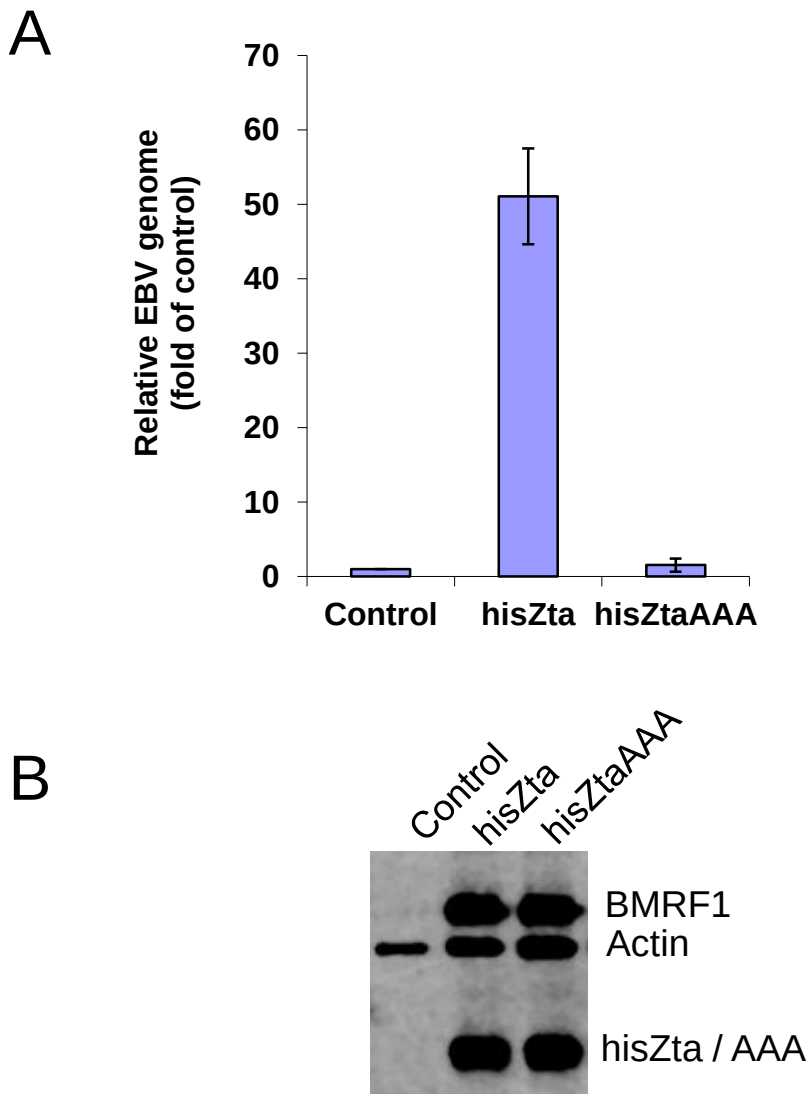


Figure 5.4 hisZta and hisZta mutants' ability to initiate genome replication by qPCR. HEK293-BZLF1-KO cells were transfected with hisZta or mutants and qPCR was used to detect the presence of the EBV genome and β-globin after 96 hours. **A** Only hisZta wild type can initiate lytic replication. hisZtaAAA can initiate BMRF1 expression but the mutant cannot replicate the EBV genome. The standard error is shown here. **B** Western blot for hisZta and mutant expression. Proteins were separated by SDS-PAGE. The antibodies used here were BZ1, anti-BMRF1 and anti-actin.

A qPCR assay was performed to detect the level of virus genome. This was measured using primers for the viral DNA polymerase gene *BALF5*. Primers for β-globin were used to detect the human genome copy numbers. This was used to standardise the samples.

Transfection with hisZta into these cells led to lytic replication, an increase in the level of EBV genome (Figure 5.4A). There was no increase in EBV genome levels with transfection of hisZtaAAA. The EBV genome levels are similar to the background control with transfection of pCDNA3 control. This baseline level is the detection of the latent viral genome copies already present in the cell. The fold change between control and viral load for these cells was 19.3.

A western blot for protein detection from transfected cells with hisZta and hisZtaAAA was performed (Figure 5.4B). The transfected proteins were readily expressed. BMRF1, the viral early lytic protein is also detected here suggesting that the beginning of lytic cycle is induced by both hisZta and hisZtaAAA, although lytic replication of EBV is not present with the hisZtaAAA transfection. As both hisZta and hisZtaAAA contain the transactivation domain, the lytic gene for the DNA processivity factor BMRF1 is expressed upon both hisZta and hisZtaAAA transfection. The hisZtaAAA mutant protein can transactivate at least one of the lytic genes, but cannot promote replication of the EBV genome. Therefore the last three amino acids of the C-terminus are essential for EBV genome replication.

The transactivation and replication functions were classified here and were then taken forward for further analysis in the context of EBV lytic replication. Size exclusion chromatography was then used to investigate the molecular complexes formed in these transfected cells undergoing full lytic cycle and lytic cycle without viral genome replication. A comparison between the lytic and defective lytic samples would indicate whether cellular and viral Zta-complexes differ in cells when they undergo EBV lytic cycle.

5.2.2. Superose 6 10/300GL size exclusion column calibration and determination of molecular weight standards

The Superose 6 10/300 GL (GE Healthcare) column is used for high performance size exclusion chromatography of protein samples. The separation range is from 5,000Da to 5,000,000Da. This column is easy to manipulate for the separation of molecular complexes. It is chemically stable in most common buffers and pH ranges, and connects readily to ÄKTA purifier systems (GE Healthcare).

To determine the column performance control, molecular weight standards were applied to the column. The column was connect to an ÄKTA purifier and washed with dH2O to remove 20% ethanol the column is stored in. The column was equilibrated in two column volumes of 20mM Tris 100mM NaCl at pH 8.0 buffer. This supports protein stability and avoid non-specific ionic interactions with the column gel. Molecular weight standards were injected into the system and separated by size exclusion.

The elution profile of these standards with the UV absorbance units is displayed in milli Absorbance unit (mAU) (Figure 5.5). The first peak at fraction H4/13ml is thyroglobulin, followed by conalbumin and a large peak at 17.5ml for ovalbumin. RibonucleaseA elutes last from the column at 19ml. This profile agrees with the instruction protocol provided by GE Healthcare. This demonstrated that the column is functional and will perform as expected to the manufacturers' standards.

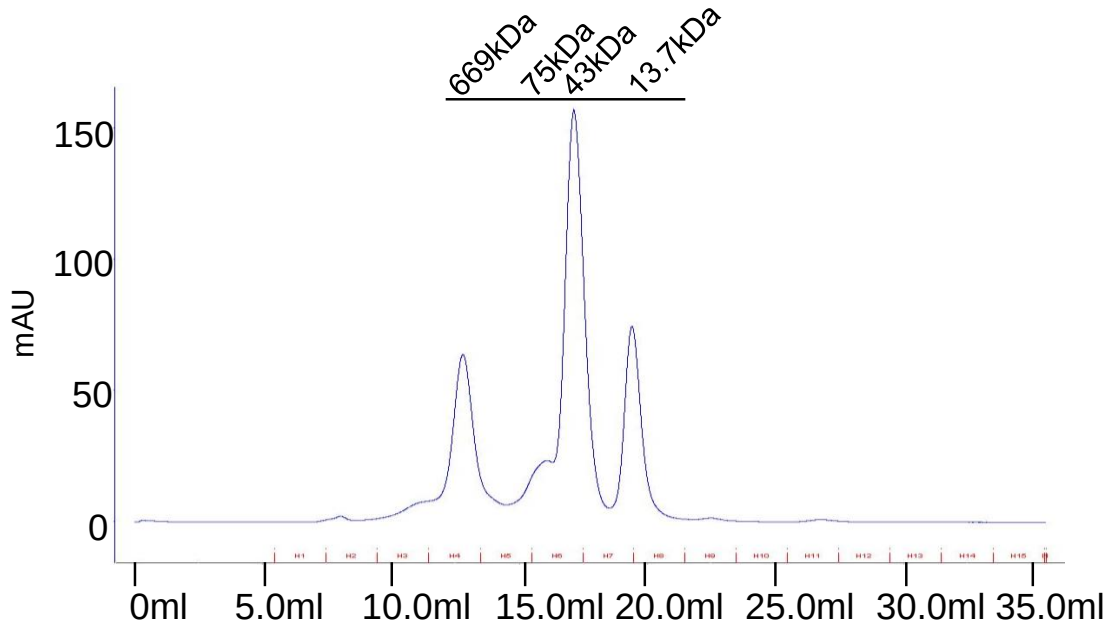


Figure 5.5 Determination of Superose 6 10/300 GL elution profile using molecular weight standards. Molecular weight standards to identify the molecular weight of fractions eluted from the column. Thyroglobulin 669,000Da at eluted 13ml, Conalbumin 75,000Da eluted 16ml, Ovalbumin 43,000Da eluted 17.5ml, RibonucleaseA 13,700Da eluted at 19ml. The standards included thyroglobulin 669,000Da at 5mg/ml, conalbumin 75,000Da at 3mg/ml, ovalbumin 43,000Da at 4mg/ml, ribonucleaseA 13,700Da at 3mg/ml.

5.2.3. Determination of the elution profile of Zta in U2OS cells

It was not known where Zta would elute from if a cell sample extract expressing Zta would be separated through a size exclusion chromatography column. By transfecting EBV negative cells with Zta expression vector, and passing a cell lysate through a column and collecting the elution, it is possible to identify what fraction Zta elutes into. The control, hisZta and hisZtaAAA expression vectors were transfected into U2OS cells and incubated for 72 hours. These cells were harvested and checked for protein expression including actin and Zta expression levels (Figure 5.6). The cells were transfected successfully and the cells from this transfection were kept for future analysis. These were applied to size exclusion chromatography studies.

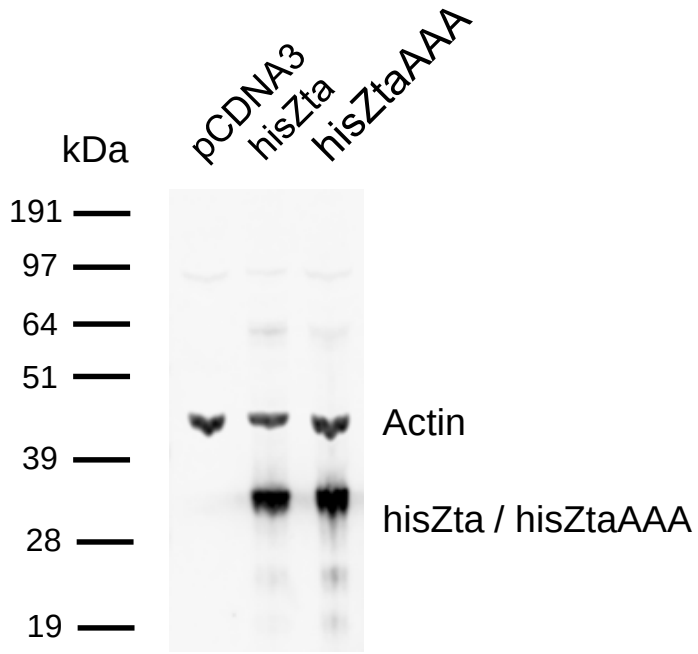


Figure 5.6 Protein expression of hisZta and mutants in U2OS cells. U2OS cells were transfected with control expression vector, hisZta, and hisZtaAAA and harvested after 72 hours. Cells were lysed in cell lytic buffer and mixed with protein sample buffer before being run on a protein gel. The proteins were separated by SDS-PAGE and a western blot for Zta performed. The antibodies used here were BZ1 and anti-actin.

U2OS cell transfected with hisZta expression vector were demonstrated to have hisZta wild type and mutant protein expression (Figure 5.6). These cells were used to create a protein extract. As detailed in Chapter 3, Cell Lytic reagent including phosphatase inhibitors and benzonase were used for extraction. The supernatant was then taken and applied to the ÄKTA purifier connected to the equilibrated Superose 6 10/300GL column. The 0.5ml elution fractions from the column were then combined afterwards, every four fractions totalling to 2mls. These fractions were subjected to acetone precipitation overnight at -20^{c} before being centrifuged and the protein pellet resuspended in protein sample buffer. The elution profile of hisZta and hisZtaAAA was processed as a western blot (Figure 5.7). The first fraction hisZta eluted into was B5-B9 at 8 to 10mls and eluted into consecutive fractions until C9-C12 at 16.5ml to 18.5ml. The peak actin eluted between C9-C12. The elution profile of hisZta across multiple fractions suggests that Zta is forming various complexes within the cell of different molecular weights ranging between 1MDa and 43kDa. The function and components of these complexes is unknown but as the cells here are U2OS cells that are EBV negative, the complexes may be linked to transcription rather than replication.

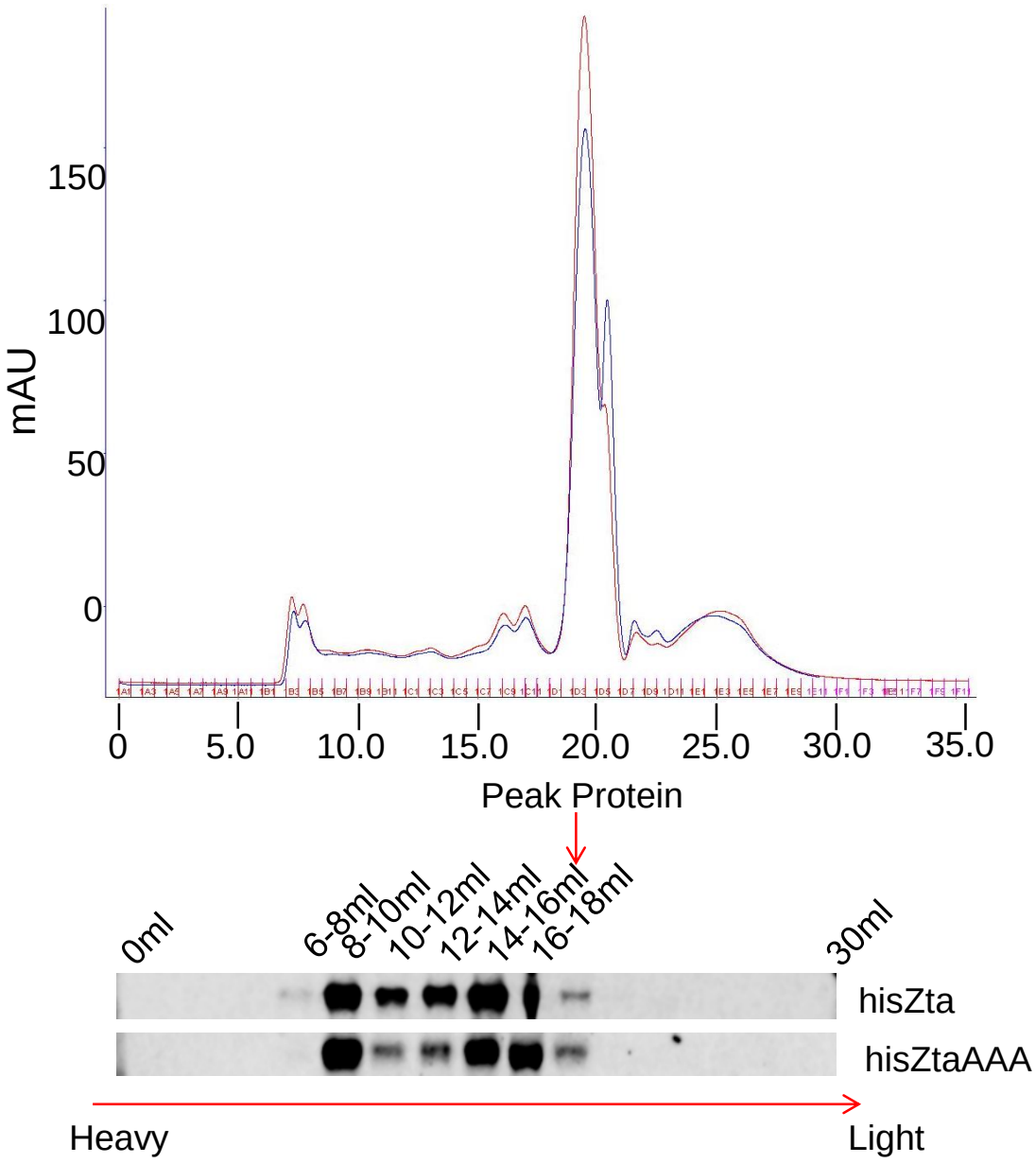


Figure 5.7 Elution profile of hisZta and hisZtaAAA transfected into U2OS cells. U2OS cells were transfected with hisZta or hisZtaAAA and harvested after 72 hours. Cell extract was applied to the Superose 6 column and 0.5ml fractions were eluted. Every 4 fractions were combined and subjected to acetone precipitation before being resuspended in sample buffer. These proteins in sample buffer were then separated by SDS-PAGE and detected by western blot for where hisZta eluted. Heavier molecular weight fractions eluted first and lighter molecular weight fractions later. The antibody used here was BZ1.

A fraction of the proteins were separated by SDS-PAGE. The mAU profile of both cell lysates as an overlap is shown (Figure 5.7). Both hisZta and hisZtaAAA cell lysates follow a near identical mAU elution profile. The first proteins detected are eluted in the same fractions and the peak protein also elutes in the same fraction.

This data demonstrates that hisZta elutes into multiple fractions, suggesting hisZta forms multiple protein complexes of varying molecular weights. HisZtaAAA follows a similar elution profile here and may be forming similar molecular complexes to hisZta.

5.2.4. Determination of the elution profile of Zta in HEK293-BZLF1-KO cells

EBV positive cells (HEK293-*BZLF1*-KO) that contain the EBV genome without a functional *BZLF1* gene were transfected with hisZta or hisZtaAAA. Only hisZta can activate the full lytic cycle, hisZtaAAA cannot activate full lytic cycle.

HEK293-*BZLF1*-KO cells were transfected with control vector, hisZta or hisZtaAAA expression vector and harvested after 96 hours to allow the reactivation of EBV from latency. The proteins separated by SDS-PAGE analysis on a fraction of the cells to detect protein expression hisZta and hisZtaAAA (Figure 5.8). Both hisZta and hisZtaAAA were expressed readily and BMRF1 was expressed in both sets of transfections. Cells from this transfection were taken forward for further size exclusion analysis.

The same protocol from U2OS cells was followed to look at the elution profile hisZta and hisZtaAAA in HEK293-*BZLF1*-KO cells, hisZta then hisZtaAAA run through the column directly afterwards. This would enable a comparison of the elution between the wild type and mutant proteins of hisZta and determine if there are any differences in molecular complexes formed, associated with Zta or ZtaAAA.

Figure 5.8 Expression of transfected hisZta and hisZtaAAA in HEK293-BZLF1-KO cells. HEK293-BZLF1-KO cells were transfected with control vector, hisZta or hisZtaAAA vector and harvested after 96 hours. Cells were lysed in cell lytic reagent and proteins were separated by SDS-PAGE. Western blot was performed to detect hisZta and hisZtaAAA expression, and actin and BMRF1. The antibodies used here were BZ1, anti-BMRF1 and anti-actin.

The transfected HEK293-*BZLF1*-KO cell pellets for hisZta and hisZtaAAA were lysed in cell lytic reagent as before and the supernatant applied to the Superose 6 column before being collected into 500µl fractions as previous. The 500µl fractions were pooled into 2ml and acetone precipitation performed before the protein pellet was resuspended in sample buffer. A fraction of the proteins were separated by SDS-PAGE. Figure 5.9 shows the mAU profile of both cell lysates overlapping. Both hisZta and hisZtaAAA cell lysates follow a near identical mAU elution profile. The first proteins detected are eluted in the same fractions and the peak protein also elutes in the same fraction. From fraction 10-12ml the elution profile is delayed slightly for hisZtaAAA.

The western blot demonstrates that hisZta transfected into HEK293-*BZLF1*-KO cells displays a similar profile of elution in U2OS cells but Zta eluted into a more prominent extra higher molecular weight complex between 6-8ml. This prominence was not present within U2OS cells. Also hisZtaAAA does not elute with the same extra complex. The hisZtaAAA elution is concentrated into a heavier molecular weight fraction.

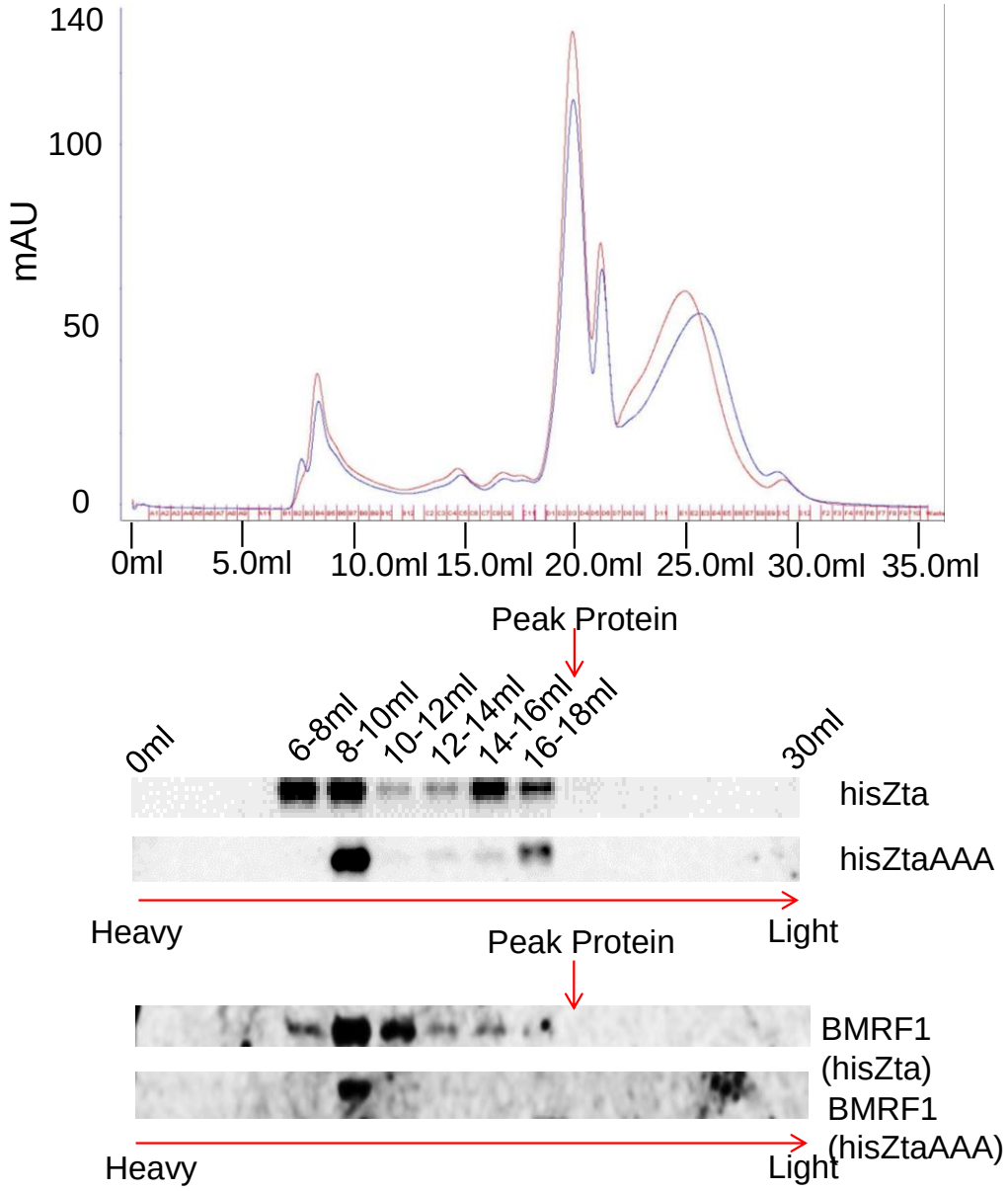


Figure 5.9 Elution profile of hisZta and hisZtaAAA transfected into HEK293-BZLF1-KO cells
HEK293-BZLF1-KO cells were transfected with control vector, hisZta or hisZtaAAA vector and harvested after 96 hours. Cell extract was applied to the Superose 6 column and 0.5ml fractions were eluted. Every 4 fractions were combined and subjected to acetone precipitation before being resuspended in sample buffer. The proteins were then separated by SDS-PAGE and hisZta, hisZtaAAA and BMRF1 were detected by western blot. Antibodies used were BZ1, anti-actin and anti-BMRF1.

BMRF1 was expressed herein in both hisZta and hisZtaAAA transfected cells, and as this protein is a replication factor the western blots were probed for this protein to see what fraction the protein elutes into. BMRF1 was detected in the heavier molecular weight fraction with most of the protein concentrated in the second and third fractions for hisZta transfected cells. BMRF1 was only detected prominently in one fraction for hisZtaAAA transfection here.

It is interesting to see the heavier molecular weight fraction in cells undergoing lytic cycle. This fraction elutes earlier than the largest molecular weight standard, indicating that this fraction may contain a very high molecular weight complex. The proteins contained within this elution fraction may contribute towards lytic replication of the virus

5.2.5. A more detailed fractionation of molecular complexes eluted in HEK293-BZLF1-KO cells

A transfected hisZta HEK293-*BZLF1*-KO cell lysate was separated using the Superose 6 column and elutions were collected into 500µl fractions. These fractions were processed into protein pellets by acetone precipitation and the pellets were resuspended in protein sample buffer. As the region of where Zta elutes has been determined, only a selection of fractions were separated by SDS-PAGE. This was performed as a preliminary investigation to see the elution profile of all proteins by coomassie staining. A coomassie stain of a gel enabled a visualisation of the overall elution profile of the proteins present in the sample (Figure 5.10). This meant that the proteins could be processed as 500µl fractions to have a more detailed insight into where hisZta elutes.

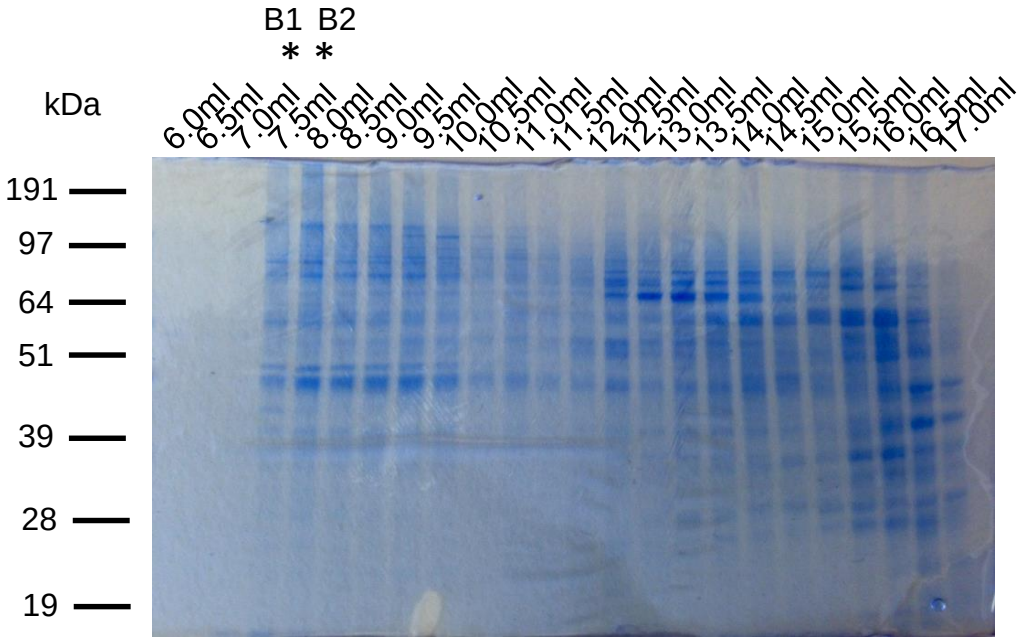


Figure 5.10 Coomassie stain of transfected HEK293-BZLF1-KO cells with hisZta expression vector. Fractions were separated into 500µl and stained with SimplyBlue safestain to visualise the overall elution profile of the proteins present in the sample. The transfected HEK293-*BZLF1*-KO cell pellets for a control, hisZta and hisZtaAAA were lysed in cell lytic reagent as before and the supernatant applied to the Superose 6 column before being collected into 500µl fractions as previously performed. The 500µl fractions were processed into protein pellets by acetone precipitation before being resuspended in protein sample buffer

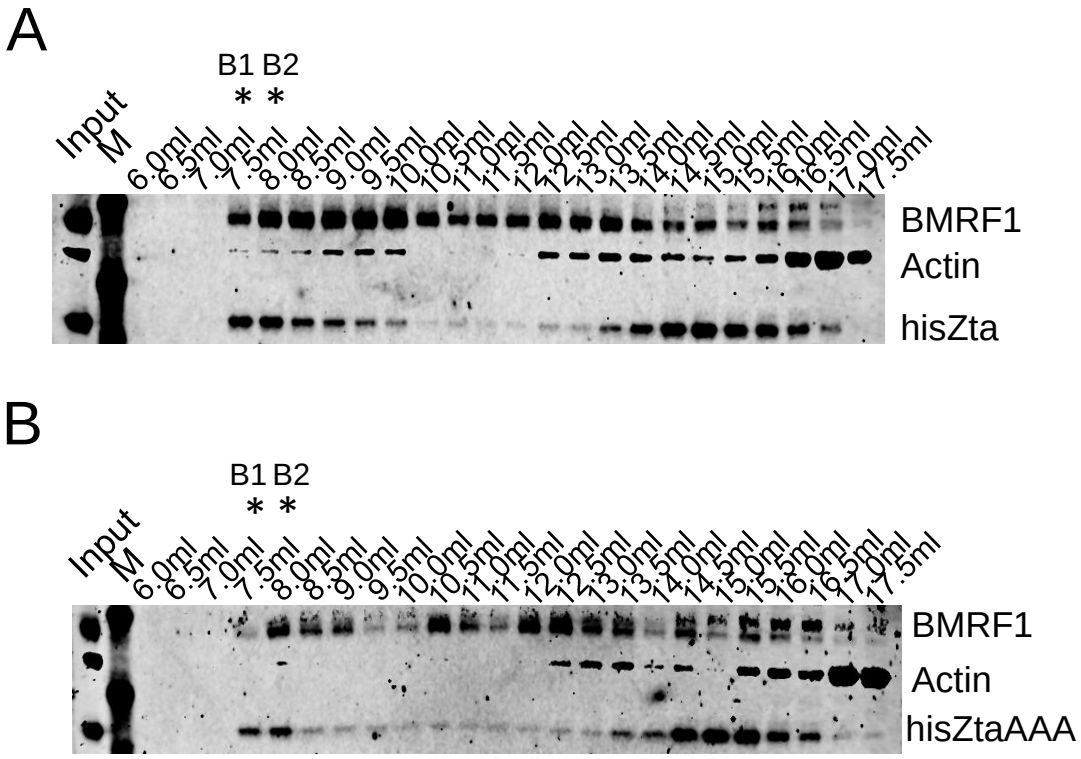


Figure 5.11 Western blot of elution profile of hisZta and hisZtaAAA in transfected HEK293-BZLF1-KO cells. HEK293-BZLF1-KO cells were transfected with **A** hisZta or **B** hisZtaAAA vector and harvested after 96 hours. Cell extracts were processed and was applied to the Superose 6 column and 0.5ml fractions were eluted. Acetone precipitation of elution fractions before proteins were resuspended in protein sample buffer. A fraction of these samples were separated by SDS-PAGE before western blot analysis. Antibodies used were BZ1, anti-actin and anti-BMRF1. **A** hisZta transfected cells **B** hisZtaAAA transfected cells

Fractions between A13/6.5ml and C6/18ml were separated by SDS-PAGE (Figure 5.11). The western blot shows the elution profile of the precipitated control, hisZta and hisZtaAAA for Zta, actin and BMRF1 expression.

HisZta itself eluted into two distinct populations of heavier and lighter molecular weight fractions, 7.5ml – 10.0ml and 13.5ml – 17.0ml. The elution was less prominent between these two populations. BMRF1 stayed relatively constant in elutions from heavy to light molecular weight fractions but more prominent between 7.5ml – 10.0ml.

HisZtaAAA eluted into mainly a lighter molecular weight fraction, some protein eluting into two of the heavier molecular weight fractions, 7.5ml – 8.0ml and 13.5ml – 16.5ml. The elution was minimal between these two populations. BMRF1 was more intermittent in elution across all fractions. It was noted that the higher molecular weight region between fraction B1 and B3 was of significance as the abundance level of hisZtaAAA was not as high as hisZta. Also in fraction B1 for the hisZtaAAA sample, BMRF1 is not present at the same level as the wild type sample.

5.2.6. SILAC labelling and determination of the elution profile of Zta in SILAC labelled HEK293-BZLF1-KO cells

To identify the proteins that eluted into the heavier fractions, SILAC proteomics with mass spectrometry was used. By using mass spectrometry on eluted fractions, it would be possible to begin to identify what proteins are present in these fractions. SILAC labelling of these cell lysates will allow a quantitative analysis of proteins between samples of the same elution fractions from mass spectrometry analysis.

HEK293-*BZLF1*-KO cells were labelled with SILAC media for 5 passages before transfection. Control cells were labelled with R0K0 (light) media (Dundee Cell). Transfected hisZta that allowed the initiation of the lytic cycle were labelled with R6K4 (medium) media. Transfected hisZtaAAA that were unable to activate the lytic cycle were labelled with R10K8 (heavy) media. The light media containing

unlabelled arginine and lysine amino acids (R0K0), medium media containing 13C labelled arginine and 2D labelled lysine amino acids (R6K4), the heavy media containing 13C and 15N labelled arginine and 13C and 15N labelled lysine (R10K8). This allows a mass shift of peptides to be detected by mass spectrometry analysis.

As there were three samples with three SILAC labels, the SILAC ratio analysis enables three ways to determine any differences between the samples: hisZta/Control, hisZtaAAA/control and hisZta/hisZtaAAA.

SILAC labelled HEK293-*BZLF1*-KO cells were transfected with control vector, hisZta or hisZtaAAA expression vectors and harvested after 96 hours to allow the reactivation of EBV from latency. Figure 5.12 represents an SDS-PAGE analysis on a fraction of the cells to assess protein expression. HisZta and hisZtaAAA were expressed readily and *BMRF1* was expressed in both sets of transfections. Cells from this transfection were taken forward for further size exclusion analysis.

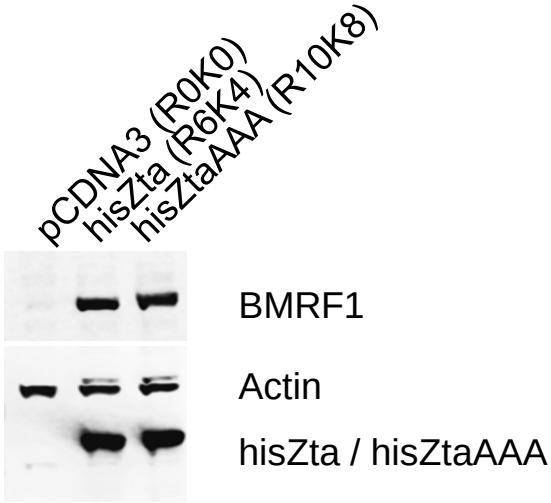


Figure 5.12 SILAC labelling of HEK293-BZLF1-KO cells and detection of transfected hisZta and mutant proteins. HEK293-BZLF1-KO cells were labelled with R0K0 (for control expression), R6K4 (for hisZta expression) and R10K8 (for hisZtaAAA expression). After labelling, the cells were transfected with control vector, hisZta or hisZtaAAA vector and harvested after 96 hours. Cell extracts were processed into cell lytic reagent supernatant and an equal volume of each transfection was mixed together. A fraction of these samples were separated by SDS-PAGE before western blot analysis. Antibodies used were BZ1, anti-actin and anti-BMRF1.

The SILAC labelled transfected HEK293-*BZLF1*-KO cells transfected with control, hisZta and hisZtaAAA were combined equally before being lysed with cell lytic reagent as described previously. A fraction of the combined cell lysates was separated by size exclusion chromatography for further analysis.

The transfected SILAC HEK293-*BZLF1*-KO cell pellets for hisZta and hisZtaAAA were lysed in cell lytic reagent as before and the supernatant applied to the Superose 6 column before being collected into 500µl fractions as previous. The 500µl fractions were processed into protein pellets by acetone precipitation before being resuspended in protein sample buffer.

The proteins eluted in the hisZta sample can be compared against the eluted proteins in the control sample. Any proteins that have a higher ratio could be attributed to the presence of Zta or located in a complex with Zta, or that specific complex is present in the fraction could be due to Zta inducing the lytic cycle.

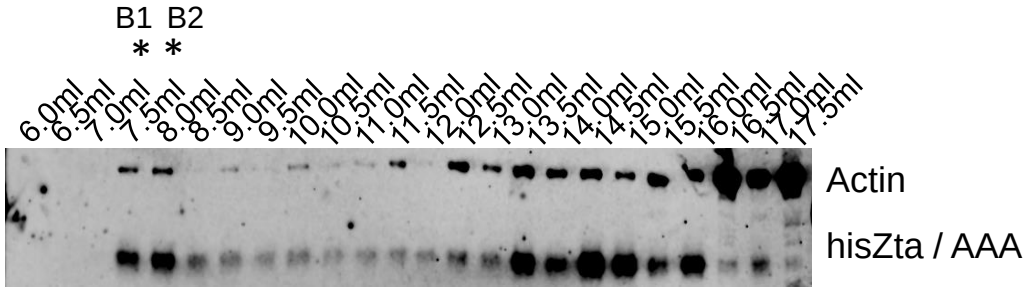


Figure 5.13 SILAC transfected HEK293-BZLF1-KO cell extracts were combined at an equal ratio and separated by FPLC. HEK293-BZLF1-KO cells were labelled with R0K0 (for control expression), R6K4 (for hisZta expression) and R10K8 (for hisZtaAAA expression). After labelling, the cells were transfected with control vector, hisZta or hisZtaAAA vector and harvested after 96 hours. Cell extracts were processed into cell lytic reagent supernatant and an equal volume of each transfection was mixed together. This mixed cell extract was applied to the Superose 6 column and 0.5ml fractions were eluted. Every fraction was subjected to acetone precipitation before being resuspended in sample buffer. The proteins were then separated by SDS-PAGE and detected by western blot for hisZta, hisZtaAAA and actin. Antibodies used were BZ1 and anti-actin.

A fraction of the resuspended SILAC labelled proteins was separated by SDS-PAGE (Figure 5.13). The western blot was probed for Zta and actin proteins. The fractions from Figure 5.11 indicated that elution fractions B1 and B2 were a distinct difference between hisZta and hisZtaAAA. There was more hisZta protein in B1 than hisZtaAAA in B1; and BMRF1 present in the hisZta B1 fraction and hardly any BMRF1 protein in hisZtaAAA B1. It was proposed to follow a mass spectrometry analysis on fraction B1 and B2.

5.2.7. Mass spectrometry of SILAC labelled elutions performed at the University of Sussex

As SILAC labelled fractions B1 and B2 had been previously processed into protein sample buffer, the samples were again separated by SDS-PAGE, along with a BSA control (Figure 5.14). Both fractions B1 and B2 were cut into 5 slices each, with BSA protein band cut into one slice. These bands were processed as described previously into peptides before being analysed by mass spectrometry.

I processed the peptides from the five slices of the B1 and B2 fractions on an LTQ-OrbitrapXL at Sussex University. Maxquant was utilized here. Maxquant can be used as a tool for analysis of mass spectrometry data (Cox et al. 2009). The .RAW output files were processed using Maxquant software based on the parameters entered. This returned a list of proteins formulated from the peptides.

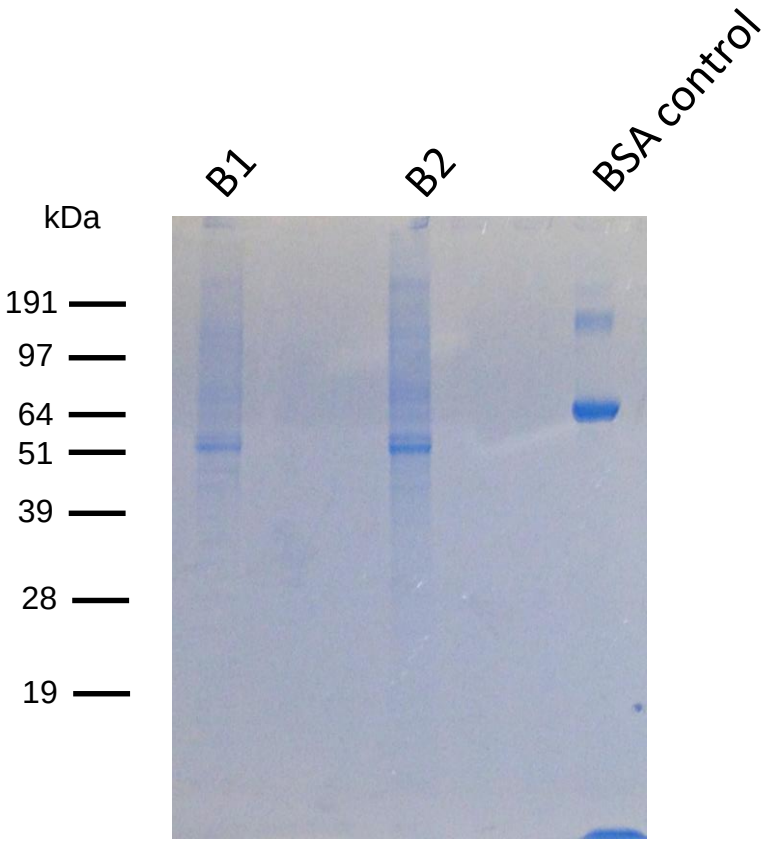


Figure 5.14 Fractions B1 and B2 from FPLC processed for mass spectrometry analysis. 7.5µl of elution B1 in protein sample buffer and 7.5µl of elution B2 in protein sample buffer were loaded onto a protein gel. The gel was stained with SimplyBlue safestain. The lanes containing B1 and B2 were cut into 5 gel pieces each before being processed for mass spectrometry. BSA was also analysed as a single band.

Accession	Protein	Unique peptides	Ratio Zta / Control normalized	Ratio ZtaAAA / Control normalized	Ratio ZtaAAA / Zta normalized
P08670	Vimentin OS	35	1.0406	0.96934	0.90718
P13645	Keratin, type I cytoskeletal 10 OS	22	NaN	NaN	NaN
P35527	Keratin, type I cytoskeletal 9 OS	20	NaN	NaN	NaN
P04264	Keratin, type II cytoskeletal 1 OS	18	NaN	NaN	NaN
P35908	Keratin, type II cytoskeletal 2 epidermal OS	16	NaN	NaN	NaN
P68363	Tubulin alpha-1B chain OS	13	0.87219	0.94791	1.0174
P38159	RNA-binding motif protein	8	0.7805	0.60235	1.1777
P63261	Actin, cytoplasmic 2 OS	6	1.011	1.0771	0.89108
P10809	60 kDa heat shock protein, mitochondrial OS	5	0.90357	1.0431	1.1045
P02533	Keratin, type I cytoskeletal 14 OS	5	NaN	NaN	NaN
>P00761	TRYP_PIG Trypsin - Sus scrofa (Pig).	5	NaN	NaN	NaN
Q5JP53	Tubulin beta chain OS	4	0.69983	0.79312	1.1174
P08107	Heat shock 70 kDa protein 1A/1B OS	4	0.9204	1.3186	1.4421
Q99879	Histone H2B type 1-M OS	3	0.9495	0.98108	1
Q07065	Cytoskeleton-associated protein 4 OS	3	NaN	NaN	NaN
P52272	Heterogeneous nuclear ribonucleoprotein M	3	NaN	NaN	NaN
P68371	Tubulin beta-4B chain OS	2	NaN	NaN	NaN
P13647	Keratin, type II cytoskeletal 5 OS	2	NaN	NaN	NaN
K7EMV3	Histone H3 OS	2	0.82972	0.83334	0.99349
F8W079	ATP synthase subunit beta, mitochondrial (Fragment) OS	2	NaN	NaN	NaN
C9J0D1	Histone H2A OS	2	NaN	NaN	NaN
S4R373	YTH domain-containing family protein 3 OS	1	NaN	NaN	NaN
Q9Y6R1	Electrogenic sodium bicarbonate cotransporter 1 OS	1	NaN	NaN	NaN
Q9BVC6	Transmembrane protein 109 OS	1	NaN	NaN	NaN
Q8AZJ5	LF1_EBVB9 Uncharacterized LF1	1	NaN	NaN	NaN
Q86YZ3	Hornerin	1	NaN	NaN	NaN
Q6ZN08	Putative zinc finger protein 66	1	NaN	NaN	NaN
Q401N2	Zinc-activated ligand-gated ion channel	1	NaN	NaN	NaN
Q04941	Proteolipid protein 2 OS	1	NaN	NaN	NaN
P81605	Dermcidin OS	1	NaN	NaN	NaN
P17066	Heat shock 70 kDa protein 6 OS	1	NaN	NaN	NaN
P07196	eurofilament light polypeptide	1	NaN	NaN	NaN
O75164	Lysine-specific demethylase 4A OS	1	NaN	NaN	NaN
O60282	Kinesin heavy chain isoform 5C OS	1	NaN	NaN	NaN
O43790	Keratin, type II cuticular Hb6 OS	1	NaN	NaN	NaN
H7BZ69	Protein PTHB1 (Fragment) OS	1	NaN	NaN	NaN
H0Y565	Glycerophosphocholine phosphodiesterase GPCPD1 (Fragment) OS	1	NaN	NaN	NaN
F5H6Y5	Nuclear mitotic apparatus protein 1 (Fragment)	1	NaN	NaN	NaN
E9PBF6	Lamin-B1 OS	1	NaN	NaN	NaN
A0A075B6E2	40S ribosomal protein S19 OS	1	NaN	NaN	NaN
>P12763	Alpha-2-HS-glycoprotein precursor	1	NaN	NaN	NaN
F8W0C6	Keratin, type II cytoskeletal 5 (Fragment) OS	0	NaN	NaN	NaN

Table 5.1 Data returned from mass spectrometry performed at University of Sussex of Fraction B1. Fraction B1 was processed into peptides and analysed by MaxQuant. NaN here stands for 'Not a Number', representing an undefined value. A non-quantifiable ratio.

Accession	Protein	Unique peptides	Ratio Zta / control	Ratio ZtaAAA / Control	Ratio ZtaAAA / Zta
P08670	Vimentin OS	33	1.0491	0.97949	0.98118
P13645	Keratin, type I cytoskeletal 10 OS	22	NaN	NaN	NaN
P04264	Keratin, type II cytoskeletal 1 OS	16	NaN	NaN	NaN
P35527	Keratin, type I cytoskeletal 9 OS	16	NaN	NaN	NaN
P35908	Keratin, type II cytoskeletal 2 epidermal OS	14	NaN	NaN	NaN
Q13813	Spectrin alpha chain, non-erythrocytic 1	14	1.0914	1.1083	1.1775
P08107	Heat shock 70 kDa protein 1A/1B OS	6	0.95245	1.2446	1.2158
P10809	60 kDa heat shock protein, mitochondrial OS	6	0.94536	1.0168	0.97957
P06576	ATP synthase subunit beta, mitochondrial OS	5	0.88325	1.1863	1.4786
>P00761		5	NaN	NaN	NaN
Q5JP53	Tubulin beta chain OS	4	0.76114	0.96369	1.0442
E7EMV2	Neurofilament medium polypeptide	4	NaN	NaN	NaN
Q01082	Spectrin beta chain, non-erythrocytic 1 OS	4	NaN	NaN	NaN
P68363	Tubulin alpha-1B chain OS	2	0.87177	0.92494	1.0449
P68371	Tubulin beta-4B chain OS	2	NaN	NaN	NaN
Q99880	Histone H2B type 1-L OS	2	0.81351	0.75806	0.89261
P48047	ATP synthase subunit O, mitochondrial OS	2	NaN	NaN	NaN
O75947	ATP synthase subunit d, mitochondrial OS	2	NaN	NaN	NaN
P12270	Nucleoprotein TPR	2	NaN	NaN	NaN
D6RDI7	Coiled-coil alpha-helical rod protein 1 (Fragment) OS	1	NaN	NaN	NaN
Q6P0N6	DST protein	1	NaN	NaN	NaN
P02533	Keratin, type I cytoskeletal 14 OS	1	NaN	NaN	NaN
K7EQH4	ATP synthase subunit alpha, mitochondrial (Fragment) OS	1	NaN	NaN	NaN
Q71U36	Tubulin alpha-1A chain OS	1	NaN	NaN	NaN
M0R0F0	40S ribosomal protein S5 (Fragment) OS	1	NaN	NaN	NaN
Q5VTE0	Putative elongation factor 1-alpha-like 3 OS	1	NaN	NaN	NaN
F5H6Y5	Nuclear mitotic apparatus protein 1 (Fragment)	1	NaN	NaN	NaN
Q5QNZ2	ATP synthase F(0) complex subunit B1, mitochondria	1	NaN	NaN	NaN
E2QRG8	Receptor expression-enhancing protein 5 (Fragment) OS	1	NaN	NaN	NaN
E9PPJ0	Splicing factor 3B subunit 2 OS	1	NaN	NaN	NaN
G3V5X4	Nesprin-2	1	NaN	NaN	NaN
A0A075B785	LisH domain and HEAT repeat-containing protein KIAA1468 OS	1	NaN	NaN	NaN
P07196	Neurofilament light polypeptide OS	1	NaN	NaN	NaN
P62269	40S ribosomal protein S18 OS	1	NaN	NaN	NaN
P81605	Dermcidin OS	1	NaN	NaN	NaN
Q401N2	Zinc-activated ligand-gated ion channel	1	NaN	NaN	NaN
Q4V328	GRIP1-associated protein 1 OS	1	NaN	NaN	NaN

Table 5.2 Data returned from mass spectrometry performed at University of Sussex of Fraction B2. Fraction B2 was processed into peptides and analysed by MaxQuant. NaN here stands for ‘Not a Number’, representing an undefined value. A non-quantifiable ratio.

Table 5.1 and Table 5.2 show the results of all proteins identified from fractions B1 and B2. The majority of proteins identified did not have a SILAC ratio attributed and were abundant proteins such as actin and tubulins, with some heat shock proteins and histone H3 and H2B. The limited numbers of proteins identified here were reviewed. Previously in Chapter 3, the proteomics results from the pull down obtained from the University of Bristol proved to return an abundant list of proteins. The remaining samples for the B1 and B2 fraction were sent to the University of Bristol Proteomics Facility.

5.2.8. Mass spectrometry of SILAC labelled elutions performed at the University of Bristol

The protein samples B1 and B2 were sent to the University of Bristol Proteomics facility for mass spectrometry analysis. These were run separately on a protein gel processed for mass spectrometry by in-gel digestion. The peptides were analysed by an LTQ-Orbitrap mass spectrometer and processed by Proteome Discoverer (Thermo Scientific) software at a false discovery rate (FDR) of 5%.

The data was returned for both fractions B1 and B2. Proteins identified in the fraction B1 were investigated first. A total of 1553 proteins with at least one peptide were identified from fraction B1. The amount of proteins that were identified here were in a much greater number compared to the data that was returned from my analysis at the University of Sussex.

The most abundant proteins identified for fraction B1 are ranked by PSM value is displayed (Table 5.3).

Accession	Description	# Total PSMs
P08670	Vimentin	249
F8VUJ7	Tubulin beta chain	135
P68371	Tubulin beta-4B chain	128
Q13813	Spectrin alpha chain, non-erythrocytic 1	124
Q9BVA1	Tubulin beta-2B chain	111
Q13885	Tubulin beta-2A chain	109
P08107	Heat shock 70 kDa protein 1A/1B	88
B2RBD5	cDNA, FLJ95457, highly similar to Homo sapiens tubulin, beta, 4 (TUBB4), mRNA	81
Q01082	Spectrin beta chain, non-erythrocytic 1	76
Q07065	Cytoskeleton-associated protein 4	71
P35908	Keratin, type II cytoskeletal 2 epidermal	68
P12270	Nucleoprotein TPR	65
B4DVQ0	cDNA FLJ58286, highly similar to Actin, cytoplasmic 2	63
H6VRF8	Keratin 1	63
H6VRG2	Keratin 1	62
A5YM63	NEFM protein	58
P13645	Keratin, type I cytoskeletal 10	58
P20700	Lamin-B1	57
Q0QEN7	ATP synthase subunit beta (Fragment)	56
Q86UP2	Kinectin	55
E9PKE3	Heat shock cognate 71 kDa protein	55
B2RCQ9	cDNA, FLJ96225, highly similar to Homo sapiens heat shock 70kDa protein 1-like (HSPA1L), mRNA	50

Table 5.3 Most abundant proteins identified by mass spectrometry analysis from fraction B1 at the University of Bristol. Proteins were ordered by the number of peptide spectrum matches. Heat shock proteins and tubulins were most abundantly identified.

Vimentin was the most abundant protein identified in fraction B1 with a PSM value of 249. Many structural proteins were identified including tubulins, and cytoskeletal-associated proteins.

Three comparisons were made between the proteins identified because three SILAC labels were used to label control, hisZta and hisZtaAAA transfected cells. The hisZta sample can be compared to control; this can show the abundance of proteins that are present only because of hisZta or the presence of lytic cycle. The hisZtaAAA sample can be compared to control this can show the abundance of proteins that are present only because of hisZtaAAA or onset of lytic cycle that cannot replicate the viral genome. The third comparison includes hisZta compared to hisZtaAAA so the abundance of proteins present here are specific to lytic cycle and the presence of viral genome replication in the cell.

The proteins identified in fraction B1 were ranked by SILAC ratio comparing hisZta and control sample (Table 5.4). They were identified with one or more SILAC peptides attributed to the protein. Proteins were identified with a SILAC ratio of 100.0 to 1.0. Proteins identified with a SILAC ratio between 2.0 and 100.0 were ranked by SILAC ratio (Table 5.4A). Proteins identified with a SILAC ratio between 1.0 and 2.0 were ranked by SILAC ratio (Table 5.4B).

A

Accession	Description	Zta/ Control	Zta/ Control Count
H0YEX1	Cytoplasmic dynein 2 heavy chain 1 (Fragment)	100.00	1
H7BYV1	Interferon-induced transmembrane protein 2 (Fragment)	47.78	1
J3QRU4	Vesicle-associated membrane protein 2	42.61	1
H7C4M4	Protein HEG homolog 1 (Fragment)	12.79	1
E7EU96	Casein kinase II subunit alpha	4.88	2
Q9BWJ5	Splicing factor 3B subunit 5	3.32	2
A8KA50	cDNA FLJ78617 (Fragment)	2.70	4
H0YJ40	SRA stem-loop-interacting RNA-binding protein, mitochondrial (Fragment)	2.36	2

B

Accession	Description	Zta/ Control	Zta/ Control Count
Q59HB8	Nuclear mitotic apparatus protein 1 variant (Fragment)	1.85	1
A6NDZ3	SAM domain and HD domain-containing protein 1	1.73	1
B3KYB6	Phosphatidylinositol transfer protein beta isoform (Fragment)	1.50	1
Q14151	Scaffold attachment factor B2	1.28	1
Q8WVW9	Similar to signal recognition particle 9kD (Fragment)	1.24	1
Q96NC0	Zinc finger matrin-type protein 2	1.23	1
Q96CS3	FAS-associated factor 2	1.14	1
Q4LE61	SYMPK variant protein (Fragment)	1.12	1
Q86TR6	Full-length cDNA 5-PRIME end of clone CS0DF014YL24 of Fetal brain of Homo sapiens (human) (Fragment)	1.00	2

Table 5.4 Proteins identified from B1 mass spectrometry analysis from the University of Bristol showing the difference in abundance of hisZta and control. Proteins returned that had a SILAC ratio from 1.00 to 100.00, and with greater than or equal to one SILAC peptide between hisZta and Control. **A** Proteins with a SILAC ratio greater than 2.0 **B** Proteins with a SILAC ratio greater than 1.0 to 2.0

The protein with the highest SILAC ratio was cytoplasmic dynein 2 heavy chain 1. This was only identified in in the Zta fraction when compared to the control, therefore the SILAC ratio was 100.0 Proteins with more than one SILAC peptides identified and a high SILAC ratio included Casein kinase II alpha, with a ratio of 4.88, Splicing factor 3B subunit alpha, with a ratio of 3.32, cDNA fragment FLJ78617 (53BP1) with a ratio of 2.70 and SRA stem-loop-interacting RNA binding protein, mitochondrial, with a ratio of 2.36.

The proteins identified in fraction B1 were ranked by SILAC ratio between hisZtaAAA and control sample (Table 5.5). They were identified with one or more SILAC peptides attributed to the protein. Proteins were identified with a SILAC ratio of 100.0 to 1.0. Proteins identified with a SILAC ratio between 2.0 and 100.0 were ranked by SILAC ratio (Table 5.5A). Proteins identified with a SILAC ratio between 1.0 and 2.0 were ranked by SILAC ratio (Table 5.5B). Splicing factor 3B subunit 5 was identified with the highest difference and this was only identified in the hisZtaAAA sample (SILAC ratio 100.0).

Some of the same proteins were identified as seen in Table 5.4. Cytoplasmic dynein 2 heavy chain 1 was identified with a SILAC ratio of 25.67, Casein kinase II alpha, with a ratio of 3.42 and SRA stem-loop-interacting RNA binding protein, mitochondrial, with a ratio of 2.36.

The proteins identified in fraction B1 were ranked by SILAC ratio between hisZta and hisZtaAAA sample (Table 5.6). They were identified with one or more SILAC peptides attributed to the protein. Proteins were identified with a SILAC ratio of 100.0 to 1.0. Proteins identified with a SILAC ratio between 100.0 and 10.0 were ranked by SILAC ratio here.

A

Accession	Description	ZtaAAA/ Control	ZtaAAA/ Control Count
Q9BWJ5	Splicing factor 3B subunit 5	100.00	1
H0YEX1	Cytoplasmic dynein 2 heavy chain 1 (Fragment)	58.67	1
H0YJ40	SRA stem-loop-interacting RNA-binding protein, mitochondrial (Fragment)	4.83	2
E7EU96	Casein kinase II subunit alpha	3.42	2

B

Accession	Description	ZtaAAA/ Control	ZtaAAA/ Control Count
A6NDZ3	SAM domain and HD domain-containing protein 1	1.14	1
Q969I0	KRT8 protein (Fragment)	1.08	2
H7BYV1	Interferon-induced transmembrane protein 2 (Fragment)	1.00	1
J3QRU4	Vesicle-associated membrane protein 2	1.00	1
H7C4M4	Protein HEG homolog 1 (Fragment)	1.00	1

Table 5.5 Proteins identified from B1 mass spectrometry analysis from the University of Bristol showing the difference in abundance of hisZtaAAA and control. Proteins returned that had a SILAC ratio from 1.00 to 100.00, and with greater than or equal to one SILAC peptide between hisZtaAAA and Control. **A** Proteins with a SILAC ratio greater than 2.0 **B** Proteins with a SILAC ratio greater than 1.0 to 2.0.

Accession	Description	Zta/ ZtaAAA	Zta/ ZtaAAA Count
D6RAX2	C-terminal-binding protein 1 (Fragment)	90.59	1
B4DJA5	Ras-related protein Rab-5A	69.63	1
B4DWS6	cDNA FLJ61181, highly similar to Homo sapiens hydroxysteroid (17-beta) dehydrogenase 12 (HSD17B12), mRNA	61.25	1
Q04941	Proteolipid protein 2	55.97	1
Q15717	ELAV-like protein 1	52.41	1
J3QRU4	Vesicle-associated membrane protein 2	42.61	1
Q96HX3	Similar to ribophorin I (Fragment)	28.74	4
B4DYW3	cDNA FLJ56752, highly similar to 28S ribosomal protein S15, mitochondrial (S15mt)	24.60	1
Q53GE7	Tetratricopeptide repeat domain 1 variant (Fragment)	24.38	1
J3QSX4	Mitotic checkpoint protein BUB3	20.44	1
Q9UI09	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 12	14.73	1
B7Z2F4	T-complex protein 1 subunit delta	13.49	2
H7BYV1	Interferon-induced transmembrane protein 2 (Fragment)	13.01	2
H7C4M4	Protein HEG homolog 1 (Fragment)	12.79	1
B4DR63	26S protease regulatory subunit 4	12.11	2
B4E2C0	Secreted glypican-4	10.41	1
Q9NQ51	Putative ATPases	10.39	1
Q86TR6	Full-length cDNA 5-PRIME end of clone CS0DF014YL24 of Fetal brain of Homo sapiens (human) (Fragment)	10.00	2
Q6FG43	FLOT2 protein	10.00	2

Table 5.6 Proteins identified from B1 mass spectrometry analysis from the University of Bristol showing the difference in abundance of hisZta and hisZtaAAA. Proteins returned that had a SILAC ratio from 1.00 to 100.00, and with greater than one or equal to one SILAC peptide between hisZta and hisZtaAAA. SILAC ratios here were reversed by 1÷(hisZtaAAA/hisZta).

The original data returned from the analysis included hisZtaAAA/hisZta, or the heavy/medium sample. To analyse the proteins in abundance with reference to hisZta to give a hisZta/hisZtaAAA the SILAC ratios returned were divided into one. (1 / hisZtaAAA/hisZta).

C-terminal-binding protein 1 (fragment) had the highest ratio of 90.59. Proteins with a peptide count greater than one included a protein similar to ribophorin 1 (fragment) with a ratio of 28.74. T-complex protein 1 subunit delta with a ratio of 13.49. Interferon-induced transmembrane protein 2 (fragment) with a ratio of 13.01 and 26S protease regulatory subunit 4 with a ratio of 12.11.

Accession	Description	# Total PSMs
Q13813	Spectrin alpha chain, non-erythrocytic 1	313
P08670	Vimentin	293
Q01082	Spectrin beta chain, non-erythrocytic 1	209
P12270	Nucleoprotein TPR	176
B4DTV8	cDNA FLJ61399, highly similar to Spectrin alpha chain, brain	164
F8VUJ7	Tubulin beta chain	151
Q9UG16	Putative uncharacterized protein DKFZp564P0562 (Fragment)	144
P68371	Tubulin beta-4B chain	140
Q9BVA1	Tubulin beta-2B chain	132
Q13885	Tubulin beta-2A chain	129
P04350	Tubulin beta-4A chain	110
P08107	Heat shock 70 kDa protein 1A/1B	107
Q13509	Tubulin beta-3 chain	90
P20700	Lamin-B1	87
A5YM63	NEFM protein	86
Q15075	Early endosome antigen 1	82
P11142	Heat shock cognate 71 kDa protein	81
G3V1U9	Tubulin alpha-1A chain	72
Q14980	Nuclear mitotic apparatus protein 1	70
B3KPS3	cDNA FLJ32131 fis, clone PEBLM2000267, highly similar to Tubulin alpha-ubiquitous chain	70
Q07065	Cytoskeleton-associated protein 4	64
B3GQS7	Mitochondrial heat shock 60kD protein 1 variant 1	60
Q0QEN7	ATP synthase subunit beta (Fragment)	56
Q59HB8	Nuclear mitotic apparatus protein 1 variant (Fragment)	53
Q86UP2	Kinectin	53
B7Z4V2	cDNA FLJ51907, highly similar to Stress-70 protein, mitochondrial	52
B2RCQ9	cDNA, FLJ96225, highly similar to Homo sapiens heat shock 70kDa protein 1-like (HSPA1L), mRNA	52

Table 5.7 Most abundant proteins identified by mass spectrometry analysis from sample B2 from the University of Bristol. Proteins were ordered by the number of peptide spectrum matches. Heat shock proteins and tubulins were most richly identified.

The data was returned and the most abundant proteins identified for fraction B2 ranked by PSM value is displayed (Table 5.7).

Spectrin alpha chain, non-erythrocytic 1 was the most abundant protein identified in fraction B2 with a PSM value of 313. Vimentin was again identified like fraction B1, here the second most abundant protein with 293 PSMs. Many structural proteins were identified including tubulins, and cytoskeletal-associated proteins.

Again**,** three comparisons were made between the proteins identified because three SILAC labels were used to label control, hisZta and hisZtaAAA transfected cells. The hisZta sample can be compared to control; this can show the abundance of proteins that are present only because of hisZta or the presence of lytic cycle. The hisZtaAAA sample can be compared to control this can show the abundance of proteins that are present only because of hisZtaAAA or the initiation of lytic cycle without viral replication. The third comparison includes hisZta compared to hisZtaAAA so the abundance of proteins present are specific to the presence of viral replication.

The proteins identified in fraction B2 were ranked by SILAC ratio comparing hisZta and control sample (Table 5.8). They were identified with one or more SILAC peptides attributed to the protein. Proteins were identified with a SILAC ratio of 100.0 to 1.0. Proteins identified with a SILAC ratio between 2.0 and 100.0 were ranked by SILAC ratio (Table 5.8A). Proteins identified with a SILAC ratio between 1.0 and 2.0 were ranked by SILAC ratio (Table 5.8B).

A

Accession	Description	Zta/ Control	Zta/ Control Count
Q5T851	Novel protein (Fragment)	100.00	1
Q15542	Transcription initiation factor TFIID subunit 5	100.00	1
O75197	Low-density lipoprotein receptor-related protein 5	100.00	1
Q9BQN1	Protein FAM83C	46.46	1
O00264	Membrane-associated progesterone receptor component 1	24.70	1
E7EU96	Casein kinase II subunit alpha	24.56	1
E5RHG8	Transcription elongation factor B polypeptide 1 (Fragment)	21.22	1
H7BZ50	Mitotic-spindle organizing protein 2B (Fragment)	11.73	1
B4DMA8	cDNA FLJ53841, highly similar to Kinesin-like protein KIFC1	7.66	2
Q96GX3	KIAA0118 protein (Fragment)	6.89	1
B2REB1	Spastic ataxia of Charlevoix-Saguenay (Sacsin) (Fragment)	5.69	2
P62304	Small nuclear ribonucleoprotein E	5.27	2
O96011	Peroxisomal membrane protein 11B	2.68	2
C9J6L4	Probable serine carboxypeptidase CPVL (Fragment)	2.54	2

B

Accession	Description	Zta/ Control	Zta/ Control Count
Q9HCD5	Nuclear receptor coactivator 5	1.25	1
B4DUG4	cDNA FLJ51308	1.13	1
F5GWR1	Coiled-coil domain-containing protein 91 (Fragment)	1.10	1
B4DMJ8	cDNA FLJ52116, highly similar to Rho-related GTP-binding protein RhoB	1.06	1
B3KPC1	cDNA FLJ31582 fis, clone NT2RI2002117, highly similar to Protein pelota homolog	1.04	1
H0YDB2	Stromal interaction molecule 1 (Fragment)	1.02	1
Q5CZB5	Putative uncharacterized protein DKFZp686M0430	1.00	2

Table 5.8 Proteins identified from B2 mass spectrometry analysis from the University of Bristol showing the difference in abundance of hisZta and control. Proteins returned that had a SILAC ratio from 1.00 to 100.00, and with greater than or equal to one SILAC peptide between hisZta and Control. **A** Proteins with a SILAC ratio greater than 2.0 **B** Proteins with a SILAC ratio greater than 1.0 to 2.0

A protein with 100.0 SILAC ratio included a novel protein (Q5T851) identified as Poly (A) RNA polymerase, mitochondrial, TAF5, and Low-density lipoprotein receptor-related protein 5. These were only identified in the hisZta sample compared to the control.

Proteins with more than one SILAC peptides identified and a high SILAC ratio included cDNA FLJ53841 similar to KIFC1, with a ratio of 7.66, Sacsin with a ratio of 5.69, Small nuclear ribonucleoprotein E with a ratio of 5.27, Peroxisomal membrane protein 11B with a ratio of 2.68, Probable serine carboxypeptidase CPVL (Fragment), with a ratio of 2.54.

Interestingly, Casein Kinase II alpha was also detected in fraction B2 with a SILAC ratio of 34.56. 53BP1 was not found in this fraction.

Proteins that were identified in fraction B2 were also ranked by SILAC ratio between hisZtaAAA and control sample (Table 5.9). They were identified with one or more SILAC peptides attributed to the protein. Proteins were identified with a SILAC ratio of 100.0 to 1.0. Proteins identified with a SILAC ratio between 2.0 and 100.0 were ranked by SILAC ratio (Table 5.9A). Proteins identified with a SILAC ratio between 1.0 and 2.0 were ranked by SILAC ratio (Table 5.9B).

A

Accession	Description	ZtaAAA /Control	ZtaAAA/ Control Count
O75197	Low-density lipoprotein receptor-related protein 5	100.00	1
E7EU96	Casein kinase II subunit alpha	26.75	1
E9PH24	Glutamate receptor delta-2 subunit	3.54	1
P62304	Small nuclear ribonucleoprotein E	3.13	2

B

Accession	Description	ZtaAAA /Control	ZtaAAA/ Control Count
H7C2B5	Telomere-associated protein RIF1 (Fragment)	1.99	1
D6RAA2	GTP-binding protein SAR1b (Fragment)	1.51	1
Q5JPU0	Pyruvate dehydrogenase (Lipoamide) alpha 1	1.19	1
A0PJ47	SAFB2 protein (Fragment)	1.01	1
Q5T851	Novel protein (Fragment)	1.00	1
Q15542	Transcription initiation factor TFIID subunit 5	1.00	1
Q9BQN1	Protein FAM83C	1.00	1
O00264	Membrane-associated progesterone receptor component 1	1.00	1
E5RHG8	Transcription elongation factor B polypeptide 1 (Fragment)	1.00	1
H7BZ50	Mitotic-spindle organizing protein 2B (Fragment)	1.00	1
Q96GX3	KIAA0118 protein (Fragment)	1.00	1

Table 5.9 Proteins identified from B2 mass spectrometry analysis at the University of Bristol showing the difference in abundance of hisZtaAAA and control. Proteins returned that had a SILAC ratio from 1.00 to 100.00, and with greater than or equal to one SILAC peptide between hisZtaAAA and Control. **A** Proteins with a SILAC ratio greater than 2.0 **B** Proteins with a SILAC ratio greater than 1.0 to 2.0

Proteins with more than one SILAC peptides identified and a high SILAC ratio included only Small nuclear ribonucleoprotein E, with a SILAC ratio of 3.13
A protein with 100.0 SILAC ratio included low-density lipoprotein receptor-related protein 5. Proteins with more than one SILAC peptides identified and a high SILAC ratio included small nuclear ribonucleoprotein E with a ratio of 3.13. Again, Casein Kinase II alpha was detected here with a SILAC ratio of 26.75.

The proteins identified in fraction B2 were ranked by SILAC ratio between hisZta and hisZtaAAA sample (Table 5.10). Proteins were identified with a SILAC ratio of 100.0 to 1.0. Proteins were identified with one or more SILAC peptides attributed to the protein. The original data returned had the ratio orientated with hisZtaAAA / hisZta. In order to show this data with a high ratio representing the abundance of protein due to hisZta, the ratios were divided into 1.

Accession	Description	Zta/ ZtaAAA	Zta/ ZtaAAA Count
Q5T851	Novel protein (Fragment)	100.00	1
Q15542	Transcription initiation factor TFIID subunit 5	100.00	1
Q6MZH3	Putative uncharacterized protein DKFZp686I05132	100.00	1
B7Z462	cDNA FLJ58016, highly similar to Polypeptide N-acetylgalactosaminyltransferase2 (EC 2.4.1.41)	100.00	1
Q108N2	Acyl-CoA synthetase long-chain family member 1 isoform c (Fragment)	100.00	1
B0AZP7	Synaptic vesicle membrane protein VAT-1 homolog	91.01	1
B3KXZ9	cDNA FLJ46477 fis, clone THYMU3025118, highly similar to Cell surface glycoprotein MUC18	71.78	1
F5H7X1	26S proteasome non-ATPase regulatory subunit 9	66.51	1
A8K2R3	cDNA FLJ75083, highly similar to Homo sapiens amine oxidase (flavin containing) domain 2 (A	59.13	1
Q53T09	Putative uncharacterized protein XRCC5 (Fragment)	50.11	1
Q9BQN1	Protein FAM83C	46.46	1
B7ZBQ1	Mediator complex subunit 20 subunit 20Trf (TATA binding protein-related factor)-proximal homolog (Drosophila) (Fragment)	43.96	1
B4DWZ5	Mitochondrial carnitine/acylcarnitine carrier protein	41.81	1
Q9HD33	39S ribosomal protein L47, mitochondrial	38.61	1
O00264	Membrane-associated progesterone receptor component 1	24.70	1
E5RHG8	Transcription elongation factor B polypeptide 1 (Fragment)	21.22	1
B4DMA8	cDNA FLJ53841, highly similar to Kinesin-like protein KIFC1	20.80	2
Q5VV42	Threonylcarbamoyladenosine tRNA methylthiotransferase	19.11	1
B0QYA5	Eukaryotic translation initiation factor 3 subunit D (Fragment)	14.97	1
B3KMS5	cDNA FLJ12482 fis, clone NT2RM1001085, highly similar to CTTNBP2 N-terminal-like protein	14.48	2
Q96S55	ATPase WRNIP1	13.10	2
H7BZ50	Mitotic-spindle organizing protein 2B (Fragment)	11.73	1
P32856	Syntaxin-2	10.63	1
B2REB1	Spastic ataxia of Charlevoix-Saguenay (Sacsin) (Fragment)	10.22	2
F8VRD8	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 12	10.15	2
Q5CZB5	Putative uncharacterized protein DKFZp686M0430	10.00	2
O60762	Dolichol-phosphate mannosyltransferase	10.00	2

Table 5.10 Proteins identified from B2 mass spectrometry analysis at the University of Bristol showing the difference in abundance of hisZta and hisZtaAAA. Proteins returned that had a SILAC ratio from 1.00 to 100.00, and with greater than one or equal to one SILAC peptide between hisZta and hisZtaAAA. SILAC ratios here were reversed by 1÷(hisZtaAAA/hisZta).

The protein with the highest ratio identified with 100.0 was Novel protein (Fragment) identified as Poly(A) RNA polymerase, TAF5, Putative uncharacterized protein DKFZp686I05132, cDNA FLJ58016, highly similar to Polypeptide N-acetylgalactosaminyltransferase2 (EC 2.4.1.41), and Acyl-CoA synthetase long-chain family member 1 isoform c (Fragment).

Proteins with more than one SILAC peptides identified and a high SILAC ratio included: cDNA FLJ53841 similar to KIFC1, with a ratio of 20.80, Eukaryotic translation initiation factor 3 subunit D (Fragment) with a ratio of 14.97, cDNA FLJ12482 fis, clone NT2RM1001085 highly similar to CTTNBP2 N-terminal-like protein with a ratio of 14.48, ATPase WRNIP1 with a ratio of 13.10. Sacsin with a ratio of 10.22, NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 12. Putative uncharacterized protein DKFZp686M0430 and Dolichol-phosphate mannosyltransferase both had a ratio of 10.00

5.2.9. Attempt to determine elution positions of proteins identified from the mass spectrometry proteomics analysis

From the SILAC analysis, casein kinase II alpha is more abundant in fractions B1 and B2 for both hisZta and hisZtaAAA compared to control cell extracts. Casein kinase II alpha was chosen because it changed two to five fold in fraction B1, and twenty-four to twenty-six fold in fraction B2 by both Zta and ZtaAAA. Mitochondrial Poly(A) RNA polymerase (Q5T851) protein was identified in fraction B2 and most abundant in the hisZta transfected elution, and not in the control or hisZtaAAA transfected elution. Mitochondrial Poly(A) RNA polymerase was chosen because it is Zta specific in fraction B2. These two proteins were initially chosen to follow up the elution profile, to attempt to confirm their abundance by western blot analysis.

First, Casein Kinase II alpha was chosen from fraction B1 and its presence in fraction B2. A new transfection of cells were prepared into cell lytic buffer as previous and fractionated by the Superose 6 column separately. Elution fractions from A14 to B4 for control, hisZta and hisZtaAAA were processed by acetone precipitation and resuspended into protein sample buffer. This range was chosen

to cover elution fractions B1 and B2. The proteins were separated by SDS-PAGE. From the SILAC data returned in Table 5.4, casein kinase II alpha had a Zta/control ratio of 4.88. In Table 5.5 casein kinase II alpha had a ZtaAAA/Control ratio of 3.42. There should be more abundance of this protein in fraction B1 in both hisZta and hisZtaAAA transfected fractions compared to the control. From the SILAC data returned in Table 5.8, casein kinase II alpha had a Zta/control ratio of 24.56. In Table 5.9 casein kinase II alpha had a ZtaAAA/Control ratio of 26.75. There should be more abundance of this protein in fraction B2 in both hisZta and hisZtaAAA transfected fractions compared to the control.

A western blot of the eluted fractions was performed (Figure 5.15A). The input for each sample was included and fractions A14-B4 were separated, hisZta and hisZtaAAA were detected, casein kinase II alpha and actin in the input by antibodies. The presence of casein kinase II alpha was detected and two protein bands are present. Both of these bands were investigated as upper and lower because it was not possible to distinguish if casein kinase II alpha was being phosphorylated, resulting in possibly a protein with a higher molecular weight on the gel.

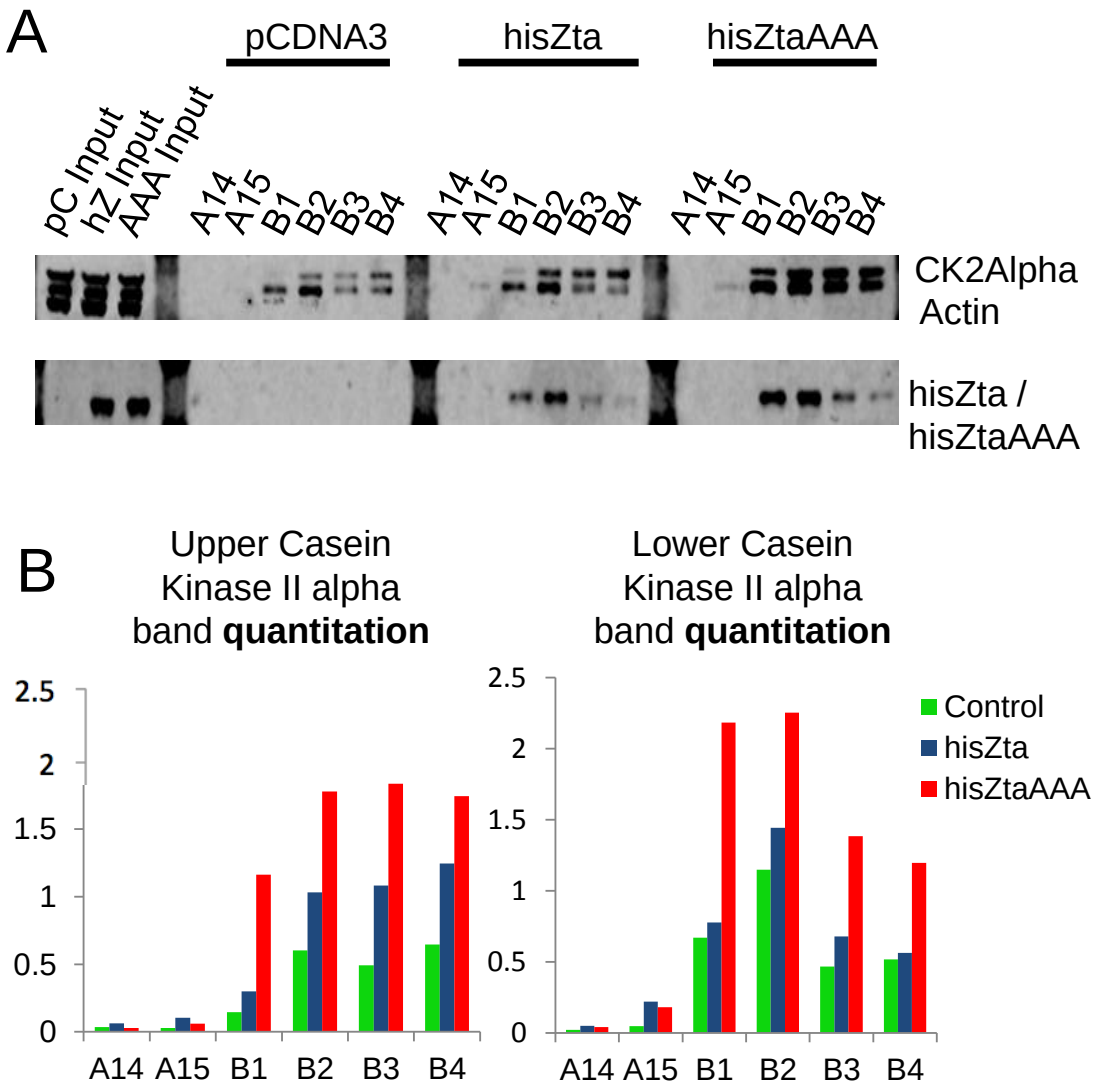


Figure 5.15 CKII alpha abundance in both hisZta and hisZtaAAA fractions between A14 and B4 elutions. A new transfection of HEK293-BZLF1-KO cells were transfected with either control, hisZta or hisZtaAAA expression vectors. These cells were harvested after 96 hours and lysed in cell lytic reagent. Cell extracts were processed and was applied to the Superose 610/300GL column and 0.5ml fractions were eluted. Acetone precipitation of elution fractions before proteins were resuspended in protein sample buffer. **A** A fraction of these samples were separated by SDS-PAGE before western blot analysis. Antibodies used were BZ1 and anti-Casein Kinase II alpha. **B** Upper Casein Kinase II alpha band quantitation. Lower Casein Kinase II alpha band quantitation. Quantitation performed using ImageStudio (Licor).

There is slightly more protein abundance in the hisZta B1 fraction compared to control fraction in fraction B1 for the upper band quantitated (Figure 5.15B). There is more casein kinase II alpha in the hisZtaAAA B1 fraction. Both of these bands agree with the trend mass spectrometry data. There is slightly more protein abundance in the hisZta B1 fraction compared to control fraction in fraction B1 for the lower band quantitated (Figure 5.15B). There is more casein kinase II alpha in the hisZtaAAA B1 fraction. Both of these bands agree with the trend mass spectrometry data.

There is slightly more protein abundance in the hisZta B2 fraction compared to control fraction in fraction B1 for the upper band quantitated (Figure 5.15B). There is more casein kinase II alpha in the hisZtaAAA B2 fraction. Both of these bands agree with the trend mass spectrometry data. There is slightly more protein abundance in the hisZta B2 fraction compared to control fraction in fraction B1 for the lower band quantitated (Figure 5.15B). There is more casein kinase II alpha in the hisZtaAAA B2 fraction. Both of these bands agree with the trend mass spectrometry data.

From the SILAC analysis Poly(A) RNA polymerase (mitochondrial) was also chosen to follow up the elution profile as this protein was present in B2 mass spectrometry analysis identified as novel protein Q5T851. A new transfection of cells were prepared into cell lytic buffer and fractionated by the Superose 6 column separately. Elution fractions from A14 to B4 for control, hisZta and hisZtaAAA were processed by acetone precipitation and resuspended into protein sample buffer. The proteins were separated by SDS-PAGE. From the SILAC data returned (Table 5.8), Poly(A) RNA polymerase (mitochondrial) had a Zta/control ratio of 100.0. The protein was only in the hisZta transfected sample compared to control. Also Poly(A) RNA polymerase (mitochondrial) had a Zta/ZtaAAA ratio of 100.0 (Table 5.10). The protein was only in the hisZta transfected sample compared to hisZtaAAA. Poly(A) RNA polymerase (mitochondrial) is present in all B2 fractions, with more protein abundance in hisZtaAAA B2 (Figure 5.16). Also Poly(A) RNA polymerase (mitochondrial) is only in hisZtaAAA B1 fraction and not in hisZta B1 or control B1 fractions.

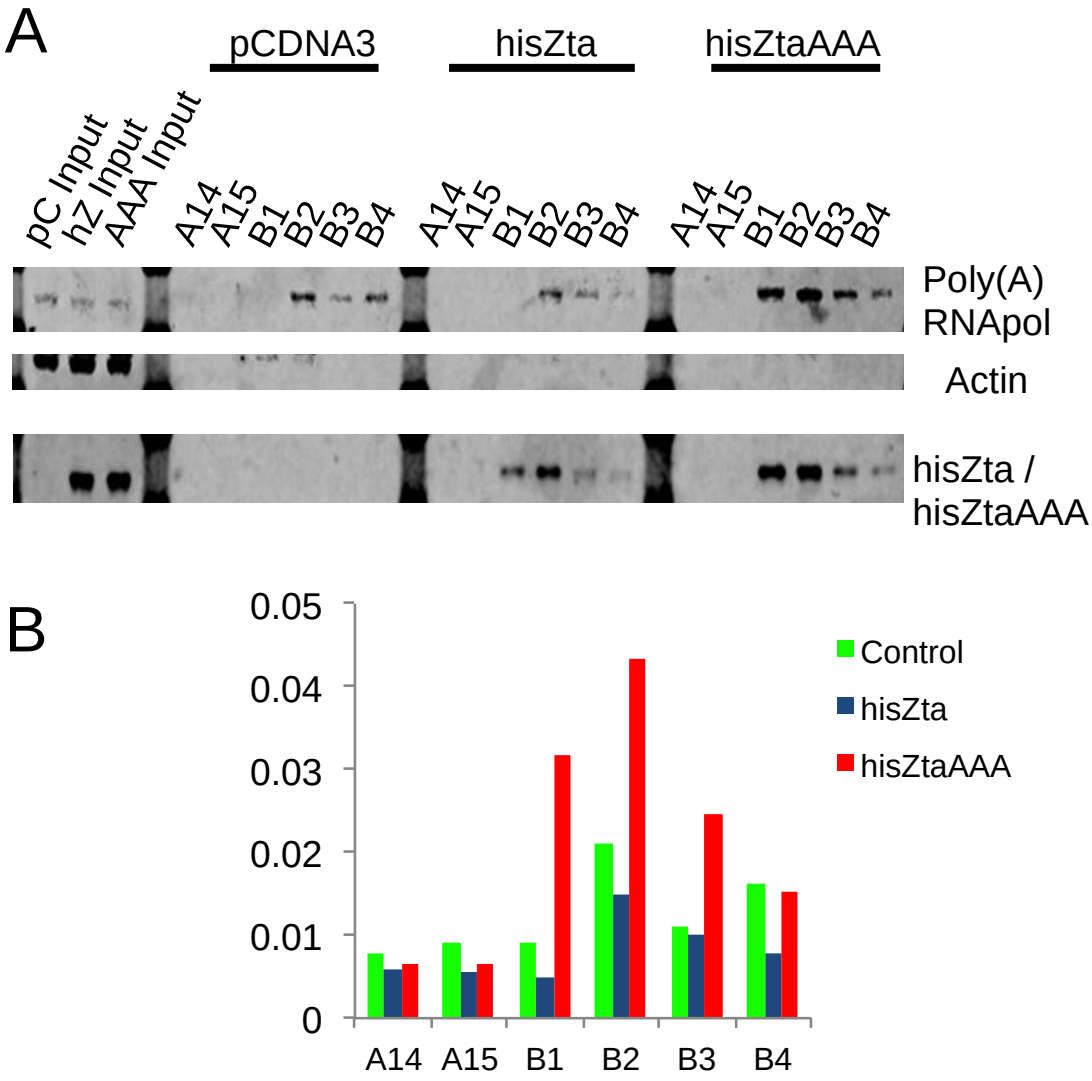


Figure 5.16 Poly(A) RNA polymerase (mitochondrial) abundance in both hisZta and hisZtaAAA fractions between A14 and B4 elutions. A new transfection of HEK293-BZLF1-KO cells were transfected with either control, hisZta or hisZtaAAA expression vectors. These cells were harvested after 96 hours and lysed in cell lytic reagent. Cell extracts were processed and was applied to the Superose 610/300GL column and 0.5ml fractions were eluted. Acetone precipitation of elution fractions before proteins were resuspended in protein sample buffer. **A** A fraction of these samples were separated by SDS-PAGE before western blot analysis. Antibodies used were BZ1 and anti- Poly (A) RNA Polymerase (mitochondrial) **B** Poly (A) RNA Polymerase (mitochondrial) band quantitation. Quantitation performed using ImageStudio (Licor).

A western blot of the eluted fractions was performed (Figure 5.16A). The input for each sample was included and fractions A14-B4 were separated, hisZta and hisZtaAAA were detected and poly(A) RNA polymerase (mitochondrial) in the input by antibodies. The presence of poly(A) RNA polymerase (mitochondrial) was in detected and one protein bands are present.

Poly(A) RNA polymerase (mitochondrial) is not present in control or hisZta B1 fraction, but present in hisZtaAAA B1 (Figure 5.16A). Although there is slightly more protein abundance in the control B1 fraction compared to hisZta fraction in fraction B1 this may be background levels of quantitation (Figure 5.16B). There is more poly(A) RNA polymerase (mitochondrial) in the hisZtaAAA B1 fraction. The protein was chosen for to be investigated based on fraction B2. Poly(A) RNA polymerase (mitochondrial) should only be in the hisZta B2 fraction (Table 5.8a, Table 5.10).

Poly(A) RNA polymerase (mitochondrial) was present in all eluted B2 samples here. There is slightly more protein abundance in the control B2 fraction compared to hisZta fraction in fraction B2 for the lower band quantitated (Figure 5.16B). There is more Poly(A) RNA polymerase (mitochondrial) in the hisZtaAAA B2 fraction. This abundance does not agree with the trend mass spectrometry data.

Further analysis of the mass spectrometry data returned indicated 53BP1 (A8KA50) might be implicated to be changing abundance. The protein was identified and had an increased SILAC ratio of 2.70 in the hisZta B1 fraction compared to control B1 fraction (Table 5.4A). Also the SILAC ratio for 53BP1 (A8KA50) between hisZta and hisZtaAAA was 5.62. Therefore, this protein should be increased in abundance only for hisZta sample.

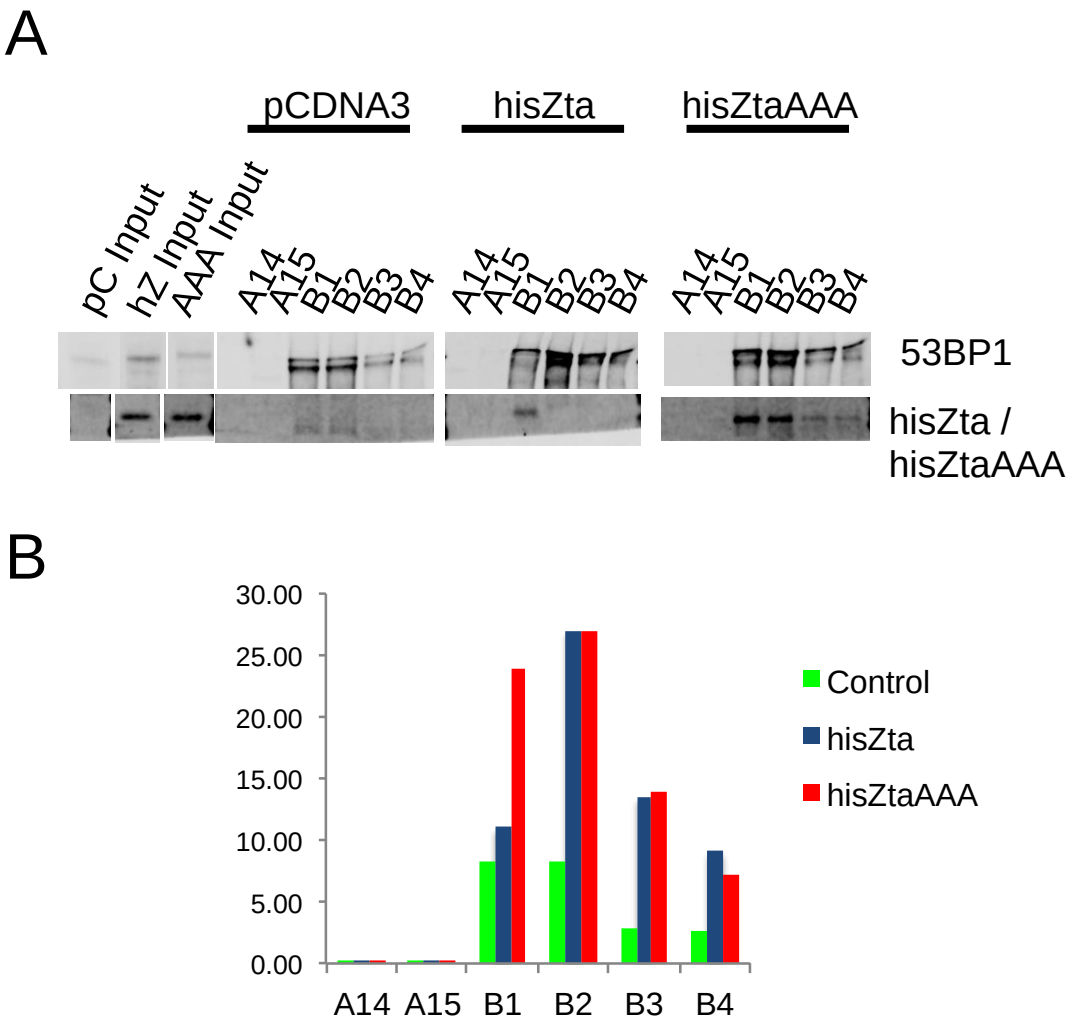


Figure 5.17 53BP1 abundance in both hisZta and hisZtaAAA fractions between A14 and B4 elutions. A new transfection of HEK293-BZLF1-KO cells were transfected with either control, hisZta or hisZtaAAA expression vectors. These cells were harvested after 96 hours and lysed in cell lytic reagent. Cell extracts were processed and was applied to the Superose 610/300GL column and 0.5ml fractions were eluted. Acetone precipitation of elution fractions before proteins were resuspended in protein sample buffer. **A** Afraction of these samples were separated by SDS-PAGE before western blot analysis. Antibodies used were BZ1 and anti- 53BP1. Antibodies used were BZ1 and anti- 53BP1 **B** 53BP1 band quantitation. Quantitation performed using ImageStudio (Licor).

Another size separation of transfected HEK293-BZLF1-KO cells was performed. The elutions were acetone precipitated and resuspended in protein sample buffer before being separated by SDS-PAGE. A western blot of the eluted fractions was performed (Figure 5.17A). The input for each sample was included and fractions A14-B4 were separated, hisZta and hisZtaAAA were detected and 53BP1 in the input by antibodies. The presence of 53BP1 was in detected as a high molecular weight protein. Two distinct bands are present.

53BP1 is present in control, hisZta and hisZtaAAA B1 fraction, (Figure 5.17A). Although there is slightly more protein abundance in the control B1 fraction compared to hisZta fraction in fraction B1 (Figure 5.17B). There is more 53BP1 in the hisZtaAAA B1 fraction. This abundance does not agree with the trend mass spectrometry data.

5.3. Discussion

It was decided to investigate protein complexes associated with Zta that may play a contributing role in EBV lytic replication. This study demonstrates that combined SEC/MS analysis can be used for the analysis of protein complexes and to predict potential interactions.

First, the function of transcription and replication of Zta and Zta mutants were demonstrated. The transcriptional activity of Zta was validated and activated the BHLF1 promoter. The ZtaAAA protein could also activate the BHLF1 promoter. The distal C-terminal region is not required for transactivation function of Zta in vivo (Schelcher et al. 2007). This study demonstrates that the last three amino acids mutated to alanine have no effect on the transactivation of the BHLF1 promoter, while the mutant maintains the ability for transactivation of genes.

The initiation of EBV genome replication was assessed by qPCR. Zta was demonstrated to initiate the lytic cycle in HEK293-*BZLF1*-KO cells, ZtaAAA could not initiate the lytic cycle in these cells (Figure 5.4). Although the early lytic protein, BMRF1 was expressed from cells transfected with Zta or ZtaAAA, therefore ZtaAAA retains its transactivation activity but cannot replicate the viral genome. The deletion of the final three carboxyl-terminal amino acids of Zta can

abolish EBV genome replication in the same cell system (Bailey et al. 2009). The ZtaAAA mutant agrees and follows an identical profile here with the published data, where the viral genome is not initiated for lytic replication. The final three amino acids of Zta are leucine, asparagine and phenylalanine (LNF). The expression vector for ZtaAAA contains three alanine substituted into these positions. Leucine and phenylalanine are non-polar amino acids, whereas asparagine is polar. The X-ray crystal structure of the Zta dimerization and DNA binding domains bound to DNA revealed that Zta amino acids in the C-terminus from Pro223 fold back against the coiled coil (Petosa et al. 2006). Cross-linking experiments demonstrate that the entire C-terminal region lies adjacent to the zipper, amino acids 221 to 230 enhance the stability of the coiled coil (Schelcher et al. 2007). Residues in the zipper region of Zta are essential (McDonald et al. 2009). The ZtaAAA proteins final amino acid region is important for the stability of Zta structure, and the mutant used in this study agrees with the literature here.

The Superose 6 10/300GL size exclusion column function was calibrated before its use. A suitable buffer was chosen and size exclusion molecular weight markers were placed in this buffer and eluted from the column. This enabled molecular weights to be applied to the elution fractions. The native molecular weight determination by gel filtration from cell extracts was performed and the elution profiles determined. Zta and the EBV lytic cycle has not been studied by size exclusion chromatography before. The single stranded DNA binding protein BALF2 has been assessed previously for its function with the viral DNA polymerase by size exclusion chromatography but not in the context if lytic replication (Tsurumi 1993).

A successful elution of hisZta and proteins from cell extracts were stable in the buffer for elution and capable of being precipitated in acetone and resuspended in protein sample buffer. This established suitable gel filtration conditions for Zta cell extracts. A suitable buffer, and acetone precipitation of elutions was developed. Also the elution profile of Zta and cellular proteins was determined. This was repeatable and consistent across multiple transfected cell extracts

Zta elutes in same fractions as U2OS cells, displaying a near identical elution profile in HEK293-*BZLF1*-KO cells, between 8-10ml and 16-18ml. A heavier fraction was evident for hisZta cells undergoing lytic replication, between 6-8mls (Figure 5.9). This region became a region of interest as this was not evident in the hisZtaAAA elution profile. HisZtaAAA did first elute into the 8-10ml fraction with a lower abundance until 16-18ml. An extremely low abundance of Zta can be seen within the 6-8ml elution in U2OS cells, this is more prominent in HEK293-*BZLF1*-KO cells. From the inputs, there was an equivalent of hisZta and hisZtaAAA. Therefore the difference seen in the heaviest fraction may be a lytic cycle difference with virus undergoing lytic replication, the molecular complexes may not be fully formed in EBV negative cells, but in EBV positive cells undergoing lytic replication these proteins are present. Therefore this elution region was investigated further.

A finer elution of the fractions were processed and separated. The one large elution fractions were separated into four fractions each for a greater resolution. The region where Zta first elutes into sees an abundance of Zta (7.5ml, Figure 5.11A). BMRF1 is also present across all fractions, and is prominent from the 7.5ml fraction.

The presence of hisZtaAAA in 7.5ml may be a slight shift into this fraction, as the abundance of BMRF1 is very minimal here (Figure 5.11B). ZtaAAA is prominent in 8.0ml fraction, which agrees with the previous data. The abundance is minimal until the 14ml fraction where the high abundance returns. The first fraction indicated a difference between lytic replicating and non-lytic replicating cells. Therefore fractions B1 and B2 were used as a focus point for distinguishing differences between control, hisZta and hisZta expressed in HEK293-*BZLF1*-KO cells.

HEK293-*BZLF1*-KO cells were labelled with SILAC light, medium and heavy media for control, hisZta and hisZtaAAA transfection respectively. Mass spectrometry of fractions B1 and B2 gave a high data set from the University of Bristol. The mass spectrometry attempt at the University of Sussex could not allow any meaningful conclusions from the data.

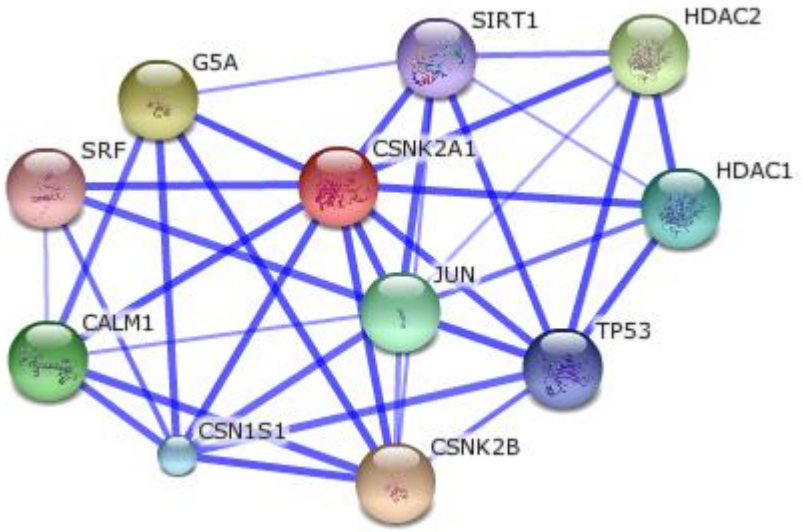
HisZtaAAA that is unable to initiate EBV lytic replication, so this allows another level of analysis. The proteins that are present in the elutions will not be present due to viral replication, and may be present due to interactions with the transactivation domain of Zta, or from the onset of lytic cycle. Some lytic genes can still be expressed by hisZtaAAA. A difference in protein abundance was identified from the control, hisZta and hisZtaAAA samples indicated by a SILAC ration. The abundance changed for some proteins for all conditions.

The proteins present in the elution fraction may not be directly interacting with Zta, but associating with elements related with EBV lytic replication. These may be forming protein complexes or associating with proteins interacting with Zta. Proteins were investigated for having a high SILAC ratio in the hisZta sample and a low SILAC ratio in the control and hisZtaAAA sample. For fraction B1, these were casein kinase II alpha, 53BP1, which were found in high abundance in the hisZta sample against the control. Casein kinase II alpha also had a high ratio for hisZtaAAA against control sample.

For fraction B2, mitochondrial Poly(A) RNA polymerase was only detected in the hisZta fraction. Casein kinase II alpha was also shown to have a high abundance in the hisZta and hisZtaAAA fractions, as seen in fraction B1. Although the protein abundance was much higher in fraction B2.

An attempt to determine elution positions of the proteins identified from the mass spectrometry analysis was made. The western blot results trying to assess the mass spectrometry data were mixed, as the abundance of the target proteins were quantitated and there was some differences from the mass spectrometry analysis. These are outlined below.

Casein kinase II alpha was assessed first. There is slightly more casein kinase II alpha abundance in the hisZta B1 fraction compared to control fraction in fraction B1 for the upper band quantitated (Figure 5.15B). There is more casein kinase II alpha in the hisZtaAAA B1 fraction. Both of these bands agree with the trend mass spectrometry data.



Predicted Functional Partners:

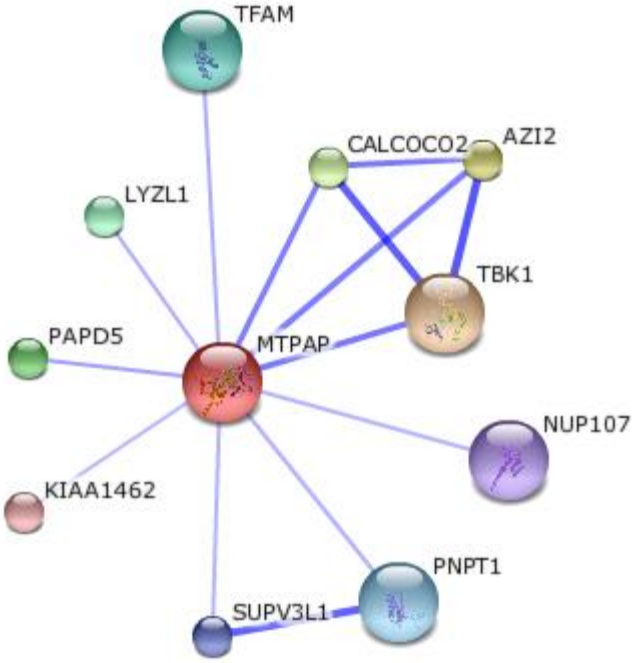
CSNK2B	casein kinase 2, beta polypeptide (234 aa)
G5A	Lymphocyte antigen 6 complex locus protein G5b Precursor ; Participates in Wnt signaling (By si [...] (234 aa)
HDAC2	histone deacetylase 2; Responsible for the deacetylation of lysine residues on the N-terminal p [...] (582 aa)
CALM1	calmodulin 1 (phosphorylase kinase, delta); Calmodulin mediates the control of a large number o [...] (149 aa)
JUN	jun oncogene; Transcription factor that recognizes and binds to the enhancer heptamer motif 5'- [...] (331 aa)
HDAC1	histone deacetylase 1; Responsible for the deacetylation of lysine residues on the N-terminal p [...] (482 aa)
CSN1S1	casein alpha s1; Important role in the capacity of milk to transport calcium phosphate (185 aa)
TP53	tumor protein p53; Acts as a tumor suppressor in many tumor types; induces growth arrest or apo [...] (393 aa)
SIRT1	sirtuin (silent mating type information regulation 2 homolog) 1 (S. cerevisiae); NAD-dependent [...] (747 aa)
SRF	serum response factor (c-fos serum response element-binding transcription factor); SRF is a tra [...] (508 aa)

Figure 5.18 Protein-protein interaction networks of casein kinase II alpha. Taken from STRING v10 (Szklarczyk et al. 2015)

The proteins may still form protein complexes or interact with proteins identified in the elution fraction. Casein kinase II alpha can interact with a number of proteins, as outlined in Figure 5.18. The STRING database allows a visualisation of known protein interactions and complexes (Szklarczyk et al. 2015). These proteins were cross-referenced in the SILAC dataset for both B1 and B2 fractions. From the 10 proteins that Casein kinase II alpha is reported to interact with, only 2 were found in fraction B1: CSNK2B and HDAC1. CSNK2B was the only protein found in fraction B2. Casein kinase II subunit beta and HDAC1 all had SILAC ratios of less than 0.6 for the Zta/Control analysis, indicating that there was less abundance of this protein in the Zta sample because the ratio was below 1.0. There was a more abundance of these proteins in the hisZta sample than the hisZtaAAA sample, but these SILAC ratios ranged from 1.2 to 1.68.

The function of casein kinase II alpha may play importance towards the initiation of the lytic cycle. Zta itself is phosphorylated at casein kinase II sites and is important for lytic cycle control (El-Guindy & Miller 2004). This is important for repressing Rta activation of late lytic genes. Also the viral kinase BGLF4 targets the nucleus and binds to nuclear pore complex proteins. This study also predicted the structure of BGLF4 based on the known structure of casein kinase II alpha, indicating sequence homology between the viral kinase and this host protein of interest (Chang et al. 2012).

The other protein identified from the mass spectrometry analysis was poly(A) RNA polymerase (mitochondrial). Poly(A) RNA polymerase (mitochondrial) is not present in control or hisZta B1 fraction, but present in hisZtaAAA B1. There is more poly(A) RNA polymerase (mitochondrial) in the hisZtaAAA B2 fraction. This abundance does not agree with the trend mass spectrometry data.



Predicted Functional Partners:

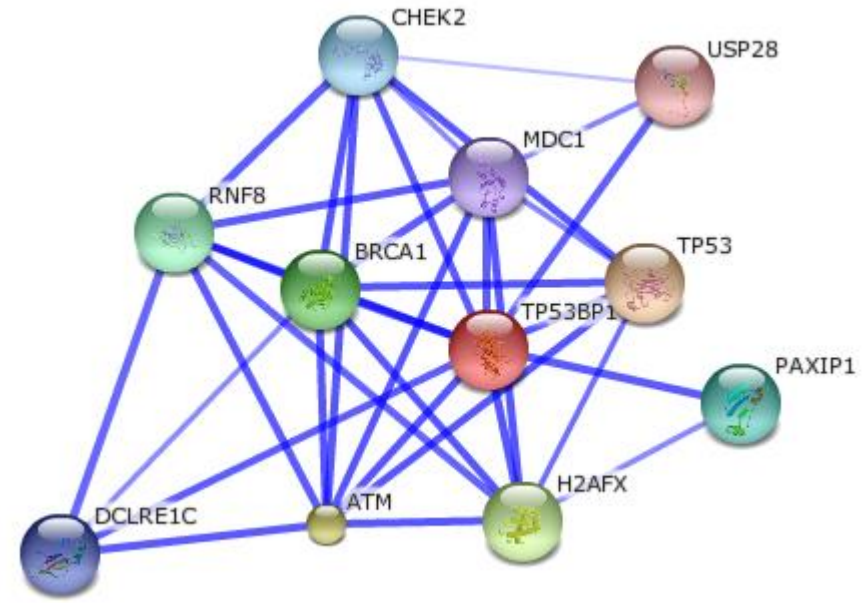
TBK1	TANK-binding kinase 1; Serine/threonine protein involved in the signaling cascade converging to [...] (729 aa)
AZI2	5-azacytidine induced 2; Activates serine/threonine-protein kinase TBK1 and facilitates its oli [...] (392 aa)
CALCOCO2	calcium binding and coiled-coil domain 2; May play a role in ruffle formation and actin cytoske [...] (446 aa)
PAPD5	PAP associated domain containing 5; Plays a role in replication-dependent histone mRNA degradat [...] (698 aa)
LYZL1	lysozyme-like 1 (194 aa)
TFAM	transcription factor A, mitochondrial; Involved in mitochondrial transcription regulation. Requ [...] (246 aa)
PNPT1	polyribonucleotide nucleotidyltransferase 1; Involved in mRNA degradation. Hydrolyzes single-st [...] (783 aa)
SUPV3L1	suppressor of var1, 3-like 1 (S. cerevisiae); ATPase and DNA/RNA helicase able to unwind DNA/DN [...] (786 aa)
NUP107	nucleoporin 107kDa; Essential component of nuclear pore complex. Required for the assembly of p [...] (925 aa)
KIAA1462	KIAA1462 (1359 aa)

Figure 5.19 Protein-protein interaction networks of poly(A) RNA polymerase (mitochondrial). Taken from STRING v10 (Szklarczyk et al. 2015)

Again, the proteins that may interact with this protein in a complex were investigated in the dataset (Figure 5.19). TFAM was in the B1 fraction with a Zta/control ratio of 0.662 and NUP107 was in the B2 fraction with a Zta/control ratio of 0.710. There was a marginal increase in abundance of these proteins in the Zta sample compared to the ZtaAAA sample in B1 and B2. Again the protein that may interact with poly(A) RNA polymerase (mitochondrial) were not detected at a matching abundance levels.

The function of poly(A) polymerase (mitochondrial) includes the Polyadenylation of RNA. This is essential for the regulation of translation of mRNAs, by completing the stop codon of RNA. These poly(A) tails formed contrastingly lead to histone mRNA degradation by this protein (Mullen & Marzluff 2008)

The final protein assessed was 53BP1. This protein is present in control, hisZta and hisZtaAAA B1 fraction, (Figure 5.17A). There is slightly more protein abundance in the control B1 fraction compared to hisZta fraction in fraction B1 (Figure 5.17B). There is more 53BP1 in the hisZtaAAA B1 fraction. This abundance does not agree with the trend mass spectrometry data



Predicted Functional Partners:

TP53	tumor protein p53; Acts as a tumor suppressor in many tumor types; induces growth arrest or apo [...] (393 aa)
ATM	ataxia telangiectasia mutated; Serine/threonine protein kinase which activates checkpoint signa [...] (3056 aa)
H2AFX	H2A histone family, member X; Variant histone H2A which replaces conventional H2A in a subset o [...] (143 aa)
BRCA1	breast cancer 1, early onset; The BRCA1-BARD1 heterodimer coordinates a diverse range of cellul [...] (1863 aa)
RNF8	ring finger protein 8; E3 ubiquitin-protein ligase required for assembly of repair proteins to [...] (485 aa)
PAXIP1	PAX interacting (with transcription-activation domain) protein 1; Functions as a key component [...] (1069 aa)
CHEK2	CHK2 checkpoint homolog (S. pombe); Regulates cell cycle checkpoints and apoptosis in response [...] (586 aa)
DCLRE1C	DNA cross-link repair 1C (PSO2 homolog, S. cerevisiae); Required for V(D)J recombination, the p [...] (692 aa)
MDC1	mediator of DNA-damage checkpoint 1 (2089 aa)
USP28	ubiquitin specific peptidase 28 (1077 aa)

Figure 5.20 Protein-protein interaction networks of 53BP1. Taken from STRING v10 (Szklarczyk et al. 2015)

Investigating 53BP1 interactions (Figure 5.20), only a BRCA1 (Fragment) had peptide coverage in fraction B2 that may have an interaction with 53BP1, but as no SILAC ratios were determined for this protein this protein could not be considered as part of any complex detected.

53BP1 is involved in binding p53 and regulates the cellular response to DNA double stranded breaks promoting non-homologous end-joining-mediated DSB repair (Panier & Boulton 2014). Zta has been demonstrated to interact with 53BP1 and 53BP1 is essential for EBV lytic replication (Bailey et al. 2009). A recent insight into the role of Zta and 53BP1 in the DNA damage response indicates Zta causes the mis-localisation of DNA damage proteins to site of damage sites (Yang et al. 2015). This impairment may play a role in altering the localisation of host proteins to facilitate EBV lytic cycle.

Follow up experiments would include attempting to analyse more fractions across from B1. Also a finer analysis of more gel slices would give a greater sensitivity. Once a protein is identified, the fractions can be finer eluted into smaller fractions and investigate potential complexes and their contribution to lytic cycle.

Overall there was insight into size exclusion fractions from EBV positive cells undergoing full lytic cycle compared against latency and cells that could not replicate the viral genome.

6. Analysis of the viral proteome during lytic replication

6.1. Introduction

Large-scale proteomic based approaches have previously been attempted to provide a means of understanding of the proteomes in cells infected with Herpesviruses. The viral proteins have numerous interactions with the host cell and this may alter the cellular proteome to assist the life cycle of the virus. It is understood that the change in mRNA abundance does not reflect the expression at protein level (Wynne et al. 2014). Numerous studies have demonstrated that EBV has a significant effect on the cellular transcriptome, whereas there has not been a focus on the cellular proteome.

Many groups have attempted to dissect the role of EBV lytic proteins and how they contribute to lytic cycle. Analysis of the EBV virion composition identified viral and cellular proteins by mass spectrometry (Johannsen et al. 2004). EBV follows the model other herpesvirus with respect to virion formation, with cellular proteins included in the virion, such as heat shock proteins, actin and tubulins that tightly associate with the viral capsid. Several previously uncharacterized genes have also been validated at both transcript and protein levels in mass spectrometry analysis of KSHV and EBV in virally infected cells (Dresang et al. 2011). For KSHV, heat shock chaperones were identified in the viral tegument playing important roles in the structure of the virion, possibly also playing a part in early infection (Zhu et al. 2005; Rozen et al. 2008)

By taking a whole cell proteomics approach to address the change in proteome in the context of EBV, I undertook an experiment to identify whether there is a difference in protein expression between latency and lytic cycle. An attempt to identify changes in cellular and viral expression from latency to lytic cycle was performed. I used a mass spectrometry based proteomic approach to enrich a population of cells undergoing lytic cycle coupled with SILAC labelling. Mass spectrometry was performed to identify and quantitate proteins from the total cell lysate to attempt to determine any differences between latent and lytic protein expression of host and viral proteins.

6.2. Results

6.2.1. Cell system to enrich cells undergoing EBV lytic cycle and proteomic analysis of EBV lytic cycle

Akata cells (Burkitts Lymphoma cell line) carry the EBV genome, but only one to two percent of EBV-positive cells are in lytic cycle and express lytic proteins (Takada et al. 1991). Even after the induction of lytic cycle with anti-IgG, this only rises to ten to twenty percent. In order to study the lytic cycle an engineered Akata cell system to enrich cells undergoing EBV lytic replication was utilised (Ramasubramanyan et al. 2015).

Akata cells that contain a bi-directional promoter sequence to express Zta, green fluorescent protein (GFP) and nerve growth factor receptor (NGFR) upon treatment with doxycycline were used (Ramasubramanyan, et al, 2015). A control cell line contained the Zta sequence in reverse orientation within the bi-directional promoter sequence (Figure 6.1).

The Akata control cells were metabolically labelled with R0K0 SILAC media (light) and the Zta inducible cells were labelled with R6K4 SILAC media (medium). The cells were passaged for five cell divisions to ensure all cells were efficiently metabolically labelled with the isotopes in each media. In the R0K0 labelled media, the amino acids arginine and lysine contain normal carbon (C) and hydrogen (H) isotopes. In the R6K4 media labelled media, amino acids arginine and lysine contain 13C arginine and 2H lysine amino acids. This slight difference in mass is differentiated using mass spectrometry when the samples are combined together and analysed. The differences in peptide mass can be identified and quantitated from all peptides that uptake the labelled amino acids.

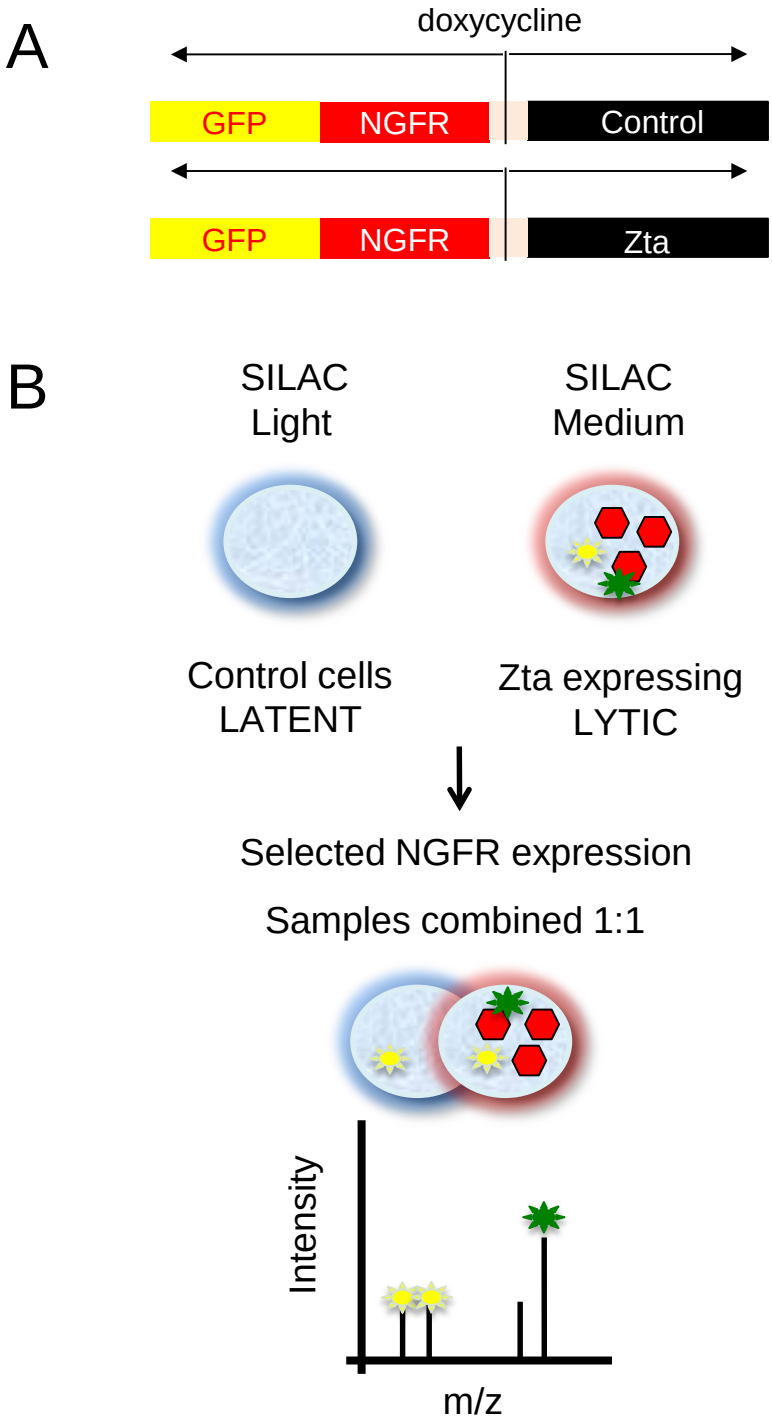


Figure 6.1 Inducible cell system initiating lytic cycle in Akata cells. A Diagram that represents the bidirectional promoter within the expression vector for Zta under doxycycline regulation. The control vector contains the same Zta sequence in the opposite direction. **B** Akata cells were SILAC labelled control (latent) and Zta expressing cells (lytic) were enriched by magnetic beads coupled with anti-NGFR (62% control and 57% Zta-expressing cells by FACS analysis). The isolated cell populations were combined equally and mass spectrometry performed.

The addition of doxycycline to both SILAC labelled control (light) and Zta expression cells (medium) leads to the expression of GFP, NGFR from both the control and Zta inducible cell lines. Upon the addition of doxycycline, Zta was expressed in the Zta expressing cell line (Figure 6.1B). The cells that express NGFR were then selected by immunomagnetic selection, using anti-NGFR antibodies coupled to magnetic beads (Figure 6.1B). To identify how many cells were isolated, the GFP expression was assessed. Fluorescent activated cell analysis (FACS) of positive GFP expression of these cells of the enriched control and Zta-expressing population showed isolated populations of 62% control and 57% Zta-expressing cells. (Ramasubramanyan, et al, 2015)

6.2.2. Mass spectrometry results of SILAC labelled Akata cells performed at the University of Sussex

The control and Zta-expressing cells was first processed myself within the University of Sussex mass spectrometry centre. The control sample and Zta-expressing sample were initially separated by SDS-PAGE and stained with SimplyBlue safestain (Invitrogen). 200ng and 500ng of BSA were also separated and run in parallel with the SILAC samples. These samples were cut for quality control (Figure 6.2A). The two samples were subjected to western blot analysis to quantitate expression of actin and Zta. Zta was expressed after activation of the promoter by the addition of doxycycline (Figure 6.2B). The actin bands were quantitated using ImageStudio (Licor v3.1.4). The actin in the cell samples detected here were equivalent in the western blot.

For the analysis, an equal amount of each sample was mixed together. 7.5µl of control sample and 7.5µl Zta-expressing cells were separated by SDS-PAGE (Figure 6.3A). These were joined with 200ng BSA and 500ng BSA bands that were cut previously (Figure 6.2A). Bands were cut and processed for mass spectrometry analysis. The ten gel slices were separated by SDS-PAGE (Figure 6.3A). The gel was stained with SimplyBlue stain and ten sections of equal size were cut and subjected to trypsin in-gel digestion. The final peptides were analysed at University of Sussex mass spectrometry. A brief summary of the in gel digestion protocol performed is demonstrated (Figure 6.3B).

The LTQ-OrbitrapXL instrument is quality controlled itself by running an internal BSA control before loading any sample. After the instrument was optimised for use, the 500ng BSA processed peptide sample was the first sample applied to the LTQ-OrbitrapXL instrument. This acts as a quality control for the in-gel digestion. The mass/charge ratio peptides were analysed by the Mascot search engine (Perkins et al. 1999). A screenshot of the BSA data returned using the Mascot software is shown (Figure 6.4). The BSA obtained a score of 66. The threshold here is set at 36 for a P>0.05. As the BSA score (66) is above the threshold it can be said that BSA peptides were detected by the mass spectrometer.

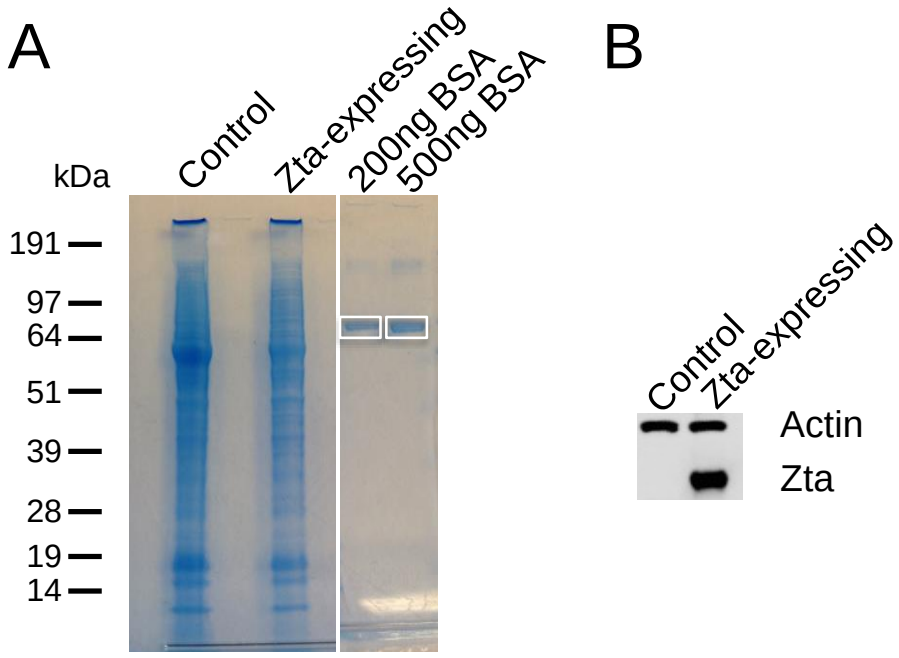


Figure 6.2 Control samples prepared for mass spectrometry analysis and western blot to check protein expression A Cellular proteins were separated by SDS-PAGE and analysed by Coomassie blue staining. Total cell extracts were prepared in protein sample buffer. 7.5µl of control sample and 7.5µl of Zta-expressing cells were run on a protein gel with 200ng BSA and 500ng BSA. Bands within the white boxes were cut and removed for mass spectrometry processing. **B** Western blot to confirm the expression of Zta of the control and Zta –expressing cells. Zta is expressed after induction with doxycycline and detected by BZ1 antibody. Actin is shown as a loading control and is detected by α-actin antibody.

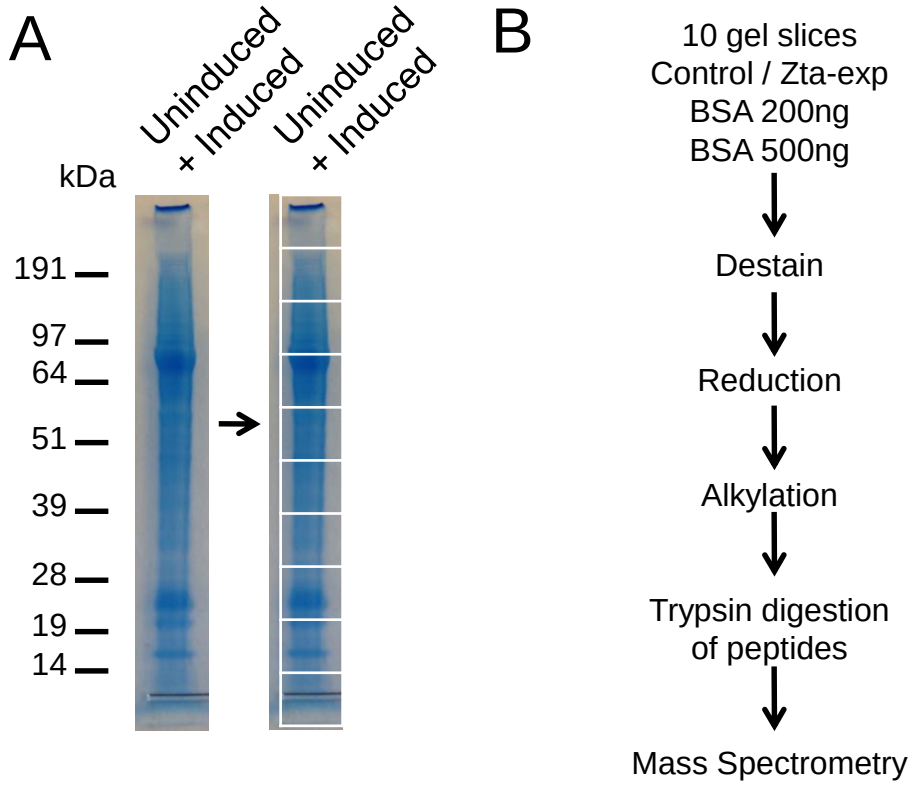
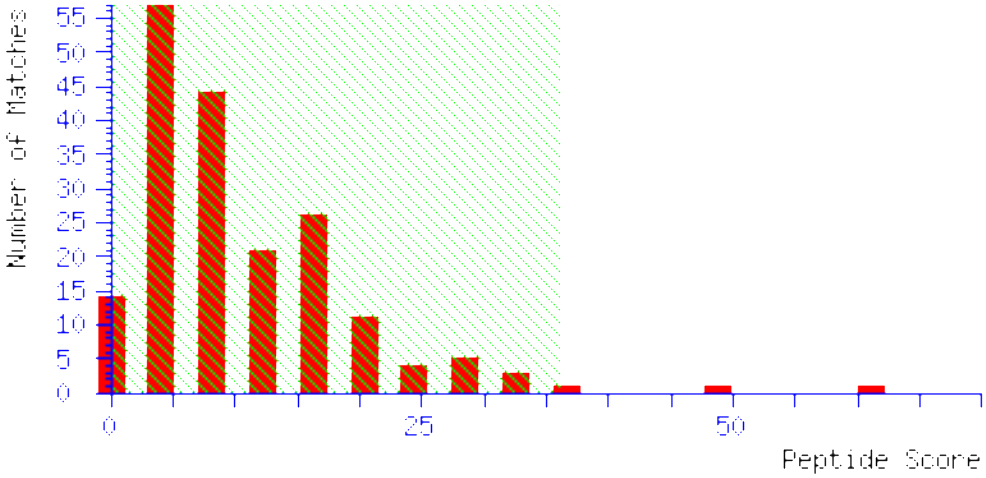
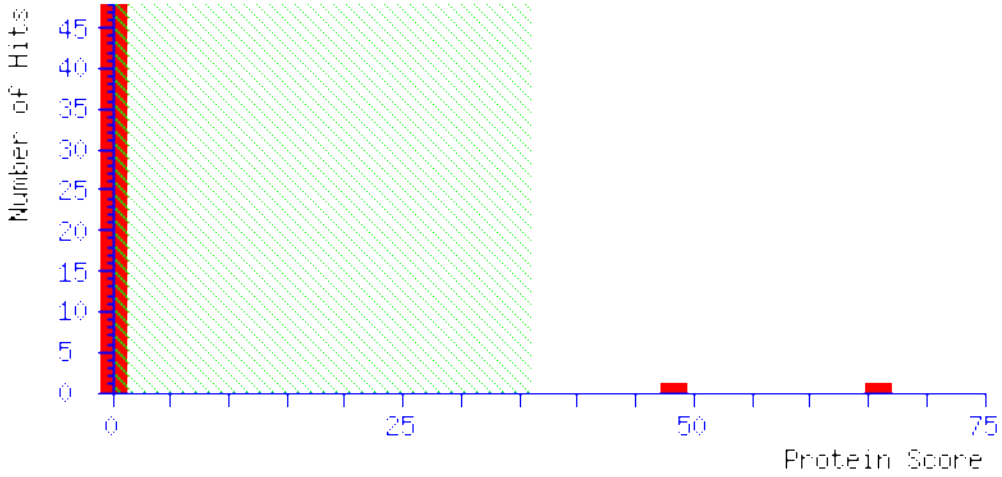


Figure 6.3 SILAC labeled proteins combined for mass spectrometry analysis and protocol flow diagram. A Latent and lytic cells sorted by NGFR expression were combined equally and separated by SDS-PAGE before coomassie staining. 10 sections were cut, shown by the white boxes, and these were processed for mass spectrometry as in B. **B** Flow diagram of in gel digestion performed on the gel slices.

A



B



C

	Score	Mass	Matches	Sequences	emPAI
ALBU_BOVIN	66	71244	36 (4)	32 (4)	0.19
K2C1B_HUMAN	50	61935	1 (1)	1 (1)	0.05

Figure 6.4 Mascot data of BSA peptide analysis. A quality check of digested peptides of the 500ng BSA (control) was performed using the Mascot Server database. **A** Peptide score distribution. The peptide score threshold was 36. **B** Score distribution for the proteins identified. The threshold protein score was 36 for P>0.05. **C** Quantitation overview of the two proteins identified. BSA was detected at Mascot score of 66. The threshold protein score was 36 for P>0.05.

The 10 gel slices processed by in gel digestion were individually analysed by mass spectrometry. The .raw output files are generated for each gel slice by the Xcalibur program (v2.0). These .raw files were applied to Maxquant (v1.5.3.) software (Cox & Mann 2008). The parameters included an FDR of 1% for protein detection from the peptides identified.

The data returned from the mass spectrometry analysis at University of Sussex revealed a limited amount of data from the samples. SILAC labelled proteins with a ratio between 1.2 and 100.0 shown in Table 6.1A. EBV lytic proteins BMRF1, BALF2, BaRF1 and BFRF1 were identified to have an increased SILAC ratio between 16.67 and 4.23. BMRF1 was the protein identified that most changed in abundance between the two samples, with a ratio of 16.67. Cellular proteins that were identified included four histones variants and ribonucleoproteins.

Proteins that decreased in abundance from latent to lytic cycle were displayed (Table 6.1B). The SILAC ratio ranges from 0.80 and below to 0.0. Interestingly, HLA class II histocompatibility antigen gamma chain (CD74) was identified to have a SILAC ratio of 0.50. This protein has been identified to be downregulated upon lytic cycle upon Zta expression in lytic cycle (Zuo et al, 2011). The mass spectrometry data here would agree with this statement.

A

Protein names	# Proteins identified	Peptides	SILAC Ratio
EBV BMRF1	3	9	16.67
EBV BALF2	3	8	12.77
EBV BaRF1	3	2	10.40
EBV BFRF1	3	4	4.23
Protein disulfide-isomerase	8	2	1.41
Histone H2B	12	2	1.35
Polyubiquitin-C	23	1	1.35
T-complex protein 1 subunit alpha	3	2	1.31
PC4 and SFRS1-interacting protein	11	2	1.31
Heterogeneous nuclear ribonucleoprotein U	2	2	1.31
Histone H3.2	10	2	1.30
Heterogeneous nuclear ribonucleoprotein Q	6	4	1.30
14-3-3 protein zeta/delta	11	2	1.28
Elongation factor 1-alpha 1	5	6	1.28
Histone H4	1	9	1.28
10 kDa heat shock protein, mitochondrial	4	3	1.25
Multifunctional protein ADE2	4	4	1.24
Myosin-9	2	2	1.21
Core histone macro-H2A.1;Histone H2A	7	3	1.21
14-3-3 protein epsilon	6	2	1.21

B

Protein names	# Proteins identified	Peptides	SILAC Ratio
RNA-binding protein Raly	6	4	0.79
Ezrin	6	3	0.78
Isocitrate dehydrogenase [NADP]	6	3	0.73
60S ribosomal protein L6	7	9	0.72
Cytochrome c oxidase subunit 6C	1	2	0.69
Heterogeneous nuclear ribonucleoprotein A/B	7	3	0.66
Medium-chain specific acyl-CoA dehydrogenase, mito	6	2	0.54
HLA class II histocompatibility antigen gamma chain	7	2	0.50
Tumor necrosis factor receptor superfamily member 16	2	4	0.20
Serine/threonine-protein kinase 36	3	1	0.03

Table 6.1 Identification of proteins by MaxQuant through University of Sussex mass spectrometry analysis. The MaxQuant analysis included a false discovery rate FDR 1% (0.01) **A** Proteins that were identified with a SILAC ratio greater than 1.2. These proteins were identified to have an increased expression in lytic cycle. **B** Proteins that were identified with a SILAC ratio less than 0.8. These proteins were identified to have a decreased expression in lytic cycle

A total of 375 proteins were identified from this study. Overall, the Maxquant analysis performed by myself at the University of Sussex identified some lytic proteins that increased expression in lytic cycle but only a few viral proteins were identified. This study was decided to be not in-depth or detailed enough to draw any conclusions. For further analysis, the remaining total cell extracts were sent to University of Bristol where a similar investigation was performed.

6.2.3. Mass spectrometry results of SILAC labelled Akata cells performed at the University of Bristol

The two SILAC labelled Akata samples (Control and Zta-expressing) were sent to University of Bristol Proteomics facility where they were mixed together equally and run on a protein gel. The proteins were cut from the gel into 6 bands and processed using in gel digestion. The peptides were applied to a LTQ-Orbitrap Velos mass spectrometer under Xcalibur software control. The .raw data was processed using Proteome Discoverer software (v1.2) and finalised into an Excel file including SILAC ratios for identified proteins.

There was an abundance in proteins identified by the mass spectrometry performed at the University of Bristol. 3016 proteins were identified with a SILAC ratio. Some of the proteins from this study and some of the proteins were investigated further.

These SILAC ratios were then converted by myself into Log2 ratios and entered into GraphPad Prism software (v6.0). A histogram and a Gaussian distribution graph were created to illustrate the total outline of the data between the latent and lytic proteins identified (Figure 6.5). A global view of the proteins shows the distribution of the proteins decreased slightly giving a normal distribution just below zero (Figure 6.5A). Some proteins had a very low Log2 ratio, which indicated they were only identified in latency and not cells undergoing lytic replication, or were quite possibly environmental contaminants. Some proteins had a very high Log2 ratio, which indicated they were only identified in lytic cycle and not expressed in latency (Figure 6.5B).

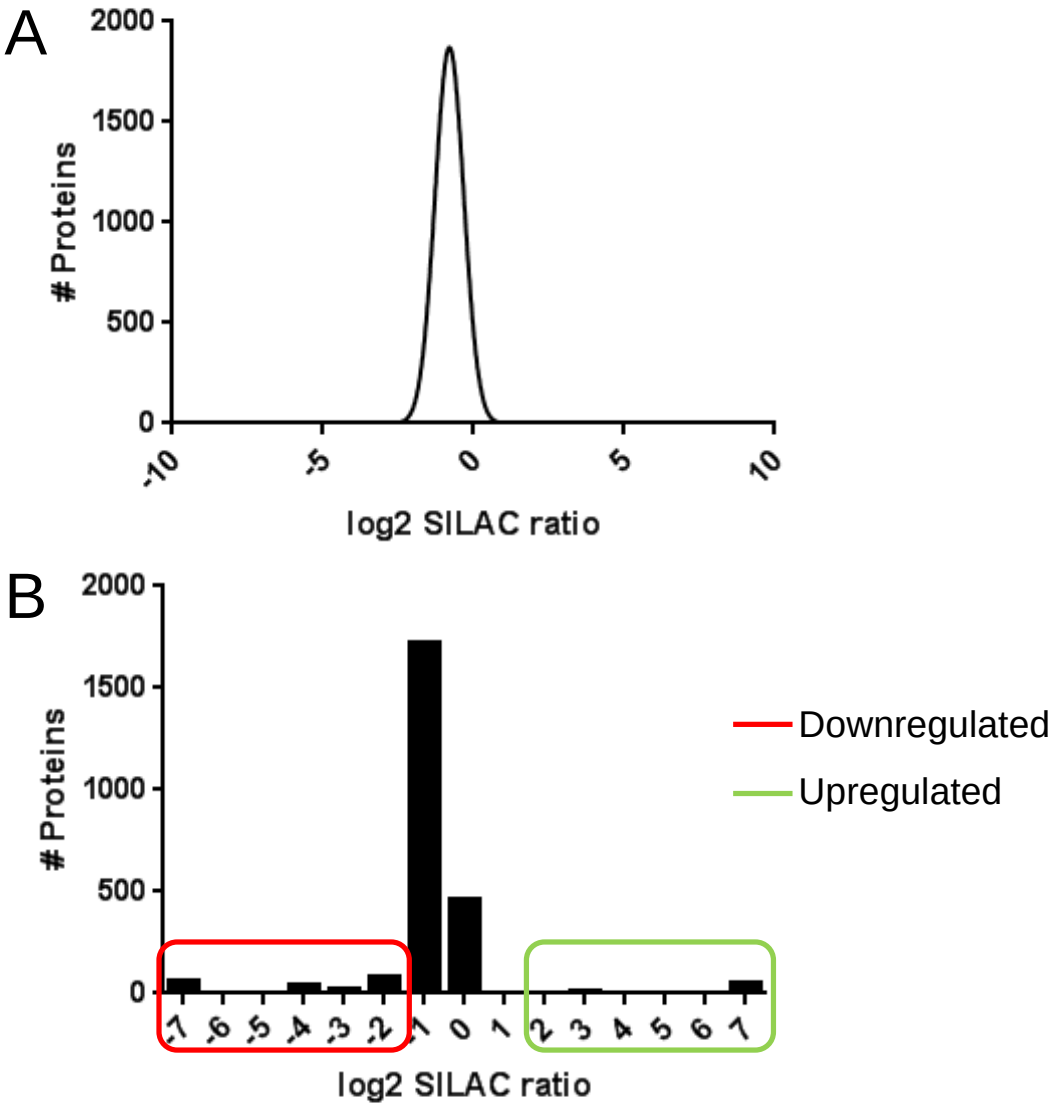


Figure 6.5 Representation of SILAC Akata proteins returned from mass spectrometry identification from the University of Bristol using log2 of the SILAC ratios Akata lytic / latent. **A** Frequency distribution curve of all proteins identified by mass spectrometry. Proteins had increased or decreased expression in lytic cycle. A distribution around -0.5 to 0 was observed **B** Histogram of all proteins identified by mass spectrometry. Proteins had increased or decreased expression in lytic cycle. Some proteins were identified to be only in latent cells or only lytic cells.

6.2.4. Cellular proteins identified by the mass spectrometry analysis

The most abundant proteins identified with a SILAC ratio were determined by the number of peptide spectrum matches (PSMs) from the data (Table 6.2). The PSM value is a total count where the experimental peptides spectra are compared with a theoretical spectrum of the specific peptide sequence. The most abundantly identified peptides corresponded to heat shock protein HSP 90-beta with a PSM value of 330. Nine other proteins were identified with a PSM value above 210. These included three heat shock proteins and six tubulins. Heat shock proteins play multiple roles in the cellular response including acting as chaperones, stabilizing unfolded proteins or transporting proteins throughout the cell (Lindquist & Craig 1988). Six tubulin variants were also identified. The cytoskeleton is composed of many structural proteins including an abundance of tubulins.

Protein ID	Cellular protein name	PSM #	SILAC Ratio
P08238	Heat shock protein HSP 90-beta	330	0.71
P07437	Tubulin beta chain	296	0.68
P68371	Tubulin beta-4B chain	294	0..730
P11142	Heat shock cognate 71 kDa protein	268	0.78
Q9BVA1	Tubulin beta-2B chain	263	0.62
P07900	Heat shock protein HSP 90-alpha	262	0.74
Q13885	Tubulin beta-2A chain	261	0.66
P10809	60 kDa heat shock protein, mitochondrial	224	0.72
P04350	Tubulin beta-4A chain	211	0.64
Q13509	Tubulin beta-3 chain	210	0.71

Table 6.2 Most abundant proteins identified by University of Bristol mass spectrometry analysis. SILAC labelled proteins were ordered by peptide spectrum matches. Heat shock proteins and tubulins were most abundantly identified.

To investigate the expression of the heat shock proteins between latent and lytic cycles, a western blot was performed to determine Heat shock protein HSP 90-alpha expression (Figure 6.6). This protein had a PSM value of 262 and a SILAC ratio of 0.74.

Total cell extracts of control Akata cells and Zta-expressing Akata cells after induction with doxycycline were assessed. This allowed the activation of lytic cycle compared to cells in latency. The Heat shock protein HSP 90-alpha showed similar expression levels between the two life cycles. The western blot agreed with the mass spectrometry data. A quantitation of the HSP90 protein bands was performed using ImageStudio (Li-cor, v3.1.4.). The quantitation levels for Heat shock protein HSP90-alpha was 1.99 for latent Akata cells and 1.35 for lytic Akata cells. The ratio of expression between these is 0.68. The SILAC mass spectrometry ratio was 0.74; indicating the expression of this protein reduces or the protein is targeted for degradation upon lytic cycle.

A

Protein ID	Cellular protein name	PSM #	SILAC Ratio
P07900	Heat shock protein HSP 90-alpha	262	0.74

B

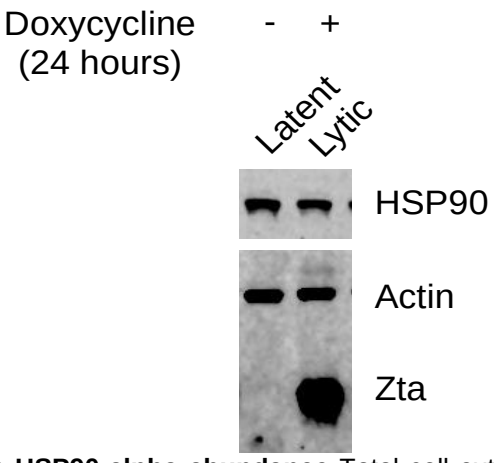


Figure 6.6 Heat shock protein HSP90-alpha abundance Total cell extracts of control Akata cells and Zta-inducible Akata cells were created after cells were induced with doxycycline after 24 hours. An equal amount of samples were separated by SDS-PAGE. **A** SILAC ratio and PSM value for Heat shock protein HSP 90-alpha identified from the mass spectrometry data. **B** Western blot to confirm heat shock protein HSP90-alpha (P07900) expression. The quantitation levels for Heat shock protein HSP90-alpha is 1.99 for latent Akata cells and 1.35 for lytic Akata cells. HSP90 is detected by α-HSP90 (ab13492, Abcam) Zta is detected by BZ1 antibody. Actin is shown as a loading control and is detected by α-actin antibody (Sigma).

Protein ID	Protein	PSM #	SILAC Ratio
Q58FF6	Putative heat shock protein HSP 90-beta 4	46	100.0
Q14568	Putative heat shock protein HSP 90-alpha A2	35	100.0
Q14222	EEF1A protein (Fragment)	37	100.0
A0A024RAC9	Zinc finger, UBR1 type 1, isoform CRA_c	4	100.0
Q5M9N0	Coiled-coil domain-containing protein 158	3	100.0
B4DNN8	cDNA FLJ60025, highly similar to Mus musculus PR domain containing 6 (Prdm6), mRNA	2	100.0
A0JP10	Excision repair cross-complementing rodent repair deficiency, complementation group 6	3	100.0
O94986	Centrosomal protein of 152 kDa	2	100.0
P78363	Retinal-specific ATP-binding cassette transporter	2	100.0
B2RUU2	ATP-binding cassette, sub-family A (ABC1), member 1	2	100.0
F8WC62	Protein FAM71F1	2	100.0
H0Y9K5	Latrophilin-3 (Fragment)	2	100.0
Q6UXR4	Putative serpin A13	2	100.0
Q9UHB4	NADPH-dependent diflavin oxidoreductase 1	2	100.0

Table 6.3 The most abundant proteins identified only in lytic cycle. Cellular proteins that were identified with a SILAC ratio of 100.0 ordered by peptide spectrum matches. These proteins were only identified in the Zta-expressing cells.

The mass spectrometry data included proteins that were only identified in lytic cycle. These proteins were detected only within the Zta-expressing, EBV lytic phase of the cell (Table 6.3). The proteins were ranked by PSM value and all proteins were given a SILAC ratio of 100.0 as only R6K4 labelled peptides were identified. The top two proteins with the most PSM values of 46 and 35 were Putative heat shock protein HSP 90-beta 4 (Q58FF6) and Putative heat shock protein HSP 90-alpha A2 (Q14568). The third most abundant protein identified with a PSM value of 37 and a SILAC ratio of 100.0 was the elongation factor EEF1A (Q14222). The heat shock proteins were of significant interest as some HSP have been implicated in facilitating lytic proteins in the lytic pathway (Kawashima et al, 2013). Also heat shock proteins can interact with the viral protein kinase BGLF4 and this interaction is important for the kinase to phosphorylate proteins essential for lytic replication (Sun et al, 2013).

6.2.5. Heat shock proteins that were identified only in lytic cycle

All of the heat shock proteins identified were collated together from a range of 100.0 to 0.44 SILAC ratios (Table 6.4). The data was ranked from 100.0 SILAC ratio being the highest ranked protein identified. Some proteins were abundantly identified with some PSM values over 200. As the highest ranked heat shock proteins returned were heat shock protein HSP 90-beta 4 and heat shock protein HSP 90-alpha A2 (Both with a SILAC ratio of 100.0, these were investigated further.

Protein ID	Protein	# PSMs	Lytic/ Latent Ratio
Q58FF6	Putative heat shock protein HSP 90-beta 4	46	100.00
Q14568	Putative heat shock protein HSP 90-alpha A2	35	100.00
Q0VDF9	Heat shock 70 kDa protein 14	5	0.92
P08107	Heat shock 70 kDa protein 1A/1B	79	0.82
Q12931	Heat shock protein 75 kDa, mitochondrial	88	0.81
B2R6X5	cDNA, FLJ93166, highly similar to Homo sapiens heat shock 70kDa protein 6 (HSP70B') (HSPA6), mRNA	52	0.81
P11142	Heat shock cognate 71 kDa protein	268	0.78
P07900	Heat shock protein HSP 90-alpha	262	0.74
P10809	60 kDa heat shock protein, mitochondrial	224	0.72
A0A024R3X7	Heat shock 10kDa protein 1 (Chaperonin 10), isoform CRA_d	31	0.72
P08238	Heat shock protein HSP 90-beta	330	0.71
A0A024RDQ0	Heat shock 105kDa/110kDa protein 1, isoform CRA_a	81	0.69
P34932	Heat shock 70 kDa protein 4	109	0.64
B3KNQ9	cDNA FLJ30200 fis, clone BRACE2001455, highly similar to HEAT SHOCK FACTOR PROTEIN 1	1	0.53

Table 6.4 All heat shock proteins identified including proteins identified in only lytic cycle. Proteins were ranked by SILAC ratio with 100.0 the highest ratio obtained.

Due to the limited availability of specific heat shock protein antibodies and the similarity in amino acid sequences of heat shock proteins, the Abcam antibody used previously for HSP90 (P07900) was investigated for reactivity with the HSP of interest. The HSP90 protein is a heat shock protein of 732 amino acids in length, of about 85kDa molecular weight. The amino acid sequence of HSP90 (P07900) was retrieved from Uniprot database (http://www.uniprot.org/uniprot/P07900) (Figure 6.7). The antibody epitope was retrieved from the manufacturers' website (http://www.abcam.com/hsp90-antibody-ac88-ab13492.html). The antibody epitope region occupies amino acids 604aa-697aa. The epitope for the AbCam HSP90 antibody (ab13492) used previously for the western blot (Figure 6.6) was mapped to the amino acid sequence and displayed in red (Figure 6.7).

Once the antibody epitope was identified then the amino acid sequences of the three proteins investigated were compared. The amino acid sequences of HSPO90AA1 HSP90AB4P and HSP90AA2 were retrieved from the NCBI database and later compared using NCBI BLAST (Figure 6.8). HSP90AB4P is an HSP of 505 amino acids in length and 58kDa molecular weight. HSP90AA2 is an HSP of 343 amino acids in length and 39kDa molecular weight. These differ significantly from the larger HSP90 (P07900). The sequences were compared using NCBI blast.

```
 10         20         30         40         50
MPEETQTQDQ PMEEEEVETF AFQAEIAQLM SLIINTFYSN KEIFLRELIS
         60         70         80         90        100
NSSDALDKIR YESLTDPSKL DSGKELHINL IPNKQDRTLT IVDTGIGMTK
        110        120        130        140        150
ADLINNLGTI AKSGTKAFME ALQAGADISM IGQFGVGFYS AYLVAEKVTV
        160        170        180        190        200
ITKHNDDEQY AWESSAGGSF TVRTDTGEPM GRGTKVILHL KEDQTEYLEE
        210        220        230        240        250
RRIKEIVKKH SQFIGYPITL FVEKERDKEV SDDEAEEKED KEEEKEKEEK
        260        270        280        290        300
ESEDKPEIED VGSDEEEEKK DGDKKKKKKI KEKYIDQEEL NKTKPIWTRN
        310        320        330        340        350
PDDITNEEYG EFYKSLTNDW EDHLAVKHFS VEGQLEFRAL LFVPRRAPFD
        360        370        380        390        400
LFENRKKKNN IKLYVRRVFI MDNCEELIPE YLNFIRGVVD SEDLPLNISR
        410        420        430        440        450
EMLQQSKILK VIRKNLVKKC LELFTELAED KENYKKFYEQ FSKNIKLGIH
        460        470        480        490        500
EDSQNRKKLS ELLRYYTSAS GDEMVSLKDY CTRMKENQKH IYYITGETKD
        510        520        530        540        550
QVANSAFVER LRKHGLEVIY MIEPIDEYCV QQLKEFEGKT LVSVTKEGLE
        560        570        580        590        600
LPEDEEEKKK QEEKKTKFEN LCKIMKDILE KKVEKVVVSN RLVTSPCCIV
        610        620        630        640        650
TSTYGWTANM ERIMKAQALR DNSTMGYMAA KKHLEINPDH SIIETLRQKA
        660        670        680        690        700
EADKNDKSVK DLVILLYETA LLSSGFSLED PQTHANRIYR MIKLGLGIDE
        710        720        730
DDPTADDTSA AVTEEMPPLE GDDDTSRMEE VD
```

Figure 6.7 Amino acid sequence of HSP90 (P07900) and where the selected antibody epitope identifies. The Abcam P07900 antibody (ab13492) epitope has been mapped to amino acid residues 604-697 of the human Hsp90 sequence highlighted in red.

A

FASTA Sequence for HSP90 protein P07900
http://www.ncbi.nlm.nih.gov/protein/P07900

>sp|P07900|HS90A_HUMAN Heat shock protein HSP 90-alpha
OS=Homo sapiens GN=HSP90AA1 PE=1 SV=5

MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKI
RYESLTDPSKLDSGKELHINLIPNKQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAF
MEALQAGADISMIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRT
DTGEPMGRGTKVILHLKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSD
DEAEEKEDKEEEKEKEEKESEDKPEIEDVGSDEEEEKKDGDKKKKKKIKEKYIDQEE
LNKTKPIWTRNPDDITNEEYGEFYKSLTNDWEDHLAVKHFSVEGQLEFRALLFVPRR
APFDLFENRKKKNNIKLYVRRVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQ
QSKILKVIRKNLVKKCLELFTELAEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELL
RYYTSASGDEMVSLKDYCTRMKENQKHIYYITGETKDQVANSAFVERLRKHGLEVIY
MIEPIDEYCVQQLKEFEGKTLVSVTKEGLELPEDEEEKKKQEEKKTKFENLCKIMKDI
LEKKVEKVVVSNRLVTSPCCIVTSTYGWTANMERIMKAQALRDNSTMGYMAAKKHL
EINPDHSIIETLRQKAEADKNDKSVKDLVILLYETALLSSGFSLEDPQTHANRIYRMIKL
GLGIDEDDPTADDTSAAVTEEMPPLEGDDDTSRMEEVD

B

FASTA Sequnce for HSP90 Q58FF6
http://www.ncbi.nlm.nih.gov/protein/Q58FF6

>sp|Q58FF6|H90B4_HUMAN Putative heat shock protein HSP 90-beta 4
OS=Homo sapiens GN=HSP90AB4P PE=5 SV=1

MSLIINTFYSNKEIFLQELISNASDALDKIRYESLTDPSKLDGGKELKIDIIPNPRECIL
TLVNTGIGMTKADLINNLGAIAKSGTEAFMEAFQSCAEISMIGQFGVGFYSAYLVAE
KVAITKHNDEEQYSWVSSAGSSFTLHVDHGEPIDRDTKVILHLKEDQTEYLEERWV
KEVVKKHPQFIGCLIAVYLEKEPEKEISDDEEEKGEKEEEDKDDKEKPKTEDVGSDE
EDDTDKNNKKKTKKIKEKYTDREELNQTKPIWTRNPDDITQEECGEFYKSLTSAWE
DHLAVKQFPVEEQEENEQLCVHHVWIMDSFDDLMPEYVGFVREDKENNKKLDEVF
SKISWLGIHEDSINWRHLSELLWSHTFQSGDEMTSLSEYVSCMKEAQKSICDIIGEC
KEQVANSAFVEQEWKKGFEVIYMSEPIDEYCVQQLKEFDGKSLLSVTKEGLELPED
EEEKKIMEESNVKFENLCRLMKEILDKKVERVTISSRLVSSPCRIVTSTYS

C

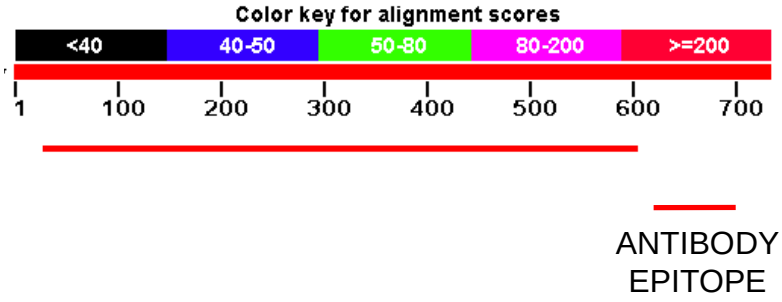
FASTA Sequnce for HSP90 Q14568
http://www.ncbi.nlm.nih.gov/protein/Q14568.2

>sp|Q14568|HS902_HUMAN Putative heat shock protein HSP 90-alpha A2
OS=Homo sapiens GN=HSP90AA2 PE=1 SV=2

MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDK
IWYESLTDPSKLDSGKELHINLIPNKQDQTLTIVDTGIGMTKADLINNLGTIAKSGTKA
FMEALQAGADISMIGQFGVSFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTV
RTDTGERMGRGTKVILHLKEDQTEYLEEQRIKEIVKKHSQLIGYPITLFVEKECDKEV
SDDETEEKEDKEEEKEKEEKESKDKPEIEDVGSDEEEEKKDGDKKKKKTKEKYIDQ
EELNKTKPIWTRNPDDITNEEYGEFCKNLTNDWEDHLAVKHFSVEGQLEFRALLFV
P

Figure 6.8 Amino acid sequences of heat shock proteins retrieved from NCBI database. A HSP90AA1 **B** HSP90AB4P **C** HSP90AA2

A



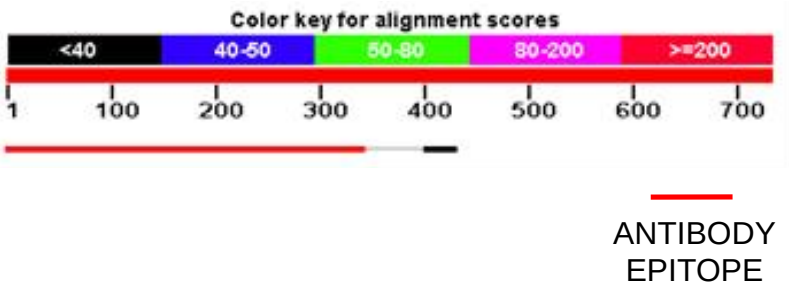
B

Score	Expect	Method	Identities	Positives	Gaps
636 bits(1641)	0.0	Compositional matrix adjust.	362/576(63%)	423/576(73%)	71/576(12%)

```
Query  30   MSLIINTFYSNKEIFLRELISNSSDALDKIRYESLTDPSKLDSGKELHINLIPNKQDRTL  89
            MSLIINTFYSNKEIFL+ELISN+SDALDKIRYESLTDPSKLD GKEL I++IPN ++  L
Sbjct  1    MSLIINTFYSNKEIFLQELISNASDALDKIRYESLTDPSKLDGGKELKIDIIPNPRECIL  60

Query  90   TIVDTGIGMTKADLINNLGTIAKSGTKAFMEALQAGADISMIGQFGVGFYSAYLVAEKVT  149
            T+V+TGIGMTKADLINNLG IAKSGT+AFMEA Q+ A+ISMIGQFGVGFYSAYLVAEKV
Sbjct  61   TLVNTGIGMTKADLINNLGAIAKSGTEAFMEAFQSCAEISMIGQFGVGFYSAYLVAEKVA  120

Query  150  VITKHNDDEQYAWESSAGGSFTVRTDTGEPMGRGTKVILHLKEDQTEYLEERRIKEIVKK  209
             ITKHND+EQY+W SSAG SFT+  D GEP+ R TKVILHLKEDQTEYLEER +KE+VKK
Sbjct  121  -ITKHNDEEQYSWVSSAGSSFTLHVDHGEPIDRDTKVILHLKEDQTEYLEERWVKEVVKK  179

Query  210  HSQFIGYPITLFVEKERDKEVSDDEAEEKEDKEEEKEKEEKESEDKPEIEDVGSDEEEEK  269
            H QFIG  I +++EKE +KE+SDDE      ++ EKE+E+K+ ++KP+ EDVGSDEE++
Sbjct  180  HPQFIGCLIAVYLEKEPEKEISDDEE-----EKGEKEEEDKDDKEKPKTEDVGSDEEDDT  234

Query  270  KDGDKKKKKKIKEKYIDQEELNKTKPIWTRNPDDITNEEYGEFYKSLTNDWEDHLAVKHF  329
               +KKK KKIKEKY D+EELN+TKPIWTRNPDDIT EE GEFYKSLT+ WEDHLAVK F
Sbjct  235  DKNNKKKTKKIKEKYTDREELNQTKPIWTRNPDDITQEECGEFYKSLTSAWEDHLAVKQF  294

Query  330  SVEGQLEFRALLFVPRRAPFDLFENRKKKNNIKLYVRRVFIMDNCEELIPEYLNFIRGVV  389
             VE Q E                          N +L V  V+IMD+ ++L+PEY+ F+R
Sbjct  295  PVEEQEE----------------------NEQLCVHHVWIMDSFDDLMPEYVGFVR---  328

Query  390  DSEDLPLNISREMLQQSKILKVIRKNLVKKCLELFTELAEDKENYKKFYEQFSKNIKLGI  449
                                                EDKEN KK  E FSK   LGI
Sbjct  329  ----------------------------------------EDKENNKKLDEVFSKISWLGI  349

Query  450  HEDSQNRKKLSELLRYYTSASGDEMVSLKDYCTRMKENQKHIYYITGETKDQVANSAFVE  509
            HEDS N + LSELL  +T  SGDEM SL +Y + MKE QK I  I GE K+QVANSAFVE
Sbjct  350  HEDSINWRHLSELLWSHTFQSGDEMTSLSEYVSCMKEAQKSICDIIGECKEQVANSAFVE  409

Query  510  RLRKHGLEVIYMIEPIDEYCVQQLKEFEGKTLVSVTKEGLELPEDEEEKKKQEEKKTKFE  569
            +  K G EVIYM EPIDEYCVQQLKEF+GK+L+SVTKEGLELPEDEEEKK  EE   KFE
Sbjct  410  QEWKKGFEVIYMSEPIDEYCVQQLKEFDGKSLLSVTKEGLELPEDEEEKKIMEESNVKFE  469

Query  570  NLCKIMKDILEKKVEKVVVSNRLVTSPCCIVTSTYG  605
            NLC++MK+IL+KKVE+V +S+RLV+SPC IVTSTY
Sbjct  470  NLCRLMKEILDKKVERVTISSRLVSSPCRIVTSTYS  505
```

Figure 6.9 BLAST search comparing HSP90AB4P against HSP90 (P07900). The BLAST search showed a sequence similarity of 63% between HSP90AB4P against HSP90 (P07900). There is no overlap with the 604-697 amino acid antibody epitope region

A



B

Score	Expect	Method	Identities	Positives	Gaps
663 bits(1711)	0.0	Compositional matrix adjust.	331/344(96%)	335/344(97%)	1/344(0%)

```
Query  1    MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIR  60
            MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKI
Sbjct  1    MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIW  60

Query  61   YESLTDPSKLDSGKELHINLIPNKQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFME  120
            YESLTDPSKLDSGKELHINLIPNKQD+TLTIVDTGIGMTKADLINNLGTIAKSGTKAFME
Sbjct  61   YESLTDPSKLDSGKELHINLIPNKQDQTLTIVDTGIGMTKADLINNLGTIAKSGTKAFME  120

Query  121  ALQAGADISMIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPM  180
            ALQAGADISMIGQFGV FYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGE M
Sbjct  121  ALQAGADISMIGQFGVSFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGERM  180

Query  181  GRGTKVILHLKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKED  240
            GRGTKVILHLKEDQTEYLEE+RIKEIVKKHSQ IGYPITLFVEKE DKEVSDDE EEKED
Sbjct  181  GRGTKVILHLKEDQTEYLEEQRIKEIVKKHSQLIGYPITLFVEKECDKEVSDDETEEKED  240

Query  241  KEEEKEKEEKESEDKPEIEDVGSDEEEEKKDGDKKKKKKIKEKYIDQEELNKTKPIWTRN  300
            KEEEKEKEEKES+DKPEIEDVGSDEEEEKKDGD KKKKK KEKYIDQEELNKTKPIWTRN
Sbjct  241  KEEEKEKEEKESKDKPEIEDVGSDEEEEKKDGD-KKKKKTKEKYIDQEELNKTKPIWTRN  299

Query  301  PDDITNEEYGEFYKSLTNDWEDHLAVKHFSVEGQLEFRALLFVP  344
            PDDITNEEYGEF K+LTNDWEDHLAVKHFSVEGQLEFRALLFVP
Sbjct  300  PDDITNEEYGEFCKNLTNDWEDHLAVKHFSVEGQLEFRALLFVP  343
```

Figure 6.10 BLAST search comparing HSP90AA2 against the HSP90 (P07900). The BLAST search showed a sequence similarity of 96%. There is no overlap with the 604-697 amino acid antibody epitope region.

The HSP90AB4P amino acid sequence was entered into the NCBI BLAST against the amino acid sequence of HSP90 (P07900). The two proteins were compared to identify if they had any similarity, focused on the epitope region the antibody identifies. The BLAST between the two sequences, the two proteins have a sequence similarity of 63% (Figure 6.9). HSP90AB4P is composed of 505 amino acids with a BLAST search against HSP90 (P07900) of 732 amino acids. Unfortunately, there is no overlap region with the 604-697 amino acid sequence recognised by the antibody (ab13492).

The HSP90AA2 amino acid sequence was entered into the NCBI BLAST against the amino acid sequence of HSP90 (P07900). The two proteins were compared to identify if they had any similarity, focused on the epitope region the antibody identifies. After the BLAST search between the two sequences, the two proteins have a sequence similarity of 96% (Figure 6.10). HSP90AA2 is composed of 343 amino acids with a BLAST search against HSP90 (P07900) is 732 amino acids. Again, there is no overlap region with the 604-697 amino acid sequence recognised by the antibody (ab13492).

The proteins HSP90AB4P and HSP90AA2 amino acid sequences were investigated against other potential antibody candidates. An extensive search was performed for heat shock antibodies and the epitope they recognise. The majority of antibodies available that recognised a HSP90 protein recognised P07900 only.

Therefore, the conclusion was that the heat shock proteins that were identified from only the lytic induced cells could not be followed up with any available antibodies. The sequences and proteins were searched on antibody databases and it was not possible to identify an available antibody.

6.2.6. Epstein-Barr virus proteins identified in lytic cycle

In addition to the cellular proteins that were identified by mass spectrometry, viral peptides were also detected and identified. 40 lytic Epstein-Barr virus proteins were identified in this study from the mass spectrometry performed at University of Bristol. Table 6.5 represents the data returned from the analysis.

Some of these proteins have been identified in previous studies (Johannsen et al. 2004; Dresang et al. 2011; Koganti et al. 2015). The most abundant EBV protein identified here was the major DNA-binding protein BALF2 with a PSM value of 164. Other replication factors identified include BMRF1 (PSM 69), BALF5 (PSM 21), BBLF3 (PSM 22), BZLF1 (PSM 5) and BSLF1 (PSM 1).

Other capsid proteins and tegument proteins were identified that include the major capsid protein, major tegument protein, envelope glycoprotein B, capsid proteins VP23 and VP26, and packaging proteins. Proteins that manipulate the cellular response to replication and posttranslational modification of cellular and viral proteins include the serine/threonine kinase BGLF4, apoptosis regulator BHRF1 and the EBV uracil-DNA glycosylase BKRF3.

Protein ID	Description	# PSMs	SILAC ratio
P03227	Major DNA-binding protein	164	100
P03191	DNA polymerase processivity factor BMRF1	69	100
P03190	Ribonucleoside-diphosphate reductase large subunit	60	100
P03226	Major capsid protein	47	100
P03185	Virion egress protein UL34 homolog	39	100
P03177	Thymidine kinase	35	100
P0CAP6	Ribonucleoside-diphosphate reductase small chain	32	100
P13288	Serine/threonine-protein kinase BGLF4	34	100
Q04360	mRNA export factor ICP27 homolog	25	100
P03198	DNA polymerase catalytic subunit	21	100
P03179	Major tegument protein	15	100
Q8AZJ7	Putative BBLF3 protein (Fragment)	22	100
Q66541	scaffold protein BdRF1	13	100
P03217	Shutoff alkaline exonuclease	19	100
P03182	Apoptosis regulator BHRF1	10	100
P03186	Deneddylase	3	100
P03188	Envelope glycoprotein B	9	100
P25214	Triplex capsid protein VP23 homolog	9	100
P03206	Trans-activator protein BZLF1	5	100
P14348	Capsid protein VP26	5	100
P03210	Tegument protein BRRF2	6	100
P03195	Deoxyuridine 5'-triphosphate nucleotidohydrolase	5	100
P03187	Triplex capsid protein VP19C homolog	2	100
P03207	Transcriptional activator BRRF1	2	100
P0CK53	Capsid-binding protein UL16 homolog	2	100
P03224	Probable membrane antigen GP85	2	100
P03193	DNA primase	1	100
P03200-2	Isoform GP220 of Envelope glycoprotein GP350	2	100
P29882	Virion egress protein UL7 homolog	2	100
P0CK49	Tegument protein BSRF1	1	100
P03205	Glycoprotein 42	1	100
P03197	Tegument protein BLRF2	5	77.217
P0CK47	Virion egress protein BFLF2	5	50.953
P03184	Packaging protein UL32 homolog	3	48.533
P12888	Uracil-DNA glycosylase	5	39.089
P03231	Envelope glycoprotein H	1	10.272
P0C727	Uncharacterized protein LF3	1	0.159
P30119	Uncharacterized protein BTRF1	1	
Q8AZJ5	Uncharacterized LF1 protein	1	
P03233	Virion-packaging protein UL25 homolog	1	

Table 6.5 All EBV proteins only identified in lytic cycle. Proteins ranked by SILAC ratio. 40 EBV lytic proteins were identified by mass spectrometry analysis. 31 proteins had a SILAC ratio of 100.0.

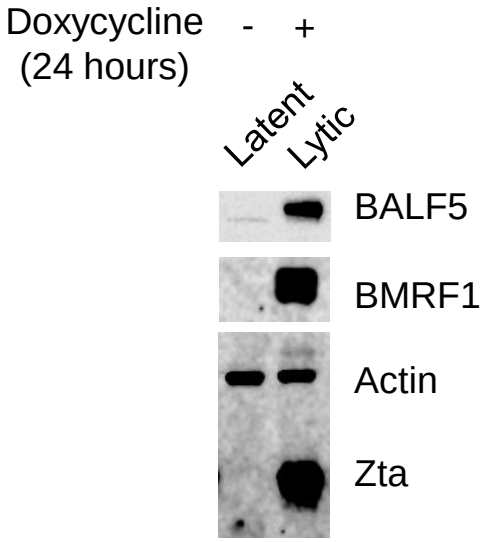


Figure 6.11 Western blot to confirm expression of lytic protein BALF5 and BMRF1. The DNA polymerase BALF5 is only expressed in lytic cycle after induction with doxycycline. The DNA processivity factor BMRF1 is only expressed in lytic cycle after induction with doxycycline. BALF5 is detected by α-BALF5 antibody, BMRF1 is detected by α-BMRF1 (ab30541, Abcam) antibody, Zta is detected by BZ1 antibody. Actin is shown as a loading control and is detected by α-actin antibody (Sigma).

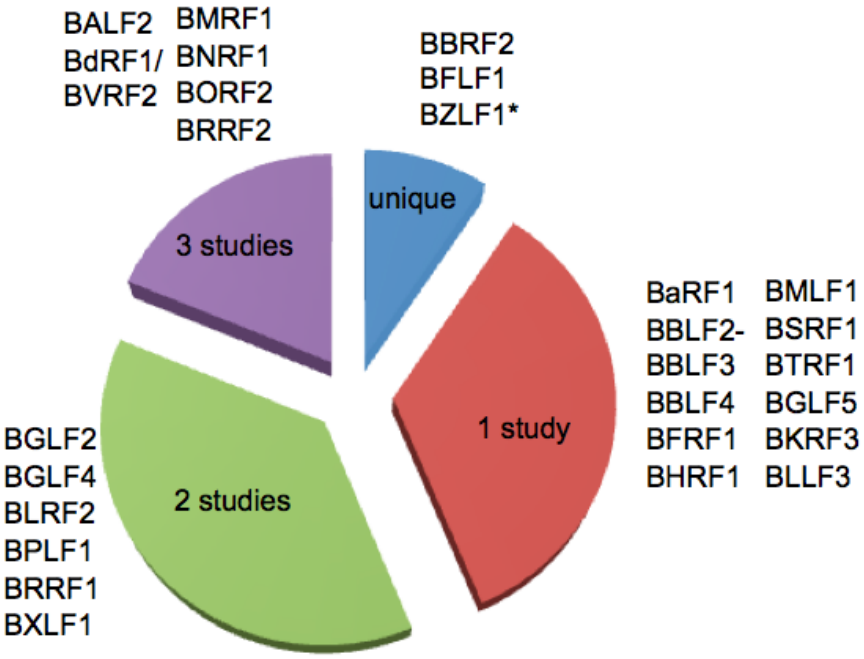


Figure 6.12 SILAC MS analysis of proteins EBV proteins detected in Akata cells during lytic cycle from (Traylen et al, unpublished). The EBV proteins identified are shown in relation to previously published studies. BZLF1 is marked with*, its expression is driven by the doxycycline induced expression vector in these cells. The other studies include Johannsen et al 2004, Dresang et al, 2011 and Koganti et al, 2015.

To demonstrate the expression of some the lytic proteins expressed in lytic cycle, a western blot for the DNA polymerase processivity factor BMRF1 was performed (Figure 6.11). Also a western blot for the DNA polymerase BALF5 was performed (Figure 6.11). The expression of both the DNA polymerase and processivity factor agrees with the mass spectrometry data of only being detected in the Zta-expressing cells. This would be expected, as these two proteins are expressed in the lytic cycle, they are components of the replisome, and are essential for efficient lytic replication of EBV.

Proteins that were not previously identified by mass spectrometry analysis include the virion egress protein BFLF1. BFLF1 is an early lytic protein that is a homolog of HSV1 UL32 and may play a role in cleavage and packaging of the viral genome (Granato et al, 2008). UL32 is involved in the efficient localisation of capsids to nuclear DNA replication compartments (Lamberti & Weller 1998). The *BFLF1* gene has been shown to be essential for the correct packaging of DNA into virion particles, as a knockout virus for the gene shows defective packaging of viral DNA (Pavlova et al. 2013).

Another protein only detected in this study was BBRF2 is a homolog of the HSV1 virion egress protein UL7 (Figure 6.12). HSV UL7 is present within the tegument layers of mature virions, and that the viral protein is localized cytoplasmic domains of infected cells, although it is also detected transiently in the nucleus (Nozawa et al. 2002).

6.3. Discussion

A SILAC mass spectrometry based proteomic approach was implemented to identify and quantitate viral and host proteins from a total cell lysate. Using an enriched population of cells undergoing lytic cycle coupled with SILAC labelling of the cells, mass spectrometry was performed to determine any differences between latent and lytic protein expression of host and viral proteins.

This global proteomics view aimed to detect differences in protein abundance in EBV lytic cycle. SILAC labelled BL cells that could be inducible for Zta expression by this bidirectional inducible cell system allowed the induction of the EBV lytic

cycle. Zta was expressed from the bi-directional promoter with NGFR and GFP. These cells were sorted for NGFR expression to isolate cells undergoing the lytic cycle (Ramasubramanyan et al. 2015).

SILAC proteomics of sorted cells were performed to indicate the expression of viral and cellular proteins by mass spectrometry. Their abundances were quantitated and given a SILAC ratio. The data from the University of Sussex mass spectrometry analysis that I performed was not in depth enough to draw a conclusion. This attempt returned proteins that included some abundant EBV lytic proteins but overall the dataset was limited. Identified CD74 expression agreed with published data. This led to the samples being sent to the University of Bristol Proteomics facility to attempt the in gel digestion and peptide analysis. The samples sent to the University of Bristol Proteomics facility for analysis returned many proteins with significant differences. The results from the study from University of Sussex identified the same abundant lytic proteins as the University of Bristol analysis. But the analysis from University of Sussex was nowhere near the same level of specificity concerning the protein abundance of PSM detected or SILAC ratio. This was either due to a more efficient processing of the sample or a more sensitive detection. The experimental results identified many cellular proteins and also identified half of the EBV lytic proteins characterised as open reading frame expressed during the lytic cycle.

Many abundant cellular proteins were abundantly identified. These included heat shock proteins that were only identified after the expression of Zta and the subsequent initiation of the lytic cycle. Heat shock proteins were shown to be components of the tegument of the EBV virion (Johannsen 2004). The heat shock proteins enriched in lytic cycle included HSP90AB4P and HSP90AA2. An attempt at confirming these two most abundant cellular proteins only identified in the Zta-expressing samples was not possible. There were no commercial antibodies that were able to identify HSP90AB4P and HSP90AA2. It would be interesting to assess if these individual proteins contribute to the lytic cycle through further studies, whether they are only expressed in the lytic cycle, part of the virion or aiding viral protein folding or packaging.

A number of EBV lytic proteins were identified in this study. Some of these proteins have been identified in previous proteomics studies (Johannsen et al. 2004; Dresang et al. 2011; Koganti et al. 2015). Many of the EBV lytic genes and proteins are homologous to the other members of the Human Herpesvirus family, where HSV is used as the reference genome. The identification of some EBV lytic proteins has not been demonstrated at the protein level. This study identified two lytic proteins that had not previously been identified by any proteomics study and mass spectrometry, BFLF1 and BBFR2. Therefore this study was successful by being able to confirm with confidence that these proteins are expressed in vivo. (Traylen et al, submitted). This proteomics study identified proteins that were previously undetected by MS and shown to be expressed at the protein level.

7. Discussion

A novel approach to understand interactions of a key lytic EBV gene and the host during the lytic cycle was undertaken here. An extraction method that was compatible with affinity chromatogaphy and size exclusion chromatography, coupled to SILAC proteomics was developed to attempt to study Zta interactions with potential cellular partners.

Firstly, polyhistidine tagged Zta was transfected into cells and stably expressed. Attempts at extracting this transfected protein from cells using a nuclear extract protocol revealed that the majority of the protein remained in the final pellet. This extraction method was optimized using the addition of a nuclease (benzonase). This allowed the release of the protein into the supernatant. The polyhistidine tag binding to the nickel agarose allowed some protein to be purified.

Optimization of the pull down coupled with SILAC allowed quantitated proteins to be identified after mass spectrometry. The targets of interest returned some promising results. Many repeated pull downs were attempted to confirm these interactions.

FANCA and BRD4 and ELP3 western blots displayed possible isoform variants when probed with the antibody. These were not at the correct molecular weight for the protein therefore it was not possible to assess an interaction with Zta. ELP3 was identified at the expected molecular weight in EBV negative cells but this could not be confirmed in EBV positive cells.

Repeated pull downs with various conditions demonstrated that the attempt to pull down interacting proteins was challenging. The two proteins may interact and perform a vital function in EBV lytic cycle, but to discover the optimal pull down condition where the protein of interest does not bind to the control affinity gel, and only be present in the target protein affinity gel sample proved to be difficult.

As poly-histidine tagged proteins also interact with nickel agarose under denaturing conditions, I explored whether this could aid the affinity purification

step. In order to find Zta interacting proteins a cross-linking step was required. I successfully optimised this process on a small scale. An attempt to scale up this approach revealed a problem with the elution conditions, identifying the next step of optimisation that would be required.

Zta interactions with the host proteome under denaturing conditions to identify potential interactions in EBV positive cells undergoing lytic replication. Although the optimization of this technique was found on a small scale, where hisZta was eluted and precipitated successfully, an attempt to maximize the protein recovery from the pull down sample was not successful in a large scale pull down effort.

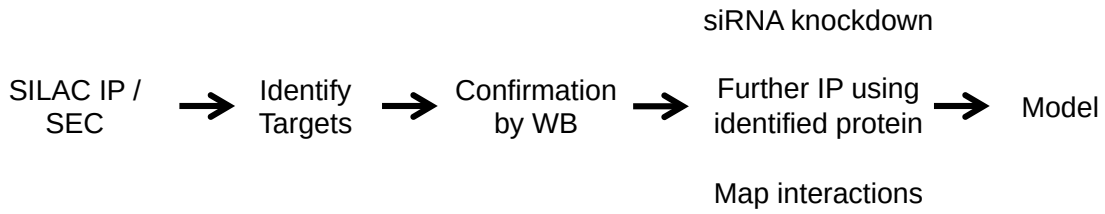
Once a potential protein interaction is assessed with confidence to interact with Zta, a development of this SILAC method could be taken further. Protein interaction screening by quantitative immunoprecipitation combined with knockdown (QUICK) is a tool that takes the SILAC principle with RNA interference and co-immunoprecipitation. This would attempt to identify endogenous proteins that may interact together through a knockdown of the target protein (Selbach & Mann 2006). This method could be used to investigate other cell types where EBV is undergoing lytic cycle.

In a separate approach, analysis of protein complexes by size exclusion chromatography for EBV negative cells transfected with Zta indicated that Zta was associating with multiple complexes of varying molecular weight. Compared against the ZtaAAA mutant, which is replication-defective, different complexes are formed. Mass spectrometry revealed different sub-sets of proteins. This suggests Zta and ZtaAAA are forming different complexes through the C-terminal domain. ZtaAAA may not be able to interact with as many proteins as the wild type Zta, therefore not being able to form complexes essential for lytic replication (Figure 5.9, Figure 5.11)

The elutions with the heaviest complexes were investigated EBV positive cells remaining latent as a control, undergoing lytic cycle with hisZta expression or undergoing the beginning of lytic cycle with viral genome replication with hisZtaAAA expression were investigated. The results of further size exclusion

chromatography experiments were interesting. The abundance of casein kinase II alpha indicates that this protein is part or forms complexes that are of different molecular weight via interactions with Zta in lytic cycle, and also with ZtaAAA during the initiation of lytic cycle. This suggests that casein kinase II alpha may facilitate lytic replication, by being part of a complex or phosphorylating proteins important for viral lytic replication or transcription. In contrast, the relative complex size of Poly (A) RNA polymerase (mitochondrial) western blot did not agree with the mass spectrometry results, while the detection of 53BP1 in the Zta elution fraction was interesting, the western blot did not agree with the mass spectrometry results.

A systematic approach to further answer the global question includes identifying proteins that may interact through either a SILAC labelled immunoprecipitation assay or SILAC labelled size exclusion chromatography for native complexes. After a confirmation by western blot these potential interactions may be assessed through further assays that can contribute to the model of EBV lytic replication.



In order to identify changes in the proteome of cellular and viral genes during the EBV lytic cycle, a SILAC whole cell proteomics approach was performed, utilizing an enrichment method of cells undergoing the lytic cycle. This provided evidence for the detection of two viral proteins by mass spectrometry for the first time. This enrichment method was previously proven to isolate cells expressing Zta and therefore cells undergoing lytic cycle (Ramasubramanyan et al. 2015). The proteins identified in this study included BBRF2, a homologue of the HSV1 virion egress protein UL7. UL7 plays a role in linking tegument proteins of HSV1, therefore BBFR2 may have a similar function for EBV tegument proteins. The identification of BFLF1 protein supports the involvement of BFLF1 protein in cleavage and packaging of the viral genome (Granato et al. 2008). Interestingly,

BFLF1 is a homologue of the HSV1 UL32 gene, which plays a role in HSV1 encapsidation (Lamberti & Weller 1998).

The definitive identification of 40 EBV proteins in BL cells undergoing EBV replication strengthens the understanding of EBV lytic replication and may highlight different targets for future strategies to enable the development of therapeutic interventions to manipulate EBV replication.

A limitation of interpreting a global proteomics study is that some proteins do not generate peptides that can be unambiguously identified. Transcriptome analysis of the viral genes transcribed has identified highly abundant EBV mRNAs (eg BMRF1, BMLF1 and BHRF1) (Dresang et al. 2011; O'Grady et al. 2014; Tierney et al. 2015) but some proteins are not identified from any mass spectrometry study performed for EBV analysis (BBFR2, BFLF1).

Overall this study attempted to study the lytic cycle including Zta at a fine detail through direct protein-protein interactions, and then investigating at protein interactions at a protein-complex level. A novel approach began to investigate this are of interest. A whole cell proteomics analysis did identify proteins that had not been previously identified by mass spectrometry before, giving confidence to these proteins are expressed in the lytic cycle, while confirming and agreeing with the current literature about these viral genes.

8. Bibliography

Adams, A., 1987. Replication of latent Epstein-Barr virus genomes in Raji cells. *Journal of Virology*, 61(5), pp.1743–1746. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC254169/.

Adamson, A.L. & Kenney, S., 2001. Epstein-barr virus immediate-early protein BZLF1 is SUMO-1 modified and disrupts promyelocytic leukemia bodies. *J Virol*, 75(5), pp.2388–2399.

Adamson, A.L. & Kenney, S., 1999. The Epstein-Barr virus BZLF1 protein interacts physically and functionally with the histone acetylase CREB-binding protein. *Journal of virology*, 73(8), pp.6551–6558.

Aho, S. et al., 2000. Ubinuclein, a novel nuclear protein interacting with cellular and viral transcription factors. *The Journal of cell biology*, 148(6), pp.1165–1176.

Alfieri, C., Birkenbach, M. & Kieff, E., 1991. Early events in Epstein-Barr virus infection of human B lymphocytes. *Virology*, 181(2), pp.595–608.

Amon, W., White, R.E. & Farrell, P.J., 2006. Epstein-Barr virus origin of lytic replication mediates association of replicating episomes with promyelocytic leukaemia protein nuclear bodies and replication compartments. *J Gen Virol*, 87(Pt 5), pp.1133–1137.

Anisimova, E. et al., 1984. Effects of n-butyrate and phorbol ester (TPA) on induction of Epstein-Barr virus antigens and cell differentiation. *Archives of virology*, 81(3-4), pp.223–237.

Armstrong, A.A. et al., 1998. Epstein-Barr virus and Hodgkin's disease: further evidence for the three disease hypothesis. *Leukemia*, 12(8), pp.1272–1276.

Asai, R., Kato, A. & Kawaguchi, Y., 2009. Epstein-Barr virus protein kinase BGLF4 interacts with viral transactivator BZLF1 and regulates its transactivation activity. *J Gen Virol*, 90(Pt 7), pp.1575–1581.

AuCoin, D.P. et al., 2002. Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8) contains two functional lytic origins of DNA replication. *Journal of virology*, 76(15), pp.7890–7896.

Baer, R. et al., 1984. DNA sequence and expression of the B95-8 Epstein-Barr virus genome. *Nature*, 310(5974), pp.207–211.

Bailey, S.G. et al., 2009. Functional interaction between Epstein-Barr virus replication protein Zta and host DNA damage response protein 53BP1. *J*

Virol, 83(21), pp.11116–11122. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19656881.

Balandraud, N. et al., 2003. Epstein-Barr virus load in the peripheral blood of patients with rheumatoid arthritis: accurate quantification using real-time polymerase chain reaction. *Arthritis and rheumatism*, 48(5), pp.1223–1228.

Balfour, H.H.J., Dunmire, S.K. & Hogquist, K.A., 2015. Infectious mononucleosis. *Clinical & translational immunology*, 4(2), p.e33.

Barth, S. et al., 2008. Epstein-Barr virus-encoded microRNA miR-BART2 down-regulates the viral DNA polymerase BALF5. *Nucleic acids research*, 36(2), pp.666–675.

Baumann, M. et al., 1998. Activation of the Epstein-Barr virus transcription factor BZLF1 by 12-O-tetradecanoylphorbol-13-acetate-induced phosphorylation. *Journal of virology*, 72(10), pp.8105–8114.

Baumann, M. et al., 1999. Cellular transcription factors recruit viral replication proteins to activate the Epstein-Barr virus origin of lytic DNA replication, oriLyt. *EMBO J*, 18(21), pp.6095–6105.

Baumforth, K.R. et al., 1999. The Epstein-Barr virus and its association with human cancers. *Molecular pathology : MP*, 52(6), pp.307–322.

Bell, P., Lieberman, P.M. & Maul, G.G., 2000. Lytic but not latent replication of epstein-barr virus is associated with PML and induces sequential release of nuclear domain 10 proteins. *J Virol*, 74(24), pp.11800–11810.

Ben-Sasson, S.A. & Klein, G., 1981. Activation of the Epstein-Barr virus genome by 5-aza-cytidine in latently infected human lymphoid lines. *International Journal of Cancer*, 28(2), pp.131–135. Available at: http://www.scopus.com/inward/record.url?eid=2-s2.0-0019410144&partnerID=40&md5=2cbd6ae2cc78087dc8580ed8a506b149.

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

Bhende, P.M. et al., 2005. BZLF1 activation of the methylated form of the BRLF1 immediate-early promoter is regulated by BZLF1 residue 186. *Journal of virology*, 79(12), pp.7338–7348.

Bhende, P.M. et al., 2004. The EBV lytic switch protein, Z, preferentially binds to and activates the methylated viral genome. *Nature genetics*, 36(10), pp.1099–1104.

Brooks, L. et al., 1992. Epstein-Barr virus latent gene transcription in nasopharyngeal carcinoma cells: coexpression of EBNA1, LMP1, and LMP2 transcripts. *Journal of virology*, 66(5), pp.2689–2697.

Burkitt, D.P., 1958. A sarcoma involving the jaws in African children. *The British journal of surgery*, 46(197), pp.218–223.

Burkitt, D.P., 1983. The discovery of Burkitt's lymphoma. *Cancer*, 51(10), pp.1777–1786.

Calderwood, M.A. et al., 2007. Epstein-Barr virus and virus human protein interaction maps. *Proc Natl Acad Sci U S A*, 104(18), pp.7606–7611.

Castagna, M. et al., 1982. Direct activation of calcium-activated, phospholipid-dependent protein kinase by tumor-promoting phorbol esters. *The Journal of biological chemistry*, 257(13), pp.7847–7851.

Cesarman, E. & Mesri, E.A., 1999. Virus-associated lymphomas. *Current opinion in oncology*, 11(5), pp.322–332.

Chagin, V.O., Stear, J.H. & Cardoso, M.C., 2010. Organization of DNA replication. *Cold Spring Harbor perspectives in biology*, 2(4), p.a000737.

Chang, C.-W. et al., 2012. Epstein-Barr virus protein kinase BGLF4 targets the nucleus through interaction with nucleoporins. *Journal of virology*, 86(15), pp.8072–8085.

Chang, L.K. et al., 2010. MCAF1 and synergistic activation of the transcription of Epstein-Barr virus lytic genes by Rta and Zta. *Nucleic Acids Res*, 38(14), pp.4687–4700.

Chang, Y.N. et al., 1990. The Epstein-Barr virus Zta transactivator: a member of the bZIP family with unique DNA-binding specificity and a dimerization domain that lacks the characteristic heptad leucine zipper motif. *Journal of Virology*, 64(7), pp.3358–3369. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC249580/.

Chee, M.S., Lawrence, G.L. & Barrell, B.G., 1989. Alpha-, beta- and gammaherpesviruses encode a putative phosphotransferase. *The Journal of general virology*, 70 (Pt 5), pp.1151–1160.

Chen, C.C. et al., 2011. Enhancement of Zta-activated lytic transcription of Epstein-Barr virus by Ku80. *J Gen Virol*, 92(Pt 3), pp.661–668.

Chen, M.-R. et al., 2000. A Protein Kinase Activity Associated with Epstein-Barr Virus BGLF4 Phosphorylates the Viral Early Antigen EA-D In Vitro. *Journal of Virology*, 74(7), pp.3093–3104. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC111808/.

Cheung, A. & Kieff, E., 1982. Long internal direct repeat in Epstein-Barr virus DNA. *Journal of virology*, 44(1), pp.286–294.

Chevallier-Greco, A. et al., 1986. Both Epstein-Barr virus (EBV)-encoded trans-acting factors, EB1 and EB2, are required to activate transcription from an

EBV early promoter. *EMBO J*, 5(12), pp.3243–3249. Available at: http://www.ncbi.nlm.nih.gov/pubmed/3028777?dopt=Citation.

Chi, T. & Carey, M., 1993. The ZEBRA activation domain: modular organization and mechanism of action. *Molecular and cellular biology*, 13(11), pp.7045–7055.

Chiu, Y.-F., Sugden, A.U. & Sugden, B., 2013. Epstein-Barr Viral Productive Amplification Reprograms Nuclear Architecture, DNA Replication and Histone Deposition. *Cell host & microbe*, 14(6), pp.607–618. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3995538/.

Chua, H.H. et al., 2012. p53 and Sp1 cooperate to regulate the expression of Epstein-Barr viral Zta protein. *J Med Virol*, 84(8), pp.1279–1288.

Cohen, J.I., 2000. Epstein-Barr virus infection. *The New England journal of medicine*, 343(7), pp.481–492.

Collas, P., 2010. The current state of chromatin immunoprecipitation. *Molecular biotechnology*, 45(1), pp.87–100.

Countryman, J. & Miller, G., 1985. Activation of expression of latent Epstein-Barr herpesvirus after gene transfer with a small cloned subfragment of heterogeneous viral DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 82(12), pp.4085–4089. Available at: http://www.scopus.com/inward/record.url?eid=2-s2.0-1642457437&partnerID=tZOtx3y1.

Cox, J. & Mann, M., 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature biotechnology*, 26(12), pp.1367–1372.

Crawford, D.H., Rickinson, A.B. & Johannessen, I., 2014. Making Sense of a Human Cancer Virus. In *Cancer Virus: The Story of Epstein-Barr-Virus*. OUP, Oxford, p. 172.

Daikoku, T. et al., 2005. Architecture of replication compartments formed during Epstein-Barr virus lytic replication. *J Virol*, 79(6), pp.3409–3418.

Daikoku, T. et al., 2006. Postreplicative mismatch repair factors are recruited to Epstein-Barr virus replication compartments. *J Biol Chem*, 281(16), pp.11422–11430.

Dambaugh, T.R. & Kieff, E., 1982. Identification and nucleotide sequences of two similar tandem direct repeats in Epstein-Barr virus DNA. *Journal of virology*, 44(3), pp.823–833.

Davison, A.J., 2007. *Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis.*,

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

Dolan, A. et al., 2006. The genome of Epstein-Barr virus type 2 strain AG876. *Virology*, 350(1), pp.164–170.

Dresang, L.R. et al., 2011. Coupled transcriptome and proteome analysis of human lymphotropic tumor viruses: insights on the detection and discovery of viral genes. *BMC genomics*, 12, p.625.

Duckworth, D.H., 1976. "Who discovered bacteriophage?". *Bacteriological reviews*, 40(4), pp.793–802.

Durkop, H. et al., 1999. Tumor necrosis factor receptor-associated factor 1 is overexpressed in Reed-Sternberg cells of Hodgkin's disease and Epstein-Barr virus-transformed lymphoid cells. *Blood*, 93(2), pp.617–623.

El-Guindy, A. et al., 2013. Essential role of Rta in lytic DNA replication of Epstein-Barr virus. *J Virol*, 87(1), pp.208–223.

El-Guindy, A., Heston, L. & Miller, G., 2010. A subset of replication proteins enhances origin recognition and lytic replication by the Epstein-Barr virus ZEBRA protein. *PLoS Pathog*, 6(8), p.e1001054.

El-Guindy, A.S. & Miller, G., 2004. Phosphorylation of Epstein-Barr virus ZEBRA protein at its casein kinase 2 sites mediates its ability to repress activation of a viral lytic cycle late gene by Rta. *Journal of virology*, 78(14), pp.7634–7644.

Epstein, M.A., Achong, B.G. & Barr, Y.M., 1964. VIRUS PARTICLES IN CULTURED LYMPHOBLASTS FROM BURKITT'S LYMPHOMA. *Lancet (London, England)*, 1(7335), pp.702–703.

Faggioni, A. et al., 1986. Calcium modulation activates Epstein-Barr virus genome in latently infected cells. *Science (New York, N.Y.)*, 232(4757), pp.1554–1556.

Farrell, P.J. et al., 1989. Epstein-Barr virus BZLF1 trans-activator specifically binds to a consensus AP-1 site and is related to c-fos. *EMBO J*, 8(1), pp.127–132.

Feederle, R. et al., 2007. Epstein-Barr virus B95.8 produced in 293 cells shows marked tropism for differentiated primary epithelial cells and reveals interindividual variation in susceptibility to viral infection. *International journal of cancer. Journal international du cancer*, 121(3), pp.588–594.

Feederle, R. et al., 2009. The Epstein-Barr virus alkaline exonuclease BGLF5 serves pleiotropic functions in virus replication. *J Virol*, 83(10), pp.4952–4962.

Feederle, R. et al., 2000. The Epstein-Barr virus lytic program is controlled by the co-operative functions of two transactivators. *The EMBO journal*, 19(12), pp.3080–3089.

Fernandez, A.F. et al., 2009. The dynamic DNA methylomes of double-stranded DNA viruses associated with human cancer. *Genome research*, 19(3), pp.438–451.

Fingeroth, J.D. et al., 1984. Epstein-Barr virus receptor of human B lymphocytes is the C3d receptor CR2. *Proceedings of the National Academy of Sciences of the United States of America*, 81(14), pp.4510–4514.

Fixman, E.D., Hayward, G.S. & Hayward, S.D., 1992. trans-acting requirements for replication of Epstein-Barr virus ori-Lyt. *J Virol*, 66(8), pp.5030–5039.

Flemington, E. & Speck, S.H., 1990a. Autoregulation of Epstein-Barr virus putative lytic switch gene BZLF1. *Journal of virology*, 64(3), pp.1227–1232.

Flemington, E. & Speck, S.H., 1990b. Evidence for coiled-coil dimer formation by an Epstein-Barr virus transactivator that lacks a heptad repeat of leucine residues. *Proceedings of the National Academy of Sciences of the United States of America*, 87(23), pp.9459–9463.

Flemington, E.K. et al., 1992. Characterization of the Epstein-Barr virus BZLF1 protein transactivation domain. *Journal of virology*, 66(2), pp.922–929.

Flower, K. et al., 2011. Epigenetic control of viral life-cycle by a DNA-methylation dependent transcription factor. *PLoS One*, 6(10), p.e25922. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22022468.

Flower, K. et al., 2010. Evaluation of a prediction protocol to identify potential targets of epigenetic reprogramming by the cancer associated Epstein Barr virus. *PLoS One*, 5(2), p.e9443. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20195470.

Fox, C.P., Shannon-Lowe, C. & Rowe, M., 2011. Deciphering the role of Epstein-Barr virus in the pathogenesis of T and NK cell lymphoproliferations. *Herpesviridae*, 2, p.8. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3180299/.

Fox, R.I., Pearson, G. & Vaughan, J.H., 1986. Detection of Epstein-Barr virus-associated antigens and DNA in salivary gland biopsies from patients with Sjogren's syndrome. *Journal of immunology (Baltimore, Md. : 1950)*, 137(10), pp.3162–3168.

Francis, A. et al., 1999. Amino acid substitutions reveal distinct functions of serine 186 of the ZEBRA protein in activation of early lytic cycle genes and synergy with the Epstein-Barr virus R transactivator. *Journal of virology*, 73(6), pp.4543–4551.

Fraser, K.B. et al., 1979. Increased tendency to spontaneous in-vitro lymphocyte transformation in clinically active multiple sclerosis. *Lancet (London, England)*, 2(8145), pp.175–176.

Fredricks, D.N. & Relman, D.A., 1996. Sequence-based identification of microbial pathogens: a reconsideration of Koch's postulates. *Clinical microbiology reviews*, 9(1), pp.18–33.

Fust, G., 2011. The role of the Epstein-Barr virus in the pathogenesis of some autoimmune disorders - Similarities and differences. *European journal of microbiology & immunology*, 1(4), pp.267–278.

Gaipl, U.S. et al., 2007. Clearance deficiency and systemic lupus erythematosus (SLE). *Journal of autoimmunity*, 28(2-3), pp.114–121.

Gallagher, A. et al., 1999. Detection of Epstein-Barr virus (EBV) genomes in the serum of patients with EBV-associated Hodgkin's disease. *International journal of cancer. Journal international du cancer*, 84(4), pp.442–448.

Gao, Z. et al., 1998. The Epstein-Barr virus lytic transactivator Zta interacts with the helicase-primase replication proteins. *J Virol*, 72(11), pp.8559–8567.

Gavin, A.-C. et al., 2006. Proteome survey reveals modularity of the yeast cell machinery. *Nature*, 440(7084), pp.631–636.

Gradoville, L. et al., 2002. Protein kinase C-independent activation of the Epstein-Barr virus lytic cycle. *Journal of virology*, 76(11), pp.5612–5626.

Granato, M. et al., 2008. Deletion of Epstein-Barr virus BFLF2 leads to impaired viral DNA packaging and primary egress as well as to the production of defective viral particles. *Journal of virology*, 82(8), pp.4042–4051.

Greenspan, J.S. et al., 1985. Replication of Epstein-Barr virus within the epithelial cells of oral "hairy" leukoplakia, an AIDS-associated lesion. *The New England journal of medicine*, 313(25), pp.1564–1571.

Grogan, E. et al., 1987. Transfection of a rearranged viral DNA fragment, WZhet, stably converts latent Epstein-Barr viral infection to productive infection in lymphoid cells. *Proceedings of the National Academy of Sciences of the United States of America*, 84(5), pp.1332–1336.

Gross, A.J. et al., 2005. EBV and systemic lupus erythematosus: a new perspective. *Journal of immunology (Baltimore, Md. : 1950)*, 174(11), pp.6599–6607.

Gruffat, H. et al., 2002. Epstein-Barr virus mRNA export factor EB2 is essential for production of infectious virus. *Journal of virology*, 76(19), pp.9635–9644.

Gutsch, D.E. et al., 1994. The bZIP transactivator of Epstein-Barr virus, BZLF1, functionally and physically interacts with the p65 subunit of NF-kappa B. *Mol Cell Biol*, 14(3), pp.1939–1948.

Hammerschmidt, W. & Sugden, B., 1988. Identification and characterization of oriLyt, a lytic origin of DNA replication of Epstein-Barr virus. *Cell*, 55(3), pp.427–433.

Hardwick, J.M., Lieberman, P.M. & Hayward, S.D., 1988. A new Epstein-Barr virus transactivator, R, induces expression of a cytoplasmic early antigen. *Journal of virology*, 62(7), pp.2274–2284.

Harris, N.L. et al., 1994. A revised European-American classification of lymphoid neoplasms: a proposal from the International Lymphoma Study Group. *Blood*, 84(5), pp.1361–1392.

Zur Hausen, H. et al., 1970. EBV DNA in biopsies of Burkitt tumours and anaplastic carcinomas of the nasopharynx. *Nature*, 228(5276), pp.1056–1058.

Zur Hausen, H. et al., 1978. Persisting oncogenic herpesvirus induced by the tumour promotor TPA. *Nature*, 272(5651), pp.373–375.

Henderson, S. et al., 1993. Epstein-Barr virus-coded BHRF1 protein, a viral homologue of Bcl-2, protects human B cells from programmed cell death. *Proceedings of the National Academy of Sciences of the United States of America*, 90(18), pp.8479–8483.

Henle, G. & Henle, W., 1966. Immunofluorescence in cells derived from Burkitt's lymphoma. *Journal of Bacteriology*, 91(3), pp.1248–1256.

Henle, G., Henle, W. & Diehl, V., 1968. Relation of Burkitt's tumor-associated herpes-ytpe virus to infectious mononucleosis. *Proceedings of the National Academy of Sciences of the United States of America*, 59(1), pp.94–101.

Henle, W. et al., 1987. Antibody responses to Epstein-Barr virus-determined nuclear antigen (EBNA)-1 and EBNA-2 in acute and chronic Epstein-Barr virus infection. *Proceedings of the National Academy of Sciences of the United States of America*, 84(2), pp.570–574.

Hernando, H. et al., 2013. The B cell transcription program mediates hypomethylation and overexpression of key genes in Epstein-Barr virus-associated proliferative conversion. *Genome biology*, 14(1), p.R3.

Heston, L. et al., 2006. Amino acids in the basic domain of Epstein-Barr virus ZEBRA protein play distinct roles in DNA binding, activation of early lytic gene expression, and promotion of viral DNA replication. *Journal of virology*, 80(18), pp.9115–9133.

Hicks, M.R. et al., 2001. Biophysical analysis of natural variants of the multimerization region of Epstein-Barr virus lytic-switch protein BZLF1. *Journal of virology*, 75(11), pp.5381–5384.

Hicks, M.R., Al-Mehairi, S.S. & Sinclair, A.J., 2003. The zipper region of Epstein-Barr virus bZIP transcription factor Zta is necessary but not sufficient to direct DNA binding. *Journal of virology*, 77(14), pp.8173–8177.

Hislop, A.D. et al., 2007. A CD8+ T cell immune evasion protein specific to Epstein-Barr virus and its close relatives in Old World primates. *J Exp Med*, 204(8), pp.1863–1873. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17620360.

Hislop, A.D., 2015. Early virological and immunological events in Epstein-Barr virus infection. *Current opinion in virology*, 15, pp.75–79.

Hjalgrim, H. et al., 2010. HLA-A alleles and infectious mononucleosis suggest a critical role for cytotoxic T-cell response in EBV-related Hodgkin lymphoma. *Proceedings of the National Academy of Sciences of the United States of America*, 107(14), pp.6400–6405.

Hoagland, R.J., 1955. The transmission of infectious mononucleosis. *The American journal of the medical sciences*, 229(3), pp.262–272.

Hoebe, E.K. et al., 2012. Epstein-Barr virus-encoded BARF1 protein is a decoy receptor for macrophage colony stimulating factor and interferes with macrophage differentiation and activation. *Viral immunology*, 25(6), pp.461–470.

Hong, G.K. et al., 2004. The BRRF1 early gene of Epstein-Barr virus encodes a transcription factor that enhances induction of lytic infection by BRLF1. *Journal of virology*, 78(10), pp.4983–4992.

Hori, K., Sen, A. & Artavanis-Tsakonas, S., 2013. Notch signaling at a glance. *Journal of cell science*, 126(Pt 10), pp.2135–2140.

Horst, D. et al., 2009. Specific targeting of the EBV lytic phase protein BNLF2a to the transporter associated with antigen processing results in impairment of HLA class I-restricted antigen presentation. *Journal of immunology (Baltimore, Md. : 1950)*, 182(4), pp.2313–2324.

Hummel, M., Thorley-Lawson, D. & Kieff, E., 1984. An Epstein-Barr virus DNA fragment encodes messages for the two major envelope glycoproteins (gp350/300 and gp220/200). *Journal of virology*, 49(2), pp.413–417.

Hung, S.C., Kang, M.S. & Kieff, E., 2001. Maintenance of Epstein-Barr virus (EBV) oriP-based episomes requires EBV-encoded nuclear antigen-1 chromosome-binding domains, which can be replaced by high-mobility

group-I or histone H1. *Proceedings of the National Academy of Sciences of the United States of America*, 98(4), pp.1865–1870.

Jaffe, E.S. & Ralfkiaer, E., 2001. Mature T-cell and NK-cell neoplasms. *World Health Organization Classification of Tumors. Pathology and Genetics of Tumors of Haematopoietic and Lymphoid Tissues. IARC Press, Lyon, France*, pp.191–194.

Jenkins, P.J., Binne, U.K. & Farrell, P.J., 2000. Histone acetylation and reactivation of Epstein-Barr virus from latency. *Journal of virology*, 74(2), pp.710–720.

Jin, X.W. & Speck, S.H., 1992. Identification of critical cis elements involved in mediating Epstein-Barr virus nuclear antigen 2-dependent activity of an enhancer located upstream of the viral BamHI C promoter. *Journal of virology*, 66(5), pp.2846–2852.

Johannsen, E. et al., 2004. Proteins of purified Epstein-Barr virus. *Proceedings of the National Academy of Sciences of the United States of America*, 101(46), pp.16286–16291.

Jones, P.A. & Baylin, S.B., 2007. The epigenomics of cancer. *Cell*, 128(4), pp.683–692.

Jonsson, R. et al., 2011. The complexity of Sjogren's syndrome: novel aspects on pathogenesis. *Immunology letters*, 141(1), pp.1–9.

Karlsson, Q.H. et al., 2008. Methylated DNA recognition during the reversal of epigenetic silencing is regulated by cysteine and serine residues in the Epstein-Barr virus lytic switch protein. *PLoS Pathog*, 4(3), p.e1000005. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18369464.

Kato, K. et al., 2003. Identification of protein kinases responsible for phosphorylation of Epstein-Barr virus nuclear antigen leader protein at serine-35, which regulates its coactivator function. *The Journal of general virology*, 84(Pt 12), pp.3381–3392.

Kenney, S.C. et al., 1992. The cellular oncogene c-myb can interact synergistically with the Epstein-Barr virus BZLF1 transactivator in lymphoid cells. *Molecular and cellular biology*, 12(1), pp.136–146.

Kiehl, A. & Dorsky, D.I., 1995. Bipartite DNA-binding region of the Epstein-Barr virus BMRF1 product essential for DNA polymerase accessory function. *J Virol*, 69(3), pp.1669–1677.

Kintner, C.R. & Sugden, B., 1979. The structure of the termini of the DNA of Epstein-Barr virus. *Cell*, 17(3), pp.661–671.

Kirchmaier, A.L. & Sugden, B., 1995. Plasmid maintenance of derivatives of oriP of Epstein-Barr virus. *Journal of virology*, 69(2), pp.1280–1283.

Kirkwood, K.J. et al., 2013. Characterization of Native Protein Complexes and Protein Isoform Variation Using Size-fractionation-based Quantitative Proteomics. *Molecular & Cellular Proteomics : MCP*, 12(12), pp.3851–3873. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3861729/.

Koganti, S. et al., 2015. Cellular STAT3 functions via PCBP2 to restrain Epstein-Barr Virus lytic activation in B lymphocytes. *Journal of virology*, 89(9), pp.5002–5011.

Kouzarides, T. et al., 1991. The BZLF1 protein of EBV has a coiled coil dimerisation domain without a heptad leucine repeat but with homology to the C/EBP leucine zipper. *Oncogene*, 6(2), pp.195–204.

Krogan, N.J. et al., 2006. Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature*, 440(7084), pp.637–643.

Kudoh, A. et al., 2009. Homologous recombinational repair factors are recruited and loaded onto the viral DNA genome in Epstein-Barr virus replication compartments. *Journal of virology*, 83(13), pp.6641–6651.

Kutok, J.L. & Wang, F., 2006. Spectrum of Epstein-Barr virus-associated diseases. *Annual review of pathology*, 1, pp.375–404.

Laichalk, L.L. & Thorley-Lawson, D.A., 2005. Terminal differentiation into plasma cells initiates the replicative cycle of Epstein-Barr virus in vivo. *Journal of virology*, 79(2), pp.1296–1307.

Lamberti, C. & Weller, S.K., 1998. The herpes simplex virus type 1 cleavage/packaging protein, UL32, is involved in efficient localization of capsids to replication compartments. *Journal of virology*, 72(3), pp.2463–2473.

Landschulz, W.H., Johnson, P.F. & McKnight, S.L., 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science (New York, N.Y.)*, 240(4860), pp.1759–1764.

Laux, G., Economou, A. & Farrell, P.J., 1989. The terminal protein gene 2 of Epstein-Barr virus is transcribed from a bidirectional latent promoter region. *The Journal of general virology*, 70 (Pt 11, pp.3079–3084.

Lemon, S.M. et al., 1977. Replication of EBV in epithelial cells during infectious mononucleosis. *Nature*, 268(5617), pp.268–270.

Levine & Enquist, 2007. *Fields Virology*, Philadelphia: Lippincott Williams & Wilkins.

Levine, P.H. et al., 1971. Elevated antibody titers to epstein-barr virus in Hodgkin's disease. *Cancer*, 27(2), pp.416–421. Available at: http://dx.doi.org/10.1002/1097-0142(197102)27:2<416::AID-CNCR2820270227>3.0.CO.

Li, L. et al., 2012. Methylation profiling of Epstein-Barr virus immediate-early gene promoters, BZLF1 and BRLF1 in tumors of epithelial, NK- and B-cell origins. *BMC cancer*, 12, p.125.

Li, Q. et al., 1997. Epstein-Barr virus uses HLA class II as a cofactor for infection of B lymphocytes. *Journal of virology*, 71(6), pp.4657–4662.

Li, Q.X. et al., 1992. Epstein-Barr virus infection and replication in a human epithelial cell system. *Nature*, 356(6367), pp.347–350.

Liao, G., Wu, F.Y. & Hayward, S.D., 2001. Interaction with the Epstein-Barr virus helicase targets Zta to DNA replication compartments. *J Virol*, 75(18), pp.8792–8802.

Lieberman, P.M. et al., 1986. Promiscuous trans activation of gene expression by an Epstein-Barr virus-encoded early nuclear protein. *Journal of virology*, 60(1), pp.140–148.

Lieberman, P.M. & Berk, A.J., 1994. A mechanism for TAFs in transcriptional activation: activation domain enhancement of TFIID-TFIIA--promoter DNA complex formation. *Genes & development*, 8(9), pp.995–1006.

Lieberman, P.M. & Berk, A.J., 1990. In vitro transcriptional activation, dimerization, and DNA-binding specificity of the Epstein-Barr virus Zta protein. *Journal of virology*, 64(6), pp.2560–2568.

Lieberman, P.M. & Berk, A.J., 1991. The Zta trans-activator protein stabilizes TFIID association with promoter DNA by direct protein-protein interaction. *Genes & development*, 5(12 B), pp.2441–2454.

Lin, Z. et al., 2013. Whole-genome sequencing of the Akata and Mutu Epstein-Barr virus strains. *J Virol*, 87(2), pp.1172–1182.

Lindquist, S. & Craig, E.A., 1988. The heat-shock proteins. *Annual review of genetics*, 22, pp.631–677.

Liu, F. & Zhou, H., 2007. Comparative virion structures of human herpesviruses. In A. Arvin, C.-F. G, & et al Mocarski E, eds. *Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis*. Cambridge: Cambridge University Press; 2007.

Liu, S. et al., 1997. Binding of the ubiquitous cellular transcription factors Sp1 and Sp3 to the ZI domains in the Epstein-Barr virus lytic switch BZLF1 gene promoter. *Virology*, 228(1), pp.11–18.

Lo, A.K.-F. et al., 2013. Inhibition of the LKB1-AMPK pathway by the Epstein-Barr virus-encoded LMP1 promotes proliferation and transformation of human nasopharyngeal epithelial cells. *The Journal of pathology*, 230(3), pp.336–346.

Longnecker, R. & Neipel, F., 2007. *Introduction to the human γ-herpesviruses. In: Arvin A, Campadelli-Fiume G, Mocarski E, et al., editors. Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis.*,

Loren, A.W. et al., 2003. Post-transplant lymphoproliferative disorder: a review. *Bone marrow transplantation*, 31(3), pp.145–155.

Luka, J., Kallin, B. & Klein, G., 1979. Induction of the Epstein-Barr virus (EBV) cycle in latently infected cells by n-butyrate. *Virology*, 94(1), pp.228–231.

Lustig, A. & Levine, A.J., 1992. One hundred years of virology. *Journal of virology*, 66(8), pp.4629–4631.

Magrath, I., 2012. Epidemiology: Clues to the pathogenesis of Burkitt lymphoma. *British Journal of Haematology*, 156(6), pp.744–756.

Makhov, A.M. et al., 2004. The Epstein-Barr virus polymerase accessory factor BMRF1 adopts a ring-shaped structure as visualized by electron microscopy. *Journal of Biological Chemistry*, 279(39), pp.40358–40361.

Marechal, V. et al., 1999. Mapping EBNA-1 domains involved in binding to metaphase chromosomes. *Journal of virology*, 73(5), pp.4385–4392.

McDonald, C.M., Petosa, C. & Farrell, P.J., 2009. Interaction of Epstein-Barr virus BZLF1 C-terminal tail structure and core zipper McDonald, C.M., Petosa, C. & Farrell, P.J., 2009. Interaction of Epstein-Barr virus BZLF1 C-terminal tail structure and core zipper is required for DNA replication but not f. *J Virol*, 83(7), pp.3397–3401.

McGeoch, D.J., 1989. The genomes of the human herpesviruses: contents, relationships, and evolution. *Annual review of microbiology*, 43, pp.235–265.

McShane, M.P. & Longnecker, R., 2004. Cell-surface expression of a mutated Epstein-Barr virus glycoprotein B allows fusion independent of other viral proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 101(50), pp.17474–17479.

Miller, N. & Hutt-Fletcher, L.M., 1992. Epstein-Barr virus enters B cells and epithelial cells by different routes. *Journal of Virology*, 66(6), pp.3409–3414. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC241121/.

Mishra, G.R. et al., 2006. Human protein reference database--2006 update. *Nucleic acids research*, 34(Database issue), pp.D411–4.

Molesworth, S.J. et al., 2000. Epstein-Barr virus gH is essential for penetration of B cells but also plays a role in attachment of virus to epithelial cells. *Journal of virology*, 74(14), pp.6324–6332.

Moore, P.S. & Chang, Y., 2010. Why do viruses cause cancer? Highlights of the first century of human tumour virology. *Nature reviews. Cancer*, 10(12), pp.878–889.

Morgenstern, J.P. & Land, H., 1990. Advanced mammalian gene transfer: high titre retroviral vectors with multiple drug selection markers and a complementary helper-free packaging cell line. *Nucleic acids research*, 18(12), pp.3587–3596.

Mullen, T.E. & Marzluff, W.F., 2008. Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5' to 3' and 3' to 5'. *Genes & development*, 22(1), pp.50–65.

Murayama, K. et al., 2009. Crystal structure of epstein-barr virus DNA polymerase processivity factor BMRF1. *J Biol Chem*, 284(51), pp.35896–35905.

Murray, P.G. et al., 2001. Expression of the tumour necrosis factor receptor-associated factors 1 and 2 in Hodgkin's disease. *The Journal of pathology*, 194(2), pp.158–164.

Murray, P.G. et al., 1992. Immunohistochemical demonstration of the Epstein-Barr virus-encoded latent membrane protein in paraffin sections of Hodgkin's disease. *The Journal of pathology*, 166(1), pp.1–5.

Nemerow, G.R. et al., 1985. Identification and characterization of the Epstein-Barr virus receptor on human B lymphocytes and its relationship to the C3d complement receptor (CR2). *Journal of virology*, 55(2), pp.347–351.

Neuhierl, B. & Delecluse, H.J., 2006. The Epstein-Barr virus BMRF1 gene is essential for lytic virus replication. *J Virol*, 80(10), pp.5078–5081.

Nonoyama, M. & Pagano, J.S., 1972. Separation of Epstein-Barr virus DNA from large chromosomal DNA in non-virus-producing cells. *Nature: New biology*, 238(84), pp.169–171. Available at: http://www.scopus.com/inward/record.url?eid=2-s2.0-0015499229&partnerID=tZOtx3y1.

Norio, P. & Schildkraut, C.L., 2001. Visualization of DNA replication on individual Epstein-Barr virus episomes. *Science (New York, N.Y.)*, 294(5550), pp.2361–2364.

Norton, V.G. et al., 1989. Histone acetylation reduces nucleosome core particle linking number change. *Cell*, 57(3), pp.449–457.

Nozawa, N. et al., 2002. Identification and characterization of the UL7 gene product of herpes simplex virus type 2. *Virus genes*, 24(3), pp.257–266.

Nutter, L.M. et al., 1987. Induction of virus enzymes by phorbol esters and n-butyrate in Epstein-Barr virus genome-carrying Raji cells. *Cancer research*, 47(16), pp.4407–4412.

O'Grady, T. et al., 2014. Global bidirectional transcription of the Epstein-Barr virus genome during reactivation. *Journal of virology*, 88(3), pp.1604–1616.

Oda, T. et al., 2000. Epstein-Barr virus lacking glycoprotein gp85 cannot infect B cells and epithelial cells. *Virology*, 276(1), pp.52–58.

Orem, J. et al., 2007. Burkitt's lymphoma in Africa, a review of the epidemiology and etiology. *African health sciences*, 7(3), pp.166–175.

Palermo, R.D., Webb, H.M. & West, M.J., 2011. RNA polymerase II stalling promotes nucleosome occlusion and pTEFb recruitment to drive immortalization by Epstein-Barr virus. *PLoS pathogens*, 7(10), p.e1002334.

Pallesen, G. et al., 1991. Expression of Epstein-Barr virus latent gene products in tumour cells of Hodgkin's disease. *Lancet (London, England)*, 337(8737), pp.320–322.

Palser, A.L. et al., 2015. Genome diversity of Epstein-Barr virus from multiple tumor types and normal infection. *Journal of virology*, 89(10), pp.5222–5237.

Panier, S. & Boulton, S.J., 2014. Double-strand break repair: 53BP1 comes into focus. *Nat Rev Mol Cell Biol*, 15(1), pp.7–18. Available at: http://dx.doi.org/10.1038/nrm3719.

Papior, P. et al., 2012. Open chromatin structures regulate the efficiencies of pre-RC formation and replication initiation in Epstein-Barr virus. *J Cell Biol*, 198(4), pp.509–528.

Park, R. et al., 2008. Mutations of amino acids in the DNA-recognition domain of Epstein-Barr virus ZEBRA protein alter its sub-nuclear localization and affect formation of replication compartments. *Virology*, 382(2), pp.145–162.

Parker, B.D. et al., 1990. Sequence and transcription of Raji Epstein-Barr virus DNA spanning the B95-8 deletion region. *Virology*, 179(1), pp.339–346.

Pathmanathan, R. et al., 1995. Clonal proliferations of cells infected with Epstein-Barr virus in preinvasive lesions related to nasopharyngeal carcinoma. *The New England journal of medicine*, 333(11), pp.693–698.

Pavlova, S. et al., 2013. An Epstein-Barr Virus Mutant Produces Immunogenic Defective Particles Devoid of Viral DNA. *Journal of Virology*, 87(4), pp.2011–2022. Available at: <Go to ISI>://WOS:000314072000009.

Penn, I., 1994. The problem of cancer in organ transplant recipients: an overview. *Transplantation science*, 4(1), pp.23–32.

Perkins, D.N. et al., 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*, 20(18), pp.3551–3567.

Perumal, S.K. et al., 2010. Single-molecule studies of DNA replisome function. *Biochimica et biophysica acta*, 1804(5), pp.1094–1112.

Petosa, C. et al., 2006. Structural basis of lytic cycle activation by the Epstein-Barr virus ZEBRA protein. *Mol Cell*, 21(4), pp.565–572.

Pfeffer, S. et al., 2004. Identification of virus-encoded microRNAs. *Science (New York, N.Y.)*, 304(5671), pp.734–736.

Poole, B.D. et al., 2009. Aberrant Epstein-Barr viral infection in systemic lupus erythematosus. *Autoimmunity reviews*, 8(4), pp.337–342.

Prince, S. et al., 2003. Latent membrane protein 1 inhibits Epstein-Barr virus lytic cycle induction and progress via different mechanisms. *Journal of virology*, 77(8), pp.5000–5007.

Ragoczy, T. & Miller, G., 1999. Role of the epstein-barr virus RTA protein in activation of distinct classes of viral lytic cycle genes. *Journal of virology*, 73(12), pp.9858–9866.

Ramasubramanyan, S. et al., 2015. Epstein-Barr virus transcription factor Zta acts through distal regulatory elements to directly control cellular gene expression. *Nucleic acids research*, 43(7), pp.3563–3577.

Ramasubramanyan, S. et al., 2012. Genome-wide analyses of Zta binding to the Epstein-Barr virus genome reveals interactions in both early and late lytic cycles and an epigenetic switch leading to an altered binding profile. *Journal of virology*, 86(23), pp.12494–12502.

Rawlins, D.R. et al., 1985. Sequence-specific DNA binding of the Epstein-Barr virus nuclear antigen (EBNA-1) to clustered sites in the plasmid maintenance region. *Cell*, 42(3), pp.859–868.

Rea, T.D. et al., 2001. Prospective study of the natural history of infectious mononucleosis caused by Epstein-Barr virus. *The Journal of the American Board of Family Practice / American Board of Family Practice*, 14(4), pp.234–242.

Reichart, P.A. et al., 1989. Oral hairy leukoplakia: observations in 95 cases and review of the literature. *Journal of oral pathology & medicine : official publication of the International Association of Oral Pathologists and the American Academy of Oral Pathology*, 18(7), pp.410–415.

Rennekamp, A.J. & Lieberman, P.M., 2011. Initiation of Epstein-Barr virus lytic replication requires transcription and the formation of a stable RNA-DNA hybrid molecule at OriLyt. *J Virol*, 85(6), pp.2837–2850.

Rennekamp, A.J. & Lieberman, P.M., 2010. Initiation of lytic DNA replication in Epstein-Barr virus: search for a common family mechanism. *Future Virol*, 5(1), pp.65–83.

Ritzi, M. et al., 2003. Complex protein-DNA dynamics at the latent origin of DNA replication of Epstein-Barr virus. *Journal of cell science*, 116(Pt 19), pp.3971–3984.

Robertson, K.D., 2005. DNA methylation and human disease. *Nature reviews. Genetics*, 6(8), pp.597–610.

Roizmann, B. et al., 1992. The family Herpesviridae: an update. The Herpesvirus Study Group of the International Committee on Taxonomy of Viruses. *Archives of virology*, 123(3-4), pp.425–449.

Rous, P., 1911. A SARCOMA OF THE FOWL TRANSMISSIBLE BY AN AGENT SEPARABLE FROM THE TUMOR CELLS. *The Journal of experimental medicine*, 13(4), pp.397–411.

Rous, P., 1910. A TRANSMISSIBLE AVIAN NEOPLASM. (SARCOMA OF THE COMMON FOWL.). *The Journal of experimental medicine*, 12(5), pp.696–705.

Rowe, M. et al., 1987. Differences in B cell growth phenotype reflect novel patterns of Epstein-Barr virus latent gene expression in Burkitt's lymphoma cells. *The EMBO journal*, 6(9), pp.2743–2751.

Rowe, M. et al., 2007. Host shutoff during productive Epstein-Barr virus infection is mediated by BGLF5 and may contribute to immune evasion. *Proceedings of the National Academy of Sciences of the United States of America*, 104(9), pp.3366–3371.

Rowe, M., Fitzsimmons, L. & Bell, A.I., 2014. Epstein-Barr virus and Burkitt lymphoma. *Chinese journal of cancer*, 33(12), pp.609–619.

Rowe, M., Raithatha, S. & Shannon-Lowe, C., 2014. Counteracting effects of cellular Notch and Epstein-Barr virus EBNA2: implications for stromal effects on virus-host interactions. *Journal of virology*, 88(20), pp.12065–12076.

Rozen, R. et al., 2008. Virion-Wide Protein Interactions of Kaposi's Sarcoma-Associated Herpesvirus . *Journal of Virology*, 82(10), pp.4742–4750. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2346726/.

Ruvolo, V. et al., 1998. The Epstein-Barr virus nuclear protein SM is both a post-transcriptional inhibitor and activator of gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, 95(15), pp.8852–8857.

Sample, J. et al., 1990. Epstein-Barr virus types 1 and 2 differ in their EBNA-3A, EBNA-3B, and EBNA-3C genes. *Journal of virology*, 64(9), pp.4084–4092.

Schaadt, E. et al., 2005. Epstein-Barr virus latent membrane protein 2A mimics B-cell receptor-dependent virus reactivation. *The Journal of general virology*, 86(Pt 3), pp.551–559.

Schelcher, C. et al., 2007. Atypical bZIP domain of viral transcription factor contributes to stability of dimer formation and transcriptional function. *J Virol*, 81(13), pp.7149–7155. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17459922.

Schepers, A. et al., 1993. cis-acting elements in the lytic origin of DNA replication of Epstein-Barr virus. *J Virol*, 67(7), pp.4237–4245.

Schepers, A. et al., 2001. Human origin recognition complex binds to the region of the latent origin of DNA replication of Epstein-Barr virus. *The EMBO journal*, 20(16), pp.4588–4602.

Schepers, A., Pich, D. & Hammerschmidt, W., 1996. Activation of oriLyt, the lytic origin of DNA replication of Epstein-Barr virus, by BZLF1. *Virology*, 220(2), pp.367–376.

Schones, D.E. & Zhao, K., 2008. Genome-wide approaches to studying chromatin modifications. *Nature reviews. Genetics*, 9(3), pp.179–191.

Scott, D.L., Wolfe, F. & Huizinga, T.W.J., 2010. Rheumatoid arthritis. *Lancet (London, England)*, 376(9746), pp.1094–1108.

Selbach, M. & Mann, M., 2006. Protein interaction screening by quantitative immunoprecipitation combined with knockdown (QUICK). *Nature methods*, 3(12), pp.981–983.

Shaknovich, R. et al., 2006. Identification of rare Epstein-Barr virus infected memory B cells and plasma cells in non-monomorphic post-transplant lymphoproliferative disorders and the signature of viral signaling. *Haematologica*, 91(10), pp.1313–1320.

Shannon-Lowe, C. et al., 2009. Features distinguishing Epstein-Barr virus infections of epithelial cells and B cells: viral genome expression, genome maintenance, and genome amplification. *Journal of virology*, 83(15), pp.7749–7760.

Shannon-Lowe, C. & Rowe, M., 2011. Epstein-Barr virus infection of polarized epithelial cells via the basolateral surface by memory B cell-mediated transfer infection. *PLoS pathogens*, 7(5), p.e1001338.

Shannon-Lowe, C.D. et al., 2006. Resting B cells as a transfer vehicle for Epstein-Barr virus infection of epithelial cells. *Proceedings of the National*

Academy of Sciences of the United States of America, 103(18), pp.7065–7070.

Shaw, J.E., Levinger, L.F. & Carter, C.W.J., 1979. Nucleosomal structure of Epstein-Barr virus DNA in transformed cell lines. *Journal of virology*, 29(2), pp.657–665.

Sinclair, A.J., 2003. bZIP proteins of human gammaherpesviruses. *J Gen Virol*, 84(Pt 8), pp.1941–1949. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12867624.

Sinclair, A.J. et al., 1991. Pathways of activation of the Epstein-Barr virus productive cycle. *J Virol*, 65(5), pp.2237–2244. Available at: http://www.ncbi.nlm.nih.gov/pubmed/1850009.

Sinclair, A.J. & Farrell, P.J., 1995. Host cell requirements for efficient infection of quiescent primary B lymphocytes by Epstein-Barr virus. *J Virol*, 69(9), pp.5461–5468. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7543582.

Sista, N.D. et al., 1995. Physical and functional interaction of the Epstein-Barr virus BZLF1 transactivator with the retinoic acid receptors RAR alpha and RXR alpha. *Nucleic acids research*, 23(10), pp.1729–1736.

Sixbey, J.W. et al., 1989. Detection of a second widespread strain of Epstein-Barr virus. *Lancet*, 2(8666), pp.761–765.

Speck, S.H. & Ganem, D., 2010. Viral latency and its regulation: lessons from the gamma-herpesviruses. *Cell host & microbe*, 8(1), pp.100–115.

Strockbine, L.D. et al., 1998. The Epstein-Barr virus BARF1 gene encodes a novel, soluble colony-stimulating factor-1 receptor. *Journal of virology*, 72(5), pp.4015–4021.

Sugimoto, A. et al., 2011. Spatiotemporally different DNA repair systems participate in Epstein-Barr virus genome maturation. *J Virol*, 85(13), pp.6127–6135.

Swenson, J.J., Holley-Guthrie, E. & Kenney, S.C., 2001. Epstein-Barr virus immediate-early protein BRLF1 interacts with CBP, promoting enhanced BRLF1 transactivation. *Journal of virology*, 75(13), pp.6228–6234.

Szklarczyk, D. et al., 2015. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic acids research*, 43(Database issue), pp.D447–52.

Takada, K. et al., 1991. An Epstein-Barr virus-producer line Akata: establishment of the cell line and analysis of viral DNA. *Virus genes*, 5(2), pp.147–156.

Takada, K., 1984. Cross-linking of cell surface immunoglobulins induces Epstein-Barr virus in Burkitt lymphoma lines. *International journal of cancer. Journal international du cancer*, 33(1), pp.27–32.

Takada, K. et al., 1986. trans activation of the latent Epstein-Barr virus (EBV) genome after transfection of the EBV DNA fragment. *Journal of virology*, 57(3), pp.1016–1022.

Takagi, S., Takada, K. & Sairenji, T., 1991. Formation of intranuclear replication compartments of Epstein-Barr virus with redistribution of BZLF1 and BMRF1 gene products. *Virology*, 185(1), pp.309–315.

Tarakanova, V.L. et al., 2007. Gamma-herpesvirus kinase actively initiates a DNA damage response by inducing phosphorylation of H2AX to foster viral replication. *Cell host & microbe*, 1(4), pp.275–286.

Tempera, I. & Lieberman, P.M., 2010. Chromatin organization of gammaherpesvirus latent genomes. *Biochim Biophys Acta*, 1799(3-4), pp.236–245.

Thompson, M.P. & Kurzrock, R., 2004. Epstein-Barr virus and cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 10(3), pp.803–821.

Thorley-Lawson, D.A., 2001. Epstein-Barr virus: exploiting the immune system. *Nature reviews. Immunology*, 1(1), pp.75–82.

Tierney, R.J. et al., 2015. Unexpected patterns of Epstein-Barr virus transcription revealed by a high throughput PCR array for absolute quantification of viral mRNA. *Virology*, 474, pp.117–130.

Tovey, M.G., Lenoir, G. & Begon-Lours, J., 1978. Activation of latent Epstein-Barr virus by antibody to human IgM. *Nature*, 276(5685), pp.270–272.

Tsurumi, T., 1991. Characterization of 3’-to 5'-exonuclease activity associated with Epstein-Barr virus DNA polymerase. *Virology*, 182(1), pp.376–381.

Tsurumi, T. et al., 1998. Overexpression, purification and helix-destabilizing properties of Epstein-Barr virus ssDNA-binding protein. *J Gen Virol*, 79 (Pt 5), pp.1257–1264.

Tugizov, S.M., Berline, J.W. & Palefsky, J.M., 2003. Epstein-Barr virus infection of polarized tongue and nasopharyngeal epithelial cells. *Nature medicine*, 9(3), pp.307–314.

Tzellos, S. & Farrell, P.J., 2012. Epstein-barr virus sequence variation-biology and disease. *Pathogens (Basel, Switzerland)*, 1(2), pp.156–174.

Urier, G. et al., 1989. The Epstein-Barr virus early protein EB1 activates transcription from different responsive elements including AP-1 binding sites. *The EMBO journal*, 8(5), pp.1447–1453.

Vockerodt, M. et al., 2013. Suppression of the LMP2A target gene, EGR-1, protects Hodgkin's lymphoma cells from entry to the EBV lytic cycle. *J Pathol*, 230(4), pp.399–409.

Vockerodt, M. et al., 2015. The Epstein-Barr virus and the pathogenesis of lymphoma. *The Journal of pathology*, 235(2), pp.312–322.

Vrzalikova, K. et al., 2011. Down-regulation of BLIMP1alpha by the EBV oncogene, LMP-1, disrupts the plasma cell differentiation program and prevents viral replication in B cells: implications for the pathogenesis of EBV-associated B-cell lymphomas. *Blood*, 117(22), pp.5907–5917.

Walling, D.M., Flaitz, C.M. & Nichols, C.M., 2003. Epstein-Barr Virus Replication in Oral Hairy Leukoplakia: Response, Persistence, and Resistance to Treatment with Valacyclovir. *Journal of Infectious Diseases* , 188 (6), pp.883–890. Available at: http://jid.oxfordjournals.org/content/188/6/883.abstract.

Wang, P. et al., 2005. A redox-sensitive cysteine in Zta is required for Epstein-Barr virus lytic cycle DNA replication. *Journal of virology*, 79(21), pp.13298–13309.

Wang, P. et al., 2009. Topoisomerase I and RecQL1 function in Epstein-Barr virus lytic reactivation. *J Virol*, 83(16), pp.8090–8098.

Wang, Y. et al., 2008. Kaposi's sarcoma-associated herpesvirus ori-Lyt-dependent DNA replication: involvement of host cellular factors. *J Virol*, 82(6), pp.2867–2882.

Weiss, L.M. et al., 1991. Epstein-Barr virus and Hodgkin's disease. A correlative in situ hybridization and polymerase chain reaction study. *The American journal of pathology*, 139(6), pp.1259–1265.

Whitley RJ, 1996. *Herpesviruses. Medical Microbiology. 4th edition.*,

Wiedmer, A. et al., 2008. Epstein-Barr virus immediate-early protein Zta co-opts mitochondrial single-stranded DNA binding protein to promote viral and inhibit mitochondrial DNA replication. *Journal of virology*, 82(9), pp.4647–4655.

Woellmer, A. & Hammerschmidt, W., 2013. Epstein-Barr virus and host cell methylation: regulation of latency, replication and virus reactivation. *Curr Opin Virol*, 3(3), pp.260–265.

Wong, P. et al., 2008. An evolutionary and structural characterization of mammalian protein complex organization. *BMC genomics*, 9, p.629.

Wu, F.Y. et al., 2003. CCAAT/enhancer binding protein alpha interacts with ZTA and mediates ZTA-induced p21(CIP-1) accumulation and G(1) cell cycle arrest during the Epstein-Barr virus lytic cycle. *Journal of virology*, 77(2), pp.1481–1500.

Wucherpfennig, K.W. & Strominger, J.L., 1995. Molecular mimicry in T cell-mediated autoimmunity: viral peptides activate human T cell clones specific for myelin basic protein. *Cell*, 80(5), pp.695–705.

Wynne, J.W. et al., 2014. Proteomics informed by transcriptomics reveals Hendra virus sensitizes bat cells to TRAIL-mediated apoptosis. *Genome biology*, 15(11), p.532.

Xue, S.-A. & Griffin, B.E., 2007. Complexities associated with expression of Epstein–Barr virus (EBV) lytic origins of DNA replication. *Nucleic Acids Research*, 35(10), pp.3391–3406. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1904260/.

Yang, J. et al., 2015. Epstein-Barr virus BZLF1 protein impairs accumulation of host DNA damage proteins at damage sites in response to DNA damage. *Laboratory investigation; a journal of technical methods and pathology*, 95(8), pp.937–950.

Yates, J. et al., 1984. A cis-acting element from the Epstein-Barr viral genome that permits stable replication of recombinant plasmids in latently infected cells. *Proceedings of the National Academy of Sciences of the United States of America*, 81(12), pp.3806–3810.

Young, L.S. & Murray, P.G., 2003. Epstein-Barr virus and oncogenesis: from latent genes to tumours. *Oncogene*, 22(33), pp.5108–5121.

Young, L.S. & Rickinson, A.B., 2004. Epstein-Barr virus: 40 years on. *Nat Rev Cancer*, 4(10), pp.757–768.

Yu, M.C. et al., 1986. Cantonese-style salted fish as a cause of nasopharyngeal carcinoma: report of a case-control study in Hong Kong. *Cancer research*, 46(2), pp.956–961.

Zech, L. et al., 1976. Characteristic chromosomal abnormalities in biopsies and lymphoid-cell lines from patients with Burkitt and non-Burkitt lymphomas. *International journal of cancer. Journal international du cancer*, 17(1), pp.47–56.

Zeng, M.-S. et al., 2005. Genomic sequence analysis of Epstein-Barr virus strain GD1 from a nasopharyngeal carcinoma patient. *Journal of virology*, 79(24), pp.15323–15330.

Zeng, Y., 1985. Seroepidemiological studies on nasopharyngeal carcinoma in China. *Advances in cancer research*, 44, pp.121–138.

Zhang, C.X. et al., 1988. Altered expression of two Epstein-Barr virus early genes localized in BamHI-A in nonproducer Raji cells. *Journal of Virology*, 62(6), pp.1862–1869. Available at: http://www.scopus.com/inward/record.url?eid=2-s2.0-0023940404&partnerID=40&md5=03a802ff9d43454abcad1be10b1ba901.

Zhang, Q. et al., 1996. Functional and physical interactions between the Epstein-Barr virus (EBV) proteins BZLF1 and BMRF1: Effects on EBV transcription and lytic replication. *J Virol*, 70(8), pp.5131–5142.

Zhang, Q. et al., 1997. The Epstein-Barr virus (EBV) DNA polymerase accessory protein, BMRF1, activates the essential downstream component of the EBV oriLyt. *Virology*, 230(1), pp.22–34.

Zhang, Q., Gutsch, D. & Kenney, S., 1994. Functional and physical interaction between p53 and BZLF1: implications for Epstein-Barr virus latency. *Mol Cell Biol*, 14(3), pp.1929–1938.

Zhao, C. et al., 2010. Hexahistidine-tag-specific optical probes for analyses of proteins and their interactions. *Analytical biochemistry*, 399(2), pp.237–245.

Zhu, F.X. et al., 2005. Virion proteins of Kaposi's sarcoma-associated herpesvirus. *Journal of virology*, 79(2), pp.800–811.

Zhu, J.Y. et al., 2009. Identification of novel Epstein-Barr virus microRNA genes from nasopharyngeal carcinomas. *Journal of virology*, 83(7), pp.3333–3341.

Zimber, U. et al., 1986. Geographical prevalence of two types of Epstein-Barr virus. *Virology*, 154(1), pp.56–66.

Zuo, J. et al., 2008. The DNase of gammaherpesviruses impairs recognition by virus-specific CD8+ T cells through an additional host shutoff function. *Journal of virology*, 82(5), pp.2385–2393.

Zutter, M.M. et al., 1988. Epstein-Barr virus lymphoproliferation after bone marrow transplantation. *Blood*, 72(2), pp.520–529.