Investigation of the interaction between

Ebola virus and the host

Thesis submitted in accordance with the requirements of

the University of Liverpool for the degree of Doctor in

Philosophy

By

Isabel García-Dorival (MSc)

April 2016

AUTHOR'S DECLARATION

Apart from the help and advice acknowledged, this thesis

represents the unaided work of the author

…………………………..

Isabel García-Dorival

April 2016

This research was carried out in the Department of Infection

Biology, Institute of Infection and Global Health, University of

Liverpool

Acknowledgements

There are many people I would like to thank for all the support they gave me during the last three years in the course of my PhD. It has been a great pleasure and opportunity to do my PhD in the University of Liverpool, UK.

First and foremost, I would like to thank god for giving me this great opportunity; to my parents for all their support and love during these three years. Thank you for both moral and financial support and also for the great example you gave me throughout my life.

Big thanks to my supervisor Prof. Julian Hiscox for giving me the opportunity to work with him, it has been a pleasure to being part of his group. Also many thanks to all our collaborators in Public Health England, University of Bristol, University of Leeds and the Center Genomic Research in the University of Liverpool. Thanks to Prof.Milles Carroll, Dr. Roger Hewson, Dr. Stuart Dowall, Dr. David Mathews and Prof.John Barr for all your support and advice during my PhD. I would also like to thank all the lab members of the Respiratory and Emerging Viruses group for their friendship and constant support over the last three years.

Particularly thanks to Dr. Weining Wu, Dr. Olivier Touzelet, Dr. Stuart Armstrong, Miss Catherine Hartley, Dr. Simon Clegg, Miss Jill Hudson and Dr. Diane Munday for all their advices and support. A heartfelt thank you to all my friends in the IC2 for their friendship and scientific support over all this time;

Abstract

Ebola viruses cause viral haemorrhagic fever in humans and non human primates. Due to the severity of their symptoms, the lack of an approved treatment or vaccine, and the mechanism of transmission, these viruses have been classified as containment level 4 (CL4) pathogens. In December 2013, an Ebola virus outbreak initiated in West Africa, which lasted more than two years, and has been considered the worst outbreak in the history of the Ebola viruses. More than twenty-seven thousand people had a positive diagnosis of infection with more than eleven thousand confirmed dead. This outbreak highlights not only the urgent necessity for new treatments and vaccines against this virus but also that there is still a large knowledge gap for Ebola viruses compared to other pathogens. This thesis focuses on using high resolution approaches to characterise the interaction between the *Zaire ebolavirus* species (EBOV) and the host. The work focuses on using high resolution sequencing to characterise the evolution of EBOV throughout the outbreak in West Africa and also in adapting to grow in a novel host. The work also focuses on two viral proteins that are critical to the virus life cycle. The VP24 protein that is an interferon antagonist and the nucleoprotein (NP) which encapsidates the viral genome. Quantitative proteomics was used to identify host cell proteins that interacted with these viral proteins and inhibitors were used to ablate cellular protein function and monitor the effect on viral biology. The thesis is written in the style of 'by publication' and the publications associated with this work are reproduced in the Appendix.

Tables of contents

List of Figures

List of Tables

List of abbreviations

ATP1A1………………………………………………..ATPase subunit alpha-1

BDBV…………………………………………………...Bundibugyo virus

BHK……………………………………………………..Baby hamster cells

CGR………………………………………………Centre Genomic Research

CID…………………………………………………Collision-induced dissociation

CL4………………………………………………………..containment level 4

Ct………………………………………………………....Cycle threshold

DCR…………………………………………….Democratic Republic of Congo

DNAJ1…………………..DnaJ heat shock protein family (HSP40) member A1

DMEM……………………………………..Dulbecco's modified Eagle's medium

DTT……………………………………………………..Dithiothreitol

EBOV……………………………………………………………Ebola virus

ECACC……………………………………………European collection of cell cultures

EGFP…………………………………………..Enhanced green fluorescent protein

EMLab…………………………………………….European Mobile Laboratory

FBS……………………………………………………..Fetal bovine serum

FCS………………………………………………………..Foetal calf serum

FDR………………………………………………………....False discovery rate

GFP………………………………………………...Green fluorescent protein

GP………………………………………………………………Glycoprotein

GP₁,₂…………………………………………………………..Full length glycoprotein

HDK 293T…………………………………..Human embryonic kidney 293 cells

HRP……………………………………………………..Horseradish peroxidase

HSP70…………………………………………………Heat shock protein 70

HSP90…………………………………………………Heat shock protein 90

ICTV…………………………International Committee on Taxonomy of Viruses

IFN……………………………………………………….Interferon

IGRs…………………………………………………..Intergenic regions

IRF…………………………………………………...Interferon regulatory factor

KPNA6………………………………………………...Karyopherin alpha-6

L…………………………………………………………….Polymerase

LFQ………………………………………………………..label free quantification

LLOV……………………………………………………..Llovio virus

LMNB1………………………………………………..Lamin B1

LTQ………………………………………………………linear trap quadruple

mAbs………………………………………………….Monoclonal antibodies

MARV……………………………………………………Marburg virus

MQ………………………………………………………..Max Quant

MRC-5………………………………………Human foetal lung fibroblast cells

MS………………………………………………………Mass spectrometry

MS/MS……………………………………………….Tandem mass spectrometry

NC…………………………………………………….......Nucleocapsid

NNS…………………………………………..Non-segmented, negative strand

NP…………………………………………………………….Nucleoprotein

NTRs……………………………………………….Non-translated regions

ORF…………………………………………………...Open reading frame

PCR…………………………………………………Polymerase chain reaction

PHE……………………………………………………Public Health England

CHAPTER 1: Introduction

1.1 Introduction to the family *Filoviridae* that contains Ebola viruses.

The International Committee on Taxonomy of Viruses (ICTV) has recommended a hierarchy of viral taxa as follows: order, family, (subfamily), genus and species. Five families compose the order *Mononegavirales*; the family *Filoviridae* is one of them (ICTV report 2009; Feldmann et al., 1993). The family *Filoviridae* is composed of three genera: *Cuevavirus, Marburgvirus* and *Ebolavirus* (ICTV report 2009; Bukreyev et al., 2014; Kuhn et al., 2010). Filoviruses are enveloped viruses that posses a non-segmented, negative-strand (NNS) RNA molecule as a genome (Feldmann et al., 2013).

The *Cuevavirus* genus contains a single species, *Llovio cuevavirus* (LLOV), which has been recently discovered from dead insectivorous bats in Cueva del Lloviu, Asturias, Spain (Negredo et al., 2011). The *Marbugvirus* genus contains only one species as well, *Marburg Marburgvirus* (MARV), which was discovered in 1967 in Marburg, Germany when laboratory workers were handling contaminated tissue samples from monkeys imported from Uganda; this was the first filovirus discovered (Slenczka and Klenk., 2007).

The *Ebolavirus* genus contains five species: *Bundibugyo ebolavirus* (BDBV), *Reston ebolavirus* (RESTV), *Sudan ebolavirus* (SUDV), *Taï Forest ebolavirus* (TAFV) and *Zaire ebolavirus* (EBOV) (*Table 1.1*) (Feldmann et al., 2013; Bukreyev et al., 2014). All of these viruses have different case fatality rates in humans. RESTV has been reported to be not pathogenic in humans, but

highly pathogenic in nonhuman primates, causing viral haemorrhagic fever symptoms in monkeys (Jahrling et al., 1990, Miranda et al., 1999; Feldmann et al., 2013). EBOV is the most pathogenic for humans with up to a 90% case fatality rate (Miranda et al., 1999; Bowen et al., 1977).

EBOV was first discovered in 1976 in Zaire, now the Democratic Republic of the Congo (DRC) (Bowen et al., 1977). SUDV was discovered in the same year in southern Sudan (Report WHO/International Study Team; 1978). BDBV was discovered in 2007 in Bundibugyo District in western Uganda (Towner et al., 2008) and TAFV was discovered in 1994 in Côte d'Ivore from a non-fatal human case (Le Guenno et al., 1995; Formenty et al., 1999). Finally, RESTV was discovered in 1989/1990 in the United States from a group of monkeys imported from the Philippines into United States (Jahrling et al., 1990).

Table 1. 1   *Classification of Family Filoviridae according to the virus Taxonomy of the ICTV released in 2014.*

FAMILY: ***FILOVIRIDAE***

	Genus: *Cuevavirus*
	Species: *Lloviu cuevavirus*
	Genus: *Marburgvirus*
	Species: *Marburg marburgvirus*
	Genus: *Ebolavirus*
	Species: *Bundibugyo ebolavirus*
	Species: *Reston ebolavirus*
	Species: *Sudan ebolavirus*
	Species: *Taï Forest ebolavirus*
	Species: *Zaire ebolavirus*

1.2   *Ebolavirus* **genus outbreaks**

EBOV and SUDV are the viruses that have been responsible for most of the *Ebolavirus* outbreaks in Africa in the last 40 years (Rougeron et al., 2015; Paessler et al., 2013; Feldmann et al., 2013). The first EBOV outbreak reported was in 1976 in Zaire, now DRC, with 318 cases and a fatality rate of 88% (Bowen et al., 1977; Burke et al., 1978); after that, several EBOV outbreaks have been reported *(Table 1.2)*. Some of the largest EBOV outbreaks, prior to the 2014-2015 West Africa crisis, were in 1995 in Kikwit, DRC with 315 cases and in Kasai Occidental province, DRC in 2007 with 264 cases (Muyembe-Tamfum et al.,1999; Muyembe-Tamfum et al.,2012).

SUDV, BDBV and TAFV has been also responsible for outbreaks in Africa; however, the mortality rate was lower compared to EBOV. One example is the SUDV outbreak in 1976 in Sudan with 284 cases and a case fatality rate of 53% (Report of a WHO/International Study Team; 1978). Another example is the SUDV outbreak in Uganda in 2000 with 425 cases and a case fatality rate of 53% (Lamunu et al., 2004). For BDBV and TAFV, only one outbreak has been reported so far; one for BDBV in Uganda in 2007 (Towner et al., 2008) and another for TAFV in Ivory Coast in 1994 (Le Guenno et al., 1995). However, the biggest EBOV outbreak so far recorded is the 2014-2015 EBOV outbreak in West Africa. This outbreak started in Guinea in December 2013 when, it is assumed, a two-year-old boy had contact with a bat (Baize et al., 2014) and from where the virus was spread throughout Guinea, Sierra Leone and Liberia. Up to submission of this thesis, there have been 28,603 cases

and 11,301 fatalities (WHO. Ebola virus disease, West Africa—update. 2016).

Although the outbreak has been brought under control, sporadic cases still

occur, threatening emergence of the disease in the region again.

Table 1.2 *Examples of Ebolavirus outbreaks in the past 40 years.*

Filovirus (species)	Year	Outbreak location	Place of origin	Human cases (% mortality)
EBOV	1976	Yambuku, DRC	DRC	318 (88)
	1977	Tandala, DRC	DRC	1 (100)
	1994	Ogooué-Invido province, Gabon	Gabon	52 (60)
	1995	Kikwit, DRC	DRC	315 (79)
	1996	Mayibout, Gabon	Gabon	37 (57)
	1996	Booue, Gabon	Gabon	60 (75)
	2001-2001	Ogooué-Invido province, Gabon; Cuvette region, DRC	Gabon?	124 (79)
	2002-2003	Cuvette region, DRC; Ogooué-Invido province, Gabon	DRC?	143 (90)
	2003	Mboma and Mbanza, RC	DRC	35 (83)
	2005	Etoumbi and Mbomo in Cuvette region, DRC	DRC	12 (75)
	2007	Kasai Occidental province, DRC	DRC	264 (71)
	2008-2009	Kasai Occidental province, DRC	DRC	32 (47)
	2013-2015	Guinea, Sierra Leone, Liberia (West Africa)	Guinea	28,603 (40)
SUDV	1976	Nzara, Maridi, Tembura, Juba, Sudan	Southern Sudan	284 (53)
	1979	Nzara, Maridi, Sudan	Southern Sudan	34 (65)
	2000-2001	Gulu District, Mbarrara, and Masindi, Uganda	Uganda	425 (53)
	2004	Yambio County, Sudan	Southern Sudan	17 (41)
	2011	Uganda (central)	Uganda	1 (100)
BDBV	2007	Bundibugyo district, Uganda	Uganda	149 (25)
TIFV	1994	Tai Forest, Ivory Coast, and Basel, Switzerland	Ivory Coast	1 (0)
	1995	Liberia	Liberia?	1 (0)
RESTV	1989	Reston, Virginia	Philippines	4 (0)
	1992	Siena, Italy	Philippines	0 (0)
	1996	Alice, Texas	Philippines	0 (0)
	2008	Philippines	Philippines	0 (0)

1.3 *Ebolavirus* virion structure, genome organization and proteins.

1.3.1 Virion structure

Ebolavirus virions are pleomorphic with a filamentous or U-shape form, with a diameter of approximately 80nm and up to 1400nm long (*Figure 1.1A*, *1.1C*). Virions contain a helical nucleocapsid (NC) of approximately 50nm in diameter, a matrix and an envelope (derived from the host-cell plasma membrane). The virion surface has projections, which are caused by the virus encoded glycoproteins (GPs) that are anchored to the membrane; the glycoproteins are 10nm long and spaced 10nm apart (Feldmann et al., 2013; Ascenzi et al., 2008; Fedlmann et al., 1993).

1.3.2 Genome structure and organization

The *Ebolavirus* genome consists of a NNS RNA molecule of approximately 19,000 bases in length, making this one of the largest virus amongst the *Mononegavirales*. This RNA molecule was found to constitute approximately 1% of the total virion mass (Regnery et al., 1980; Fedlmann et al., 1993; Fedlmann et al., 1992; Feldmann et al., 2013; Kuhn et al., 2008). The genome of the filoviruses is very similar in organization to paramyxoviruses and rhabdoviruses, which are other members of the *Mononegavirales*. The *Ebolavirus* genome consists of seven genes starting with the nucleoprotein

(NP), virion protein 35 (VP35), virion protein 40 (VP40), glycoprotein (GP), virion protein 30 (VP30), virion protein 24 (VP24) and polymerase (L) (Elliott et al., 1985; Sanchez et al., 1993) (Figure 1.1B).

All *Ebolavirus* genes encode for one protein with the exception of the GP gene which encodes for three proteins; the full length glycoprotein (GP$_{1,2}$), the secreted glycoprotein (sGP) and the soluble secreted glycoproteins (ssGP). In order to do this, RNA editing occurs during transcription of the GP gene (Mehedi et al., 2011; Sanchez et al., 1996). At the extreme 3' and 5' end of the genome, there are extragenic regions, called *leader* and *trailer* regions. These regions contain the promoters for the replication and transcription process (Neumann et al., 2009; Weik et al., 2005). Each gene in the *Ebolavirus* genome contains an open reading frame (ORF) flanked by nontranslated sequences ranging from 57 to 684 nucleotides (Mühlberger E., 2007). Among the mononegavirales, this characteristic, of having long nontranslated sequences, has been only seen in filoviruses and Nipah and Hendra viruses (Wang et al., 2001).

Each *Ebolavirus* gene is flanked by highly conserved transcription start and stop signals (Neumann et al., 2009; Mühlberger E., 2007). The transcription stop and start signals of NNS RNA viruses are separated by intergenic regions (IGRs). For *Ebolavirus* the IGRs differ in length being relatively short, 4 or 5 nucleotides, for the IGRs of NP/VP35, VP40/GP and VP24/L; or long for the IGR of VP30/VP24 consisting of 144 nucleotides (Sanchez et al., 1993; Mühlberger E., 2007; Neumann et al., 2009). The modification or deletion of the non-translated regions (NTRs), IGRs, or transcription stop and start signals

ultimately affects protein expression levels to various extents, through modulation of mRNA abundance (Neumann et al., 2009).

A.

- Glycoprotein(GP)
- Viral RNAss (-)
- Nucleoprotein (NP)
- Matrix protein (VP40)
- Polymerase (L)
- Viral protein 24 (VP24)
- Viral protein 30 (VP30)
- Viral protein 35 (VP35)

B.

(-) strand RNA genome

Cleavage by host furin

GP1   GP2

3'— OH — NP — VP35 — VP40 — sGP — VP30 — VP24 — L — 5'

Leader

GP

ssGP

mRNA editing

Trailer
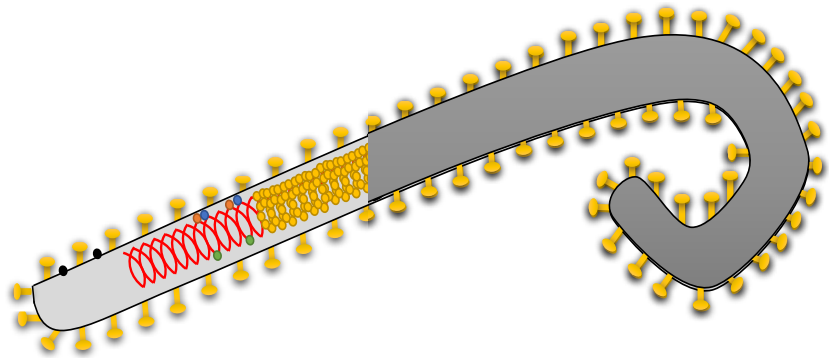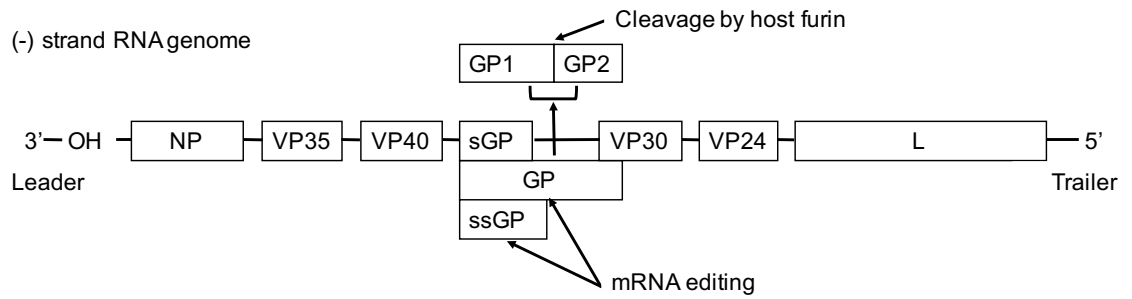
C.
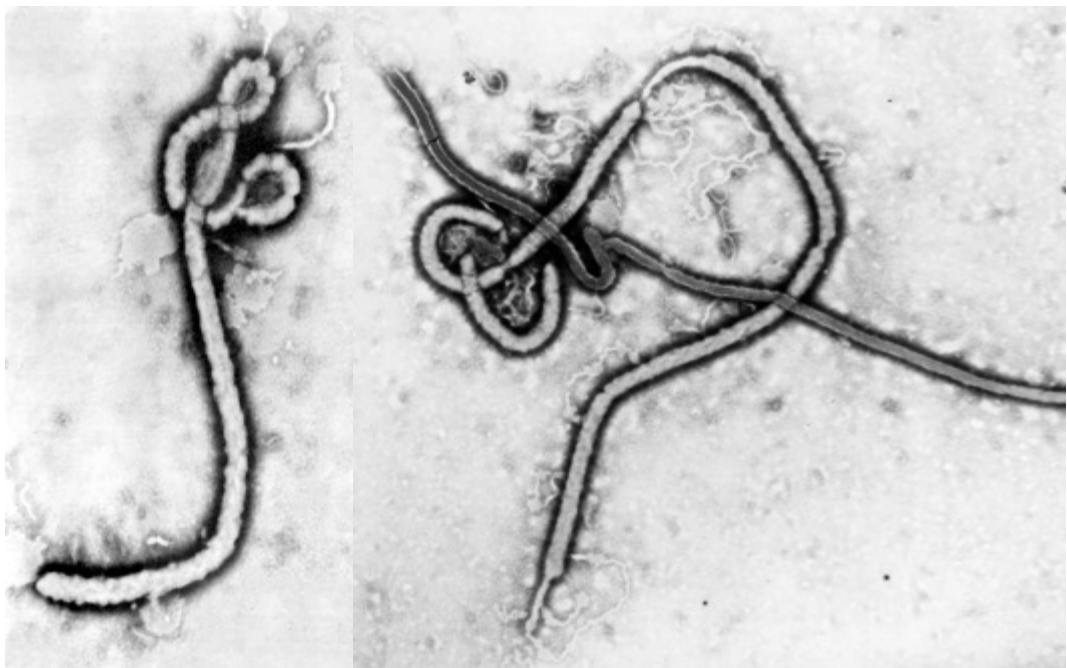
Figure. 1.1 Ebolaviruses and their genome organization. (A) *Schematic illustration of a filovirus particle. Four proteins are involved in the formation of the ribonucleoprotein complex: polymerase or large protein (L), nucleoprotein (NP), virion structural protein 30 (VP30), VP35. The glycoprotein (GP) is a type 1 trans membrane protein and is anchored with the carboxyl-terminal part in the virion membrane. The soluble GP (sGP) is a non-structural glycoprotein secreted from infected cells and is only secreted by Ebolaviruses. VP40 and VP24 are membrane-associated proteins.* ***(B)*** *schematic illustration of a filovirus genome* ***(C)*** *Negatively stained virions. Magnification: approximately x60,000 diagnostic specimen from the first passage in Vero cells of a specimen from a human patient—this image is from the first isolation and visualization of Ebola virus, 1976. Micrograph from F. A. Murphy, University of Texas Medical Branch, Galveston, Texas.*

1.3.3 Virus proteins

The *Ebolaviru*s genome encodes two non-structural proteins (sGP and ssGP) and seven structural proteins which are; NP, VP35, VP40, GP, VP30, VP24 and L (Table 1.3) (Elliott et al., 1985; Feldmann et al., 1993). While NP, VP30, VP35, VP40, GP and L are counterparts of proteins that have previously been found in others NNS RNA viruses; VP24 is unique for *Ebolavirus* (Kuhn et al., 2008; Mühlberger E., 2007; Biacchesi et al., 2003). *Ebolavirus* proteins are quite conserved among species with an amino acid identity that ranges between 55% and 81%. GP is the more variable protein with only 55% of identity (when comparing EBOV and RESTV) and VP24 is the more conserved with 81% of amino acid identity (when comparing EBOV and RESTV) (Ascenzi et al., 2008).

Ebolavirus proteins can be grouped in two categories, those that form the ribonucleoprotein (RNP) complex and those that are associated with the envelope. The proteins that form the RNP complex are involved in the replication and transcription process while the proteins associated with the envelope are more involved in the entry process or the assembly of the virus or modulation of the host response (Ascenzi et al., 2008; Kuhn et al., 2008; Feldmann et al., 2013).

Table.1.3.   *Filoviruses genes, gene products and their function (from Feldmann et al., 2013).*

Gene	Protein Function	MW (kd)[a]
Nucleoprotein (NP)	Major nucleoprotein; RNA encapsidation	90-104
VP35	Polymerase complex cofactor; interferon antagonist	35
VP40	Matrix protein; virion assembly and budding	35-40
Glycoprotein (GP)	GP[b] = Virus entry (surface peplomer); receptor binding and membrane fusion	150-170
	sGP[c] = Soluble glycoprotein; probably antibody decoy	50-55
	ssGP[c]= Small soluble glycoprotein; function unknown	50-55
VP30	Minor nucleoprotein; RNA binding and transcription activation	27-30
VP24	Minor matrix protein; virion assembly; interferon antagonist	24, 25
Polymerase (L)	RNA-dependent RNA Polymerase; enzymatic component of polymerase complex	~270

[a]Molecular weight (MW) is approximated and based on SDS-PAGE migration.
[b]Peplomer = trimer of the GP_1-GP_2 heterodimer; MW is for the heterodimer.
[c]Expressed only by EBOV; MW is for the monomeric form.

1.3.3.1   **Nucleoproteins**

NP and VP30 are the major and minor nucleoproteins. *Ebolavirus* NP, the largest NP among the mononegavirales, is encoded by the first (3') gene of *Ebolavirus* genome (Sanchez et al., 1992). NP is responsible for the formation of the NC and binds tightly to the *Ebolavirus* RNA (Ruigrok et al., 2011). This protein is 739 amino acids in length with a predicted molecular weight of 83.3kD.   Some of regions in the *Ebolavirus* NP gene are very conserved among the mononegavirales including paramyxoviruses and rhabdoviruses (Sanchez et al., 1992; Barr et al., 1991; Kuhn et al., 2008). The *Ebolavirus* NP is phosphorylated and only phosphorylated NP is incorporated to the virions; the phosphorylation of NP might be responsible for self-association (Kuhn et al., 2008; Watanabe et al., 2006).

The *Ebolavirus* VP30 is encoded by the fifth (3') gene from the *Ebolavirus* genome, this protein is 288 amino acids in length with a molecular weight of 32kDa (Kuhn et al., 2008). VP30 acts as a transcription activator factor playing an important role in transcription, however, this depends on the phosphorylation state of VP30. Phosphorylation of VP30 positively regulates its binding to NP but reduces its activity as a transcriptional activator (Modrof et all., 2002). VP30 functions as a transcriptional activator and is enhanced by a RNA secondary structure formed by nucleotides 54 to 80 in the leader region of the EBOV genome (Weik et al., 2002). VP30 binds directly to *Ebolavirus* RNA similar to NP suggesting that this interaction might play a role in the transcriptional regulation of mRNA synthesis *Ebolavirus* (John et al., 2007).

14

1.3.3.2   Polymerase and the polymerase cofactor

L is encoded by the last (3') gene of the *Ebolavirus* genome; and is translated into a 2,212 amino acid protein with a calculated molecular mass of 252.7 kDa (Feldmann et al., 2013; Elliott et al., 1985). *Ebolavirus* gene transcription is catalysed by L; furthermore, a comparison between the *Ebolavirus* L protein with other L proteins from other mononegavirales showed a close relationship amongst them (Volckov et al., 1999; Poch et al., 1990). L contains six conserved regions and three conserved RNA-depended RNA-polymerase sequence motifs: an RNA binding element, an RNA template recognition element and a nucleotide triphosphate-binding element. The arrangement of these conserved elements and the order of these three regions are similar among the mononegavirales (Poch et al., 1990; Kuhn et al., 2008).

Viral RNA replication and transcription take place in the cytoplasm and is carry-out by viral proteins L, NP, VP30 and VP35, which form the ribonucleoprotein complex (RNP). However, viral RNA replication and transcription can be affected by some cellular factors. It has been shown that inhibition of some cellular proteins can affect polymerase function and therefore affect viral replication or transcription (Lai., 1998). An example of this is the finding that the inhibition of cellular chaperones (HSP90 and HSP70) affected the stability of L and the synthesis of viral RNA for respiratory syncytial virus (RSV) (Munday et al., 2015). Another example is the effect of cellular chaperones (HSP70) and co-chaperones (DNAJ1 and HSP40) in the synthesis of influenza

virus RNA and the maintenance of the RNP (Cao et al., 2014; Manzoor et al., 2014).

The *Ebolavirus* VP35 is encoded by the second (3') gene of the *Ebolavirus* genome and is translated to produce a 321 amino acid protein of 35kDa. VP35 has two primary functions, acting as a polymerase cofactor for L and also as a type I interferon (IFN) antagonist (Feldmann et al., 2013; Basler et al., 2000). VP35 affects the mode of the RNA synthesis (transcription and replication) similar to the P proteins of others NNS RNA viruses, and is functionally linked to L and NP proteins (Mülhberger et al., 1999; Trunschke et al., 2013; Leugng et al., 2015; Kirchdoerfer et al., 2015). VP35 also plays an important role in the evasion of the immune response for Ebola viruses by acting as a type I IFN antagonist by blocking the RIG-I-like receptor (RLR) pathways. The protein also inhibits the the activation of interferon regulatory factor 3 (IRF-3) (Basler et al., 2000; Basler et al., 2003) and the production of IFN beta induced by RIG-I signalling, due to ability to bind to double strand RNA (Cardenas et al., 2006).

1.3.3.3   Matrix proteins

VP40 and VP24 are matrix proteins; VP40 is encoded by the third (3') gene of the *Ebolavirus* genome and is translated into a 326 amino acid protein. VP40 functions as the major matrix protein and it is the most abundant protein in the virion (Feldmann et al., 2013; Kuhn et al., 2008). VP40 is also capable of inducing virus like particles and binds to membranes (Jasenosky et al., 2001;

Noda et al., 2002); this protein is also involved in virus assembly and budding (Timmins et al., 2001).

VP24 is encoded by the sixth (3') gene in the *Ebolavirus* genome and is translated into a protein of 241 amino acids. VP24 is considered a minor matrix protein and little VP24 is incorporated to the virion compared to its relative abundance in the cell (Feldmann et al., 2013). VP24 acts as an interferon type I antagonist via interaction with karyopherin alpha-6 (KPNA 6) and therefore prevents the phosphorylation of the signal transducer and activator transcription factor 1 (STAT1) (Mateo et al., 2010; Xu et al., 2014). Furthermore, VP24 has several potential functions in the biology of Ebolavirus; including a role in nucleocapsid formation. Both VP40 and VP24 are also involved in the replication and transcription processes (Hoenen et al., 2010).

1.3.3.4   Structural and non-structural glycoproteins

The GP gene of the *Ebolavirus* encodes for three different glycoproteins: $GP_{1,2}$, sGP and ssGP (Sanchez et al., 1992; Mehedi et al., 2011). This happen due to a transcriptional RNA editing phenomenon, where several GP-specific mRNAs are synthesized (Sanchez et al., 1996). The primary product of the GP gene is the sGP, therefore no transcriptional editing needs to be done. In order to express $GP_{1,2}$ transcriptional RNA editing occurs at a series of seven uridines within the genomic RNA, resulting in an additional adenosine (A) residue in the transcript (Mehedi et al., 2011). This A addition results in an extended ORF which encodes the $GP_{1,2}$ (Sanchez et al., 1996). The addition

of another A during the transcriptional editing of the GP gene leads to the expression of the ssGP (Mehedi et al., 2011). In the GP gene, only about 20% of transcripts are edited while the remaining 80% are not, therefore there is more production of sGP than $GP_{1,2}$ in infected cells.

GP undergoes post-translational modifications. The protein is cleaved by furin into two subunits; GP1 and GP2, the latter of which contains the transmembrane domain. Mature GP ($GP_{1,2}$), exists as a trimer in the viral particle; in this mature $GP_{1,2}$, the GP1 and GP2 are connected by a disulphide bridge (Volchkov et al., 1998; Feldmann et al., 2013). $GP_{1,2}$ is one of the most studied proteins of the *Ebolavirus* virion due to the critical role of this protein in virus entry and pathogenesis (Martinez et al., 2013; Feldman et al., 2013). The sGP has been linked to an "antigenic subversion" function, in which it is proposed sGP induces a host antibody response that focuses on the epitopes it shares with $GP_{1,2}$ and therefore allowing it to compete for anti-$GP_{1,2}$ antibodies (Mohan et al., 2012). The function of ssGP is still unknown.

1.4   Replication cycle

1.4.1 Binding and entry

Filoviruses can infect different cell types; these viruses enter into cells by macropinocytosis and this is mediated by the interaction of $GP_{1,2}$ with host-cell surface molecules followed by uptake of the virus into endosomes (Feldmann et al., 2013; Saeed et al., 2013; Lee et al., 2009; Chandran et al., 2005). Once

$GP_{1,2}$ is processed by endosomal cysteine proteases, it interacts with the cholesterol transporter, Niemann-Pick C1 (NPC1), a protein that mediates cholesterol transport (Carette et al., 2012; Côté et al., 2011). Finally, the fusion between viral and endosome membranes occurs and this results in the release of the viral RNP into the cytoplasm (Miller et al., 2012).

Several cell surface proteins have been identified that interact with $GP_{1,2}$ and therefore facilitate the entry of the virus into cells. Some examples of these molecules are lectins (Alvarez et al., 2002; Takada et al., 2004), the tyrosine kinases receptors Tyro3/Axl/Mer (TAM) (Shimojima et al., 2006), the T cell immunoglobulin and mucin domain (TIM) proteins (Kondratowicz et al., 2011) and integrins (Takada et al., 2000). Furthermore, several studies on $GP_{1,2}$ have confirmed the importance of this protein and its interactions for the viral entry (Martinez et al., 2013; Groseth et al., 2012). Some studies showed that a mutation in GP can restrict viral entry into a specific cell type (Martinez et al., 2013). However, even when $GP_{1,2}$ contributes to viral entry, is not sufficient for virulence in vivo (Groseth et al., 2012).

1.4.2 Transcription and Replication

After release of the RNP, transcription of viral mRNAs and replication of the genome occurs. For the *Ebolavirus,* as well as for the other members of the order *Mononegavirales*, these processes take place in the cytoplasm of the infected cells. Due to the genetic material Mononegavirales posses (NNS RNA) these viruses carry their own RNA-dependant RNA polymerase in the

virus particle that is associated with the RNP, upon release, RNA synthesis events then occur (Feldmann et al., 2013; Brauburger et al., 2012; Whelan et al., 2004). During transcription, new polyadenylated monocistronic messenger RNAs (mRNAs) are synthesized in a 3' to 5' order (Mühlberger et al., 2007; Brauburger et al., 2012; Feldmann et al., 2013). Because of the release of the polymerase complex from the RNA template, reinitiation at downstream located genes are attenuated. Therefore, the NP mRNA is the most abundant mRNA and the L mRNA is the least abundant (Mühlberger et al.,2007). This is a feature common to the Mononegavirales (Whelan et al., 2004).

Filovirus genes have highly conserved transcription start and stop signals, genes that do not overlap and are separated by IGRs (Sanchez et al., 1993; Mühlberger E., 2007; Neumann et al., 2009). These IGRs have been shown to affect filovirus transcription (Neumann et al., 2009). Filovirus mRNAs are capped and polyadenylated (Mühlberger et al., 1996) and can be detected in cells as early as 6 hours post infection (hpi) (Nambo et al., 2013). The proteins necessary for the transcription of filovirus may vary according to the genera. For *Marburgvirus* is been shown that only NP, VP35 and L are necessary for the transcription while for *Ebolavirus* VP30 is also needed. *Ebolavirus* VP30 acts as a transcription activator (Mühlberger et al., 1999).

Replication of the filovirus genome starts with the synthesis of a complementary antigenomic RNA. As in the transcription, this is also done by L; the antigenome serves as a template for the generation of the new filoviral genome. Due to the high degree of sequence complementarity at the ends of

the genome, stem-loop structures are formed at the 3' and 5' ends of the filoviral genome and antigenome. These structures are believed to be essential for filovirus replication (Crary et al., 2002). The replication promoter sequence in the *Ebolavirus* genome is located at the 3' termini of the genome and is bipartite (Weik et al., 2005). The first part is located in the leader region (nucleotides 1 to 55) and the second part is located in the NTR of the NP gene (nucleotides 81 to 128) (Mühlberger et al., 2007). The sequence between these two promoters is critical for replication and only insertions or deletions of six or multiples of six nucleotides can be tolerated. This indicates that the replication promoter obeys the rule of six even when the length of the total genome does not (Weik et al., 2005; Mühlberger et al., 2007).

The initial expression of NP is thought to trigger a switch from transcription to replication which occurs due to the encapsidation of the antigenomic RNA. This encapsidated antigenomic RNA serves as a template for the synthesis of the filoviral genome which is also rapidly encapsidated. A reduction in the amount of the capsid proteins is believed to cause a return to transcription and *vice versa* (Feldmann et al., 2013). Furthermore, filoviral induced inclusion bodies, have been recognized as the site of the virus genome replication (Hoenen et al., 2012; Nambo et al., 2012) (Figure 1.2).

Different methodologies have been developed in order to study the viral replication/transcription of these viruses. An example of these are the mini-genome systems which allow the study of viral replication/transcription without using a high containment facility which is usually required for filoviruses

(Mülherber et al., 1998; Mülherber et al., 1999). Mini genome systems provide a way to identify and test new compounds and small molecule inhibitors that affect the viral replication/ transcription of the virus and therefore can be used as a possible therapy (Ueblhoer et al., 2014; Hoenen et al., 2014; Jasenosky et al., 2010), as well as the delineation of viral and cellular factors involved in viral RNA synthesis. Further refinements have developed virus like particles that contain mini-genomes (Hoenen et al., 2011; Edwards et al., 2015), and thus recapitulate many stages of the virus life cycle.
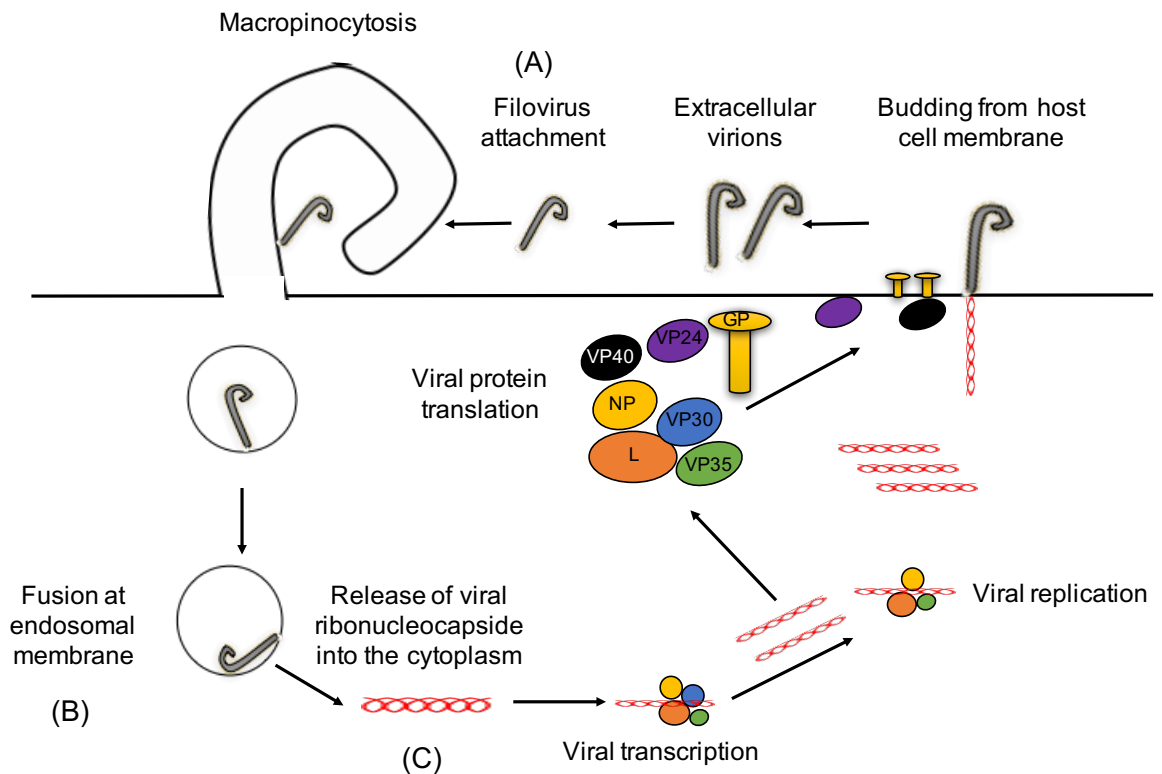
Figure 1.2. Different stages in the Filovirus replication cycle. *(A) First, the filovirus attaches to the cell membrane by its surface GP. The virus is then taken up by macroponicytosis. (B) Then GP and the host protein NPC1 interact, which is a prerequisite for the fusion of viral and endosomal membranes. GP mediates fusion of the viral and the endosomal membrane, (C) releasing the viral ribonucleocapsid into the cytoplasm, where the NNS RNA genome undergoes transcription and replication. Production of 5'-capped, 3'-polyadenylated mRNAs from individual viral genes occurs, and genome replication follows. Translation of viral proteins occurs, and new viral particles are assembled at the plasma membrane. VP40 functions as the viral matrix protein and directs budding of particles from the cell surface. GP, a type I transmembrane protein, is incorporated into the budding particles, as are viral nucleocapsids containing the viral genome **(from Messaoudi, et al., 2015)**.*

1.4.3 Assembly and Release

Once there is a sufficient amount of viral proteins, new RNA genomes are encapsidated and virion assembly initiates. The assembly and release of viral particles involves the transport of several viral proteins to the plasma membrane. Once the inclusion bodies, the sites of replication of the virus, become larger and closer to the nucleus, VP40 interacts with the RNP in order to facilitate trafficking to the plasma membrane (Nambo et al., 2012; Jasenosky and Kawaoka, 2004). VP40 traffics to the plasma membrane by interaction with different cellular factors. Some examples of these are the ubiquitin ligase NEDD4; the TGS101, which is a component of the vesicular protein sorting machinery and the clathrin-coated vesicle associated AP-2 protein complex (Jasenosky and Kawaoka., 2004). Membrane/lipid rafts have been identified as platforms for virion assembly (Feldmann et al., 2013). Once $GP_{1,2}$ is produced and transported to the plasma membrane this may interact with VP40 and assemble virus particles (Jasenosky and Kawaoka.,2004)

1.5 Pathogenesis and host response

Filoviruses cause severe viral haemorrhagic fevers in humans and non-human primates. The incubation period for these viruses can vary from 2 to 21 days, and then, infected individuals start to develop symptoms. These symptoms are non-specific and similar to influenza virus (flu-like symptoms) and usually start with fever, headache, muscle pain, and diarrhoea (Zampieri et al., 2007). These symptoms can get worse according to the development of the disease

and can include gastrointestinal problems, such a vomiting and also neurological symptoms as confusion, delirium and seizure. Haemorrhagic symptoms do not appear in all patients, and if they do, these occur at the peak of the infection. These haemorrhagic symptoms can finally end in multi organ failure and death (Zampieri et al., 2007; Rougeron et al., 2015).

In order to study the pathogenesis of filoviruses in controlled conditions (e.g. the absence of co-infection and defined virus and known times of infection) several animal models have been used. Non-human primates have been the model of choice for many years and have been used to map the immune response and evaluate potential vaccines. Some examples of non-human primates used for these studies include macaques, African green monkeys, rhesus monkey and cynomolgus monkeys (Geisbert et al., 2003a, Geisbert et al., 2000). Rodent models have also been used to study filovirus infection; particularly therapeutic and virus adaptation studies (Volchkov et al., 2000; Volchkova et al., 2011), and they have the advantage of being cheaper and easier to use than non-human primates and can provide large experimental group sizes if needed. Some examples of rodent models include the guinea pig (Connolly et al., 1999; Volchkov et al., 2000; Volchkova et al., 2011; Mateo et al., 2011; Wong et al., 2014), mice (Ebihara et al., 2006) and the Syrian hamster (Ebihara et al., 2012).

Filoviruses can infect different cell types including cells of the immune system. Once a person gets infected with the virus, via direct contact with fluids from infected humans or animals, the virus can infect dendritic cells and monocytes

/macrophages. These cells are the primary target for virus replication and responsible for dissemination of the virus in the body enabling the virus to migrate to different cell types/organs and therefore causing a systemic infection (Geisbert et al., 2003a; Bray et al., 2005; Rougeron et al., 2015). One of the characteristics of the filoviral pathogenesis is the uncontrolled replication of the virus in the body. This has an enormous effect on the body, not only for cytopathic effect but also on the host immune response (Figure 1.3) (Zampieri et al., 2007).

Filovirus can cause an impairment/dysregulation of the immune response due to the infection of the dendritic cells and monocytes /macrophages. These infected cells initiate inflammation and coagulation by releasing cytokines and chemokines (IFN-α, INF-β, IL-6, IL-18, TNFα among others), recruiting more cells to the site of infection (Geisbert et al., 2003a, Bray et al., 2005; Zampieri et al., 2007). Due to the large amount of cytokines/chemokines release, apoptosis and cell lysis occurs; which causes vascular damage and coagulation abnormalities (Geisbert et al., 2000; Geisbert et al., 2003b; Gupta et el., 2007). As discussed earlier Filoviruses have also the capacity to counteract the host immune response through specific encoded viral proteins. EBOV VP35 (Basler et al., 2003; Basler et al., 2000; Edwards et al., 2016) and EBOV VP24 (Mateo et al., 2010) act as interferon antagonists and soluble GP acts as an antibody decoy.

Contact with skin lesions/abrasion, mucosal surface

Dendritic cell

Macrophage/Monocyte

Immune response

Tissue damages

Liver

- Multifocal necrosis/infection
- Dysfunction and decreased production of clotting factors
- Hepatic failure

Spleen

- Defective adaptive immunity (humoral and cellular)

Adrenal

- Necrosis of adrenal cortical cells
- Impaired synthesis of steroids

Inflammatory cytokines and chemokines

- Increased endothelial permeability
- Vascular leakage
- Coagulopathy
- Bystander tissue damage

Inhibition of INF production (VP35, VP24)

- Lack of viral replication control

Limited panel of cytokines

- Defective humoral response

Interactions with affected APC. Soluble factors release?

- T lymphocytes and natural killer cells apoptosis

Figure 1.3. Pathogenesis model based on findings with EBOV. Viruses enter the body through lymphatic and/or blood vessels. Dendritic cells and macrophages are the first cells to be infected by filoviruses. The virus spreads from the initial site of infection to secondary lymphoid organs and liver where intense replication takes place in other cells of the host organism (hepatocytes, endothelial cells, fibroblasts and epithelial cells). The extensive infection of antigen-presenting cells (APC) leads to altered inflammatory responses and uncontrolled released of mediators. This contributes to their pathogenesis by attracting inflammatory responses followed by further and uncontrolled released of mediators (from Rougeron et al., 2015).

1.6 Treatment and vaccines

The latest EBOV outbreak in West Africa has been the biggest and deadliest outbreak since EBOV was discovered in 1976. This recent outbreak highlights the importance on EBOV research and the urgent necessity for novel therapeutic strategies. Due to the lack of an effective treatment, clinical trials and studies for new possible treatments and vaccines have been performed in the current outbreak to assess their ability to control and reduce mortality rates (Mendoza et al., 2016).

Examples includes antibodies from convalescent patients to treat EBOV infections (Mupapa et al., 1999; Maruyama et al., 1999). This method has been used recently in the latest EBOV outbreak in West Africa to treat patients. Transfusion of blood or plasma, from patients who survived the disease, was administrated to patients in order to help them to counteract the disease. Although this treatment alone was not very successful, a combined treatment of this with a monoclonal antibody (mAbs) cocktail against $GP_{1,2}$ showed promising results (Zeitlin et al., 2016). One of the concerns with the antibody treatment for *Ebolavirus* infections has been the cross reactions of these antibodies among the Ebolaviruses and if these antibodies can confer protection against different species of *Ebolavirus.* Some of the mAbs characterized from patients from the BDBV outbreak in Uganda in 2007, have neutralizing activity against EBOV (Flyak et al., 2016).

Cocktails of mAbs that target EBOV GP have been used in the recent EBOV outbreak. An example of this is the ZMAb cocktail, which is a mixture of three different antibodies against different regions of EBOV GP: the GP base and glycan cap (Davison et al., 2015; Audet et al., 2014; Qiu et al., 2013). ZMAb has been shown to be effective in nonhuman primates (Qiu et al., 2013) and in patients during the EBOV outbreak in West Africa, especially when it is combined with plasma treatment (Mendoza et al., 2016).

Different types of vaccines have been evaluated, and have been investigated in animal models (Hart et al., 2003; Marzi et al., 2014). However, due to the impact of the recent EBOV outbreak, some vaccines have been approved for trials in humans. An example of this is the recombinant replication-competent vesicular stomatitis virus-based vaccine expressing a surface EBOV GP (rVSV-EBOV), which has been evaluated in Guinea (Henao-Restrepo et al., 2015). A small trial of the vaccine looking for secondary and adverse effects was conducted with approximately 50 patients and no safety concerns were reported (Regules et al., 2015). After that, a bigger trial, with more than seven thousand participants was conducted in Guinea. This trial concluded that the vaccine might be highly efficient and safe in preventing EBOV disease (Henao-Restrepo et al., 2015). Small-molecule drugs have been also tried against EBOV infection *in vitro* and *in vivo*. An example of this is favipiravir, a pyrazine carboxmide derivate which is used against influenza infection; brincidobir, a nucleotide analog used against cytomegalovirus; and amiodarone, an antiarrhythmic drug, which is approved for human use (Mendoza et al., 2016).

1.7   Aims and objectives

My PhD aims were to investigate the interaction of EBOV with host cells and the potential evolution of the virus *in vivo*. This work comprised three parts: The first part focused on the characterization of the EBOV evolution using high resolution sequencing in adapting to grow in a novel host (guinea pig). During the course of this research, the West African outbreak occurred, and I deployed to the European Mobile Laboratory in Guinea to aid in the diagnostic effort and also to process samples for laboratory analysis back in the UK and other consortia partners. The work on the guinea pig model allowed us to analyse potential changes in EBOV evolution during this and future outbreaks. This work also identified amino acid changes that occurred during virus adaptation to a new host and during the EBOV outbreak in West Africa. This part of the work is presented in Chapter 3. These results contributed to and have been published in two primary publications; one in Nature and the other one in Genome Biology (Carroll et al., 2015 and Dowall et al., 2014). This work also helped validate the MinIon sequencing work that was published in Nature and I am one of the co-authors on this paper (Quick et al. 2016).

The second part of this work focused on the EBOV VP24. It aimed to define the cellular interactome of EBOV VP24 using quantitative proteomics. Several cellular proteins were identified as possible interaction partners for EBOV VP24; however, ATP1A1 was one of the most significant hits. A novel potential inhibitor against ATP1A1, ouabain, was presented in this part of the work that had anti-viral activity against EBOV. This part of the work is presented in

Chapter 4 and has been published in the Journal of Proteome Research (Garcia et al., 2014).

The third part of this thesis focuses on EBOV NP and RESTV NP and it aimed to define the cellular interactome of these two viral proteins using quantitative proteomics. The role of molecular chaperones on the EBOV replication/transcription identified and investigated in this study by developing mini genome system for EBOV-Makona. The results based on this work are presented in Chapter 5 and have been just accepted for publication in the Journal of Proteome Research (Garcia et al., 2016).

CHAPTER 2: Materials and Methods

All virus preparations, animal experiments and virus infections were done by our collaborators in Public Health England (PHE), Porton Down (UK) in their CL4 facilities, where I assisted in the downstream processing. It was not possible to undertake the CL4 training in the time period of the project. The RNA sequence and bioinformatics analysis (sequencing and sequence assembly) was done by the team of Centre Genomic Research (CGR) at University of Liverpool and Dr David Mathews in University of Bristol, and I also assisted in this and processed the downstream data. The rest of the analysis and experiments have been done by myself in University of Liverpool.

2.1 Virus preparation

Due to the high containment level classification of EBOV, all virus preparation was done by our collaborators in PHE in their CL4 facilities at Porton Down. The EBOV Zaire ME718 strain that was used in this work was provided by PHE; this virus was isolated during an outbreak in October 1976 (Report WHO 1978) in Yambuku, Mongala Province (Democratic Republic of the Congo) (Bowen et al., 1977; Johnson et al., 1977; Pattyn et al., 1977). It was sourced from an acute phase serum sample obtained from a 42-year-old Belgian nun (Sister M.E.) who became ill on September 23, 1976, in Yambuku, Equateur Province Zaire. This strain is equivalent to the Mayinga strain, as Sister M.E. nursed Sister Mayinga (who the strain is named after) and most likely contracted her infection. Virus stocks used for this work were grown in VeroE6 cells European Collection of Cell Cultures, UK (ECACC) cultured in Leibovitz's L15 (L15) media containing 2% foetal calf serum (FCS), and aliquots were stored at -80°C. Virus titres were determined by 100-fold dilution with L15 media without any FCS added. A total of 100 µL of each dilution was overlaid onto semi-confluent cell monolayers in four replicate 12.5 cm^2 tissue culture flasks and left to absorb for 1 h. A volume of 5 mL of media was then added to the cells and those were incubated at 37°C for 7 to 8 days. Cytopathic effects were determined by microscopy, and the results from each dilution were used to calculate 50% tissue culture infective dose (TCID$_{50}$) using the Reed-Muench method (Reed et al., 1938). Infection experiments were planned jointly.

2.2   Animal studies

Due to the high containment level classification of EBOV, all the animal studies including the animal challenge, observations and monitoring, necropsy and tissue collection and RNA preparation was done by our collaborators in PHE in their CL4 facilities at Porton Down. Infection experiments were planned jointly.

2.2.1 Animals used in this study

For animal infection studies, female Dunkin-Hartley guinea pigs with weights of 250 g to 350 g (Harlan Laboratories, UK) were used. Before procedures involving the manipulation of animals, guinea pigs were anesthetized with 1.5% to 2% isoflurane in an induction change until full sedation was achieved. Animals infected with EBOV were housed within an isolator under climate-control conditions in an animal CL4 room; food and sterile water were available as they needed. There were 10 guinea pigs depending on the group for each passage of the virus and a control group. From a practical standpoint of working at CL4 this number also represented the maximum number of animals that could be processed at the time. Of these animals, four were killed at day 7 post infection for preparation of virus and six to eight were carried on and used to measure clinical parameters. The study was performed under a UK

Home Office Project License conforming to the Animal Procedures Act. Ethical

review was performed by the PHE Animal Welfare and Ethical Review Board.

2.2.2 Animal challenge

EBOV stock was diluted in sterile phosphate-buffered saline (PBS) solution to

prepare the relevant dose of virus in a 0.2 mL volume. For passaging

experiments (required for virus adaptation), the dose delivered was 10^4

$TCID_{50}$. Surplus inoculation was made to confirm concentration via back

titration in cell culture. Guinea pigs were sedated, and subcutaneously

inoculated with the virus suspension in the lower right quadrant of the back.

Then, the guinea pigs were returned to their cages and monitored for adverse

effects caused by the injection of the anaesthetic until the animals fully

recovered. Negative control groups were injected with the same volume of

PBS.

2.2.3 Observations and monitoring

Animals were monitored at least twice daily, and observations (swelling at

injection site, movement, breathing, food intake, water intake and appearance)

recorded for the duration of the study. A set of humane clinical end points were

defined (20% weight loss, or 10% weight loss and a clinical symptom) which

indicated that the animal would be euthanized to prevent any unnecessary suffering. Weights of the animals were taken daily, and temperatures recorded using a pre-inserted temperature chip.

2.2.4 Necropsy and tissue collection

Animals were humanely euthanized by intraperitoneal injection of 200 mg/mL pentobarbital sodium. Necropsies were performed within a flexible-film isolator in the animal CL4 facility. Spleens were removed from four out of the 10 guinea pigs at each passage and stored at -80°C. For processing, samples were transported to the *in vitro* CL4 laboratory and thawed at room temperature. Spleens were homogenised by vigorously passing through a 500 mm cell strainer (Corning, UK). The resultant suspension was clarified by centrifugation at 400 g for 10 minutes to remove cell debris. The supernatant was collected, aliquoted and stored at -80°C. One vial was used to assess virus concentration by TCID$_{50}$ assay.

2.2.5 RNA preparation of the animal challenge experiment

After the animal challenge, spleen from each passage was homogenate and then added to 560 μL of AVL buffer (Qiagen, UK) containing carrier RNA, which is poly-A tails that will help to the precipitation of the RNA and therefore

help to the RNA extraction. The mixture was then incubated for 10 minutes at room temperature and then 560μl of 96-100% ethanol was added. The samples were then removed from the CL4 laboratory for processing. RNA was then isolated using a viral RNA extraction kit (Qiagen; 52904) following the manufacturer's instructions (Qiagen, UK). Confirmation of EBOV-specific RNA extraction was achieved by use of a block-based PCR assay (at PHE) and a conventional RT-PCR *in house* (at University of Liverpool) using primers that detect the NP and L regions.

2.2.6 RT-PCR for confirmation of EBOV-specific RNA extraction

The reverse transcription (RT) step was done using a Platinium TaqDNA polymerase from the ThermoScript™RT-PCR System (Invitrogen: 11146-024) and gene-specific primers (forward primers for EBOV NP and EBOV L) following the manufacturer's instructions (Invitrogen; UK) (a list of primers used in this study are in the Appendix). A mixture of 500ng of RNA from the samples was mixed with 20μM of EBOV NP or EBOV L forward primers and 2μL of 10mM dNTP mix (Invitrogen).The mixture was then incubated at 65°C for 5 minutes. The mixture was then cooled on ice for few minutes and 8μl of the cDNA synthesis mix was added (4μL of 5X buffer, 1 μL of DTT (0.1M), 1μL of DEPC treated water, 1 μL of RNase OUT™ and 1μL of ThermoScript™RT). The mixture was incubated then at 55°C for 60 minutes and finally 1μL of

RNaseH was added to the mixture and incubated at 37°C for 20 minutes. The conventional PCR was done using a Paq5000 DNA polymerase (Aligent Technologies: 600682) following the manufacturer's instructions (Aligent Technologies, UK). Each PCR reaction was performed at conditions shown in Table 2.1. For the PCR 2μL of the cDNA (RT steep product) plus the following ingredients were used:

- 5μL of 10X Paq5000™ Reaction Buffer (Agilent)

- 4μL of 100mM dNTP mix (25mM each dNTP)

- 1μL of Forward primer (10μM)

    - EBOV-NP Forward: 5'-GATGATGGAAGCTACGGCGA-3'

    - EBOV-L Forward: 5'-TACGCAGGACCATCACCCTA-3'

- 1μL of Reverse primer (10μM)

    - EBOV-NP Reverse: 5'-TGGATTGTGTCTGTCTGCCC-3'

    - EBOV-L Reverse: 5'-GTGCGTGTTCCTTCCAGGTA-3

- 0.5μL of Paq5000 DNA Polymerase (5U/μL) (Agilent)

- 36.5μL of nuclease free water

Table 2.1 *PCR conditions used for the confirmation of EBOV-specific RNA extraction (University of Liverpool).*

N°. of Cycles	Temperature	Duration
1	95°C	2 minutes
30	95°C	20 sec
	55°C	20 sec
	72°C	1 minute
1	72°C	5 minutes

2.2.7 Agarose gel electrophoresis

Agarose gel electrophoresis was used to detect RT-PCR products. In order to do this 1% (W/V) of agarose molecular grade (BIOLINE: BIO-4126) was melted in 1X TBE buffer (89mM Tris base, 89mM Boric acid, 2mM EDTA, pH 8.3) by heating. After that, the necessary amount of SYBR Safe DNA Gel Stain *10,000X concentrate in DMSO (Invitrogen: S33102) was added to briefly cooled agarose to a final concentration 1X. DNA. Samples were made up with 5X DNA Loading Buffer (BIOLINE:) and 5μL of HyperLadder™I (BIOLINE:

BIO-33053) was loaded into the agarose gel and electrophoresis was run at 120V constant voltage for 40 minutes. Once the presence of the EBOV products was confirmed by RT-PCR, the RNA samples were pooled from individual guinea pigs for each passage for sequencing analysis.

2.2.8 Next generation sequencing analysis of the guinea pig samples

The following experiment was done by our collaborators in the CGR (University of Liverpool) and by Dr. David Mathews from Bristol University. In order to do the next generation sequencing analysis, the RNA samples were DNase treated with Turbo DNase (Ambion) using a rigorous protocol. RNA–Seq libraries were prepared from the resultant RNA using the Epicentre ScriptSeq v2 RNA-Seq Library Preparation Kit. RNA (50 ng) was used as input and following 12 cycles of amplification, libraries were purified using AMPure XP beads. Each library was quantified using Qubit and the size distribution assessed using the Agilent 2100 Bioanalyser. These final libraries were pooled in equimolar amounts using the Qubit and Bioanalyzer data. The quantity and quality of the pool was assessed by Bioanalyzer and subsequently by qPCR using the Illumina Library Quantification Kit from Kapa on a Roche Light Cycler LC480II according to manufacturer's instructions. The pool of libraries was sequenced on one flow cell of the MiSeq at 2 × 150 based paired-end sequencing with v2 chemistry.

Due to the high-expected levels of polyA RNA introduced during the RNA extraction 15% of phiX library was loaded onto the flow cell for sequencing to produce a more balance base composition. The Fastq files were QC filtered as follows. Initial processing and quality assessment of the sequence data were performed using an in-house pipeline (developed by Richard Gregory, CGR). Briefly, base calling and de-multiplexing of indexed reads was performed by CASAVA version 1.8.2 (Illumina). The raw fastq files were trimmed to remove Illumina adapter sequences using Cutadapt version 1.2.1 (Martin M. C., 2011). The option '-O 3' was set, so the 3' end of any reads which matched the adapter sequence over at least 3 bp was trimmed off. The reads were further trimmed to remove low quality bases, using Sickle version 1.200. After trimming, reads shorter than 10 bp were removed. If both reads from a pair passed this filter, each was included in the R1 (forward reads) or R2 (reverse reads) file. If only one of a read pair passed this filter, it was included in the R0 (unpaired reads) file.

There was significant variation in the amount of viral RNA present in each sample. This presented challenges when obtaining a consensus alignment for the viral genetic material. Therefore, the sequence data were aligned using several approaches before settling on two distinct types of analysis, a low stringency and high stringency approach. Initially we used a low specificity approach in order to gain maximum data return while accepting that some of the alignment output may be of low quality.

In this approach, the sequence reads were simply aligned to the viral genome using TopHat (Trapnell et al., 2012). In a second, highly stringent approach, we used Bowtie2 to map the reads to the viral genome. We then parsed the output to discard reads that did not map in a proper mate pair before removing PCR duplicates both steps were done using Sequence Alignment Map (SAM) tools. At this stage any reads less than 25 nucleotides were not considered. In either case, the output files were converted into SAM files to allow manual searching for sequence reads corresponding to the region of the G glycoprotein known to contain non-template insertions. In addition, the reads aligned to the virus genome were analysed using QuasiRecomb (Töpfer et al., 2013). This generated a coverage file indicating how many sequence reads covered each nucleotide of the genome and the frequency of each of the four nucleotides at each position on the virus genome. From this output we were able to generate a consensus sequence for the virus at each passage, which was later analysed using CLUSTAL in MEGA 6.06 (Tamura et al., 2013; Larkin et al., 2007).

2.3   Study with samples of patients from the EBOV outbreak 2014-2015 in West Africa

As a part of the response for the EBOV outbreak 2014-2015 in West Africa; the European Mobile Laboratory (EMLab)(http://www.emlab.eu/) deployed several teams to Guinea and Sierra Leone to provide molecular diagnostics for the EBOV outbreak. The samples collected during this study were provided

by the EMLab, where I deployed (Coyah team #19). In order to use patient's samples for this study, the National Committee of Ethics in Medical Research of Guinea approved the use of diagnostic leftover samples and corresponding patient data for this study (permit no. 11/CNERS/14). As the samples had been collected as part of the public health response to contain the outbreak in Guinea, informed consent was not obtained from patients and not required under the ethical agreement.

2.3.1 RNA preparation of patient samples

The viral genome sequence was derived from the RNA extracted for diagnostic purposes from blood samples in the field with no pre-amplification of the viral genome. These samples were processed by the EMLab in the Ebola Treatment Centre in Guinea. The RNA extraction of the samples were done using 140µL of EDTA-blood or serum samples and the viral RNA extraction kit (Qiagen; 52904) following the manufacturer's instructions. Samples were shipped to EMLab centre in Hamburg and then to our laboratory at the University of Liverpool, where we had Home Office approval to store and analyse these samples.

2.3.2 Next generation sequencing analysis of the patient's samples

Similar to the next generation sequencing analysis of the guinea pig samples; the following experiment was done by our collaborators in CGR (University of Liverpool) and by Dr. David Mathews at the University of Bristol. In order to perform the next generation sequencing analysis extracted RNA was DNase treated with Turbo DNase (Ambion) according to instructions of the manufacture. RNA sequencing libraries were prepared from the resultant RNA using the Epicentre ScriptSeq v2 RNA-Seq Library Preparation Kit. Following 10–15 cycles of amplification, libraries were purified using AMPure XP beads. Each library was quantified using Qubit and the size distribution assessed using the Agilent 2100 Bioanalyzer. These final libraries were pooled in equimolar amounts using the Qubit and Bioanalyzer data with 9–10 libraries per pool. The quantity and quality of the pool was assessed by Bioanalyzer and subsequently by qPCR using the Illumina Library Quantification Kit from Kapa on a Roche Light Cycler LC480II according to manufacturer's instructions. Each pool of libraries was sequenced on one lane of a HiSeq2500 at 2 × 125-bp paired-end sequencing with v4 chemistry. The trimmed fastq files were first aligned to a copy of the human genome using Bowtie2 (Langmead et al., 2012) and the unaligned reads were then mapped with Bowtie2 to a list of 3731 known viral genomes excluding EBOV genomes.

The reads that were still unmapped were then aligned to the EBOV genome either the prototype strain isolated in Zaire in 1976 (AF086833.2/ NC_002549) or a strain isolated during the current outbreak (KJ660348.2). For this step we again used Bowtie2 and the resultant alignment files were filtered with samtools to remove unmapped reads and reads with a mapping quality score below 11, followed by filtering with mark-up to remove PCR duplicates. The resultant BAM file was then analysed by Quasirecomb (Töpfer et al., 2013) to generate a phred-weighted table of nucleotide frequencies which were parsed with a custom perl script to generate a consensus genome in fasta format.

This consensus genome was then used as a reference genome to which we remapped the sequence reads which did not map to the human genome or other viruses in order to generate a second consensus. In this way we were able to manually determine if the reference genome used by Bowtie2 influenced the process of calling a consensus genome. The pipeline is entirely open source and implemented in the Galaxy environment (Goecks et al., 2010). Sequence alignment maps were manually inspected and curated over regions with consistent low coverage (for example, at the 5′ ends).

2.4 Analysis of the amino acid variations for the guinea pig experiment and for the patient's samples from the EBOV outbreak 2014-2015 in West Africa

Once the consensus sequence was generated for each passage in the guinea pig experiment, an analysis of the nucleotides and amino acid variations in each passage was done using the Molecular Evolutionary Genetics Analysis version 6.0 (MEGA 6.06) (Tamura et al., 2013) by CLUSTAL W analysis (Larkin et al., 2007; Thompson et al., 1994) running a pairwise alignment followed of a multiple alignment and a phylogenetic analysis. The variations in the amino acid sequences in each passages for guinea pig samples were compared to the input EBOV reference sequence (Zaire ebolavirus isolate Ebola virus/H.sapiens-tc/COD/1976/Yambuku-Mayinga, complete genome NC_002549). This analysis allowed to determinate the amino acid changes in the consensus sequence of each passage for the guinea pig experiment. For the patient samples from the EBOV outbreak 2014-2015 in West Africa, a similar analysis was done.

Once the 179 consensus sequences of the samples were generated, the sequences were aligned by CLUSTAL analysis using MEGA6.06 program. In order to detect mutations in the consensus sequences of the 179 samples, the EBOV reference sequence (NCBI reference sequence number: NC_002549) and two other EBOV sequences from the beginning of the outbreak (NCBI number: KJ660346.2 and KJ660347.2) were compared to the 179 sample

sequences. The sequences KJ660346.2 and KJ660346.2 are the complete genomes from Zaire ebolavirus isolate H. sapiens-wt/GIN/2014/Makona-Kissidougou-C15 and C07 respectively. Once the sequences were aligned, a comparison of the amino acid sequences of each sample was done; then the percentage of mutation in each EBOV protein was calculated.

2.5   Plasmids used for the expression of *Ebolavirus* proteins and for the Ebola virus (Makona strain) mini-genome system.

Codon optimized (for expression in human cells) cDNA sequence encoding for EBOV VP24 protein (NCBI reference sequence number: NP_066250.1); EBOV NP protein (NCBI reference sequence number: NP_066243.1); RESTV VP24 (NCBI reference sequence number:NP_690586.1) or RESTV NP (NCBI reference sequence number: NP_690580.1) were cloned into the multiple cloning site (MCS) region of the pEGFP-C1 and pEGFP-N1 vectors (Clontech and Gene Art-Life Technologies) (*Figure 2.1A*) to generate EBOV VP24, EBOV NP, RESTV VP24 or RESTV NP with C-terminal EGFP tag or N-terminal EGFP tag. Once the sequence was inserted in the plasmids sequencing analysis was done to confirm the correct insertion of the sequence; all inserted sequences are listed in details in the Appendix.

Plasmids used for the mini-genome system for Ebola virus (Makona strain) were designed following the mini-genome system model developed by Mülhgerber in 1999 with some modifications (Mülhgerber et al., 1999). Five

plasmids were used for the EBOV mini-genome system; four support plasmids that express the VP30, VP35, L and NP. To construct all the plasmids a pUC57-A338 plasmid (Gene Art-Life Technologies) was used as a backbone (*Figure 2.1B*). For the mini-genome plasmid, plasmid expressing an artificial viral genomic RNA that contains the leader and trailer sequences from EBOV (NCBI sequence number KJ660347.2) with the luciferase gene (NCBI sequence number AF058756.1) as a reporter gene was made (Gene Art-Life Technologies). In order to do this, a T7 promoter sequence was inserted along with the last 740 nt of the 5 end of the EBOV genome (NCBI number: KJ660347.2) trailer sequence from 18220 to 18959 (bottom strand from 5' to 3'). Then, a firefly luciferase gene (NCBI sequence number AF058756.1) was inserted along with the first 469 nt of the 3 end of the EBOV genome including the translational start codon of the NP gene (bottom strand from 5' to 3'). Finally, a hepatitis delta virus ribozyme sequence and the T7 terminator sequence were inserted (*Figure 2.1C*). For the support plasmids, a T7 promoter sequence was inserted along with a encephalomyocarditis virus internal ribosomal entry site (EMCV IRES), the ORF of the EBOV proteins (NP, VP30, VP35 or L) and a T7 terminator sequence (*Figure 2.1D*).

The EBOV sequence used in this study to design all the constructs used in the mini-genome system was the sequence from Zaire ebolavirus isolate H. sapiens-wt/GIN/2014/Makona-Gueckedou-C07 complete genome (NCBI sequence number KJ660347.2); this sequence corresponds to an isolate from early 2014 from the EBOV outbreak in West Africa.

A.

pEGFP-C1
4.7 kb

Ase I (8)
ApaL I (4360)
pUC ori
*Eco*O109 I (3854)
HSV TK poly A
Kan^r/Neo^r
SV40 ori
P_{SV40_e}
Stu I (2577)
$P_{CMV\ IE}$
SnaB I (341)
Nhe I (592)
*Eco*47 III (597)
Age I (601)
EGFP
SV40 poly A
BsrG I (1323)
MCS (1330–1417)
f1 ori
P
Mlu I (1642)
Dra III (1872)

EGFP 1330 1340 1350 1360 1370 1380 1390 1400 STOPs
TAC AAG TCC GGA CTC AGA TCT CGA GCT CAA GCT TCG AAT TCT GCA GTC GAC GGT ACC GCG GGC CCG GGA TCC ACC GGA TCT AGA TAA CTG ATC A
BspE I   *Bgl* II *Xho* I   *Hind* III *Eco*R I   *Pst* I   *Sal* I   *Kpn* I   *Apa* I   *Bam*H I   *Xba* I*   *Bcl* I*
Sac I   *Acc* I   *Asp*718 I *Bsp*120 I   *Xma* I
*Ecl*136 II   *Sac* II   *Sma* I

pEGFP-N1
4.7 kb

Ase I
ApaL I
pUC ori
*Eco*O109 I
HSV TK poly A
Kan^r/Neo^r
SV40 ori
P_{SV40_e}
Stu I
SnaB I
$P_{CMV\ IE}$
MCS
EGFP
SV40 poly A
BsrG I
Not I
Xba I*
f1 ori
P
Afl II
Dra III

B.

pUC57
(2710 bp)

*Eco*O109I 2698
*Aat*II 2641
*Ssp*I 2525
*Pdm*I 2318
*Bcg*I 2239
*Sca*I 2201
bla (Ap^R)
2510 146
2236
2510 1650 1490
*Gsu*I 1808
*Cfr*10I 1803
*Eco*31I 1790
*Eam*1105I 1718
*Cai*I 1241
rep (pMB1)
*Bse*YI 1134
*Afl*III, *Psc*I 830
*Bsa*XI 683
*Sap*I 707
876
493
476
396
236
lacZ
MCS
*Pfo*I 46
*Bst*API 179
*Nde*I 183
*Ehe*I 235

M13/pUC sequencing primer (-20),17- mer   396   476
*Eco*R I   *Ecl*136II   *Acc*65I   *Mwa*1269I   *Ctr*9I   *Hinc* II   *Alf*I
*Xap*I   *Sac*I   *Kpn*I   *Bsp*68I   *Mph*103I   *Xba* I   *Eco*RV   *Bam*HI   *Eco*88I   *Apa* I   *Sal*I   *Pst*I   *Eco*147I   *Pae* I   *Hind*III
Sma I   *Bsp*120I   *Xmi* I

5′ G TAA AAC GAC GGC CAG TGA ATT CGA GCT CGG TAC CTC GCG AAT GCA TCT AGA TAT CGG ATC CCG GGC CCG TCG ACT GCA GAG GCC TGC ATG CAA GCT TGG
3′ C ATT TTG CTG CCG GTC ACT TAA GCT CGA GCC ATG GAG CGC TTA CGT AGA TCT ATA GCC TAG GGC CCG GGC AGC TGA CGT CTC CGG ACG TAC GTT CGA ACC
LacZ ◄── Val Val Ala Leu Ser Asn Ser Ser Pro Val Glu Arg Ile Cys Arg Ser Ile Pro Asp Arg Ala Arg Arg Ser Cys Leu Gly Ala His Leu Ser Pro

CGT AAT CAT GGT CAT AGC TGT TTC CTG 3′
GCA TTA GTA CCA GTA TCG ACA AAG GAC 5′
Thr Ile Met Thr Met
M13/pUC reverse sequencing primer (-20),17- mer

C.

D.

Figure 2.1 *Representation of all plasmids used in this study. (A) pEGFP_c1 and pEGFP-N1 vectors used for the expression of EBOV VP24/ NP and RESTV VP24/NP. (B) pUC57 plasmid used for the mini-genome system and (C) diagram of the insert for mini-genome plasmid and the (D) support plasmid; L support plasmid*

2.6 Cell Culture

Mainly two cell lines were used in this study; human embryonic kidney 293 cells (HEK 293T) which were obtain from the ECACC, UK and BSR-T7, a baby hamster kidney (BHK) derived cell line which stably express T7 polymerase; this cells which were kindly provided by Dr. John Barr from the University of Leeds.

2.6.2 Calcium phosphate transfection and EBOV proteins expression.

HEK 293T cells were grown in Dulbecco's modified Eagle's medium (DMEM; Sigma-Aldrich) supplemented with 10% fetal bovine serum (FBS; Sigma-Aldrich) and 1% penicillin-streptomycin (Sigma-Aldrich) at 37 °C with 5% CO_2. To transfect HEK 293T cells, two 145 cm^2 dishes were seeded with 4×10^6 cells 24h prior to calcium phosphate transfection to achieved approximately 50% confluence. 25.6 μg of plasmid DNA (encoding EGFP, EGFP-EBOV/VP24, EBOV/VP24-EGFP, EGFP-RESTV/VP24, RESTV/VP24-EGFP, EGFP-EBOV/NP, EBOV/NP-EGFP, EGFP-RESTV/NP or RESTV/NP-EGFP) and 156 μL of $CaCl_2$ were mixed and made up to 1280μL with nuclease-free water, and added in a drop wise manner to 1280μL of 2 X HBS (274nM NaCl, 50 mM HEPES, 1.5 mM Na_2HPO_4, pH adjusted to 7.0). The mixture was incubated then for 30 minutes at room temperature and then added drop wise

into cells. Twenty-four-hour post-transfection the cells were harvested, lysed, and immunoprecipitated using a GFP-Trap (Chromotek).

2.6.3 Lipofectamine transfection and expression of EBOV Makona mini-genome system.

BSR-T7 cells were grown in DMEM (Sigma-Aldrich) with 10% FBS (Sigma-Aldrich) plus 1X Glutamax (Invitrogen) and 600 μg/mL of G418 (Life technologies) added at every second passage (stock at 100 mg/ml, add 90 μl to 15ml of medium in a T75 flask). For transfection of a 24-well multi well dish (Appleton woods), 8×10^4 cells were seeded overnight to achieve desired cell density of approximately 90%. Lipofectamine 2000 (Invitrogen) was used to transfect the BSR-T7 cells; different amount of the mini-genome plasmids (miniG-Luc 0.5 μg, N 0.25 μg, VP35 0.125 μg, VP30 0.125 μg and L 0.125 μg) were add into 50 μL of OPTI-MEM (Invitrogen) while 2.5μL of Lipofectamine 2000 was add into another 50μl of OPTI-MEM (Invitrogen). Lipofectamine 2000 was then incubate for 5 minutes at room temperature in the OPTI-MEM media and then the mixture of Lipofectamine 2000 plus the mini-genome plasmids was added; the mixture was then mix well by pipetting and incubated for 20 minutes at room temperature and finally added drop wise onto cells maintained in 1 mL of growth media (DMEM with 10% FBS). Lipofectamine 2000 was removed by changing medium at 4 h post-transfection and then the cells were leave it incubating at 37C$^\circ$, 5% CO_2 for 24 h before the analysis. Luciferase detection was then done using a western blot analysis and a Dual-Luciferase$\circledR$ Reporter Assay System (Promega; E1910).

2.7   EGFP Coimmunoprecipitations

EGFP,   EGFP-EBOV/VP24,   EBOV/VP24-EGFP,   EGFP-RESTV/VP24,

RESTV/VP24-EGFP, EGFP-EBOV/NP, EBOV/NP-EGFP, EGFP-RESTV/NP

and RESTV/NP-EGFP immunoprecipitations were performed using a GFP-

Trap_A (Chromotek), and cell pellet was resuspended in 200 μL of lysis buffer

(10 mM Tris/Cl pH 7.5; 150 mM NaCl; 0.5 mM EDTA; 0.5%NP40)

supplemented with Halt Protease Inhibitor Cocktail EDTA-Free (Thermo

Scientific) and incubated for 30 minutes on ice. The lysate was clarified by

centrifugation at 14,000g and diluted five-fold with dilution buffer (10 mM

Tris/Cl pH 7.5; 150 mM NaCl; 0.5 mM EDTA) supplemented with Halt Protease

Inhibitor Cocktail EDTA-Free (Thermo Scientific). The GFP-Trap agarose

beads were equilibrated with ice-cold dilution buffer supplemented with Halt

Protease Inhibitor Cocktail EDTA-Free (Thermo Scientific) and then incubated

with diluted cell lysate overnight at 4 °C on a rotator, followed by centrifugation

at 2700*g* for 2 minutes. The bead pellet was washed two times with wash buffer

(10 mM Tris/Cl pH 7.5; 150 mM NaCl; 0.5 mM EDTA) supplemented with Halt

Protease Inhibitor Cocktail EDTA-Free (Thermo Scientific). After the removal

of wash buffer, the beads were resuspended with 25 μL of elution buffer (200

mM Glycine pH 2.5) and incubated for 10 minutes at room temperature in a

thermal shaker to elute bound proteins. The beads were then collected by

centrifugation, and the eluted proteins were transferred to a 1.5 mL centrifuge

tube. This step was repeated four times to ensure the maximum elution and

all of the eluates were combined; 10 μL of Tris-base (1M) buffer (pH 10.4) was

added to neutralize the eluate. Immunoprecipitated samples were then

analysed using label-free mass spectrometry. For western blot analysis the elution step was done in 100µL of 2xSDS-sample buffer (120mM Tris/Cl pH6.8; 20% glycerol; 4% SDS, 0.04% bromophenol blue; 10% β-mercaptoethanol).

2.8   RNase treatment and coimmunoprecipitations

To analyse wheather the interaction of EBOV/NP with selected cellular proteins was real or due to interaction of some of those cellular proteins to the RNA, which NP binds; an RNAse treatment was done before the co-immunoprecipitations with GFP-Trap. In order to do this, the cell pellet was resuspended in 200µl of lysis buffer (10mM Tris/Cl pH 7.5; 150mM NaCl; 0.5mM EDTA; 0.5%NP40) supplemented with HaltTM Protease Inhibitor Cocktail EDTA-Free (Thermo Scientific) and then incubated for 30 minutes on ice. The lysate was then clarified by centrifugation at 14000 g and diluted five-fold with dilution buffer (10mM Tris/Cl pH 7.5; 150mM NaCl; 0.5mM EDTA) supplemented with HaltTM Protease Inhibitor Cocktail EDTA-Free (Thermo Scientific) and 15 units of RNAse (QIAGEN; 19101); the rest of steps for the co-immunoprecipitation protocol was done as described in section 2.7.

2.9   Reverse coimmunoprecipitation

Immunoprecipitations for KPNA6, LMNB1, EBOV/NP, HSP70 and BAG-2 were performed using the immobilized recombinant protein G resin (Generon) and specific antibodies against KPNA6 (E-11) (Santa Cruz Biotechnology; sc-

390055), LMNB1 (ZL-5) (Santa Cruz Biotechnology; sc-56145), EBOV/NP (IBT BIOSERVICES; 01-012), HSP70 (Abcam; ab2787) and BAG-2 (Abcam; ab47106). The cell pellets were incubated for 30 minutes on ice with 200 µl of lysis buffer; the lysate was clarified by centrifugation and diluted five-fold with dilution buffer prior to adding 2 µg of the primary antibody and then incubated at 4 °C on a rotator for 2 h. The protein G resin (Generon) was equilibrated with ice-cold dilution buffer and then incubated at 4 °C on a rotator with diluted cell lysate containing the antibody overnight at 4 °C on a rotator, followed by centrifugation at 2700g for 2 minutes to remove nonbounds. The wash and elution steps were preformed as previously described in section 2.7.

2.10 Label-Free Mass Spectrometry

The following experiment along with the label free analysis was done by Dr. Stuart Armstrong in the Department of Infection Biology (University of Liverpool). Eluted samples for EGFP, EGFP-EBOV/VP24, EBOV/VP24-EGFP, EGFP-EBOV/NP, EBOV/NP-EGFP, EGFP-RESTV/NP and RESTV/NP-EGFP were diluted two-fold with 25 mM ammonium bicarbonate. Rapigest (Waters) was added to a final concentration of 0.05% (w/v), and the sample was heated to 80 °C for 10 minutes. Proteins were reduced with 3 mM dithiothreitol (DTT) (Sigma) at 60 °C for 10 minutes, then alkylated with 9 mM iodoacetimide (Sigma) at room temperature for 30 minutes in the dark. Proteomic grade trypsin (Sigma) was added (0.2 µg), and samples were incubated at 37 °C overnight. The Rapigest was removed by adding trifluoroacetic acid (TFA) to a final concentration of 1% (v/v) and incubating at

37 °C for 2 h. Peptide samples were centrifuged at 12 000*g* for 60 minutes (4 °C) to remove precipitated Rapigest. Each digest was concentrated and desalted using C18 Stage tips (Thermo Scientific), then dried down using a centrifugal vacuum concentrator (Jouan) and resuspended in a 0.1% (v/v) TFA, 3% (v/v) acetonitrile solution.

For the mass spectrometry (MS) analysis two different approaches were done for the different group of proteins. For EGFP, EGFP-EBOV/VP24 and EBOV/VP24-EGFP, the peptide mixtures (2 µl) were analysed by online nano flow liquid-chromatography (LC) using the nano ACQUITY-nLC system (Waters MS technologies, Manchester, U.K.) coupled to an LTQ-Orbitrap Velos (Thermo Fisher Scientific, Bremen, Germany) mass spectrometer equipped with the manufacturer's nanospray ion source. The analytical column (Nano ACQUITY UPLCTM BEH130 C18 15 cm × 75 µm, 1.7 µm capillary column) was maintained at 35 °C and a flow-rate of 300nL/min. The gradient consisted of 3–40% acetonitrile in 0.1% formic acid for 50 min then a ramp of 40–85% acetonitrile in 0.1% formic acid for 3 min. Full-scan MS spectra (*m/z* range 300–2000) were acquired by the Orbitrap at a resolution of 30 000. Analysis was performed in data-dependent mode. The top-20 most intense ions from MS1 scan (full MS) were selected for tandem mass spectrometry (MS/MS) by collision-induced dissociation (CID), and all product spectra were acquired in the LTQ ion trap.

In the case of EGFP-EBOV/NP, EBOV/NP-EGFP, EGFP-RESTV/NP, RESTV/NP-EGFP and its control EGFP; peptides were analysed by on-line

nanoflow LC using the Thermo EASY-nLC 1000 LC system (Thermo Fisher Scientific) coupled with Q-Exactive mass spectrometer (Thermo Fisher Scientific). Samples were loaded on a 50cm Easy-Spray column with an internal diameter of 75 μm, packed with 2 μm C_{18} particles, fused to a silica nano-electrospray emitter (Thermo Fisher Scientific). The column was operated at a constant temperature of 35°C. Chromatography was performed with a buffer system consisting of 0.1 % formic acid (buffer A) and 80 % acetonitrile in 0.1 % formic acid (buffer B). The peptides were separated by a linear gradient of 3.8 – 50 % buffer B over 30 minutes at a flow rate of 300 nl/minutes. The Q-Exactive was operated in data-dependent mode with survey scans acquired at a resolution of 70,000. Up to the top 10 most abundant isotope patterns with charge states +2, +3 and/or +4 from the survey scan were selected with an isolation window of 2.0Th and fragmented by higher energy collisional dissociation with normalized collision energies of 30. The maximum ion injection times for the survey scan and the MS/MS scans were 250 and 100ms, respectively, and the ion target value was set to 1E6 for survey scans and 1E4 for the MS/MS scans. Repetitive sequencing of peptides was minimized through dynamic exclusion of the sequenced peptides for 20s.

2.10.1   Label-Free Analysis

Label-free quantitation was performed using MaxQuant (MQ) software (version 1.3.0.5.) with its internal search engine Andromeda. Precursor mass and fragment mass were searched with mass tolerance of 6 ppm and 0.5 Da, respectively. All other settings were default. The search included variable

modifications of methionine oxidation and N-terminal acetylation and fixed modification of carbamidomethyl cysteine. Enzyme specificity was set to trypsin, minimal peptide length was set to seven amino acids, and a maximum of two mis-cleavages was allowed. The false discovery rate (FDR) was set to 0.01 for peptide and protein identifications. The Andromeda search engine was configured for a database containing human proteins (Uniprot release-2013_03), EBOV/VP24-EGFP constructs (56601 entries), EBOV/NP-EGFP or RESTV/NP-EGFP constructs. The MQ software further included a decoy database as well as containing common contaminants to determine the FDR and to exclude false-positive hits due to contamination by proteins from different species. For label-free quantification (LFQ) analysis, "multiplicity" was set to one. Feature matching between raw files was enabled using a retention time window of 2 minutes. "Discard unmodified counterpart peptides" was unchecked. Only unmodified and unique peptides were utilized. Averaged LFQ intensity values were used to calculate protein ratios.

2.10.2   Bioinformatics Analysis

Label-free mass spectrometry results were processed and analysed using the Perseus software (MaxQuant); this software was used to perform the statistical analysis and to differentiate background proteins (those cellular proteins that interacted with EGFP alone) from interacting proteins (those cellular proteins that interacted with EGFP-VP24, VP24-EGFP, EGFP-EBOV/NP, EBOV/NP-EGFP, EGFP-RESTV/NP or RESTV/NP-EGFP). LFQ intensity values were analysed using a T-Test with a confidence of 95%. Volcano plots graphic and

a table were generated showing the statistical significant proteins, those proteins that had a high probability of interacting with EBOV/VP24, EBOV/NP or RESTV/NP. The String 9.05. program was used to predict the protein–protein interactions of the statistically significant proteins and also to group proteins according function.

2.11 Sodium-Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE) and Western Blot Analysis

After the co-immunoprecipitation or reverse co-immunoprecipitation, beads were then resuspended and boiled in 100 µl of 2× SDS-sample buffer at 95° for 10 minutes to elute the bound proteins. The beads were then collected by centrifugation, and SDS-PAGE was performed with the supernatant (elution faction) and the whole cell lysate (input sample). To perform the SDS-PAGE 10% SDS polyacrylamide gels (*Table 2.2*) were used; gel was run at 200V for approx. 1 hour in SDS-running buffer (25mM Tris-HCl (pH8.3), 250mM glycine, 0.1% (w/v) SDS); then, gels were transferred to methanol activated polyvinylidene difluoride (PVDF) membranes (Millipore) in Towbin transfer buffer (25mM Tris, 192mM Glycine and 20% (v/v) methanol) at 15V for 1.5 h using a semidry transfer system.

Transferred membranes were blocked in 5% skimmed milk powder dissolved in tris-buffered saline (TBS) 0.1% Tween (TBS-T) (50 mM Tris-HCl (pH8.3), 150 mM NaCl, and 0.5% (v/v) Tween-20) buffer for 60 minutes at room temperature. Primary antibodies (Table 2.3) were diluted 1:1000 in blocking

buffer and then incubated at 4 °C overnight. After three washes, blots were incubated with appropriate horseradish peroxidase (HRP) secondary antibodies (Table 2.4) were diluted in blocking buffer at a 1:2000 for 1 h at room temperature. Blots were developed using enhanced chemiluminesence reagent (BioRad) and detected using a BioRad Imaging system.

Table 2.2 *Recipe of SDS-polyacrylamide gel (resolving gel and stacking gel).*

Resolving gel (10ml)				Stacking gel (5%)	
Concentration	12%	10%	7.5%	Concentration	5%
30% acrylamide	4 ml	3.3ml	2.5ml	30% acrylamide	0.83mL
1.5 M Tris-HCl pH 8.8	2.5ml	2.5ml	2.5ml	1. M Tris-HCl pH 6.8	0.63mL
ddH$_2$O	3.3ml	4ml	4.8ml	ddH$_2$O	3.4ml
10% (w/v) SDS	100µl	100µl	100µl	10% (w/v) SDS	50µl
10% (w/v) APS	100µl	100µl	100µl	10% (w/v) APS	50µl
TEMED	10µl	10µl	10µl	TEMED	5µl
Resolution	10-70 kDa	15-100 kDa	25-250 kDa		

Table 2.3 *Primary antibodies used in this study. The dilutions used for each antibody for western blot are shown in this table.*

Antigen	Cat n°	Manufacturer	Species	Dilution
GFP	sc-8334	Santa Cruz Biotechnology	Rabbit polyclonal	1/2000
Karyopherin α6	sc-390055	Santa Cruz Biotechnology	Mouse monoclonal	1/1000
Lamin B1	sc-56145	Santa Cruz Biotechnology	Mouse monoclonal	1/1000
Na^+/K^+-ATPase α	sc-48345	Santa Cruz Biotechnology	Mouse monoclonal	1/1000
VDAC1/Porin	ab14734	Abcam;	Mouse monoclonal	1/1000
EBOV/NP	0301-012	IBT bioservices	Rabbit polyclonal	1/1000
DnaJA2	sc-136515	Santa Cruz Biotechnology	Mouse monoclonal	1/1000
PCNA	ab29	Abcam;	Mouse monoclonal	1/1000
AIF	ab32516	Abcam;	Rabbit monoclonal	1/1000
BAG-2	ab47106	Abcam;	Rabbit polyclonal	1/1000
STAT1	ab3987	Abcam;	Mouse monoclonal	1/1000
HSP70	ab2787	Abcam;	Mouse monoclonal	1/1000
HSP90	ab1429	Abcam;	Mouse monoclonal	1/1000
GAPDH	ab8245	Abcam;	Mouse monoclonal	1/5000
Firefly Luciferase	ab185923	Abcam;	Rabbit monoclonal	1/1000

Table 2.4 *Secondary antibodies used in this study; dilutions used for each antibody for western blot are shown in this table.*

Antigen	Catalog n°	Manufacturer	Species	Dilution
Mouse	A4116	Sigma	Goat	1/2000
Rabbit	A6154	Sigma	Goat	1/2000

2.12 EBOV infection and treatment with ouabain

MRC-5 human fetal lung cells (ECACC; 05081101) were maintained in Leibovitz's L-15 media (Invitrogen; 31415-086) supplemented with 10% FBS (Sigma; F9665). Media was removed from the cell monolayers, and the cells were infected with Ebola Zaire virus (strain ME718) at a multiplicity of infection of ~0.1, with virus left to adsorb for 1 h at 37 °C. Leibovitz's L-15 media containing 2% fetal bovine serum was added, and the infected cells were incubated at 37 °C for 1 h. In triplicate flasks, Ouabain Octahydrate (Sigma; O3125) was added to the media to a final concentration of 20 nM; a further three flasks were left as an untreated control. At 24 and 48 h post infection, a sample of media from each flask was added to AVL inactivation buffer (Qiagen; 19073) prior to removal from the CL4 laboratory for nucleic acid extraction.

The QIAamp Viral RNA Mini Kit from Qiagen (cat. no. 52904) was used for nucleic acid extraction, as this facilitated transport of the nucleic acid samples

from CL4 to CL2. Following the extraction of viral RNA, the cDNA was synthesized using specific primers against the gene encoding the nucleoprotein of Zaire Ebola virus. The cDNA was synthesized using the Thermo Script RT-PCR System from Invitrogen (cat. no. 11146-024) following the manufacturer's instructions. PCR was performed with Paq5000 DNA polymerase (Agilent Technologies) (data not shown). Finally, a qPCR was performed to compare the RNA levels in the infected cells treated and untreated with Ouabain experiment. This was performed with the iTaq Universal SYBR Green Supermix (BIO RAD).

2.12.1 Ouabain and its effect on EBOV/VP24 expression.

For EBOV/VP24 transfection, 5×10^4 HEK293T cells were seeded overnight in a 24-well multi well dish (Appleton woods), to achieve desired cell density of approximately 70%. Calcium phosphate transfection was used to transfect the HEK293T cells as was described before in section 2.6. The expression of EBOV/VP24 was then analysed using a western blot analysis using a highly specific EGFP antibody (Santa Cruz Biotechnology; sc-8334)

2.13 Evaluation of Luciferase production by the EBOV Makona mini-genome system using western blot analysis and a Dual-Luciferase® Reporter Assay System

Twenty-four hours post transfection the cells were harvested and lysed using

and passive lysis buffer 5X (Promega); in order to do this the media was removed form each well and cells were washed with 500μl of PBS; then 100μl of passive lysis buffer were added to each well and incubated at room temperature on a rocking platform for 15 minutes; the lysate was then was transfered into an 1.5ml and stored at -70 C° for further analysis. In order to detect luciferase an antibody against Firefly luciferase (abcam; ab185923) was used in a western blot analysis as was described before. A Dual-Luciferase® Reporter Assay System (Promega; E1910) was used as well to detect the production of firefly luciferase by the EBOV Makona mini-genome system. The Dual-Luciferase® Reporter Assay System helped us to improve the experimental accuracy by using a co-transfected "control" (plasmid expressing *Renilla luciferase;* promega; pGL4.74[*hRluc/TK*]) reporter which provided an internal control that serve as the baseline control. The Dual-Luciferase® Reporter Assay System (Promega; E1910) was used according to instructions of the manufacture.

2.14 Inhibition of Hsp70 using VER-15508 and its effect on the replication and transcription of EBOV using the mini-genome.

Expression of EBOV Makona mini-genome system using BSR-T7 cells was done as described before; at 4 h post-transfection, Lipofectamine 2000 was removed by changing medium which contained different concentrations of the HSP-70 inhibitor (VER-155008). VER-15508 was diluted in DMSO to make a stock solution of 10mM; then the stock solution was diluted in medium and added to the cells in different concentrations (5μM, 10μM, 20μM, 40μM, 60μM,

80μM and 100 μM). Cells were then incubating at 37C°, 5% CO_2 for 24 h before the analysis. Luciferase detection was then done using a western blot analysis and a Dual-Luciferase® Reporter Assay System (Promega; E1910).

2.14.1 Inhibition of HSP70 using VER-15508 and its effect on EBOV/NP expression.

For EBOV/NP transfection, $8x10^4$ BSR-T7 cells were seeded overnight in a 24-well multi well dish (Appleton woods), to achieve desired cell density of approximately 90%. Lipofectamine 2000 (Invitrogen) was used to transfect the BSR-T7 cells as was described before, 0.8 μg of EGFP, EGFP-EBOV/NP or EBOV/NP-EGFP DNA plasmid were added into 50 μl of OPTI-MEM (Invitrogen) while 2.0μl of Lipofectamine 2000 was add into another 50μl of OPTI-MEM (Invitrogen). Lipofectamine 2000 was then incubated for 5 minutes at room temperature in the OPTI-MEM media and then the mixture of Lipofectamine 2000 plus the EGFP, EGFP-EBOV/NP or EBOV/NP-EGFP plasmids was added; the rest of the procedure was followed as described and the EBOV/NP expression was then analysed using a western blot analysis using a highly specific EGFP antibody (Santa Cruz Biotechnology; sc-8334) and an EBOV/NP antibody (IBT bioservices; 0301-012).

CHAPTER 3: Elucidating variations in the nucleotide sequence of Ebola virus during adaptation in an animal model and during the 2014-2015 outbreak in West Africa.

3.1　Introduction

Ebolavirus causes severe haemorrhagic fever in humans and non-human primates. The case fatality rate is related to the species of Ebola viruses; EBOV has the highest case fatality rate (up to 90%) while RESTV is not pathogenic for humans. However, RESTV can cause viral haemorrhagic fever in non-human primates and illustrates the potential zoonotic threat of Ebola viruses (Feldmann et al., 2013; Weingartl et al., 2013; Marsh et al., 2011). EBOV outbreaks occur sporadically, with the most recent occurring in Southern Guinea in 2014 (Baize et al., 2014), with fatality rates over 50%. This epidemic spread to Liberia, Sierra Leone and Nigeria (Baden et al., 2014), with one case reported in Senegal.

EBOV is an enveloped non-segmented negative-sense single-stranded RNA virus with a genome of 19 kb in length encoding several proteins including: NP, VP35, VP40, the surface glycoprotein ($GP_{1,2}$), VP30, VP24 and the RNA-dependent RNA polymerase (L). The L protein is the catalytic subunit that forms an integral part of the polymerase complex that transcribes viral mRNA and replicates the viral genome. The combination and action of the EBOV gene products and their interactions with the host cell contribute to the severe haemorrhagic fever (Martinez et al., 2013).

There are several properties of the polymerase complex that contribute to virulence. The mRNA encoding $GP_{1,2}$ is transcribed through two disjointed reading frames in the genome. These two reading frames are brought together

by slippage of the viral polymerase complex at an editing site (a run of seven A residues) to insert an eighth A residue, generating an mRNA transcript that allows read-through translation of $GP_{1,2}$ (Volchkov et al., 1995; Sanchez et al., 1996) During EBOV infection and transcription of GP, approximately 20% of transcripts are edited, while the other 80% of unedited transcripts possess a premature stop codon. This results in the synthesis of a number of different products (Mehedi et al. 2011) including a truncated GP product (soluble GP or sGP), which is then secreted into the extracellular space. In a concept described as 'antigenic subversion' sGP has been proposed to induce a host antibody response that targets epitopes that sGP has in common with $GP_{1,2}$, thereby allowing sGP to bind and compete for anti-$GP_{1,2}$ antibodies (Mohan et al., 2012). The editing has been observed in both *in vitro* and *in vivo* models of infection (Shabmann et al., 2014; Volchkova et al., 2011).

Similar to other viruses with RNA genomes and which code an RNA dependent RNA polymerase, the L protein is unlikely to have extensive error correction, if at all. The molecular evolutionary rate for EBOV has been estimated at being between 2.2×10^{-4} and 7.06×10^{-4} nucleotide substitutions/site/year (Carroll et al., 2013), whereas for nuclear genomes with error correction, it is approximately 10^{-9} nucleotide substitutions/site/year (Liu et al., 2003). Therefore, the selection of phenotypic and hence genotypic variants of Ebola virus that can occupy new niches is facilitated through the high nucleotide substitution rate.

EBOV is proposed to be a zoonotic infection. Fruit bats are believed to be the natural reservoir and are not thought to develop disease (Leroy et al., 2005).

In fact, the origin of the virus that started the current outbreak is thought to have been a zoonotic transmission from a bat to a year-old boy in December 2013 (Baize et al., 2014). From this index case the virus spread by human-to-human contact throughout Guinea, Sierra Leone and Liberia (Gire et al., 2014). One major concern during the outbreak was that with increasing numbers of humans being infected in the current outbreaks, EBOV variants could be selected which were better adapted to be spread through human-to-human transmission.

Therefore, in this study, to investigate these factors, the potential evolution of EBOV was studied in two ways. First, a guinea pig model of infection was used to simulate the transmission to and adaptation of EBOV in a new host. Second, 179 patient samples from the 2014-2015 West African outbreak were used to estimate the rate of evolution of the virus in a human host and also look for which amino acid mutations were present in the different EBOV proteins along the outbreak.

The sequencing pipeline and gene analysis developed for the guinea pig model of infection was also used to analyse changes in EBOV evolution on the ongoing EBOV outbreak in West Africa. This kind of analysis will help us to gain insight into the evolution of the EBOV during the current outbreak. This analysis pipeline was also then applied to samples from fatal and non-fatal infection from the West African outbreak.

This work was conducted in close collaboration with Public Health England and also the European Mobile Laboratory (EMLab). Note, that I also deployed

to Guinea in 2015 as part of EMLab to gather samples and develop sequencing methodologies in support of this project.

Data from this chapter has been published or contributed towards the following publications:

Dowall, S. D., Matthews, D. A., **García-Dorival, I.,** Taylor, I., Kenny, J., Hertz-Fowler, C., ... & Hiscox, J. A. (2014). ***Elucidating variations in the nucleotide sequence of Ebola virus associated with increasing pathogenicity.*** *Genome biology*, *15*(11), 540.

Carroll, M. W., Matthews, D. A., Hiscox, J. A., Elmore, M. J., Pollakis, G., Rambaut, A., Hewson, R., **García-Dorival, I.,** …& Günter, S. (2015). ***Temporal and spatial analysis of the 2014-2015 Ebola virus outbreak in West Africa***. *Nature*. 524, 97–101 (06 August 2015).

3.2 Results

3.2.1 Adaptation of EBOV to grow in guinea pigs.

To investigate whether genetic changes in EBOV became established in response to adaptation in a different host, a guinea pig model of infection was used. Previous work had demonstrated that EBOV could adapt to grow in small animal models, including mouse (Edhibara et al., 2006), Syrian hamster (Edibhara et al., 2012) and guinea pig (Volchkova et al., 2011; Volchkov et al., 2000; Connolly et al., 1999). In this experimental system, guinea pigs were infected with EBOV and the virus was serially passaged to develop uniform lethality in guinea pigs (*Figure 3.1*). Dr. Roger Hewson, our collaborator at PHE, led this phase of the work. There were 10 guinea pigs per passage. Four animals were used for the preparation of spleen homogenate for subsequent virus infection (culled 7 days post challenge) and six were taken forward for measuring survival rates and clinical parameters (for up to 14 days post challenge). Adaptation of EBOV to growth in the guinea pigs was achieved with serial passage involving a subcutaneous injection of 10^4 $TCID_{50}$ EBOV, with spleens harvested 7 days post infection (as a source of progeny virus). Virus titer was determined and a new inoculum prepared before administering 10^4 $TCID_{50}$ EBOV to a new group of guinea pigs. This was repeated until there was clinical and virological evidence that the virus adapted to the guinea pig host. Animals were observed for 2 weeks post infection. Weight data indicated that guinea pigs showed a minimal response to the initial challenge, whereas with subsequent passages weight loss exceeding 10% was observed.

Similarly, with temperatures, the same responses were observed, where only after initial passage in the guinea pigs were temperature increases of between 1°C and 2.5°C observed. At passage two, several animals that met humane clinical endpoints displayed symptoms of hypothermia prior to being euthanized. Six animals from each passage study that were scheduled to last 14 days post infection were used to assess mortality. By five passages, 75% mortality was observed with a challenge dose of 10^4 $TCID_{50}$. There was also no increase in viral titer in the spleen collected from animals culled at day 7 (Table 3.1) compared with the previous passage, indicating that the viral burden had peaked. The minimum lethal dose of the passaged virus was determined to be 10^3 $TCID_{50}$.

10^4 TCID$_{50}$ virus
delivered subcutaneously

Virus concentration determined by
TCID$_{50}$ assay

Infected animals culled 7 days
post-infection

Spleens harvested and
homogenised
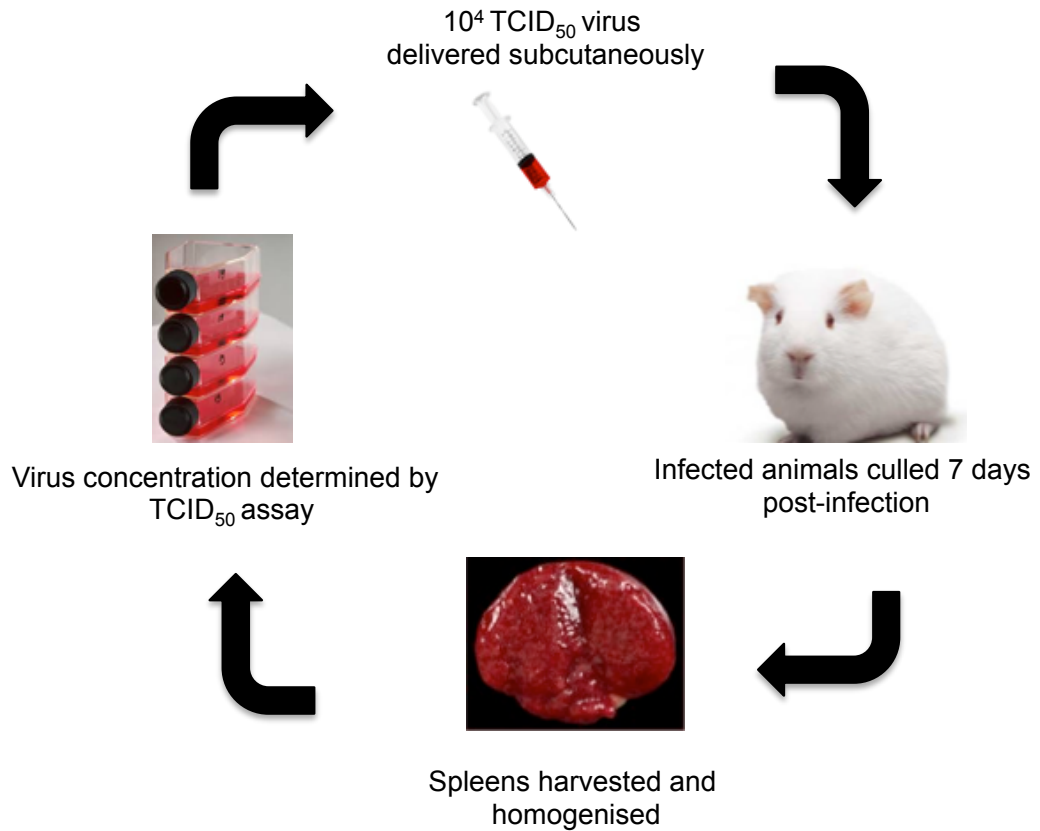
Figure 3.1. Passaging of virus in vivo. *In order to give a reproducible model of infection EBOV was passage five times in guinea pig in a forced evolution model. There were 10 animals per group, where four animals where used for harvesting spleens for virus preparation and the remaining six animals used to measure clinical parameters.*

3.2.2 Analysis of EBOV genome sequence with passage.

To investigate whether any nucleotide changes in the EBOV genome were potentially associated with adaptation, viral RNA was purified from spleens isolated from four guinea pigs from passage one through to five using a Qiagen Viral isolation kit. This allowed the safe transfer of nucleic acid from CL4 to CL2 and CL1 facilities. For further analysis, RNA was pooled from each passage and sequenced using MiSeq to avoid potential problems (associated with HiSeq) with the polyA carrier in the viral isolation kit. Sequence analysis indicated an increased proportion of sequence reads mapping to the EBOV genome with passage (*Table 3.1*). By using low stringency mapping approach, there were 4,298 reads in RNA sequenced from passage one material compared to 12,060 reads in RNA sequenced from passage five material. The decrease in percentage reads mapping to the EBOV genome in passage five compared to passage four may in part be due to the greater proportion of total reads mapped in passage five compared to passage four. Alternatively, these represent pooled samples and there was likely to be variation among individual clinical samples. Similarly, using a high stringency analysis of the alignments (see methods) we found 478 reads and 7,142 reads at passages one and five, respectively. This correlated well with the increase in viral titers observed with each passage (*Table 3.1*).

Table 3.1. *The titer of EBOV in the spleens isolated from guinea pigs. Four guinea pigs were taken from each passage increased, and then reached a plateau indicating that the virus had become adapted to grow in the guinea pig model. Note that PF are the reads that pass filter or have good quality enough for the analysis.*

Passage number	Virus titer from spleen preparation ($TCID_{50}$)	Total PF reads	Reads mapping to the EBOV genome	Reads mapping to the EBOV genome (%)	Mortality (number of deaths/ number challenged) (%)
1	2.1×10^4/spleen	2,429,959	4,298	0.18	0% (0/4)
2	3.0×10^7/spleen	3,505,156	5,655	0.16	50% (2/4)
3	5.8×10^7/spleen	3,453,615	9,736	0.28	0% (0/4)
4	6.1×10^7/spleen	2,696,262	13,783	0.51	25% (1/4)
5	6.1×10^7/spleen	3,264,859	12,060	0.37	75% (3/4)

3.2.3 Increased editing in the GP gene with passage.

Sequence analysis identified editing in the mRNA encoding $GP_{1,2}$, suggesting that viral mRNA might be co-purified or that there is editing of the genome itself during viral replication. Previous work has shown that approximately 20% of GP mRNA analysed from Ebola virus infected cells can be edited (Volchkov et al., 1995). When TopHat (Trapnell et al., 2012) was used to align the sequencing reads to the viral genome and then searched the aligned reads for evidence of insertions within sequence reads mapping to the appropriate area on the viral genome. Analysis of the sequence data from each of the passages revealed that at passage one there were no insertions (0/23 sequence reads mapped to that region), at passage two approximately 15% of sequence reads had insertions (3/22 reads), similarly at passage three approximately 15% of reads had insertions (10/68 reads) but by passage four there was an increase to approximately 30% (36/124) and by passage five there were insertions in approximately 25% of the sequence reads mapping to that region (25/99 reads). Although the numbers of sequence reads were low, the data suggested that the proportion of full length mRNA encoding $GP_{1,2}$ increased with sequential passage and this may be associated with the gain of virulence observed with sequential passage in the guinea pig model. One interpretation of this data is that the amount of $GP_{1,2}$ was limiting in early passages. As the proportion of edited mRNA increased so more $GP_{1,2}$ was available for virus assembly and on virus particles and this contributed to the increase in progeny virus observed in the later passages. However, a more stringent analysis of sequence reads that mapped to the relevant region of the glycoprotein gene showed additional A residues at the following rates: 0 out of 2 mapped reads

for passage one, 2 out of 3 (66%) mapped reads at passage two, 17 out of 45 (37.8%) reads at passage three, 45 out of 115 (39.1%) mapped reads for passage four and 37 out of 91 (40.7%) mapped reads for passage 5. Thus, an analysis of those reads mapped at a higher stringency indicated an increased rate of editing both overall and increasing with passage. However, the low number of reads overall means these observations must be treated with caution.

3.2.4 Nucleotide substitutions become established with passage.

EBOV sequence at each passage was analysed for coverage and variants using QuasiRecom (Topfer et al., 2013), allowing the consensus nucleotide to be determined at each position and also to derive a map of the minor variant frequency. Thus a consensus sequence for the virus was determined at each passage and this was compared to the published EBOV sequence at each particular passage. To determine if sequence/amino acid changes being established the frequency of minor variants was studied. For example, a minor variant at passage three that is not present at passage four indicates the sequence change did not become established. Thus whether a coding change became established with passage or not could be distinguished. There were two major types of substitutions; those that appeared in passages two to four and were selected against by passage five (*Table 3.2*) or alternatively substitutions that became established by passage five (*Table 3.3*). Some of these correlated with previous point mutational analysis, thus placing

confidence in the approach. One example of this is at the amino acid position

163 in VP24, where a Lys is change for an Arg (Kugelman et al., 2012).

The method used in this study to detect amino acid variations that might be

related with an increasing pathogenicity in a guinea pig model could be also

used to analyse changes in EBOV evolution during the ongoing 2014-2015

EBOV outbreak in West Africa that started during this PhD.

Table 3.2. Amino acid substitutions that are the predominant change in the virus population analysed during an individual passage. *The protein name is indicated, as is the amino acid position on the top of the table. Ebola virus (Mayinga strain 1976) was used as a control and therefore the amino acid present in the input sequence.*

NP

	191	323	414	566
Ebola virus	W	V	L	N
P1	R	D	R	N
P2	?	V	L	S
P3	W	V	L	S
P4	W	V	L	S
P5	W	V	L	S

VP35

	129	204	246
Ebola virus	S	N	I
P1	P	D	A
P2	S	D	I
P3	S	D	I
P4	S	D	I
P5	S	D	I

VP40

	15	66	259
Ebola virus	**E**	**P**	**M**
P1	Q	S	M
P2	?	P	R
P3	E	P	M
P4	E	P	M
P5	E	P	M

	sGP/ssGP											
ssGP/sGP/GP	**GP**											
	1	11	49	92	187	203	323	493	544	639	646	661
Ebola virus	**M**	**R**	**D**	**V**	**P**	**V**	**P**	**I**	**I**	**G**	**T**	**V**
P1	?	?	?	L	P	?	?	I	T	E	I	A
P2	K	?	?	V	L	I	T	T	T	G	T	?
P3	M	H	N	V	P	I	P	I	T	G	T	V
P4	M	R	N	V	P	I	P	I	T	G	T	V
P5	M	R	N	V	P	I	P	I	T	G	T	V

VP30

Ebola virus	214	248
	L	**Q**
P1	P	?
P2	L	R
P3	L	Q
P4	L	Q
P5	L	Q

VP24

Ebola virus	26	29	43	163	218
	L	**F**	**A**	**K**	**K**
P1	F	V	P	?	R
P2	F	F	A	K	?
P3	F	F	A	R	K
P4	F	F	A	K	K
P5	F	F	A	K	K

L

Ebola virus	30	38	161	525	537	538	669	705	707	811
	G	**N**	**R**	**N**	**K**	**L**	**I**	**M**	**G**	**Q**
P1	?	?	W	N	R	P	?	?	?	?
P2	G	N	R	N	K	L	S	?	?	K
P3	W	N	R	D	K	L	I	T	?	Q
P4	?	?	R	N	K	L	I	M	G	Q
P5	G	K	R	N	K	L	I	M	A	Q

L

	826	868	879	930	940	943	993	1096	1271	1280
Ebola virus	**S**	**S**	**F**	**T**	**L**	**I**	**T**	**L**	**Y**	**F**
P1	Y	S	?	?	?	?	?	?	?	?
P2	S	S	F	T	L	I	A	S	Y	F
P3	S	S	F	T	L	I	T	L	Y	F
P4	S	P	L	T	L	R	T	L	Y	F
P5	S	S	F	A	P	I	T	L	**STOP**	S

L

	1308	1478	1546	1733	1763	1949	1998	2144	2151	2186
Ebola virus	**S**	**N**	**A**	**F**	**L**	**H**	**S**	**N**	**F**	**M**
P1	?	?	?	?	?	?	?	N	-	-
P2	P	I	A	F	L	Q	S	N	V	M
P3	S	N	A	F	L	H	S	N	V	M
P4	S	N	A	Y	P	?	?	N	V	M
P5	?	H	E	F	L	H	T	K	V	T

L

	2197
Ebola virus	**L**
P1	-
P2	L
P3	L
P4	P
P5	L

Table 3.3. Amino acid substitutions that are the predominant change in the virus population analysed at passage 5. *The protein name is indicated, as is the amino acid position. Ebola virus (Mayinga strain 1976) was used as a control and therefore the amino acid present in the input sequence.*

NP

	566
Ebola virus	**N**
P1	N
P2	S
P3	S
P4	S
P5	S

VP35

	204
Ebola virus	**N**
P1	D
P2	D
P3	D
P4	D
P5	D

No amino acid mutations for VP40

GP/sGP/ssGP	sGP/ssGP		
	GP		
	49	203	544
Ebola virus	**D**	**V**	**I**
P1	?	?	T
P2	?	I	T
P3	N	I	T
P4	N	I	T
P5	N	I	T

No amino acid mutations for VP30

VP24

	26
Ebola virus	**L**
P1	F
P2	F
P3	F
P4	F
P5	F

L

	38	707	930	940	1271	1280	1478	1546	1998	2144	2151
Ebola virus	**N**	**G**	**T**	**L**	**Y**	**F**	**N**	**A**	**S**	**N**	**F**
P1	?	?	?	?	?	?	?	?	?	N	-
P2	N	?	T	L	Y	F	N	A	S	N	V
P3	N	?	T	L	Y	F	N	A	S	N	V
P4	N	G	T	L	Y	F	N	A	?	N	V
P5	K	A	A	P	**STOP**	S	H	E	T	K	V

L

	2186
Ebola virus	**M**
P1	-
P2	M
P3	M
P4	M
P5	T

3.2.5 Amino acid variations of the EBOV in the 179 genomes analysed from patients during the West African outbreak (2014-2015)

A deep sequencing approach was used to gain insight into the evolution of EBOV in Guinea from the ongoing West African outbreak. This was an approach based on analysis pipelines developed for a guinea pig model of EBOV infection and Hendra virus infection of human and bat cells (Dowall et al., 2014; Wynne et al., 2014). Here this approach was used to derive consensus EBOV genomes from individual patient samples that can be used to study viral genome evolution during the course of the outbreak. Viral genomes were derived primarily from blood samples that had been taken from patients in Guinea and sent to the EMLab, deployed by the World Health Organisation (WHO) within the Médecins Sans Frontières Ebola Treatment Centre Guéckédou in March 2014 to aid the diagnostic effort.

The viral genome sequence was derived from RNA sequencing analysis of the patient samples with no pre-amplification of the viral genome. In general, we selected a range of samples from both males and females of different ages and a fair representation of sequences for each month; and with cycle threshold (*C*t) values less than 20 for EBOV RNA. In this selected patient cohort, with a relatively high viral load, there was approximately 80% mortality; the amount of viral load as determined by quantitative reverse-transcription PCR (qRT–PCR).

Given the error-prone nature of EBOV genome replication, the potential amino acid variation in EBOV proteins was examined from the start of the sample collection in March 2014 to January 2015. The location of amino acid changes on EBOV proteins and their relative representation in the 179 assembled genomes were compared to an isolate identified in March 2014 (Baize et al., 2014) (*Figure 3.2*). While there is amino acid variation in all of the genomes sampled, there were very few changes in VP30, VP40 and VP24, and these changes are only in less than ~2% of the genomes sampled. However, a single amino acid substitution in VP24 is associated with adaptation to a new host (Dowall et al., 2014; Mateo et al., 2011) and this may be due to interactions with host-cell proteins (Garcia-Dorival et al., 2014; Basler et al., 2009). While some of the variation may be attributed to a purely random molecular clock pattern, in GP, VP35, NP and L there are some amino acid variations that are present in over ~15% of the genomes sampled. For example, in GP there is an A to V substitution in ~70.5% of the genomes sampled compared to the reference genome. Implications of the mutations within GP in relation to immune escape of therapeutics and vaccines will need to be assessed in pseudo type neutralization assays using EBOV monoclonal antibodies and serum from people who have been vaccinated.
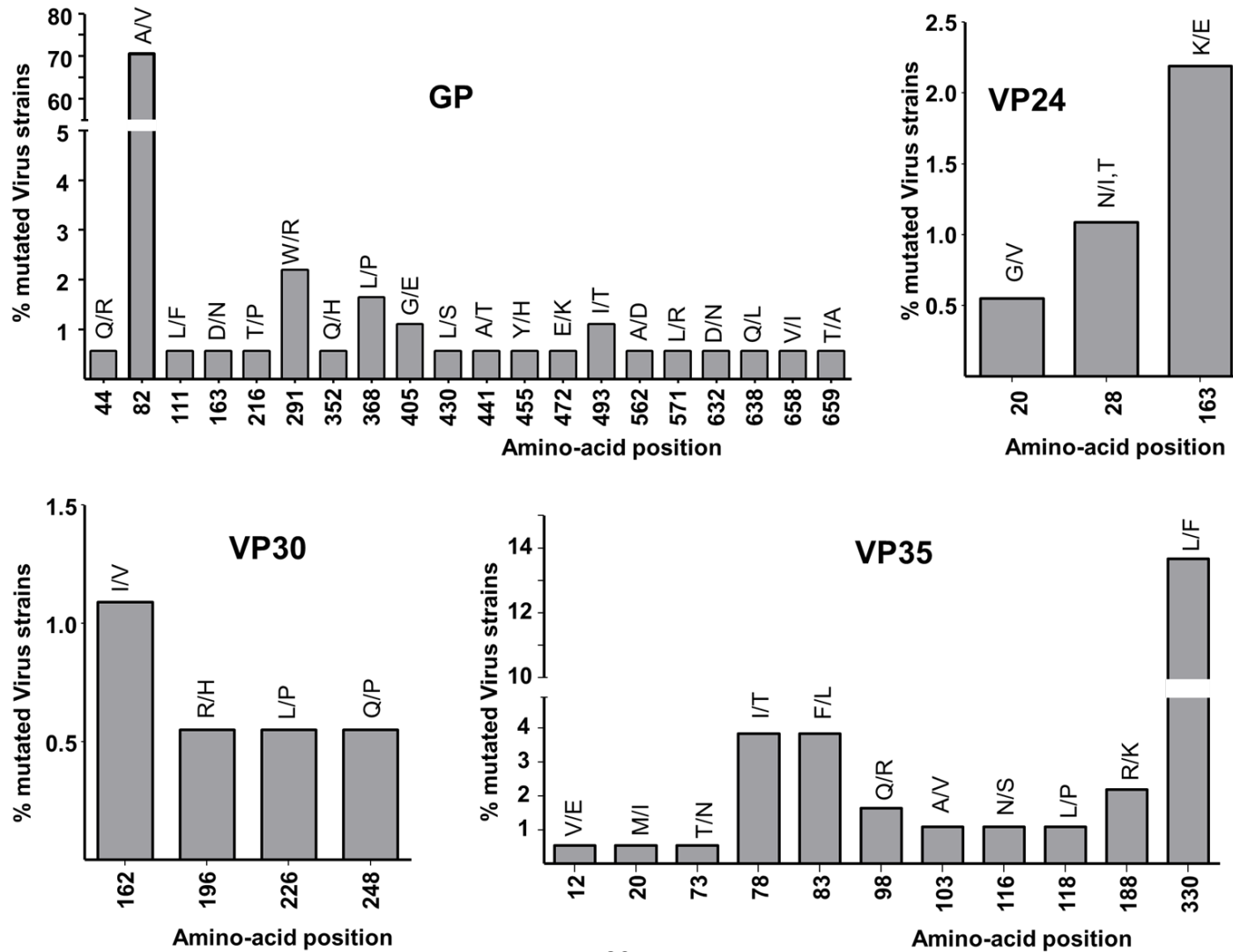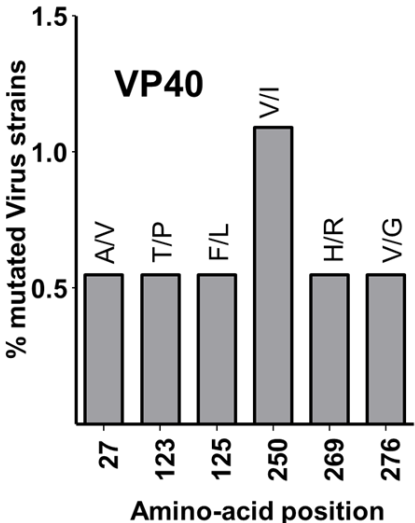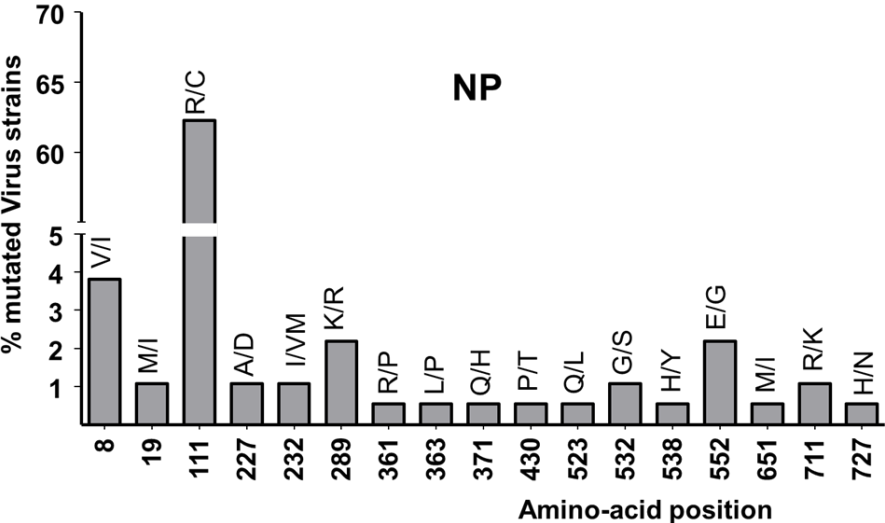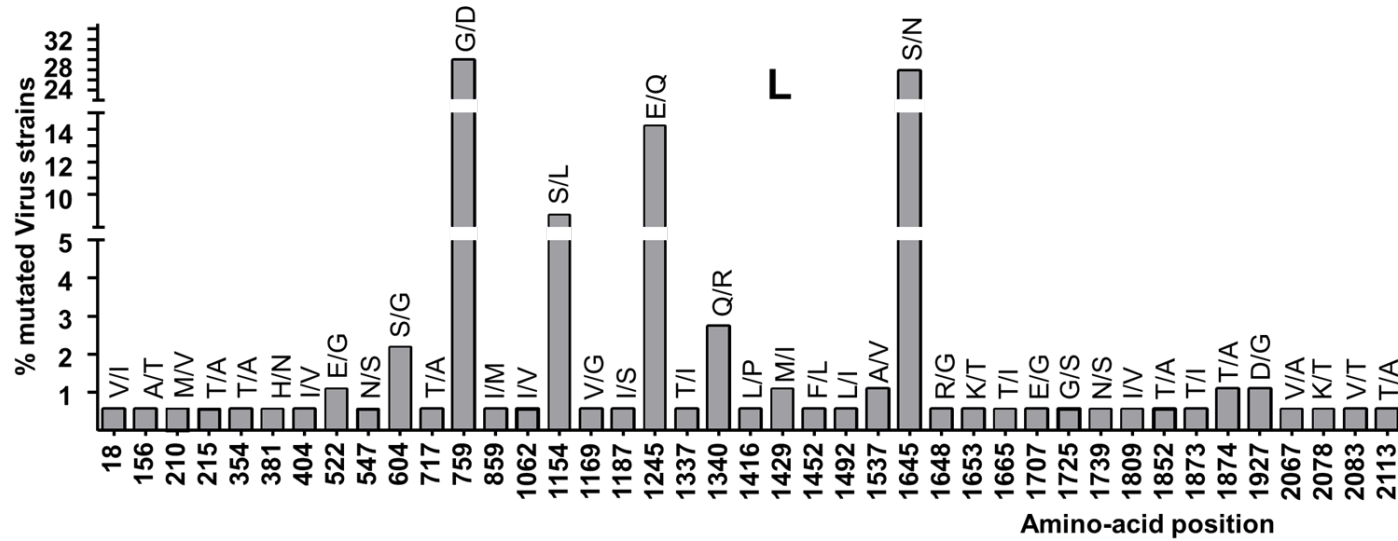
Figure 3.2. Position of non-synonymous amino acid variations in the 179 genomes analysed in this study compared to a reference sequence taken from March 2014 (KJ660346.2). *Shown is the frequency of all amino acid positions that had variability and the substitution that occurred with the first single letter position indicating the reference sequence and the second position showing the variation. The percentage frequency in the 179 genomes is shown on the y-axis.*

3.3   Discussion

As it was mention before, in this study, the potential evolution of EBOV was studied in two ways. The first one, a guinea pig model of infection was used to simulate the transmission to and adaptation of EBOV in a new host. Second, patient samples from the 2014-2015 West African outbreak were used to look for which amino acid mutations were present in the different EBOV proteins along the outbreak.

For the guinea pig model of infection, others have used this method of adapting EBOV before (Connolly et al., 1999; Volchkov et al., 2000). In these studies, mortality was first shown to occur during passages three to four (Chepurnov et al., 2003; Ryabchikova et al., 1996; Subbotina et al., 2010). Complete lethality was then detected soon after, but ranged from passage four to seven (Connolly et al., 1999; Volchkov et al., 2000; Ryabchikova et al., 1996; Subbotina et al., 2010). While 50% lethality was seen in the second passage in the current study, this was most likely due to the low titers in the passage one material requiring a higher concentration of spleen homogenate to be delivered to the guinea pigs in order to achieve challenge with 10^4 TCID$_{50}$. This amount of material would have had adverse impacts due to lipid peroxidation, and protein oxidation and pro-apoptotic factors through cellular damage during preparation of the homogenate.

In the analysis of the adaptation of EBOV to guinea pigs it was found that some viral proteins accumulated substitutions whereas other did not. An example of

this is that no substitutions were observed in either VP40 or VP30 by passage five. As its been mention before, VP40 is a viral matrix protein with multiple roles in the virus life cycle, associating with both cellular and other viral proteins including the RNP complex. It is also involved in virus assembly and release (for example, Adu-Gyamfi et al., 2013; Hoenen et al.,2005), and thus may be evolutionary constrained. Likewise, VP30 is a transcription factor and modulates interaction with NP and VP35 (for example, Biedenkopf et al., 2013) and may operate independently of the host cell for function.

In contrast, some substitutions were present in early passages but were lost by passage five. For example, in VP30 at passage two, the predominant amino acid at position 248 was an arginine (R) rather than the glutamine (Q) found in the input sequence. By passages three, four and five this was again a Q. In VP40 at passage two the predominant amino acid at position 259 was a R rather than the methionine (M) found at passages one and the input sequence. By passages three, four and five, this had become an M again. Some of these changes have been previously associated with virulence. For example, in VP24, the predominant amino acid at position 163 changes from a lysine (K) to R in passage three (and then K becomes dominant again in passages four and five), which is a conserved substitution. This substitution was described previously by Kugelman et al., who postulated that this amino acid change in VP24 might modulate interaction with other proteins instead of having an effect on structural stability (Kugelman et al., 2012).

A number of amino acid substitutions became established during adaptation

and were present in passage five (Table 3.2 and 3.3). For example, the predominant amino acid at position 26 in VP24 becomes a phenylalanine (F) in place of a leucine (L). This substitution has previously been identified as being responsible for an increase in virulence in the guinea pig model using reverse genetics (Mateo et al., 2011), and thus places confidence in the analysis of consensus sequence to detect biologically relevant variations. Using data from the stringent analysis, the frequency of amino acid substitution (amino acid changes/number of amino acid in the ORF) appeared to be of the same order of magnitude for each of the proteins that had an amino acid substitution(s) by passage five (Table 3.3). For NP this was 0.001, VP35 was 0.003, sGP was 0.006, VP24 was 0.004 and L was 0.005. However, there were no coding mutations in VP40 and VP30. In this study, it was observed 11 coding changes by passage five in the L protein, a study investigating adaptions of EBOV to a mouse model highlighted two coding changes and one silent change. This may reflect a difference in adaption of the virus to the two hosts.

As noted by Ebihara et al., additional mutations in the L protein are likely to contribute to virulence by affecting viral RNA synthesis (Ebihara et al., 2006). These may also mediate both viral and host cell interactions. A similar situation has been described for influenza A virus where a cellular protein is critical for polymerase activity and transmission between different species and interaction of this protein with the polymerase is determined by a single amino acid substitution (Hudjetz et al., 2012; Gabriel et al., 2011). However, with that said, data indicate that VP24 associates with a number of different cellular

proteins that may be critical for its function inside virus infected cells (for example, Garcia-Dorival et al., 2014; Xu et al., 2014; Reid et al., 2006). It is interesting to speculate that the adaptation of the L protein to the new host may be correlated with the increase in RNA editing activity to transcribe the $GP_{1,2}$ mRNA. This editing activity of the polymerase complex maybe associated with and mediated by a host cell factor, although both cis-acting factors on the EBOV genome and VP30 have been implicated in this process (Mehedi et al., 2013).

In this study, a virus (EBOV) that was not initially virulent in guinea pigs was serially passaged and became more pathogenic in its new host. Several alterations in the amino acid coding sequence were associated with this increase in virulence. A reverse genetics approach would precisely characterize these changes and their linkage with pathogenicity and virulence. Interestingly, this approach was used in a mouse model study of EBOV to investigate the molecular determinants of virulence (Edibhara et al., 2006). There are several implications of our research for the biology of EBOV and associated outbreaks. EBOV causes a zoonotic infection (Feldmann et al., 2013; Groseth et al., 2007), and humans have been considered a dead end host, with the long-term survival of EBOV in nature likely being dependent on its ability to persist in its natural host. Our data suggest that the selection pressures at the initial stages of replication in a new host are different from those when the virus becomes established, and this may be dependent on population size, density and route of transmission. Therefore, as EBOV is so pathogenic in humans, one possibility for a sustained human-to-human

transmission scenario might be selection of variants that are less pathogenic and that could lead to a more long-term infection of the population, thus allowing EBOV to persist. However, balanced with this are the social aspects of infection. For example, where humans tend to gather to grieve, suggesting that a reduction in pathogenicity is not the only way the virus can be maintained long term in the human population. Our study can be considered a model for the initial jump of EBOV from a reservoir into a new host, where selection pressure may be at its highest.

For the analysis done during the current outbreak in West Africa, amino acid changes were found in some of the viral proteins. Some of these changes are really low but it doesn't mean it may not be important for the virus evolution and adaptation. It has been already describe before that a single amino acid change in one single protein (VP24) is necessary to become lethal and increasing the pathogenesis in an animal model (Mateo et al 2011); furthermore several amino acid mutation were found in the glycoprotein and the RNA polymerase of the EBOV and those proteins are known to be important and crucial for the virus adaptation to a new host.

While some of the variations in the amino acids may be attributed to a purely random molecular clock pattern, in GP, VP35 and L there are some amino acid variations that are present in over 15% of the genome samples. One example is in the GP where there is an alanine (A) to valine (V) substitution in 70.5% of the genomes sampled compared to the reference genome. Implications of the mutations within GP in relation to immune escape of therapeutics and vaccines

will need to be assessed in pseudo type neutralization assay using EBOV monoclonal antibodies and serum from people who have been vaccinated.

When the amino changes found in this study were compared to those amino acid changes found in the forced evolution model (guinea pig) only one position in the protein VP30 (amino acid position 248) was found similar for both studies. However, this mutation found in the guinea pig model was not persistent though all passages and therefore only present in passage 4. Furthermore, the amino acid change itself, even when it was found in the same position, was a different amino acid.

The data presented in this study, using a forced evolutionary and transmission model, would suggest that the initial evolutionary trajectory of EBOV in a new host leads to a gain in virulence. Given the circumstances of the sustained transmission of EBOV in the current outbreak in West Africa, increases in virulence maybe associated with prolonged and uncontrolled epidemics of EBOV.

CHAPTER 4: Elucidation of the Ebola virus VP24 cellular interactome and disruption of virus biology through targeted inhibition of host cell protein function

4. 1   Introduction

Viral pathogenesis in the infected cell is a balance between antiviral responses and subversion of host-cell processes. Many viral proteins specifically interact with host cell proteins to promote virus infection. Understanding these interactions can lead to knowledge gains about infection and provide potential targets for anti-viral therapy. One such virus is EBOV, which has profound consequences for human health and causes viral haemorrhagic fever where case fatality rates can approach 90% (Feldmann et al., 2013). EBOV is an enveloped non-segmented negative sense single stranded RNA virus with a genome of 19kb in length consisting of seven genes (NP, VP35, VP40, GP, VP30, VP24 and L) encoding nine proteins (Feldmann et al., 2013; Feldmann & Geisbert., 2011). The combination and action of these gene products and their interactions with the host cell cause the severe haemorrhagic fever.

One of the proteins thought to contribute to disease pathology is VP24 and this may be a major factor in virulence (Mateo et al., 2011A). This protein is considered a secondary matrix protein as only a small amount of VP24 is incorporated into viral particles. However, VP24 has affinity for the plasma membrane and is associated with the envelope of the virion (Han et al., 2003; Eliot et al., 1985). VP24 has been shown to have several functions, including an involvement in viral nucleocapsid formation and the regulation of replication (Mateo et al., 2011B; Watanabe et al., 2007).

Critically, VP24 is also involved in modulation of the host response to infection through evasion of the host immune response. VP24 may disrupt interferon

signalling by binding to KNP-α (Reid et al., 2007) which blocks the nuclear accumulation of tyrosine-phosphorylated STAT1 (PY-STAT) (Reid et al., 2006). STAT1 is a major signalling molecule in initiating the anti-viral response (Randall et al., 2008). Nuclear translocation of STAT1 is essential for transcriptional activation of numerous interferon responsive genes. VP24 can also inhibit IFN-α/β and IFN-γ-induced gene expression (Mateo et al., 2010). The mechanism of action through which VP24 modulates the immune response is unknown. However, other work suggested that VP24 might be structurally similar to importin-α/β and exportin (Lee et al., 2009). Thus, VP24 may be involved in the mimicry of host transporter/cargo interactions (Lee et al., 2009; Ramanan et al., 2011).

Mutations in VP24 have been linked to the adaptation of EBOV in mice and guinea pigs to produce lethal disease (Mateo et al., 2011A; Volchkov et al., 2000; Ebihara et al., 2006). Therefore, VP24 may form critical protein-protein interactions with host cell proteins to modulate host cell pathways; this is the case of the interaction of VP24 with KPNA, which is required for the inhibition of the interferon-signalling pathway. Therefore, I hypothesized that the elucidation of these interactions in detail would provide a deeper understanding of EBOV infection and also potentially present anti-viral chemo-therapeutic targets to disrupt virus biology. The study of the EBOV VP24 interactome will provide further insights into the virus biology, which in turn may prove useful to discover new potential therapeutic drug targets against Ebola virus.

To determine which cellular proteins or complexes interact with VP24 and to predict function, high-affinity purification coupled to a label-free mass-spectrometry-based approach was used. Using conservative selection criteria approximately 50 cellular proteins were identified that had a high probability of interacting with VP24. Some of the candidate proteins had been identified in previous analyses, confirming that the approach was able to identify known interactions, thus increasing confidence that the analysis was capable of identifying new cellular proteins in the same experiment. One of the novel hits identified was sodium/potassium-transporting ATPase subunit alpha-1 (ATP1A1). This enzyme can be inhibited by the small molecule inhibitor ouabain, which is used as therapeutic agents to increase the force of cardiac contractions in the treatment of heart disease (Fürstenwerth, H. 2010; Vemuri et al 1989). Here ouabain was used to see the effect of the inhibition of the function of ATP1A1 in cells infected with EBOV. The data indicated that progeny virus decreased in treated infected cells compared to untreated infected cells.

Data from this chapter has been published or contributed towards the following publication:

García-Dorival, I., Wu, W., Dowall, S., Armstrong, S., Touzelet, O., Wastling, J., ... & Hiscox, J. A. (2014). **Elucidation of the Ebola virus VP24 cellular interactome and disruption of virus biology through targeted inhibition of host-cell protein function**. *Journal of Proteome Research*, *13*(11), 5120-5135.

4.2   Results

4.2.1 Expression of VP24 in 293T cells

To obtain a more complete picture of the interaction partners of *Ebolavirus* VP24 quantitative proteomics coupled to an immunoprecipitation strategy was done based on expressing VP24 as an EGFP-fusion in human 293Tcells and utilizing a GFP-trap to selectively precipitate the fusion protein and interacting partners (*Figure 4.1A*). 293-T cells are derived from embryonic kidney cells and were chosen due to their ability to sustain EBOV replication, their high transfection efficiencies using calcium-phosphate, and the well-annotated human databases that aid with protein identification and function assignment.

To identify interacting cellular partners specific for VP24 label-free LC-MS/MS approach was used in order to analyse binding partners from cells expressing either EGFP or the EGFP-VP24 or VP24-EGFP fusion proteins. Subsequent comparison of the immunoprecipitated interacting partners from these conditions allowed the identification of cellular components that specifically bound to the VP24 moiety within the EGFP/VP24 fusion protein. This general approach has been shown to improve sensitivity and allow discrimination of specific from nonspecific interactions with the target protein (Boulon et al, 2010; Trinkle-Mulcahy et al., 2008) and can be readily applied to the analysis of the cellular interactomes of viral proteins (Wu et al., 2012; Emmott et al., 2013; Joudan et al., 2012).

To identify cellular proteins that interact with VP24 a gene encoding EGFP was cloned either 5′ or 3′ of a gene encoding codon-optimized VP24 from the EBOV (Mayinga Strain 1976) or RESTV. Such that, when the construct was transfected in cells the EGFP moiety was expressed as either an N- or C-terminal fusion of VP24, respectively. In this way, we mitigated against potential nonbinding to cellular targets caused by steric hindrance of the EGFP moiety. Proteins were precipitated using the anti-EGFP antibody, and both input (cell lysate) and eluted fraction were compared using immunofluoresence (*Figure 4.1B*) and Western blot (*Figure 4.1C*). The data indicated that both EBOV/VP24 (as a EGFP-EBOV/VP24 and EBOV/VP24-EGFP) and RESTV/VP24 (as a EGFP-RESTV/VP24 and RESTV/VP24-EGFP) were expressed with high efficiency at the expected molecular weight (~55 kDa).

A

Calcium phosphate transfection of the
three constructs in 293T cells

Whole cell lysate

GFP-Trap pull-down

Protein identification and quantification by
Label free Mass Spectrometry

B

C
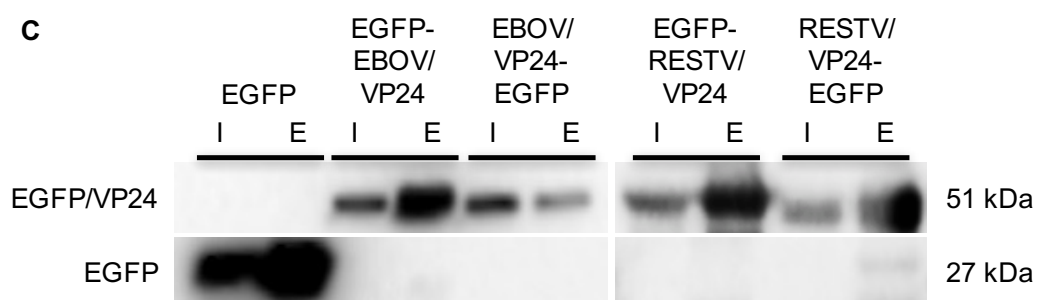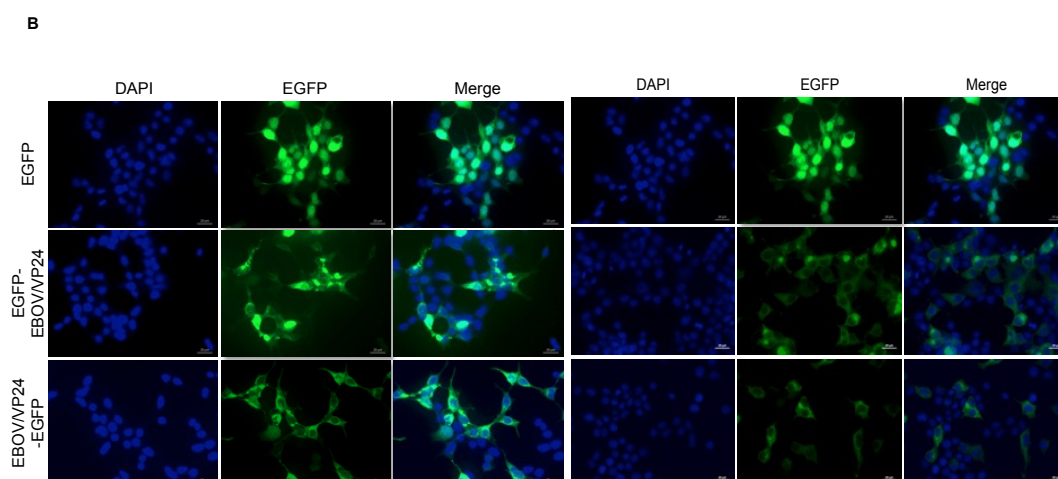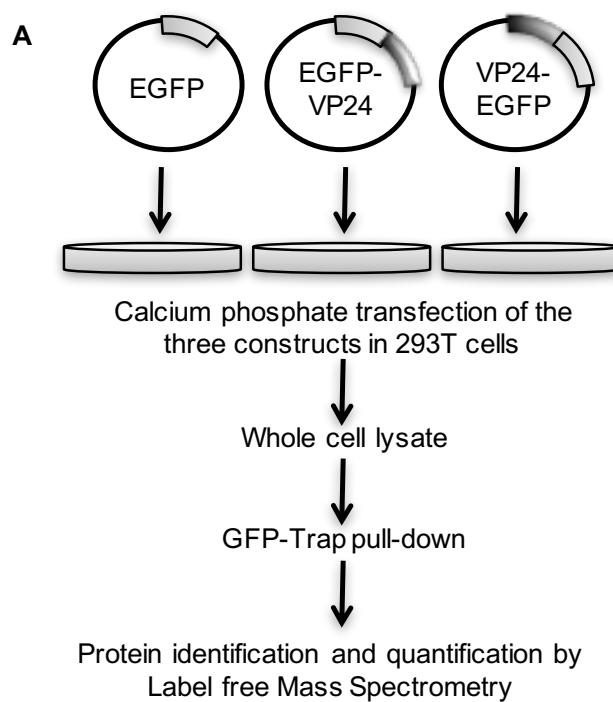
Figure 4.1. Expression of VP24 in 293T cells. (A) *Schematic representation of the methodology used in this study. HEK 293T cells were grown in DMEM supplemented with 10% FBS and 1% penicillin-streptomycin at 37°C with 5% CO_2. Two 145cm^2 dishes were seeded with $4x10^6$ cells 24 hours prior to calcium phosphate transfection with 25.6 µg of plasmid DNA encoding EGFP, EGFP-EBOV/VP24, EBOV/VP24-EGFP, EGFP-RESTV/VP24 and RESTV/VP24-EGFP respectively. 24 hours post transfection the cells were harvested, lysed and immunoprecipitated using a GFP-Trap (Chromotek). Label free Mass Spectrometry analysis on the eluted samples was then carried out.* ***(B)*** *Expression of EGFP-EBOV/VP24, EBOV/VP24-EGFP, EGFP-RESTV/VP24, RESTV/VP24-EGFP and the control EGFP in 293T cells was confirmed by immunofluorescence using confocal microscopy; the panels show that expression of the four constructs were similar, however EGFP-VP24 show a higher expression than VP24-EGFP in both strains.* ***(C)*** *Analysis of the pull down products using a western blot confirm the presence of EGFP-VP24 and VP24-EGFP (for both strain, EBOV and RESTV) and the control EGFP with the expected molecular weight, confirming the expression of these proteins in 293T cells*

4.2.2 Identification of the potential cellular interacting partners of EBOV VP24

EGFP-EBOV/VP24, EBOV/VP24-EGFP and EGFP were then overexpressed in 293T cells, and the cellular binding partners were immunoprecipitated using the GFP-trap. These proteins were identified by mass spectrometry. To differentiate those cellular proteins that formed specific interactions with EBOV/VP24 versus those proteins that interacted with EGFP or formed nonspecific interactions with the binding matrix, we repeated the immune precipitations and mass spectrometry five times. Approximately 600 cellular proteins were initially identified and quantified, which represented both the specific and nonspecific interactions. To differentiate between these possibilities, we took several conservative processing steps. The 15 data sets were analysed by the Perseus software; proteins identified by a single peptide were removed, and a p value was set at <0.01 for the *t*-test analysis, where theoretically only 1:100 proteins were misidentified. The data was then organized in the form of volcano plots for EGFP-EBOV/VP24 and EBOV/VP24-EGFP (*Figure 4.2 A, B*). These show P values ($-\log_{10}$) for confidence in peptide identification versus fold difference in binding of a protein between the EBOV/VP24 fusion protein and EGFP only; dots represent individual proteins. Those dots inside the volcano plots represent proteins that did not associate with EBOV/VP24 with statistical significance. Proteins that had a binding ratio greater than 2 and were statistically significant are shown in the right-hand quadrant. For EGFP-EBOV/VP24 and EBOV/VP24-EGFP, 48 and 51 proteins (*Table 4.1*) respectively, were identified as interacting with

statistical significance. Both fusion proteins had 40 proteins identified in common; 8 proteins were unique for EGFP-EBOV/VP24 and 11 proteins were unique for EBOV/VP24-EGFP. A map of this protein–protein interactions were performed using String program (version 9.05) and IPA (QIAGEN) and are presented in *Figure 4.3*.

Cellular proteins that had a higher statistical likelihood of associating with EBOV/VP24 included KNPA1 and KNPA6, LMNB1, VDAC-1 and ATP1A1. KNPA6 has been previously shown to interact with VP24 (Reid et al., 2006), and therefore indicated that our approach identified known interactions and increased confidence in the analysis of cellular proteins found in the same experiment. Comparison of proteins identified as interacting with VP24 and their abundance in an average human cell (data taken from PaxDb: Protein Abundance Across Organisms database) (*Table 4.1*) indicated that the likelihood of binding was not proportional to their abundance in a cell, thus increasing confidence that the amount of binding was not merely a reflection of overall protein abundance.

A EGFP-EBOV/VP24

-Log$_{10}$ P-value

Log$_2$ Fold change difference (EGFP vs. EGFP-EBOV/VP24)

ATP1A1
KPNA6
KPNA1
LMNB1
VDAC2
RAB11B/
RAB11A
VDAC1

B EBOV/VP24-EGFP

-Log$_{10}$ P-value

Log$_2$ Fold change difference (EGFP vs. EBOV/VP24-EGFP)

KPNA6
ATP1A1 KPNA1
LMNB1
RAB11B/
RAB11A
VDAC1
VDAC2
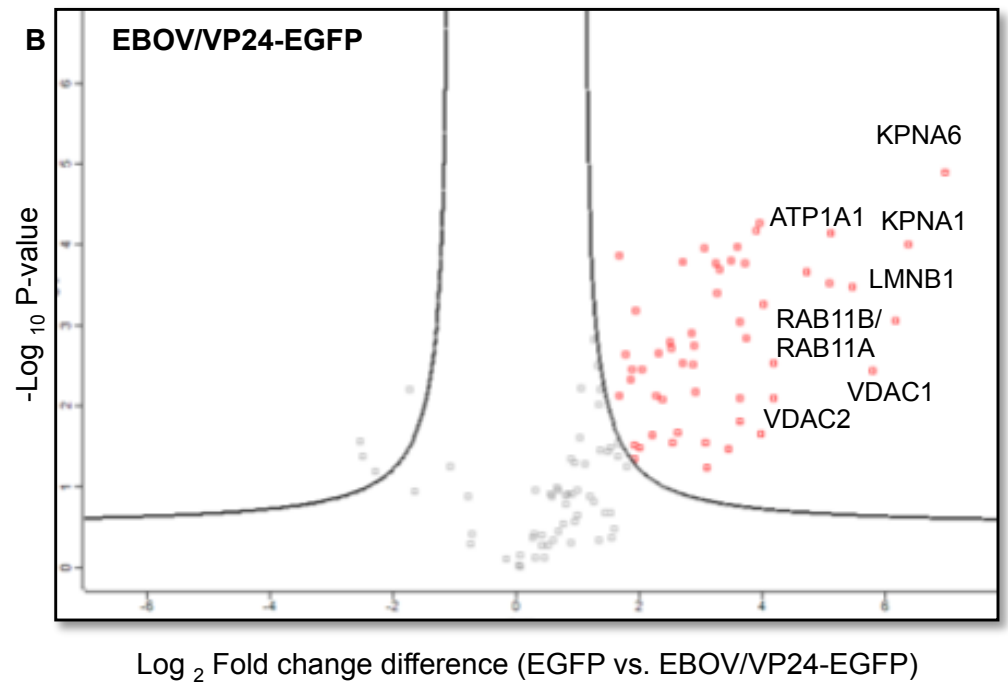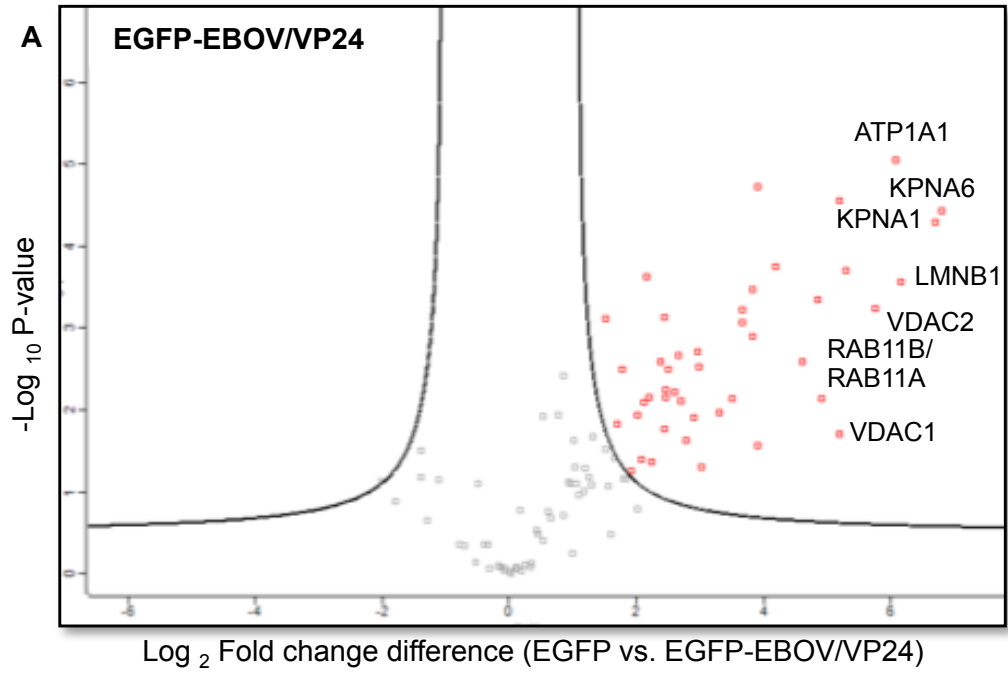
Figure 4.2. Volcano plot representing results of the label free pull-down of the complex EGFP-EBOV/VP24 (A) or EBOV/VP24-EGFP (B) versus the control EGFP performed in triplicate. *For any potential protein interaction partner with EBOV/VP24, the value of its abundance co-immunoprecipitated with EGFP-EBOV/VP24 or EBOV/VP24-EGFP was compared to the value of co-immunoprecipitated with EGFP alone. Logarithmic ratio of protein intensities is plotted against the negative logarithmic p values of the t-test. The dashed curve indicates the region of significant interactions with a false discovery rate of 1% and the red dots in the upper right corners are the proteins that have the highest probability of interact with EBOV/VP24. This analysis indicated that the proteins, that may potentially interact with the EBOV/VP24 protein are: KPNA6, LMNB1. These latter proteins are components of the nuclear lamina.*

A.
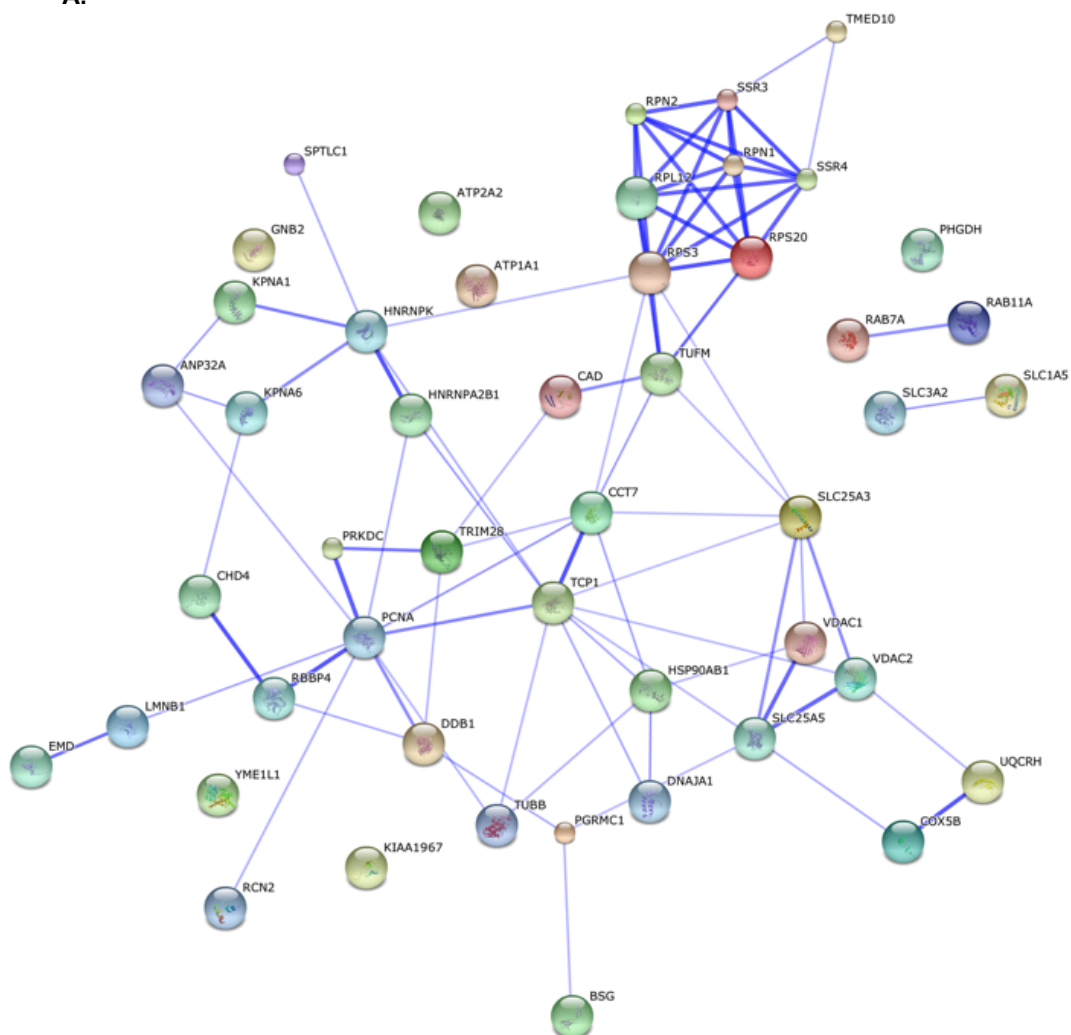
B.
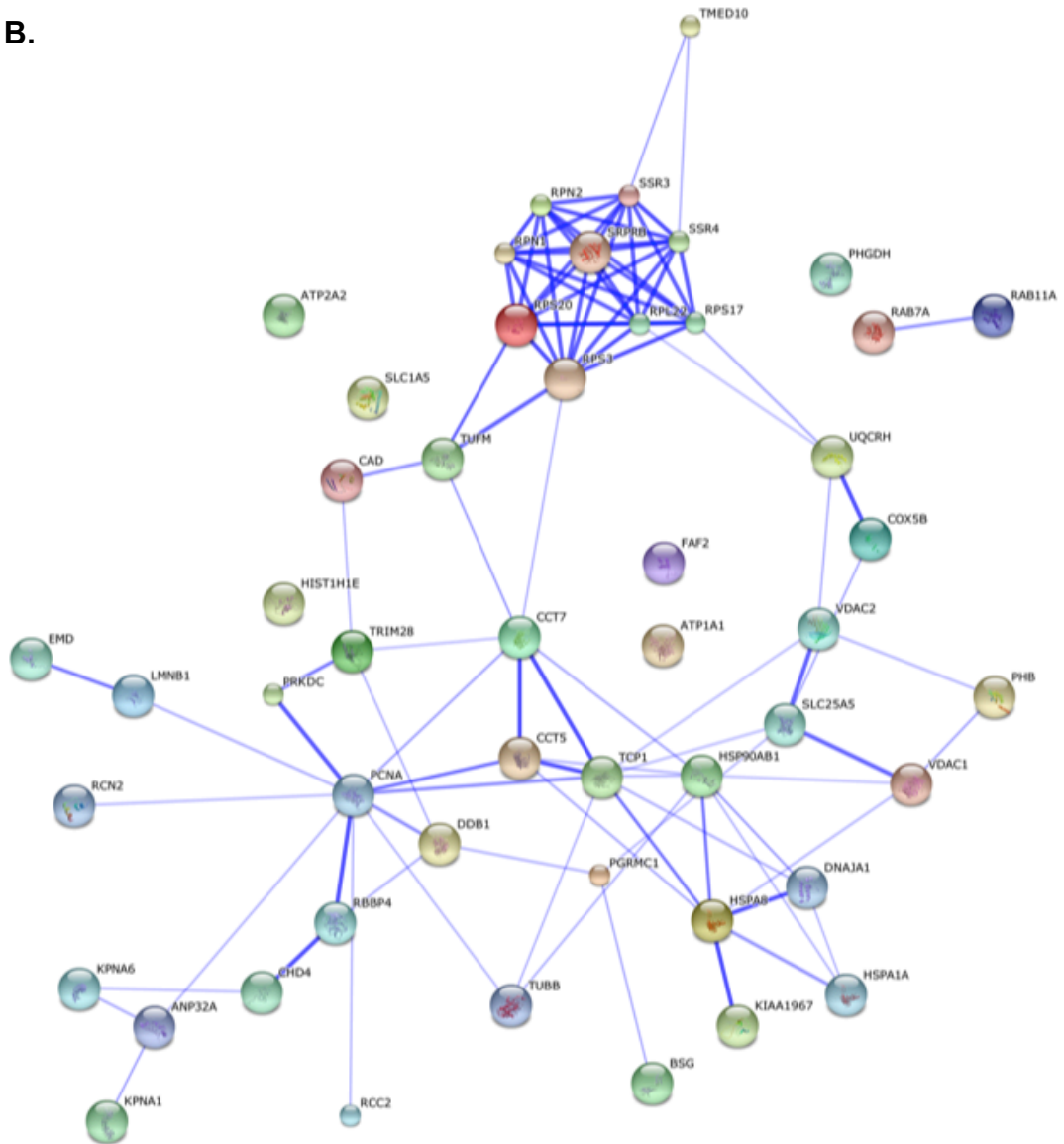
Figure 4.3. *Map showing the group of proteins that interact with VP24: (A) EGFP-VP24 and (B) VP24-EGFP, using String 9.1 software. Most of the proteins detected were located in the membranes and in the perinuclear region, as is VP24. In the protein analysis three main proteins groups were detected; ribosome proteins, mitochondrial proteins and proteins related to the nucleus.*

Table 4.1: *Cellular proteins that have a higher probability of forming protein-protein interactions with EBOV/VP24. Shown are candidate proteins identified using the EGFP-EBOV/VP24 (A) and EBOV/VP24-EGFP (B) fusion proteins and identified using label free quantitative proteomics. Protein identifier, protein name and gene names are indicated Total and unique peptides used to identify the protein are indicated. The –LogP value is a comparison of the cellular protein between the EBOV/VP24 pull down and EGFP, where the higher the number means the higher probability of interacting, and a threshold above 2.0 has been selected. The t-test Difference is the difference of the means of the intensities of the cellular proteins in the EBOV/VP24 pull down and EGFP. PEP is posterior error probability and percentage of sequence coverage of the protein identified using the peptides is indicated SC%. The abundance (ppm) of the protein in an average human cell is listed – data taken from the PaxDb: Protein Abundance Across Organisms database.*

TABLE 4.1 A: *Significant proteins that interact with EGFP-EBOV/VP24.*

	Protein identifier	Protein names	Gene name	Peptides	Unique peptides	-Log P value	t-test Difference (EGFP-VP24 vs. EGFP)	PEP	SC(%)	Abundance (ppm)
1	P20700; E9PBF6	Lamin-B1	LMNB1	40	36	7.1	6.7	0	61.3	105 ppm (top 5%)
2	O60684; F5GYL8	Importin subunit alpha-7	KPNA6	16	12	6.5	6.1	4.89E-243	40.5	4.51 ppm
3	P52294; C9JYI4	Importin subunit alpha-5	KPNA1	13	10	6.7	6.0	7.9E-159	29.2	5.90 ppm
4	Q15907; P62491	Ras-related protein Rab-11B;Ras-related protein Rab-11A	RAB11B;RAB11A	6	6	7.1	5.8	1.44E-86	22.9	42.0 ppm (top 25%)
5	F5H3A1; P05023	Sodium/potassium-transporting ATPase subunit alpha-1	ATP1A1	23	21	7.7	5.6	0	28	46.8 ppm (top 25%)

6	P39687; F2Z3H3	Acidic leucine-rich nuclear phosphoprotein 32 family member A	ANP32A	6	5	5.8	5.6	1.72E-67	27.7	305 ppm (top 5%)
7	P16615; Q93084	Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	ATP2A2	15	15	5.3	5.1	4.61E-133	20.2	30.1 ppm(top 25%)
8	P21796; C9JI87	Voltage-dependent anion-selective channel protein 1	VDAC1	10	10	2.9	4.9	2.03E-168	44.2	370 ppm (top 5%)
9	P35613; A6NJW1	Basigin	BSG	7	7	5.4	4.9	1.81E-124	21.8	126 ppm (top 10%)
10	F8WCY5 ;Q14257	Reticulocalbin-2	RCN2	7	7	7.1	4.5	2.98E-83	26.3	2.74 ppm
11	Q09028; B4DRT0	Histone-binding protein RBBP4	RBBP4	8	5	3.3	4.3	4.44E-82	22.8	67.4 (top 25%)
12	P45880; B4DKM5	Voltage-dependent anion-selective channel protein 2	VDAC2	9	9	3.6	4.2	1.2E-78	54.4	341 ppm (top 5%)
13	P51571; A6NLM8	Translocon-associated protein subunit delta	SSR4	7	7	5.9	4.0	6E-63	41	37.5 ppm (top 25%)
14	P07919	Cytochrome b-c1 complex subunit 6, mitochondrial	UQCRH	3	3	4.1	3.8	2.96E-73	29.7	200 ppm (top 10%)
15	O00264; B7Z1L3	Membrane-associated progesterone receptor component 1	PGRMC1	8	7	5.3	3.8	2.33E-60	19	78.5 ppm (top 25%)
16	B4DW28; P60866	40S ribosomal protein S20	RPS20	4	4	2.9	3.7	6.92E-35	28.2	126 ppm (top 25%)
17	P50402; Q5HY57	Emerin	EMD	5	5	7.3	3.7	4.75E-24	24.4	8.00 ppm
18	P12004	Proliferating cell nuclear antigen	PCNA	7	7	3.6	3.7	5.42E-51	34.5	182 ppm (top 10%)
19	P05141; P12235	ADP/ATP translocase 2	SLC25A5	13	8	2.7	3.5	1.18E-175	29.2	213 ppm (top 5%)
20	P04843; F5H6I5	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1	RPN1	7	7	4.5	3.4	1.16E-66	13.5	57.7 ppm (top 10%)

21	P51149; C9J8S3	Ras-related protein Rab-7a	RAB7A	6	6	2.8	3.4	6.14E-126	30.4	51.6 ppm (top 10%)
22	P04844; Q5JYR6	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 2/Ribophorin II	RPN2	5	5	4.5	3.4	1.88E-60	11.9	97.7 ppm (top 10%)
23	O43175; Q5SZU1	D-3-phosphoglycerate dehydrogenase	PHGDH	15	15	2.4	3.1	8.01E-223	36.4	84.6 (top 10%)
24	P31689; F5GZ88	DnaJ homolog subfamily A member 1	DNAJA1	6	6	3.6	3.0	9.26E-57	24.2	54.2 ppm (top 25%)
25	P23396; E9PL09	40S ribosomal protein S3	RPS3	8	8	2.4	2.8	1.45E-104	30	136 ppm (top 5%)
26	Q13263	Transcription intermediary factor 1-beta	TRIM28	6	6	3.9	2.8	3.89E-43	10.4	56.5 ppm (top 10%)
27	B4E2P2; Q9UNL2	Translocon-associated protein subunit gamma	SSR3	3	3	1.9	2.7	1.14E-62	14.6	10.1 ppm
28	P10606	Cytochrome c oxidase subunit 5B, mitochondrial	COX5B	4	4	3.8	2.7	2.63E-08	30.2	14.4 ppm
29	F8VVM2; F8VWQ0	Phosphate carrier protein, mitochondrial	SLC25A3	4	2	1.9	2.6	1.48E-44	14.2	40.8 ppm (top 25%)
30	F5H4J7; P08195;	4F2 cell-surface antigen heavy chain	SLC3A2	4	4	3.7	2.6	6.55E-52	9.8	45.3 (top 25%)
31	P49755	Transmembrane emp24 domain-containing protein 10	TMED10	5	5	2.4	2.5	6.21E-45	22.4	37.2 ppm (top 25%)
32	P27708; F8VPD4	CAD protein; Glutamine-dependent carbamoyl-phosphate synthase; Aspartate carbamoyltransferase; Dihydroorotase	CAD	6	6	3.8	2.5	4.72E-47	3.2	10.7 ppm
33	Q14839; F5GWX5	Chromodomain-helicase-DNA-binding protein 4	CHD4	8	8	2.4	2.4	1.09E-56	6.7	7.83 ppm
34	P49411	Elongation factor Tu, mitochondrial	TUFM	11	11	1.8	2.3	2.62E-114	28.3	45.3 ppm (top 10%)
35	P78527; E7EUY0	DNA-dependent protein kinase catalytic subunit	PRKDC	10	10	3.4	2.3	1.83E-101	3.8	34.9 ppm (top 25)

36	P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	HNRNPA2B1	6	6	3.5	2.2	4.16E-27	22.9	347 ppm (top 5%)
37	Q16531; E7EPB0	DNA damage-binding protein 1	DDB1	6	6	2.8	2.2	3.02E-107	8.9	22.7 ppm (top 25%)
38	P17987; E7ERF2	T-complex protein 1 subunit alpha	TCP1	13	13	4.7	2.1	7.32E-171	34.9	334 ppm (top 5%)
39	P08238; Q58FF7	Heat shock protein HSP 90-beta	HSP90AB1	14	9	2.5	2.1	2.46E-115	18.1	713 ppm (top 5%)
40	P30050	60S ribosomal protein L12	RPL12	5	5	1.9	2.1	1.62E-81	49.7	425 (top 5%)
41	Q15758; E9PC01	Neutral amino acid transporter B	SLC1A5	3	3	4.1	2.1	6.72E-34	8.3	21.6 ppm (top 25%)
42	Q96TA2; B4DNM1	ATP-dependent zinc metalloprotease YME1L1	YME1L1	4	4	3.5	2.0	2.44E-50	10.1	14.6 ppm (top 25%)
43	P62879; C9JXA5	Guanine nucleotide-binding protein G (I)/G (S)/G (T) subunit beta-2	GNB2	3	3	3.0	1.9	9.24E-17	11.2	121 ppm (Top 5%)
44	O15269; Q96IX6	Serine palmitoyltransferase 1	SPTLC1	5	5	3.8	1.9	3.87E-91	14.8	83.1 ppm (top 10%)
45	P07437; F8VW92	Tubulin beta chain	TUBB	21	4	3.8	1.8	0	57.4	223 ppm (top 5%)
46	Q8N163; E5RHJ4	Cell cycle and apoptosis regulator protein 2, DBIRD complex subunit KIAA1967	KIAA1967	6	6	3.1	1.8	6.22E-92	10.4	150 ppm (top 5%)
47	Q99832; B8ZZC9	T-complex protein 1 subunit eta	CCT7	3	3	2.3	1.7	6.95E-14	8.1	116 ppm (top 5%)
48	P61978; Q5T6W5	Heterogeneous nuclear ribonucleoprotein K	HNRNPK	7	7	3.9	1.7	1.48E-60	21.2	511 ppm (top 5%)

	Protein identifier	Protein names	Gene names	Peptides	Unique peptides	-Log P value	t-test Difference VP24-EGFP vs. EGFP)	PEP	SC(%)	Abundance (ppm)
		TABLE 4.1 B: *Significant proteins that interact with EBOV/VP24-EGFP*								
1	P20700; E9PBF6	Lamin-B1	LMNB1	40	36	5.9	6.8	0	61.3	105 ppm (top 5%)
2	O60684; F5GYL8	Importin subunit alpha-7;Importin subunit alpha	KPNA6	16	12	8.6	6.2	4.89E-243	40.5	4.51 ppm
3	P52294; C9JYI4	Importin subunit alpha-1	KPNA1	13	10	6.3	5.9	7.90E-159	29.2	5.90 ppm
4	Q15907; P62491	Ras-related protein Rab-11B;Ras-related protein Rab-11A	RAB11B;RAB11A	6	6	6.7	5.4	1.44E-86	22.9	42.0 ppm (top 25%)
5	P35613; A6NJW1;	Basigin	BSG	7	7	5.2	5.3	1.81E-124	21.8	126 ppm (top 10%)
6	F5H3A1; P05023;	Sodium/potassium-transporting ATPase subunit alpha-1	ATP1A1	23	21	9.9	5.3	0	28	46.8 ppm (top 25%)
7	Q09028; B4DRT0	Histone-binding protein RBBP4	RBBP4	8	5	3.2	4.8	4.44E-82	22.8	67.4 (top 25%)
8	P39687; F2Z3H3	Acidic leucine-rich nuclear phosphoprotein 32 family member A	ANP32A	6	5	2.4	4.2	1.72E-67	27.7	305 ppm (top 5%)
9	F8WCY5; Q14257	Reticulocalbin-2	RCN2	7	7	7.9	4.1	2.98E-83	26.3	2.74 ppm
10	B4DW28; P60866;	40S ribosomal protein S20	RPS20	4	4	3.3	3.9	6.92E-35	28.2	126 ppm (top 25%)
11	P05141; P12235	ADP/ATP translocase 2	SLC25A5	13	8	2.6	3.9	1.18E-175	29.2	213 ppm (top 5%)
12	P51571; A6NLM8	Translocon-associated protein subunit delta	SSR4	7	7	8.6	3.8	6.00E-63	41	37.5 ppm (top 25%)

13	P07919	Cytochrome b-c1 complex subunit 6, mitochondrial	UQCRH	3	3	3.9	3.8	2.96E-73	29.7	200 ppm (top 10%)
14	P12004	Proliferating cell nuclear antigen	PCNA	7	7	3.4	3.8	5.42E-51	34.5	182 ppm (top 10%)
15	P10606	Cytochrome c oxidase subunit 5B, mitochondrial	COX5B	4	4	4.0	3.7	2.63E-08	30.2	14.4 ppm
16	P50402; Q5HY57	Emerin	EMD	5	5	5.9	3.6	4.75E-24	24.4	8.00 ppm
17	P21796; C9JI87	Voltage-dependent anion-selective channel protein 1	VDAC1	10	10	3.5	3.5	2.03E-168	44.2	370 ppm (top 5%)
18	P16615; Q93084	Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	ATP2A2	15	15	5.5	3.5	4.61E-133	20.2	30.1 ppm(top 25%)
19	P45880; B4DKM5	Voltage-dependent anion-selective channel protein 2	VDAC2	9	9	3.5	3.3	1.20E-78	54.4	341 ppm (top 5%)
20	O43175; Q5SZU1	D-3-phosphoglycerate dehydrogenase	PHGDH	15	15	2.9	3.2	8.01E-223	36.4	84.6 (top 10%)
21	B4E2P2; Q9UNL2	Translocon-associated protein subunit gamma	SSR3	3	3	4.2	3.2	1.14E-62	14.6	10.1 ppm
22	P49411	Elongation factor Tu, mitochondrial	TUFM	11	11	2.9	3.2	2.62E-114	28.3	45.3 ppm (top 10%)
23	P04843; F5H6I5	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1	RPN1	7	7	5.7	3.2	1.16E-66	13.5	57.7 ppm (top 10%)
24	P23396; E9PL09	40S ribosomal protein S3	RPS3	8	8	3.2	3.1	1.45E-104	30	136 ppm (top 5%)
25	P04844; Q5JYR6	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 2/Ribophorin II	RPN2	5	5	4.5	3.4	1.88E-60	11.9	97.7 ppm (top 10%)

26	O00264; B7Z1L3	Membrane-associated progesterone receptor component 1	PGRMC1	8	7	6.4	2.9	2.33E-60	19	78.5 ppm (top 25%)
27	P31689; F5GZ88	DnaJ homolog subfamily A member 1	DNAJA1	6	6	1.9	2.8	9.26E-57	24.2	54.2 ppm (top 25%)
28	P49755	Transmembrane emp24 domain-containing protein 10	TMED10	5	5	3.7	2.7	6.21E-45	22.4	37.2 ppm (top 25%)
29	P51149; C9J8S3	Ras-related protein Rab-7a	RAB7A	6	6	4.8	2.6	6.14E-126	30.4	51.6 ppm (top 10%)
30	Q8N163; E5RHJ4	DBIRD complex subunit KIAA1967	KIAA1967	6	6	4.4	2.6	6.22E-92	10.4	150 ppm (top 5%)
31	P07437; F8VW92	Tubulin beta chain	TUBB	21	4	5.5	2.5	0	57.4	223 ppm (top 5%)
32	P17987; E7ERF2	T-complex protein 1 subunit alpha	TCP1	13	13	4.9	2.4	7.32E-171	34.9	334 ppm (top 5%)
33	Q14839; F5GWX5	Chromodomain-helicase-DNA-binding protein 4	CHD4	8	8	3.4	2.3	1.09E-56	6.7	7.83 ppm
34	Q9P258	Protein RCC2	RCC2	6	6	1.8	2.3	9.96E-36	15.9	67.00 ppm (top 10%)
35	Q13263	Transcription intermediary factor 1-beta	TRIM28	6	6	1.4	2.2	3.89E-43	10.4	56.5 ppm (top 10%)
36	P35268	60S ribosomal protein L22	RPL22	2	2	2.3	2.2	1.22E-20	30.5	659 ppm (top 5%)
37	Q9Y5M8	Signal recognition particle receptor subunit beta	SRPRB	5	5	1.6	2.2	6.04E-40	26.2	21 ppm (top 25%)
38	Q96CS3; B4E2M8	FAS-associated factor 2	FAF2	5	5	2.2	2.2	1.14E-93	18.2	16.8 ppm (top 25%)
39	P35232; C9JW96	Prohibitin	PHB	7	7	2.9	2.2	5.73E-82	30.9	213 ppm (top 5%)
40	P27708; F8VPD4	CAD protein;Glutamine-dependent carbamoyl-phosphate synthase;Aspartate	CAD	6	6	4.1	2.2	4.72E-47	3.2	10.7 ppm

|		carbamoyltransferase;Dihydroorotase								
41	Q15758; E9PC01	Neutral amino acid transporter B(0)	SLC1A5	3	3	5.4	2.2	6.72E-34	8.3	22.8 ppm (top 25%)
42	O95831	Apoptosis-inducing factor 1, mitochondrial	AIFM1; PDCD8	7	7	3.4	2.1	2.87E-34	20.4	43.3 ppm (top 10%)
43	Q99832; B8ZZC9	T-complex protein 1 subunit eta	CCT7	3	3	2.0	2.1	6.95E-14	8.1	116 ppm (top 5%)
44	P10412	Histone H1.4	HIST1H1E	5	3	2.2	2.1	4.48E-43	21.5	534 ppm (top 5%)
45	Q16531; E7EPB0	DNA damage-binding protein 1	DDB1	6	6	2.6	1.9	3.02E-107	8.9	22.7 ppm (top 25%)
46	P48643; E9PCA1	T-complex protein 1 subunit epsilon	CCT5	4	4	2.9	1.9	3.45E-20	9.1	278 ppm (top 5%)
47	P08238; Q58FF7	Heat shock protein HSP 90-beta	HSP90AB1	14	9	2.9	1.9	2.46E-115	18.1	713 ppm (top 5%)
48	P08708; P0CW22	40S ribosomal protein S17;40S ribosomal protein S17-like	RPS17; RPS17L	3	3	2.6	1.8	5.71E-42	37.8	443 ppm (top 5%)
49	P78527; E7EUY0	DNA-dependent protein kinase catalytic subunit	PRKDC	10	10	3.7	1.8	1.83E-101	3.8	52.9 ppm (top 10%)
50	P11142; E9PKE3	Heat shock cognate 71 kDa protein	HSPA8	27	23	5.5	1.7	0	35.3	1362 ppm (top 5%)
51	P08107; F5GZ62	Heat shock 70 kDa protein 1A/1B	HSPA1A;HSPA1B	25	13	5.5	1.4	0	39.8	113 ppm (top 5%)

4.2.3 Validation of EBOV/VP24 interactions.

Selected cellular targets identified in the immune-precipitations to EBOV/VP24 by mass spectrometry analysis were then investigated using western blot analysis as an alternative means of identification; using western blot analysis it was also analysed whether these proteins interacted with RESTV/VP24 (*Figure 4.4 A*). These targets included KPNA6 (as a positive control), ATP1A1, LMNB B1, and VDAC-1. The presence of STAT1 was also investigated; even though it had not been identified in the immunoprecipitation, as purified truncated STAT1 and purified truncated VP24 had previously been shown to associate *in vitro* (Zhang et al., 2012). Whole cell lysate (input) and eluate (elution) from separate immunoprecipitations were separated by 1D SDS-PAGE; proteins were then transferred in Western blot and detected using a primary antibody against each selected protein. The data indicated that KPNA6, LMNB1, VDAC1 and ATP1A1 immunoprecipitated with both the EGFP-VP24 and VP24-EGFP moieties (for both strains). STAT1, which was used as a negative control, could be detected in the input fractions but not detected in the immunoprecipitations (*Figure 4.4A*).

To further validate the mass spectrometry results and potential interactions between EBOV/VP24 and cellular proteins, reverse immunoprecipitations were performed against selected cellular targets where antibody combinations allowed. These targets were KPNA6 (as a positive control) and LMNB1. EBOV/VP24-EGFP, EGFP-EBOV/VP24, or EGFP were expressed in 293T cells and cell lysates were prepared. Reverse immunoprecipitations were

performed using specific monoclonal antibodies against KPNA6 and LMNB1. Analysis of the pull-down products using western blot analysis confirmed the presence of EGFP-EBOV/VP24, EBOV/VP24-EGFP, and EGFP in the input and elution fraction. The presence of EGFP-EBOV/VP24 and EBOV/VP24-EGFP was confirmed using a specific antibody against EGFP (*Figure 4.4B*) again validating the forward immunoprecipitations.

A. WB for EGFP-Trap products
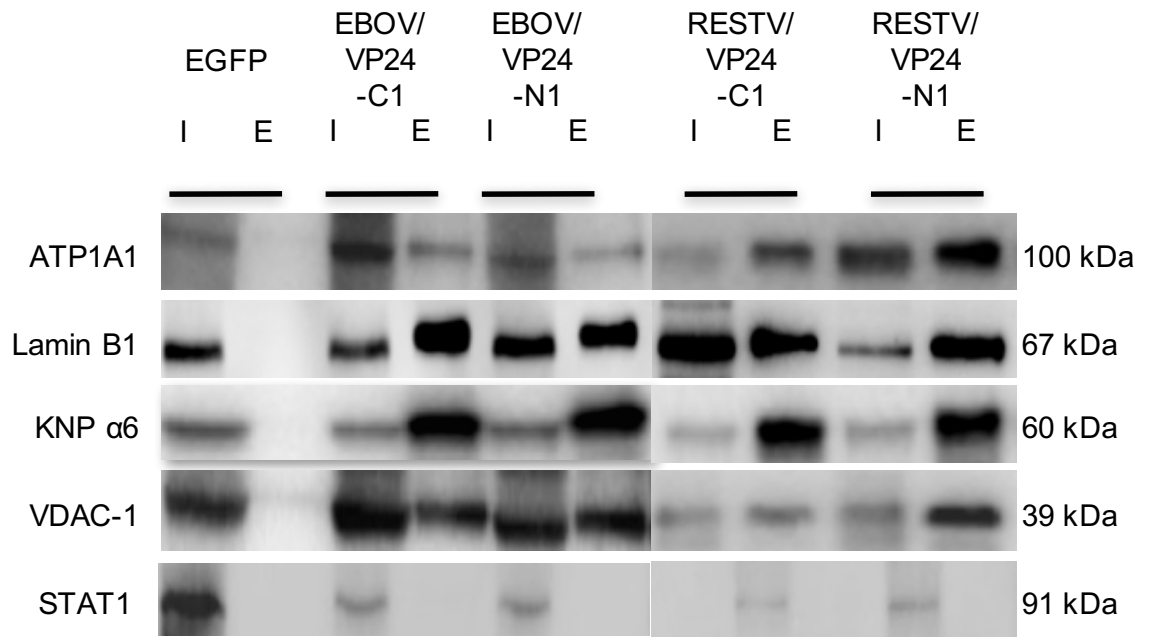
B. Reverse immunoprecipitation for Karyopherin α6 and Lamin B1.

- Karyopherin α6
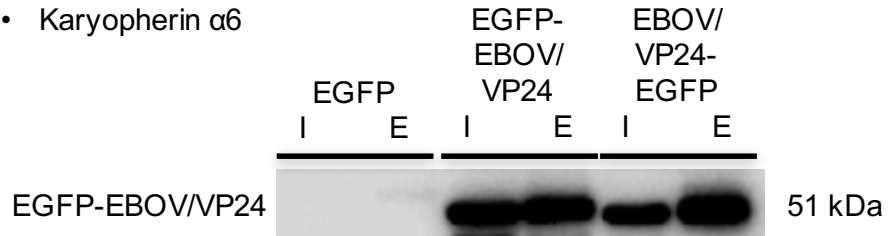
- Lamin B1
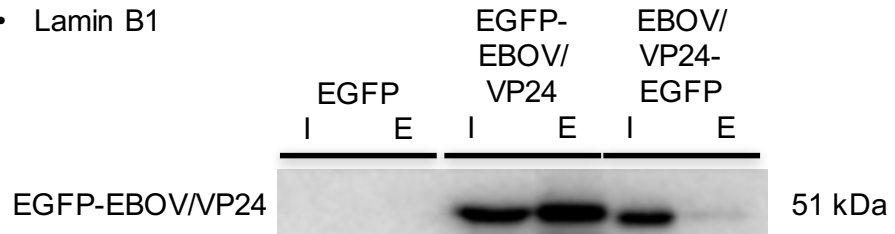
Figure 4.4. Validation of EBOV or RESTV/VP24-EGFP and EGFP-EBOV or RESTV/VP24 interactions by Western blot analysis. (A) *Confirmation of proteins detected in the label free mass spectrometry analysis by western blot analysis was done in the whole cell lysate or input sample (I) and in the eluate sample (E) for five different constructs. Specific monoclonal antibodies against KNPA6, LMNB1, ATP1A1, VDAC-1 and STAT-1 were used to the detection of the example proteins. The antibody anti-EGFP was used as a control to show the presence of the constructs EGFP-EBOV/VP24, EBOV/VP24-EGFP, EGFP-RESTV/VP24 and RESTV/VP24-EGFP in the input and in the elution samples.* ***(B)*** *Validation of EBOV/VP24-EGFP and EGFP-EBOV/VP4 interactions by reverse immunoprecipitations. To further validate the mass Spectrometry results reverse immunoprecipitations were preformed against selected cellular proteins identified by the label free mass spectrometry analysis. The selected proteins for the validation of the label free mass spectrometry by using reverse immunoprecipitations were KNPA6 and LMNB1; the presence of the protein complex EGFP-EBOV/VP24 and EBOV/VP24-EGFP was confirmed using a Western blot analysis and using a specific antibody against EGFP.*

4.2.4 Inhibition of ATP1A1 in Ebola virus infected cells.

To investigate whether ATP1A1 played a functional role in Ebola virus infected cells, the small molecule inhibitor to ATP1A1, ouabain, was used to treat human MRC-5 cells infected with EBOV strain ME718 (MOI = 0.1). This fibroblast lung culture was established from a 14-week gestation human male and is a normal diploid cell line that supports the growth of Ebola virus and has been used in antiviral drug screening studies. (Huggins et al.,1999). Cell viability assays were used to determine the affect of different concentrations of ouabain on cell biology and included two other small-molecule inhibitors (17-AAG and W-7) as positive controls (*Figure 4.5A*).

The data indicated that both 10 and 20 nM ouabain had no significant effect on cell viability compared with either untreated or cells treated with the vehicle (DMSO) only control. The 20 nM concentration of ouabain also had no apparent effect on the abundance of EGFP-VP24 in treated cells expressing this construct (*Figure 4.5C*). Cells were either infected and untreated or infected and treated with 20 nM ouabain to determine the potential effect of this drug on Ebola virus infection. The experiment was repeated in triplicate, and the use of a consistent volume of cell culture supernatant permitted the comparison of the relative abundance of the Ebola virus genome in the supernatant as a proxy measurement for virus particles. (Spurgers et al., 2010; Trombley et al., 2010). Therefore, equal volumes of supernatant containing progeny virus were taken at 24 and 48 h post infection. Viral RNA was then isolated and RNA levels were compared in triplicate using qRT-PCR between each treatment at 24 and 48 h post infection; data were normalized to the level

of viral RNA in infected but untreated cells (*Figure 4.5B*). Results indicated that

infected cells treated with 20nM ouabain produced less viral RNA compared

with infected but untreated cells.
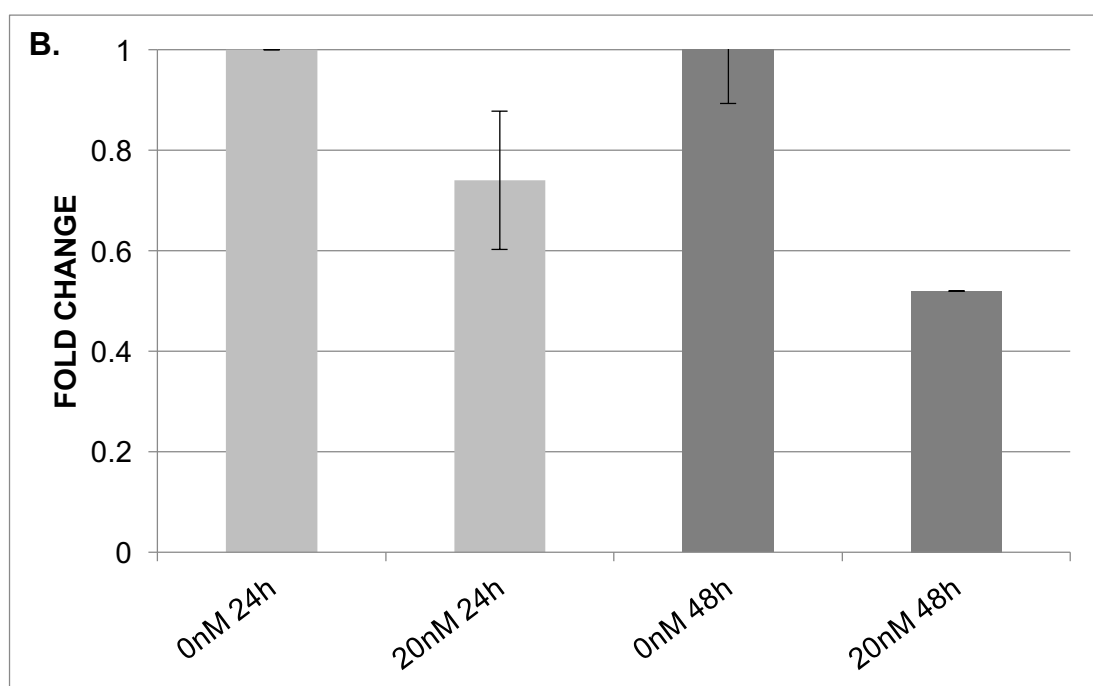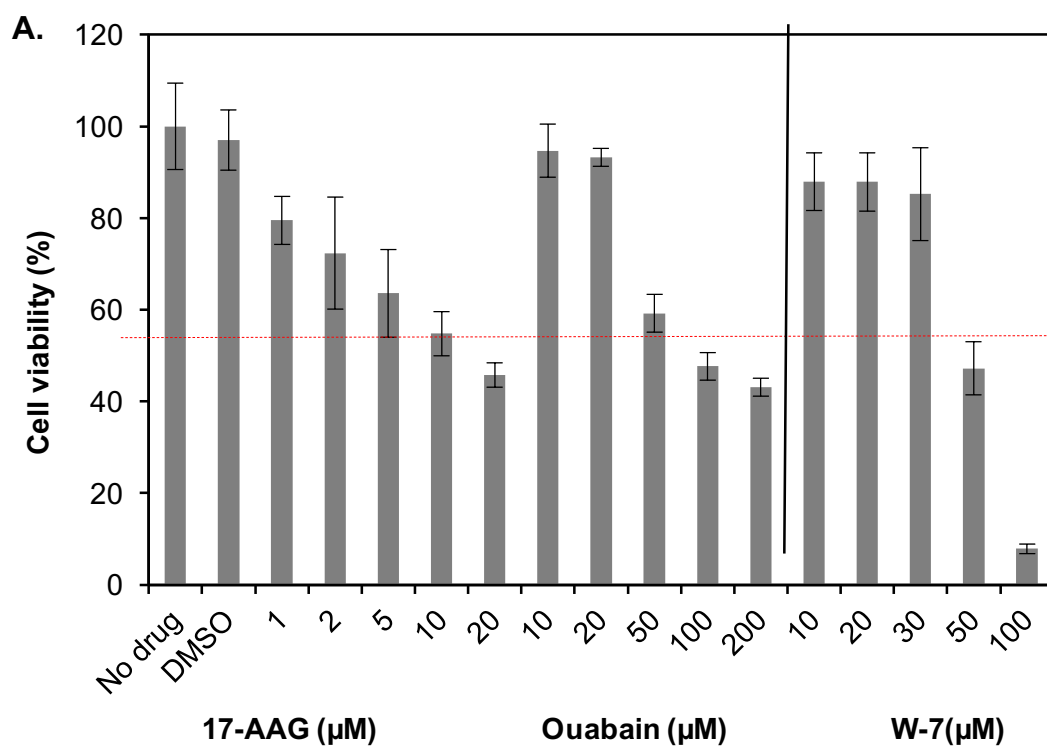
C.
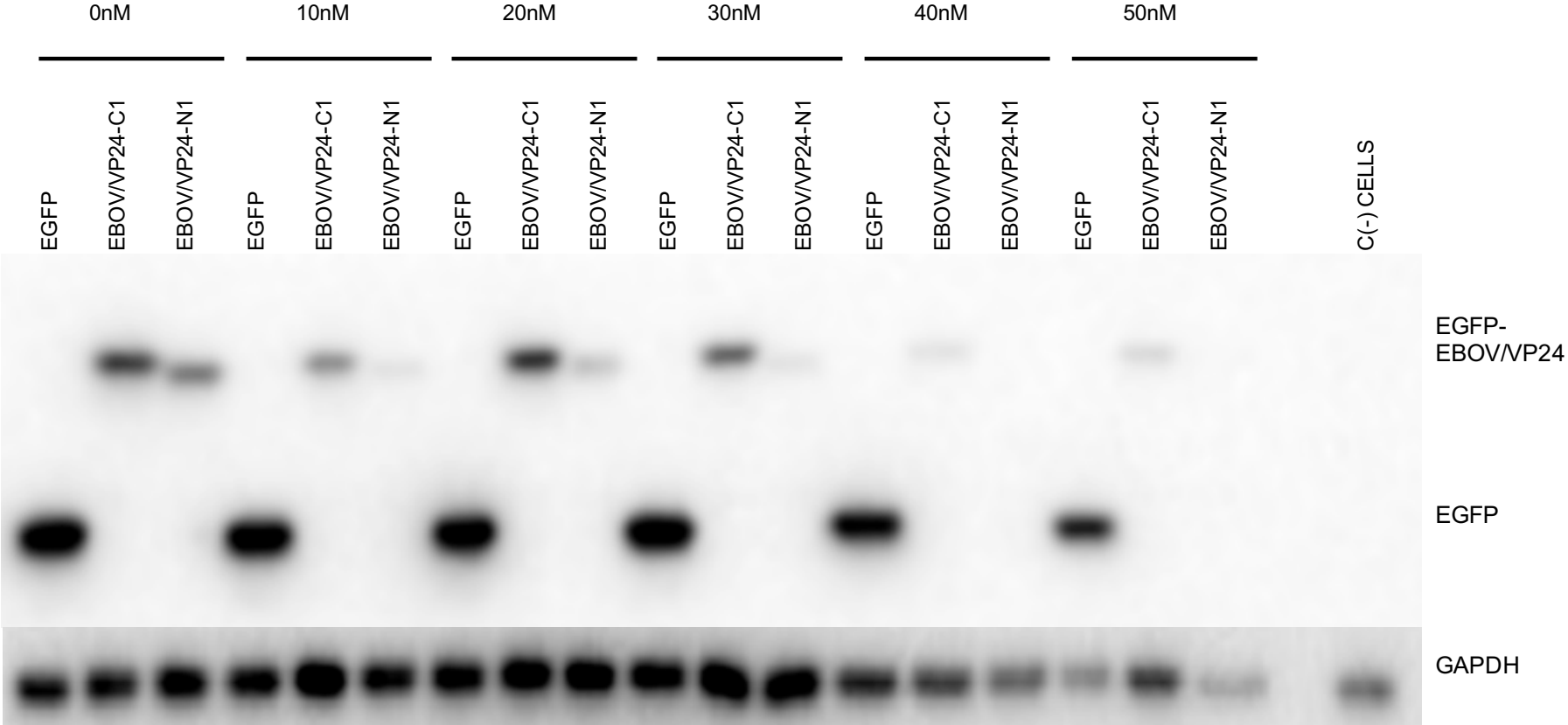
Figure 4.5*. Determining the effect of the ATP1A1 inhibitor ouabain on Ebola virus infected cells. **(A)** An MTT assay was used to determine cell viability in the absence of drug, the vehicle only control (DMSO), and various concentrations of ouabain and two controls using known small molecule inhibitors that affect the function of target cellular proteins (17-AAG and W-7). The data indicated that 10 and 20 nM ouabain and no apparent effect on cell viability compared with the untreated or vehicle only controls. **(B)** qRT-PCR was used to compare the abundance of viral RNA in RNA purified from supernatant taken from cells infected with Ebola virus at either 24 or 48 h post infection in the absence (0 nM) or presence of ouabain (20 nM). The data were normalized to viral RNA levels present in infected but untreated cells at 24 and 48 h post infection **(C)** Stability of EGFP-EBOV/VP24 and EBOV/VP24-EGFP compared with the EGFP control were assessed by Western blot in the absence and presence of various concentrations of ouabain. The abundance of GAPDH was also determined.*

4.3   Discussion

The genomes of RNA viruses have limited encoding capacity, and therefore viral proteins usually possess multiple functions in the cell and virus life cycle. The elucidation of these interactions provides a better understanding of viral pathogenesis and delivers more options for effective antiviral strategies. The elucidation of the interactome of different viral proteins has been described before, one example of this is the work done by Pilchmar *et al* 2012; where the interactome of 70 different viral proteins from 30 different viruses was described. This work showed an approach of similar cellular proteins that interact with the different viral proteins of different families suggesting similar strategies among different viruses (Pichlmair et al., 2012). A more detail study of the interactome for single viral proteins RNA viruses has been done before as well; some examples of that are the study of the interactome of the non-structural protein 2 (NSP2) from the porcine reproductive and respiratory syndrome virus (PRRSV) (Xiao et al., 2016), the NP of PRRSV (Liu et al., 2014; Jourdan et al., 2011), the non-structural protein 1 (NS1) of human respiratory syncytial virus (HRSV) (Wu et al., 2012) These studies not only provided useful information about the interactome of the different viral proteins but also possible new drug therapy targets for these viruses.

There is no much information about the interactome for the different EBOV proteins and therefore the importance of this work. EBOV/VP24 has been shown to have multiple functions in the virus life cycle and therefore would be predicted to interact with a variety of different cellular proteins. Several studies

have investigated this with data suggesting that VP24 can interact with KPNA6, STAT1, and Sec61α (Mateo et al., 2010; Zhang et al., 2012; Iwasa et al., 2011). In this study, affinity tagging coupled to label-free quantitative mass spectrometry was used to identify potential interacting partners of EBOV/VP24 and reduce or eliminate false-positive interactions. Approximately 50 cellular proteins were identified that had a high probability of interacting with EBOV/VP24. The majority of proteins identified associated with either EGFP-EBOV/VP24 or EBOV/VP24-EGFP. However, some of the cellular proteins identified were unique to the different fusion proteins, however the proteins with the highest ratio are common for both constructs This may have been due some of the interactions with the cellular proteins being sterically hindered by the EGFP moiety and emphasized the advantage of using both the N- and C-terminal tagged fusion proteins.

EBOV/VP24 has affinity for the plasma membrane and has also been observed in the cytoplasm where the protein formed cytoplasmic inclusions (Nanbo et al., 2013). These are the cellular locations of a number of the proteins identified as associating with EBOV/VP24 (*Table 4.1*). The reliability of this mass spectrometry data was confirmed using western blot analysis and reverse immunoprecipitations against selected cellular proteins. Furthermore, the precipitations and mass spectrometry analysis were repeated five times with independent samples to ensure the reproducibility of the technique and also for the statistical analysis. The immunoprecipitation technique was also optimized for identifying low-affinity and transient interactions by modulating the salt concentration in the washing step.

KPNA6 and 1 KPNA1, identified in this study with an average binding ratio of 8 to both EGFP-VP24 and VP24-EGFP, reflected previous work demonstrating an interaction between EBOV/VP24 and karyopherin α subunit 6 (Reid et al., 2006; Mateo et al., 2010). In addition, as part of a larger viral/host protein study, the interactome of VP24 was determined using tandem affinity purification, and KPNA6 and other karyopherins were identified as part of 48 cellular proteins that potentially interacted with EBOV/VP24 (Pichlmar et al., 2012). Proteins in common with this previous work and recorded in our analysis as having a greater than two-fold abundance with either EGFP-EBOV/VP24 or EBOV/VP24-EGFP were ANP32A, ANP32B, CCAR2 (KIAA1967), CGGBP1, HNRNPL, HNRNPU, KPNA1, KPNA6, MYO5A, POM121C, PRKCSH, RPS27, RPS27L, SLC25A3, TUB1C, and TUBB, 28 proteins were not detected in our study, and three proteins (EEF1A1, EEF1A2, and SPOP) were identified as proteins that formed background/nonspecific interactions.

In the present study, no interaction was detected between EBOV/VP24 and STAT1 using a cell based and mass spectrometry approach, and this was confirmed by Western blot. Likewise, no interaction between VP24 and STAT1 was described in the tandem affinity approach (Pilchmar et al., 2012). This differs from previously published work that described an interaction between VP24 and STAT1 (Zhang et al., 2012). However, the difference in these results may be due to different systems employed. In previously published work, the interaction between EBOV/VP24 and STAT1 was determined by ELISA using purified truncated forms of VP24 and STAT1 (Zhang et al., 2012). In the

present study, the EGFP moiety may have prevented interaction with STAT1, likewise with the tandem affinity approach described previously (Pichlmair et al., 2012), although the tag did not prevent interaction with the karyopherins and other cellular proteins.

SEC61α was not listed in the final data set, but proteins that form a complex with SEC61α were identified, including SSR4 and Sec61B. Notably, Sec61B has been formally demonstrated to be significant using biological approaches (Iwasa et al., 2011). Indeed, SEC61B protein was found in one of the trial runs of this study but excluded after the statistical analysis of the five independent mass spectrometry analyses of EGFP-EBOV/VP24 and EBOV/VP24-EGFP. This suggests that in the case of Sec61B our selection criteria were too conservative or affected by the EGFP moiety or the interaction with Sec61α may be a weaker or more transient interaction than those listed in Table 1. Although, we note that the interaction between VP24 and Sec61α was identified using a tandem tagging approach of VP24 (Iwasa et l., 2011) but was not identified in a more recent study (Pichlmar et al., 2012).

One cellular protein with a high binding ratio was ATP1A1, a Na+/K+-ATPase. The function of this protein is in establishing and maintaining Na^+ and K^+ electrochemical gradients across the plasma membrane and in cell signalling. Inhibiting the function of ATP1A1 with ouabain resulted in a decrease in progeny virus. Ouabain is used in treating atrial fibrillation and heart failure and has been demonstrated to have antiviral effect with other viruses including herpes simplex virus (Dodson et al., 2007) and porcine reproductive and

respiratory syndrome virus (Karuppannan et al., 2012). The effect of ouabain on ATP1A1 may not have had a direct effect on the relationship between ATP1A1 and VP24. For example, treatment of cells infected with human cytomegalovirus with ouabain prevented enlargement of cells that is associated with progeny virus production (Altamirano et al., 1994). Nevertheless, the use of this small molecule to reduce progeny virus in Ebola-virus-infected cells illustrates how existing therapeutics can be repurposed for antiviral therapy. In the case of a natural infection in humans, even a modest reduction in viral progeny production in vivo may slow the virus sufficiently for the host immune response to mount a life-saving response.

This was the first detailed interactome analysis of EBOV/VP24 using quantitative proteomics to identify cellular proteins that have a high probability of interacting with EBOV/VP24 and to eliminate or identify cellular proteins that potentially associate with a binding matrix used in immunoprecipitations. The current study complemented previous work that used small molecules to inhibit the function of cellular proteins to disrupt Ebola virus biology. For example, *S*-adenosylhomocysteine hydrolase inhibitors were evaluated both in vitro and in vivo and were detrimental to Ebola virus infection (Huggins et al., 1999). Inhibition of heat shock protein 90 using 17-AAG disrupted Ebola biology in vitro (Smith et al., 2010). At the time of publication of this manuscript, the largest EBOV outbreak so far known is occurring in West Africa and has proved difficult to control; the "repurposing" of therapeutics to reduce virus infection may tip the balance between recovery and death.

Overall, the proteomic approach demonstrated how determining the interactome of viral proteins could be used to identify cellular proteins that play important roles in the virus lifecycle and therefore increase the repertoire of potential druggable targets. Resistance is a constant problem in developing antiviral therapy to target the function of viral proteins. Transiently targeting the function of host cell proteins crucial to virus biology offers an exciting new therapeutic avenue, with the potential to solve the problem of resistance, as pro-viral cellular proteins are evolutionarily static on the time scale of lytic virus replication as well as separated from the genome that would benefit from resistance.

CHAPTER 5: Elucidation of the Ebolavirus NP cellular interactome reveals association with the protein chaperone pathway and targeted inhibition disrupts virus biology

5.1   Introduction

Viruses that belong to the order *Mononegavirales* have a negative-sense non-segmented RNA genome. In these viruses the genome does not exist as a free RNA molecule, instead the RNA is encapsidated by a viral nucleoprotein (Lamb. A R., 2013). *Ebolavirus* belongs to the *Filoviridae* family, in the order *Mononegavirales*. In the *Ebolavirus* genus, the nucleoprotein mRNA is transcribed from the first gene of the *Ebolavirus* genome and is translated to generate a protein of 739 amino acids that can be divided into a hydrophobic N-terminal half (approx. 350 aa) and a hydrophilic C-terminal half (approx. 400aa) (Sanchez et al 1989; Feldman et al., 2013). The N-terminal region is thought to be important for RNA binding via its interaction with the phosphodiester backbone of the RNA (Noda et al., 2010). When the *Ebolavirus* NP binds to the viral RNA genome it forms a highly stable NP-RNA complex structure; this NP-RNA complex is long, flexible and forms helical structures (Ruigrok et al., 2011). NP along with VP30, VP35 and (RNA-dependent RNA polymerase) L form the RNP (Feldmann et al., 2013).

Ebolavirus NP has potentially multiple functions during virus biology playing an essential role in virus replication (Shi et al., 2008; Feldman et al., 2013). NP binds to the viral RNA and protects the viral RNA from degradation; also the NP-RNA complex serves as a template for it synthesis, transcription and replication which is carried out by L (Lamb & Parks, 2007; Ruigrok et al., 2011). *Ebolavirus* NP also interacts with other viral proteins in order to achieve some

of their functions; for example, NP enhances the budding process of VP40. Co-expression of VP40 and NP have been reported to enhance the production of virus like particles (VLPs); the C-terminus of the NP may interact with VP40 (Noda et al., 2010; Licata et al., 2004). Several studies have investigated how *Ebolavirus* NP interacts with viral RNA (Noda et al., 2010) and with other viral proteins like VP35 (Huang et al., 2002; Noda et al., 2011; Kirchdoerfer et al 2015), VP40 (Licata et al., 2004) and VP24 (Huang et al., 2002). Due to the multiple roles that NP plays in the *Ebolavirus* biology, many of these may be mediated or have cellular proteins that contribute to function. Therefore, defining the cellular interactome of NP will provide new information on how NP may contribute to the life cycle of the virus. This may also identify new possible targets for antivirals against *Ebolavirus*, by targeting the function of cellular proteins that are essential for virus biology (Smith et al., 2010).

To determine which cellular proteins or cellular protein complexes interact with *Ebolavirus* NP, in this case NP for RESTV and EBOV; and to predict function of these interactions, a high-affinity immunoprecipitation coupled to a label free mass spectrometry-based approach was used. Using conservative selection criteria, approximately 120 cellular proteins were identified that had a high probability of interacting with EBOV/NP and RESTV/NP. Although not many studies regarding the interactome of *Ebolavirus* nucleoprotein have been published, similar studies have been reported with other *Ebolavirus* proteins and also other viral proteins from viruses that belong to the Order *Mononegavirales* (García et al., 2014; Pilchmar et al., 2012; Munday et al., 2015).

138

The study described in this chapter showed new possible binding proteins for *Ebolavirus* NP (EBOV/NP and RESTV/NP); including the chaperone HSP70 and its co-chaperones HSP40 and BAG2. To further investigate the role of the interaction between HSP70 and *Ebolavirus* NP, and the effect on viral RNA synthesis, a small molecule inhibitor of HSP70, VER 155008, was used. In order to do this, a mini-replicon system for EBOV (Makona strain) was developed; this mini-genome system recapitulated the essential stages of virus replication and transcription and expressed a luciferase gene reporter which can be used to study the effect of different compounds in the EBOV biology (Hoenen et al., 2013). The data presented in this work indicated that amount of luciferase produced by the mini-genome system decreased in cells treated with the HSP70 inhibitor compound VER-155008 compared to untreated cells, therefore the replication of EBOV was reduced by inhibition of HSP70. These results suggest that HSP70 may play an important role for the EBOV virus replication probably by providing stability to NP.

5.2. **Results**

5.2.1 Expression of *Ebolavirus* nucleoprotein in 293T cells

To determine the cellular interacting partners of EBOV/NP and RESTV/NP a high affinity EGFP co-immunoprecipitation (co-IP) coupled to a label-free mass spectrometry approach was used (*Figure 5.1.A*). In order to do this, the target proteins (EBOV/NP and RESTV/NP) were expressed as EGFP-fusion proteins. This technique has been shown to improve the sensitivity and also to allow the discrimination between specific and non-specific interaction to the target proteins (Trinkle-Mulcahy et al., 2008). The gene encoding the NP from EBOV or RESTV was cloned either at the 5' or 3' of an EGFP gene (EBOV/NP-EGFP and RESTV/NP-EGFP or EGFP-EBOV/NP and EGFP-RESTV/NP), creating a contiguous open read frame and the expression of a fusion protein. The expression of EGFP at either the N or C-terminal of NP was done in order to mitigate for any steric hindrance caused by the fusion EGFP molecule. The genes that encode for EBOV/NP and RESTV/NP were then codon optimized in order to facilitate expression in human cells (293T cells), as was described previously for EBOV/VP24 in chapter 4.

Once EBOV/NP-EGFP and RESTV/NP-EGFP were overexpressed in cells, proteins were extracted from lysed cells and used for immunoprecipitation. To identify the potential interacting partners of EBOV/NP and RESTV/NP, the eluate or bound fraction was then analysed by label free mass spectrometry. To reduce the false positive rate, EGFP alone was also expressed in 293T

cells as a control. The 293T cells (human embryonic kidney cell line) were selected for this study due to their high efficiency for calcium phosphate mediated transfection, annotation of the human protein database and because this cell line is permissive for *Ebolavirus* infection (Khan et al., 2013; Jordan et al., 2004; Kuhn et al., 2008). Protein expression was confirmed in 293T cells using immunofluorescence; the efficiency of transfections for all plasmids was approximately 70%, the panels show that expression of the four constructs were similar, however EGFP-EBOV/NP and EGFP-RESTV/NP show a higher and more uniform expression in cells than EBOV/NP-EGFP and RESTV/NP-EGFP. (*Figure 5.1B*). Possible interacting partners for EBOV/NP and RESTV/NP were then immunoprecipitated using a GFP-Trap (Nano trap technology), which used highly specific single chain anti-EGFP antibody. After immunoprecipitation, both input and bound samples were analysed by western blotting; a protein corresponding to the molecular weight of EBOV/NP-EGFP or RESTV/NP-EGFP (approx. 131 kDa) in both samples was detected using an anti-EGFP Ab and an anti-EBOV/NP antibody (*Figure 5.1.C*). Label free mass spectrometry and quantitative proteomics was then used to distinguish the EBOV/NP-EGFP, EGFP-EBOV/NP, RESTV/NP-EGFP and RESTV/NP-EGFP interactomes from the control (EGFP alone), which led to the identification of possible interacting partners for EBOV/NP and RESTV/NP.
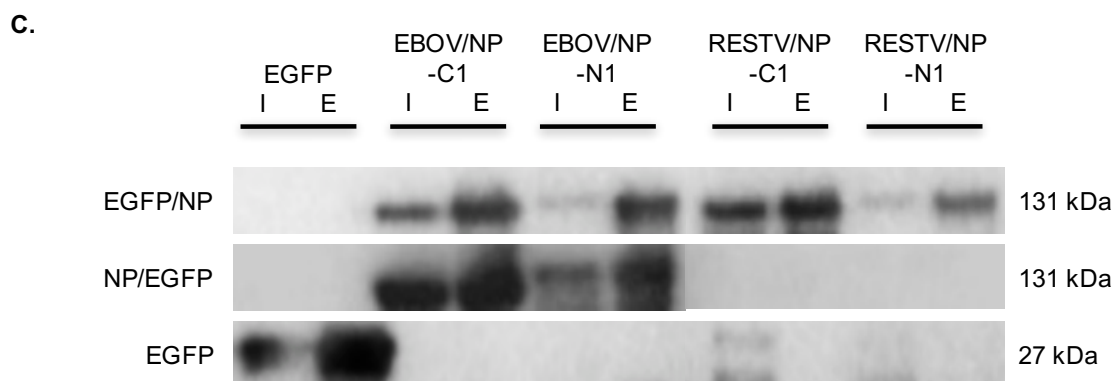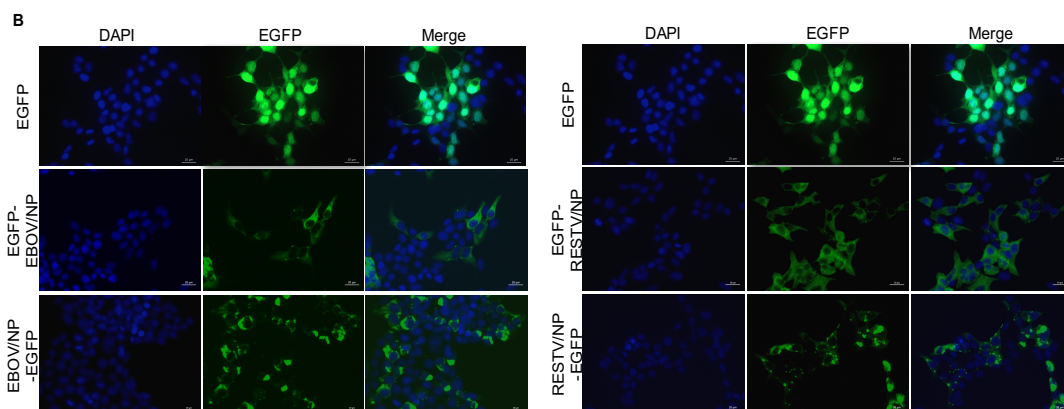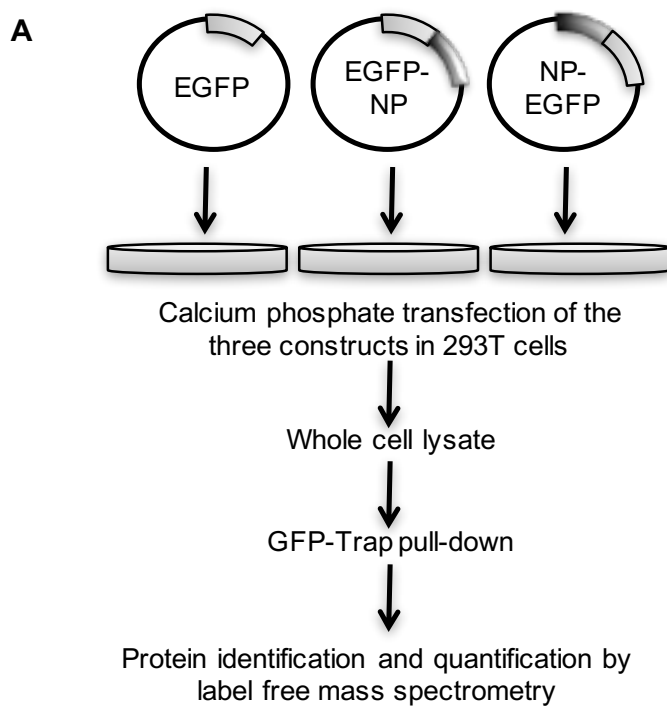
Figure 5.1. Expression of Ebolavirus NP in 293T cells. (A) *Schematic representation of the methodology used in this study. 293T cells were grown in DMEM supplemented with 10% FBS and 1% penicillin-streptomycin at 37°C with 5% CO2. Two 145cm2 dishes were seeded with $4x10^6$ cells 24 hours prior to calcium phosphate transfection with 25.6 µg of plasmid DNA encoding EGFP, EGFP-EBOV/NP, EBOV/NP-EGFP, EGFP-RESTV/NP and RESTV/NP-EGFP respectively. 24 hours post transfection the cells were harvested, lysed and immunoprecipitated using a GFP-Trap (Chromotek). Label free Mass Spectrometry analysis on the eluted samples was then carried out. (B)* *Expression of EGFP-EBOV/NP, EBOV/NP-EGFP, EGFP-RESTV/NP, RESTV/NP-EGFP and the control EGFP in 293T cells was confirmed by immunofluorescence using confocal microscopy. (C)* *Analysis of the pull down products using a western blot confirm the presence of EGFP-NP and NP-EGFP (for both strains) and the control EGFP with the expected molecular weight, confirming the expression of these proteins in 293T cells*

5.2.2 Identification of the potential cellular interacting partners of *Ebolavirus* nucleoprotein

After overexpression of EGFP-EBOV/NP, EBOV/NP-EGFP, EGFP-RESTV/NP, RESTV/NP-EGFP and EGFP in 293T cells, the cellular proteins were extracted and then immunoprecipitated using a GFP-trap. Potential interacting partners for EBOV/NP and RESTV/NP were identified by label free mass spectrometry. A statistical analysis approach was done in order to help distinguish the EBOV/NP and RESTV/NP interacting partners from the background non-specific binding proteins (those cellular proteins that bind to the binding matrix or EGFP alone) (Trinkle-Mulcahy et al., 2008). The experiment was repeated three times. Approximately 1050 proteins were identified by mass spectrometry after the immunoprecipitation. After statistical analysis 150 proteins were filtered to have a high probability (95% of confidence) of interacting with the EBOV/NP and RESTV/NP.

The statistical analysis of the mass spectrometry results and all data sets were calculated using the Perseus program algorithm (Max Plank Institute); with this software a t-Test analysis was done with a p-value <0.01; proteins identified with a single peptide were removed to give strength to the analysis. In the t-Test analysis each group of proteins detected in the MS for all the constructs that express *Ebolavirus* NP (EBOV/NP-EGFP, EGFP-EBOV/NP, RESTV/NP-EGFP and EGFP-RESTV/NP) was compared to the control (plasmid expressing EGFP alone). The data was organized then in the form of volcano plots (*Figure 5.2*); in these volcano plots the logarithmic ratio of protein

intensities (x-axis) was plotted against the negative logarithmic p-values of the t test (y-axis). The proteins that were significant or had a high probability of interacting with *Ebolavirus* NP had a high log ratio and therefore are located in the right-up hand quadrant. Each dot outside the volcano represents a protein with a high probability of interacting with *Ebolavirus* NP; those proteins inside the volcano are the background proteins or the proteins that do not interact with *Ebolavirus* NP. Proteins that had a binding ratio greater than 2 and were statistically significant are shown in the right-up hand quadrant (Trinkle-Milcahy et al., 2012). For EBOV/NP and RESTV/NP, approximately 150 and 130 proteins were identified as interacting partners with statistical significance (*Tables A.1 and A.2 in Appendix*); some of the proteins with the higher probability to interact with EBOV/NP or RESTV/NP were selected for following studies (*Tables 5.1 and 5.2*).
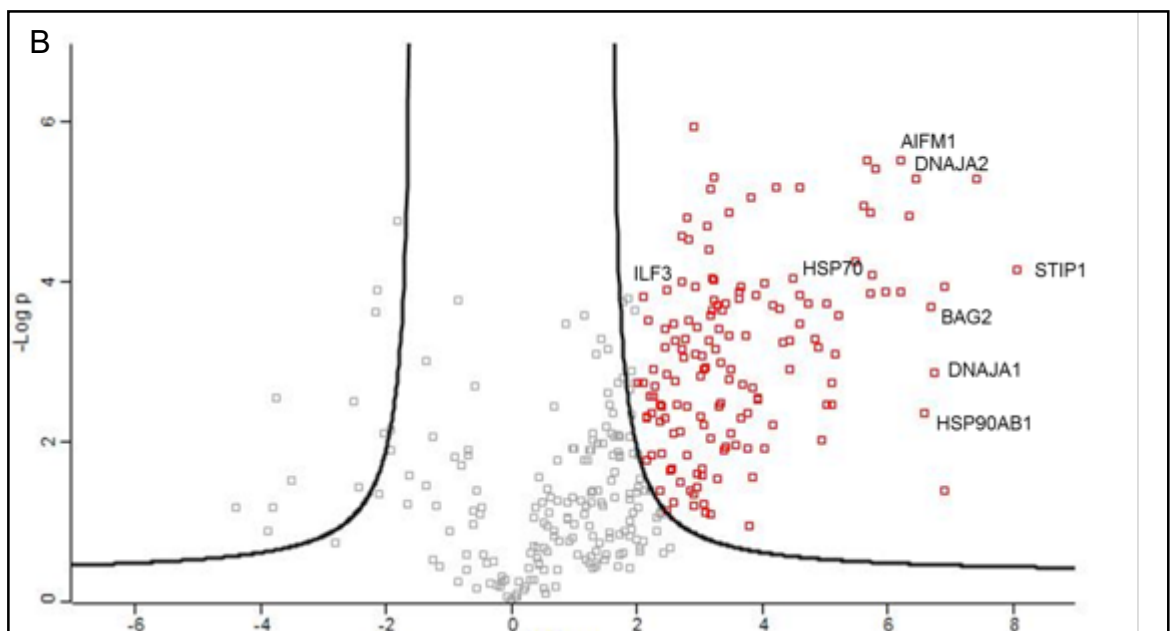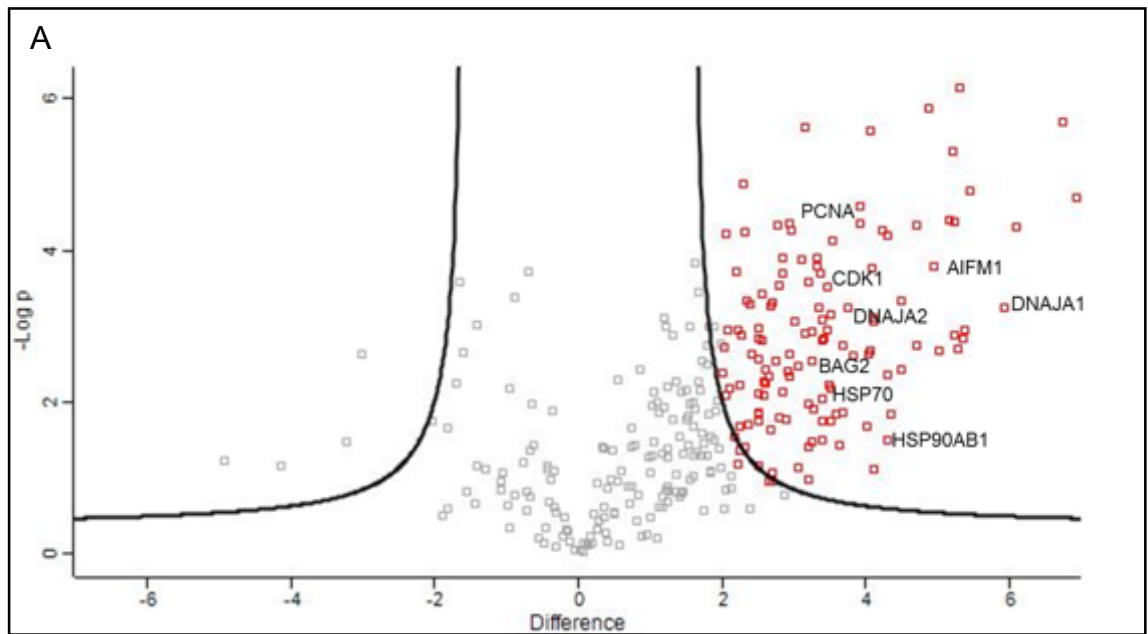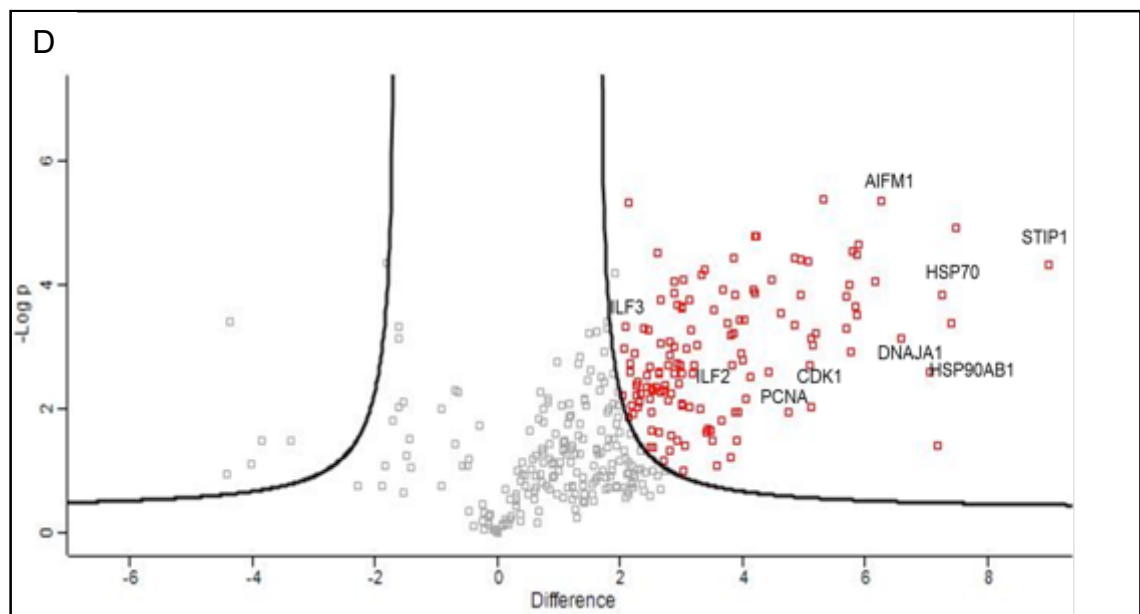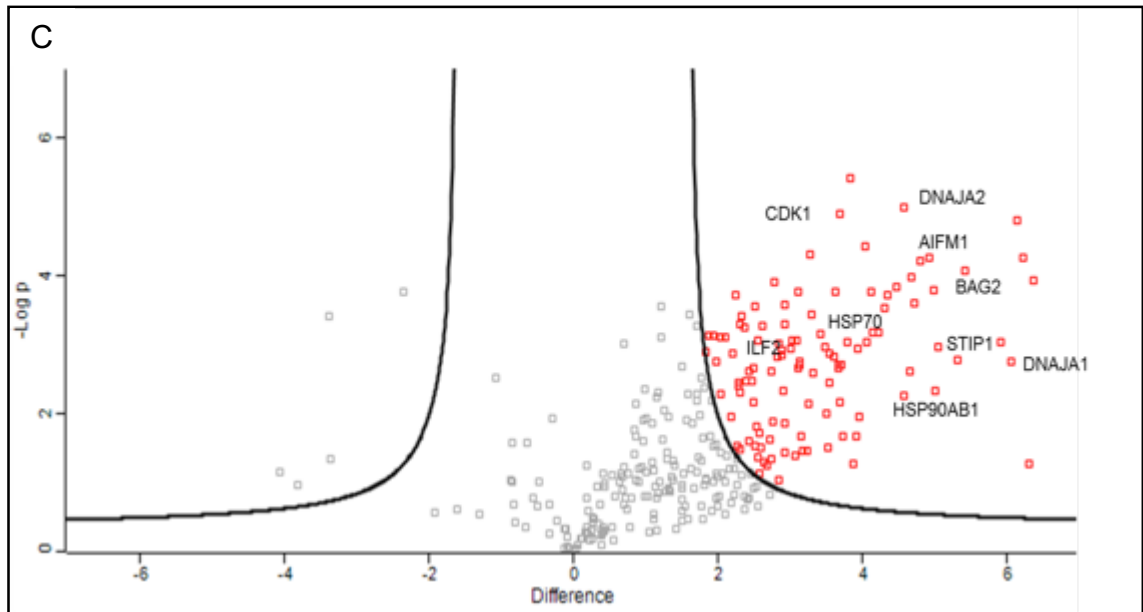
Figure 5.2 *Volcano plot representing results of the MS and the statistical analysis, the pull down and label free mass spectrometry was done in triplicate for a more accurate analysis. In these volcano plots the dots outside the volcano represents the potential protein interacting partners for* **EGFP-EBOV/NP (A), EBOV/NP-EGFP (B), EGFP-RESTV/NP (C) and RESTV/NP-EGFP (D).** *For any potential protein interaction partner with EBOV/NP or RESTV/NP, the value of its abundance co-immunoprecipitated with any of Ebolavirus nucleoproteins was compared to the value of co-immunoprecipitated with the control (EGFP alone).*

Table 5.1: Cellular proteins that have a higher probability of forming protein-protein interactions with EBOV/NP and RESTV/NP. *Shown are candidate proteins identified using the (A) EGFP-EBOV/NP, (B) EBOV/NP-EGFP, (C) EGFP-RESTV/NP, and (D) RESTV/NP-EGFP fusion proteins and identified using label free quantitative proteomics. Protein identifier, protein name and gene names are indicated. Total and unique peptides used to identify the protein are indicated. The –LogP value is a comparison of the cellular protein between the EBOV/NP or RESTV/NP pull down and EGFP, where the higher the number means the higher probability of interacting, and a threshold above 2.0 has been selected. The t-test Difference is the difference of the means of the intensities of the cellular proteins in the EBOV/NP or RESTV/NP pull down and EGFP. The percentage of sequence coverage of the protein identified using the peptides is indicated SC%. The abundance (ppm) of the protein in an average human cell is listed – data taken from the PaxDb: Protein Abundance Across Organisms database*

TABLE 5.1A: *Significant proteins that interact with EGFP-EBOV/NP.*

	Protein IDs	Protein names	Gene names	Peptides	Unique peptides	-Log t-test P value	T-test Difference (EGFP-EBOV/NP vs EGFP)	SC [%]	Abundance (PPM)
1	Q5JP53;P07437	Tubulin beta chain	TUBB	42	8	1.3	7.3	78.4	1041 ppm
2	Q13825;B4DYI6	Methylglutaconyl-CoA hydratase, mitochondrial	AUH	8	8	4.7	6.9	28.0	27.3 ppm
3	Q16342;F5H4V9	Programmed cell death protein 2	PDCD2	11	11	5.7	6.7	39.0	NA
4	Q3ZCM7;F5H0I4	Tubulin beta-8 chain	TUBB8	13	2	4.3	6.1	25.2	175 ppm
5	P31689;B7Z5C0	DnaJ homolog subfamily A member 1	DNAJA1	25	25	3.2	5.9	63.5	15.2 ppm

6	Q96EY1;I3L1T6	DnaJ homolog subfamily A member 3, mitochondrial	DNAJA3	10	1	4.8	5.5	26.9	1.19 ppm
7	P49411;H3BNU3	Elongation factor Tu, mitochondrial	TUFM	23	23	3.0	5.4	59.7	991 ppm
8	O43175;Q5SZU1	D-3-phosphoglycerate dehydrogenase	PHGDH	33	33	2.8	5.4	64.4	100 ppm
9	Q14257;F8WCY5	Reticulocalbin-2	RCN2	12	12	6.1	5.3	50.2	112 ppm
10	Q9BUF5;K7ESM5	Tubulin beta-6 chain	TUBB6	23	10	2.7	5.3	67.0	183 ppm
11	Q9BVA1	Tubulin beta-2B chain	TUBB2B	36	1	4.4	5.2	73.9	543 ppm
12	P30153;F5H3X9	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform	PPP2R1A	14	14	2.9	5.2	36.8	184 ppm
13	P17987;E7ERF2	T-complex protein 1 subunit alpha	TCP1	26	26	5.3	5.2	62.2	114 ppm
14	Q9H7B4;B0QZ88	Histone-lysine N-methyltransferase SMYD3	SMYD3	7	7	4.4	5.2	17.5	NA
15	Q10567;P63010	AP-1 complex subunit beta-1	AP1B1	13	13	2.7	5.0	13.8	127 ppm
16	O95831;E9PMA0	Apoptosis-inducing factor 1, mitochondrial	AIFM1	13	13	3.8	4.9	32.6	688 ppm
17	K7EJL1;B4DDG7	AP-1 complex subunit mu-1	AP1M1	8	8	5.9	4.9	26.2	3.03 ppm
18	H0YEN5;P15880	40S ribosomal protein S2	RPS2	10	10	2.7	4.7	49.7	131 ppm
19	Q9BQE3;F5H5D3	Tubulin alpha-1C chain	TUBA1C	37	2	2.4	4.5	76.8	691 ppm
20	F5GZS6;J3KPF3	4F2 cell-surface antigen heavy chain	SLC3A2	8	8	3.3	4.5	17.9	147 ppm
21	P68363;A8MUB1	Tubulin alpha-1B chain;Tubulin alpha-4A chain	TUBA1B	39	4	1.8	4.4	76.5	749 ppm
22	Q9Y230;B3KQ59	RuvB-like 2	RUVBL2	28	28	2.4	4.3	74.7	23.4 ppm
23	P56192;A6NC17	Methionine--tRNA ligase, cytoplasmic	MARS	12	12	4.2	4.3	20.0	6.11 ppm
24	P04843;B7Z4L4	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1	RPN1	10	10	1.5	4.3	21.7	179 ppm
25	Q9GZT9	Egl nine homolog 1	EGLN1	3	3	4.3	4.2	14.6	NA
26	P85037;E9PM37	Forkhead box protein K1	FOXK1	8	6	3.1	4.1	11.5	NA
27	O60884;H3BMW5	DnaJ homolog subfamily A member 2	DNAJA2	10	10	3.1	4.1	29.1	3.39 ppm
28	P27824;B4DGP8	Calnexin	CANX	17	17	1.1	4.1	34.3	1213 ppm
29	B7Z9I1;Q5T4U5	Medium-chain specific acyl-CoA dehydrogenase, mitochondrial	ACADM	8	8	3.8	4.1	29.1	828 ppm
30	P35613;R4GN83	Basigin	BSG	3	3	5.6	4.1	12.5	21.7 ppm
31	P11177;C9J634	Pyruvate dehydrogenase E1 component subunit beta, mitochondrial	PDHB	9	9	2.7	4.1	27.9	393 ppm
32	H7C4H2;Q9Y5M8	Signal recognition particle receptor subunit beta	SRPRB	6	6	2.6	4.0	46.2	17.5 ppm
33	P08238;Q58FF7	Heat shock protein HSP 90-beta	HSP90AB1	35	26	1.7	4.0	49.0	693 ppm

34	Q9Y5J9;G3XAN8	Mitochondrial import inner membrane translocase subunit Tim8 B	TIMM8B	6	6	4.6	3.9	60.2	NA
35	Q9BSD7;Q5TDF0	Cancer-related nucleoside-triphosphatase	NTPCR	7	7	4.4	3.9	56.8	4.71 ppm
36	P49368;B4DUR8	T-complex protein 1 subunit gamma	CCT3	15	15	2.6	3.8	37.6	139 ppm
37	P04792;F8WE04	Heat shock protein beta-1	HSPB1	7	7	3.2	3.7	45.4	2459 ppm
38	P36542;B4DL14	ATP synthase subunit gamma, mitochondrial	ATP5C1	6	6	2.7	3.7	22.8	811 ppm
39	P05141;P12235	ADP/ATP translocase 2	SLC25A5	17	10	1.9	3.7	53.4	999 ppm

TABLE 5.1B: *Significant proteins that interact with EBOV/NP-EGFP.*

	Protein IDs	Protein names	Gene names	Peptides	Unique peptides	-Log P value	t-test Difference (EBOV/NP-EGFP vs EGFP)	SC [%]	Abundance (PPM)
1	P31948;G3XAD8	Stress-induced-phosphoprotein 1	STIP1	58	58	4.2	8.1	74	120 ppm
2	Q14257;F8WCY5	Reticulocalbin-2	RCN2	12	12	5.3	7.4	50.2	112 ppm
3	Q5JP53;P07437	Tubulin beta chain	TUBB	42	8	1.4	6.9	78.4	1041 ppm
4	B4E2W0;P55084	Trifunctional enzyme subunit beta, mitochondrial;3-ketoacyl-CoA thiolase	HADHB	21	21	3.9	6.9	51.5	777 ppm
5	P31689;B7Z5C0	DnaJ homolog subfamily A member 1	DNAJA1	25	25	2.9	6.7	63.5	15.2 ppm
6	O95816;B4DXE2	BAG family molecular chaperone regulator 2	BAG2	15	15	3.7	6.7	77.3	5.24 ppm
7	P08238;Q58FF7	Heat shock protein HSP 90-beta	HSP90AB1	35	26	2.4	6.6	49	693 ppm
8	Q3ZCM7;F5H0I4	Tubulin beta-8 chain	TUBB8	13	2	5.3	6.5	25.2	175 ppm
9	P40939;B4DYP2	Trifunctional enzyme subunit alpha, mitochondrial;Long-chain enoyl-CoA hydratase	HADHA	16	16	4.8	6.3	30.9	1175 ppm
10	P49411;H3BNU3	Elongation factor Tu, mitochondrial	TUFM	23	23	3.9	6.2	59.7	991 ppm
11	P17066;P48741	Heat shock 70 kDa protein 6	HSPA6	19	2	5.5	6.2	20.5	104 ppm
12	Q9BUF5;K7ESM5	Tubulin beta-6 chain	TUBB6	23	10	3.9	6.0	67	183 ppm
13	Q9BVA1	Tubulin beta-2B chain	TUBB2B	36	1	5.4	5.8	73.9	543 ppm
14	Q16342;F5H4V9	Programmed cell death protein 2	PDCD2	11	11	4.1	5.8	39	NA
15	P17987;E7ERF2	T-complex protein 1 subunit alpha	TCP1	26	26	3.9	5.7	62.2	114 ppm

TABLE 5.1C*: Significant proteins that interact with EGFP-RESTV/NP*

	Protein IDs	Protein names	Gene names	Peptides	Unique peptides	-Log t-test P value	t-test Difference (EGFP-RESTV/NP vs EGFP)	SC [%]	Abundance (PPM)
1	Q3ZCM7;F5H0I4	Tubulin beta-8 chain	TUBB8	13	2	3.9	6.4	25.2	175 ppm
2	Q5JP53;P07437	Tubulin beta chain	TUBB	42	8	1.3	6.3	78.4	1041 ppm
3	Q16342;F5H4V9	Programmed cell death protein 2	PDCD2	11	11	4.2	6.2	39	NA
4	P17987;E7ERF2	T-complex protein 1 subunit alpha	TCP1	26	26	4.8	6.1	62.2	114 ppm
5	P31689;B7Z5C0	DnaJ homolog subfamily A member 1	DNAJA1	25	25	2.7	6.0	63.5	15.2 ppm
6	P49411;H3BNU3	Elongation factor Tu, mitochondrial	TUFM	23	23	3.0	5.9	59.7	991 ppm
7	O95816;B4DXE2	BAG family molecular chaperone regulator 2	BAG2	15	15	4.1	5.4	77.3	5.24 ppm
8	P31948;G3XAD8	Stress-induced-phosphoprotein 1	STIP1	58	58	2.8	5.3	74	120 ppm
9	P30153;F5H3X9	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform	PPP2R1A	14	14	2.9	5.0	36.8	184 ppm
10	P11940;E7EQV3	Polyadenylate-binding protein 1	PABPC1	32	21	2.3	5.0	48.7	21.0 ppm
11	Q96EY1;I3L1T6	DnaJ homolog subfamily A member 3, mitochondrial	DNAJA3	10	1	3.8	5.0	26.9	1.19 ppm
12	Q14257;F8WCY5	Reticulocalbin-2	RCN2	12	12	4.3	4.9	50.2	112 ppm
13	O95831;E9PMA0	Apoptosis-inducing factor 1, mitochondrial	AIFM1	13	13	4.2	4.8	32.6	688 ppm
14	P17066;P48741	Heat shock 70 kDa protein 6	HSPA6	19	2	3.6	4.7	20.5	104 ppm
15	Q9BUF5;K7ESM5	Tubulin beta-6 chain	TUBB6	23	10	4.0	4.7	67	183 ppm

Table 5.1D: *Significant proteins that interact with RESTV/NP-EGFP.*

	Protein IDs	Protein names	Gene names	Peptides	Unique peptides	-Log t-test P value	t-test Difference RESTV/NP-EGFP vs EGFP)	Sc [%]	Abundance (PPM)
1	P31948;G3XAD8	Stress-induced-phosphoprotein 1	STIP1	58	58	4.3	9.0	74	120 ppm
2	Q3ZCM7;F5H0I4	Tubulin beta-8 chain	TUBB8	13	2	4.9	7.5	25.2	175 ppm
3	O95816;B4DXE2	BAG family molecular chaperone regulator 2	BAG2	15	15	3.4	7.4	77.3	5.24 ppm
4	P17066;P48741	Heat shock 70 kDa protein 6	HSPA6	19	2	3.8	7.3	20.5	104 ppm
5	Q5JP53;P07437	Tubulin beta chain	TUBB	42	8	1.4	7.2	78.4	1041 ppm
6	P08238;Q58FF7	Heat shock protein HSP 90-beta	HSP90AB1	35	26	2.6	7.1	49	693 ppm
7	P31689;B7Z5C0	DnaJ homolog subfamily A member 1	DNAJA1	25	25	3.1	6.6	63.5	15.2 ppm
8	O95831;E9PMA0	Apoptosis-inducing factor 1, mitochondrial	AIFM1	13	13	5.4	6.3	32.6	688 ppm
9	P50502;H7C3I1	Hsc70-interacting protein;Putative protein FAM10A5	ST13	14	14	4.0	6.2	34.4	179 ppm
10	P34932	Heat shock 70 kDa protein 4	HSPA4	23	21	4.6	5.9	37.5	46.3 ppm
11	Q9BUF5;K7ESM5	Tubulin beta-6 chain	TUBB6	23	10	3.5	5.9	67	183 ppm
12	Q16342;F5H4V9	Programmed cell death protein 2	PDCD2	11	11	4.5	5.9	39	NA
13	Q9UNE7;H3BS86	E3 ubiquitin-protein ligase CHIP	STUB1	16	16	3.6	5.9	71	3.78 ppm
14	Q14257;F8WCY5	Reticulocalbin-2	RCN2	12	12	4.5	5.8	50.2	112 ppm
15	Q9Y230;B3KQ59	RuvB-like 2	RUVBL2	28	28	2.9	5.8	74.7	23.4 ppm
16	Q9BVA1	Tubulin beta-2B chain	TUBB2B	36	1	4.0	5.7	73.9	543 ppm
17	P07900;Q86U12	Heat shock protein HSP 90-alpha	HSP90AA1	21	13	3.8	5.7	36.3	841 ppm
18	P49411;H3BNU3	Elongation factor Tu, mitochondrial	TUFM	23	23	3.3	5.7	59.7	991 ppm

5.2.3 Validation of *Ebolavirus* NP interactions.

Possible interacting partners of *Ebolavirus* NP were then confirmed by western blot analysis. Reverse immunoprecipitations were also done to confirm the interaction between selected cellular proteins and *Ebolavirus* NP. Antibodies were used against some selected proteins which had the highest probability of interacting with EBOV/NP. In reverse immunoprecipitations the selected cellular proteins were immunoprecipitated and then the presence of EBOV/NP in the elution fraction was confirmed using western blot analysis validating the MS results.

Some of the selected cellular targets identified in the immune-precipitations of *Ebolavirus* NP were HSP70, HSP90, DNAJA2, PCNA, AIF and BAG2. These targets were selected due to their common functions, probability of interacting with *Ebolavirus* NP (proteins with high binding ratio), the availability of small molecules inhibitors and the fact that some of the proteins selected were identified before to interact with other NPs of other *Mononegavirales* (Lahaye et el., 2012; Zhang et al., 2005). The presence of these proteins was confirmed by western blot (*Figure 5.3*). Whole cell lysate (input) and eluate (bound) sample from immunoprecipitations analysis (GFP-Trap) were separated by 1D-SDS PAGE, proteins were then transferred in western blot and detected using a primary antibody against each selected protein. The data indicated that HSP70, HSP90, DNAJA2, PCNA, AIF and BAG2 associated with EGFP-EBOV/NP, EBOV/NP-EGFP, EGFP-RNP and RNP-EGFP. By western blot, HSP70 and BAG-2 showed to have a greater association for both *Ebolavirus*

NP; especially for EBOV/NP-EGFP and RNP-EGFP constructs. The cellular protein STAT1 was used as a negative control for the pull downs and mass spectrometry results as it was not identified by LC-MS/MS, but was identified in the input fractions (*Figure 5.3*).
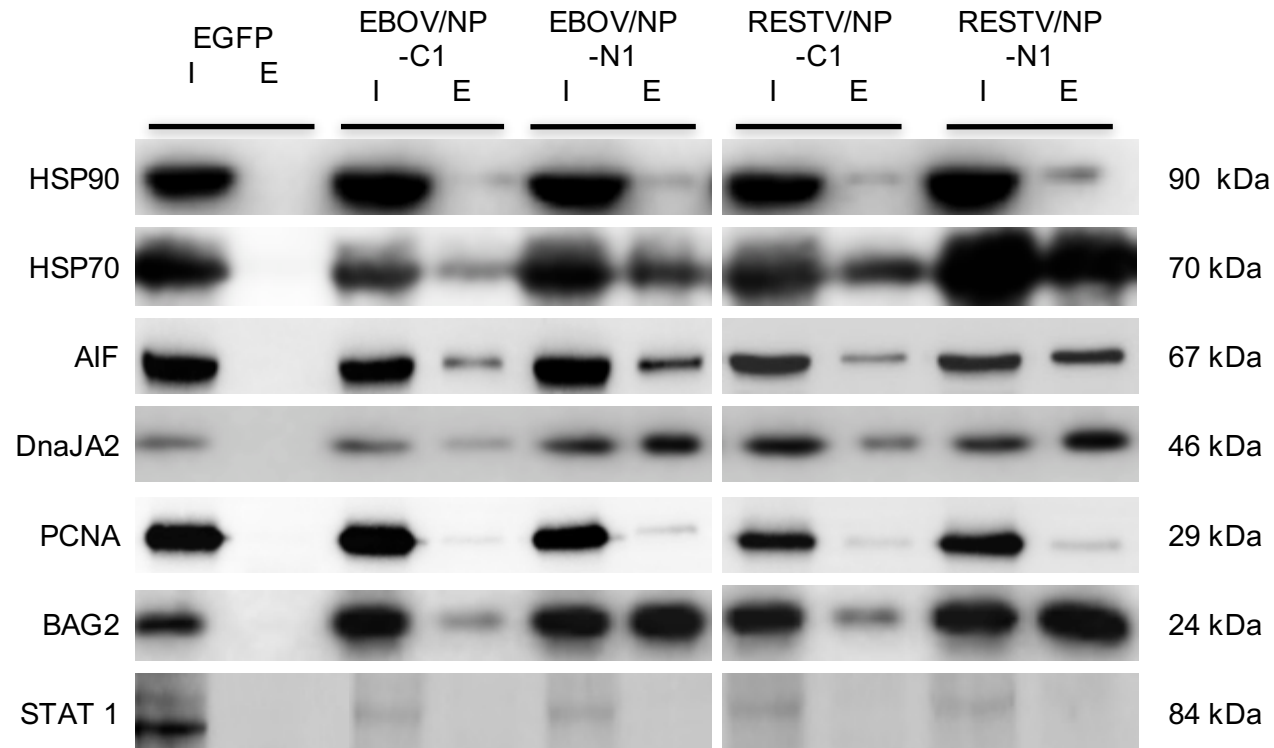
Figure 5.3. Validation of possible interacting partners of Ebolavirus NP by western blot analysis. *Confirmation of proteins detected in the label free mass spectrometry analysis by western blot analysis was done in the whole cell lysate or i nput samples (I) and in the eluate sample (E) for three different constructs. Specific monoclonal antibodies against HSP90, HSP70, AIF, DNAJA2, PCNA, BAG2 and STAT-1 were used to the detection of the mentioned proteins. The antibody anti-EGFP was used as a control to show the presence of the constructs EGFP-VP24 and VP24-EGFPin the input and in the elution samples.*

To corroborate the potential interaction partners for *Ebolavirus* NP and to determine whether these were mediated by RNA, input samples were treated with RNAse before immunoprecipitation (*Figure 5.4A*). After immunoprecipitation of the treated input samples, the bound samples were analysed again by western blot using specific antibodies against selected proteins HSP70, BAG2 and DNAJA2 (*Figure 5.4B*). The selected target proteins were found in the bound fraction after the RNAse treatment, which indicated that the interactions were not mediated by RNA.

A.

B.
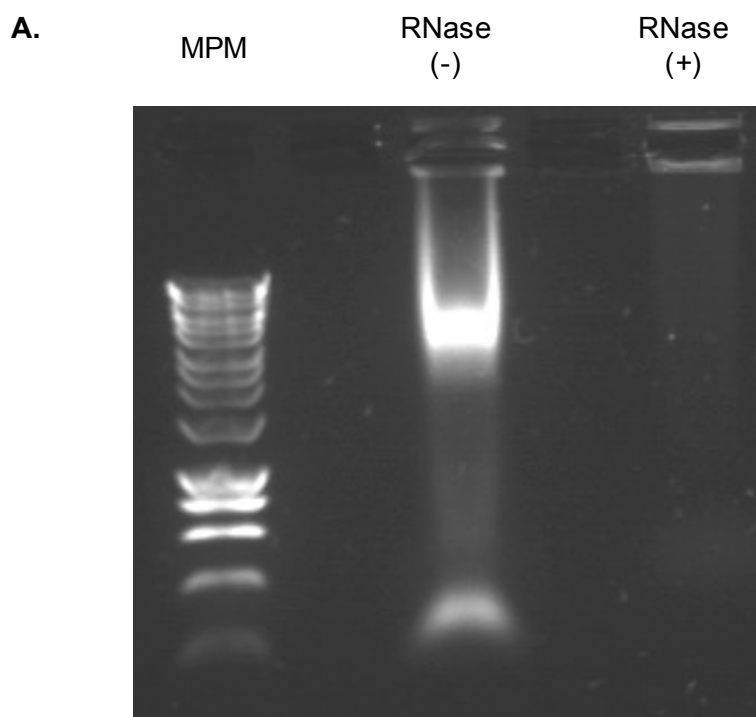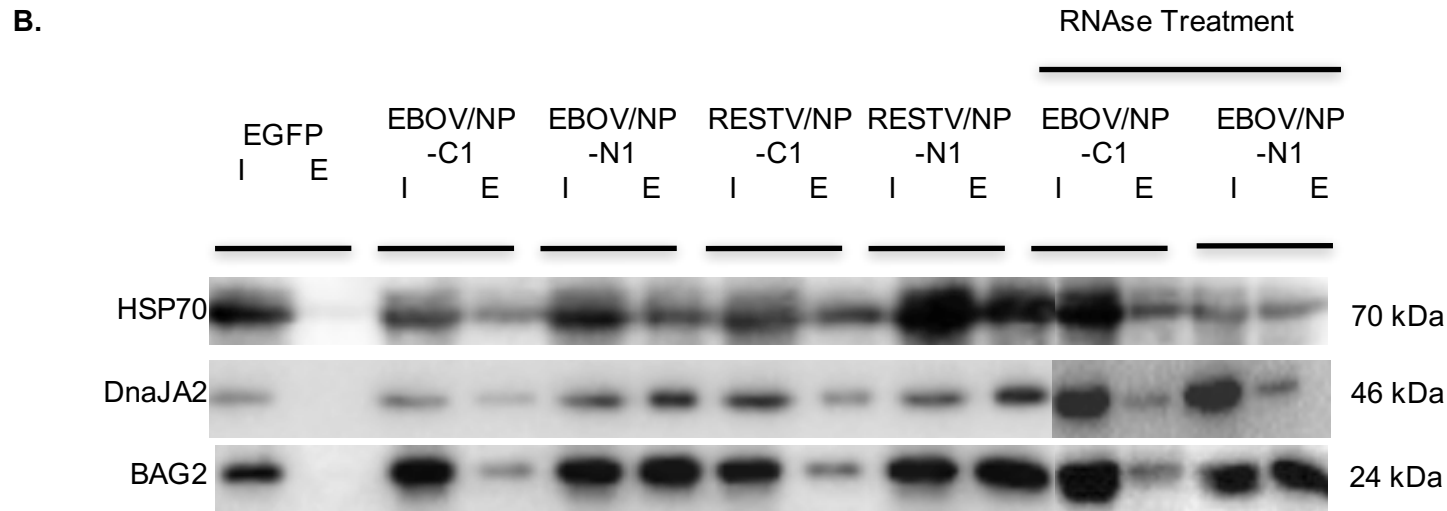
***Figure 5.4. Validation of Ebolavirus NP interactions using RNAse treatment and by western blot analysis. (A)** Agarose gel showing the RNAse treatment of the input samples before doing the immunoprecipitations. The first line shows the molecular marke (MPM) r, the second line show the band tat correspond the RNA in the untreated samples or RNase (-) ; the last line shows no RNA band due to the treatment with RNAse or RNase (+) **(B)** Confirmation of proteins detected in the label free mass spectrometry analysis RNAse treatment of the input sample following by immunoprecipitation and western blot analysis. The Input samples (I) and in the eluate sample (E) for the five different constructs were shown with the RNAse treated Input samples and the elution samples for the RNAse treatment. Specific monoclonal antibodies against HSP70, DnaJA2 and BAG2 were used to the detection of the selected proteins.*

To further validate the mass spectrometry results and the potential interacting partners of EBOV⁄NP, reverse immunoprecipitation was done against selected cellular targets where antibody combinations allowed. These targets were EBOV-NP (as a positive control), HSP70 and BAG2. In order to do the reverse immunoprecipitations, *Ebolavirus* NPs were overexpressed in 293T cells and then cellular proteins were extracted. Reverse immunoprecipitations were then performed using protein G beads and specific monoclonal antibodies against the selected protein targets EBOV-NP, HSP70 and BAG2. The bound samples obtained from the reverse immunoprecipitations were then analysed by western blot, which confirmed the presence of EGFP-EBOV/NP, EBOV/NP-EGFP and EGFP in the input and elution fraction, this was done using a specific antibody against EBOV-NP and EGFP (*Figure 5.5*). This confirmed and validated the results from the forward immunoprecipitation approach.

A. Reverse pull down for EBOV/NP
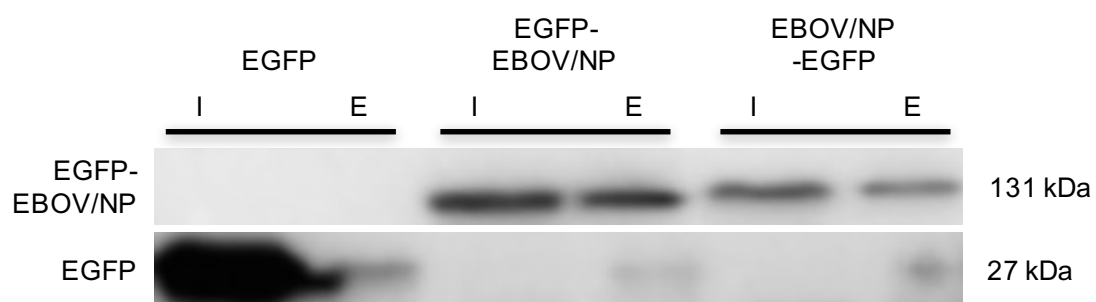
B. Reverse pull down for HSP70.
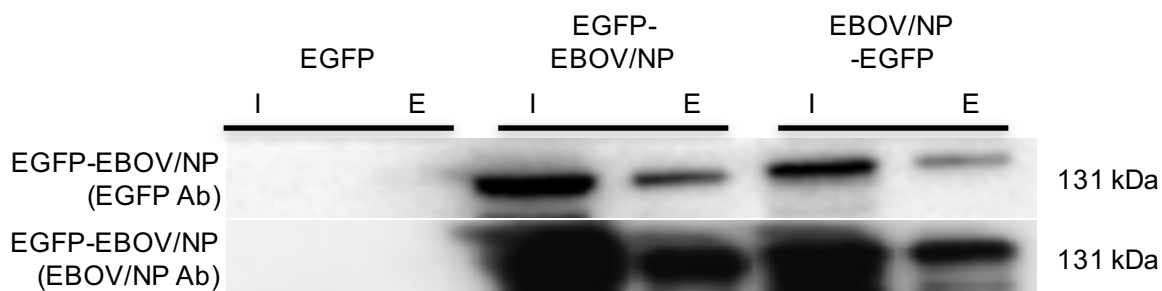
C. Reverse pull down for BAG2.
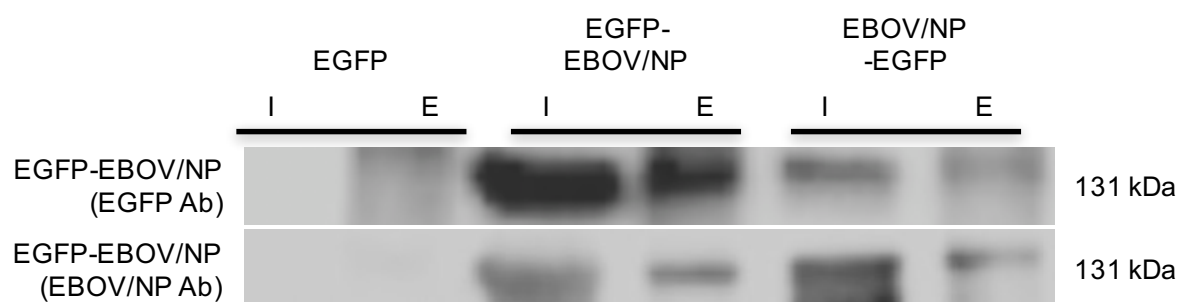
Figure 5.5. Validation of EBOV/NP interactions using reverse immunoprecipitation. *To further validate the mass Spectrometry results, reverse immunoprecipitation was preformed against selected cellular proteins identified by the label free mass spectrometry and the statistical analysis. The selected proteins for the validation of the label free mass spectrometry were EBOV/NP (as a positive control), HSP70 and BAG2; the presence of the protein complex EBOV/NP were confirmed using a western blot analysis and using a specific antibody against EGFP (for reverse pull down for EBOV-NP), an EBOV/NP and EGFP antibody in for the reverse pull down for HSP70 and BAG2.*

5.2.4 Development of a mini-genome system for EBOV and its use to test small molecular inhibitors

Based on the essential role of NP in viral RNA synthesis, the hypothesis that protein chaperones are important for this role was tested. To dissect this precise function and recapitulate this safely at CL-2, a mini-genome system for EBOV was developed. The mini-genome system for *Ebolavirus* is a well known method to study viral transcription and replication. Furthermore, it is a very convenient way to test new compounds against EBOV; due to the safety of this method, the use of high containment facilities is not necessary (Hoenen et al 2011; Uebelhoer et al 2014). The mini-genome system for Ebola virus (Makona strain) used in this study was designed following a similar design that was used by Mülhgerber in 1999, specially for the support plasmids (Mülhgerber et al.,1999). For the mini-genome plasmid some modifications were made; an example of these modifications is the use of another reporter gene for the mini-genome plasmid (in this case we used luciferase) (see Chapter 2).

The EBOV mini-genome system used in this study consists of four support plasmids that express VP30, VP35, L and NP; plus, a mini-genome plasmid expressing an artificial viral genomic RNA that contains the leader and trailer sequences from EBOV with the luciferase (luc) gene as a reporter gene. All the plasmids used for the mini-genome system were under control of the T7 promoter (Chapter 2). The amount of each plasmid used for the mini-genome system to produce luciferase was optimized, as well as the time for the

expression of luciferase to obtain the maximum signal-to-noise ratio (*Table 5.3; Figure 5.5*). Once the time and the amounts of each plasmid were optimized the mini-genome system was used to test the hypothesis that NP interacted with host protein chaperones to facilitate synthesis of viral RNA.

The appropriate concentration of plasmids (*Table 5.3*) that formed the mini-genome system were transfected to BSR-T7 using lipofectamine 2000 (lipid mediated transfection); 24 and 48 hours post transfection the whole cell lysates were extracted and then analysed using by western blot. To detect the luciferase produced in the mini-genome system an antibody against *Firefly luciferase* was used (*Figure 5.6*). As control, cells transfected with all the plasmids apart from L was used, as well as cells treated with only Lipofectamine to assess the toxicity effect of Lipofectamine reagent in cells. An antibody against GAPDH was used to confirm the loading of equal amounts of the samples *(Figure 5.6c).*

Table 5.2. Optimization of amount of plasmids for the mini-genome system: Table showing the different concentrations used in order to optimize the expression of Luciferase for the mini-genome system; several controls were used (columns 1 to 4) and also different concentration of plasmids were tried for the optimization; the condition number 9 was the best and therefore selected to the following experiments.

EBOV Makona Mini-genome											
(μg)	1	2	3	4	5	6	7	8	9	10	11
Mini G	0	0.5	0.5	0	0.25	0.25	0.5	0.5	0.5	0.5	0.5
N	0	0.25	0.25	0.25	0.125	0.25	0.125	0.25	0.25	0.25	0.25
VP35	0	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125
VP30	0	0.125	0	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125
L	0	0	0.125	0.125	0.125	0.125	0.125	0.0625	0.125	0.25	0.5

A.

B.

Fold change	*/Mock
Mock	1
No L	0.265741014
With L	224.9873404
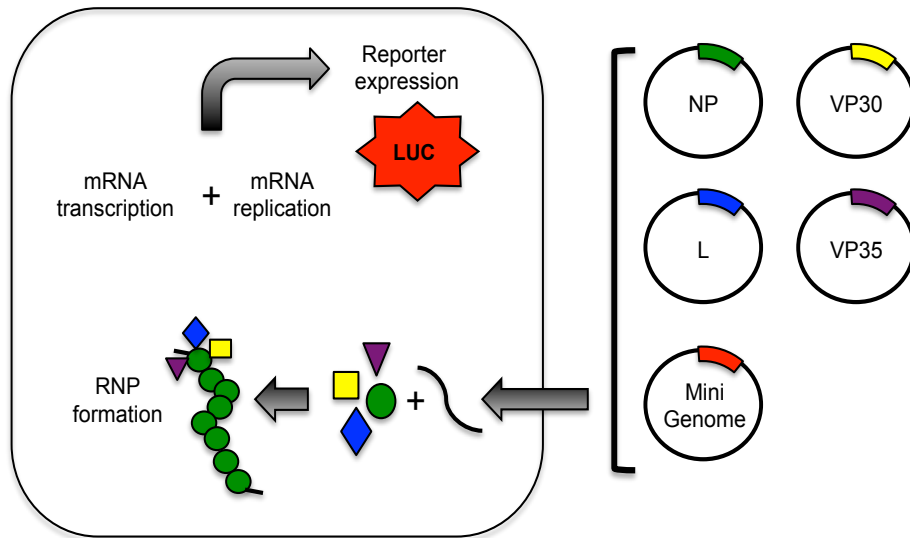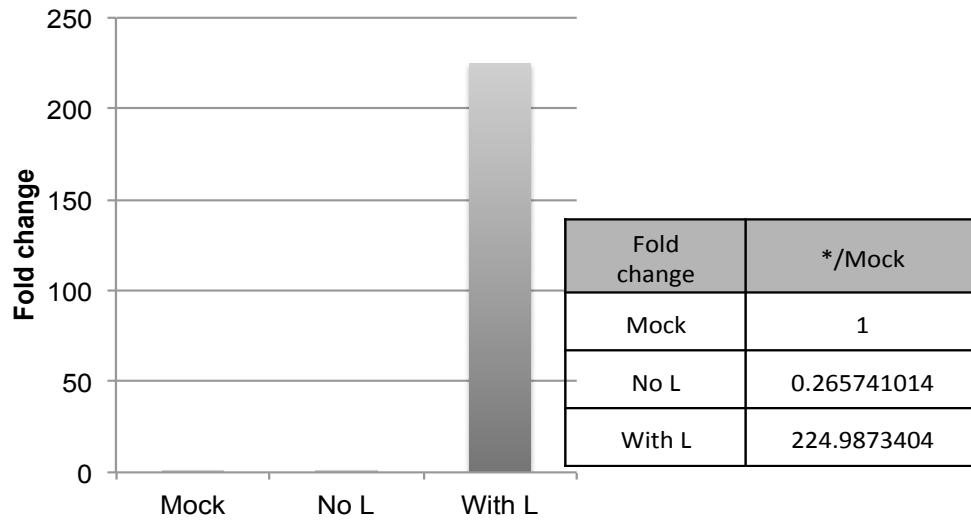
Mini Genome Makona EBOV

C.
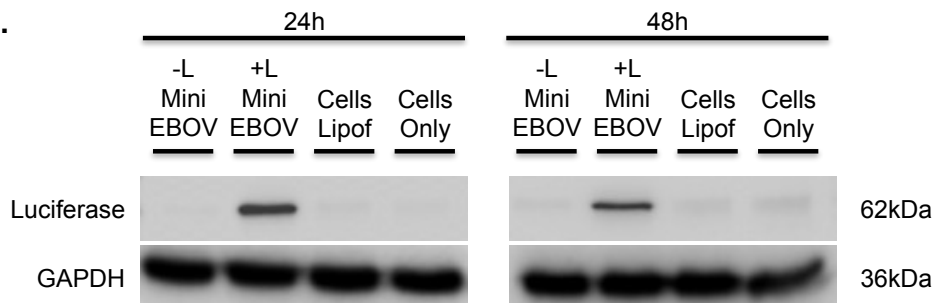
Figure 5.6 Optimization of amount of plasmids for the mini-genome system. (A) *Schematic representation of the mini-genome system. The EBOV (Makona) mini-genome system consist of five plasmids: four supportive plasmids that express the NP, VP30, VP35 and L; and the mini-genome plasmid which contains the leader and trailer sequences of EBOV (Makona strain) and has sequences that encode Luciferase as a reporter gene.* ***(B)*** *Measure of the luciferase expression using the Dual Luciferase assay (Promega).* ***(C)*** *Western blot of luciferase expressed by the mini-genome system at 24 and 48 hours; for the western blot analysis a highly specific antibody against Firefly Luciferase was used.*

5.2.5 The inhibition of HSP70 affects the transcription and replication of EBOV

Heat shock proteins are essential components of the cellular network of molecular chaperones, assisting a wide range of folding processes including folding of newly synthesized proteins. The HSP70 and its co-chaperones are essential for a proper folding process (Mayer et al 2005a). The importance of the interaction between viral proteins and cellular chaperones has been described by several studies (Taguna et al., 2015; Xiao et al 2010; Mayer et al 2005b). In order to determine if the inhibition of HSP70 affected viral replication and transcription, a mini-genome system was used. The mini-genome designed for this study expressed Luciferase as a reporter gene (as was described above).

For the inhibition assay, VER-155008, a small inhibitor compound against HSP70 was selected. VER-155008 targets the ATPase binding domain of HSP70 competing with ATP and preventing the binding of ATP to the nucleotide binding site and therefore prevents the interaction of HSP70 with the target protein (Wen et al., 2014; Williamson et al., 2009). Different concentrations of VER-155008 were used according to similar previous studies with other viruses from the Order *Mononegavirales* (Munday et al., 2015).

An MTT assay was also done in BSR-T7 cells to determine the maximum concentration permitted for VER 155008 without inducing cytotoxic effects was

60 μM, which had no apparent effect was shown on cell viability compared with the untreated or vehicle only controls. (*Figure 5.7A*); the concentrations used in this study for VER-155008 were 5μM, 10μM, 20μM, 40μM, 60μM, 80μM and 100 μM. BSR-T7 cells were transfected with the mini-genome and a plasmid control expressing *Renilla* luciferase; after 4 hours post transfection the media was removed and new media with the different concentrations of the drug were added. Twenty-four hours post transfection proteins were extracted from cells and then the dual Luciferase assay (Promega) was used to measure the luciferase activity expressed in each condition. The results showed a decrease in the Luciferase expressed by the EBOV mini-genome with increased concentration of VER-155008 (*Figure 5.7B*). *In the x-axes we can see the different concentration of the drug used in the experiment; in the y-axes the fold change of Luciferase expression compared to the control (Figure 5.7B).* Similar results were observed when the lysed samples were analysed by western blot (*Figure 5.7C*).

A.

MTT VER-15508 at 24h in BSRT7 cells

B.

Effect of VER-155008 on Luciferase expression

C.
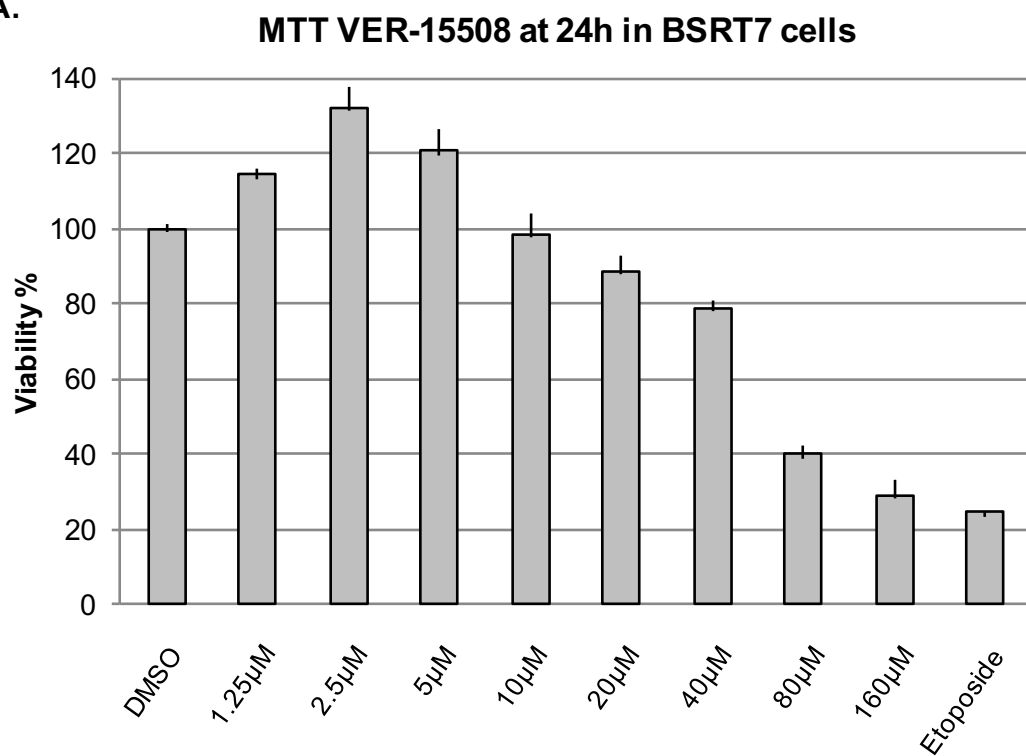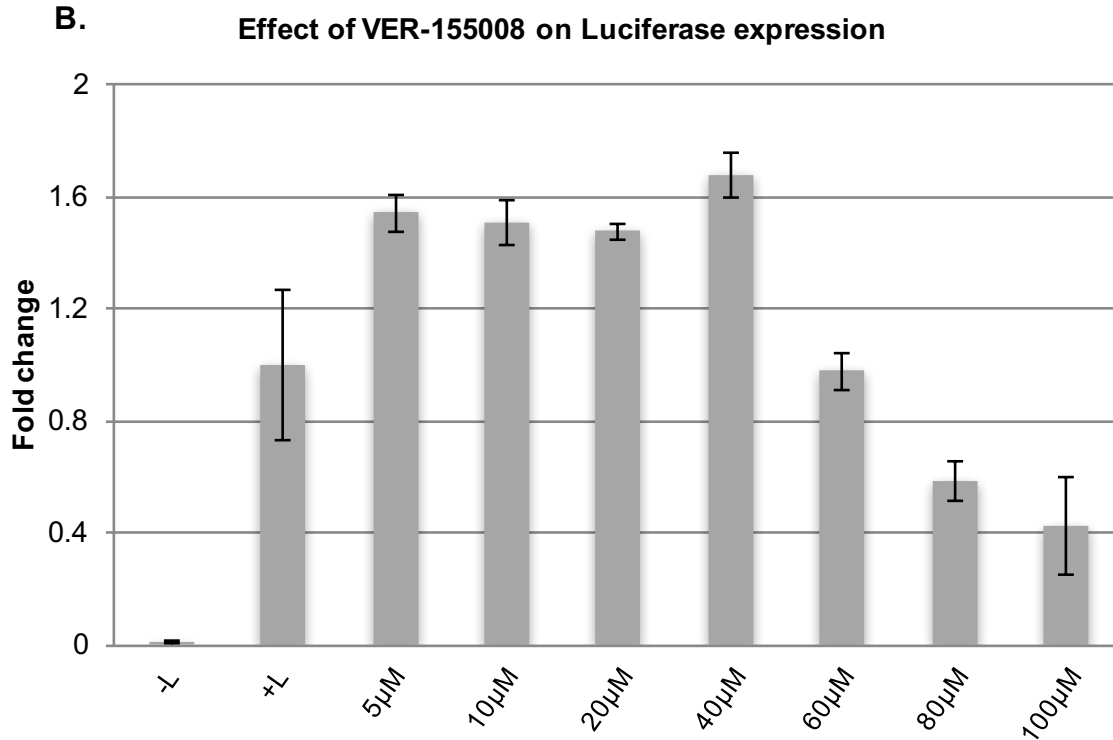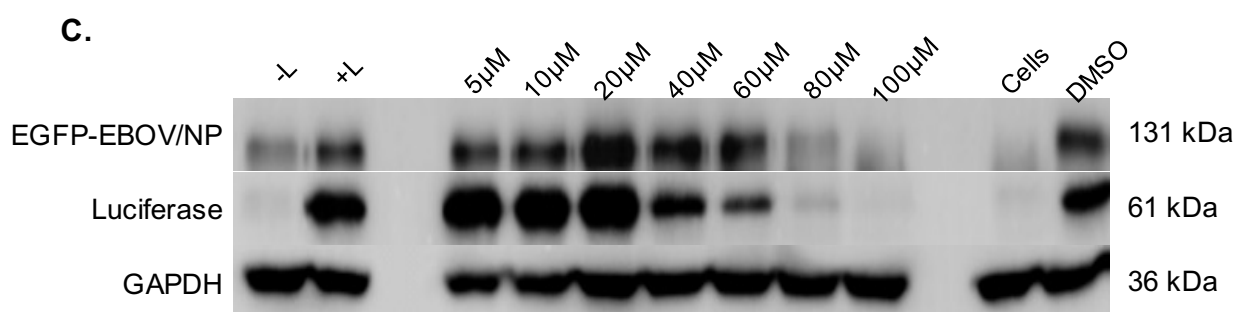
Figure 5.7. Effect of the HSP70 inhibitors in the replication and transcription of Ebola virus. A mini-genome system for Ebola virus that express Luciferase was used to measure the effect of VER-155008 to see their effect in the replication-transcription of Ebola virus. (A) An MTT assay was used to determine cell viability in the absence of drug, the vehicle only control (DMSO), and various concentrations of VER-155008. (B) For the experiment with the mini-genome system, seven different concentrations of VER-15508 were used for the inhibition experiment. The Dual Luciferase Assay (Promega) was used to measure the amount of Luciferase produce for the mini-genome system. (C). Results of these experiments confirm by western blot analysis using a highly specific antibody against Firefly Luciferase.

In order to further validate this results, the effect of VER-155008 in the expression of EBOV-NP was tested. The prediction being that in the absence of chaperone activity NP should be destabilized and potentially targeted for degradation. To do this EGFP-EBOV/NP, EBOV/NP- EGFP and EGFP plasmids were transfected in BSR-T7 cells and treated with the different inhibitors at similar concentrations as have been done for the mini-genome experiment to see if the expression of the EBOV-NP is affect by the inhibitors (Figure 5.8). In the presence of VER-155008 the expression of EBOV-NP was decreased.
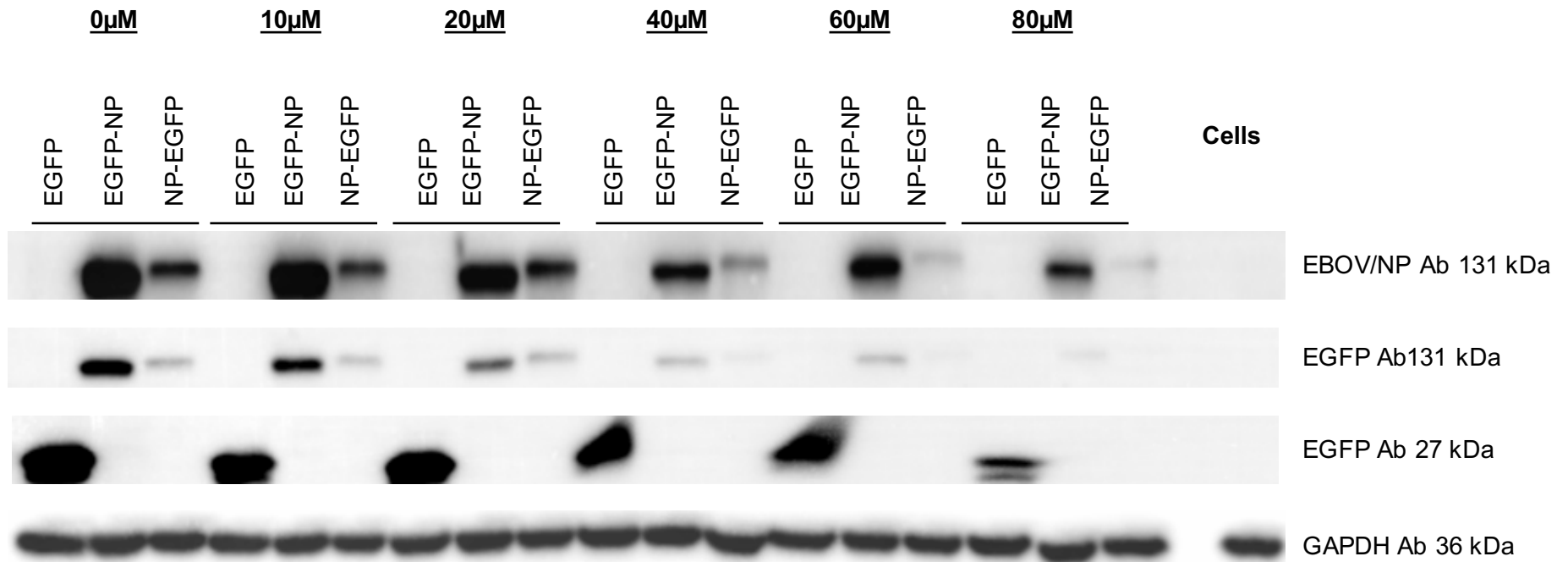
Figure 5.8. *Effect of VER-155008 on the expression of EBOV/NP in BSR-T7 cells. Different concentrations of VER-155008 were tested to see its effect in the expression of EBOV/NP. A decrease in the expression of EBOV/NP was observed when the cells were treated with VER-155008.*

5 .3   Discussion

There is little information available regarding the interactions of *Ebolavirus* NP with cellular proteins. Elucidation of these interactions will provide a better understanding of the function of *Ebolavirus* NP in virus biology and also will provide highlights of cellular targets for potential novel antivirals. The study of interacting partners of different viral proteins has been used before in order to identify new targets for drug therapy against Ebola virus and other RNA viruses (Garcia et al., 2014; Munday et al., 2015; Wu et al., 2012; Emmott et al., 2013). Some of the proteins found in this study such as HSP90 have been proven to be important for the *Ebolavirus* biology (Smith et al., 2010). However, most of the interacting partners of *Ebolavirus* NP found in this study have been never described before, some examples are: HSP70, BAG2, DNAJA2, AIF and PCNA. Furthermore, some of the interacting partners of *Ebolavirus* NP found in this study have been found to interact with NPs of other viruses that belong to the order *Mononegavirales.* For example, the NP of rabies virus and measles virus (Lahaye et el., 2012; Zhang et al., 2005), thus indicating a general pattern.

Among all the interacting partners of *Ebolavirus* NP found in this study, the cellular chaperone HSP70 and its co-chaperon BAG2 were selected for further analysis and validation of their interaction, first by reverse pull down analysis (HSP70 and BAG2) and by using specific small molecules inhibitors (for HSP70). HSP70 is involved in several cellular processes including protein

folding processes, regulation of stress response, and control of the activity of regulation proteins among others (Mayer et al., 2005a). Due to the several functions in the cells, HSP70 may also play an important role in the viral survival strategies (Meyer et al., 2005b, Xiao et al., 2010; Kim et al., 2012) such as degradation of certain viral proteins (Katoh et al., 2014; Lahaye et al 2012); envelope protein/folding maturation (de Silva et al 1990) among others. The production of large amounts and accumulation of viral proteins during the virus life cycle make chaperones essential for the production of viral proteins and also can provide new possible drug target therapy against viruses (Geller et al., 2012; Smith et al., 2010; Necker et al., 2008). A recent study of HSP70 in relation with dengue virus showed that HSP70 is crucial playing an important role in virus biology (Taguwa et al., 2015).

In this study, VER-155008, a small molecule inhibitor against HSP70 was tested to study the effect of HSP70 inhibition in the replication of the *Ebolavirus* using a mini-genome system. The mini-genome system has been used before in other studies to identify new inhibitors for Ebola viruses (Uebelhoer et al., 2014), this kind of approach has been very useful for research because it can be used in BSL-2 facilities. Treatment with VER-155008, a small molecule inhibitor against HSP70, showed an effect in the expression of luciferase reporter gene in the EBOV mini-genome system. Treatment with VER-155008 showed a reduction in the expression of luciferase when increasing the concentration of the drug. Similar results were found for RSV (another member of the *Mononegavirales* family); in this case, four inhibitors against HSP70 including VER-155008 among others, but only VER-155008 showed a

reduction in the replication of RSV (Munday et al 2015). Another example is Dengue virus where also small doses of VER-155008 was showed to affect the virus replication and therefore progeny virus production; in this study HSP70 was showed to be required for Dengue virus life cycle (Taguwa et al., 2015).

VER-155008 interacts with the ATPase domain of the Hsp70 and therefore blocks ATP binding to HSP70 (Schlecht et al., 2013). Therefore, these results indicated that EBOV-NP might interact near to the ATPase domain of HSP70 and the interaction of EBOV/NP with HSP70 can be disrupted by VER-155008. However, more studies are still need in order to look into this further. Taken together, the fact that HSP70 inhibition affects EBOV RNA synthesis confirms that the interaction of EBOV-NP with HSP70 might be important for the virus biology. The HSP70 inhibitor, VER 155008, has been shown also to affect the stability of EBOV/NP; showing to have negative effect on EBOV/NP stability, as the doses of the compound are increased, the amount of EBOV/NP expressed decrease. This negative effect of VER 155008 in the stability of EBOV/NP can be the reason why the EBOV mini-genome is affected by VER 155008 and therefore the reduced expression of luciferase by the EBOV mini-genome system.

This study showed that the interaction of EBOV-NP with chaperones (HSP 70, HSP 90) and co-chaperons (BAG-2, DNAJA2) might play an important role in the virus biology and, for that reason, further studies in the role of these cellular

chaperones and co-chaperones in the Ebolavirus biology are discussed in the next chapter. Overall the elucidation of the *Ebolavirus* NP interactome allowed the discovery of new possible drug targets therapy for Ebola viruses..

CHAPTER 6: Final discussion and future perspectives.

EBOV is one of the most pathogenic virus in the order Mononegavirales, causing viral haemorrhagic fever in humans and non-human primates. Due to the lack of approved treatment or vaccine and its high pathogenicity this virus has been classified as a CL4 pathogen. The recent EBOV outbreak in West Africa has been the biggest outbreak since its discovery, causing more than eleven thousand deaths and lasting more than two years. This outbreak highlighted the urgent and unmet needed of antivirals against this virus and effective vaccines. The virus has an RNA genome and no real evidence of error correction during the replication cycle. EBOV is also thought to be a zoonotic infection, so how the virus adapts to grow into a new host needs to be elucidated.

Currently licenced antivirals are targeted against viral proteins to inhibit their function. However, experience with treating HIV and influenza virus demonstrates that resistant viruses can soon be selected. An emerging area in virology is to transiently target host cell proteins that play critical pro-viral roles in virus biology, especially for acute infections. This has the advantage that the protein being targeted is evolutionary removed from the genome of the virus. Several groups have shown that, at least in vitro, selection for resistant genotypes does not occur when host cell proteins have been targeted (Xiao et al., 2015). Also, many of these therapeutics may have been already licenced for use in humans and thus could be rapidly be repurposed as anti-viral compounds.

In order to identify and characterise the role of host cell proteins in the biology of EBOV and to select potential drugable targets, the interactome of two EBOV proteins, VP24 and NP was defined. The study of the virus interactome has become very popular in the virology field in the last few years, through advances in pulldowns and mass spectrometry. Viral proteomics has been used extensively to provide a better understanding of the virus biology of several viruses, including RSV (Wu et al., 2012; Munday et al., 2015), IBV (Emmott et al., 2013), among others (Pilchmair et al., 2012). Viral proteomics has been also used as a tool to look for new targets for inhibitors that can disrupt the virus life cycle (Munday et al., 2012; Maxwell and Frappier., 2007).

The thesis was divided into three main parts: the study of viral evolution in an animal model and in a human population during the West African outbreak (Chapter 3); the study of the EBOV VP24 interactome (Chapter 4) and the study of the EBOV NP and RESTV NP interactome (Chapter 5).

In the first part of Chapter 3 EBOV viral evolution was studied in a guinea pig model. This model has been used before by other researchers (Volchkov et al., 2000). However, in these other studies, an RNA sequence analysis of the last passage in the animal experiment was performed, which offered limited information of the EBOV evolution during sequential passage to gain virulence. Therefore, it was not possible to determine in which passage these mutations were established, and if these mutations appeared in conjunction with the symptoms in the animal model (Volchkov et al., 2000; Mateo et al., 2011; Conolly et al., 1999).

In this work, sequential passages of EBOV were performed in guinea pigs until the virus became pathogenic for this host (5 passages); then, an RNA extraction and RNA sequence analysis was performed during each passage. After that, an analysis of the amino acid sequence was done to detect and compare mutations that become established by the end of the five passages or in earlier passages. An example of an established mutation was the amino acid position 26 in the EBOV VP24, which appeared in early passages and remained stable along all passages till passage five. This mutation has been previously identified and has also linked with an increase in virulence in the guinea pig model by using reverse genetics (Mateo et al., 2011; Volchkov et al 2000; Conolly et al., 1999). Thus establishing the approach of examining viral sequence at each passage. Future work in the area would involve reverse genetics to test the pathogenicity of mutations found in other EBOV genes during passage including those in the polymerase. Certainly, these could also be evaluated using a mini-replicon system similar to that described in Chapter 5, but centred on replication in guinea pig cells.

In the second part of Chapter 3, a study of the genetic evolution of EBOV in a human population was performed looking for amino acid mutations in the different EBOV proteins during the EBOV outbreak in West Africa (2013-2015). In order to do this, several samples from Guinea were taken at different times during the outbreak, which I was involved with by deploying as part of the European Mobile Laboratory. This study was one of the first studies showing amino acid mutations for the different EBOV proteins during the course of the outbreak (Carroll et al., 2015). In this part of the study, most of the viral proteins

showed mutations, some examples of these proteins are: NP, GP, L and VP35. All these proteins play a crucial role for the EBOV biology. Therefore, it will be very interesting to further analyse the role of these mutations in EBOV pathogenesis, particularly through the use of the mini-genome and ultimately reverse genetics. It will also very interesting to determine if any of these mutations are present, and to what extent, in patients that survived the disease, or if these they are strictly found in patients experiencing a fatal outcome. Particularly to assay changes in the polymerase complex proteins, that can be assessed rapidly using the mini-genome.

In Chapter 4, the EBOV VP24 interactome of was studied. This protein was selected due to the many important functions that it plays in EBOV biology and also for its importance in acquiring pathogenesis in the guinea pig model (Volchkov et al 2001), which was also demonstrated in Chapter 3 (Dowall et al., 2014). The current study is the most complete analysis of the EBOV VP24 interactome. Other studies present a more general approach of the interactome of various viral proteins, including EBOV VP24 (Pichlmair et al., 2012). Nevertheless, those works only presented a general idea of the possible EBOV VP24 cellular interactome, and did not confirm any of the possible interacting partners with further validations.

The current study for the EBOV VP24 interactome presented more than 50 possible cellular proteins with a high probability of interacting with EBOV VP24 including KPNA6, which has been already described to interact with EBOV VP24 (Xu et al., 2014), and thus validated the approach. Other nuclear import

factors were identified suggesting that VP24 generally targets these molecules to disrupt the STAT1 activated signalling cascade. This work also found no interaction between STAT1 and VP24, which had been postulated by one group, based on structural predictions, to mediate STAT1 inhibition (Zhang et al., 2012). Thus, the work described in this thesis and the accompanying publication (Dowall et al. 2014) favours the model in which VP24 out competes STAT1 for binding to nuclear import factors (Reid et al., 2007).

Among these proteins that interacted with VP24, ATP1A1 was selected for an inhibition study due the high probability that this protein interacts with EBOV VP24, and the availability of a highly specific and commercially available small molecule inhibitor against ATP1A1. After several validations of the interaction EBOV VP24 and ATP1A1, an inhibition analysis was done. To do this, a small molecule inhibitor against ATP1A1, ouabain, was used. Ouabain was shown to affect the viral biology by decreasing the amount of EBOV viral RNA in cells infected with EBOV. A decrease in EBOV RNA was observed as the amount of ouabain increased. Therefore, it was concluded that the inhibition of ATP1A1 can affect the EBOV life cycle (Garcia et al., 2014). This is the first study that showed the inhibition of a cellular interacting partner of EBOV VP24 can be used to disrupt the EBOV viral biology and therefore be used as a possible therapeutic treatment. Preliminary experiments by our collaborators at PHE applied this drug in the guinea pig model of infection. Whilst the compound maintained anti-viral activity *in vivo*, several other compounds displayed better activity during the down select process and were taken

forward. Nevertheless, this demonstrates the general application of identifying host cell proteins that are critical for virus biology and inhibiting their function.

In Chapter 5, the interactome of EBOV NP and RESTV NP were analysed. This viral protein was selected due to its important function in forming the RNP, and because it is one of the components of the mini genome system for EBOV, and thus its function can be rapidly assessed. The EBOV mini genome system was used to test potential inhibitors against EBOV that targeted the cellular proteins associating with NP. EBOV NP was also selected because it was one of the proteins with a higher percentage of mutations in the viral evolution analysis in the previously described analysis of EBOV evolution in the West African population (Chapter 3). This study also investigated and compared the cellular interacting partners of NP for two species of the *Ebola viruses*, EBOV and RESTV, with this being a novel area of investigation. There was not a large difference in the interacting partners that were identified for both proteins (EBOV NP and RESTV NP). Furthermore, most of the cellular interacting partners were similar for both proteins with the only difference being the amount of cellular protein detected that interacts with each viral protein. Chaperones (HSP70, HSP90) were showed to play an important role for the *Ebolavirus* NP, as these proteins and their co-chaperones (BAG-2; HSP40) were detected as an interacting partners for both *Ebolavirus* NP. However, this is the first time HSP70 has been described as an interacting partner for *Ebolavirus* NP.

HSP90, another cellular chaperone has previously shown to be important for EBOV virus biology, and was postulated to associate with the L protein (Smith et al., 2010). However, this is the first time that HSP70 has been shown to be important for EBOV virus biology and also the first time to be described as a possible interacting partner for *Ebolavirus* NP. In order to determine the importance of HSP70 for the EBOV biology, a specific molecular inhibitor against HSP70, VER-155008 was selected for the inhibition analysis. This time, an EBOV mini genome system was used to investigate the effect of the small molecule inhibitor in viral replication. The mini genome system has been used by several groups to identify new possible inhibitors against EBOV (Uebelhoer et al., 2014). However, this is the first time a mini genome system was developed against the new strain of EBOV, Makona EBOV, which was responsible for the EBOV outbreak in West Africa in 2013-2015. The results showed that the inhibition of the chaperone HSP70 can affect the viral replication of EBOV, and therefore presents new therapeutic approaches as many of these compounds are in phase II trials as anti-cancer therapies. The work also adds to the growing body of evidence that cellular chaperones play critical roles in maintaining the structure and function of viral proteins (Munday et al., 2015; Taguwa et al., 2015).

Overall, this study showed the application of a highly specific system, EGFP-trap for pull down coupled to LC-MS/MS to study EBOV proteins possible interacting partners. This, along with a high number of samples analysed and the statistical analysis together with multiple validations done in this study

ensured a high coverage in the identification of possible cellular interacting partners for selected *Ebolavirus* proteins.

As it was discussed before, future studies need to be done in order to better understand the impact and function of the different interactomes of each EBOV viral proteins. Likewise, it will very interesting to compare the interactome of specific viral proteins for different species of Ebolaviruses and if this changes in the interactome may contribute to the severity of the disease. An example of this will be to compare the interactome of the Ebolavirus viral proteins for EBOV and RESTV.

Further studies need to be done as well in order to fully understand the importance of mutations in the viral proteins of the EBOV in patients who died or survived from the disease. To examine and understand these mutations is important for the development of possible new drugs for therapies and also to study the number of mutations that are necessary for an increase in virulence. This will help to select more effectively possible treatments against the disease to avoid drug resistance.

In summary, this thesis presented a new sequencing pipeline approach that was developed in an animal model for EBOV, and was then applied in the EBOV outbreak. This gave us the possibility to analyse and evaluate the different amino acid mutations in the EBOV proteins in order to see the importance of these mutations during an outbreak. This thesis also has given a new insight into the study of the cellular interacting partners for *Ebolavirus*

proteins. This thesis provides evidence that the study of the cellular interacting partners for viral proteins can be crucial for the development and research of new possible drug therapies against viral infections, and obviously provides additional information on the biology of the virus.

CHAPTER 7: Bibliography

- Adu-Gyamfi, E., Soni, S. P., Xue, Y., Digman, M. A., Gratton, E., & Stahelin, R. V. (2013). The Ebola Virus Matrix Protein Penetrates into the Plasma Membrane a key step in viral protein 40 (vp40) oligomerization and viral egress. *Journal of Biological Chemistry*, *288*(8), 5779-5789.

- Altamirano, A. A., Fons, M. P., Russell, J. M., Cragoe, E. J., & Albrecht, T. (1994). Human cytomegalovirus infection increases the number of ouabain-binding sites in human fibroblasts. *Virology*, *199*(1), 151-159.

- Alvarez, C. P., Lasala, F., Carrillo, J., Muñiz, O., Corbí, A. L., & Delgado, R. (2002). C-type lectins DC-SIGN and L-SIGN mediate cellular entry by Ebola virus in cis and in trans. *Journal of virology*, *76*(13), 6841-6844.

- Ascenzi, P., Bocedi, A., Heptonstall, J., Capobianchi, M. R., Di Caro, A., Mastrangelo, E., ... & Ippolito, G. (2008). Ebolavirus and Marburgvirus: insight the Filoviridae family. *Molecular aspects of medicine*, *29*(3), 151-185.

- Audet, J., Wong, G., Wang, H., Lu, G., Gao, G. F., Kobinger, G., & Qiu, X. (2014). Molecular characterization of the monoclonal antibodies composing ZMAb: a protective cocktail against Ebola virus. *Scientific reports*, *4*.

- Baden, L. R., Kanapathipillai, R., Campion, E. W., Morrissey, S., Rubin, E. J., & Drazen, J. M. (2014). Ebola—an ongoing crisis. *New England Journal of Medicine*, *371*(15), 1458-1459.

- Baize, S., Pannetier, D., Oestereich, L., Rieger, T., Koivogui, L., Magassouba, N. F., ... & Günther, S. (2014). Emergence of Zaire Ebola virus disease in Guinea. *New England Journal of Medicine*, *371*(15), 1418-1425.

- Barr, J., Chambers, P., Pringle, C. R., & Easton, A. J. (1991). Sequence of the major nucleocapsid protein gene of pneumonia virus of mice: sequence comparisons suggest structural homology between nucleocapsid proteins of pneumoviruses, paramyxoviruses, rhabdoviruses and filoviruses. *Journal of general virology*, *72*(3), 677-685.

- Basler, C. F., & Amarasinghe, G. K. (2009). Evasion of interferon responses by Ebola and Marburg viruses. *Journal of Interferon & Cytokine Research*, *29*(9), 511-520.

- Basler, C. F., Mikulasova, A., Martinez-Sobrido, L., Paragas, J., Mühlberger, E., Bray, M., ... & García-Sastre, A. (2003). The Ebola virus VP35 protein inhibits activation of interferon regulatory factor 3. *Journal of virology*, *77*(14), 7945-7956.

- Basler, C. F., Wang, X., Mühlberger, E., Volchkov, V., Paragas, J., Klenk, H. D., ... & Palese, P. (2000). The Ebola virus VP35 protein functions as a type I IFN antagonist. *Proceedings of the National Academy of Sciences*, *97*(22), 12289-12294.

- Biacchesi, S., Skiadopoulos, M. H., Boivin, G., Hanson, C. T., Murphy, B. R., Collins, P. L., & Buchholz, U. J. (2003). Genetic diversity between human metapneumovirus subgroups. *Virology*, *315*(1), 1-9.

- Biedenkopf, N., Hartlieb, B., Hoenen, T., & Becker, S. (2013). Phosphorylation of Ebola virus VP30 influences the composition of the viral nucleocapsid complex impact on viral transcription and replication. *Journal of Biological Chemistry*, *288*(16), 11165-11174.

- Boulon, S., Ahmad, Y., Trinkle-Mulcahy, L., Verheggen, C., Cobley, A., Gregor, P., ... & Lamond, A. I. (2010). Establishment of a protein frequency library and its application in the reliable identification of specific protein interaction partners. *Molecular & Cellular Proteomics*, *9*(5), 861-879.

- Bowen, E. T. W., Lloyd, G., Harris, W. J., Platt, G. S., Baskerville, A., & Vella, E. E. (1977). Viral haemorrhagic fever in southern Sudan and northern Zaire: preliminary studies on the aetiological agent. *The Lancet*, *309*(8011), 571-573.

- Brauburger, K., Hume, A. J., Mühlberger, E., & Olejnik, J. (2012). Forty-five years of Marburg virus research. *Viruses*, *4*(10), 1878-1927.

- Bray, M., & Geisbert, T. W. (2005). Ebola virus: the role of macrophages and dendritic cells in the pathogenesis of Ebola hemorrhagic fever. *The international journal of biochemistry & cell biology*, *37*(8), 1560-1566.

- Bukreyev, A. A., Chandran, K., Dolnik, O., Dye, J. M., Ebihara, H., Leroy, E. M., ... & Kuhn, J. H. (2014). Discussions and decisions of the 2012–2014 International Committee on Taxonomy of Viruses (ICTV) Filoviridae study group, January 2012–June 2013. *Archives of virology*, *159*(4), 821-830.

- Burke, J., Declerq, R., Ghysebrechts, G., Pattyn, S. R., Piot, P., Ronsmans, M., ... & Comm, I. (1978). Ebola hemorrhagic-fever in Zaire, 1976-report of an international-commission. Bulletin of the World Health Organization, 56(2), 271-293.

- Cao, M., Wei, C., Zhao, L., Wang, J., Jia, Q., Wang, X., ... & Deng, T. (2014). DnaJA1/Hsp40 is co-opted by influenza A virus to enhance its viral RNA polymerase activity. *Journal of virology*, *88*(24), 14078-14089.

- Cardenas, W. B., Loo, Y. M., Gale, M., Hartman, A. L., Kimberlin, C. R., Martínez-Sobrido, L., ... & Basler, C. F. (2006). Ebola virus VP35 protein binds double-stranded RNA and inhibits alpha/beta interferon

production induced by RIG-I signaling. *Journal of virology*, *80*(11), 5168-5178.

- Carette, J. E., Raaben, M., Wong, A. C., Herbert, A. S., Obernosterer, G., Mulherkar, N., ... & Dal Cin, P. (2011). Ebola virus entry requires the cholesterol transporter Niemann-Pick C1. *Nature*, *477*(7364), 340-343.

- Carroll, S. A., Towner, J. S., Sealy, T. K., McMullan, L. K., Khristova, M. L., Burt, F. J., ... & Nichol, S. T. (2013). Molecular evolution of viruses of the family Filoviridae based on 97 whole-genome sequences. *Journal of virology*, *87*(5), 2608-2616.

- Chandran, K., Sullivan, N. J., Felbor, U., Whelan, S. P., & Cunningham, J. M. (2005). Endosomal proteolysis of the Ebola virus glycoprotein is necessary for infection. *Science*, *308*(5728), 1643-1645.

- Chepurnov, A. A., Zubavichene, N. M., & Dadaeva, A. A. (2003). Elaboration of laboratory strains of Ebola virus and study of pathophysiological reactions of animals inoculated with these strains. *Acta tropica*, *87*(3), 321-329.

- Connolly, B. M., Steele, K. E., Davis, K. J., Geisbert, T. W., Kell, W. M., Jaax, N. K., & Jahrling, P. B. (1999). Pathogenesis of experimental Ebola virus infection in guinea pigs. *Journal of Infectious Diseases*, *179*(Supplement 1), S203-S217.

- Côté, M., Misasi, J., Ren, T., Bruchez, A., Lee, K., Filone, C. M., ... & Cunningham, J. (2011). Small molecule inhibitors reveal Niemann-Pick C1 is essential for Ebola virus infection. *Nature*, *477*(7364), 344-348.

- Crary, S. M., Towner, J. S., Honig, J. E., Shoemaker, T. R., & Nichol, S. T. (2003). Analysis of the role of predicted RNA secondary structures in Ebola virus replication. *Virology*, *306*(2), 210-218.

- Davidson, E., Bryan, C., Fong, R. H., Barnes, T., Pfaff, J. M., Mabila, M., ... & Doranz, B. J. (2015). Mechanism of binding to Ebola virus glycoprotein by the ZMapp, ZMAb, and MB-003 cocktail antibodies. *Journal of virology*,*89*(21), 10982-10992.

- Dodson, A. W., Taylor, T. J., Knipe, D. M., & Coen, D. M. (2007). Inhibitors of the sodium potassium ATPase that impair herpes simplex virus replication identified via a chemical screening approach. *Virology*, *366*(2), 340-348.

- Dowall, S. D., Matthews, D. A., García-Dorival, I., Taylor, I., Kenny, J., Hertz-Fowler, C., ... & Barr, J. N. (2014). Elucidating variations in the nucleotide sequence of Ebola virus associated with increasing pathogenicity. *Genome biology*, *15*(11), 540.

- Ebihara, H., Takada, A., Kobasa, D., Jones, S., Neumann, G., Theriault, S., ... & Kawaoka, Y. (2006). Molecular determinants of Ebola virus virulence in mice. *PLoS Pathog*, *2*(7), e73.

- Ebihara, H., Zivcec, M., Gardner, D., Falzarano, D., LaCasse, R., Rosenke, R., ... & Feldmann, H. (2012). A Syrian golden hamster model recapitulating ebola hemorrhagic fever. *Journal of Infectious Diseases*, jis626.

- Edwards, M. R., Liu, G., Mire, C. E., Sureshchandra, S., Luthra, P., Yen, B., ... & Amarasinghe, G. K. (2016). Differential Regulation of Interferon Responses by Ebola and Marburg Virus VP35 Proteins. *Cell Reports*.

- Edwards, M. R., Pietzsch, C., Vausselin, T., Shaw, M. L., Bukreyev, A., & Basler, C. F. (2015). High-Throughput Minigenome System for Identifying Small-Molecule Inhibitors of Ebola Virus Replication. *ACS infectious diseases*, *1*(8), 380-387.

- Elliott, L. H., Kiley, M. P., & McCormick, J. B. (1985). Descriptive analysis of Ebola virus proteins. *Virology*, *147*(1), 169-176.

- Emmott, E., Munday, D., Bickerton, E., Britton, P., Rodgers, M. A., Whitehouse, A., ... & Hiscox, J. A. (2013). The cellular interactome of the coronavirus infectious bronchitis virus nucleocapsid protein and

functional implications for virus biology. *Journal of virology*, *87*(17), 9486-9500.

- Feldmann, H., & Geisbert, T. W. (2011). Ebola haemorrhagic fever. *The Lancet*, *377*(9768), 849-862.

- Feldmann, H., Klenk, H. D., & Sanchez, A. (1993). *Molecular biology and evolution of filoviruses* (pp. 81-100). Springer Vienna.

- Feldmann, H., Mühlberger, E., Randolf, A., Will, C., Kiley, M. P., Sanchez, A., & Klenk, H. D. (1992). Marburg virus, a filovirus: messenger RNAs, gene order, and regulatory elements of the replication cycle. *Virus research*, *24*(1), 1-19.

- Feldmann, H., Sanchez, A., & Geisbert, T. W. (2013). Filoviridae: Marburg and Ebola viruses. *Fields virology. 6th ed. Philadelphia: Lippincott Williams & Wilkins*, 923-56.

- Flyak, A. I., Shen, X., Murin, C. D., Turner, H. L., David, J. A., Fusco, M. L., ... & Branchizio, A. (2016). Cross-Reactive and Potent Neutralizing Antibody Responses in Human Survivors of Natural Ebolavirus Infection. Cell. Cell, 164(3), 392-405.

- Formenty, P., Boesch, C., Wyers, M., Steiner, C., Donati, F., Dind, F., ... & Le Guenno, B. (1999). Ebola virus outbreak among wild

chimpanzees living in a rain forest of Cote d'Ivoire. *Journal of Infectious Diseases*, *179*(Supplement 1), S120-S126.

- Fürstenwerth, H. (2010). Ouabain–the insulin of the heart. *International journal of clinical practice*, *64*(12), 1591-1594.

- Gabriel, G., Klingel, K., Otte, A., Thiele, S., Hudjetz, B., Arman-Kalcek, G., ... & Keiner, B. (2011). Differential use of importin-[alpha] isoforms governs cell tropism and host adaptation of influenza virus. *Nature communications*, *2*, 156.

- García-Dorival, I., Wu, W., Dowall, S., Armstrong, S., Touzelet, O., Wastling, J., ... & Hiscox, J. A. (2014). Elucidation of the Ebola virus VP24 cellular interactome and disruption of virus biology through targeted inhibition of host-cell protein function. *Journal of proteome research*, *13*(11), 5120-5135.

- Geisbert, T. W., Hensley, L. E., Gibb, T. R., Steele, K. E., Jaax, N. K., & Jahrling, P. B. (2000). Apoptosis induced in vitro and in vivo during infection by Ebola and Marburg viruses. *Laboratory investigation*, *80*(2), 171-186.

- Geisbert, T. W., Hensley, L. E., Larsen, T., Young, H. A., Reed, D. S., Geisbert, J. B., ... & Davis, K. J. **(2003 A).** Pathogenesis of Ebola hemorrhagic fever in cynomolgus macaques: evidence that dendritic

cells are early and sustained targets of infection. *The American journal of pathology*, *163*(6), 2347-2370.

- Geisbert, T. W., Young, H. A., Jahrling, P. B., Davis, K. J., Kagan, E., & Hensley, L. E**. (2003 B).** Mechanisms underlying coagulation abnormalities in ebola hemorrhagic fever: overexpression of tissue factor in primate monocytes/macrophages is a key event. *Journal of Infectious Diseases*,*188*(11), 1618-1629.

- Geller, R., Taguwa, S., & Frydman, J. (2012). Broad action of Hsp90 as a host chaperone required for viral replication. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 1823(3), 698-706.

- Gire, S. K., Goba, A., Andersen, K. G., Sealfon, R. S., Park, D. J., Kanneh, L., ... & Wohl, S. (2014). Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. *Science*, *345*(6202), 1369-1372.

- Goecks, J., Nekrutenko, A., & Taylor, J. (2010). Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol*, *11*(8), R86.

- Groseth, A., Feldmann, H., & Strong, J. E. (2007). The ecology of Ebola virus. *Trends in microbiology*, *15*(9), 408-416.

- Groseth, A., Marzi, A., Hoenen, T., Herwig, A., Gardner, D., Becker, S., ... & Feldmann, H. (2012). The Ebola virus glycoprotein contributes to but is not sufficient for virulence in vivo. *PLoS pathogens*, *8*(8), e1002847.

- Gupta, M., Spiropoulou, C., & Rollin, P. E. (2007). Ebola virus infection of human PBMCs causes massive death of macrophages, CD4 and CD8 T cell sub-populations in vitro. *Virology*, *364*(1), 45-54.

- Han, Z., Boshra, H., Sunyer, J. O., Zwiers, S. H., Paragas, J., & Harty, R. N. (2003). Biochemical and functional characterization of the Ebola virus VP24 protein: implications for a role in virus assembly and budding. *Journal of virology*, *77*(3), 1793-1800.

- Hart, M. K. (2003). Vaccine research efforts for filoviruses. *International journal for parasitology*, *33*(5), 583-595.

- Henao-Restrepo, A. M., Longini, I. M., Egger, M., Dean, N. E., Edmunds, W. J., Camacho, A., ... & Enwere, G. (2015). Efficacy and effectiveness of an rVSV-vectored vaccine expressing Ebola surface glycoprotein: interim results from the Guinea ring vaccination cluster-randomised trial. *The Lancet*,*386*(9996), 857-866.

- Hoenen, T., & Feldmann, H. (2014). Reverse genetics systems as tools for the development of novel therapies against filoviruses. *Expert review of anti-infective therapy*, *12*(10), 1253-1263.

- Hoenen, T., Groseth, A., Callison, J., Takada, A., & Feldmann, H. (2013). A novel Ebola virus expressing luciferase allows for rapid and quantitative testing of antivirals. *Antiviral research*, *99*(3), 207-213.

- Hoenen, T., Groseth, A., de Kok-Mercado, F., Kuhn, J. H., & Wahl-Jensen, V. (2011). Minigenomes, transcription and replication competent virus-like particles and beyond: reverse genetics systems for filoviruses and other negative stranded hemorrhagic fever viruses. *Antiviral research*, *91*(2), 195-208.

- Hoenen, T., Jung, S., Herwig, A., Groseth, A., & Becker, S. (2010). Both matrix proteins of Ebola virus contribute to the regulation of viral genome replication and transcription. *Virology*, *403*(1), 56-66.

- Hoenen, T., Shabman, R. S., Groseth, A., Herwig, A., Weber, M., Schudt, G., ... & Feldmann, H. (2012). Inclusion bodies are a site of ebolavirus replication. *Journal of virology*, *86*(21), 11779-11788.

- Hoenen, T., Volchkov, V., Kolesnikova, L., Mittler, E., Timmins, J., Ottmann, M., ... & Weissenhorn, W. (2005). VP40 octamers are

essential for Ebola virus replication. *Journal of virology*, *79*(3), 1898-1905.

- Huang, Y., Xu, L., Sun, Y., & Nabel, G. J. (2002). The assembly of Ebola virus nucleocapsid requires virion-associated proteins 35 and 24 and posttranslational modification of nucleoprotein. *Molecular cell*, *10*(2), 307-316.

- Hudjetz, B., & Gabriel, G. (2012). Human-like PB2 627K influenza virus polymerase activity is regulated by importin-alpha1 and-alpha7. *PLoS Pathog*, *8*(1), e1002488.

- Huggins, J., Zhang, Z. X., & Bray, M. (1999). Antiviral drug therapy of filovirus infections: S-adenosylhomocysteine hydrolase inhibitors inhibit Ebola virus in vitro and in a lethal mouse model. *Journal of Infectious Diseases*, *179*(Supplement 1), S240-S247.

- ICTV report 2009; virus taxonomy release 2014; http://www.ictvonline.org/virusTaxonomy.asp.

- Iwasa, A., Halfmann, P., Noda, T., Oyama, M., Kozuka-Hata, H., Watanabe, S., ... & Kawaoka, Y. (2011). Contribution of Sec61α to the life cycle of Ebola virus. *Journal of Infectious Diseases*, *204*(suppl 3), S919-S926.

- Jahrling, P. B., Geisbert, T. W., Johnson, E. D., Peters, C. J., Dalgard, D. W., & Hall, W. C. (1990). Preliminary report: isolation of Ebola virus from monkeys imported to USA. *Lancet*, 335(8688), 502-505.

- Jasenosky, L. D., & Kawaoka, Y. (2004). Filovirus budding. Virus research,106(2), 181-188.

- Jasenosky, L. D., Neumann, G., & Kawaoka, Y. (2010). Minigenome-based reporter system suitable for high-throughput screening of compounds able to inhibit Ebolavirus replication and/or transcription. *Antimicrobial agents and chemotherapy*, *54*(7), 3007-3010.

- Jasenosky, L. D., Neumann, G., Lukashevich, I., & Kawaoka, Y. (2001). Ebola virus VP40-induced particle formation and association with the lipid bilayer. *Journal of virology*, *75*(11), 5205-5214.

- John, S. P., Wang, T., Steffen, S., Longhi, S., Schmaljohn, C. S., & Jonsson, C. B. (2007). Ebola virus VP30 is an RNA binding protein. *Journal of virology*,*81*(17), 8967-8976.

- Johnson, K. M., Lange, J. V., Webb, P. A., & Murphy, F. A. (1977). Isolation and partial characterisation of a new virus causing acute haemorrhagic fever in Zaire. *Lancet*, *309*(8011), 569-571.

- Jordan, M., Schallhorn, A., & Wurm, F. M. (1996). Transfecting mammalian cells: optimization of critical parameters affecting calcium-phosphate precipitate formation. *Nucleic acids research*, *24*(4), 596-601.

- Jourdan, S. S., Osorio, F., & Hiscox, J. A. (2012). An interactome map of the nucleocapsid protein from a highly pathogenic North American porcine reproductive and respiratory syndrome virus strain generated using SILAC-based quantitative proteomics. *Proteomics*, *12*(7), 1015-1023.

- Karuppannan, A. K., Wu, K. X., Qiang, J., Chu, J. J. H., & Kwang, J. (2012). Natural compounds inhibiting the replication of Porcine reproductive and respiratory syndrome virus. *Antiviral research*, *94*(2), 188-194.

- Katoh, H., Kubota, T., Kita, S., Nakatsu, Y., Aoki, N., Mori, Y., ... & Kidokoro, M. (2015). Heat shock protein 70 regulates degradation of the mumps virus phosphoprotein via the ubiquitin-proteasome pathway. *Journal of virology*, *89*(6), 3188-3199.

- Khan, K. H. (2013). Gene expression in mammalian cells and its applications. Advanced pharmaceutical bulletin, 3(2), 257.

- Kim, M. Y., & Oglesbee, M. (2012). Virus-heat shock protein interaction and a novel axis for innate antiviral immunity. *Cells*, *1*(3), 646-666.

- Kirchdoerfer, R. N., Abelson, D. M., Li, S., Wood, M. R., & Saphire, E. O. (2015). Assembly of the Ebola Virus Nucleoprotein from a Chaperoned VP35 Complex. *Cell reports*, *12*(1), 140-149.

- Kondratowicz, A. S., Lennemann, N. J., Sinn, P. L., Davey, R. A., Hunt, C. L., Moller-Tank, S., ... & Sandersfeld, L. M. (2011). T-cell immunoglobulin and mucin domain 1 (TIM-1) is a receptor for Zaire Ebolavirus and Lake Victoria Marburgvirus. *Proceedings of the National Academy of Sciences*, *108*(20), 8426-8431.

- Kugelman, J. R., Lee, M. S., Rossi, C. A., McCarthy, S. E., Radoshitzky, S. R., Dye, J. M., ... & Warren, T. K. (2012). Ebola virus genome plasticity as a marker of its passaging history: a comparison of in vitro passaging to non-human primate infection. *PloS one*, *7*(11), e50316

- Kuhn, J. H., Becker, S., Ebihara, H., Geisbert, T. W., Johnson, K. M., Kawaoka, Y., ... & Jahrling, P. B. (2010). Proposal for a revised taxonomy of the family Filoviridae: classification, names of taxa and viruses, and virus abbreviations. *Archives of virology*, *155*(12), 2083-2103.

- Kuhn, J., & Calisher, C. H. (2008). *Filoviruses: a compendium of 40 years of epidemiological, clinical, and laboratory studies* (Vol. 20). Springer Science & Business Media.

- Lahaye, X., Vidy, A., Fouquet, B., & Blondel, D. (2012). Hsp70 protein positively regulates rabies virus infection. *Journal of virology*, *86*(9), 4743-4751.

- Lai, M. M. (1998). Cellular factors in the transcription and replication of viral RNA genomes: a parallel to DNA-dependent RNA transcription. *Virology*,*244*(1), 1-12.

- Lamb A. Robert. (2013). Mononegavirales. *. Fields virology. 6th ed. Philadelphia: Lippincott Williams & Wilkins*, 747-94.

- Lamb, R. A., & Parks, G. D. (2013). Paramyxoviridae. I n: Knipe, DM, Howley, PM (Eds.), *Fields virology. 6th ed. Philadelphia: Lippincott Williams & Wilkins*, 957-995

- Lamunu, M., Lutwama, J. J., Kamugisha, J., Opio, A., Nambooze, J., Ndayimirije, N., & Okware, S. (2004). Containing a haemorrhagic fever epidemic: The Ebola experience in Uganda (October 2000–January 2001). *International journal of infectious diseases*, *8*(1), 27-37.

- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature methods*, *9*(4), 357-359.

- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., ... & Thompson, J. D. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, *23*(21), 2947-2948.

- Le Guenno, B., Formenty, P., Wyers, M., Gounon, P., Walker, F., & Boesch, C. (1995). Isolation and partial characterisation of a new strain of Ebola virus. *Lancet*, *345*(8960), 1271-1274.

- Lee, J. E., & Saphire, E. O. (2009). Ebolavirus glycoprotein structure and mechanism of entry. *Future virology*, *4*(6), 621-635.

- Lee, M. S., Lebeda, F. J., & Olson, M. A. (2009). Fold prediction of VP24 protein of Ebola and Marburg viruses using de novo fragment assembly. *Journal of structural biology*, *167*(2), 136-144.

- Leroy, E. M., Kumulungui, B., Pourrut, X., Rouquet, P., Hassanin, A., Yaba, P., ... & Swanepoel, R. (2005). Fruit bats as reservoirs of Ebola virus. *Nature*, *438*(7068), 575-576.

- Leung, D. W., Borek, D., Luthra, P., Binning, J. M., Anantpadma, M., Liu, G., ... & Shabman, R. S. (2015). An intrinsically disordered peptide from Ebola Virus VP35 controls viral RNA synthesis by modulating nucleoprotein-RNA interactions. *Cell reports*, *11*(3), 376-389.

- Licata, J. M., Johnson, R. F., Han, Z., & Harty, R. N. (2004). Contribution of Ebola virus glycoprotein, nucleoprotein, and VP24 to budding of VP40 virus-like particles. *Journal of virology*, *78*(14), 7344-7351.

- Liu, G., Zhao, S., Bailey, J. A., Sahinalp, S. C., Alkan, C., Tuzun, E., ... & NISC Comparative Sequencing Program. (2003). Analysis of primate genomic variation reveals a repeat-driven expansion of the human genome. *Genome research*, *13*(3), 358-368.

- Liu, L., Lear, Z., Hughes, D. J., Wu, W., Zhou, E. M., Whitehouse, A., ... & Hiscox, J. A. (2015). Resolution of the cellular proteome of the nucleocapsid protein from a highly pathogenic isolate of porcine reproductive and respiratory syndrome virus identifies PARP-1 as a cellular target whose interaction is critical for virus biology. *Veterinary microbiology*, *176*(1), 109-119.

- Manzoor, R., Kuroda, K., Yoshida, R., Tsuda, Y., Fujikura, D., Miyamoto, H., ... & Takada, A. (2014). Heat shock protein 70 modulates influenza A virus polymerase activity. *Journal of Biological Chemistry*, *289*(11), 7599-7614.

- Marsh, G. A., Haining, J., Robinson, R., Foord, A., Yamada, M., Barr, J. A., ... & Rollin, P. E. (2011). Ebola Reston virus infection of pigs: clinical significance and transmission potential. *Journal of Infectious Diseases*, *204*(suppl 3), S804-S809.

- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. journal*, *17*(1), pp-10.

- Martinez, O., Ndungo, E., Tantral, L., Miller, E. H., Leung, L. W., Chandran, K., & Basler, C. F. (2013). A mutation in the Ebola virus envelope glycoprotein restricts viral entry in a host species-and cell-type-specific manner. *Journal of virology*, *87*(6), 3324-3334.

- Maruyama, T., Rodriguez, L. L., Jahrling, P. B., Sanchez, A., Khan, A. S., Nichol, S. T., ... & Burton, D. R. (1999). Ebola virus can be effectively neutralized by antibody produced in natural human infection. *Journal of virology*, *73*(7), 6024-6030.

- Marzi, A., & Feldmann, H. (2014). Ebola virus vaccines: an overview of current approaches. *Expert review of vaccines*, *13*(4), 521-531.

- Mateo, M., Carbonnelle, C., Martinez, M. J., Reynard, O., Page, A., Volchkova, V. A., & Volchkov, V. E. **(2011 A).** Knockdown of Ebola virus VP24 impairs viral nucleocapsid assembly and prevents virus replication. *Journal of Infectious Diseases*, *204*(suppl 3), S892-S896.

- Mateo, M., Carbonnelle, C., Reynard, O., Kolesnikova, L., Nemirov, K., Page, A., ... & Volchkov, V. E. **(2011 B).** VP24 is a molecular determinant of Ebola virus virulence in guinea pigs. *Journal of Infectious Diseases*, *204*(suppl 3), S1011-S1020.

- Mateo, M., Leung, L. W., Basler, C. F., & Volchkov, V. E. (2010). Ebolavirus VP24 binding to karyopherins is required for inhibition of interferon signaling. *Journal of virology*, *84*(2), 1169-1175.

- Mayer, M. P. (2005 B). Recruitment of Hsp70 chaperones: a crucial part of viral survival strategies. In *Reviews of physiology, biochemistry and pharmacology* (pp. 1-46). Springer Berlin Heidelberg.

- Mayer, M. P., & Bukau, B. (2005A). Hsp70 chaperones: cellular functions and molecular mechanism. *Cellular and molecular life sciences*, *62*(6), 670-684.

- Mehedi, M., Falzarano, D., Seebach, J., Hu, X., Carpenter, M. S., Schnittler, H. J., & Feldmann, H. (2011). A new Ebola virus nonstructural glycoprotein expressed through RNA editing. *Journal of virology*, *85*(11), 5406-5414.

- Mehedi, M., Hoenen, T., Robertson, S., Ricklefs, S., Dolan, M. A., Taylor, T., ... & Feldmann, H. (2013). Ebola virus RNA editing depends on the primary editing site sequence and an upstream secondary structure. *PLoS Pathog*, *9*(10), e1003677.

- Mendoza, E. J., Qiu, X., & Kobinger, G. P. (2016). Progression of Ebola Therapeutics During the 2014–2015 Outbreak. *Trends in molecular medicine*.

- Messaoudi, I., Amarasinghe, G. K., & Basler, C. F. (2015). Filovirus pathogenesis and immune evasion: insights from Ebola virus and Marburg virus. *Nature Reviews Microbiology*, *13*(11), 663-676.

- Miller, E. H., Obernosterer, G., Raaben, M., Herbert, A. S., Deffieu, M. S., Krishnan, A., ... & Ruthel, G. (2012). Ebola virus entry requires the host-programmed recognition of an intracellular receptor. *The EMBO journal*, *31*(8), 1947-1960.

- Miranda, M. E., Ksiazek, T. G., Retuya, T. J., Khan, A. S., Sanchez, A., Fulhorst, C. F., ... & Peters, C. J. (1999). Epidemiology of Ebola (subtype Reston) virus in the Philippines, 1996. *Journal of Infectious Diseases*, *179* (Supplement 1), S115-S119.

- Modrof, J., Mühlberger, E., Klenk, H. D., & Becker, S. (2002). Phosphorylation of VP30 impairs Ebola virus transcription. *Journal of Biological Chemistry*, *277*(36), 33099-33104.

- Mohan, G. S., Li, W., Ye, L., Compans, R. W., & Yang, C. (2012). Antigenic subversion: a novel mechanism of host immune evasion by Ebola virus. *PLoS Pathog*, *8*(12), e1003065.

- Mühlberger, E. (2007). Filovirus replication and transcription. *Future virology*,*2*(2), 205.

- Mühlberger, E., Lötfering, B., Klenk, H. D., & Becker, S. (1998). Three of the four nucleocapsid proteins of Marburg virus, NP, VP35, and L, are sufficient to mediate replication and transcription of Marburg virus-specific monocistronic minigenomes. *Journal of virology*, *72*(11), 8756-8764.

- Mühlberger, E., Trommer, S., Funke, C., Volchkov, V., Klenk, H. D., & Becker, S. (1996). Termini of all mRNA species of Marburg virus: sequence and secondary structure. *Virology*, *223*(2), 376-380.

- Mühlberger, E., Weik, M., Volchkov, V. E., Klenk, H. D., & Becker, S. (1999). Comparison of the transcription and replication strategies of Marburg virus and Ebola virus by using artificial replication systems. *Journal of virology*,*73*(3), 2333-2342.

- Munday, D. C., Wu, W., Smith, N., Fix, J., Noton, S. L., Galloux, M., ... & Easton, A. J. (2015). Interactome analysis of the human respiratory syncytial virus RNA polymerase complex identifies protein chaperones as important cofactors that promote L-protein stability and RNA synthesis. *Journal of virology*, *89*(2), 917-930.

- Mupapa, K., Massamba, M., Kibadi, K., Kuvula, K., Bwaka, A., Kipasa, M., ... & Muyembe-Tamfum, J. J. (1999). Treatment of Ebola

hemorrhagic fever with blood transfusions from convalescent patients. *Journal of Infectious Diseases*, *179*(Supplement 1), S18-S23.

- Muyembe-Tamfum, J. J., Kipasa, M., Kiyungu, C., & Colebunders, R. (1999). Ebola outbreak in Kikwit, Democratic Republic of the Congo: discovery and control measures. *Journal of Infectious Diseases*, *179*(Supplement 1), S259-S262.

- Muyembe-Tamfum, J. J., Mulangu, S., Masumu, J., Kayembe, J. M., Kemp, A., & Paweska, J. T. (2012). Ebola virus outbreaks in Africa: past and present. *Onderstepoort Journal of Veterinary Research*, *79*(2), 06-13.

- Nanbo, A., Watanabe, S., Halfmann, P., & Kawaoka, Y. (2013). The spatio-temporal distribution dynamics of Ebola virus proteins and RNA in infected cells. *Scientific reports*, *3*.

- Neckers, L., & Tatu, U. (2008). Molecular chaperones in pathogen virulence: emerging new targets for therapy. Cell host & microbe, 4(6), 519-527.

- Negredo, A., Palacios, G., Vázquez-Morón, S., González, F., Dopazo, H., Molero, F., ... & Tenorio, A. (2011). Discovery of an ebolavirus-like filovirus in europe. *PLoS Pathog*, *7*(10), e1002304.

- Neumann, G., Watanabe, S., & Kawaoka, Y. (2009). Characterization of Ebolavirus regulatory genomic regions. *Virus research*, *144*(1), 1-7.

- Noda, T., Hagiwara, K., Sagara, H., & Kawaoka, Y. (2010). Characterization of the Ebola virus nucleoprotein–RNA complex. *Journal of General Virology*, *91*(6), 1478-1483.

- Noda, T., Kolesnikova, L., Becker, S., & Kawaoka, Y. (2011). The importance of the NP: VP35 ratio in Ebola virus nucleocapsid formation. *Journal of Infectious Diseases*, *204*(suppl 3), S878-S883.

- Noda, T., Sagara, H., Suzuki, E., Takada, A., Kida, H., & Kawaoka, Y. (2002). Ebola virus VP40 drives the formation of virus-like filamentous particles along with GP. *Journal of virology*, *76*(10), 4855-4865.

- Paessler, S., & Walker, D. H. (2013). Pathogenesis of the viral hemorrhagic fevers. *Annual Review of Pathology: Mechanisms of Disease*, *8*, 411-440.

- Pattyn, S., Vander Groen, G., Jacob, W., Piot, P., & Courteille, G. (1977). Isolation of Marburg-like virus from a case of haemorrhagic fever in Zaire. *The Lancet*, *309*(8011), 573-574.

- Pichlmair, A., Kandasamy, K., Alvisi, G., Mulhern, O., Sacco, R., Habjan, M., ... & Bürckstümmer, T. (2012). Viral immune modulators

perturb the human molecular network by common and unique strategies. *Nature*, *487*(7408), 486-490.

- Poch, O., Blumberg, B. M., Bougueleret, L., & Tordo, N. (1990). Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: theoretical assignment of functional domains. *Journal of General Virology*, *71*(5), 1153-1162.

- Qiu, X., Audet, J., Wong, G., Fernando, L., Bello, A., Pillet, S., ... & Kobinger, G. P. (2013). Sustained protection against Ebola virus infection following treatment of infected nonhuman primates with ZMAb. *Scientific reports*, *3*.

- Quick, J., Loman, N. J., Duraffour, S., Simpson, J. T., Severi, E., Cowley, L., ... & Ouédraogo, N. (2016). Real-time, portable genome sequencing for Ebola surveillance. Nature, 530(7589), 228-232.

- Ramanan, P., Shabman, R. S., Brown, C. S., Amarasinghe, G. K., Basler, C. F., & Leung, D. W. (2011). Filoviral immune evasion mechanisms. *Viruses*, *3*(9), 1634-1649.

- Randall, R. E., & Goodbourn, S. (2008). Interferons and viruses: an interplay between induction, signalling, antiviral responses and virus countermeasures. *Journal of General Virology*, *89*(1), 1-47.

- Reed, L. J., & Muench, H. (1938). A simple method of estimating fifty per cent endpoints. *American journal of epidemiology*, *27*(3), 493-497.

- Regnery, R. L., Johnson, K. M., & Kiley, M. P. (1980). Virion nucleic acid of Ebola virus. *Journal of virology*, *36*(2), 465-469.

- Regules, J. A., Beigel, J. H., Paolino, K. M., Voell, J., Castellano, A. R., Muñoz, P., ... & Gutiérrez, R. L. (2015). A recombinant vesicular stomatitis virus Ebola vaccine—preliminary report. *New England Journal of Medicine*.

- Reid, S. P., Leung, L. W., Hartman, A. L., Martinez, O., Shaw, M. L., Carbonnelle, C., ... & Basler, C. F. (2006). Ebola virus VP24 binds karyopherin alpha1 and blocks STAT1 nuclear accumulation. *Journal of virology*, *80*(11), 5156-5167.

- Reid, S. P., Valmas, C., Martinez, O., Sanchez, F. M., & Basler, C. F. (2007). Ebola virus VP24 proteins inhibit the interaction of NPI-1 subfamily karyopherin alpha proteins with activated STAT1. *Journal of virology*, *81*(24), 13469-13477.

- Report of a WHO/International Study Team. (1978). Ebola haemorrhagic fever in Sudan, 1976. Bulletin of the World Health Organization, 56(2), 247.

- Rougeron, V., Feldmann, H., Grard, G., Becker, S., & Leroy, E. M. (2015). Ebola and Marburg haemorrhagic fever. Journal of Clinical Virology, 64, 111-119.

- Ruigrok, R. W., Crépin, T., & Kolakofsky, D. (2011). Nucleoproteins and nucleocapsids of negative-strand RNA viruses. *Current opinion in microbiology*, *14*(4), 504-510.

- Ryabchikova, E., Kolesnikova, L., Smolina, M., Tkachev, V., Pereboeva, L., Baranova, S., ... & Rassadkin, Y. (1996). Ebola virus infection in guinea pigs: presumable role of granulomatous inflammation in pathogenesis. *Archives of virology*, *141*(5), 909-921.

- Saeed, M. F., Kolokoltsov, A. A., Albrecht, T., & Davey, R. A. (2010). Cellular entry of ebola virus involves uptake by a macropinocytosis-like mechanism and subsequent trafficking through early and late endosomes. *PLoS Pathog*, *6*(9), e1001110.

- Sanchez, A., Kiley, M. P., Holloway, B. P., & Auperin, D. D. (1993). Sequence analysis of the Ebola virus genome: organization, genetic elements, and comparison with the genome of Marburg virus. *Virus research*, *29*(3), 215-240.

- Sanchez, A., Kiley, M. P., Holloway, B. P., McCormick, J. B., & Auperin, D. D. (1989). The nucleoprotein gene of Ebola virus: cloning, sequencing, and in vitro expression. *Virology*, *170*(1), 81-91.

- Sanchez, A., Kiley, M. P., Klenk, H. D., & Feldmann, H. (1992). Sequence analysis of the Marburg virus nucleoprotein gene: comparison to Ebola virus and other non-segmented negative-strand RNA viruses. *Journal of General Virology*, *73*(2), 347-357.

- Sanchez, A., Trappier, S. G., Mahy, B. W., Peters, C. J., & Nichol, S. T. (1996). The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. *Proceedings of the National Academy of Sciences*, *93*(8), 3602-3607.

- Schlecht, R., Scholz, S. R., Dahmen, H., Wegener, A., Sirrenberg, C., Musil, D., ... & Bukau, B. (2013). Functional analysis of Hsp70 inhibitors. *PloS one,*8(11), e78443.

- Shabman, R. S., Jabado, O. J., Mire, C. E., Stockwell, T. B., Edwards, M., Mahajan, M., ... & Basler, C. F. (2014). Deep sequencing identifies noncanonical editing of Ebola and Marburg virus RNAs in infected cells. *MBio*, *5*(6), e02011-14.

- Shi, W., Huang, Y., Sutton-Smith, M., Tissot, B., Panico, M., Morris, H. R., ... & Yang, Z. Y. (2008). A filovirus-unique region of Ebola virus nucleoprotein confers aberrant migration and mediates its incorporation into virions. *Journal of virology*, *82*(13), 6190-6199.

- Shimojima, M., Takada, A., Ebihara, H., Neumann, G., Fujioka, K., Irimura, T., ... & Kawaoka, Y. (2006). Tyro3 family-mediated cell entry of Ebola and Marburg viruses. *Journal of virology*, *80*(20), 10109-10116.

- Slenczka, W., & Klenk, H. D. (2007). Forty years of Marburg virus. *Journal of Infectious Diseases*, *196*(Supplement 2), S131-S135.

- Smith, D. R., McCarthy, S., Chrovian, A., Olinger, G., Stossel, A., Geisbert, T. W., ... & Connor, J. H. (2010). Inhibition of heat-shock protein 90 reduces Ebola virus replication. *Antiviral research*, *87*(2), 187-194.

- Spurgers, K. B., Alefantis, T., Peyser, B. D., Ruthel, G. T., Bergeron, A. A., Costantino, J. A., ... & DelVecchio, V. G. (2010). Identification of essential filovirion-associated host factors by serial proteomic analysis and RNAi screen. *Molecular & Cellular Proteomics*, *9*(12), 2690-2703.

- St Patrick Reid, C. V., Martinez, O., Sanchez, F. M., & Basler, C. F. (2008). Ebola Virus VP24 Proteins Inhibit the Interaction of NPI-1

Subfamily Karyopherin α Proteins with Activated STAT1. *Journal of virology*, *82*(6), 3163.

- Subbotina, E., Dadaeva, A., Kachko, A., & Chepurnov, A. (2010). Genetic factors of Ebola virus virulence in guinea pigs. *Virus research*, *153*(1), 121-133.

- Taguwa, S., Maringer, K., Li, X., Bernal-Rubio, D., Rauch, J. N., Gestwicki, J. E., ... & Frydman, J. (2015). Defining Hsp70 Subnetworks in Dengue Virus Replication Reveals Key Vulnerability in Flavivirus Infection. *Cell*, *163*(5), 1108-1123.

- Takada, A., Fujioka, K., Tsuiji, M., Morikawa, A., Higashi, N., Ebihara, H., ... & Kawaoka, Y. (2004). Human macrophage C-type lectin specific for galactose and N-acetylgalactosamine promotes filovirus entry. *Journal of virology*, *78*(6), 2943-2947.

- Takada, A., Watanabe, S., Ito, H., Okazaki, K., Kida, H., & Kawaoka, Y. (2000). Downregulation of β1 integrins by Ebola virus glycoprotein: implication for virus entry. *Virology*, *278*(1), 20-26.

- Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular biology and evolution*, *30*(12), 2725-2729.

- Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research*, *22*(22), 4673-4680.

- Timmins, J., Scianimanico, S., Schoehn, G., & Weissenhorn, W. (2001). Vesicular release of Ebola virus matrix protein VP40. *Virology*, *283*(1), 1-6.

- Töpfer, A., Zagordi, O., Prabhakaran, S., Roth, V., Halperin, E., & Beerenwinkel, N. (2013). Probabilistic inference of viral quasispecies subject to recombination. *Journal of Computational Biology*, *20*(2), 113-123.

- Towner, J. S., Sealy, T. K., Khristova, M. L., Albariño, C. G., Conlan, S., Reeder, S. A., ... & Okware, S. (2008). Newly discovered ebola virus associated with hemorrhagic fever outbreak in Uganda. PLoS Pathog, 4(11), e1000212.

- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., ... & Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols*, *7*(3), 562-578.

- Trinkle-Mulcahy, L., Boulon, S., Lam, Y. W., Urcia, R., Boisvert, F. M., Vandermoere, F., ... & Lamond, A. (2008). Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. *The Journal of cell biology*, *183*(2), 223-239.

- Trinkle-Mulcahy, L. (2012). Resolving protein interactions and complexes by affinity purification followed by label-based quantitative mass spectrometry. *Proteomics*, *12*(10), 1623-1638.

- Trombley, A. R., Wachter, L., Garrison, J., Buckley-Beason, V. A., Jahrling, J., Hensley, L. E., ... & Kulesh, D. A. (2010). Comprehensive Panel of Real-Time TaqMan™ Polymerase Chain Reaction Assays for Detection and Absolute Quantification of Filoviruses, Arenaviruses, and New World Hantaviruses. *The American journal of tropical medicine and hygiene*, *82*(5), 954-960.

- Trunschke, M., Conrad, D., Enterlein, S., Olejnik, J., Brauburger, K., & Mühlberger, E. (2013). The L–VP35 and L–L interaction domains reside in the amino terminus of the Ebola virus L protein and are potential targets for antivirals. *Virology*, *441*(2), 135-145.

- Uebelhoer, L. S., Albariño, C. G., McMullan, L. K., Chakrabarti, A. K., Vincent, J. P., Nichol, S. T., & Towner, J. S. (2014). High-throughput, luciferase-based reverse genetics systems for identifying inhibitors of Marburg and Ebola viruses. *Antiviral research*, *106*, 86-94.

- Vemuri, R., Longoni, S., & Philipson, K. D. (1989). Ouabain treatment of cardiac cells induces enhanced Na+-Ca2+ exchange activity. *American Journal of Physiology-Cell Physiology*, *256*(6), C1273-C1276.

- Volchkov, V. E., Becker, S., Volchkova, V. A., Ternovoj, V. A., Kotov, A. N., Netesov, S. V., & Klenk, H. D. (1995). GP mRNA of Ebola Virus Is Edited by the Ebola Virus Polymerase and by T7 and Vaccinia Virus Polymerases 1. *Virology*, *214*(2), 421-430.

- Volchkov, V. E., Chepurnov, A. A., Volchkova, V. A., Ternovoj, V. A., & Klenk, H. D. (2000). Molecular characterization of guinea pig-adapted variants of Ebola virus. *Virology*, *277*(1), 147-155.

- Volchkov, V. E., Feldmann, H., Volchkova, V. A., & Klenk, H. D. (1998). Processing of the Ebola virus glycoprotein by the proprotein convertase furin. *Proceedings of the National Academy of Sciences*, *95*(10), 5762-5767.

- Volchkov, V. E., Volchkova, V. A., Chepurnov, A. A., Blinov, V. M., Dolnik, O., Netesov, S. V., & Feldmann, H. (1999). Characterization of the L gene and 5'trailer region of Ebola virus. *Journal of General Virology*, *80*(2), 355-362.

- Volchkova, V. A., Dolnik, O., Martinez, M. J., Reynard, O., & Volchkov, V. E. (2011). Genomic RNA editing and its impact on Ebola virus adaptation during serial passages in cell culture and infection of guinea pigs. *Journal of Infectious Diseases*, *204*(suppl 3), S941-S946.

- Wahl-Jensen, V., Bollinger, L., Safronetz, D., de Kok-Mercado, F., Scott, D. P., & Ebihara, H. (2012). Use of the Syrian hamster as a new model of ebola virus disease and other viral hemorrhagic fevers. *Viruses*, *4*(12), 3754-3784.

- Wang, L. F., Harcourt, B. H., Yu, M., Tamin, A., Rota, P. A., Bellini, W. J., & Eaton, B. T. (2001). Molecular biology of Hendra and Nipah viruses. *Microbes and infection*, *3*(4), 279-287.

- Watanabe, S., Noda, T., & Kawaoka, Y. (2006). Functional mapping of the nucleoprotein of Ebola virus. *Journal of virology*, *80*(8), 3743-3751.

- Watanabe, S., Noda, T., Halfmann, P., Jasenosky, L., & Kawaoka, Y. (2007). Ebola virus (EBOV) VP24 inhibits transcription and replication of the EBOV genome. *Journal of Infectious Diseases*, *196*(Supplement 2), S284-S290.

- Weik, M., Enterlein, S., Schlenz, K., & Mühlberger, E. (2005). The Ebola virus genomic replication promoter is bipartite and follows the rule of six. *Journal of virology*, *79*(16), 10660-10671.

- Weik, M., Modrof, J., Klenk, H. D., Becker, S., & Mühlberger, E. (2002). Ebola virus VP30-mediated transcription is regulated by RNA secondary structure formation. *Journal of virology*, *76*(17), 8532-8539.

- Weingartl, H. M., Nfon, C., & Kobinger, G. (2013). Review of Ebola virus infections in domestic animals. In Vaccines and Diagnostics for Transboundary Animal Diseases (Vol. 135, pp. 211-218). Karger Publishers.

- Wen, W., Liu, W., Shao, Y., & Chen, L. (2014). VER-155008, a small molecule inhibitor of HSP70 with potent anti-cancer activity on lung cancer cells lines. *Experimental Biology and Medicine*, *239*(5), 638-645.

- Whelan, S. P. J., Barr, J. N., & Wertz, G. W. (2004). Transcription and replication of nonsegmented negative-strand RNA viruses. In *Biology of Negative Strand RNA Viruses: The Power of Reverse Genetics* (pp. 61-119). Springer Berlin Heidelberg.

- WHO (World Health Organization). Ebola virus disease, West Africa—update. 2016.Available at: URLhttp://apps.who.int/ebola/current-situation/ebola-situation-report-17-february-2016.

- Wong, G., Qiu, X., Richardson, J. S., Cutts, T., Collignon, B., Gren, J., ... & Kobinger, G. P. (2015). Ebola virus transmission in guinea pigs. *Journal of virology*, *89*(2), 1314-1323.

- Wu, W., Tran, K. C., Teng, M. N., Heesom, K. J., Matthews, D. A., Barr, J. N., & Hiscox, J. A. (2012). The interactome of the human respiratory syncytial virus NS1 protein highlights multiple effects on host cell biology. *Journal of virology*, *86*(15), 7777-7789.

- Wynne, J. W., Shiell, B. J., Marsh, G. A., Boyd, V., Harper, J. A., Heesom, K., ... & Todd, S. (2014). Proteomics informed by transcriptomics reveals Hendra virus sensitizes bat cells to TRAIL mediated apoptosis. *Genome biology*, *15*(11), 532.

- Xiao, A., Wong, J., & Luo, H. (2010). Viral interaction with molecular chaperones: role in regulating viral infection. *Archives of virology*, *155*(7), 1021-1031.

- Xiao, A., Wong, J., & Luo, H. (2010). Viral interaction with molecular chaperones: role in regulating viral infection. *Archives of virology*, *155*(7), 1021-1031.

- Xiao, Y., Wu, W., Gao, J., Smith, N., Burkard, C., Xia, D., ... & Zhou, E. M. (2015). Characterization of the interactome of the porcine reproductive and respiratory syndrome virus (PRRSV) NSP2 protein reveals the hyper variable region as a binding platform for association with 14-3-3 proteins. *Journal of proteome research*.

- Xu, W., Edwards, M. R., Borek, D. M., Feagins, A. R., Mittal, A., Alinger, J. B., ... & Pappu, R. V. (2014). Ebola virus VP24 targets a unique NLS binding site on karyopherin alpha 5 to selectively compete with nuclear import of phosphorylated STAT1. *Cell host & microbe, 16*(2), 187-200.

- Zampieri, C. A., Sullivan, N. J., & Nabel, G. J. (2007). Immunopathology of highly virulent pathogens: insights from Ebola virus. *Nature immunology,8*(11), 1159-1164.

- Zeitlin, L., Whaley, K. J., Olinger, G. G., Jacobs, M., Gopal, R., Qiu, X., & Kobinger, G. P. (2016). Antibody therapeutics for Ebola virus disease. *Current opinion in virology, 17*, 45-49.

- Zhang, A. P., Bornholdt, Z. A., Liu, T., Abelson, D. M., Lee, D. E., Li, S., ... & Saphire, E. O. (2012). The ebola virus interferon antagonist VP24 directly binds STAT1 and has a novel, pyramidal fold. *PLoS Pathog, 8*(2), e1002550.

- Zhang, X., Bourhis, J. M., Longhi, S., Carsillo, T., Buccellato, M., Morin, B., ... & Oglesbee, M. (2005). Hsp72 recognizes a P binding motif in the measles virus N protein C-terminus. *Virology, 337*(1), 162-174.

Appendix

Directory list of CD contents:

- **Chapter 2:** Supplementary information.
  - List of sequences used for the plasmids of the mini genome system:
    - Supportive plasmids
    - Mini genome plasmid

- **Chapter 3:** Supplementary information; papers used for this chapter:

  - Dowall, S. D., Matthews, D. A., García-Dorival, I., Taylor, I., Kenny, J., Hertz-Fowler, C., ... & Barr, J. N. (2014). Elucidating variations in the nucleotide sequence of Ebola virus associated with increasing pathogenicity. *Genome biology*, *15*(11), 1-12.

  - Carroll, M. W., Matthews, D. A., Hiscox, J. A., Elmore, M. J., Pollakis, G., Rambaut, A., ... & Abdellati, S. (2015). Temporal and spatial analysis of the 2014-2015 Ebola virus outbreak in West Africa. Nature.

- **Chapter 4:** Supplementary information; papers used for this chapter.

o García-Dorival, I., Wu, W., Dowall, S., Armstrong, S., Touzelet, O., Wastling, J., ... & Hiscox, J. A. (2014). Elucidation of the Ebola virus VP24 cellular interactome and disruption of virus biology through targeted inhibition of host-cell protein function. *Journal of proteome research*, *13*(11), 5120-5135.

- **Chapter 5:** Supplementary information,
    o **Table A1.** *Complete table of significant proteins that interact with EGFP-EBOV/NP.*
    o **Table A2.** *Complete table of significant proteins that interact with EBOV/NP-EGFP.*
    o **Table A3.** *Complete table of significant proteins that interact with EGFP-RESTV/NP.*
    o **Table A4.** *Complete table of significant proteins that interact with RESTV/NP-EGFP.*