Identification and characterisation of candidate monoclonal antibody *Ebolavirus* therapeutics

A thesis submitted in fulfilment of

the requirements for the degree

of Doctor of Philosophy

Francesca R. Donnellan

Lincoln College, University of Oxford

Nuffield Department of Medicine
The Jenner Institute

April 2021

Word count: ~48,000 words main text

Abstract

Identification and characterisation of

candidate monoclonal antibody *Ebolavirus* therapeutics

By Francesca R Donnellan

Multiple species of *Ebolaviruses* can cause outbreaks of devastating disease with high mortality. The 2014 outbreak in West Africa, focused attention on efforts to develop effective vaccines and therapies to save lives. Since then, vaccines and first generation therapies have been developed and approved for use. The only therapies to have shown any efficacy in human clinical trials are monoclonal antibodies (mAbs) targeting the Glycoprotein (GP) on the virus surface. However, they are limited in potency and breadth of activity as they target only *Zaire ebolavirus* (EBOV), yet *Sudan ebolavirus* (SUDV), *Bundibugyo ebolavirus* (BDBV) and *Tai Forest ebolavirus* (TAFV) also cause human disease. The aims of this thesis were to generate a panel of mAbs that could bind to GP from all species of *Ebolaviruses* that cause human disease, then to characterise the mAbs to select candidates to enter the pipeline of next generation therapies progressing to *in vivo* study. Using high throughput methods a panel of nineteen anti-GP mAbs were isolated, six of which could recognise GP from all four *Ebolaviruses* that have caused human disease. These mAbs were characterised in a range of *in vitro* assays focusing on their ability to neutralise pseudoviruses coated with *Ebolavirus* GPs. The panel of mAbs was subjected to a range of complementary biochemical assays to map their epitopes, and the epitope of one mAb, 11886, was characterised in more detail. 11886 is of interest to progress to *in vivo* study as it is able to neutralise all pseudoviruses so far tested and investigation of its epitope suggests that it binds a similar footprint to other broadly protective mAbs in the literature, but neutralises the virus via a different mechanism or mode of binding.

Dedication

For Hilda and Edwin, who passed away before it was finished.

And for Nene, Sosuke, Harvey, Ada and Kit,

three of whom were born since it was started.

Acknowledgements

First, I would like to thank my supervisors Simon Draper, Alain Townsend and Daniel Lightwood. I have been immensely fortunate to have spent time in each of their lab groups. Simon, the supervision and mentorship I have received from you has been unwavering; the motivation and successes of the Draper Lab are testament to your dedication, management and insight. I would also like to thank you for your continual reassurances, and for being willing to read this entire Thesis in the last few weeks before submission. Alain, I feel greatly privileged to have been able to spend time with you and your group. The model you provide of collaboration, conversation and good humour is one to which I can only aspire. You'll be guiding me every time I do a 'Friday afternoon experiment'. Dan, you have only ever opened doors for me; I am very grateful for your willingness to let me loose in your lab at UCB, for continually providing more resources, and for letting me keep on coming back long after the initial three months.

From the Draper Lab I would like to thank the scientists, past and present, who have given me so much of their time, and always been willing to answer my questions; in particular Martino Bardelli, Carolyn Nielsen, Dave Pattinson, Gen Labbé and Kirsty McHugh. Thanks also to Kate Skinner for constantly fighting battles with bureaucracy and finance on my behalf. The last six months of this DPhil involved a lab move and during COVID Lockdown the friendship and teamwork of those of who ended up in Biochemistry has been especially important; so thanks to Kirsty, Giacomo, Jordan, Dimitra, Martino, Amelia, Doris, Ana and Vicky. Thanks also to the Foster and Berks groups in who helped us settle in. Barney and Rob, I hope you are as proud as I am to be a member of the Draper Lab graduating class of 2020-2021; I couldn't have asked for better fellow students. Somewhere along the line your combined sportiness rubbed off and I suppose I should thank you for that too. Dan and Sarah, thank you for taking me under your wings from when I was first a Part II in the group.

The generosity and ingenuity of the Townsend Lab really knows no bounds; my sincere thanks to Pramila Rijal, Jack Tan, Julie Xiao, Lisa Schimanski and Holly Sadler. This Thesis would not exist without the reagents, protocols and coffee room breaks you each have shared with me. It has been a

great pleasure to be part of a group that understands good food as a motivator; next round of momos are on me.

At UCB, I would especially like to thank Vicky O'Dowd for all her patient guidance and good sense throughout this DPhil. Thanks to Sarfaraj Topia for his help with protein purification; Anastasios Spiliotopoulos for teaching me how to conduct peptide phage display and for sharing his phage libraries; and Becki Munro and Leah Manning for their help with RAB-9 cells. During my time at UCB, I was made to feel truly welcome and included; Leah and Alice deserve special mention, but my thanks go to the whole Core Antibody Discovery Team for their camaraderie and expertise.

I am privileged to have spent the last five years working somewhere I consider so many to be friends not just colleagues, my thanks to all at the Jenner Institute who made it such a social and supportive atmosphere. I would especially like to thank some current and past members for their support and friendship, in and out of work; Sean, Bec, Darren and Iona, as well as Hannah and Ciaran. My thanks also to Julie Furze for keeping the whole operation running.

I would be remiss not to thank my collaborators Erica Saphire and Amar Parvate; getting to see the EM structure during a COVID lockdown, in which I was not able to work in the lab on my DPhil, was particularly exciting.

Special thanks go to my housemates, who had to put up with me every day of this DPhil; Chris, Ash (who gets special credit for putting up with me through a global pandemic and the final months of this Thesis all by himself), Emma, George and Dom. Thanks also to Brigitte and Emma (again) for the PhD friendship support group.

Finally, I would like to thank my family. I am grateful to Margaret and Tom, my aunt and uncle, who proof-read sections of this Thesis. The love and support I have always had from my brothers Jonathan, Ben and Ed, means the world to me, and I thank you for all the ways you've made life better for your little sister – much of what I have so far achieved has been made easier by your accomplishments. The same goes for Zoe and Michelle; thank you both for your unfailing kindness and encouragement. Last, but not least, thank you to Mum and Dad, quite simply, for everything.

Authorship Statement

All animal care and procedures were undertaken by staff at UCB. James Snowden, UCB, assigned predicted germline sequences to each rabbit monoclonal antibody variable region in Chapter 3. Sorting of MDCK SIAT-1 cell lines generated throughout this Thesis was conducted by Craig Waugh, WIMM FACS facility. Dr Terry Baker, UCB, designed and provided the synthetic fusion loop peptide used in Chapter 4. Dr Anastasios Spiliotopoulos, UCB, processed the NGS data and conducted Z-score analyses for sequences derived from peptide display in Chapter 5. The cryo-electron microscopy data collection and model building in Chapter 5 was conducted by Dr Amar Parvate, La Jolla Institute for Immunology, using Fab produced by me.

These contributions have been clearly designated in the main text.

1 Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Full name
4NPP	4-Nitrophenyl phosphate
ADCC	Antibody dependent cellular cytotoxicity
ADCD	Antibody dependent complement deposition
ADCP	Antibody dependent cellular phagocytosis
ADNP	Antibody dependent neutrophil phagocytosis
BDBV	*Bundibugyo ebolavirus*
BOMV	*Bombali ebolavirus*
BSA	Bovine serum albumin
CADP	Core Antibody Discovery Platform
CDR	Complementarity-determining region
CFU	Colony forming units
DC	Dendritic cell
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol
E/M	EBOV GP with MARV GP fusion loop
EBOV	*Zaire ebolavirus*
ED	Ebola disease
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbance assay
EM	Electron microscopy
FACS	Fluorescence Activated Cell Sorting
FBS or FCS	Fetal bovine serum or Fetal calf serum
Fc	Fragment crystallisable
FcR	Fc Receptor
FITC	Fluorescein
FL	Fusion loop
FP	Fusion peptide
GALT	Gut-associated Lymphoid Tissue

GC	Glycan cap
GP	Glycoprotein
h	hour (s)
HA	Hemagglutinin
HCAb	Heavy Chain antibody
HEK	Human Embryonic Kidney
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HM buffer	HEPES MES buffer
HIV	Human Immunodeficiency virus
HPLC	High performance liquid chromatography
HR	Heptad repeat
IC50	Half maximal inhibitory concentration
IgNAR	Ig-based New Antigen Receptor
IgG	Immunoglobulin G
IgM	Immunoglobulin M
LN_2	Liquid nitrogen
M/E	MARV GP with EBOV GP fusion loop
mAb	Monoclonal antibody
MDCK	Madin Darby Canine Kidney
MES	2-(N-morpholino)ethanesulfonic acid
min	minutes
MLD	Mucin-like domain
MPER	Membrane proximal external region
MWCO	Molecular weight cut off
NGS	Next generation sequencing
NHP	Non-human primate
NK	Natural Killer
NPC1-C	Neimann Pick C1 protein domain C
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PFU	Plaque forming units
RBR	Receptor binding region

RESTV	*Reston ebolavirus*
rpm	Revolutions per minute
RPMI	Roswell Park Memorial Institute media
RT	Room temperature
RT-PCR	Reverse transcription polymerase chain reaction
s	seconds
SDM	Site-directed mutagenesis
SDS	Sodium dodecyl sulphate
SEC	Size Exclusion Chromatography
SOC	Super Optimal broth with Catabolite repression
SUDV	*Sudan ebolavirus*
TAE	Tris base, acetic acid and EDTA
TAFV	*Tai Forest ebolavirus*
TBS	Tris buffered saline
TCID50	Median tissue culture infectious dose
THL	Thermolysin
TM	Transmembrane
TMB	3,3',5,5' Tetramethylbenzidine
UPLC	Ultra performance liquid chromatography
V	volts
V region	Variable region
VGM	Viral growth media
VIC	Viral Hemorrhagic Fever Immunotherapeutic Consortium
VLP	Virus-like particle
VSV	Vesicular stomatitis virus
WGA	Wheat germ agglutinin

1. Introduction

1.1 *Ebolaviruses* and Ebola Disease

Ebolaviruses are zoonotic, enveloped, filamentous viruses of the family *Filoviridae* (Figure 1). Approximately 19kb in size, *Ebolavirus* genomes are linear, non-segmented, single-stranded, negative-sense RNA genomes, and encode seven structural proteins; Nucleoprotein (NP), VP35, VP40, Glycoprotein (GP), VP30, VP40, VP24 and an RNA-dependent RNA polymerase (L). Additional open reading frames in the GP gene generated by RNA editing result in production of two additional non-structural proteins; soluble GP (sGP) and small soluble GP (ssGP) [13-15].

Of the six *Ebolavirus* species described to date, *Zaire ebolavirus* (EBOV), *Sudan ebolavirus* (SUDV), *Bundibugyo ebolavirus* (BDBV) and *Tai Forest ebolavirus* (TAFV) can infect and cause disease in humans. EBOV, SUDV and BDBV have caused recurrent outbreaks of fatal Ebola Disease (ED). Although not yet confirmed by isolation of infectious virus, evidence for multiple bat species as asymptomatic reservoirs of EBOV includes viral RNA and *Ebolavirus* antibodies in wild bats, experimental infection of bat species and *in vitro* infection of bat cell lines [16-19]. EBOV infection of other non-human primate (NHP) species reflects human disease and in the wild has been caused by spillover from reservoir species [20, 21]. Previous outbreaks of ED have been linked to an initial zoonotic event, followed by human-to-human transmission via close contact with infected bodily fluids of ED patients or the bodies of those who have died from ED [21-24].

TAFV, whilst causing fatal haemorrhagic disease in NHPs, has caused only one known case of non-fatal human disease [25].

Reston ebolavirus (RESTV) was discovered after outbreaks of fatal disease in NHPs, and its identification in captive domestic swine populations alongside experimental infection of pigs suggests pigs can serve as an asymptomatically infected reservoir of RESTV [26-28].

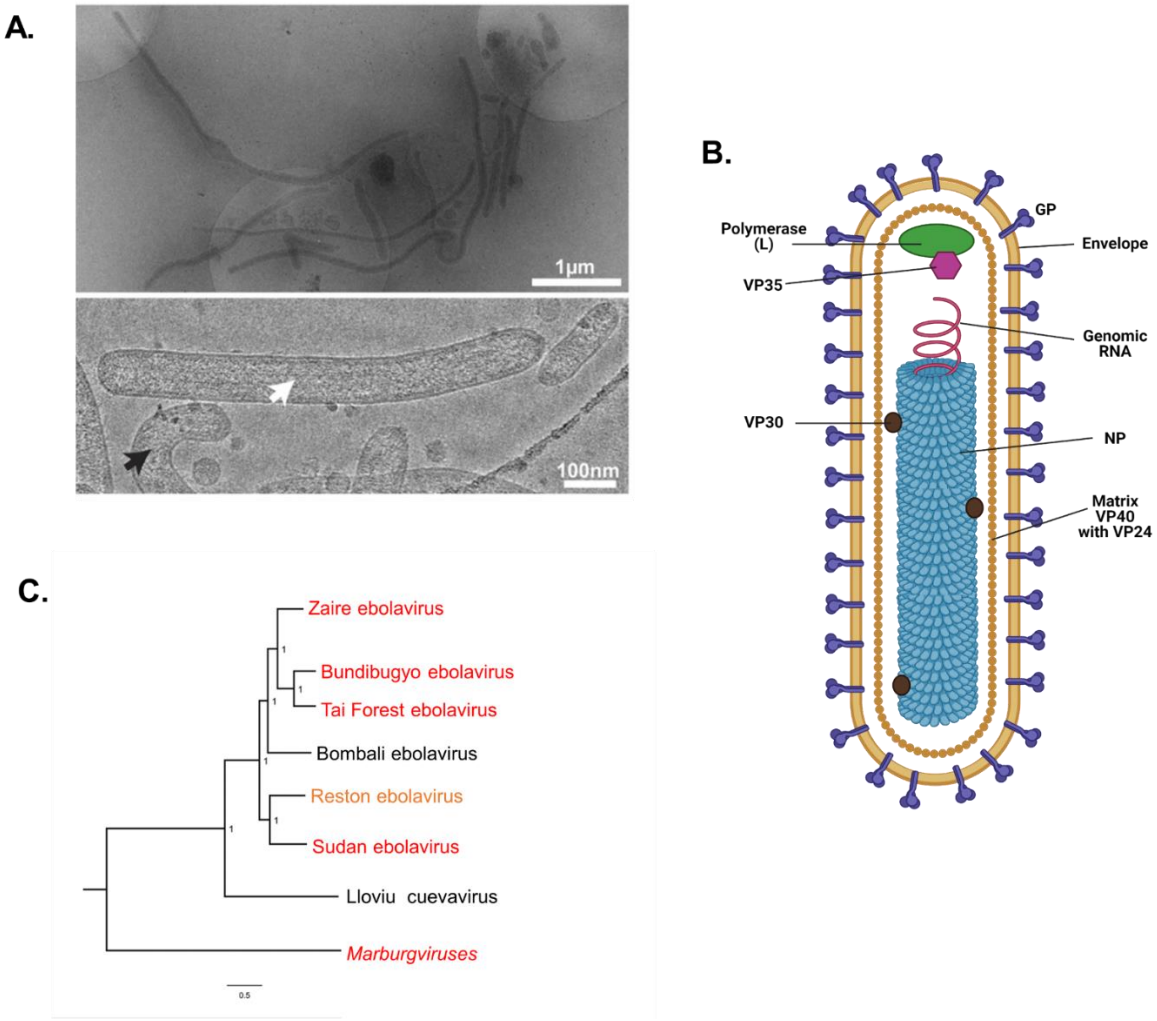


Figure 1. Structure and phylogeny of Ebolaviruses. A. Electron micrographs of purified EBOV*, reproduced from Bharat et al. [7].* ***B. Schematic of an Ebolavirus filamentous particle*** *showing genomic RNA surrounded by nucleocapsid, drawn using BioRender.com.* ***C. Phylogenetic relationship of Filoviruses*** *including all known Ebolaviruses based on complete genomes. Red; cause disease in humans. Orange; known asymptomatic infection of humans. Posterior probabilities are shown at the nodes. Scale bar indicates genetic distance. Adapted from Forbes et al. [9].*

No human disease has been identified despite evidence of infection by seroconversion in humans in contact with RESTV-infected animals [29].

The recently discovered *Bombali ebolavirus* (BOMV) has been identified via presence of viral RNA in two bat species (*Mops condylurus* and *Chaerephon pumilus*) in Sierra Leone, Guinea and Kenya indicating a wide geographic range for the virus [9, 30, 31]. No evidence of infection was found in febrile humans in the area surrounding the study site in Kenya, despite reported contact with bats and *M. condylurus* commonly nesting in human structures [9]. However, virus pseudotyped with GP from BOMV is able to infect human cells *in vitro* and cell entry is dependent on the presence of human Niemann Pick C1 protein (NPC1), which has previously been defined as the cellular receptor for other *Ebolaviruses* [30].

1.1.1 Outbreaks

The 2014-16 West African EBOV outbreak killed more than 11,000 people; however, since the discovery of *Ebolaviruses* in 1976, a year in which both EBOV and SUDV emerged in central Africa, there have been twenty-eight ED outbreaks caused by EBOV, SUDV and BDBV, each with high mortality (Figure 2). Along with the continued emergence of new *Ebolaviruses* with the potential to infect humans, this highlights that, although EBOV has been responsible to date for the greatest number of deaths (due to the 2014-2016 and 2018-2020 outbreaks), strategies to devise prophylactics or therapeutics for ED must consider multiple *Ebolavirus* species and the risk of new emerging species.

The 2014 outbreak illustrates the potential of *Ebolaviruses* to cause large international epidemic health emergencies in resource poor settings. Worryingly, there have been almost annual outbreaks of ED since the year 2000, with the Democratic Republic of Congo recently suffering the second largest recorded outbreak of EBOV [32], which despite having been declared over in June 2020 appears to have led to a case in January 2021 as determined by sequencing [33]. Additionally, EBOV has re-emerged in Guinea causing a cluster of cases in February 2021; sequencing of the virus in this instance suggests it is the same virus that was circulating in the previous 2014 outbreak in the same

A.

Year	Country	Cases*	Deaths	Case fatality
2021-ongoing	**Guinea**	**23**	**12**	**52%**
2021-ongoing	**DRC (North Kivu)**	**12**	**6**	**50%**
2020	DRC (Équateur Province)	130	55	42%
2018-2020	DRC (North Kivu)	3470	2287	66%
2018	DRC (Bikoro)	54	33	61%
2017	DRC	8	4	50%
2014-2016	Guinea, Sierra Leone, Liberia**	28,616	11,310	40%
2012	DRC	57	29	51%
2012	Uganda	7	4	57%
2012	Uganda	24	17	71%
2011	Uganda	1	1	100%
2008	DRC	32	14	44%
2007	Uganda	149	37	25%
2007	DRC	264	187	71%
2005	Congo	12	10	83%
2004	Sudan	17	7	41%
2003 (Nov-Dec)	Congo	35	29	83%
2003 (Jan-Apr)	Congo	143	128	90%
2001-2002	Congo	59	44	75%
2001-2002	Gabon	65	53	82%
2000	Uganda	425	224	53%
1996	South Africa (ex-Gabon)	1	1	100%
1996 (Jul-Dec)	Gabon	60	45	75%
1996 (Jan-Apr)	Gabon	31	21	68%
1995	DRC	315	254	81%
1994	Cote d'Ivoire	1	0	0%
1994	Gabon	52	31	60%
1979	Sudan	34	22	65%
1977	DRC	1	1	100%
1976	Sudan	284	151	53%
1976	DRC	318	280	88%

B.

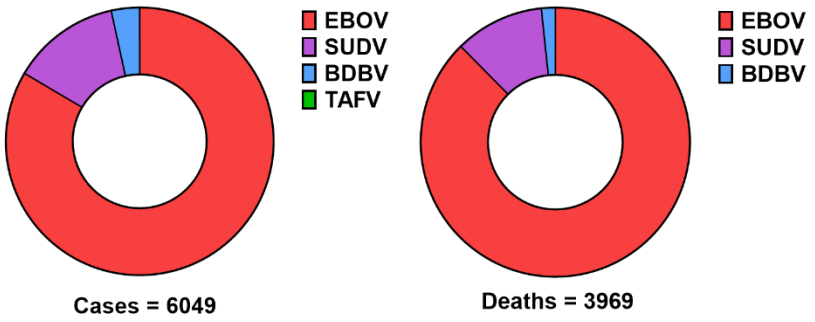


Figure 2. Recorded outbreaks of Ebola disease. A. Summary of recorded outbreaks from information provided by WHO and CDC*. Bold; Ongoing outbreak. *Suspect, probable and confirmed cases. **Italy, Spain, DRC, USA, UK, Senegal, Mali, Nigeria. Accurate as of 5th March 2021.* ***B. Contributions of each species to total recorded cases and deaths caused by Ebolaviruses excluding the 2014-2016 West African outbreak and ongoing outbreaks.*** *Coloured by causative species of Ebolavirus. Red; EBOV. Purple; SUDV. Blue; BDBV. Green; TAFV.*

country [34]. Together with increasing evidence of asymptomatic infection, viral persistence in ED survivors and sexual transmission months after initial infection, the risk of ED outbreaks re-occurring in human populations without a new spillover event from animal reservoirs increases the challenge of controlling ED outbreaks [35-40].

Concurrently, the population at risk, the geographic range, and length of outbreaks, have also increased due to a range of factors including increasing connectivity between sparsely populated remote areas and densely populated urban areas [41]. The response to ED therefore requires continued robust surveillance, education, access to effective vaccines, and effective therapeutics to treat those that become infected and develop ED. Before 2014, there were no licensed vaccines or therapeutics for *Ebolavirus* infection, however the situation has changed in the last five years due to research efforts given momentum by the 2014 outbreak. Vaccines approved as a result of clinical trials conducted during that outbreak have been deployed to limit the impact of subsequent outbreaks in the DRC and Guinea [42-44]. A clinical trial during the 2018-2020 outbreak in the DRC, has also led to the licensure of two antiviral therapies to treat ED [45-47]. These therapies, both comprising of monoclonal antibodies, will be reviewed further in Section 1.4.3.1.

1.1.2 Ebola disease

Ebola disease (ED) is characterised by systemic viral replication, infection of immune cells, immune-mediated damage, electrolyte imbalances, fluid loss, hypovolemic shock, organ failure and high mortality. ED begins with an asymptomatic incubation period of up to twenty-one days, followed by fever, headache, myalgia, malaise, weakness, dizziness, progressing to diarrhoea, abdominal pain and vomiting with large fluid loss [48]. At this acute stage some patients recover, however, others progress to hypovolemic shock and may present with haemorrhagic symptoms including maculopapular rash, conjunctival bleeding, petechiae, mucosal bleeding especially gastrointestinal bleeding, and continued oozing after venepuncture [49-52]. Patients can recover from shock, however, in fatal cases, mean time to death after symptom onset is approximately ten days, and higher viral loads are strongly associated with poorer patient outcomes [53]. High rates of maternal and perinatal

mortality have been recorded and EBOV is able to cross the placenta [54]. The majority of available clinical data describe EBOV infection as it was the cause of the most recent and largest outbreaks of ED, however SUDV and BDBV cause similar disease [55, 56]. Survivors of ED can experience sequelae including althragias, uveitis leading to vision loss, and hearing loss [57-59].

The proportion of patients with haemorrhagic symptoms has varied between outbreaks and may be affected by the exact mode of transmission [60]: e.g. the 1976 EBOV outbreak had a high number of fatal cases with haemorrhagic symptoms correlating with a high number of cases associated with contaminated needle transmission [61]. As reviewed by Munoz-Fontela and McElroy, in the 2014 West African outbreak, less than 15% of patients in Sierra Leone and Liberia showed haemorrhagic symptoms (although reports were higher from some centres in Guinea) and there was a lack of correlation between degree of haemorrhage and disease severity [60]. This has culminated in a recent renaming of the disease caused by *Ebolaviruses* to ED, dropping the 'haemorrhagic fever' moniker found in earlier designations.

1.1.3 *Ebolavirus* pathogenesis

Ebolaviruses enter the body through mucosal surfaces, breaks in the skin or injection. The initial infection of dendritic cells and macrophages leads to dissemination of virus and immune dysregulation, which in turn may cause endothelial dysfunction leading to hypovolemic shock and systemic infection of multiple organs. The virus has a broad cell tropism which ultimately leads to multiple organ failure and onward transmission of virus via multiple bodily fluids. In the following sections, this current hypothesis for the pathogenesis of *Ebolavirus* infection is discussed.

1.1.3.1 Immune dysregulation

Dendritic cells (DCs) and macrophages are early targets of EBOV in NHPs and support EBOV replication *in vitro* [62-64]. DCs migrate to the draining lymph nodes, potentially disseminating the virus from the initial point of entry to draining lymph nodes and to the liver and spleen via the lymphatic and vascular systems [65]. EBOV-infected *in vitro*-derived DCs show reduced activation and antigen-presentation

[63]. Poor activation of antigen-presenting cells has been associated with more severe ED, and conversely, higher levels of circulating markers of DC stimulation correlated with improved SUDV outcomes [66-68]. Histopathology of severe ED cases shows lymphoid depletion of spleen and lymph nodes; lymphocytes appear not to be direct targets of infection, but undergo bystander apoptosis leading to depletion of Natural Killer (NK) and T lymphocytes, reflected in the circulating cell populations [69].

Aside from directly infecting cells, the virus structural proteins VP24 and VP35 are thought to have roles in evading host innate responses and type I interferon response [70]. The immune response to *Ebolavirus* infection is discussed in more detail in section 1.2.

1.1.3.2 Endothelial Dysfunction

Increased vascular permeability by breakdown of endothelial junctions is necessary to allow cells and molecules to reach sites of inflammation, however, widespread activation of this process in fatal ED can lead to extravasation of fluid into extravascular spaces and hypovolemic shock [51]. Pro-inflammatory chemokines and cytokines required to induce permeability may come from infected antigen-presenting cells [71]. Biomarkers of activated endothelium and barrier breakdown are increased in the plasma of patients with severe ED caused by EBOV and SUDV [66, 67, 72]. Endothelial cells themselves are infected during the terminal phase of ED, after the barrier has already been dysregulated [62, 73].

1.1.3.3 Systemic infection and organ damage

Study of ED pathology by autopsy is limited due to the necessary biosafety requirements to keep personnel safe. However, formalin-fixed tissues are not infectious and so histology and immunohistochemistry (IHC) using monoclonal antibodies can be used to study ED pathology safely [73].

Filoviruses are hepatotropic, with liver necrosis characteristic of infection, and surrounding hepatocytes showing viral inclusion bodies in EBOV, SUDV and BDBV infected individuals [73, 74]. Infection of cells in the lamina propria throughout the gastrointestinal tract correlates with human-to-

human transmission via contact with bloody stools [73]. Infection of endothelial cells of the kidney correlates with virus in the urine and necrotic damage to the kidneys contributes to renal failure in ED patients [22, 73]. In fatal cases of EBOV-related ED, infected alveolar lung macrophages and endothelial cells have been identified in sites of lung oedema and haemorrhage [73]. Systemic infection of immune cells and endothelial cells extends to the skin; IHC shows viral antigens in epidermal DCs, endothelial cells and connective tissue fibroblasts, as well as sweat and sebaceous gland epithelium [22, 74]. In an infected monkey, IHC and electron microscopy confirmed the presence of virus in seminiferous tubules of the testes, which is consistent with reports of virus in the semen and sexual transmission of the virus [75].

1.1.3.4 Electrolyte imbalances

The volume of fluid loss and renal damage together likely explain the electrolyte imbalances reported in ED patients. These imbalances have a range of negative consequences for the patient, and their monitoring and management may lead to much improved survival rates alone [60].

1.1.3.5 Haemorrhage and coagulopathy

Whilst occurring in only a subset of ED patients, haemorrhagic symptoms indicate some form of dysregulation of blood clotting, as in NHPs haemorrhage is not simply a result of endothelial cell cytolysis [62]. As platelet counts have been found to be normal in ED patients, the mechanism remains unclear [76]. Thrombomodulin acts as an anticoagulant and is released by activated endothelial cells, and von Willebrand Factor mediates interactions between platelets and damaged endothelium; high levels of both proteins are found in the plasma of SUDV and EBOV-related ED patients with severe outcomes, and associates coagulopathy with the endothelial dysfunction previously discussed [66, 67, 72, 77].

1.2 Immune response to *Ebolavirus* infection

With innate immune cells an early target of *Ebolavirus* infection, immunopathology and immunomodulation are major characteristics of ED. The immune response is implicated in both pathogenesis and protection in ED; patients mount a robust adaptive immune response, although the

character of the response may be associated with different outcomes (Figure 3). The immune response to *Ebolaviruses* is distinct from the response to other viruses such as Lassa virus, an *Arenavirus*, despite both targeting similar cell types and organs in humans [78].

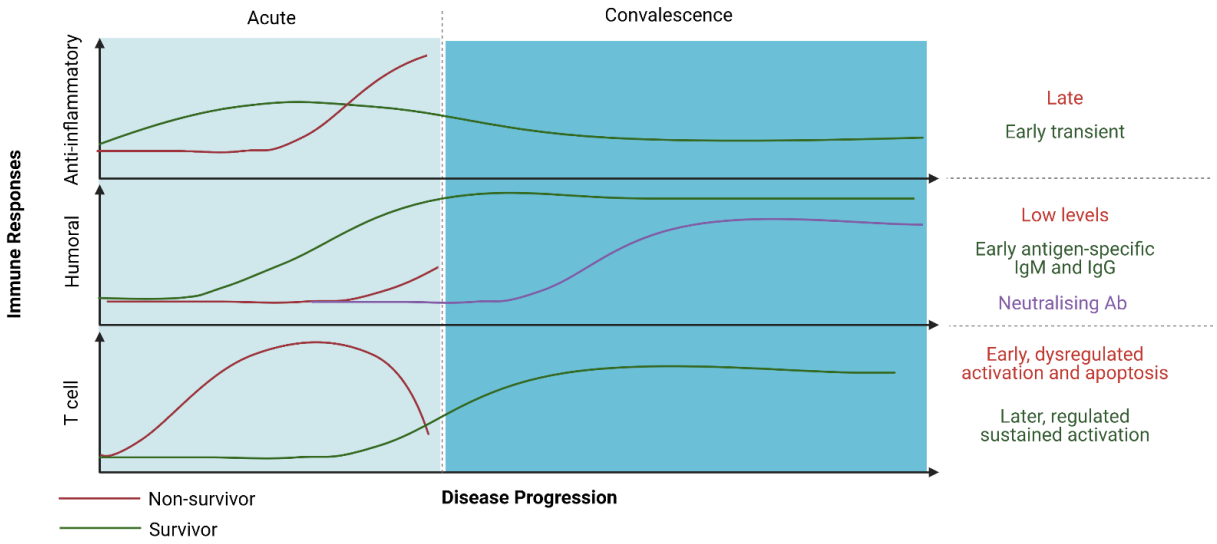


Figure 3. Adaptive immune responses to Ebolavirus infection in survivors and non-survivors*. Drawn using BioRender.com*

1.2.1 Innate response

Studies of cytokine and chemokines released by myeloid cells (DCs and macrophages) that are the initial targets of infection in ED suggest a general dysregulation of immune signalling, which potentially drives the recruitment of further myeloid cells to sites of infection and may increase viral dissemination [69, 79], although this could also be a result of uncontrolled viral replication rather than simply a driver. Sustained high levels of circulating innate inflammatory chemokines and cytokines are correlated with fatal outcomes in ED [67, 80]. However, in asymptomatic cases, very high upregulation of inflammatory chemokines and cytokines is followed by downregulation associated with control of viral replication, suggesting that sustained inflammation may be mediated by uncontrolled viral replication [60, 81]. Comparison of wild type (WT) EBOV infection in mice before and after virus adaptation to the host (MA-EBOV) suggests viral clearance of the WT virus with infection restricted to myeloid cells, but progression to hepatocytes by MA-EBOV, suggesting that ability of the virus to modulate the response of myeloid cells is critical to its dissemination [78].

The role of the antiviral interferon response in ED is unclear, with the dynamics of the response requiring further investigation due to the link between inflammation and controlled or uncontrolled viral replication [60]. However, evidence for a role of type I IFN responses in protection comes from the lethality of EBOV and SUDV in type I IFN receptor deficient mice, where wild type mice are resistant to disease [82]. In addition, the virus has evolved mechanisms to directly antagonise IFN responses. Viral sensing by host pattern-recognition receptors, in particular retinoic acid-inducible gene-I (RIG-I)-like receptors (RLRs), stimulates expression of IFN and other cytokines that in turn amplify expression of interferon-stimulated genes (ISGs) necessary for the innate antiviral response, including upregulation of expression of co-stimulatory molecules on DCs [83]. This pathway is antagonised by EBOV VP35 via multiple mechanisms: binding to double stranded viral RNA to prevent recognition by the RLRs; acting as decoy substrate to components of the signalling pathway downstream of RLRs; binding to a positive regulator of the RLRs to reduce activation; and promoting sumoylation of the activated transcription factors to prevent expression of ISGs [70]. EBOV VP24 also antagonises IFN signalling in infected cells to prevent the upregulation of ISGs [84].

In addition, innate immunity is antagonised by EBOV by apoptotic mimicry. The EBOV envelope contains phosphatidylserine (PS), which is bound by the TYRO3/AXL/MER (TAM) family of receptor tyrosine kinases on target innate cells, such as macrophages, which use PS sensing to identify apoptotic cells. In addition to enhancing uptake of virus into macrophages, TAM receptor binding also suppresses type I IFN signalling [85].

In contrast to high levels of circulating pro-inflammatory markers, necrotic damage in the tissues of multiple organs with direct evidence of infection seems often accompanied by limited inflammation, which has been argued to reflect the overall dysregulation of the immune response to *Ebolavirus* infection [73].

1.2.2 Cellular immunity

DCs act as the antigen-presenting cells for activation of naïve CD8+ T cells which are required for clearance of virus infected cells [86]. Infected DCs, although able to migrate to lymph nodes, are

arrested in an immature state with reduced antigen presentation [64]. However some antigen presentation can still occur via macrophages which, unlike DCs, are activated despite infection [64]. Robust T cell activation has been observed both for fatal cases and for survivors of ED, during both acute disease, and up to two years post-infection; therefore the effect of EBOV infection on DCs does not prevent T cell activation, but T cell activation does not necessarily lead to viral clearance [87, 88]. Increased levels of negative immune checkpoint markers PD-1 and CTLA-4 in peripheral blood T cells of ED patients were associated with fatal outcomes, suggesting dysregulation after activation [87]. As described in Section 1.1.3.1, T cells are not infected by EBOV, but initial early expansion of lymphocytes is followed by lymphopenia likely caused by their apoptosis, impacting the ability of patients to mount an effective response [69, 79]. In an outbreak of EBOV in Gabon, initial indicators of T cell activation were followed by large decreases in markers of activated T cells (CD3, CD8 and T cell receptor mRNA) and down-regulation of BCL-2 (anti-apoptotic protein) in peripheral blood mononuclear cell (PBMC) samples from ED patients who developed fatal disease compared to those who survived [79]. However, some loss of circulatory T cells may occur by their recruitment from the periphery into infected tissues [78].

Serum levels of IFN-γ and IL-2, which are common markers of Th1 cell and NK cell responses, were observed to be lower during the symptomatic phase of disease for EBOV-related ED survivors, and increased in fatal cases, although not for SUDV-related ED [67, 69, 79, 89, 90]. High levels of IFN-γ from activated T cells, may ultimately lead to T cell apoptosis, and high IFN-γ levels early in the disease course correlates with worse outcomes [79, 89]. In addition, higher levels of IL-10, an anti-inflammatory cytokine produced by DCs and macrophages, are measured in fatal cases later in infection, but early transient IL-10 responses are associated with survival [67, 79, 81, 87, 89]. An early hyper-activation of T cells in response to *Ebolavirus* infection therefore seems to be dysregulated, and countered by apoptosis and immune suppression later in the disease course in those with worse ED outcomes. In comparison, survival is associated with regulated, lower initial T cell activation, that transitions to an inflammatory response that does not trigger a corresponding anti-inflammatory response, leading to sustained T cell activation rather than lymphopenia [78] (Figure 3).

1.2.3 Humoral immunity

In survivors of ED, IgM appears a few days after symptom onset and disappears within a few months, and specific IgG can be detected a week after symptom onset and persist for years [91-94]. The primary targets of naturally acquired antibody are GP and NP [95]. Early induction of IgM and class-switching to IgG correlated with viral clearance and positive outcomes in related ED, and absent or low IgM and IgG responses (consistent with poor induction due to dysregulated innate and T cell responses) have been reported in fatal ED cases [79, 80, 95, 96]. Asymptomatic infection defined by detection of EBOV specific IgM and IgG, and has been detected approximately three weeks post estimated exposure [39, 81, 91]. Together this suggests that a robust early humoral response is part of a protective response to *Ebolavirus* infection.

In contrast to this hypothesis, however, limited data from small field studies report cases of SUDV-related ED survivors with no IgG detected by ELISA [97], and that fatal and non-fatal cases showed no significant differences in humoral responses; both groups showed late IgG responses with no association with viral clearance [98]. In general though, the weight of evidence is that successful induction of a robust, but regulated, adaptive immune response, encompassing T cells and humoral immunity, is associated with positive ED outcomes.

Consistent with a role for humoral immunity, ED patients can produce antibodies that neutralise the virus; neutralising monoclonal antibodies against GP that are protective against disease in animal models have been isolated from survivors of EBOV and BDBV infection [93, 99-102] and detected in the serum when tested three months to 11 years after recovery [101, 103]. Naturally occurring neutralising antibodies from ED survivors that cross-react to other *Ebolaviruses* are among those reported and will be discussed in more detail in Section 1.4.3.2. However, when sampling serum throughout convalescence, neutralising antibody responses in three acute ED survivors, but not total EBOV GP specific IgG, were initially low and increased over time up to nine months after recovery [104]. As acute patients show high circulating levels of plasmablasts [80], this suggests non-neutralising antibodies contributes more to initial protection. ED survivors can generate broader B cell clonality than individuals infected with HIV-1 or influenza, which along with development of higher

neutralising antibody in convalescence than in acute illness, may reflect ongoing affinity maturation of antibody due to the viral persistence already discussed, and may contribute towards protection from future infection [60, 101].

This potential for cross-reactive antibody-mediated immune memory to provide protection from other *Ebolaviruses* is important for vaccination efforts. It also strengthens the arguments made for pursuing monoclonal antibody therapies and rational vaccine design efforts to increase cross-protection against currently known and emerging *Ebolaviruses*. Non-naturally acquired antibodies in protection from ED will be discussed in greater detail in section 1.4.

1.3 Glycoprotein

The GP is the only transmembrane *Ebolavirus* protein and target for neutralising antibody. GP is a Class I viral membrane protein which is responsible for entry of virus into host cells as the mediator of cell attachment, cell receptor binding and viral and host cell membrane fusion. As such, it is a major target for vaccines and therapeutics, both licensed and in development.

1.3.1 Structure

GP is post-translationally cleaved by host furin proteases into GP1 and GP2 subunits [105]. Disulphide bonded GP1,2 heterodimers associate non-covalently to form a trimer mediated by GP1-GP2 and GP2-GP2 contacts [1]. GP1 (residues 30-501) is responsible for cell attachment and receptor binding; it contains the GP core, receptor binding region (RBR), glycan cap (GC) domain and mucin-like domain (MLD) and forms a chalice-like bowl architecture in the pre-fusion conformation [12, 106]. GP2 (502-676) cradles GP1 and is responsible for fusion and contains the cytoplasmic tail, transmembrane (TM) domain, membrane proximal external region (MPER), heptad repeats (HR1 and HR2) and fusion loop (FL) (Figure 4) [1, 107, 108]. During the infection cycle, the GP undergoes cleavage events followed by major conformational rearrangements upon receptor binding and pH triggers, from a pre-fusion to post-fusion structure via a fusion competent intermediate [108-113]. These rearrangements will be introduced here and their role in the infection cycle addressed in Section 1.3.2.

The MLD is heavily glycosylated and highly variable between species. Its overall disordered structure remains unresolved in high resolution structural studies, and is often removed to facilitate crystallisation, however electron microscopy shows it covers the GC and largely occludes the RBR [114]. It is dispensable for viral entry and the total mass of the MLD from one monomer is 75kDa including protein and oligosaccharide, which is approximately equal to the rest of the monomer excluding the TM domain [1].

The GC sits at the apex of the GP chalice and contains four predicted N-linked glycans. Formed by a continuous run of residues 227-310, it forms an alpha helix and four-stranded β-sheet. The β17-β18 loop of the GC (residues 279-298) is mostly disordered in crystal structures and extends down the outside of the vertex of each GP1,2 monomer to interact with the 3_{10} pocket on GP1 [115].

The rest of the GP1, the GP1 core, is sometimes divided into the base and head subdomains formed by discontinuous portions of the GP1 sequence; the base consists of two beta sheets that contact the FL and GP2 helices, and the head sits between the base and the GC and contains the RBR [1].

The conserved RBR (residues 54-201) inside the bowl of the GP chalice is shielded by the MLD and GC until they are removed during infection by the host proteases Cathepsins B and L allowing it to interact with the host receptor Niemann Pick C1 [1, 12]. This processed form of GP without the GC or MLD is referred to as GP_{CL} (Figure 4Biii). Niemann Pick C1 protein has been identified as the cellular receptor of all *Filoviruses* and the lumenal C domain (NPC1-C) alone is required for infection although this interaction appears to be relatively low affinity [12, 116-118]. NPC1 is a ubiquitously expressed 13 transmembrane protein involved in cholesterol transport which is found primarily in the late endosomes and lysosomes; mutations or absence of NPC1 causes fatal cholesterol storage disease [119]. NPC1-C binding to the GP_{CL} mimics hydrophobic interactions between the GC and the GP1 core remaining in GP_{CL} [12].

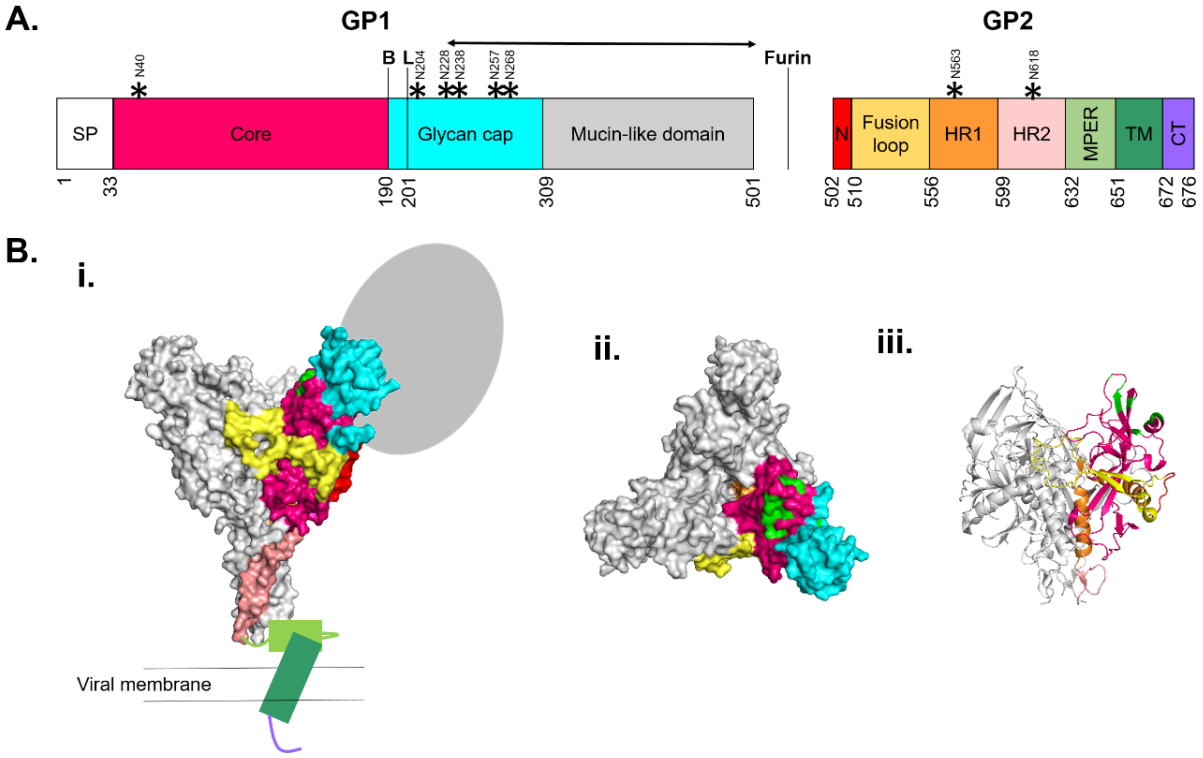


Figure 4. Pre-fusion GP structure. A. Domain schematic of GP. *B and L denote cathepsin cleavage sites. SP; signal peptide. N; GP2 N terminal domain. HR1; Heptad repeat 1. HR2; Heptad repeat 2. MPER; Membrane proximal external region. TM; Transmembrane domain. CT; cytoplasmic tail. * represent predicted N glycosylation sites [1]. Arrow indicates area predicted to contain O-linked glycosylation sites [8].* ***B. i. and ii. Side and top views of molecular surface of pre-fusion GP*** *(PDB: 5JQ7, [6]) with additional domains not in crystal structure shown.* ***iii. Cartoon representation of GP$_{CL}$.*** *(PDB: 5HJ3, Fab hidden [11]). One monomer coloured as per schematic in part A. GP1 core residues identified as part of receptor binding site highlighted in fluorescent green [1, 12]*

The FL consists of a continuous sequence of residues (511-556) close to the N terminus of GP2. In the pre-fusion conformation, it forms a disulphide-bonded antiparallel β-strand cradled against the side of the GP chalice that extends around to pack against the neighbouring GP1,2 monomeric unit [1, 6]. The hydrophobic fusion peptide (FP) is situated at the tip of the FL and is inserted into the host cell endosomal membrane [120]. The β13-14 loop (residues 190-213) connecting the GC and two portions of the core domain of GP1 extends over the base of the FL paddle and is the cleavage site for removal of the GC [110].

In the pre-fusion conformation the HRs form the cradle for GP1 (HR1) and the stem of the chalice (HR1 and HR2) contributing to the trimer interface [6]. Pre-fusion HR1 is divided into four segments (HR1$_{A-D}$), the first two of which (residues 554-575) form a 40° kinked α helix, which may act as a conformational switch [113]. At the trimer interface, the hydrophobic faces of amphipathic HR1$_D$ helices from each monomer pack together [1].

The MPER is immediately adjacent to the TM domain and, independent of pH changes, they together form a helix-turn-helix motif, with the MPER consisting of the first helix and the TM domain the turn and second helix [107].

1.3.2 Role in infection

1.3.2.1 Cell attachment and uptake

Many host cell surface proteins have been implicated in viral attachment to target cells including; DC-SIGN (dendritic-cell-specific ICAM3-grabbing non-integrin), L-SIGN (liver and lymph node SIGN), β1 integrins, T cell immunoglobulin and mucin domain-containing (TIM) proteins and TYRO3/AXL/MER (TAM) family proteins [121-125]. None are essential and these proteins likely represent degenerate mechanisms for initial attachment.

Macropinocytosis is a major mechanism of uptake of micro-organisms and apoptotic cells by the phagocytic and antigen presenting cells that are the initial targets of *Ebolavirus* infection. Multiple studies showing co-localisation of virus with markers of macropinocytosis, and inhibition of viral uptake in the presences of inhibitors of macropinocytosis, indicate *Ebolavirus* uptake occurs predominantly by this pathway [126-128] (Figure 5A). Clathrin-mediated endocytosis, but not caveolae, may be an alternative route dependent on cell origin and viral particle size [126].

The large number of attachment co-factors and the ubiquitous expression of the NPC1 receptor in the endosomal pathway provides an explanation for the broad cell tropism of *Ebolaviruses*.

1.3.2.2 Processing and receptor binding in the endosomal pathway

Entry into the endocytic pathway exposes the GP to lower pH, to host proteases Cathepsins B and L that cleave the GP to reveal the RBR, and to its cellular receptor NPC1. Along with a postulated final trigger these events are necessary for induction of fusion of viral and endosomal membranes [129] (Figure 5B).

GP processing is dependent on endosomal Cathepsins B and L [111, 130, 131]. Cleavage removes the MLD, then GC and the β14 strand in a processive manner, via a 50 kDa intermediate to a 20 kDa then 19 kDa metastable core, by two cleavage events in the β13-β14 loop [110, 112, 132]. The

processive cleavage of GP lowers the energy barrier to conformational change of the GP and the 19 kDa GP1 core, which is still associated with intact GP2 (together forming GP_{CL}), and can mediate infection at endosomal pH (~pH 5); the 20 kDa intermediate can also trigger fusion, but less favourably [110]. This suggests the presence of the GC, β14 strand and/or β13-β14 loop impedes conformational changes of the full length GP, 50 kDa and 20 kDa intermediates. The role of pH in triggering fusion has been debated, but has been demonstrated to have an active role in triggering the conformational changes of primed 19 kDa GP core [110]. Recently, in a forward genetic screen, major histocompatibility complex (MHC) class II transactivator (CIITA) has been shown to have antiviral activity against *Ebolaviruses*, by upregulating expression of the p41 isoform of MHC-II invariant chain (CD74) which inhibits endogenous endosomal cathepsins; thereby identifying a host pathway for resistance to *Ebolavirus* infection [133].

NPC1 appearance in endosomes always precedes viral fusion, which can occur in intermediate to late endosomes [129, 134]. Comparison of crystal structures of unliganded GP and of GP_{CL} in complex with NPC1-C shows that, upon receptor binding, conformational changes are induced in the RBR that result in a cascade of small conformational changes that causes the 3_{10} helix in the β3-α1 loop to also be pulled upwards. This likely abolishes a hydrogen bond between N27 and K510 at the N terminus of the FL, and introduces repulsion between K510 and R559, which may help release the FL [12]. An A82V polymorphism in the α1 helix that was associated with higher mortality during the 2014 West African outbreak, was found to reduce the threshold of NPC1 activation; affinity for NPC1-C was unchanged, but the stability of the GP was reduced, suggesting that this priming of GP_{CL} by receptor binding is important for infectivity [135].

Spence *et al*., suggest that cathepsins may have an additional role after receptor binding; infection by GP_{CL} bearing virions remains sensitive to cathepsin inhibitors, potentially due to interference with fusion pore formation at the six-helix-bundle conformational transition [129].

1.3.2.3 Membrane fusion

The metastable, primed GP_{CL} undergoes global rearrangements to first insert the FL into the cell membrane and then bring the viral and cell membranes together and form a fusion pore that allows the release of the viral RNA into the cell cytoplasm (Figure 5C). The FL takes on a 'hydrophobic fist' conformation which is necessary for the FL to be embedded in the target membrane [120, 136]. *In vitro* this conformational change and insertion happens only in low pH conditions, with pH changes likely affecting multiple charged residues in the FL to facilitate the conformational switch [109, 120].

Upon insertion of the FL into the host membrane, the HRs are extended (with FP at one end inserted into the host membrane, and the TM domain at the opposite end embedded in the viral membrane) and then reform into a six-helix-bundle, typical of Class I viral fusion proteins, providing the energy to bring the two membranes into contact [108, 113, 137]. This brings the FL and TM domain into close proximity, where, along with the MPER, they mediate the last stages of membrane fusion. The FL directly interacts with two tryptophan residues in the MPER region, and their mutation impedes GP-mediated membrane fusion [107].

1.3.2.4 Replication and Budding

The release of *Ebolavirus* nucleocapsid into the cytoplasm leads to expression of viral proteins, genome replication, nucleocapsid and virion assembly, and finally budding of mature virions from the inner leaflet of the infected cell membrane (Figure 5). The latter step provides the virus with its lipid envelope. Viral mRNA (positive sense) is transcribed from genomic RNA (negative sense) and released into the cytoplasm for viral protein translation. Viral perinuclear inclusion bodies are formed by VP40, NP, VP35, VP30 and VP24; where all but VP40 are necessary for nucleocapsid formation [138-140]. VP40 associates with the nucleocapsid and accumulates in lattices in the inner leaflet of the cell membrane, and is required for virion budding with regulatory roles for many host cell factors including membrane phosphatidylserine [141, 142]. GP is transported to the cell membrane independently of other viral proteins, where it co-localises with VP40 at the latter stages of virion assembly [139].

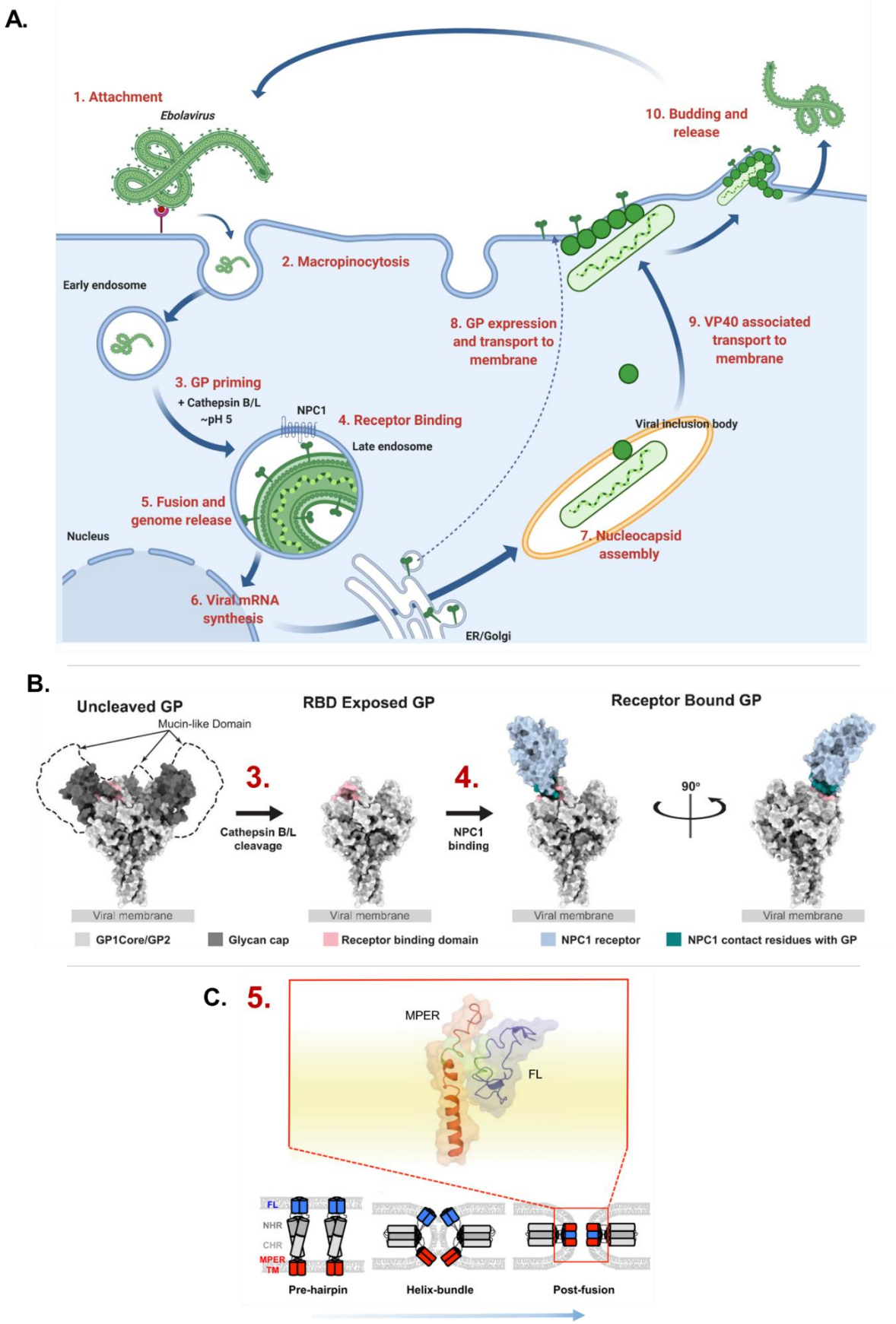


Figure 5. Ebolavirus infection and replication. A. Schematic of Ebolavirus infection and replication cycle*. Drawn using BioRender.com.* ***B. Structural detail of GP priming for fusion by cathepsin cleavage, pH change and receptor binding in the endosomal pathway.*** *Reproduced from Misasi et al. [143].* ***C. Fusion intermediates and MPER/FL interaction****. Reproduced from Lee et al. [107].*

VP40 alone can drive the formation and release of filamentous virus-like particles; however release is more efficient when VP40 is co-expressed with GP, allowing production of virus like particles (VLPs) that morphologically resemble *Ebolavirus* and can fuse with target cells [144]. Along with the ability of GP alone to induce some non-uniform, non-filamentous particles, this suggests a role for GP viral budding [144]. One mechanism may be antagonism of tetherin, a broadly antiviral host cell factor that acts by restricting viral budding. GP has been shown to be sufficient to overcome tetherin-mediated restriction of VLP budding and to directly interact with tetherin during GP maturation [145, 146].

1.3.3 Other products of the *GP* gene

The *GP* gene produces three proteins as a result of transcriptional editing of seven consecutive uridine bases in the template mRNA. Full length, membrane-anchored GP is a minor product of the *GP* gene resulting from transcriptional stuttering leading to inclusion of an eighth adenosine.

The major product of the gene, however, is a smaller dimeric soluble GP (sGP) produced from the unedited transcript [14]. The third product of the *GP* gene is another small secreted homodimer, ssGP, arising from the inclusion of a ninth adenosine to produce a 298 residue protein of unknown function, that is produced in much smaller quantities than either sGP or GP [13]. The three proteins therefore share the N terminal residues 1-295, but have distinct C terminal domains. In sGP, the 65 residue C terminal domain includes a cysteine residue that forms a disulphide bond to dimerise the protein [147]. Unlike the GP, sGP does not contain the MLD or any of GP2. The sequence shared between GP and sGP, and the high ratio of sGP:GP produced, lead to the hypothesis that sGP acts as an immune decoy; antibody raised against GP is adsorbed by soluble sGP secreted by infected cells rather than binding to virus particles [148, 149].

Like full length GP, sGP is first expressed as a precursor that undergoes furin cleavage, this produces the mature sGP and a C terminal peptide fragment, the Δ-peptide, which is also secreted

[150]. This peptide, but not mature sGP or ssGP, prevents infection of *Filovirus* permissive cells by GP coated pseudoviruses, but its functions remain unknown [151].

1.4 Non-naturally acquired antibodies in protection from ED

Antibodies that are not acquired by natural infection, i.e. that are induced by vaccination, from passive serum transfer and monoclonal antibody (mAb) therapies, can provide protection from and treatment for EBOV infection in animal models and human clinical trials. The basic rationale for such treatments is that antibodies act in multiple ways to reduce viral load which, especially early in infection, should result in better outcomes for patients, reflecting the trend in viral loads seen between survival vs fatal outcomes of ED in the absence of treatment [32, 45]. Protection mediated by different sources of non-naturally acquired antibodies against EBOV will be discussed next.

1.4.1 Vaccine-induced antibody

The rVSV/ZEBOV vaccine has been used successfully in human ED outbreaks since 2014 to protect those most at risk of infection and generated robust GP specific antibody responses [42, 152]. In NHP models, protection mediated by the vaccine against lethal EBOV infection was shown to be dependent on antibody against GP; CD4+ T cell-depleted macaques before vaccination did not survive challenge and were depleted of GP specific antibody, in contrast, macaques depleted of CD8+ T cells before vaccination survived challenge, as did macaques depleted of CD4+ cells after vaccination and the induction of GP specific antibody [153]. Additionally, both the rVSV/ZEBOV and other vaccines utilising different antigen delivery platforms assessed in early phase human clinical trials have been shown to induce neutralising antibody titres in pseudovirus and/or WT virus neutralisation assays, although the efficacy of these latter vaccines at preventing disease in humans is yet to be assessed [154-158]. Other vaccines in preclinical development can induce neutralising antibody responses in animals, with multivalent vaccines able to induce cross-neutralising immunity to multiple species of *Ebolavirus* in small animal and NHP models [159-162].

1.4.2 Passive transfer

Mupapa *et al*. report a case study during 1995 Kikwit outbreak of ED, in which seven of eight ED patients who were transfused with total blood from convalescent donors survived. Donors had detectable EBOV IgG in their blood and had experienced acute disease. At the time clinical care otherwise consisted of palliative care and treatment of renal insufficiency and cardiovascular collapse as a result of hypovolemic shock, and the overall case fatality rate of patients in the rest of the outbreak was 80%. However, the apparent efficacy of passive transfer was confounded by the higher general standard of care received by the transfer recipients than had been experienced earlier in the outbreak [163]. Nevertheless, this initial report suggested passive serum transfer was worth revisiting in a larger more controlled setting. During the 2014-2016 outbreak, van Griensven *et al*. conducted a non-randomized comparative study using convalescent plasma, with unknown neutralising antibody titre, to treat 99 ED patients; the risk of death for convalescent plasma treated patients was slightly lower, but did not reach significance [164]. NHP studies have provided a mixed picture of the efficacy of both convalescent and vaccine-generated polyclonal antibody in challenge studies [165-167].

The variability in convalescent plasma potency (especially when neutralising antibodies cannot be tested before use) and availability, as well as concerns of accidental infection or reinfection with virus have been attempted to be addressed by development of vaccine-induced hyperimmune equine and ovine serums that show definite, if caveated, efficacy in non-human primate challenge models [168, 169]. With animal-origin products there are additional concerns regarding the immunogenicity of non-human immunoglobulin leading to reduced efficacy and safety, for example by inducing serum sickness as a result of anti-drug antibody induction and immune complex formation [170]. In contrast, recombinant mAbs, whilst significantly more expensive, offer a safer, defined product, with known pharmacokinetic and efficacy profiles, although they too can still be immunogenic.

1.4.3 Monoclonal antibodies against GP

Evidence from neutralising antibody responses to natural infection (see section 1.2.3), along with the development of mAb based therapies for other infectious diseases, and the lack of other effective antivirals for *Filovirus* infections, led to a concerted effort to develop mAbs against EBOV GP during the

2014 West African outbreak in an attempt to rapidly find an effective therapy. The result of which was that towards the end of the outbreak, ZMapp, a cocktail of three mAbs against the GP, was tested in humans. However, whilst a trend towards a protective effect was seen, the threshold for showing efficacy was not met [171]. Since then, a large number of EBOV GP mAbs have been reported and reviewed in the literature [143, 172]. Contributing to this expansion of the field, a panel of 82 human *Ebolavirus* GP monoclonal antibodies were isolated from vaccinees in a collaboration between the supervisors of this thesis [3]; these mAbs are used throughout this thesis as benchmarks and controls for the evaluation of the novel rabbit mAbs generated and characterised in this study, and will be referred to as the "Oxford/UCB panel".

Two major milestones in the field have been the establishment of a large collaborative consortium (Viral Hemorrhagic Fever Immunotherapeutic Consortium (VIC)) which identified features of GP mAbs *in vitro* that correlate with *in vivo* efficacy in mouse models of EBOV infection [173]; and a second clinical trial of experimental therapies in a subsequent ED outbreak involving 681 patients, which provided evidence of efficacy leading to the recent licensure of two mAb based ED therapies [45]. The VIC analysed 171 mAbs from human and non-human origin, vaccinated and convalescent samples, published and unpublished, that were submitted by many labs across the globe. The mAbs were assessed head-to-head in a large array of assays conducted by collaborating labs to understand which mAb features, and which assays, best predicted *in vivo* efficacy in mouse challenge models. The major findings were: that both neutralisation and immune effector function can contribute to mAb-mediated protection; neutralisation *in vitro* is strongly, although not completely, correlated with *in vivo* protection in mice; mAbs that are strongly neutralising in multiple *in vitro* neutralisation assays are more likely to be protective than those that are neutralising in a single assay format; and that protection and neutralisation are not affected by cross-reactivity to sGP [173, 174].

Whilst the discovery of therapies and successful execution of a clinical study, to establish their efficacy in an extremely challenging outbreak setting are significant achievements, a key limitation of the newly licensed mAb therapies is their predicted lack of efficacy against other *Ebolavirus* species. Broadly reactive mAbs that can protect against multiple *Ebolaviruses* have become the focus of recent

efforts by researchers across the field to address the ongoing risk of infection by different known and emerging *Ebolavirus* species, and are also the focus of this thesis. Due to the number of GP mAbs now reported in the literature, the following sections will focus on reviewing mAb-based therapies that have been tested in human clinical trials (1.4.3.1), and the next generation of broadly reactive mAbs in pre-clinical development (1.4.3.2).

1.4.3.1 Monoclonal antibodies assessed in clinical trials

1.4.3.1.1 ZMapp

ZMapp is a 1:1:1 cocktail of three chimeric mAbs, c13C6, c2G4 and c4G7, originally derived from mice immunised with VSV (2G4, 4G7) or Venezuelan equine encephalitis virus (13C6) replicons coated in GP [175-177]. Neutralising mAbs c2G4 and c4G7 both bind the base of the GP and can inhibit fusion, whereas c13C6 is non-neutralising and binds to the GC [175, 178, 179]. Partial protection in animal models of fatal ED challenge conferred by c13C6 is therefore reliant on Fc effector functions, and c13C6 can recruit antibody-dependent neutrophil phagocytosis (ADNP) *in vitro* to a much greater degree than c2G4 and c4G7 [174, 177, 180]. ZMapp completely reversed symptoms of ED in a fatal EBOV (Kikwit variant) challenge NHP model when delivered in a three dose schedule of 50mg/kg, doses delivered three days apart, starting five days after infection [177]. These data led to ZMapp being delivered under compassionate use protocols and wider assessment in a randomised, controlled trial in 72 patients during the 2014 West African ED outbreak. The effect of ZMapp appeared to be beneficial (22% case fatality rate (CFR) in ZMapp treated group vs 37% in group receiving current standard of care alone), but did not meet the prespecified threshold for efficacy. This may have been due to the lower than planned enrolment due to falling ED case rate and reflect the longer delay from infection to treatment for the patients compared to the NHP study leading to the treatment window in which ZMapp is most efficacious to be missed; several deaths in the ZMapp treated group occurred before administration of the second dose [171]. Having shown protection after a significant delay to treatment in NHPs, and as the only mAb based therapy to be tested in human clinical trials, ZMapp became the gold standard comparator for all subsequent therapies.

A second, more recent, randomised, controlled trial in a large outbreak of ED in the DRC, compared ZMapp, two other mAb therapies, and antiviral Remdesivir [45]. In this trial, every patient was assigned to a treatment group, with ZMapp-treated patients serving as the control arm. Mortality in the ZMapp group was 50%, which was significantly worse than other treatments tested at the time and the CFRs reported in the previous trial. As a result, treatment with ZMapp was halted after interim trial analysis [45]. Considering that the majority of deaths in the trial (97%) occurred within the first 10 days after enrolment, the long dosing schedule of ZMapp and the fact that the cocktail had to be reconstituted on site leading to a longer treatment preparation time that often delayed the start of treatment, suggest ZMapp in its current form is less practical and efficacious than it perhaps could be in other formulations or settings.

ZMapp has been developed into a simplified humanised two component cocktail, MIL77E, containing only 13C6 and 2G4 (1:2 ratio), after studies showed that the two base binding mAbs competed for the same epitope, and that 4G7 was harder to produce in bulk [179, 181]. MIL77E has not been assessed in human clinical trials, but protected 3/3 NHPs when given three days post infection with EBOV (Makona variant) and delivered in the same regimen as ZMapp [181].

1.4.3.1.2 REGN-EB3 (Inmazeb)

REGN-EB3 is a cocktail of three EBOV GP specific mAbs (REGN3470, REGN3471, and REGN3479) derived from immunisation of VelocImmune mice with EBOV GP using DNA and recombinant protein [182]. VelocImmune mice are engineered to utilise the human immunoglobulin variable region to produce antibody [183]. Pascal *et al.*, report that the antibodies were rationally selected for inclusion in the cocktail based on *in vitro* neutralisation, *in vitro* FcγRIIIa triggering (as a proxy for their ability to induce antibody-dependent cellular cytotoxicity), and ability to simultaneously bind non-overlapping epitopes on the GP; the cocktail was then tested in guinea pig and NHP challenge models [182]. REGN3479 binds to the base of the GP, REGN3470 binds to the GC at an angle parallel to the viral membrane on the outside of the GP chalice, and REGN3471 approaches the GP perpendicular to the membrane binding top down and contacting the GP1 core. In a lentiviral pseudotype neutralisation assay, REGN-3470 and REGN-3479 were completely neutralising, and REGN-3471 partially neutralising;

however, in a wild type virus neutralisation experiment only REGN-3479 was neutralising. Individually, each antibody showed some, although incomplete, protection in a guinea pig challenge model. In NHPs, a single 150mg/kg dose of the cocktail delivered five days after infection completely reversed disease and cleared viremia in 4/5 rhesus macaques challenged with a lethal dose of EBOV (Kikwit).

After a dose escalation study to assess safety and pharmacokinetics of the cocktail in humans, REGN-EB3 was added as a later protocol amendment to the efficacy study in the DRC [184]. REGN-EB3 was delivered as a single 150mg/kg dose in a co-formulated cocktail, and reduced deaths by 17.8% (to 33.5% mortality) compared to the ZMapp treated arm of the study (when corrected for the subgroup of patients recruited to the ZMapp arm after the addition of REGN-EB3 to the trial) and decreased incidence of death for patients with both high and low viral loads at time of enrolment [45].

REGN-EB3 (brand name Inmazeb) was approved for use by the FDA to treat EBOV related ED in October 2020 [185].

1.4.3.1.3 mAb114 (Ansuvimab, Ebanga)

Derived from a survivor of the 1995 EBOV (Kikwit variant) outbreak in the DRC, mAb114 is a single RBR binding GP-specific mAb [103]. It approaches the RBR by binding inside the GP chalice at an angle perpendicular to the viral membrane and can inhibit receptor binding [186]. *In vitro*, mAb114 neutralised both EBOV GP lentivirus pseudoviruses and WT EBOV, and induced antibody-dependent cellular cytotoxicity (ADCC) in a flow cytometry-based cell killing assay [103]. When given three doses of 50mg/kg mAb114 alone on three consecutive days starting five days post infection with a lethal EBOV (Kikwit) challenge dose, 3/3 NHPs survived [103]. Safety and pharmacokinetic properties of mAb114 were determined and compassionate use approved in August 2018 during an ED outbreak in the DRC, subsequently it was assessed for efficacy alongside ZMapp and REGN-EB3 [187]. mAb114 was delivered as a single dose infusion of 50mg/kg on day one, and reduced mortality by 14.6% compared to ZMapp treated patients, regardless of viral load upon enrolment.

mAb114 (also designated Ansuvimab, and brand name Ebanga) was approved for use by the FDA to treat EBOV-related ED in December 2020 [185]

1.4.3.2 Broadly reactive antibodies

As discussed in Section 1.1.1 there is an extant threat of ED caused by SUDV and BDBV, as well as continued emergence of new *Ebolaviruses* which all utilise the same receptor to infect cells and have the potential to cause disease with high mortality in humans. Key limitations of the most advanced monoclonal antibody therapeutics discussed in Section 1.4.3.1 are their restricted range against *Ebolavirus* species and their low efficacy in cases with higher viral loads or longer delay to treatment. Broadly-reactive mAbs that bind GP from all known species of *Ebolavirus* that cause human disease have been increasingly reported in the literature. Many have demonstrated efficacy against challenge with multiple species of *Ebolavirus* in different animal models and are summarised in Table 1 and Table 2. The diversity of assays used to characterise and assess the efficacy of these mAbs makes direct comparison difficult, but trends, particularly in epitope, can be observed.

Epitopes of broadly reactive mAbs on the GP are, by necessity, conserved; these sites have not mutated as *Ebolaviruses* have diversified and may perform essential functions and be less amenable to viral escape during outbreaks or to mutation in emerging species. Focusing on finding broadly reactive mAbs therefore may not only prepare better for current and emerging threats from multiple *Ebolaviruses*, but also help find more potent therapies that will be effective for longer. Broadly reactive mAbs could also be used for designing diagnostics to distinguish *Ebolavirus* infections and identify infections caused by newly emerged viruses.

As Table 1 demonstrates, and discussed in sections1.2.3 and 1.4.1, broadly reactive antibodies are part of the response to natural infection and vaccination, but there are challenges faced when trying to generate a robust pipeline of broadly protective mAbs to treat human disease [94, 99, 100, 188-190]. In most instances, the discovery of mAbs with broad reactivity requires either high throughput B cell sampling and screening methods, or efficient sorting for cross-reactive immunoglobulin (Ig). Once a panel of broadly reactive mAbs has been generated, an even smaller proportion will demonstrate other desirable qualities such as *in vitro* neutralisation of multiple species, and not all of these will possess other qualities making them suitable for development to the clinic, such as stable expression. For example, to identify EBOV-520, Gilchuk *et al.* generated 600

lymphoblastoid cell lines, of which 44% bound EBOV GP only, 10% were cross-reactive to SUDV, BDBV and EBOV, and 20% did not bind any GP at all. From the broadly reactive IgG producing cell lines, 16 broadly reactive mAbs with different sequences were identified, only three of which neutralised EBOV, SUDV and BDBV *in vitro* [99]. To find CA45, Zhao *et al.* sorted memory B cells (mBCs) from a vaccinated macaque; 0.06% of mBCs were EBOV/SUDV cross-reactive, leading to 29 cells sorted, and recovery of 17 paired heavy and light Ig sequences, of which 12 were successfully expressed, 11 were EBOV/SUDV cross-reactive, 7 had unique sequences, and only one was ultimately able to neutralise both EBOV and SUDV GP coated VSV pseudoviruses [190]. Broadly reactive mAbs are there to be found, but their discovery requires the most recent technological advancements in antibody discovery methods.

Once characterised *in vitro*, *in vivo* efficacy can now be assessed in numerous established small animal and NHP virus challenge models. However, as seen in Table 2, efficacy in mouse or guinea pig challenge models does not always translate to efficacy in NHPs, and challenge models rarely recapitulate symptoms of human disease (especially for SUDV and BDBV infection), route of infection, or likely length of delay from infection to treatment in outbreaks of ED in humans. From a practical standpoint, the necessity of BSL-4 animal facilities means that, with the exception of the guinea pig study reported by Rijal *et al*., all the studies reported in Table 2 were conducted in one of three facilities in US or Canada (US Army Medical Research Institute of Infectious Disease, Fort Detrick; Galveston National Laboratory, UTMB; National Microbiology Laboratory, Winnipeg, Canada). Whilst these researchers are clearly extremely collaborative, the field at large is therefore limited by the capacity and resource of these institutions.

Once in the clinic, still more therapeutic candidates, like ZMapp, will likely fail to be effective in humans or practical for use in outbreaks. Testing efficacy in humans requires conducting clinical trials in complex outbreak scenarios as no safe human challenge model for *Ebolavirus* infection exists, and to rapidly initiate such trials to both benefit those at risk from the outbreak, and to achieve statistically relevant data sets before outbreaks are brought to an end is very challenging.

The generation of a robust pipeline of multiple broadly reactive candidate mAbs from a variety of sources and targeting a range of GP epitopes and functions is vital to rapid progression towards efficacy trials and better *Ebolavirus* therapies. The aims of this thesis were therefore to generate a panel of mAbs that could bind to GP from all species of *Ebolaviruses* that cause human disease, then to characterise the mAbs to select novel candidates to enter the pipeline of next-generation therapies progressing to *in vivo* study.

As shown in Table 1 and Table 2, many of the earlier broadly reactive protective mAbs reported in the literature appeared to be concentrated on a few base and fusion loop GP epitopes; these will be discussed next. However, some studies have identified broadly reactive mAbs that bind across all the major domains of GP [94, 191]. Combining non-competing mAbs targeting different (and potentially co-operative) epitopes in cocktails is an area of active research, but requires multiple candidates against a range of distinct epitopes [115, 191, 192].

1.4.3.2.1 Fusion loop antibodies

The GP FL consists of a series of close, but seemingly distinct, antigenic sites as defined by characterisation of a growing set of antibodies against this domain. The base and stem of the FL can elicit broadly neutralising mAbs, as well as mAbs with limited cross-reactivity and EBOV-specific mAbs; the more species restricted mAbs (such as K252 and c2G4) target epitopes lower on the GP that are less conserved. The FL tip appears to be a target for pan-reactive mAbs such as 6D6, FVM02, ADI-15878 and its clonal relative ADI-15742, which all share an overlapping epitope [10, 192, 193]. The footprint of EBOV-specific mAb100 also overlaps with 6D6, however it binds lower on the GP, does not bury the fusion peptide, and has a different angle of attack, consistent with the hypothesis that high position and recognition of the hydrophobic portion of the FL are important for broad neutralisation [10, 186].

1.4.3.2.2 3_{10} pocket antibodies

Two broadly-reactive, base-binding mAbs that also contact the FL base have been defined more specifically as 3_{10} pocket antibodies. EBOV-520 and ADI-15946 both contact residues 71-75 of the GP1

core, designated the 3_{10} pocket, which is usually occupied by the β17-β18 loop of the GC prior to cathepsin cleavage [115, 191]. Binding to this portion of the GP may allow mAbs to disrupt multiple conformational changes of the GP required for infection, making this a particularly vulnerable site on the GP.

1.4.3.2.3 Other epitopes

Broadly reactive mAbs that target other portions of the GP have been reported, but were less common at the inception of this project. For example, EBOV-548 was isolated as a non-competing partner for 3_{10} pocket antibody EBOV-520, and binds in the glycan cap. Whilst it is broadly reactive, and can neutralise WT EBOV, SUDV and BDBV to varying degrees *in vitro*, it was only partially protective against EBOV challenge in mice and failed to protect against SUDV challenge [191]. Other broadly reactive mAbs to the GC were generated from a survivor of BDBV infection, of which BDBV289 has been found to be protective in NHPs challenged with BDBV when treatment was delayed as late as eight days post challenge [94, 194]. MR78 cannot neutralise *Ebolaviruses*, but has a pan-filovirus epitope situated in the conserved RBR of the GP, which in *Ebolaviruses* is occluded by the MLD [4, 195]. Three broadly reactive mAbs that bind to the HR2 and MPER domains (BDBV223, BDBV317, and BDBV340) were found to be unable to neutralise SUDV *in vitro*, but provided some protection *in vivo* against EBOV and BDBV. In the same study, the authors showed that immunising rabbits with the HR2-MPER peptides elicited a cross-reactive and polyclonal response, which in some animals was cross-neutralising *in vitro* [94, 100].

1.4.3.2.4 Other sources of broadly neutralising antibodies

As can be seen from Table 1, the published literature on broadly reactive GP mAbs focusses heavily on human, macaque and murine antibodies; the human and macaque samples derived from vaccinees, survivors of infection, and tissues from animals used in virus challenge studies. However, in the search for therapeutics there are many other species that could be used to find novel anti-viral and broadly reactive antibodies. The range of immunoglobulin repertoires and structures available from other species has already been exploited to develop broadly neutralising antibodies against other viruses.

Most famously, bovine antibodies with ultra-long complementarity determining region 3 (CDR3) structures that can access a recessed epitope on HIV-1 envelope protein.

Research into broadly neutralising antibodies from HIV-1 infected individuals found mAbs with exceptionally long heavy chain CDR3 sequences of >35 amino acids (compared to an average of ~15 residues) which were shown to allow the antibody to contact a hidden site of vulnerability on the HIV-1 Envelope (Env) protein [196, 197]. Humans however, inefficiently generate long CDR3 sequences; these mAbs develop in only a small number of HIV-1 infected individuals [197]. In contrast, bovines produce a significant amount of Ig with ultra-long heavy chain CDR3 sequences of 40-70+ amino acids [198]. Cows immunised with BG505 SOSIP, a soluble cleaved form of the Env trimer, generated serum responses that were cross-neutralising despite only being presented with a single immunogen. One mAb with an ultra-long CDR3 that was isolated from that study was able to strongly neutralise 72% of 117 virus isolates tested, by contacting the conserved CD4 binding site with the tip of that single ultra-long CDR [199].

Camels and sharks produce antibodies with significantly different structures that make them amenable to recombinant display technologies for antibody discovery. Camelid nanobodies have also been used to target the CD4 binding site on HIV-1 Env [200].

Rabbits are commonly vaccinated for antibody discovery projects due both to their ease of housing and care, the availability of reagents to support rabbit antibody cloning and production, and their alternative immune repertoire diversification methods compared to humans and mice. Yet they are relatively unexploited in the search for novel anti-GP mAbs.

These alternative approaches to antibody discovery are discussed further in Section 3.1.

1.4.3.2.5 Increasing function of non-neutralising antibodies

As discussed in Section 1.4.3, virus neutralisation correlates with, but is not a pre-requisite for protection by anti-GP mAbs, and Fc effector functions can play a role in protection. This means that strategies that increase the neutralisation potency of poorly or non-neutralising mAbs, or increase the

engagement of Fc-mediated effector functions of mAbs may both improve the therapeutic potential of candidate antibodies.

Fc engineering strategies provide a wide range of ways to alter the ability of mAbs to recruit Fc effector functions. For example, IgG form hexamers on the surface of infected cells upon antigen binding; these hexamers bind to the C1 component of the classical complement cascade that ultimately leads to targeted cell killing via formation of membrane attack complexes, and opsonisation of the target cell to attract immune effector cells [201]. The propensity of IgG to form on-target hexamers and hence improve complement activation, can be increased by addition of a single amino acid changes in the Fc that promote clustering [202, 203]. Alternatively, the introduction of IgM-derived tailpieces allows for covalent hexamerisation of IgG which not only increases complement activation, but also the avidity of the antibody-antigen interaction [204]. Other strategies, such as glycan engineering can also alter the interaction of the Fc with various Fc receptors [205]. Glycan engineering of therapeutic candidate *Ebolavirus* mAbs is discussed in more detail in Section 6.4. Any technique which engineers the Fc is still ultimately reliant on the angle and footprint of binding by that antibody's variable region; for instance, if the epitope is too close the viral membrane it may preclude binding or formation of a large hexameric molecule.

The properties of an antibody variable region can be refined by *ex vivo* affinity maturation; successive rounds of mutation and selection can increase affinity or breadth of specificity, both of which may affect ability to neutralise virus. For example, mAb ADI-23774 was derived by affinity maturation of broadly reactive mAb ADI-15946; by reducing reliance on contacts within its footprint on the GP that were not conserved on SUDV, ADI-23374 has improved neutralisation of SUDV whilst maintaining neutralisation of the other *Ebolaviruses* [2].

Pairing co-operative or synergistic pairs of antibodies together can also improve neutralisation and provide a key role for non-neutralising antibodies in cocktails; mAbs that are non-neutralising in isolation, can potentiate a partner mAb by altering the kinetics of the interaction between the partner

mAb and the antigen. Examples within published broadly reactive GP mAbs include EBOV-520 and EBOV-548 as well as ADI-15946 and FVM09 [191, 193] and is discussed in more detail in Section 6.6.1.

A related, yet alternative approach, taken by Wec et al, pairs variable regions from non-neutralising antibodies into a bispecific molecule rather than a cocktail. This approach utilises the ability of the variable region of one that can recognise uncleaved pre-fusion GP to target the bispecific molecule to the GP before the virus is taken into then endosome. Once in the endosome, cleavage of the GP reveals the conserved RBR, the target of the second variable region in the bispecific molecule. Each variable region alone is unable to neutralise virus bearing uncleaved GP, but together, they can target the occluded RBR in the context of the virus infecting a cell [206].

Although outside the scope of this thesis, any antibodies identified targeting highly conserved epitopes, which are poorly neutralising or poorly protective, may still be leveraged by engineering strategies such as those discussed.

1.5 Thesis Outline

As discussed, *Ebolaviruses* have complex ecology and transmission cycles, and multiple species have caused and continue to cause disease with high mortality in humans. The GP on the virus surface represents a target for vaccines and therapies due to its many indispensable roles in the *Ebolavirus* infection cycle. Almost a decade of concerted research into mAb therapies that target GP has culminated in the recent approval of two therapies (the first for treatment of ED) that show some efficacy in reducing deaths in outbreak scenarios. The next generation of mAb therapies for ED in pre-clinical development aim for broader and more potent activity against multiple *Ebolaviruses*.

The work described in this thesis details the discovery and initial characterisation of novel broadly-reactive monoclonal antibodies (mAbs) that bind to the GP of all *Ebolavirus* species known to cause human disease. High throughput screening strategies were used to generate a panel of mAbs from immunised rabbits, and recombinantly expressed candidate mAbs were characterised in a range of *in vitro* assays to identify lead molecules for future development and testing *in vivo*.

Chapter 3 describes the generation of the mAb panel: the immunisation strategy used to generate broadly reactive immune responses in rabbits; followed by B cell culture; design and implementation of screening strategies used to identify broadly reactive IgG in B cell culture supernatants; recovery and analysis of V region sequences; and recombinant expression of monoclonal antibody. Chapter 4 describes the characterisation of the panel of antibodies in biochemical, cell-based, and pseudotype neutralisation assays to identify neutralising antibodies and characterise their mechanism of neutralisation. Chapter 5 outlines multiple methods used to define the epitopes of antibodies in the panel, with a particular focus on antibody 11886 which was identified in Chapter 4 as a potentially broadly-neutralising antibody. Epitope information is especially important in the rational design of antibody cocktails as these antibodies move into *in vivo* assessment.

mAb/ Cocktail	Origin					Binding								
	Source	Format	Original species	Immunisation	Ref	EBOV	SUDV	BDBV	TAFV	RESTV	BOMV	MARV	Assay	Ref
6D6	Mouse hybridoma	IgG	Mouse	E + S GP VLPs	Furuyama et al, 2016	Y	Y	Y	Y	Y	n.d.	N	ELISA	Furuyama et al, 2016
EBOV-520	DRC 2014 ED survivor, PBMC -> hybridomas	IgG4	Human	n/a	Gilchuk et al, 2018	Y Y	Y n.d.	Y n.d.	n.d. n.d.	n.d. n.d.	n.d. n.d.	N n.d.	ELISA Cells	Gilchuk et al, 2018
rEBOV-520 LALA	DRC 2014 ED survivor, PBMC -> hybridomas	IgG1 + LALA	Human	n/a	Gilchuk et al, 2018	Y	Y	Y	n.d.	n.d.	n.d.	n.d.	Cells	Gilchuk et al, 2020
EBOV-515	DRC 2014 ED survivor, PBMC -> hybridomas	IgG1	Human	n/a	Gilchuk et al, 2018	Y Y	Y n.d.	Y n.d.	n.d. n.d.	n.d. n.d.	n.d. n.d.	N n.d.	ELISA Cells	Gilchuk et al, 2018
EBOV-442	West African 2013-2016 EVD survivor, PBMC -> hybridomas	IgG1	Human	n/a	Gilchuk et al, 2018	Y Y	Y n.d.	Y n.d.	n.d. n.d.	n.d. n.d.	n.d. n.d.	N n.d.	ELISA Cells	Gilchuk et al, 2018
EBOV-437	West African 2013-2016 ED survivor, PBMC -> hybridomas	IgG1	Human	n/a	Gilchuk et al, 2018	Y Y	Y n.d.	Y n.d.	n.d. n.d.	n.d. n.d.	n.d. n.d.	N n.d.	ELISA Cells	Gilchuk et al, 2018
EBOV-548	ED survivor, PBMC -> hybridomas	IgG1	Human	n/a	Gilchuk et al, 2020	Y	Y	Y	n.d.	n.d.	n.d.	n.d.	Cells	Gilchuk et al, 2020
rEBOV-458 + rEBOV-520	ED survivor, PBMC -> hybridomas	IgG1	Human	n/a	Gilchuk et al, 2020	See data for individual mAbs								
CA45	Immunised macaque, Single cell sort -> RT-PCR	IgG1	Cynomolgous macaque	E, S , M GPΔmuc + adjuvant	Zhao et al, 2017	Y	Y	Y	N	Y (poor)	n.d.	N	ELISA	Zhao et al, 2017
FVM04	Immunised macaque, Single cell sort -> RT-PCR	IgG1	Cynomolgous macaque	E, S, M GPΔmuc + adjuvant E, S, M GP VLPs + adjuvant	Keck et al, 2015	Y	Y	Y	n.d.	Y	n.d.	Partial	ELISA	Keck et al, 2016
FVM02	Immunised macaque, Single cell sort -> RT-PCR	IgG1	Cynomolgous macaque		Keck et al, 2015	Y	Y	Y	n.d.	Y	n.d.	Partial	ELISA	Keck et al, 2016
ADI-15946	West African 2013-2016 ED survivor, PBMC -> Single cell sort -> RT-PCR	IgG1	Human	n/a	Bornholdt et al, 2016	Y	Poor	Y	n.d.	n.d.	n.d.	n.d.	ELISA	Wec et al, 2017
ADI-23774	Affinity maturation of ADI-15946	IgG1	Human	n/a	Wec et al, 2019	Y	Y	Y	n.d.	n.d.	n.d.	n.d.	BLI	Wec et al, 2019
ADI-15878	West African 2013-2016 ED survivor, PBMC -> Single cell sort -> RT-PCR	IgG1	Human	n/a	Bornholdt et al, 2016	Y	Y	Y	n.d.	n.d.	n.d.	n.d.	ELISA	Wec et al, 2017
ADI-15742	West African 2013-2016 ED survivor, PBMC -> Single cell sort -> RT-PCR	IgG1	Human	n/a	Bornholdt et al, 2016	Y	Y	Y	n.d.	n.d.	n.d.	n.d.	ELISA	Wec et al, 2017
BDBV223	2007 BDBV survivor PBMC	IgG1	Human	n/a	Flyak et al, 2016	Y	Y	Y	n.d.	n.d.	n.d.	n.d.	ELISA	Flyak et al 2016
BDBV317						Y	Y	Y	n.d.	n.d.	n.d.	n.d.	ELISA	Flyak et al 2016
BDBV340						Y	Y	Y	n.d.	n.d.	n.d.	n.d.	ELISA	Flyak et al 2016
BDBV289						Y	Y	Y	n.d.	n.d.	n.d.	n.d.	ELISA	Flyak et al 2016
m4B8	Mouse hybridoma	?	Mouse	25µg E, S, M GPΔmuc + adjuvant.	Holtsberg et al, 2016	Y	Y	Y	n.d.	Y	n.d.	N	ELISA	Holtsberg et al, 2016
2G1	Human vaccinee	IgG1?	Human	Ad5-EBOV	Fan et al, 2020	Y	Y	Y	n.d.	Y	n.d.	N	ELISA	Fan et al, 2020
040	Human vaccinee	IgG1	Human	ChAd3 EBOV + MVA-BN Filo	Rijal et al, 2019	Y	Y	Y	n.d.	n.d.	n.d.	n.d.	IFA	Rijal et al, 2019
66-3-9C						Y	Y	Y	n.d.	n.d.	n.d.	n.d.	IFA	Rijal et al, 2019
6662						Y	Y	Y	n.d.	n.d.	n.d.	n.d.	IFA	Rijal et al, 2019
6541						Y	Y	Y	n.d.	n.d.	n.d.	n.d.	IFA	Rijal et al, 2019

Table 1. Table of broadly reactive Ebolavirus GP monoclonal antibodies that have been assessed in published in vivo animal challenge models. *Summarising origin and binding. Table continued on next two pages. E; EBOV. S; SUDV. M; MARV PBMC; Peripheral blood mono nuclear cells. Y; Positive. N; Negative. n.d.; no data found. IFA; immunofluorescence assay. [2, 3, 10, 11, 94, 99-101, 115, 190, 191, 193, 194, 207-213] Continued on next page.*

mAb/ Cocktail	*In vitro* pseudotype neutralisation										*In vitro* WT neutralisation IC50 (μg/ml)									*In vitro* Fc Effector functions					
	EBOV	SUDV	BDBV	TAFV	RESTV	BOMV	MARV	LLOV	Assay	Ref	EBOV	SUDV	BDBV	TAFV	RESTV	BOMV	MARV	Assay	Ref	ADCP	ADNP	NK	ADCD	RFADCC	Ref
6D6	Y	Y	Y	Y	Y	n.d.	N	n.d.	VSV	Furuyama et al, 2016	Y	Y	Y	Y	Y	n.d.	N	WT	Furuyama et al, 2016	n.d.					
EBOV-520	n.d.										5.738	6.318	3.81	n.d.	n.d.	n.d.	n.d.	eGFP	Gilchuk et al, 2018	Y	N	Poor	N	Y	Gilchuk et al, 2018
rEBOV-520 LALA	Y	n.d.							VSV	Gilchuk et al, 2020	1.801	1.672	0.2	n.d.	n.d.	n.d.	n.d.	eGFP	Gilchuk et al, 2020	N	N	N	N	N	Gilchuk et al, 2018
EBOV-515	n.d.										1.224	0.891	1.458	n.d.	n.d.	n.d.	n.d.	eGFP	Gilchuk et al, 2018	Y	N	Y	N	n.d.	Gilchuk et al, 2018
EBOV-442	n.d.										0.467	P	1.489	n.d.	n.d.	n.d.	n.d.	eGFP	Gilchuk et al, 2018	Y	Y	Y	Y	n.d.	Gilchuk et al, 2018
EBOV-437	n.d.										P 62%	P, 21%	n.d.	n.d.	n.d.	n.d.	n.d.	eGFP	Gilchuk et al, 2020	Y	Y	Y	Poor	n.d.	Gilchuk et al, 2018
EBOV-548	Y	n.d.							VSV	Gilchuk et al, 2020	1.601	P	2.262	n.d.	n.d.	n.d.	n.d.	eGFP	Gilchuk et al, 2020	Y	Y	Y	n.d.	Y	Gilchuk et al, 2020
rEBOV-458 + rEBOV-520	Y	n.d.							VSV	Gilchuk et al, 2020										See data for individual mAbs					
CA45	Y	Y	Y	N	Y	n.d.	n.d.	N	VSV	Zhao et al, 2017										n.d.					
FVM04	Y	Y	P	N	n.d.	n.d.	n.d.	n.d.	VSV	Howell et al, 2016	Y	Y	P	n.d.	n.d.	n.d.	N	WT	Howell et al, 2016	n.d.					
FVM02	N	N	N	n.d.	n.d.	n.d.	n.d.	n.d.	VSV	Keck et al, 2016	N	N	n.d.	n.d.	n.d.	n.d.	n.d.	WT	Keck et al, 2016	n.d.					
ADI-15946	Y	P	Y	Y	N	n.d.	N	N	VSV	Wec et al, 2017	Y	N	Y	n.d.	n.d.	n.d.	n.d.	WT	Wec et al, 2017	n.d.					
ADI-23774	Y	Y	Y	Y	Y	n.d.	n.d.	n.d.	VSV	Wec et al, 2019	Y	Y	Y	n.d.	n.d.	n.d.	n.d.	WT	Wec et al, 2019	Poor	n.d.	Poor	n.d.	n.d.	Bornholdt et al, 2019
ADI-15878	Y	Y	Y	Y	Y	n.d.	N	N	VSV	Wec et al, 2017	Y	Y	Y	n.d.	n.d.	n.d.	n.d.	WT	Wec et al, 2017	Poor	n.d.	Poor	n.d.	n.d.	Bornholdt et al, 2019
ADI-15742	Y	Y	Y	Y	Y	n.d.	N	N	VSV	Wec et al, 2017	Y	Y	Y	n.d.	n.d.	n.d.	n.d.	WT	Wec et al, 2017	n.d.					
BDBV223	n.d.										Y	N	Y	n.d.	N	n.d.	n.d.	eGFP	Flyak et al 2018	n.d.					
BDBV317											Y	N	Y	n.d.	P	n.d.	n.d.	eGFP	Flyak et al 2018	n.d.					
BDBV340											P	N	Y	n.d.	N	n.d.	n.d.	eGFP	Flyak et al 2018	n.d.					
BDBV289											Y	N	Y	n.d.	n.d.	n.d.	Y	eGFP	Flyak et al 2016	n.d.					
m4B8	N	N	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	VSV	Holtsberg et al, 2016	N	N	n.d.	n.d.	n.d.	n.d.	n.d.	WT	Holtsberg et al, 2016	N	N	n.d.	n.d.	n.d.	Holtsberg et al, 2016
2G1	Y	Y	Y	n.d.	n.d.	n.d.	n.d.	n.d.	HIV-Luc	Fan et al, 2020	Y	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	eGFP	Fan et al, 2016	n.d.					
040	Y	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	S-FLU	Rijal et al, 2019										n.d.					
66-3-9C	N	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	S-FLU	Rijal et al, 2019										n.d.					
6662	Y	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	S-FLU	Rijal et al, 2019										n.d.					
6541	Y	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	S-FLU	Rijal et al, 2019										n.d.					

Table 1 contd. Table of broadly reactive Ebolavirus GP monoclonal antibodies that have been assessed in published in vivo animal challenge models. *Summarising in vitro neutralisation and Fc effector function. LLOV; Lloviu cuevavirus. eGFP; EBOV-eGFP and GP chimeras. ADCP; Antibody-dependent cellular phagocytosis. ADNP; Antibody-dependent neutrophil phagocytosis. ADCD; Antibody-dependent complement deposition. RFADCC; Rapid fluorometric. Antibody-dependent cellular cytotoxicity. Y; Positive. N; Negative. n.d.; no data found. P; partial.* *Continued on next page.*

mAb /Cocktail	Epitope							
	Bin	Specific/ critical residues	Mapping techniques	Ref	EMD code	PDB code	Description	Ref
6D6	FL	Tip of IFL: 528, 530,	Escape mutants, replication competent VSV pseudotype	Furuyama et al, 2016	EMD-9048 EMD-9049	6DG2	FP, bridges protomers	Milligan et al, 2018
EBOV-520	Base	IFL: N512, GP1 : E106 **E106K**	Competition, THL cleaved GP, Alanine scanning, escape mutants	Gilchuk et al, 2018	EMD-7955	n.d.	Higher up GP1 above CA45 site. Does not overlap ADI-15878. Quaternary, spans GP1 and GP2.	Gilchuk et al, 2018
rEBOV-520 LALA	Base	**F97V**, **N514Y** GP1 70–78, 106, 107, 127, 139 GP2 510–514	Escape mutants, crystal structure	Gilchuk et al, 2020	See EBOV-520	6OZ9	IFL (GP2) and 3_{10} pocket (GP1) Distinct from ADI-15878 but similar to ADI-15946	Gilchuk et al, 2020
EBOV-515	Base	FL P513, **N514**	Competition, THL cleaved GP, Alanine scanning, escape mutants	Gilchuk et al, 2018	EMD-7956	n.d.	Similar to CA45 but with different angle of approach. Does not overlap ADI-15878	Gilchuk et al, 2018
EBOV-442	GC		Competition, THL cleaved GP	Gilchuk et al, 2018	n.d.	n.d.		
EBOV-437	GC		Competition, THL cleaved GP	Gilchuk et al, 2018	n.d.	n.d.		
EBOV-548	GC	**G271E**, **S583A, L273P, A361T** T240, Y261, R266, T269, T270, I274, W275	THL cleaved GP, Alanine scanning, escape mutants	Gilchuk et al, 2020	EMD-20947	6UYE	Angle ~perpendicular to viral membrane/GP. Similar to 13C6 but with additional contacts in interior of chalice in GC domain.	Gilchuk et al, 2020
rEBOV-458 + rEBOV-520	See data for individual mAbs				EMD-20293 EMD-20301	6PCI	See data for individual mAbs	
CA45	FL	Base of FL: Y517, G546, N550 GP1: R64	Competition, THL cleaved GP, Alanine scanning	Zhao et al, 2017	EMD-8694	n.d.	Angle ~parallel to viral membrane. FL base, does not compete for binding with FVM02. Does compete KZ52. β3, β19, β20 Escape mutant variants do not contain mutations in epitope (**A101V, K588R, N643D** or **A654T** in the GP stalk)	Zhao et al, 2017
FVM04	RBR	K115, D117, G118	Alanine scanning	Howell et al, 2016		n.d.	Directly competes for residues on RBR crest that are involved in NPC1 binding.	Howell et al, 2016
FVM02	FL tip	526-535	Peptide mapping	Keck et al, 2016	n.d.	n.d.		
ADI-15946	3_{10} pocket + FL	**K510E**	Escape mutants, crystal structure	Wec et al, 2017 West et al, 2019	EMD-8701	6MAM	Base of fusion loop. Displaces β17-β18 loop and binds into GP1 3_{10} pocket.	West et al, 2019
ADI-23774	See ADI-15946							
ADI-15878	FL + HR1	**G528E**	Escape mutants, crystal structure	Wec et al, 2017 West et al, 2018	EMD-8700	6AE7 6EA5	Quaternary epitope. FL of one protomer, and cryptic pocket underneath GP2 N terminus (displaced) of neighbouring protomer. Contacts conserved glycan at N563, improved activity when glycan removed.	West et al, 2018
ADI-15742	FL	**G528E**	Escape mutants, EM structure	Wec et al, 2017	EMD-8699		Clonally related to ADI-15878.	
BDBV223	HR2 + MPER	D624, D632	EM structure, alanine scanning, binding to linear peptides, escape mutants	Flyak et al, 2018	n.d.	n.d.		
BDBV317	HR2 + MPER	K633		Flyak et al, 2018	n.d.	n.d.		
BDBV340	HR2 +MPER	D624		Flyak et al, 2018	n.d.	n.d.		
BDBV289	GC		EM structure	Flyak et al, 2016	EMD-6530	n.d.		
m4B8	GP1 core or base		Binds GP_{CL}	Holtsberg et al, 2016				
2G1	Base		Binding to truncated GPs, improved binding to GP_{CL}	Fan et al, 2020				
040	GC		Competition assay	Rijal et al, 2019				
66-3-9C	GC β17-β18 loop		Yeast peptide display	Rijal et al, 2019				
6662	RBR		Competition assay	Rijal et al, 2019				
6541	Base		Competition assay	Rijal et al, 2019				

Table 1 contd. Table of broadly reactive Ebolavirus GP monoclonal antibodies that have been assessed in published in vivo animal challenge models. *Summarising information on epitopes. Bold; escape mutation.*

mAb/ Cocktail	Virus model					Animal model	Treatment			Survival		Ref
	Species		GenBank ID/Strain	Infectious dose (PFU)	Route		Timing (dpi)	Route	Dose	Treated	Controls	
6D6	EBOV	MA	Mayinga	1000	i.p.	Balb/c mice	1	i.p.	100µg	8/8	0/8	Furuyama et al, 2016
	EBOV	WT	Yambuku 1976	1000	i.p.	IFNAR−/− mice	1	i.p.	100µg	6/6	0/6	Furuyama et al, 2016
	SUDV	WT	Nzara	1000	i.p.	IFNAR−/− mice	1	i.p.	100µg	5/6	5/6	Furuyama et al, 2016
EBOV-520 IgG4	EBOV	MA	AF49101	1000	i.p.	C57BL/6 mice	1	i.p.	100µg	5/5	0/5	Gilchuk et al, 2018
	SUDV	WT	Gulu	1000	i.p.	STAT1 KO mice	1	i.p.	200µg	3/5	0/5	Gilchuk et al, 2018
	SUDV	GA	KT878488 Boniface	1000	i.p.	Guinea pig	1 and 3	i.p.	~15mg/kg	4/5	1/5	Gilchuk et al, 2018
	BDBV	WT	200706291 Uganda	1000	i.m.	Ferret,	3 and 6	i.p.	18mg	1/4	0/4	Gilchuk et al, 2018
rEBOV-520 Ig1	EBOV	MA	AF49101	1000	i.p.	C57BL/6 mice	1	i.p.	5mg/kg	5/5	0/5	Gilchuk et al, 2018
	EBOV	MA	AF49101	1000	i.p.	Balb/c mice	1	i.p.	20µg (low dose)	0/5	0/5	Kuzmina et al, 2018
rEBOV-520 Ig1 LALA	EBOV	MA	AF49101	1000	i.p.	C57BL/6 mice	1	i.p.	5mg/kg	5/5	0/5	Gilchuk et al, 2018
	EBOV	MA	AF49101	1000	i.p.	Balb/c mice	1	i.p.	20µg (low dose)	3/5	0/5	Kuzmina et al, 2018
	SUDV	WT	Gulu	1000	i.p.	STAT1 KO mice	1	i.p.	10mg/kg	1/10	0/15	Gilchuk et al, 2020
EBOV-515	EBOV	MA	AF49101	1000	i.p.	C57BL/6 mice	1	i.p.	5mg/kg	5/5	0/5	Gilchuk et al, 2018
	SUDV	WT	Gulu	1000	i.p.	STAT1 KO mice	1	i.p.	10mg/kg	4/5	5/5	Gilchuk et al, 2018
EBOV-442	EBOV	MA	AF49101	1000	i.p.	C57BL/6 mice	1	i.p.	5mg/kg	4/5	0/5	Gilchuk et al, 2018
rEBOV-548	EBOV	MA	AF49101	1000	i.p.	Balb/c mice	1	i.p.	5mg/kg	2/5	0/5	Gilchuk et al, 2020
	SUDV	WT	Gulu	1000	i.p.	STAT1 KO mice	1	i.p.	10mg/kg	0/5	0/15	Gilchuk et al, 2020
rEBOV-520 Ig1 LALA + rEBOV-548	SUDV	WT	Gulu	1000	i.p.	STAT1 KO mice	1	i.p.	20mg/kg (1:1)	5/10	0/15	Gilchuk et al, 2020
	EBOV	WT	199510621 Kikwit	1025	i.m.	Rhesus macaque	3 and 6	i.v.	30mg/kg (1:1)	5/5	0/1 (0/10)	Gilchuk et al, 2020
CA45	EBOV	MA	AF499101, Mayinga	100	i.p.	Balb/c mice	2	i.p.	10mg/kg	8/10	0/30	Zhao et al, 2017
									5mg/kg	8/10		Zhao et al, 2017
									2.5mg/kg	19/20		Zhao et al, 2017
									1mg/kg	12/20		Zhao et al, 2017
									0.63mg/kg	4/10		Zhao et al, 2017
									0.5mg/kg	3/10		Zhao et al, 2017
	SUDV	WT	FJ968794.1, Boniface	1000	i.p.	IFNαR-/- mice	1	i.p.	10mg/kg	5/6	4/13	Zhao et al, 2017
	EBOV	GA	AF272001.1	1000 LD_{50}	i.p.	Guinea pig	3	i.p.	5mg	3/6	0/4	Zhao et al, 2017
	SUDV	GA	KT878488.1	1000 LD_{50}	i.p.	Guinea pig	3	i.p.	5mg	6/6	0/4	Zhao et al, 2017
CA45 + FVM04	SUDV	WT	FJ968794.1, Boniface	1000	i.p.	IFNαR-/- mice	1	i.p.	10mg/kg (1:1)	7/7	4/13	Zhao et al, 2017
	SUDV	WT	FJ968794.1, Boniface	1000	i.p.	IFNαR-/- mice	1	i.p.	7.5mg/kg (1:2, CA45:FVM04)	10/10	4/10	Zhao et al, 2017
	EBOV	GA	AF272001.1	1000 LD_{50}	i.p.	Guinea pig	3	i.p.	10mg (1:1)	6/6	0/4	Zhao et al, 2017
	EBOV	GA	AF272001.1	1000 LD_{50}	i.p.	Guinea pig	3	i.p.	5mg (1:1)	6/6	0/4	Zhao et al, 2017
	BDBV	WT	KR063673.1, Uganda 2007	253 $TCID_{50}$	i.m.	Ferret	3 and 6	i.p.	40mg (1:1)	4/4	0/2	Zhao et al, 2017
	EBOV	GA	Mayinga	1000 LD_{50}	i.p.	Guinea pig	3	i.p.	5mg (1:1)	6/6	0/4	Brannan et al, 2019
									2.5mg (1:1)	6/6		
									1.25mg (1:1)	3/6		
	SUDV	GA	Boniface	1000PFU	i.p.	Guinea pig	3	i.p.	5mg (1:1)	6/6	1/6	Brannan et al, 2019
							4		5mg (1:1)	6/6	0/6	Brannan et al, 2019
									2.5mg (1:1)	5/6		
	EBOV	WT	Makona	1000 $TCID_{50}$	i.m.	Rhesus macaque	4	i.v.	20mg/kg (1:1)	5/5	0/2	Brannan et al, 2019
	SUDV	WT	Boniface	2625 PFU	i.m.	Rhesus macaque	4 and 6	i.v.	20mg/kg (1:1), 8mg/kg FVM04 + 5mg/kg CA45	3/3	1/2	Brannan et al, 2019
FVM04	SUDV	WT	FJ968794.1, Boniface	1000	i.p.	IFNαR-/- mice	1	i.p.	10mg/kg	11/14	4/13	Zhao et al, 2017
	EBOV	GA	AF272001.1	1000 LD_{50}	i.p.	Guinea pig	3	i.p.	5mg	1/6	0/4	Zhao et al, 2017
	EBOV	MA	Mayinga	100	i.p.	Balb/c mice	1	i.p.	200µg	10/10	1/10	Howell et al, 2016
							2			8/10		
							3			3/10		
	EBOV	MA	Mayinga	100	i.p.	Balb/c mice	2	i.p.	200µg	10/10	0/10	Howell et al, 2016
									100µg	10/10		
									50µg	7/10		
	SUDV	WT		1000	i.p.	IFNαR-/- mice	1	i.p.	200µg	5/7	1/6	Howell et al, 2016
	EBOV	GA	Mayinga	1000 LD_{50}	i.p.	Guinea pig	1	i.p.	5mg	2/6	0/6	Howell et al, 2016
	SUDV	GA	Boniface	1000 LD_{50}	i.p.	Guinea pig	1	i.p.	5mg	6/6	0/6	Howell et al, 2016
FVM02	EBOV	MA	Mayinga	1000	i.p.	Balb/c mice	0 and 3	i.p.	25mg/kg	7/15	0/15	Keck et al, 2016
m4B8	EBOV	MA	Mayinga	1000	i.p.	Balb/c mice	0 and 3	i.p.	25mg/kg	5/5	0/15	Holtsberg et al, 2016
	EBOV	MA	Mayinga	1000	i.p.	Balb/c mice	3	i.p.	30mg/kg	4/5	0/10	Holtsberg et al, 2016

Table 2. Table summarising published in vivo animal challenge model data for broadly reactive mAbs. *Coloured by animal species. Yellow; mice. Orange; guinea pigs. Red; ferrets. Blue; NHPs. Target virus dose reported unless actual dose reported. For mice, 5 mg/kg equivalent to ~100 µg. MA; Mouse adapted virus. GA; Guinea pig adapted virus. WT; Wild type virus. PFU; Plaque forming units. LD50; Median lethal dose. TCID50; Median tissue culture infectious dose. i.p; intraperitoneal. i.m.; intramuscular. i.n.; intranasal. i.v.; intravascular. s.c.; subcutaneous.[2, 3, 94, 99, 100, 190-192, 194, 207-209, 211, 212, 214, 215]*

mAb /Cocktail	Virus model					Animal model	Treatment			Survival		Ref
	Species		GenBank ID/Strain	Infectious dose (PFU)	Route		Timing (dpi)	Route	Dose	Treated	Controls	
BDBV223	EBOV	MA	Mayinga	1000	i.p.	Balb/c mice	1	i.p.	100µg	5/5	0/5	Flyak et al, 2016
	EBOV	MA	Mayinga	1000	i.p.	Balb/c mice	3	i.p.	100µg	0/5	0/5	Flyak et al, 2016
	EBOV	GA	Mayinga	1000	i.p.	Guinea pig	1	i.p.	5mg	1/5	0/5	Flyak et al, 2016
	EBOV	GA	Mayinga	1000	i.p.	Guinea pig	1 and 3	i.p.	5mg	1/5	0/5	Flyak et al, 2016
	EBOV	GA	Mayinga	1000	i.p.	Guinea pig	1	i.p.	5mg	1/5	0/5	Flyak et al, 2018
	BDBV	WT	Uganda 2007	1000	i.m.	Ferret	3 and 6	?	20mg	2/4	0/4	Flyak et al, 2018
BDBV317	EBOV	MA	Mayinga	1000	i.p.	Balb/c mice	1	i.p.	100µg	5/5	0/5	Flyak et al, 2018
	EBOV	GA	Mayinga	1000	i.p.	Guinea pig	1	i.p.	5mg	3/5	0/5	Flyak et al, 2018
BDBV340	EBOV	MA	Mayinga	1000	i.p.	Balb/c mice	1	i.p.	100µg	0/5	0/5	Flyak et al, 2018
BDBV289	EBOV	MA	Mayinga	1000	i.p.	Balb/c mice	1	i.p.	100µg	4/5	0/5	Flyak et al, 2016
	EBOV	MA	Mayinga	1000	i.p.	Balb/c mice	3	i.p.	100µg	0/5	0/5	Flyak et al, 2016
	EBOV	GA	Mayinga	1000	i.p.	Guinea pig	1	i.p.	5mg	3/5	0/5	Flyak et al, 2016
	EBOV	GA	Mayinga	1000	i.p.	Guinea pig	1 and 3	i.p.	5mg	5/5	0/5	Flyak et al, 2016
	BDBV	WT	200706291 Uganda	825	i.m.	Rhesus macaque	8 and 11	i.v.	35mg/kg	6/6	6/10	Gilchuk et al, 2018b
BDBV289 + BDBV223	EBOV	GA	Mayinga	1000	i.p.	Guinea pig	1	i.p.	5mg (2.5mg of each)	5/5	0/5	Flyak et al, 2016
	EBOV	GA	Mayinga	1000	i.p.	Guinea pig	1 and 3	i.p.	5mg (2.5mg of each)	5/5	0/5	Flyak et al, 2016
ADI-15946	EBOV	MA	Mayinga	100	i.p.	Balb/c mice	2	i.p.	300µg	8/10	1/20 (3/10)	Wec et al, 2017
	SUDV	WT	Boneface-USAMRIID111808	1000	i.p.	IFNαR-/- mice	1 and 4	i.p.	300µg	11/20	13/20	Wec et al, 2017
ADI-23774	EBOV	MA	Mayinga	100	i.p.	Balb/c mice	3	i.p.	300µg	9/10	0/10	Wec et al, 2019
	SUDV	WT	Boneface-USAMRIID111808	1000	i.p.	IFNαR-/- mice	1 and 4	i.p.	300µg	20/20	13/20	Wec et al, 2019
ADI-15878	EBOV	MA	Mayinga	100	i.p.	Balb/c mice	2	i.p.	300µg	8/10	1/20 (3/10)	Wec et al, 2017
	SUDV	WT	Boneface-USAMRIID111808	1000	i.p.	IFNαR-/- mice	1 and 4	i.p.	300µg	20/20	12/20	Wec et al, 2017
	EBOV	GA	Mayinga	1000LD_{50}	i.p.	Guinea pig	3	i.p.	5mg	2/6	0/4	Wec et al, 2019
									2.5mg/kg	3/6		
	BDBV	WT	BDBV/H.sap-tc/UGA/07/Butalya-811250	1000 $TCID_{50}$	i.m.	Ferret	3 and 6	i.p.	15mg then 10mg	3/4	0/2 (0/8)	Wec et al, 2017
ADI-15742	EBOV	MA	Mayinga	100	i.p.	Balb/c mice	2	i.p.	300µg	10/10	1/20 (3/10)	Wec et al, 2017
	SUDV	WT	Boneface-USAMRIID111808	1000	i.p.	IFNαR-/- mice	1 and 4	i.p.	300µg	19/20	12/20	Wec et al, 2017
	BDBV	WT	BDBV/H.sap-tc/UGA/07/Butalya-811250	1000 $TCID_{50}$	i.m.	Ferret	3 and 6	i.p.	15mg then 10mg	2/4	0/2 (0/8)	Wec et al, 2017
MP134 (ADI-15787 + ADI-23774)	EBOV	GA	Mayinga	1000LD_{50}	i.p.	Guinea pig	3	i.p.	0.5 mg	0/6	0/4	Wec et al, 2019
									1.3 mg	1/6		
									2.5 mg	5/6		
									3.3 mg	6/6		
MP134^{AF} (ADI-15787^{AF} + ADI-23774^{AF})	EBOV	GA	Mayinga	1000LD_{50}	i.p.	Guinea pig	3	i.p.	0.5 mg	1/6	0/4	Wec et al, 2019
									1.3 mg	6/6		
									2.5 mg	5/6		
									3.3 mg	5/6		
	SUDV	GA	Boneface	1000LD_{50}	i.p.	Guinea pig	4	i.p.	0.3 mg	1/6	0/6	Wec et al, 2019
									0.5 mg	2/6		
									0.8 mg	5/6		
									1.3 mg	5/6		
									2.5 mg	6/6		
	SUDV	GA	Boneface	1000LD_{50}	i.p.	Guinea pig	5	i.p.	2.5 mg	4/6	0/6	Wec et al, 2019
									5 mg	6/6		
	EBOV	WT	Makona	1000	i.n.	Ferret	2 and 4	i.p.	15 mg	4/4	0/1 (0/7)	Bornholdt et al 2019
							3 and 6			4/4		
	SUDV	WT	Gulu	1000	i.n.	Ferret	3 and 6	i.p.	5 mg	0/4	0/1 (0/5)	Bornholdt et al 2019
									15 mg	4/4		
	BDBV	WT	But-811250	1000	i.n.	Ferret	3 and 6	i.p.	5 mg	3/4	0/1 (0/5)	Bornholdt et al 2019
									15 mg	4/4		
	EBOV	WT	Kikwit	988	i.m.	Rhesus macaque	4	i.v.	25mg/kg	4/4	0/2	Bornholdt et al 2019
							4 and 7		50mg/kg, 25mg/kg	4/4		
	SUDV	WT	Boniface	1750	i.m.	Rhesus macaque	5	i.v.	7.5mg/kg	4/4	2/4	Bornholdt et al 2019
									25mg/kg	4/4		
	BDBV	WT	But-811250	938	i.m.	Cynomolgous macaque	7	i.v.	25mg/kg	5/6	0/3	Bornholdt et al 2019
2G1	EBOV	MA	Mayinga	1000LD_{50}	i.p.	Balb/c mice	1	i.p.	100µg	10/10	0/10	Fan et al, 2020
040+66-3-9C+6662+6541	EBOV	GA	Mayinga	1000$TCID_{50}$	s.c.	Guinea pig	3	i.p.	10mg/kg per mAb	6/6/	0/6	Rijal et al, 2019

Table 2 contd. Table summarising published in vivo animal challenge model data for broadly reactive mAbs.

2 Materials and Methods

For buffers, media, plasmid maps, cell lines and oligonucleotide primers, please see 7.1 Appendix 1.

2.1 RAB-9 cell culture and transient transfection

RAB-9 cells were cultured to a density of approximately 1 x 10^8 cells/ 5-stack CellSTACK® Culture Chambers (Corning, 3319) in RAB-9 culture media. Media was aspirated, cells washed in PBS and lifted using StemPro™ Accutase™ Cell Dissociation Reagent (Gibco, A1110501), for 10 minutes. Cells were aspirated from flask, diluted ~10-fold in PBS, centrifuged for 10 minutes at 1400 rpm, and resuspended in PBS and placed on ice. Cells were counted and resuspended at 5 x 10^7 cells/mL in Earle's Balanced Salts (Sigma, E3024). 3 x 10^7 cells were transferred to Gene Pulser Cuvettes (BioRad, 165-2088) and incubated with 150 µg plasmid DNA. Cells were electroporated for 20 seconds at 160-170 volts, 5 amps then gently transferred into a small volume of pre-warmed RAB-9 cell culture media using a Pasteur pipette, then further diluted in in RAB-9 media and seeded in cell culture flask (either 50 mL media in T175 or 500 mL media in 5-stack flask) and incubated at 37°C, 5% CO_2. Cells were lifted the next day, counted and cell viability measured. Expression of antigens was confirmed by FACS via staining with human antigen-specific antibody and goat anti-human Fc-specific AffiniPure F(ab')$_2$ Fragment AlexaFluor-647 conjugate (Jackson 109-606-170). Fluorescence was measured using BD FACS Canto™ 11 and analysed with FlowJo3 software. Cells were frozen down prior to vaccine formulation and stored in LN_2.

2.2 Immunisation of rabbits with RAB-9 cells expressing *Ebolavirus* GP antigens

2.2.1 Vaccine formulation

On day of vaccination, cells were thawed and washed twice with PBS. 1x10^7 cells were resuspended in 750 µL PBS per dose in 1.5 mL endotoxin free, sterile tubes (Eppendorf, #22600028) and stored on ice until use. Mixed antigen doses consisted of 2.5x10^6 of cells transfected with each GP.

2.2.2 Immunisations

Rabbits were cared for in the animal facility at UCB, Slough. All animal care, welfare monitoring, and procedures were conducted by UCB staff. Three female New Zealand White rabbits were immunised four times subcutaneously at two week intervals. First dose only was adjuvanted with complete Freund's adjuvant delivered at a separate site. Serum was taken on day of each vaccination, and rabbits euthanized two weeks after final vaccination. Lymph nodes, bone marrow, blood and spleen samples were kept on ice and processed the same day.

2.3 Rabbit B cell preparation for culture

2.3.1 Splenocyte and lymph node processing

Spleens were manually trisected, and each piece in 5mL harvest media was homogenised using a gentleMACS C tube (Miltenyi Biotech, #130-093-237) and gentleMACS Dissociator (Miltenyi Biotech, #130-093-235) (program: m_spleen_04.01). Lymph nodes were similarly homogenised (program: m_spleen_03.02). Cell suspensions were passed through 70µm cell strainer (BD Falcon, #352350). Cells were centrifuged at 1400 rpm for 10 min, and resuspended in 50 mL harvest media. Cells for immediate culture were kept on ice until use, remaining cells were resuspended in10% DMSO/FCS, aliquoted in cryovials and frozen at -80°C in a CoolCell® before transfer to LN_2 for long term storage.

2.3.2 Lymphocyte and plasma preparation from whole blood

Blood was separated using Leucosep™ tubes (Greiner Bio-One, #227289) and Mammal Lympholyte (Cedarlane Laboratories Ltd, #CL5120). After removal of lymphocyte layer, plasma layer was decanted into 50mL centrifuge tube and centrifuged at 3000 g for 10 min. Plasma was decanted into clean 50mL centrifuge tube, aliquoted and stored at -20°C.

2.3.3 Processing serum

Blood samples collected in uncoated tubes were spun at 10,000rpm for 5 min, serum layer pipetted into sterile 1.5 mL Eppendorf tube and stored at -20°C.

2.4 B cell culture and primary screening

B cell culture and primary screening was conducted using the automated high throughput platform at UCB, Slough, as described by Tickle *et al*. [216].

2.4.1 B cell culture

B cells were co-cultured with irradiated EL-4-B5 murine thymoma feeder cell line in B cell media (RPMI, 10% FCS, 1% Penicillin/Streptomycin, 2% HEPES, 1% glutamine, 0.1% 2-mercaptopehtanol, 1% Normocin) with proprietary supplement, similar to culture as described by Lightwood *et al.* [217].

[redacted]

2.4.2 Fluorescent intracellular staining

Cells in suspension were fluorescently labelled with DiI or DiO using Vybrant Multicolour Cell-labelling kit (Invitrogen, V22889). Cells were resuspended in PBS at a density of ~$1.5x10^6$ cells/mL and incubated with an excess of dye (1:375 to 1:460 dilution) with shaking, 37°C, 30 min. Cells were washed twice; samples centrifuged at 1400 rpm for 5 min, supernatant discarded and cell pellet resuspended in PBS/5% FCS. Samples were protected from light throughout and kept at 4°C until used in assay on same day.

2.4.3 Homogeneous fluorescence-based binding assay using Mirrorball fluorescence cytometer device

Stained antigen expressing cells (2.4.2) were resuspended ~$1x10^5$ cells/mL in PBS/5% FCS. Alexa Fluor® 647 AffiniPure $F(ab')_2$ Fragment Goat Anti-Rabbit IgG, Fc fragment specific (Jackson ImmunoResearch, 111-606-046) was added at a dilution of 1:2000 (1:3000 in final assay volume). 10µL of B cell culture supernatant was incubated with 20µL of cells/anti-IgG fluorophore conjugate for 2 hours in barcoded, black skirted, clear bottomed, 384-well plates (Greiner Bio-one, 781090-012)

before fluorescence measured using Mirrorball® Fluorescence cytometer and Cellista software (TTP Labtech).

2.4.4 Hit-picking and freezing of B cell cultures

[redacted]

2.5 Antigen-specific B cell isolation using fluorescent foci method with solid-phase immobilised antigen

Individual antigen specific IgG secreting memory B cells were recovered using the fluorescent foci method (US Patent 7993864/ Europe EP1570267B1) [218].

[redacted]

[redacted]

2.6 Cognate heavy and light chain V region recovery from single B cells using reverse Transcription and nested PCRs

2.6.1 Reverse transcription

[redacted]

2.6.2 Nested PCRs for amplification of V region sequences from cDNA

[redacted]

[redacted]

2.6.3 Sub-cloning into IgG and Fab expression vectors.

2nd PCR V region products were purified by gel electrophoresis and QIAquick Gel Extraction Kit (QIAGEN, 28704). 1µL purified V region DNA, 1µL ~50 ng/µL pre-linearised in house expression vectors (using restriction enzyme sites: Hind III/Xho I (VH) or Hind III/BsiW I (VL)), and 2 µL 2X GeneArt™ Seamless Cloning and Assembly Enzyme Mix (Invitrogen, A14606) were incubated at RT for 20 min for homologous recombination, then on ice for 2-3 min in 8-well PCR tube strips. Rabbit IgG heavy chain V regions were further sub-cloned into in house UCB rabbit Fab expression vectors with histidine tags. 50 µL XL-1 Blue Competent Cells (Agilent, #200249) were added to each reaction and transformed via a heat shock protocol: incubated on ice for 15 min, 42°C for 45 s, then on ice for 2 min. 150µL SOC was added and bacteria allowed to recover for 1 hour at 37°C before plating on 50 µg/mL Kanamycin agar plates. Next day, individual colonies were picked for overnight liquid culture in 3 mL LB with kanamycin, plasmid DNA purified and presence of insert confirmed by sequencing. Expression vectors contained an HCMV promoter, rabbit IgG constant region cassette and poly-adenylation sequence, as well as kanamycin resistance selection marker. DNA was sequenced using a forward primer that annealed in the promoter region of the vector.

2.7 Variable region sequence analysis

DNA sequencing of recovered and sub-cloned rabbit immunoglobulin variable regions were performed by Macrogen Europe B.V., using forward and reverse primers annealing either side of insert.

Germline gene usage analysis was performed by Dr James Snowden, UCB, using an IgBLAST program (an immune repertoire specific version of BLAST) using databases of the heavy and light chain rabbit germline sequences. Germline sequence was assigned based on similarity over the v-germline

[redacted]

sequence (Heavy chain: Framework region 1 to Framework region 3, inclusive. Kappa chain: Framework region 1 to Complementarity determining region 3, inclusive).

2.8 Flow cytometry

2.8.1 High throughput antigen binding assay using iQUE flow cytometer

Transiently transfected HEK293F cells were seeded at 5,000- 10,000 cells/well in 10 µL in FACS buffer in barcoded, black skirted, clear bottomed, 384-well plates (Greiner Bio-one, 781090-012). 10 µL of culture supernatant or of purified antibody was added and plates incubated 4°C, 1 h. Cells were washed three times; 60 µL of PBS or FACS buffer was added using Xrd-384 Multidrop dispenser, then plates were centrifuged (500 g, 5 min) and supernatant aspirated to 10 µL remaining volume to avoid loss of cells. 10 µL of 1:1000 dilution of Alexa Fluor® 647 AffiniPure F(ab')$_2$ Fragment Goat Anti-Rabbit IgG (Jackson ImmunoResearch, 111-606-046) or Alexa Fluor® 647 AffiniPure F(ab')$_2$ Fragment Rabbit Anti-Human IgG (Jackson ImmunoResearch, 309-606-008), in FACS buffer was added per well and plates incubated 4°C, 1 h, protected from light. Cells were washed as previously, and cells resuspended in remaining 10 µL volume of FACS buffer and plates covered with piercable seals to reduce evaporation. Fluorescence measured using iQUE flow cytometer (Intellicyt) with automated plate stacker arm. Program included plate shaking to resuspend settled cells and 4 s CIP (cleaning in place) between samples to reduce carry over between wells. Variation of this assay with larger cell numbers and volumes was used to confirm antigen expression on HEK293F cells before use in high throughput screening or titration experiments. Podoplanin transfected cells were used as negative controls, and expression confirmed with directly labelled PE anti-human Podoplanin antibody (Biolegend, 337004).

2.8.2 Low throughput flow cytometry antigen binding assays

Transiently transfected RAB-9 cells lifted as described in section 2.1 and 5x10^5 cells resuspended in 50 µL of 10 µg/mL of antibody in PBS in FACS tube (66-3-9C, 6535 [3]) and incubated 4°C, 40min - 1 h. Cells were washed twice; 3 mL FACS buffer added, cells centrifuged 1200 rpm, 5 min, and supernatant discarded. Cells were resuspended in 50 µL of secondary antibody in FACS buffer. PE anti-human Podoplanin antibody (Biolegend, 337004) was used at 1:20 dilution, and Alexa Fluor® 647 AffiniPure

$F(ab')_2$ Fragment Goat Anti-Human IgG (Jackson ImmunoResearch, 109-606-170) was used at 1:1000. Cells were incubated 4°C, 40min - 1 h, then washed as previously and resuspended in 300 µL FACS buffer. Data was acquired using a FACS Canto™ 11 Flow Cytometer (Beckman Dickinson).

Transiently transfected HEK293F cells were seeded in 96 well V bottom plates, $1x10^6$ cells/well in 100 µL PBS. Plates were centrifuged, 500 g, 5 min and supernatant aspirated. 45 µL of antibody solution in FACS buffer was added per well, gently mixed and incubated 4°C, 1 h. Cells were washed three times; 150 µL of PBS or FACS buffer was added, then plates were centrifuged (500 g, 5 min) and supernatant aspirated. 50 µL of 1:500 secondary antibody solution in FACS buffer was added per well, gently mixed and incubated 4°C, 1 h. For rabbit samples Alexa Fluor 488 Goat Anti-Rabbit IgG (H L) Antibody (Invitrogen, A-11008) was used. For human mAb samples Alexa Fluor 488 Goat Anti-Human IgG (H L) Antibody (Invitrogen, A-11013) was used. Cells were washed as previously, then transferred into fresh FACS cluster tubes through cell strainer lids. Data was acquired using LSR FORTESSA X20 (Beckman Dickinson).

2.9 Protein expression and purification

2.9.1 Transient transfection of HEK293F cells

Expi293F cells were transiently transfected for soluble protein expression (mAb, Fab, soluble GPs) and expression of antigens on the surface of non-adherent cells for screening and FACS experiments using ExpiFectamine™ 293 Transfection Kits (Gibco, A14525). Cells were resuspended in pre-warmed Expi293™ Expression Medium (Gibco, A1435101) at a density of $3x10^6$ cells/mL.

Per mL of final culture volume to be transfected, 1 µg of plasmid DNA (0.5 µg of each heavy and light chain plasmids when expressing mAb or Fab) was mixed with 50 µL Opti-MEM® I (1x) Reduced Serum Medium (Gibco, 31985-062). 2.66 µL ExpiFectamine 293 reagent was diluted in 50 µL and incubated at RT, 5 min. DNA and ExpiFectamine 293 mixtures were gently combined and incubated at RT, 20-30 min, then added to cells. Cells were incubated overnight at 37°C, 8% CO_2, 95-225 rpm (dependent on size of culture vessel). 5µL of Transfection Enhancer 1 and 50 µL Transfection Enhancer 2 were added

per mL of culture volume and cells returned to incubator until harvest. Transfection reagent volumes and amount of DNA was linearly scaled for cultures of 1 mL to 300 mL.

2.9.2 Transient transfection of HEK293E cells

HEK293E cells were transiently transfected for soluble EBOV GPΔTM expression. Cells were resuspended in pre-warmed FreeStyle™ 293 Expression Medium (Gibco, 12338018) with geneticin (1:1000) at a density of $3x10^6$ cells/mL. Per 500 mL culture volume, 0.33 µg plasmid DNA and 0.5 mL polyethylenimine was incubated in 20 mL FreeStyle™ 293 Expression Medium and incubated at RT, 10 min, before addition to cells. Cells were incubated at 37°C, 8% CO_2, 95-225 rpm (dependent on size of culture vessel) until supernatant harvested.

2.9.3 Expression and purification of rabbit IgG

IgG was purified from cell culture supernatant 6-7 days after transfection via Protein A affinity chromatography utilising an AKTA Pure system (GE Healthcare, Life Sciences) with CETAC autosampler and using 2ml or 5 ml column HiTrap MabSelect SuRe Protein A column (GE Healthcare, Life Sciences) dependent in culture volume with 2 min contact time. Columns were washed and IgG were eluted in 0.1 M sodium citrate, pH 3.4. Eluted fractions were neutralised with 2 M Tris/HCl pH 8.0.

2.9.4 Expression and purification of Fabs

Fabs were purified from cell culture supernatant 7 days after transfection using serial affinity chromatography. Fab were initially purified using a HisTrap™ Excel 5mL column (GE Healthcare, Life Sciences), and eluted in 5mL 0.5M NaCl / 10mM PBS/ 250mM Imidazole pH7.4. Then further purified using Gammabind Plus Sepharose Protein G 25 ml column (GE Healthcare, Life Sciences), with 20 min contact time, and eluted in 0.1M glycine, pH 2.7. Eluted Fab was neutralised using Tris HCl pH 8.5.

2.9.5 Buffer exchange and quality control analysis of rabbit IgG and Fab

Small volumes of pooled eluted IgG or Fab were buffer exchanged into PBS pH 7.4 and concentrated using Amicon Ultra Spin columns with a 30 kDa cut off membrane (Millipore, UFC905008) and centrifugation at 4000 g, before sterile filtering. Eluted samples with high concentration or in excess of 10 – 20 mL volume were buffer exchanged using two HiPrep 26/10 Desalting column (GE Healthcare,

Life Sciences) columns in series, flow rate 10 mL/min. Concentration was measured by A280 (Nanodrop spectrophotometer) and HPLC. Monomer purity was assessed by size exclusion on a Acquity UPLC system with a BEH200; 1.7 µM, 4.6 mm X 300 mm column (Waters, 176003905) and developed with an isocratic gradient of 0.2 M phosphate, pH 7.0 at 0.3 mL/min.

2.9.6 Purification of soluble *Ebolavirus* glycoproteins

300 mL culture of HEK293F cells were transfected with a pENTR4LPTOS plasmid encoding SUDV GPΔTM (1-649, Sudan virus/H.sapiens-tc/UGA/2000/Gulu-808892) with C terminal CTAG peptide. After 5 days, supernatant was harvested, clarified by centrifugation and filtered.

1.5 L of HEK293E cells were transiently transfected with a pENTR4LPTOS plasmid encoding EBOV GPΔTM (1-649 H.sapiens-wt/GIN/2014/Makona-Gueckedou-C07) with C terminal CTAG peptide. After 3 days, supernatant was harvested clarified by centrifugation and filtered. Supernatant was concentrated 10 x using tangential flow filtration and immediately purified.

GP proteins were purified using CTAG affinity chromatography, using Ctag XL affinity resin column (Thermo Fisher Scientific) and eluted in CTAG elution buffer (2M $MgCl_2$, 20 mM Tris, pH 7.4). Pooled peak fractions (~5mL) were immediately subjected to size exclusion chromatography in TBS using a HiLoad 16/600 200pg column (GE Healthcare, Life Sciences). Two peaks were analysed by Native PAGE (Section 2.9.8) and reducing and non-reducing SDS-PAGE (Section 2.9.7). For EBOV GPΔTM trimeric and monomeric fractions were pooled, for SUDV GPΔTM, monomeric fraction was pooled; samples were buffer exchanged into PBS using Amicon Ultra Spin columns (Millipore, UFC905008) and sterile filtered.

2.9.7 SDS-PAGE

SDS-PAGE was conducted using pre-cast NuPAGE 4-12% Bis-Tris Midi gels (Invitrogen) in MES buffer at 200V, 40 min. Samples were mixed with Laemmeli sample reagent (Bio-Rad, 161-0737) and boiled at 95°C, 5 min with or without DTT for reducing and non-reducing experiments, then loaded into gel

alongside Colour Pre-stained Protein standards (New England BioLabs, P7712).Gels were stained overnight using Quick Coomassie stain (Generon), and de-stained in H_2O.

2.9.8 Native PAGE

Native PAGE was conducted using pre-cast NativePAGE™ Bis-Tris 3-12% gels (Invitrogen, BN1003) in 1X NativePAGE Running Buffer (Invitrogen, BN2001) and NativePAGE Cathode Additive (Invitrogen, BN2002) at 150V, ~100 min. Samples were diluted in NativePAGE 4X sample buffer (Invitrogen, BN2003) and loaded alongside NativeMark™ Unstained Protein Standard (Invitrogen, LC0725). Coomassie stain was fixed using 40% Methanol/10% acetic acid and de-stained with 8% acetic acid.

2.10 Production of plasmids encoding GP sequences for transfection and lentivirus production

2.10.1 Sub-cloning into pcDNA(-)3.1 and pHR-SIN plasmids

Full length cDNA sequences synthesised by GENEART encoding *Ebolavirus* Glycoprotein (GP) or chimeric GP, codon optimised for expression in mammals, were sub-cloned from provided pMAT vector or DNA string into both pcDNA(-)3.1 expression vector and into pHR-SIN, a self-inactivating, lentiviral, transfer plasmid, using EcoRI and NotI restriction enzyme sites. GP sequences obtained from SDM of sequences in pcDNA(-)3.1 vectors (Section 2.10.3) were further sub-cloned into pHR-SIN.

Vectors and inserts were prepared by double digest reactions as in Table 3, and incubated at 37°C, 1 h.

Table 3. Restriction enzyme digestion of pcDNA3.1(-) and pHR-SIN

Reagent	**Volume (/ Amount**
DNA	1 µg
EcoRI-HF (New England BioLabs, R3101S)	1 µL
NotI-HF (New England BioLabs, R3189S)	1 µL
10x Cutsmart buffer (New England BioLabs, B7204S)	5 µL
RNAase-free dH2O	to 50 µL

Digested vectors and inserts were separated by gel electrophoresis using 1% agarose with SYBR Safe DNA stain (Thermo Fisher Scientific) and DNA standards (Generuler 1kb, Thermo Scientific, 5M0313) in TAE buffer. Desired bands were excised and DNA purified using QiaQuick gel purification kits

(QIAGEN), and eluted in 50 µL ddH2O. Digested inserts and vectors were ligated in a ratio of 3:1 with 20 ng of vector as per Table 4. Reactions were prepared on ice, then incubated at RT, 7 min, then kept at 4°C until transformations performed with 10 µL of ligated reaction.

Table 4. Ligation reactions

Reagent	**Volume / Amount**
DNA insert	60 ng
DNA vector	20 ng
Quick Ligase (New England BioLabs, M2200)	0.5 µL
2x Quick Ligase buffer (New England BioLabs, M2200)	10 µL
RNAase-free dH2O	to 20 µL

2.10.2 Transformation of TOP10 *E.coli*

E.coli were thawed on ice, 25 µL of cells transferred to 15 mL pop capped tube, plasmid DNA added, and incubated on ice, for 30 min. Cells were incubated for 30 s, 42°C then transferred to ice for 2 min. 250 µL SOC was added per reaction and cells recovered for 1 h, 37°C, 225 rpm. Appropriate volume of cells (250 µL for SDM reactions, 50 µL for plasmid reactions) was spread onto LB agar plates with appropriate antibiotic for selection, and incubated overnight at 37°C.

Individual colonies were picked into 3 mL LB/carbenicillin (0.1 µg/mL) and incubated overnight 37°C, 225 rpm. Plasmid DNA was purified (Section 2.11) and inserts were sequenced by GATC Bioscience using primers # 7-14 (Table 8) as appropriate for vector and insert. For full coverage of GP, pairs of primers annealing wither side of multiple cloning site in the vector and a third primer annealing part way through the GP sequence were required.

2.10.3 Site-directed mutagenesis (SDM)

R134A and R136A mutations were introduced into DNA encoding full length EBOV (Makona) GP using a QuikChange II Site-Directed Mutagenesis Kit (Agilent, #200524).

Primers were designed as per enzyme manufacturer's guidelines (Primers #1-6, Table 8) and synthesised and purified via HPLC by Integrated DNA Technologies Ltd. Primers were resuspended in DNAse-free H_2O. Purified pcDNA3.1(-) plasmid encoding EBOV GP (Makona) used as template. 50 µL

reaction mixes were set up on ice in 0.2 mL PCR tubes as per enzyme manufacturer's protocol. Reactions were conducted using a thermocycler as per Table 5.

Table 5. Thermocycler reaction for site-directed mutagenesis

Cycles	Temperature	Time
1	95°C	30 s
16	95°C	30 s
	55°C	1min
	68°C	7 min 30s
/	4°C	Hold

1 µL DpnI enzyme was added per reaction and incubated 1 h, 37°C. TOP10 *E.coli* (Invitrogen, C404003) were transformed with 1 µL of reaction mixture (Section 2.10.2).

2.11 Plasmid purification

All plasmid DNA purifications from bacterial cultures were performed via standard alkaline lysis protocols using QIAGEN plasmid purification kits appropriate for volume of bacteria; for 3 mL cultures used QIAprep Spin Miniprep kit, for 100-200mL cultures used Plasmid Plus Maxi Kit, for larger volumes used Plasmid Plus Giga Kit.

2.12 Enzyme-Linked Immunosorbent assays

2.12.1 Sandwich ELISA for detection of rabbit IgG

NUNC Maxisorp 96 well plates were coated overnight at 4°C or 1 hr at 37°C with 100 µL/well of 2µg/mL capture antibody (Goat Anti-Rabbit IgG, Fc fragment specific, Jackson ImmunoResearch, 111-006-046) in carbonate coating buffer (15mM Na_2CO_3, 35mM $NaHCO_3$). Plates were washed 4x with PBS/Tween after each following incubation step until development. After washing, plates were blocked with 200 µL/well 1% BSA/PBS and incubated at RT for at least 1 hour. After washing standards (ChromPure Rabbit IgG, whole molecule, Jackson ImmunoReserach, 011-000-003) and samples were added to plate. Purified IgG standard was titrated though a half-log series from 1000 ng/mL to 1 ng/mL in 1% BSA/PBS, with 100 µL/well added in triplicate to each plate. Samples were diluted as appropriate (1:10 for TAP transfection supernatants) and titrated in half-log series using 1% BSA/PBS, with 100

µL/well added to plate. Plates were incubated RT, 1h. 100µL/well of 1:5000 dilution of Peroxidase AffiniPure F(ab')$_2$ Fragment Goat Anti-Rabbit IgG, F(ab')$_2$ fragment specific (Jackson ImmunoResearch, 111-035-047) in 1% BSA/PBS was added and plate incubated RT, 1h. Plates were developed at RT using 100 µL/well of TMB buffer. Development was stopped with of 50 µL/well 0.5M H_2SO_4 at RT for 5 min. Absorbance was measured 450nm.

2.12.2 Antigen ELISAs

96 well NUNC Maxisorp plates were coated overnight at 4°C with 50 µL/well of 2 µg/mL soluble *Ebolavirus* GP protein (in house) in carbonate coating buffer (15mM Na_2CO_3, 35mM $NaHCO_3$). Plates were washed 4 times with PBS/Tween after each following incubation step until development. After washing, plates were blocked with 1% BSA/PBS and incubated at RT for at least 1 hour. After washing, plates were incubated with 50 µL/well of culture supernatant or control antibody (10 µg/mL) in 1% BSA/PBS, RT, 1 h. After washing, 50 µL/well of peroxidase conjugated anti-IgG secondary antibody diluted 1:5000 in 1% BSA/PBS was added; for rabbit samples F(ab')$_2$ Fragment Goat Anti-Rabbit IgG, Fc fragment specific (Jackson ImmunoResearch, 111-036-046), for human samples F(ab')$_2$ Fragment Goat Anti-Human IgG, F(ab')$_2$ fragment specific (Jackson ImmunoResearch, 109-036-097). Plates were incubated at RT for 1 hr. After washing plates were developed as in Section 2.12.1.

2.12.3 Fusion loop peptide ELISA

96 well NUNC Maxisorp plates were coated with 100 µL/well 5 mg/mL Streptavidin in PBS overnight, 37°C. Plates were washed three times in PBS/Tween and blocked with 200 µL/well casein, RT, 1 - 1.5 h. Plates were washed as previously, and incubated with 100 µL/well of fusion loop peptide (Terry Baker, UCB) diluted to 20 nM in casein, RT, 1 h. Fusion peptide sequence GLAWIPYFG(AGV) corresponds to residues 528-536 of EBOV GP with additional residues shown in brackets. Plates were washed as previously and incubated with 50 µL/well of antibody in PBS, RT, 1 h. Plates were washed as previously and incubated with 50 µL/well of either Alkaline phosphatase-conjugated Anti-human IgG (Sigma-Aldrich, A3187) or anti-rabbit IgG (Sigma-Aldrich, A8025) in PBS, RT, 1 h. Plates were washed six times in PBS/Tween followed by PBS. Plates were developed by addition of 4NPP development buffer, and incubated 2h -24 h, RT protected from light, then absorbance at 450nm measured.

2.13 Biotinylation of soluble proteins

Soluble proteins were biotinylated using EZ-Link Sulfo-NHS-Biotin kit (Thermo Scientific, #21326 or #21217). Excess free biotin was removed with Zeba™ Spin Desalting columns (Thermo Scientific, # 89890 or #89882). Biotinylation of mAbs confirmed via dot blot revealed with ExtrAvidin-alkaline phosphatase conjugate (Sigma, E2636) and NBT/BCIP development.

2.14 Fluorometric microvolume assay (FMAT)

Antigen expressing cells were resuspended in 5% FCS/ PBS at a density of $2x10^5$ cells/mL. Species appropriate Alexa Fluor 647 conjugated secondary antibody was added at a 1:2000 dilution (1:3000 in final assay volume): for rabbit samples, Alexa Fluor® 647 AffiniPure $F(ab')_2$ Fragment Goat Anti-Rabbit IgG, Fc fragment specific (Jackson ImmunoResearch, 111-606-046); for human samples, $F(ab')_2$ Fragment Goat Anti-Human IgG Fcγ fragment specific (Jackson ImmunoResearch, 109-606-170). 10 µL/well of sample containing IgG was transferred into barcoded, black skirted, clear bottomed, 384-well plates (Greiner Bio-one, 781090-012), before addition of 20 µL/well of antigen-expressing cells and secondary antibody mix. Plates were incubated in the dark, 1 h, RT, before fluorescence read using FMAT 8200 Cellular Detection System (Applied Biosystems) with stacker, gating thresholds applied and data analysed using 8200 Analysis software (Applied Biosystems).

2.15 Immunofluorescence assays (IFA) using GP expressing MDCK-SIAT 1 cell lines

In 96 well flat bottom tissue culture plates, MDCK SIAT-1 cells were seeded at a density of $3x10^5$ cells /per well in 100 µL D10 media and incubated for 18 hours, 37°C, 5% CO_2. Cells were washed in PBS to remove D10 media, and 50 µL of mAb diluted in PBS added. Cells were incubated for 1 h at RT or 4°C. Plates were washed three times with PBS. 50 µL 5 µg/mL appropriate secondary antibody Alexa Fluor® 647 or 488 conjugate was added to each well and cells incubated for 1 h at RT or 4°C (Anti-rabbit IgG Alexa Fluor 647 or 488 conjugate (Invitrogen, A21244 or A11008), anti-human IgG Alexa Fluor 647 or 488 conjugate (Invitrogen, A21445 or A11013). Plates were washed at least three times with PBS and 100 uL PBS added per well. Fluorescence measured using Clariostar plate reader (BMG Labtech). If

plates could not be read immediately, cells fixed with 1% formalin/PBS. Fluorescence plate reader settings were as previously described by Xiao *et al.* [219].

2.15.1 Antibody competition IFA

Competition experiments were conducted using an adapted version of the assay described in Section 2.15. Biotinylated mAb1 at 5 µg/mL in PBS and unbiotinylated mAb2 at 50 µg/mL in PBS were mixed in equal volume prior to addition to cells. Each assay plate included competition controls where mAb1=mAb2 and additionally where mAb2 is a mAb against an irrelevant antigen or PBS only as a minimum competition control. After incubation and washing, 50 µL of 2 µg/mL of Streptavidin Alexa Fluor® 647 conjugate (Invitrogen, S21374) was added to each well. After incubation and washing, fluorescence measured and degree of competition was determined by: (X–Minimum binding)/(Maximum binding – Minimum binding), where 'X' is binding of the biotinylated mAb in presence of competing mAb, 'minimum binding' is the signal from the biotinylated mAb in the presence of self (unbiotinylated) and 'maximum binding' is the signal from the biotinylated mAb in presence of a non-competing non-GP mAb.

2.15.2 IFA with thermolysin digested-GP

Binding to thermolysin digested GP was determined using an adapted version of the assay described in Section 2.15. Cells were initially seeded in round-bottom 96 well tissue culture plates. Cells were washed and incubated in 100 µL 0.25mg/mL thermolysin (THL) (Sigma, P1512) in HM buffer or HM buffer alone for 1 h, with gentle shaking at RT or 37°C, 5% CO_2. After this and every subsequent incubation, cells were washed three times with plates spun at 500 g and supernatant aspirated between washes. Cells were incubated with 50 uL 10 µg/mL mAb or 10 µg/mL biotinylated NPC1-C protein in PBS, and incubated for 1 h at RT, then 5 µg/mL anti-rabbit IgG Alexa Fluor 647 conjugate (Invitrogen, A21244) or anti-human IgG Alexa Fluor 647 conjugate (Invitrogen, A21445) or Streptavidin Alexa Fluor® 647 conjugate (Invitrogen, S21374, and 5 µg/mL of wheat germ agglutinin (WGA) Alexa Fluor 488 conjugate (Invitrogen, W11261) in 50uL PBS for 1 h, at RT with gentle shaking, protected from light. Plates were centrifuged before fluorescence read at both 625-30/680-30 and 488-14/535-30. Gain was adjusted to give a ratio of approximately 1 in wells containing cells that had been stained

with both WGA-488 and anti-Rabbit IgG-647 conjugates only (i.e. no primary antibody added). Wells with too few cells as determined by WGA-488 signal were excluded from analysis. For inhibition of THL digestion assays, cells were pre-incubated with 50 µL 10 µg/mL mAb before incubation with THL, rather than after digestion.

2.15.3 NPC1-C competition IFA

Cells were seeded, treated with THL, and washed as described in Section 2.15.2. 10 µg/mL biotinylated NPC1-C and 50 µg/mL mAb in PBS were premixed (1:1 by volume) and 50 uL added per well of digested cells. After incubation and washing, 50 µL of 2 µg/mL of Streptavidin Alexa Fluor® 647 conjugate (Invitrogen, S21374) and 5 µg/mL of wheat germ agglutinin (WGA) Alexa Fluor 488 conjugate (Invitrogen, W11261) for 1 h, at RT with gentle shaking and protected from light. Plates were centrifuged and fluorescence read as described in Section 2.15.2.

2.16 Immunoprecipitation of thermolysin digested GP expressed on cells

MDCK SIAT-1 cells expressing GP were seeded at 8-12 x10^5 cells per well in 6 well tissue culture plate (Corning, CLS3516) in 2 mL D10 and incubated overnight, 37°C, 5% CO_2. Next day, cell surface proteins were biotinylated using EZ-Link Sulfo-NHS-Biotin (Thermo Scientific, 21217). Media was aspirated from wells and cells rinsed with PBS. 0.5 mL of 0.5 mg/mL biotin reagent was added per well and incubated at RT, 30 min. Biotin reagent was removed and free biotin quenched by addition of 100mM glycine. Glycine was removed and 0.5 mL thermolysin (THL) diluted in HM buffer or HM buffer alone added per well. Plates were incubated with gentle shaking, 1 h, RT. THL was gently aspirated, replaced with 0.5 mL lysis buffer (Alfa-Aesar, J62805.AK) and 1 µL of protease inhibitors (Sigma-Aldrich, P8340), and plate incubated with gentle shaking for 30 min, RT. Buffer and cells removed from each well to 1.5 mL Eppendorf tube and centrifuged to pellet cell debris. Supernatant (0.5 mL) was transferred to 1.5 mL Eppendorf containing 7.5 µg immunoprecipitation antibody and 100 µL Protein A Sepharose slurry (Sigma-Aldrich, P3391) (5% w/v in lysis buffer). Samples were incubated with rotation, 1 h, 4°C. Samples were centrifuged in pre-cooled microfuge, 5 min, 13,000 rpm, 4°C, and supernatant discarded. Sepharose A pellet was washed twice with 0.5 mL chilled lysis buffer and centrifuged a final

time. Supernatant discarded and final volume removed using gel loading tip (Greiner Bio-one, 770280). Protein was eluted from Sepharose A by addition of 35 µL of sample loading buffer (Sigma-Aldrich, S3401-10) and incubating 5 min, 80°C. Samples were centrifuged and supernatant loaded into Bolt 4-12% Bis-Tris Plus gel (Invitrogen) using sample loading tips. Electrophoresis was conducted in MES buffer at 100-120 V for 40 min with protein standards (NEB, P7119S, or Bio-Rad, #1610373). Proteins were transferred to nitrocellulose membrane and membranes blocked in 5% milk powder/PBS overnight, 4°C. Membranes were washed three times with TBS (incubated 5-10 min per wash), then incubated in 1 µg/mL dilution of Streptavidin Alexa Fluor 647 (Invitrogen, S21374) for 1 h, RT with rocking. Membranes were washed three times with TBS (incubated 5-10 min per wash) then imaged using iBright FL100 (Thermo Fisher Scientific) using the auto exposure setting. In studies of inhibition of THL digestion cells were pre-incubated with 500 µL 15 µg/mL Fab in PBS after biotinylation and glycine incubations for 1 h, RT.

2.17 Generation of MDCK SIAT-1 cell lines stably expressing GP at cell surface

2.17.1 Production of lentivirus

HEK293T cells were passaged using 2 mM EDTA, washed in PBS and resuspended in OPTI-MEM® I Reduced Serum Medium (Gibco, 31985-062) at 10^6 cells/mL. 1 mL of cell suspension was seeded per well of 6-well tissue culture plate (Corning). 1 µg each of envelope plasmid pCMV-VSV-G (Addgene: #8454, encoding VSV envelope protein), packaging plasmid pCMVdeltaR8.91 (encoding the HIV pol, gag, tat and rev viral genes as well as the rev-response element (RRE)) and pHR-SIN plasmids were diluted in 250 µl OPTI-MEM® I Reduced Serum Medium. 10 µl Lipofectamine® 2000 (Invitrogen, 11668-019) was diluted in 250 µL OPTI-MEM® I Reduced Serum Medium and incubated at RT, 5 min. DNA-OPTI-MEM and Lipofectamine- OPTI-MEM mixtures were combined, mixed gently and incubated at RT, 20 min, then added to cells (500 µL per well of cells). Cells were incubated at least 4 h, 37°C, 5% CO2 media removed and replaced with 2 mL D10 and returned to incubator for 48-64 h. Media containing virus was aspirated from well using a Pasteur pipette and centrifuged 1400 rpm, 5 min. Supernatant was transferred to a fresh tube and stored for up to 24 hours at 4°C before use.

2.17.2 MDCK-SIAT 1 cell culture

MDCK-SIAT 1 cells expanded, passaged using trypsin-EDTA (Sigma-Aldrich, T3924) and maintained in culture in vented cap T175 flasks at 37°C, 5% CO_2 in D10 media, or stored in 10% DMSO/FCS in LN_2 until use.

2.17.3 Transduction of MDCK SIAT-1 cells with lentivirus

On day 1, MDCK-SIAT 1 cells were seeded at 10^5 cells/well of 6-well tissue culture plate in 2 mL D10. Cells were incubated overnight with 1mL lentivirus and 8 µg/mL polybrene at 37°C, 5% CO_2. On day 2, media was removed and replaced re-infected as previously. On day 3, media was replaced with 2 mL D10 only and cells returned to incubator. When cells reached confluencey they were lifted using trypsin and expanded before freezing.

2.17.4 Sorting

Cells were thawed, washed and allowed to recover in D10 for 1 h, 37°C, 5% CO2. Cells were resuspended in 500 µL 2% FCS/PBS and transferred to FACS tube then resuspended in 500 µL antigen specific antibody at 50 µg/mL in 2% FCS/PBS and cells incubated at RT for 1 h. Cells were washed and incubated in the dark with 5 µg/mL of anti-human IgG Alexa Fluor 647 conjugate and live/dead stain, 20-30 min, RT. Cells were washed and transferred to clean FACS tube via cell strainer. ~100,000-250,000 cells were sorted into tubes containing D10 by Craig Waugh, WIMM FACS facility. Cells were seeded in T25 flask and expanded before use.

2.18 Production of S-FLU pseudotype viruses

Media was removed from confluent T175 flasks of MDCK SIAT-1 cells expressing GP, replaced with 10 mL PBS and flasks incubated at RT for 10 min. PBS was removed and replaced with 10 mL 75% Viral growth media (VGM) in PBS and seed S-FLU virus at high MOI (~ 1 $TCID_{50}$) added. Flasks were incubated for 1 h 37°C, 5% CO_2. Seed virus and media were aspirated, cells washed with PBS and ~40mL VGM added per flask. Virus was left to propagate for 48-72 h at 37°C, 5% CO_2. Media containing virus was aspirated and centrifuged at 1400 rpm for 8 min to pellet cells. Supernatant was transferred to a clean tube, aliquoted and stored at -80°C. Virus was titrated to confirm production and titre for

use in assays. Virus was titrated 1:2 in VGM across a 96 well flat bottomed tissue culture plate (Corning) and 100 µL 3 x 10^5 MDCK SIAT-1 cells/mL added per well. After 24 h, media was removed from each well and replaced with 100 µL 10% formalin for 30 minutes, 4°C. Formalin was removed and 100 µL PBS added to each well. GFP fluorescence was read using a Clariostar plate reader (BMG Labtech) as described in Xiao *et al.* [219].

2.19 S-FLU microneutralisation assay

In a 96 well, flat bottom tissue culture plate (Corning), virus diluted in 50 µL VGM was incubated with 50 µL serum or mAb diluted in PBS at titrated concentration for 2 hours, 37°C, 5% CO_2, before addition of 100 µL MDCK SIAT-1 cells in VGM (3 x 10^5 cells/mL). Virus was used at a dilution previously determined to give maximum infection of cells. After 20-24 hours incubation at 37°C, 5% CO_2, supernatant was aspirated from cells, and cells fixed with 100 µL 10% formalin for 30 minutes, 4°C. Formalin was removed and 100 µL PBS added to each well. GFP fluorescence was read using a Clariostar plate reader (BMG Labtech) as described in Xiao *et al.* [219]. Maximum infection was indicated by signal from cells infected with viruses pre-incubated with PBS only or a non-GP binding antibody. Minimum signal was indicated by uninfected cells. Inhibitory concentration at 50% (IC_{50}) and 90% (IC_{90}) was derived by linear interpolation.

2.20 Epitope mapping of monoclonal antibodies via peptide phage display

2.20.1 Biopanning

In brief, 3-8 wells of 96 well NUNC MAXISORP plates were coated with 1.5 µg of mAb in PBS and incubated overnight at 4°C. Plates were washed twice with 0.1%Tween/PBS. Plates were blocked with 3% milk powder/PBS at room temperature for at least 1 h. Phage library aliquot was thawed and incubated in approximately 2.5x volume of 6% milk powder/PBS, at RT with mixing for at least 1 h. Plates were washed twice with 0.1% Tween/PBS. Each mAb coated well was incubated with 100 µL of blocked phage (~2.5x10^{13} CFU/mL), at RT for 1 h with gentle shaking. Plates were washed ten times with 0.1% Tween/PBS. Bound phage were eluted in 0.1 M HCl (Sigma, 2104) and neutralised with 1 M Tris-HCl pH 8.0 (Sigma, T2694). Eluted phage (R1) were used to infect TG1 *E.coli* (F' *traD36 proAB*

lacIqZ ΔM15] supE thi-1 Δ(lac-proAB) Δ(mcrB-hsdSM)5(rK - mK -))(Lucigen, 60502) cells and grown on 2TY plates supplemented with carbenicillin and glucose at 30°C. R1 output bacteria were scraped (a small aliquot of each was flash frozen and stored at -80°C), grown at 37°C to OD~0.5 and rescued using M13KO7 helper phage (NEB, N0315S). Rescued cultures were resuspended in 2TY-Carbenicillin/Kanamycin without glucose and grown overnight at 30°C to produce phage for R2 panning. Next day phage were precipitated using 20% PEG8000 2.5 M NaCl, resuspended in 1-2 mL PBS and blocked with 6% Milk/PBS for a R2 biopanning including a subtraction step. For subtractive selection, 6 wells of 96 well NUNC MAXISORP plates were coated with 0.5 µg Chromopure Rabbit IgG (Jackson, 011-000-003) in PBS. For mAb 66-3-9C, Chromopure Human IgG (Jackson, 009-000-003) was used. Plates were washed and blocked as before. After washing, each Chromopure coated well was incubated with 100 µL of blocked phage at RT for 1 h with gentle shaking. Unbound phage were transferred to 3-6 wells of 96 well NUNC MAXISORP plate coated with 0.5 µg of monoclonal antibody of interest and pre-blocked with 3% Milk/PBS for the second round of panning. Phage were eluted (R2), and used to infect TG1 cells as previously for R1. Finally, R2 bacterial lawn were scraped and glycerol stock were prepared and flash frozen at -80°C.

2.20.2 Sample preparation and Ion Torrent Sequencing

R2 phagemid DNA was extracted and purified from bacterial glycerol stocks; 100 µL of each bacterial library was thawed and phagemid DNA purified using QIAprep Spin Miniprep kit (QIAGEN, 27104). DNA was eluted in 50 µL H_2O.

2.20.2.1 Addition of barcodes via PCR

PCR amplicon DNA containing peptide sequences of interest (9mer or 13mer) was generated using R2 phagemid DNA as template and barcoded primers were used that added DNA encoded barcodes to distinguish libraries generated by panning against different antibodies and that added P1 adaptor sequences designed to enable attachment of DNA to beads for emulsion PCR required for Ion Torrent sequencing platform. For 9mer phagemid libraries a single step PCR was utilised. For 13mer phagemid libraries a preparatory PCR was required to incorporate an adaptor sequence. Between PCRs samples were column purified using QiaQuick PCR Purification kit (QIAGEN) to remove primers.

2.20.2.2 Preparation of samples for Ion Torrent Sequencing

PCR products were separated by gel electrophoresis using low melting point agarose and bands at ~350bp were excised and purified using Nucleospin columns (Machery-Nagel, 740609-250). DNA concentration was quantified using A260 (NanoDrop spectrophotometer) and a Bioanalyser 2100 high sensitivity DNA chip. 100 ng of DNA from each individually barcoded library was pooled. Pooled DNA was purified using magnetic AMPure XP beads (Beckman Coulter, A63880) and DNA concentration quantified using A260.

2.20.2.3 Ion Torrent Sequencing

Pooled barcoded DNA was sequenced via an Ion Torrent PGM service (Macrogen, Inc.) on a 318 chip.

2.20.2.4 Z-score analysis and peptide ranking

NGS data was processed as described in Naqid *et al* [220].

A two proportion Z-score analysis was conducted to identify peptides specifically enriched by mAb of interest. For the 9mer library, each set of peptides enriched by a mAb of interest was compared to those enriched by mAb 11886. Peptides enriched by mAb 11886 were compared to those enriched by human GP-binding antibody 66-3-9C. For the 13mer library, each set of peptides enriched by a target mAb of interest was compared to those enriched by a control non-GP specific rabbit IgG which was panned against the library in parallel to the target GP-binding mAbs. Z-score takes into account the frequency of a peptide sequence in the total number of sequences in the sample, as well as the ratio of the number of copies of a sequence in the sample isolated against the target antibody and number of copies of the same sequence in the sample isolated against the control antibody. Z-score can be used to rank peptide sequences by relative statistical importance and is calculated using the formula:

$$z = \frac{p1 - p2}{\sqrt{\frac{p1(1-p1)}{n1} + \frac{p2(1-p2)}{n2}}}$$

Where n1 = number of peptide sequences obtained with the target sample; n2 = number of peptide sequences obtained with the control sample; p1 = the number of sequences obtained for a specific

peptide against the target sample/n1; p2 = the number of sequences obtained for a specific peptide against the control sample/n2 [220].

Peptides were ranked based on Z-score i.e. for specific enrichment in the total reads from Ion torrent sequencing by target mAb of interest. The top 100 peptides were used in all further analysis.

2.20.2.5 Identification of peptide motifs

Motifs in the top 100 ranked peptides by z-score were identified using the Multiple EM for Motif Elicitation (MEME) algorithm [5]. For MEME analysis it was assumed that each motif was expected to occur zero or one time per sequence (zoop). Algorithm was constrained such that motifs occur in a minimum number of ten peptide sequences. Motifs with an E value <0.05 were considered significant.

2.20.2.6 Alignment of peptide motifs with GP sequences

Peptide motifs were aligned with GP sequences from EBOV (ATY51135), SUDV (YP_138523.1), BDBV (YP_003815435.1) and TAFV (YP_003815426.1) using Clustal Omega algorithm and MegAlignPro software (Version: 11.2.1).

2.21 Cryo-Electron Microscopy of EBOV GPΔmuc-11886 Fab complex

This study was carried out entirely by Dr Amar Parvate and colleagues in Professor Erica Ollmann Saphire's lab, La Jolla Institute for Immunology, California, USA, using Fab produced by in the UK as described in Section 2.10.4. The methods described here are an abbreviated version of methods supplied by Dr Parvate and are included for reference.

2.21.1 Preparation of fab-GP complexes

Four fold molar excess of the fab was complexed with Zaire EBOV GPΔmuc trimer to give a molar excess of Fab (160 µg GP trimer + 180 µg Fab) at 40C for 3-4 h or overnight. Complex was purified using size exclusion (S6 column), assessed by SDS PAGE and concentrated to 0.44 mg/ml by centrifuging at 5000g for 5 min using a 10 kDa MWCO filter.

2.21.2 Preparation of grids

Approximately 3 µL of the sample was mixed with 1µL 0.02M lauryl maltose neopentyl glycol (LMNG) detergent and loaded onto a glow discharged C-flat grid (R2/1). The grids were plunge frozen using a Vitrobot (Thermo Fisher) using 9-10 s blot time and a blot force of 8. Girds were stored under liquid nitrogen till further use.

2.21.3 Cryo-EM data acquisition

Grids were loaded on an FEI Titan Krios at the Arizona State University (Tempe, AZ, USA) and movies were collected at 130,000x nominal magnification, at a pixel size 1.096 Å Gatan K2 Direct Electron Detector in counting mode. A total of 30 frames were collected per image and the total dose was 34 e^-/ $Å^2$/s. Movies were initially collected as *.tiff* and were gain corrected and converted to *.mrc* using the *unpack* command from IMOD. Movies were then imported into Cryosparc for single particle analyses.

2.21.4 Single particle analyses

Movies were subjected to motion correction and CTF estimation using standard options in Cryosparc. Particle co-ordinates were extracted using Blob Picker at a box size of 440 pixel. Micrographs were sorted using Inspect Picks option and micrographs with estimated CTF >10 Å and motion correction for >100 pixels were discarded. Co-ordinates from remaining micrographs were extracted and subjected to reference free 2D classification. A small subset of classes containing optimal top and side views were selected for *ab initio* modelling, with particles showing suboptimal Fab occupancy excluded. Classes with GP trimer showing suboptimal occupancy of the Fab.

2.21.5 3D refinement

72,000 particles were used for final refinement using the Homogenous Refinement option in Cryosparc. C3 symmetry was imposed in subsequent rounds. Particles were then subjected to a round of local and global CTF correction and again subjected to Homogenous Refinement. A progressively tighter circular mask was applied to improve the local resolution. Final resolution of the map was calculated by FSC gold standard and was estimated at 0.143 as 5.9 Å.

2.21.6 Homology modelling and visualization

All visualizations were created using UCSF Chimera (developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311, [221]. The PDB map of ADI-15946 [115] was used as a template to build the homology model for the VH and VL region of 11886. Models were generated using the Expasy online server both with and without a template, with both models converging well [222]. For the GP core, PDB map 5JQ7 was used as a template to build the homology model.

3 Antibody discovery

3.1 Introduction

High-throughput antibody discovery platforms allow the mining of the diversity of B cells for antibodies with rare characteristics. This chapter describes the discovery process for generating a panel of six rabbit monoclonal antibodies that bind Glycoprotein (GP) from all four *Ebolaviruses* known to cause human disease. An additional panel of 13 partially or non-cross-reactive anti-GP antibodies was also generated. This work was undertaken using the high throughput Core Antibody Discovery Platform (CADP) at UCB. This platform utilises B cell culture, fluorescent based screening for binding properties of secreted IgG, fluorescent foci formation for recovery of single B cells, and RT-PCR followed by nested PCRs for recovery of variable region sequences.

3.1.1 Approaches for generation of broadly reactive anti-GP antibodies from immune samples

Antigen-specific B cells can be isolated from convalescent or vaccinated samples, but each strategy encounters different hurdles for antibody discovery. Presenting immunodominant non-neutralising epitopes or shielding neutralising epitopes from the humoral immune system, amongst other directly immunomodulatory mechanisms used by pathogens, impairs the mounting an effective immune response. In *Ebolavirus* infection the glycosylated MLD and GC may act as an immunodominant glycan shield, preventing access to epitopes around the chalice of the pre-fusion GP structure and diverting the immune response from sites of vulnerability to neutralisation [223]. The cleavage of these domains is required for access to the receptor binding region, and as this only occurs within the endosomal pathway, the receptor binding domain may not be an effective target for mAbs that act in the extracellular fluid, despite the fact that the inhibition of NPC-1 binding by mAbs prevents infection [212]. In addition, it has been suggested that the shorter soluble protein sGP, which shares the N terminal domain of full length GP in the viral capsid, is produced for antigenic subversion of the immune response to GP [148, 149]. Earlier antibody discovery campaigns were concerned that cross-

reactivity between GP and sGP would make a mAb therapeutic less effective, however, antibodies that bind sGP (in addition to viral capsid GP) have been shown to be protective *in vivo* [173].

Subunit vaccines and rational vaccine design aim to avoid these immune evasion strategies and potentially focus the immune response onto vulnerable sites on an antigen [224]. However, antibodies generated by vaccination are only useful if the immunogen used reflects a relevant form of the antigen, e.g. correctly folded or glycosylated protein; this is of less concern in strategies using convalescent samples for mAb discovery, where only physiologically relevant antigen will have been encountered.

Access to convalescent samples is limited due to logistical, ethical and safety concerns; it is also highly probable that convalescent ED patients have only experienced infection with a single species of *Ebolavirus*. When vaccinating human volunteers a highly developed vaccine made to GMP is required, and there are limitations in the number of vaccinations that can be given which may impede the opportunity to explore vaccination strategies that would increase the chance of generating a highly cross-reactive antibody response to multiple species of *Ebolavirus*. With animal models it is possible to sample B cells from compartments that it would be unethical or invasive to sample from humans, such as the spleen, lymph nodes or bone marrow. Vaccinating animals with pre-clinical immunogens specifically for antibody discovery is therefore a cost-effective and safe way of trying to induce cross-reactive B cells that can be harvested from different lymphoid organs.

Another advantage of using animals to generate mAbs is the opportunity to exploit the different Ig diversification mechanisms and structures present in different species to find mAbs that can target vulnerable sites on the antigen with therapeutic potential.

As discussed in Section 1.4.3.2.4, in contrast to humans and mice, bovines produce a significant amount of Ig with ultra-long heavy chain CDR3 sequences [198]. This is likely due to their greater reliance on Ig diversification mechanisms occurring after initial germline rearrangements. Cows rely on only 12 heavy chain V gene segments, compared to >100 available to humans, but this lack of diversity is compensated for by a greater degree of somatic hypermutation (SHM) in the

subsequent diversification occurring in gut-associated lymphoid tissues (GALT) prior to antigen exposure. This is exemplified by the finding that in an next generation sequencing (NGS) study of bovine Ig repertoires, almost all ultra-long CDR H3 containing sequences utilised the same V, D and J gene segments [225, 226]. These ultra-long heavy chain CDR3s protrude from the traditional antigen binding site of the antibody and form an antiparallel β-stranded stalk tipped by a disulphide-bonded knob domain responsible for antigen binding [225, 227]. SHM in many species such as humans generates small insertions and deletions or point mutations, whereas in bovines the mechanism generates large insertions and deletions that contributes to the generation of long CDR3 sequences by diversifying the arrangement of the cysteines and disulphide bonds required for this structure [226].

Long CDR3s are also feature of camelid nanobodies. Camels and Llamas produce unconventional antibodies consisting of homodimers of heavy chains only (heavy chain antibodies, HCAb), each chain containing a variable region and two constant domains [228]. HCAbs are produced from dedicated germline genes in the same locus as conventional Ig and can be up to 50% of Ig in the serum [229]. As the N-terminal variable domain (VHH) consists of only a single chain, it contains only three CDR loops, with long (can be >20aa) CDR3 sequences that can be solely responsible for antigen binding [228, 230]. Similarly to camelid HCAbs, cartilaginous fish also produce a type of HCAb termed Ig-based New Antigen Receptors (IgNARs)[231]**.** The variable domains of these HCAbs can be expressed alone and are conducive to the production of phage display libraries for antibody discovery [232, 233]. These recombinant antigen-binding domains, or nanobodies, have been shown to be capable of neutralising various viruses including HIV-1 and Severe Acute Respiratory Syndrome virus 2 (SARS-2) [234, 235].

Alongside these examples, rabbits provide yet another system for generating monoclonal antibodies with a distinct ontogeny to human repertoires. Similar to bovines, the initial gene rearrangements of rabbits Ig repertoires are restricted, but greater diversification is be subsequent mechanisms. The advantages of using rabbits for antibody discovery and the ontogeny of their B cell repertoires is discussed next.

3.1.2 Rabbit antibody repertoire and ontogeny

Broadly reactive anti-GP mAbs have been isolated from human survivors [99, 100, 103, 191], human vaccinees [3], immunised macaques [190], and immunised mice [208, 209]. Peptide immunisation of rabbits to generate reagents for GP ELISAs [236], and more recently, immunisation of rabbits with a peptide representing the HR2-MPER domains of GP to generate broadly neutralising anti-GP antibodies in rabbits has been reported [100]. However, when this project started, the rabbit immune repertoire was yet to be widely exploited in the published literature, particularly for the purpose of identifying therapeutic candidates.

Rabbits have a long history as a model system for exploration of immunology and for development of antibody-based tools and therapeutics. As such there are many established techniques and resources to facilitate rabbit antibody discovery e.g. purification of rabbit antibody, amplification of rabbit V region DNAs, libraries of rabbit Ig germline sequences. In addition, in comparison to mice, the larger blood volume and tissue sizes of rabbits allows greater recovery of B cells from blood and lymphoid tissues. Due to their ability to produce high titres of high-affinity antibody, rabbits have long been utilised as a source of polyclonal antibody for clinical use, notably anti-thymocyte globulin for immunosuppression for organ and bone marrow transplants [237]. However, the non-human origin of rabbit mAbs requires their humanisation for use as therapeutics, in order to increase safety and efficacy by reducing induction of anti-drug antibody in patients. Rabbit mAb humanisation is now well established, with successfully chimerised or humanised rabbit mAbs assessed in human clinical trials, and Brolucizumab, a rabbit-derived humanised scFv, was approved in 2019 in the US for treatment of macular degeneration [46, 238].

As Lagomorphs, rabbit antibody repertoires have the potential to recognise epitopes that are not immunogenic to members of other mammalian orders such as rodents or primates, due to differences in germline sequences and distinct mechanism of generating antibody diversity [239].

Rabbit B cell repertoires therefore may be a source of monoclonal antibodies against epitopes not generated by species previously interrogated for anti-GP mAbs.

Rabbit B cell repertoire development occurs in three stages: neo-natal B cell lymphogenesis, post-natal primary repertoire development and finally, secondary repertoire development due to an immune response. Initial antibody gene rearrangements and the additional post-natal diversification of antibody sequences for generation of the primary repertoire set rabbits apart from other species.

In initial antibody gene rearrangements, rabbits utilise a highly restricted number of the available VH genes, with VH1 being used in ~80% of rearranged sequences [240]. However, this initial restriction is compensated for by other mechanisms, including greater diversity in the light chain sequences. During heavy and light chain recombination, N nucleotide addition at junctions increases diversity of sequences, with light chain CDR3 on average (mean ± s.d.) longer in rabbit sequences (12±2) than mice or humans (9±1) [241].

Like bovines, the development of the rabbit primary repertoire is GALT-dependent and utilises untemplated SHM, however it also uses templated somatic gene conversion (SGC) [238]. SGC utilises homologous recombination to replace portions of V region sequences with tracts of sequences from the many non-utilised V genes. Comparative analysis of SGC in rabbit sequences by Lavinder *et al*. (2014), determined tract lengths (mean ± s.d.) for heavy and light chain sequences of 59±36 and 86±48 nucleotides respectively, with SGC utilised in 23% of the rabbit Ig repertoire, but only 2.5% in humans, and undetectable in mice. The degree of SHM in rabbit antibody sequences is similar to that of humans, and greater than that of mice [242].

Therefore, due to their distinct antibody diversity generating mechanisms, rabbit immune repertoires represent an underutilised resource for generating complementary anti-GP antibodies to those generated from mice and primates that can be readily humanised for clinical use.

3.1.3 Cross-reactive antibodies to *Ebolavirus* GP

When this project began, a small number of broadly reactive mAbs against *Ebolavirus* GPs were beginning to be reported. Since then, as discussed in Section 1.4.3.2, the panel of available antibodies

in the literature has expanded, and broadly reactive mAbs have been shown to be cross-protective in *in vivo* animal models (Table 2). Concomitant with this has been the evidence from human clinical trials that mAb therapies reduced mortality in an EBOV outbreak [45]. Therefore there is a strong rationale for development of a robust pipeline of broadly-reactive and broadly protective monoclonal antibodies that can be developed into cocktails for clinical use in future outbreaks of ED caused by other *Ebolavirus* species, to protect against potential new *Ebolavirus* strains, and to increase efficacy of current therapies.

The primary aim of the research conducted in this chapter was to contribute to this pipeline by generating a panel of new mAbs that could bind to the GP of all species of *Ebolavirus* known to have caused human disease (EBOV, SUDV, BDBV and TAFV). This was achieved by vaccinating rabbits with GP from all four *Ebolavirus* species and B cell culture methods using the high throughput Core Antibody Discovery Platform at UCB.

3.2 Rabbit immunisations and comparison of sequential and mixed antigen immunisation for generation of cross-reactive antibodies against *Ebolavirus* GP

3.2.1 Production of transiently transfected RAB-9 cells

RAB-9 cells (New Zealand White rabbit fibroblast cell line) were transiently transfected with pcDNA3.1(-) plasmids encoding full length EBOV GP (NP_066246.1), SUDV GP (YP_138523.1), BDBV GP (YP_003815435.1) or TAFV GP (YP_003815426.1). These full length sequences include the endogenous viral signal peptide and transmembrane sequences that target the GP to the cell membrane in the context of natural infection. These plasmids have also been used to successfully transfect HEK293F cells with expression confirmed by extracellular staining and flow cytometry, and the same GP sequences used to transduce MDCK SIAT-1 cell lines with expression of GP at the cell surface confirmed in IFA and viral pseudotyping experiments. Concurrently, in this experiment cells were transiently transfected with a non-viral antigen (podoplanin) in a mammalian expression vector as a positive control for the transfection. The day after transfection, cell surface antigen expression was

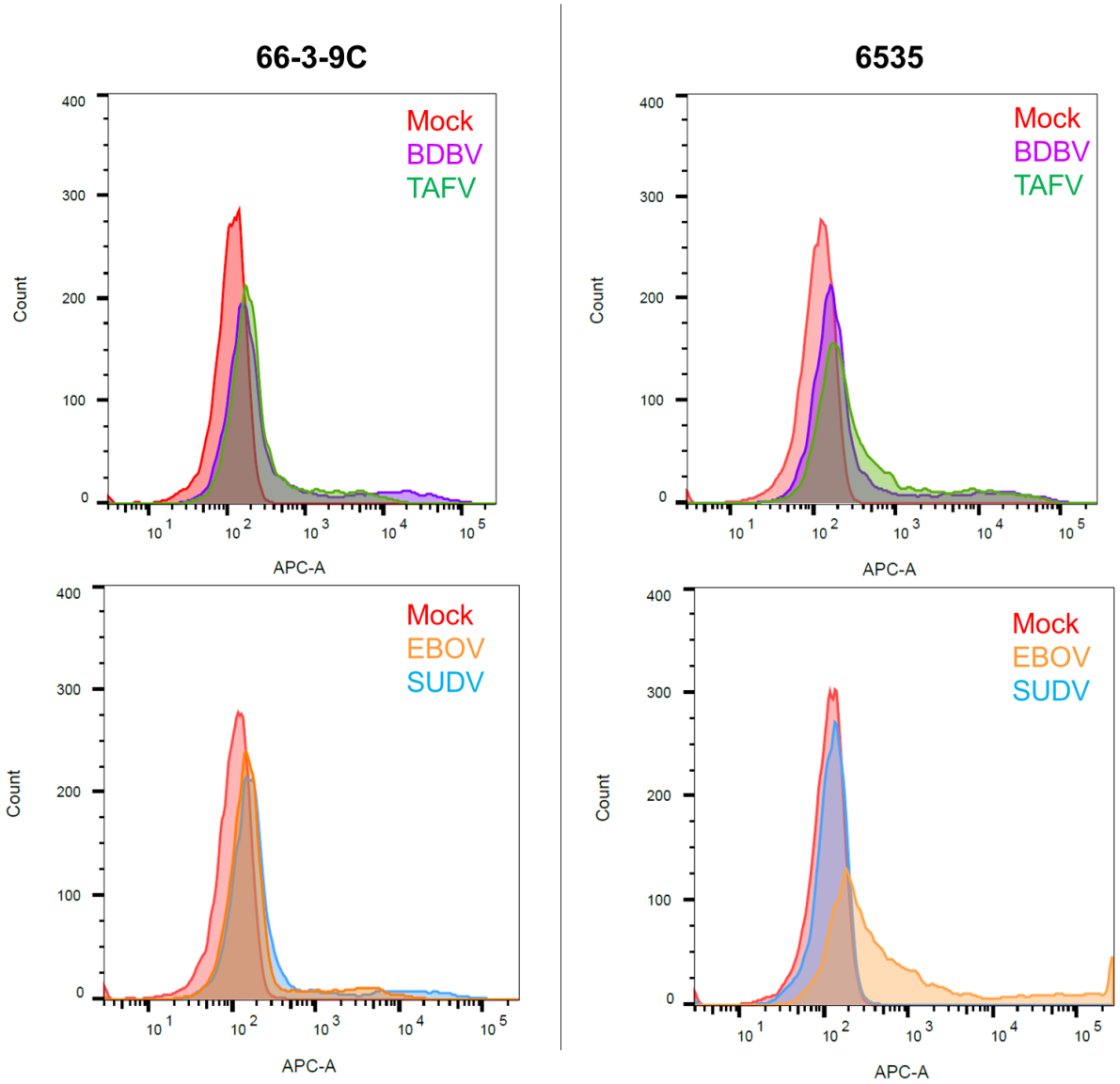


Figure 6. Confirmation of GP expression on transiently transfected RAB-9 cells prior to use in rabbit immunisations. *$5x10^5$ cells transfected with EBOV GP, SUDV GP, BDBV GP, TAFV GP or Podoplanin (mock) were stained with 10 µg/mL of 66-3-9C and 6535 followed by Alexa Fluor-647 conjugated anti-human IgG antibody, or anti-podoplanin secondary antibody (data not shown). Data were collected using BD FACS Canto™ II and BD FACSDiva software, analysis was performed with FlowJo3 software.*

confirmed by flow cytometry (

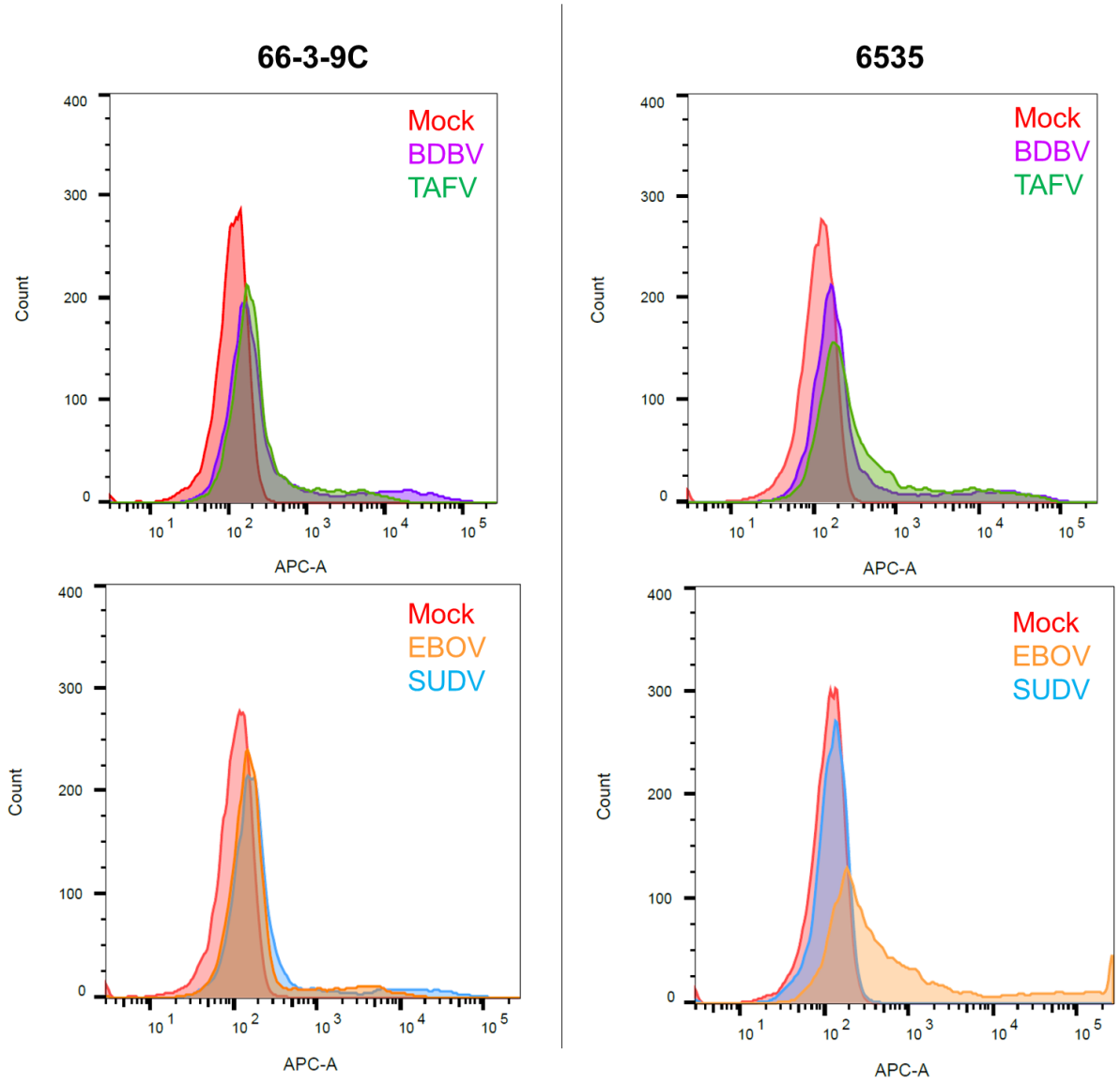


Figure *6*), and remaining cells were frozen at -80°C until day of vaccination.

Expression of GP antigens on the surface of RAB-9 cells was confirmed by indirect immunofluorescence using human anti-GP mAbs 66-3-9C and 6535 and Alexa Fluor 647-conjugated anti-human IgG (Jackson ImmunoResearch, 109-606-170) [3]. Expression of podoplanin was confirmed by direct immunofluorescence using anti-podoplanin PE conjugate (BioLegend, 337004).

Cells transfected with any GP antigen, but not podoplanin transfected cells, were specifically recognised by human mAb 66-3-9C. 66-3-9C is a known broadly-reactive anti-GP mAb with ~1 nM binding affinity for trimeric soluble GP [3]. 6535 has limited cross-reactivity allowing discrimination between the antigens. This mAb displayed a binding profile consistent with previously tested binding

to MDCK-SIAT cell lines expressing *Ebolavirus* GPs; strong binding to EBOV GP, with cross-reactivity to TAFV and BDBV GP only [3]. Podoplanin transfected cells were specifically stained by anti-podoplanin mAb only. Cells were not sorted on the basis of antigen expression before being used to vaccinate rabbits.

3.2.2 Immunogen formulation and immunisation regimes

Three immunisation regimes were selected to cover a range of strategies to generate broadly-reactive monoclonal antibodies (Figure 7). Three New Zealand White rabbits received four vaccinations at two week intervals. $1x10^7$ GP-transfected RAB-9 cells (see Section 3.2.1) per dose were delivered via the subcutaneous route. Lymph nodes and spleen were harvested two weeks after final vaccination for B cell culture. At each vaccination, rabbit 378-1 received a total dose of $1x10^7$ cells composed of equal amounts of cells transfected with one of each of the four GP antigens. Rabbits 378-2 and 378-3 were sequentially immunised with cells expressing different GP antigens. With only two rabbits it was not possible to explore all permutations of the order of GP antigens. SUDV and EBOV GP are the least homologous of the four antigens, therefore the order of vaccination with these two GPs may have the greatest impact on boosting of a cross-reactive response. Consequently, the sequence for vaccinating 378-2 and 378-3 differs in having SUDV GP as the first or last boosting antigen.

Transiently transfected cells were thawed on day of vaccination, prepared in 500 µL PBS and kept on ice until vaccination. Following UCB's protocols for rabbit immunisations, Complete Freund's adjuvant (CFA) was given via subcutaneous route at a second site with the first immunisation only, subsequent immunisations were given without adjuvant. All animal procedures and husbandry were carried out by staff on site at UCB, Slough.

	Vaccination Regime				
Rabbit	Day 0	Day 14	Day 28	Day 42	Day 56
378-1	EBOV, SUDV, BDBV, TAFV	EBOV, SUDV, BDBV, TAFV	EBOV, SUDV, BDBV, TAFV	EBOV, SUDV, BDBV, TAFV	
378-2	EBOV	SUDV	BDBV	TAFV	
378-3	EBOV	BDBV	TAFV	SUDV	
Adjuvant	CFA	None	None	None	
Sampling	Serum	Serum	Serum	Serum	**B cells**

Figure 7. Vaccination and sampling regimen for rabbits 378-1, 378-2 and 378-3 immunised with RAB-9 cells expressing Ebolavirus GP. Complete Freund's Adjuvant (CFA).

3.2.3 Rabbit serum IgG responses to *Ebolavirus* Glycoproteins

3.2.3.1 Binding to full length GP expressed on HEK cells

Expi293F HEK cells were transiently transfected with the same GP-encoding plasmids used to transfect RAB-9 cells in 3.2.1 and incubated in suspension culture for two days. Serum samples collected on the day of each vaccination and two weeks post-final vaccination, were titrated in PBS, and incubated with transfected cells. Bound serum IgG was detected using anti-rabbit IgG PE-conjugate (Jackson ImmunoResearch, 111-116-144). Antigen expression was confirmed by staining with cross-reactive human mAbs 66-3-9C, 66-4-C12, 6535 and 6451 and detection with anti-human IgG PE conjugate (Jackson ImmunoResearch, 109-116-170) or direct immunofluorescence with anti-podoplanin PE conjugate (BioLegend, 337004).

Serum from rabbit 378-1, which received mixed antigen immunisations, shows titratable response to each GP with boosting of response clear two weeks after second immunisation (Figure 8A). Maximum response to each antigen detected in this assay was achieved two weeks after second immunisation and antibody titres were maintained until day 56 when B cells were harvested for culture.

Rabbits 378-2 and 378-3 were sequentially immunised with each of the four antigens in a different order. Both rabbits show at least a low concentration of serum IgG that binds to each of the four GPs (Figure 8B, C). In general, the response to BDBV GP and TAFV GP is much lower than seen in 378-1 serum.

3.2.3.2 Neutralisation in S-FLU pseudotype virus assay

At the highest concentrations tested, polyclonal serum from RAB-9-GP immunised rabbits was unable to neutralise single-cycle S-FLU viruses pseudotyped with either EBOV GP (NP_066246.1) (Figure 9A) or SUDV GP (YP_138523.1) (Figure 9B) *in vitro* [219]. This pseudotype system will be discussed in more detail in Section 4.1.1. Infected cells express GFP and reduction in GFP signal indicates neutralisation of virus. All samples were assessed in the assay at the same time and averages of maximum (no serum added to well) and minimum (no virus added to well) signal for each plate in the assay are shown and were consistent across the assay. High concentrations of serum appeared to enhance the fluorescent signal which could indicate enhancement of infection, however as this is observed when naïve serum samples (Day 0) were tested, this cannot be attributed to the GP-specific antibody and may be an artefact of the assay. CA45 is a broadly neutralising anti-GP mAb and was included as a positive control, neutralisation of pseudotype viruses titrates as expected with dilution of CA45 [190]. Assay was conducted with diluted serum, but purification of IgG for use in the assay may improve the result.

3.3 Generation of novel rabbit monoclonal antibody panel

Considering the presence of *Ebolavirus* GP cross-reactive polyclonal IgG in the serum (Figure 8), and to preclude the chance of false negative results in the S-FLU neutralisation assay (Figure 9), the rabbit immune material was used for an antibody discovery campaign. Consistent with the 3Rs principles from National Centre for the Replacement, Refinement and Reduction of Animals in Research, this decision also avoided use of further animals.

3.3.1 B cell culture

Splenocytes and lymph node samples were harvested 14 days post final vaccination. On the day of harvest, cells were seeded in barcoded 96 well cell culture plates at a density of 5000 cells per well

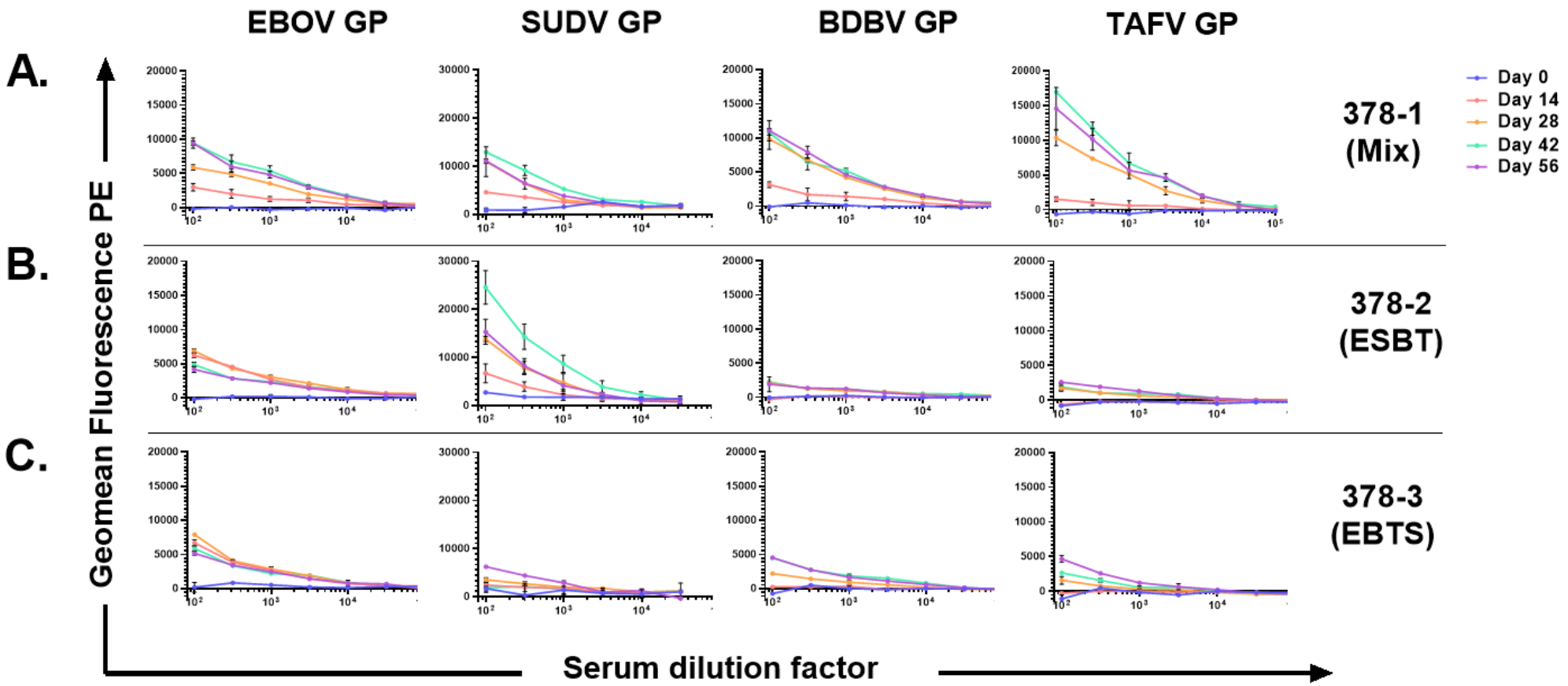


Figure 8. Serum IgG responses from rabbits immunised with Ebolavirus GPs expressed on RAB-9 cells to full length Ebolavirus GPs on transiently transfected HEK cells. *Fluorescence expressed as signal from sample binding to GP expressing cells minus signal from sample binding to cells transiently transfected with podoplanin, a non-GP antigen. 10 µL of diluted serum was incubated with 1x10^4 cells per well in 384-well plates. Bound IgG was detected using 10 µL of 1:1000 dilution of anti-rabbit IgG PE conjugate. Single cell fluorescence measured using iQUE flow cytometer (Intellicyt) and data analysed using Forcyt (Intellicyt) and GraphPad Prism 7.05. Mean and range of technical duplicates shown. BL2 with an excitation wavelength of 488 nm and detection channel of 572/28 was used.*

.

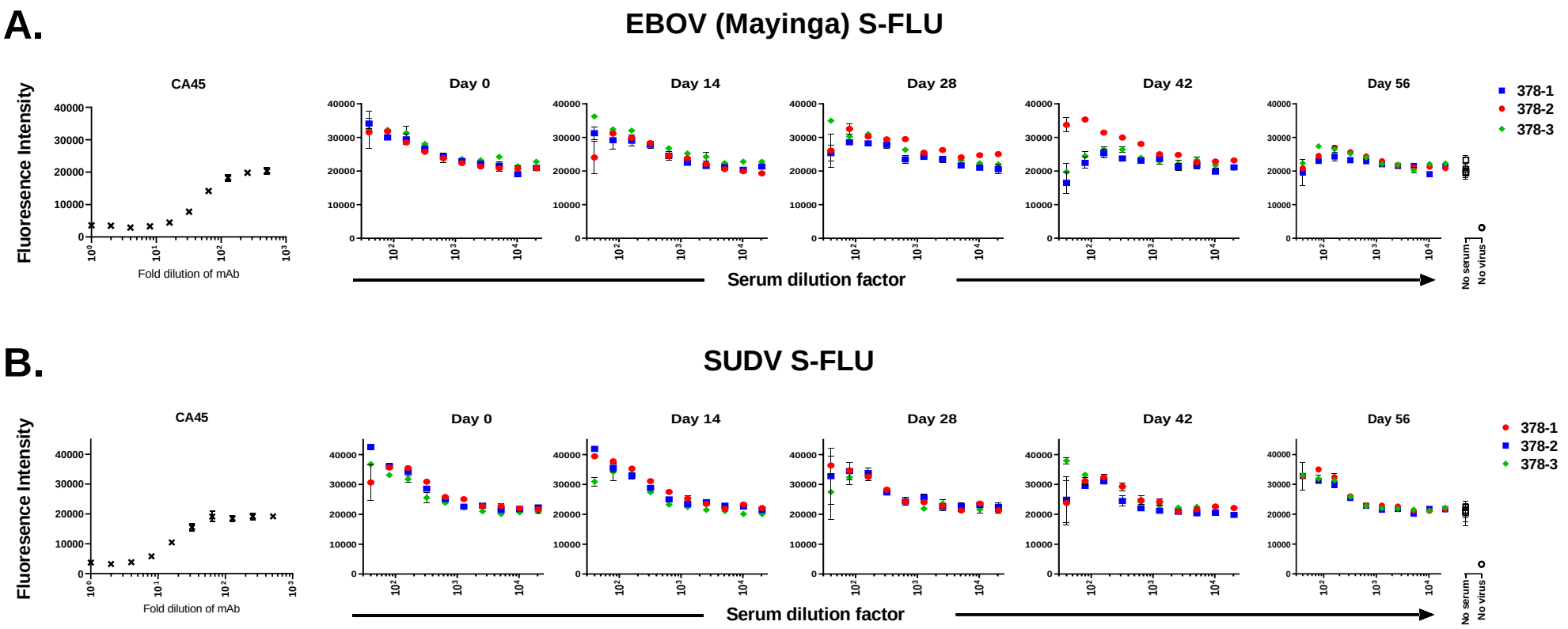


Figure 9. In vitro neutralisation of EBOV S-FLU (NP_066246.1) and SUDV S-FLU (YP_138523.1) with polyclonal serum from rabbits 378-1, 378-2 and 378-3 vaccinated with RAB-9 cells expressing Ebolavirus GPs. *Serum was collected on day of each vaccination prior to immunisation. Final serum dilution series in assay during incubation with virus began at 40-fold. CA45 was incubated with virus at concentrations starting at 25 µg/mL with a 2-fold serial dilution to 48 ng/mL in duplicate. Mean and range of fluorescence intensity shown. After removing virus and fixing cells, fluorescence intensity was measured using a Clariostar Plate Reader (BMG, Labtech) (Excitation/Detection filters: 483-8/515-8, Gain: 3000). Mean and range of 8 wells per assay plate either without serum added (maximum signal) or without virus added (minimum signal) are shown for each assay.*

using a method similar to that described by Tickle *et al*., with seeding density aiming to achieve less than a single antigen-specific antibody secreting cell per well. Culture conditions replicate the extracellular milieu required to maintain IgG secretion by B cells [216]. 28,800 culture wells were seeded. After 6 days culture, screening was performed using UCB Core Antibody Discovery Platform (CADP) Filling Module.

3.3.2 Screening of B cell culture supernatants for anti-GP IgG using fluorescence cell-based assays

Screening assays relied heavily on fluorescence based binding assays using *Ebolavirus* GPs expressed on mammalian cell lines and small volumes of B cell culture supernatant. Primary screening served to select several hundred possible GP binders and exclude the vast majority of non-GP binding IgG from secondary screening. Secondary screening assays were designed to identify IgG that were cross-reactive to multiple GP species. Single antigen-specific B cells were then isolated from these wells using the fluorescent foci method followed by variable region gene recovery.

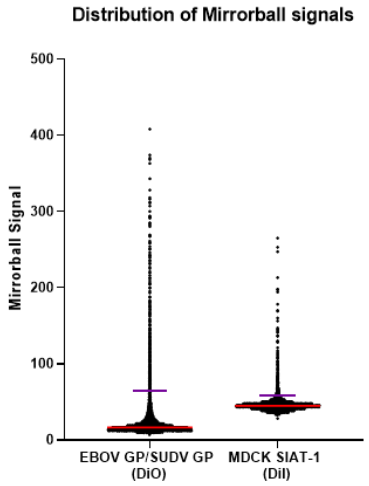
3.3.2.1 Primary screening using laser scanning fluorescence cytometry

Approximately 12,000 B cell culture supernatants were assayed over a two day period (Figure 10). Due to technical problems with the robotic handling of plates, only cells from rabbit 378-1 were assayed in high numbers.

Primary screening was conducted as a homogenous two colour counter-screening assay using cells expressing antigen and the parental cell line (as a negative control) fluorescently labelled using different lipophilic cell membrane dyes with distinct absorption and emission spectra. An equal number of MDCK SIAT-1 cells transduced to express EBOV GP (NP_066246.1) and SUDV GP (YP_138523.1) were stained with DiO, and parental MDCK SIAT-1 cells were stained with DiI (Invitrogen, V22889). The stained cells were mixed and incubated for 2 hours with 10 µL of B cell culture supernatant and anti-rabbit IgG Alexa Fluor 647 conjugated antibody (Jackson ImmunoResearch, 111-606-046) in 384-well assay plates. Fluorescence was then measured using a Mirrorball laser scanning microplate cytometer (SPT Labtech, formerly TPP Labtech). IgG bound to

Primary Screening of B cell culture supernatant IgG with Mirrorball Fluoresence Assay

A.



B.

Immune sample	Culture wells screened	MDCK SIAT-1 +	GP +	Hits (MDCK SIAT-1-/GP+)
378-1				
Splenocytes	4803	0.64 %	0.44 %	0.27 %
Lymph Node	4815	0.61 %	3.1 %	2.79 %
378-2				
Lymph Node	2408	0.22 %	0.96 %	0.90 %
Total	**12026**	**1.47 %**	**4.49 %**	**3.97 %**

C.

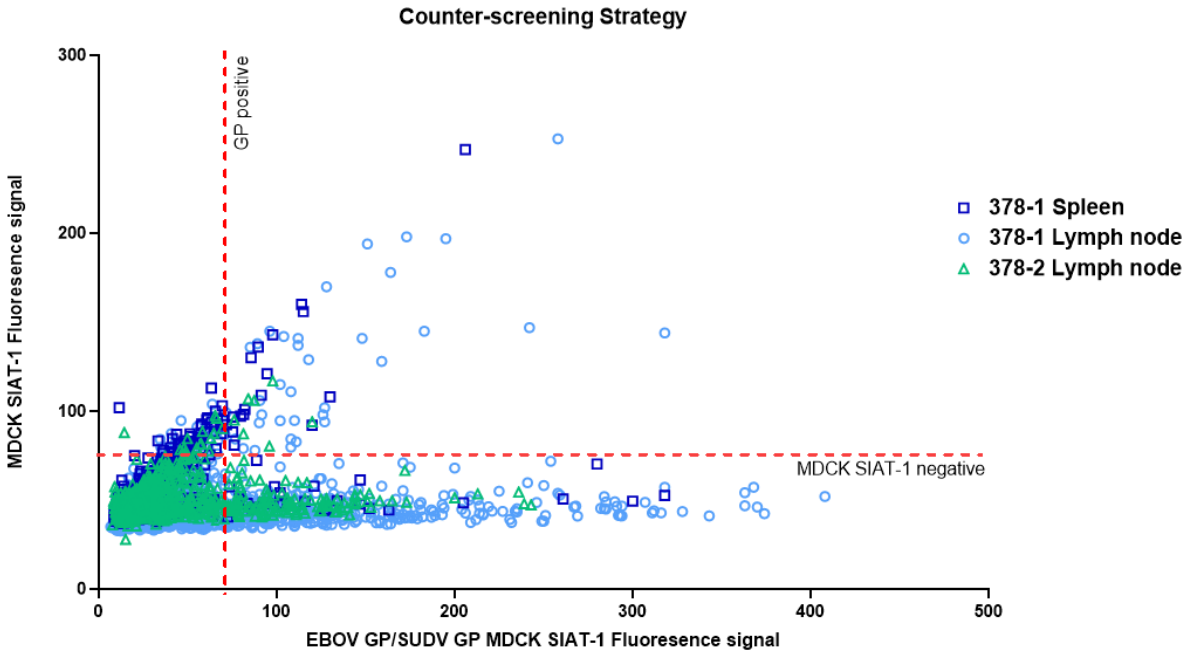


Figure 10. Primary screen of B cell culture supernatant IgG for GP binding using Mirrorball fluorescence assay. A. Distribution of fluorescent signal from secondary antibody in DiO and DiI channels. *Median and 95th or 98th percentiles indicated by red and purple lines respectively.* ***B. Summary of screening results using threshold values >70.3 for GP binding and >75 for MDCK SIAT-1 binding.*** *Percentages indicated are of all samples screened. C. Counter-screening strategy based on fluorescent signals associated with DiI-labelled MDCK SIAT-1 cells and DiO-labelled EBOV GP or SUDV GP expressing MDCK SIAT-1 cells. Thresholds indicated by red dotted lines. Hits defined by lower right hand quadrant.*

labelled cells was identified by localisation in the well of signal from fluorophore conjugated to secondary antibody with signal from either channel used to detect labelled cells.

Counts per well was used as an indication of assay quality, with counts remaining consistent between plates despite the period of time over which screening was conducted. Range of signals was similar between cell cultures from the three immune samples screened.

479 culture wells were selected for further screening. Supernatant was removed from selected wells for secondary screening assays and cells frozen *in situ* with freeze media at -80°C until individual B cell recovery. As internal positive and negative controls could not be included in this assay due to the high throughput setting, thresholds for positive binding were defined on the basis of distribution of fluorescence signals (Figure 10A), and hits defined as MDCK SIAT-1 negative/GP positive. All hits were taken from wells with GP binding signal above the 95th percentile for all wells screened (threshold value: 70.3), and MDCK SIAT-1 binding signal below the 98th percentile for all wells screened (threshold alue: 75) (Figure 10B, C). This resulted in ~4% of all wells screened carried forward. These thresholds were chosen to reduce selection of IgG that bound MDCK SIAT-1 cells or to GP at background noise levels, but to be generous enough to allow for supernatants with low concentrations of GP-specific IgG to be carried forward.

3.3.2.2 Secondary screening using flow cytometry and fluorometric microvolume assay

Secondary screening assays conducted offline from the CADP robotic handling allowed deconvolution of GP binding IgG into likely single and cross-reactive wells. All assays were conducted in a 384-well format with 10 µL of B cell supernatant. 479 IgG culture supernatants were screened in this manner, although some were not subjected to the entire screen, dependent on available volume.

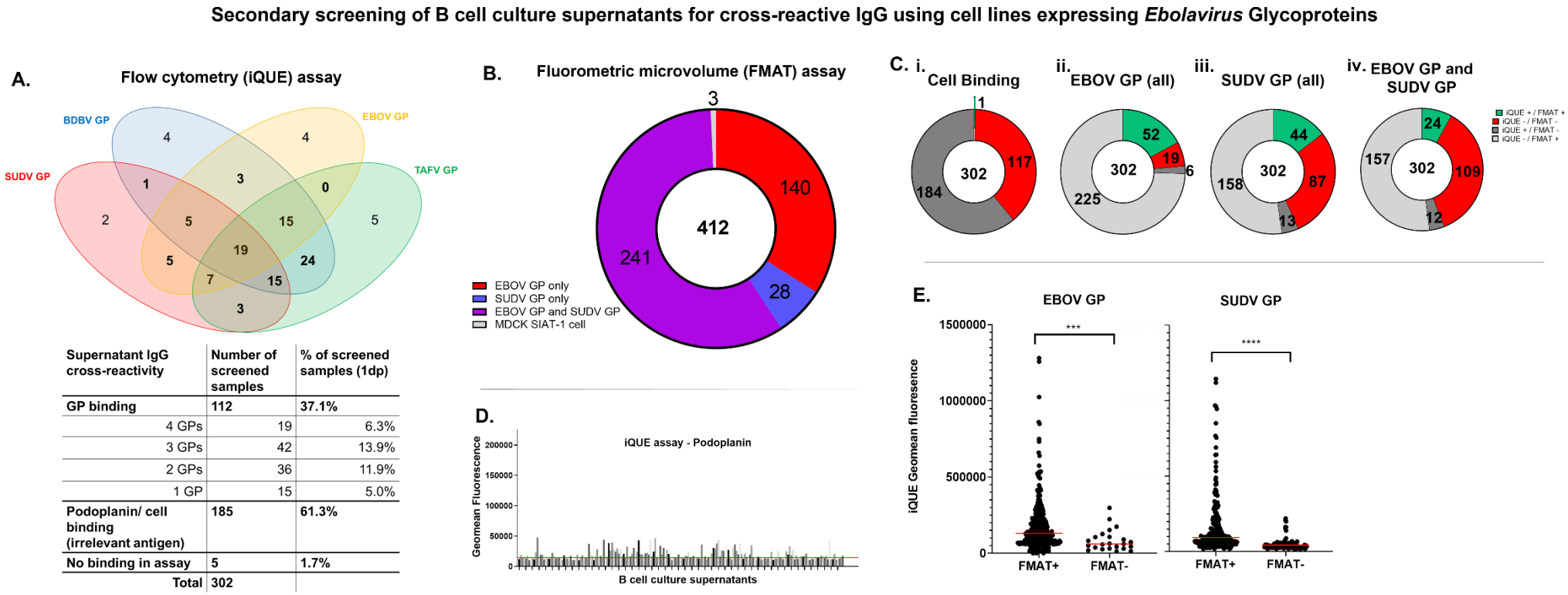
Flow cytometric assays were conducted using HEK cells transiently transfected with GP from EBOV, SUDV, BDBV and TAFV (NP_066246.1, YP_138523.1, YP_003815435.1, YP_003815426.1) or an unrelated antigen (podoplanin, selected due to availability of suitable DNA and detection reagents). Antigen expression was confirmed prior to use of cells on the day of the assay. 302 supernatants were screened against all five antigens. This assay identified many primary screen hits as potentially binding

to non-GP cell antigens (61.3%), but of those that bound GP and did not bind podoplanin transfected cells, 86.6% were identified as cross-reactive (Figure 11A), with 19 supernatants identified as containing IgG that could bind to all four GP antigens. As the primary screening only included EBOV and SUDV GP antigens the identification of IgG that appears to bind only BDBV GP or only TAFV GP is possibly due to a lower amount of EBOV GP and SUDV GP antigen on the surface of transiently transfected HEK cells compared to the stably transduced MDCK SIAT-1 cells in the primary screen; this could result in a loss of signal in relation to EBOV GP and/or SUDV GP between the two screening assays for supernatants with lower IgG concentrations or IgG with lower affinity for either of these two antigens. Wells positive for binding were defined by a signal >3x the standard deviation of 16 wells per assay plate with no culture supernatant (background fluorescence).

Fluorometric microvolume (FMAT) assays were conducted with MDCK SIAT-1 cell lines used in primary screening (3.3.2.1) without direct labelling of cells. 412 culture supernatants were screened (Figure 11B). This assay identified only 3 wells that bound the parental cell line, with more than 58% of the remaining samples positive for binding both EBOV and SUDV GP expressing cells. FMAT data are gated based on negative wells run within the assay and each well image is visually checked to ensure fluorescence is due to cell binding, not to autofluorescence from cell debris.

A comparison of normalised FMAT signal (0, negative; 1, positive) with signal in the iQUE assay between all samples tested in both assays against SUDV and EBOV GP expressing cells indicates a significant difference in the median fluorescence detected in the iQUE assay between FMAT positive and negative samples for both EBOV and SUDV GP binding assays (unpaired, non-parametric t test, Mann Whitney). This indicates that higher signals in the iQUE do correlate with positive binding in the FMAT assay (Figure 11E).

Secondary screening results were used to select 45 B cell culture wells for recovery of individual B cells and variable region sequences. Samples that were negative for cell binding and positive for binding multiple GPs in both assays were prioritised.



Supernatant IgG cross-reactivity	Number of screened samples	% of screened samples (1dp)
GP binding	**112**	**37.1%**
4 GPs	19	6.3%
3 GPs	42	13.9%
2 GPs	36	11.9%
1 GP	15	5.0%
Podoplanin/ cell binding (irrelevant antigen)	**185**	**61.3%**
No binding in assay	**5**	**1.7%**
Total	**302**	

Figure 11. Fluorescence based screens identifying cross-reactive IgG in selected B cell culture supernatants. A. Cross-reactivity of B cell culture supernatant IgG to GPs expressed on transiently transfected HEK cells. *Single cell fluorescence measured using iQUE flow cytometer (Intellicyt) and data gated and analysed using Forcyt (Intellicyt).* ***B. Cross-reactivity of B cell culture supernatant IgG to GPs expressed on MDCK SIAT-1 cells.*** *Fluorescence associated with cells measured in a homogenous fluorescence microvolume assay (FMAT) performed with Applied Biosystems 8200 cellular detection system.* ***C. Comparison of results of 302 supernatants assessed in both assays for binding to parental cell line or cells transfected with non-GP antigen (cell binding), EBOV GP and SUDV GP****. + or - indicates above or below threshold defined for indicated assay respectively. Green and red sectors indicate agreement between the two assays, grey sectors indicate samples for which assay results differed.* ***D. Examples of B cell culture supernatant responses to podoplanin transfected cells in the iQUE assay****. Red line indicates background threshold.* ***E. iQUE signals for supernatants positive and negative for binding to EBOV and SUDV GPs in FMAT assay.*** *Median in red. Non-parametric t test, Mann-Whitney (p= 0.0004 for EBOV GP, p<0.0001 for SUDV GP).*

3.3.3 Fluorescent foci picking, PCRs, cloning and expression

Individual B cells were manually recovered using the fluorescent foci method (US Patent 7993864/ Europe EP1570267B1 [218]. In brief, B cells frozen *in situ* in culture plates are thawed and incubated on a microscope slide with antigen-coated beads and a fluorescent FITC-conjugated anti-rabbit IgG antibody. Production of IgG by recovered B cells results in fluorescent halos of beads around individual cells secreting antigen-specific IgG visible using a fluorescent microscope. These cells are manually picked using a micromanipulator and capillary needle and immediately frozen on dry ice for preservation of mRNA for recovery of cognate heavy and light chain sequence pairings of antigen-specific IgG.

Recoveries from 45 B cell culture wells selected after secondary screening are shown in Figure 12. In this case, monomeric EBOV GP and SUDV GP were conjugated to two different sized beads. Visual discrimination of beads involved in fluorescent foci by size gave initial indications of cross-reactivity of IgG. Good recovery and binding to beads was seen with more than two thirds of B cells recovered from freezing producing enough IgG for foci formation.

Reverse transcription and nested PCRs using degenerate primers allowed recovery of variable region DNA sequences from single B cells. An additional PCR to incorporate V region sequences into linear DNA transcriptional cassettes (TAP) was conducted and resultant DNA used to transiently transfect HEK 293F cells in 1mL culture volumes. Of these almost all expressed IgG as detected via sandwich ELISA. Of TAP cultures expressing IgG, two thirds showed continued binding to either EBOV GP, SUDV or both in ELISA. All variable regions were further sub-cloned into expression vectors for production of recombinant antibody. Thirty matched heavy and light chain pairings were successfully cloned and used to express IgG by transient transfection of HEK 293F cells. Expression of recombinant IgG was confirmed using sandwich ELISA and protein G binding.

Confirmation of binding of recombinant IgG from expression vectors to trimeric GPs expressed on the surface of cells, antibody sequence analysis and selection of final cross-reactive antibody panel are discussed in sections 3.4 and 3.5.

Foci recovered	**33/45**
TAP IgG expressed	**31/33**
EBOV GP binding only	5/31
SUDV GP binding only	4/31
EBOV GP and SUDV GP binding	12/31
No GP binding	10/31
V region DNA cloned into expression vectors	**30/33**
Paired heavy and light chain, sequencing confirmed	30/33
Heavy chain	31/33
Light chain	
Recombinant IgG expressed	**30/30**
GP binding confirmed	**19/30**

Figure 12. Recoveries of paired heavy and light chain V region sequences for recombinant antibody expression from single B cells.

3.4 Confirmation of binding to GPs and cross-reactivity to multiple species of *Ebolavirus*

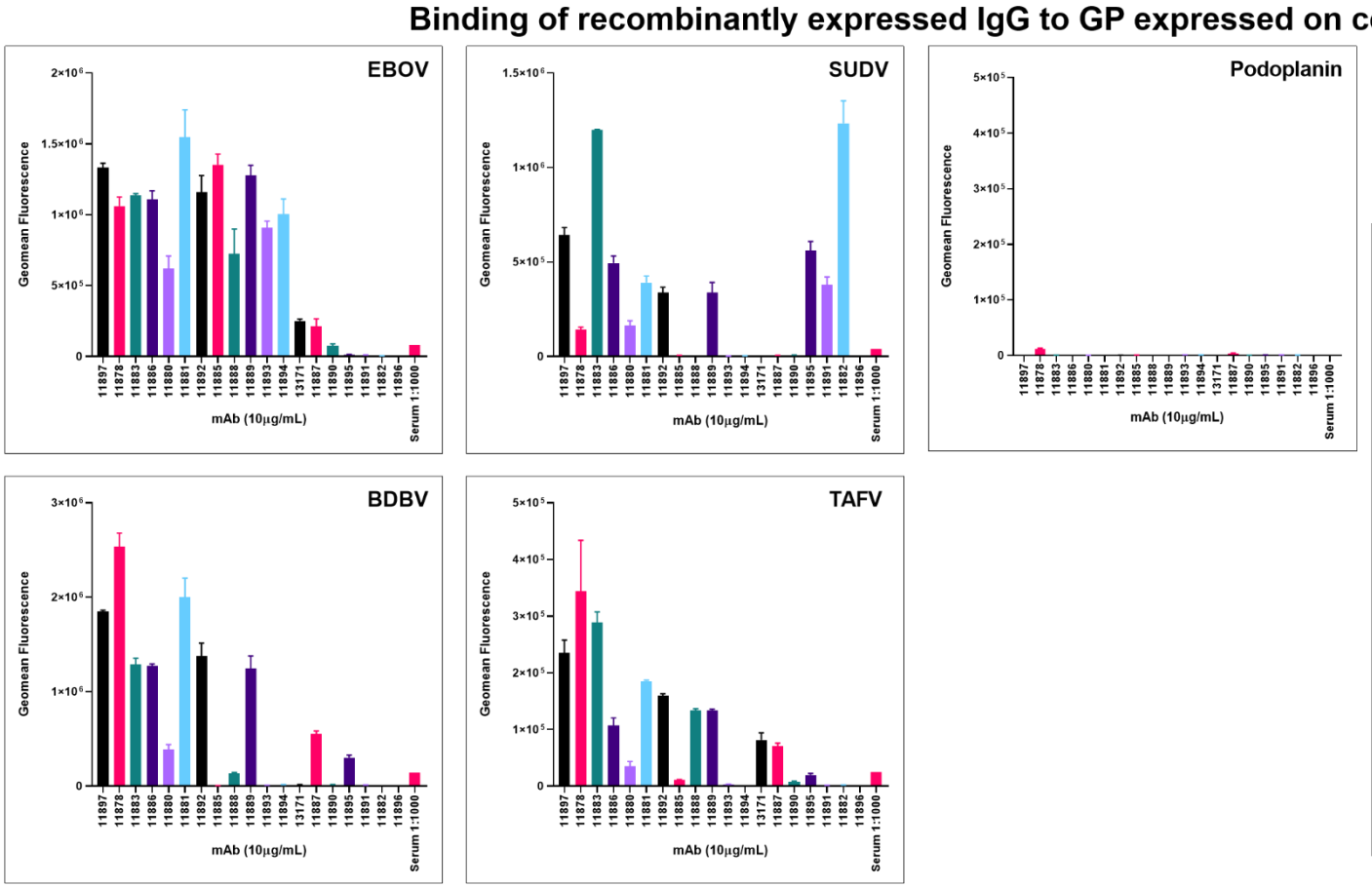
3.4.1 Rabbit mAb binding to GPs from species known to cause human disease

GP binding and cross-reactivity of the recombinant antibody panel was confirmed using both GP expressed on transiently transfected HEK cells, and stably expressed GPs on MDCK SIAT-1 cells. Binding to GPs from species of *Ebolaviruses* known to cause human disease (EBOV, SUDV, BDBV, TAFV) was assessed. A summary of multiple experiments is collated in Figure 13 for mAbs that bound GP. 19/30 mAbs produced showed binding to at least one GP in either assay. Figure 13A shows representative data using flow cytometry to describe binding of panel of mAbs at a single concentration of 10 µg/mL to GPs on HEK cells. In both assays, antibody was also titrated to confirm binding (data not shown). 13/19 recombinantly expressed mAbs that bound GP, showed reactivity to GP from more than one species of *Ebolavirus*; with 11/19 consistently binding three or more different GPs.

3.4.2 Human mAb cross-reactivity to TAFV GP

Prior to this project, the Oxford/UCB panel of EBOV GP human mAbs was generated. Although this panel was generated by selecting for EBOV GP binding, many mAbs were cross-reactive to SUDV GP and BDBV GP; however, cross-reactivity to TAFV GP was not tested [3]. At the start of this project, in collaboration with Pramila Rijal, I therefore generated a TAFV GP expressing MDCK-SIAT-1 cell line in order to test the Oxford/UCB panel for binding to TAFV GP in an immunofluorescence assay (IFA) (Figure 14). The IFA was conducted twice and 26 mAbs showed consistent binding to TAFV GP. 66-4-C12 was successfully used to stain for transduced cells when creating the cell line and was therefore used as a positive control. Z3B2 is a non-GP mAb. Determining the wider cross-reactivity of the Oxford/UCB panel allows a better comparison of the degree of cross-reactivity of the humAb panel and the broadly reactive rabbit mAb panel in this thesis.

A.



B.

	HEK					MDCK SIAT-1		
	EBOV	SUDV	BDBV	TAFV	Podoplanin	EBOV	SUDV	Untransduced
11897	+	+	+	+	-	+	+	-
11878	+	+	+	+	-	+	+	-
11883	+	+	+	+	-	+	+	-
11886	+	+	+	+	-	+	+	-
11880	+	+	+	+	-	+	+	-
11881	+	+	+	+	-	+	+	-
11892	+	+	+	+	-	+	+	-
11889	+	+	+	+	-	+	+	-
11888	+	-	+	+	-	+	+	-
11895	-	+	+	+	-	-	+	-
11887	+	-	+	+	-	+	-	-
11896	-	-	-	-	-	+	+	-
13171	+	-	-	+	-	-	-	-
11893	+	-	-	-	-	+	-	-
11894	+	-	-	-	-	+	-	-
11885	+	-	-	-	-	+	-	-
11890	(+)	-	-	-	-	(+)	-	-
11891	-	+	-	-	-	-	+	-
11882	-	+	-	-	-	-	+	-

Figure 13. Binding of recombinantly expressed IgG to GP expressed on cells. A. Purified rabbit monoclonal antibodies binding to antigens transiently expressed on HEK293F cells at a single concentration using an iQUE flow cytometer*. Rabbit IgG detected using anti-rabbit IgG Alexa Fluor-647 conjugate. Samples gated for singlets and geomean fluorescence calculated using Forcyt software. Background calculated from mean of 16 wells and subtracted from each result. Mean and range of technical duplicates shown. Podoplanin is a non-ebolavirus antigen.* ***B. Summary of binding to different antigens by rabbit mAbs across repeats of iQUE assay including titration of mAb concentration, as well as binding of recombinant mAb cell culture supernatant to adherent cells stably expressing EBOV and SUDV in plate-based fluorescence assay using FMAT****. Only mAbs showing reactivity to any GP are shown. mAbs highlighted in red text indicate samples where the two assay set ups differ in result. + binding; - no binding. 11890 displays weak binding to EBOV GP.*

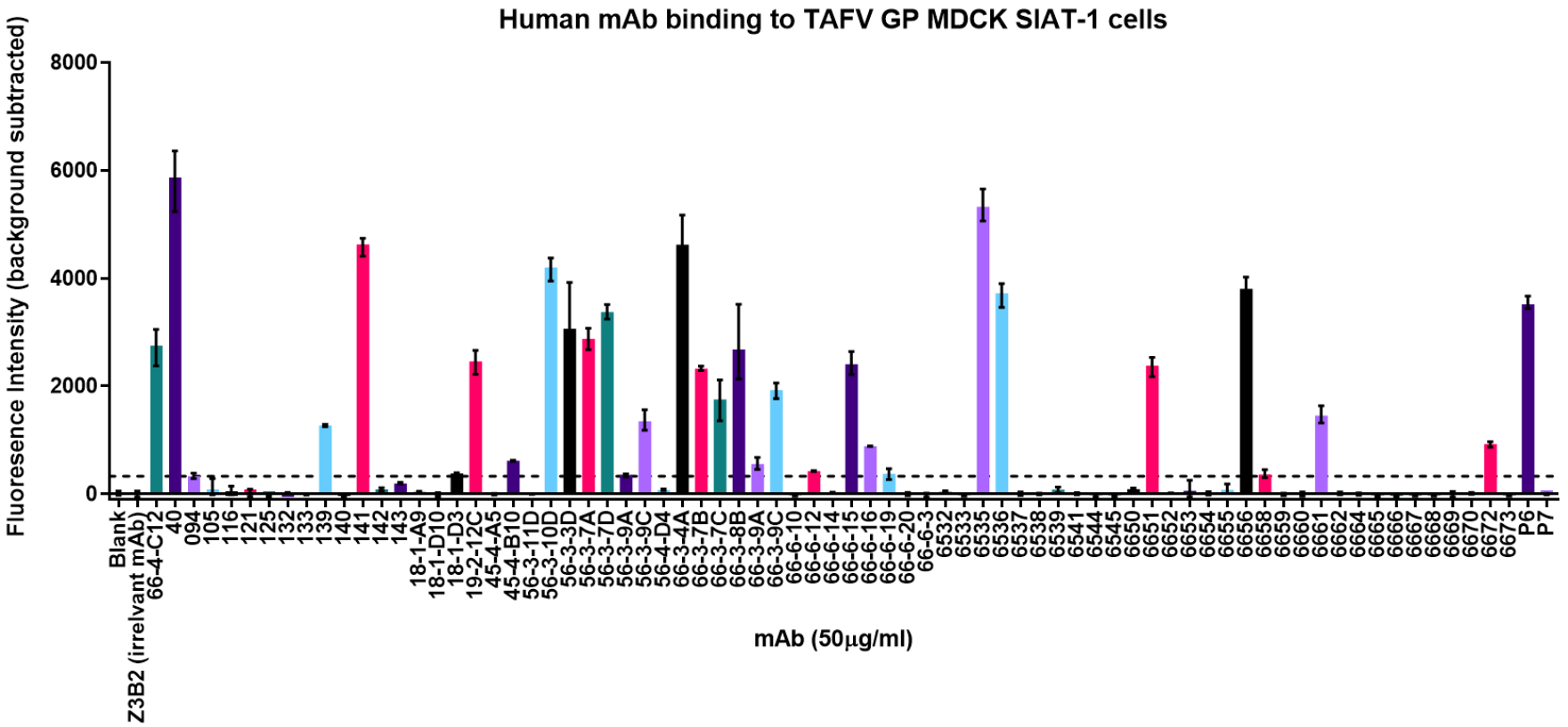


Figure 14. Binding of panel of anti-GP human monoclonal antibodies to TAFV GP expressed on MDCK SIAT-1 cells*. Dotted line represents threshold for classification as TAFV binding. Assay conducted in triplicate with 50μg/mL of mAb. Mean and range of triplicates within assay shown. Human IgG detected using anti-human IgG AlexaFluor-647 conjugate, and fluorescence measured using Clariostar plate reader. Background calculated from mean of triplicate wells and subtracted from each result.*

3.5 Variable region sequences

Once cloned into expression vectors, heavy and light chain variable regions were sequenced and sequences of confirmed GP antibodies analysed. Antibody 11880 variable region differs from 11886 at a two bases in the light chain (LC) sequence resulting in a two amino acid change in framework region 1 (FR1). These antibodies were cloned from two B cells picked from the same B cell culture well. Similarly, antibodies 11881 and 11892 differ only in 2 SNPs, one in the heavy chain is silent and the other in LC FR1. These antibodies were derived from separate B cell culture wells. For both pairs of mAbs, the sites in LC FR1 are within the region where the primers for V region amplification anneal, therefore differences likely arose due to amplification by degenerate primers and are commonly seen in these locations. As such, 11886 and 11892, but not 11880 or 11881 were included in a panel of six broadly-reactive mAbs selected for detailed characterisation. Sequences of the GP mAbs are listed in 7.1.4 Appendix 2.

Germline gene usage and degree of mutation as shown in Figure 15 were identified by comparison to New Zealand White rabbit (NZW, *Oryctolagus cuniculus)* germline sequences by James Snowden, UCB. 11881 sequences were not included in this analysis due to its similarity to 11892. IGHV1S69*01 is used in 7/18 antibodies (Figure 15), and IGHV1S45 and IGHV1S40 contribute collectively to the gene usage in 8/18 mAbs in this study. The J4 J_H germline sequence is utilised in 12/19 sequences. IGKV1S10 and IGKV1S36 were used in 3/18 and 5/18 light chain sequences respectively and all Jκ sequences arose from the same germline sequence.

			Germline V gene usage		Percent identity to germline nucleotide sequence		Number of amino acid changes from germline		Germline D gene usage	Germline J gene usage		CDR3 length (amino acids)		Cross-reactivity
ID	Animal	Tissue	V_H	V_K	V_H	V_K	V_H	V_K	D_H	J_H	J_K	H	K	
11886	**378-1**	**LN**	**IGHV1S69*01**	**IGKV1S36*01**	**94.10%**	**93.70%**	**11**	**12**		**IGHJ2*01**	**IGKJ1-2*01**	**16**	**13**	**EBOV, SUDV, BDBV, TAFV**
11880	378-1	LN	IGHV1S69*01	IGKV1S36*01	94.10%	93.40%	11	13		IGHJ2*01	IGKJ1-2*01	16	13	
11889	**378-1**	**LN**	**IGHV1S40*01**	**IGKV1S36*01**	**99%**	**92.70%**	**2**	**14**		**IGHJ4*01**	**IGKJ1-2*01**	**18**	**13**	
11883	**378-1**	**LN**	**IGHV1S45*01**	**IGKV1S49*01**	**95%**	**91.70%**	**12**	**14**		**IGHJ4*01**	**IGKJ1-2*01**	**14**	**13**	
11892	**378-1**	**LN**	**IGHV1S69*01**	**IGKV1S36*01**	**94.40%**	**91.70%**	**10**	**16**		**IGHJ4*01**	**IGKJ1-2*01**	**17**	**12**	
11897	**378-1**	**LN**	**IGHV1S69*01**	**IGKV1S10*01**	**94%**	**90.20%**	**11**	**15**		**IGHJ2*01**	**IGKJ1-2*01**	**14**	**13**	
11878	**378-1**	**LN**	**IGHV1S69*01**	**IGKV1S10*01**	**94%**	**90.10%**	**9**	**20**		**IGHJ2*01**	**IGKJ1-2*01**	**6**	**10**	
11888	378-1	SP	IGHV1S45*01	IGKV1S32*01	99.70%	89.30%	2	20		IGHJ4*01	IGKJ1-2*01	14	12	EBOV, BDBV, TAFV
11887	378-1	LN	IGHV1S47*01	IGKV1S19*01	82.80%	94.90%	30	10		IGHJ4*01	IGKJ1-2*01	19	11	
11879	378-1	LN	IGHV1S69*01	IGKV1S7*01	92.70%	93.40%	13	17		IGHJ4*01	IGKJ1-2*01	13	11	
11895	378-1	LN	IGHV1S45*01	IGKV1S2*02	96%	89.50%	9	16	IGHD3	IGHJ3*01	IGKJ1-2*01	19	11	SUDV, BDBV, TAFV
11894	378-1	LN	IGHV1S69*01	IGKV1S10*01	95.10%	91.10%	9	17		IGHJ2*01	IGKJ1-2*01	18	12	EBOV
11893	378-1	LN	IGHV1S40*01	IGKV1S42*01	96.30%	92.30%	14	9		IGHJ4*01	IGKJ1-2*01	12	10	EBOV
11890	378-1	LN	IGHV1S47*01	IGKV1S42*01	96.20%	90.20%	9	19		IGHJ4*01	IGKJ1-2*01	19	12	EBOV
13171	378-1	SP	IGHV1S45*01	IGKV1S36*01	96%	93.50%	6	10		IGHJ4*01	IGKJ1-2*01	14	13	EBOV
11885	378-1	LN	IGHV1S40*01	IGKV1S34*01	95%	90.70%	12	13		IGHJ4*01	IGKJ1-2*01	12	10	EBOV
11891	378-1	LN	IGHV1S45*01	IGKV1S34*01	92.30%	92.80%	16	15		IGHJ4*01	IGKJ1-2*01	18	13	SUDV
11882	378-2	LN	IGHV1S43*01	IGKV1S32*01	88.10%	93.10%	20	11		IGHJ4*01	IGKJ1-2*01	14	9	SUDV

Figure 15. V, D and J gene usage of heavy and light chain variable region sequences of 18 GP anti-specific antibodies isolated from vaccinated rabbits shown in 7.1 Appendix 2. *SP = spleen, LN = lymph node. Number of amino acid changes from germline shaded according to increasing number of mutations. Antibodies in bold were selected for detailed further characterisation as a panel of six broadly reactive monoclonal antibodies.*

3.6 Discussion

The central aim of this Chapter was to generate a panel of new mAbs that could bind to the GP of all species of *Ebolavirus* known to have caused human disease (EBOV, SUDV, BDBV and TAFV). The outcome of the antibody discovery process was a small panel of rabbit mAbs that was none-the-less enriched for antibodies that bound to GPs from three or more of the *Ebolavirus* species tested.

3.6.1 RAB-9 cells as immunogen and vaccination regimen

A key component of the antibody discovery process was the initial antigen selection. My approach was to use an immunogen that allowed agnostic presentation of all possible epitopes in a physiologically relevant state. The reasons for this were:

- Limited published knowledge at the time of areas of cross-reactivity between GPs which focused mainly on recently characterised fusion-loop binding mAbs, and a deliberate attempt to isolate broadly-reactive mAbs against epitopes other than the fusion loop. These mAbs could then be used in concert with fusion loop antibodies to create cocktails for therapeutic development in the future.
- As the rabbit immune repertoire had not been previously been interrogated for anti-GP mAbs using full length GP, and an advantage of using rabbits was the possibility of isolating antibodies to epitopes not immunogenic in other species, I did not want to bias which areas of the GP were presented based on epitopes that were immunogenic in other species previously used for antibody discovery.
- To generate mAbs that would be able to bind in the context of GP as on virions or virally infected cells, rather than against epitopes that may not be accessible to mAbs if given therapeutically e.g. epitopes that are revealed inside the endosomal pathway by processing of GP during invasion of cells.

To this end, full length GP expressed on the surface of mammalian cells or inactivated or pseudotyped virus, as used in human vaccination efforts, allows for production of an immunogen close to that of protein in the context of infection, without changes to GP that would come with making recombinant soluble protein; removal of transmembrane domains to promote solubility, differences in

glycosylation due to expression system, addition of exogenous domains or mutations to stabilise protein conformations. Initially, I explored the use of pseudotyped S-FLU viruses coated in GP as immunogens in preliminary experiments in mice (data not shown) [3, 219], but issues in titre for TAFV GP pseudotyped viruses and unknown amounts of GP antigen on the surface of virions meant this line of investigation was abandoned. RAB-9 cells are regularly used as a cell-based vaccine platform at UCB and provided an alternative strategy in which all the GP antigens could be readily produced as conformationally relevant, full length transmembrane proteins. The sequences used to transfect cells were based on cDNA eliminating the production of shorter secreted forms of GP made by *Ebolaviruses* [13, 14].

Three immunisation strategies were chosen to maximise the chances of generating broadly reactive antibodies with a limited number of animals. Prior animal immunisation strategies to generate cross-reactive GP mAbs had successfully worked with either sequential or mixed immunisations of *Filovirus* antigens. For example, Furuyama *et al.,* reported antibody 6D6 isolated from mice sequentially immunised with EBOV and SUDV GPs on virus-like particles (VLPs), and Keck *et al.* reported the isolation of broadly reactive macaque antibodies, including FVM02, from animals immunised three times with a mixture of EBOV, SUDV and MARV GP ectodomain lacking the MLD and then a further two times with the GPs on VLPs; the latter resulting in the later additional report of antibody CA45 derived from the same macaque material [190, 209, 210].

In this antibody discovery campaign, the effect of the different immunisation regimens (Figure 7) at the polyclonal level was assayed using serum sampled from each rabbit prior to harvesting tissues for B cell culture to ensure that a detectable response to antigen had been achieved before culling (Figure 8).

Two sequential immunisation strategies were chosen as the degree of conservation at amino acid level between GPs from different species of *Ebolavirus* ranges from approximately 55-75%. Therefore, the order of antigens, and their relatedness, may have consequences for their ability to sustain or frustrate the germinal centre reactions necessary for higher titres of broadly reactive antibody. With a limited number of animals, not all permutations of possible sequences of antigens

could be explored, however the two selected place the two most divergent GPs (EBOV and SUDV) as either the first and second immunisations, or first and final immunisations; it was hypothesised that the divergence of these two sequences would likely have the most impact on antibody generation.

The overall low response of rabbits 378-2 and 378-3 to BDBV GP and TAFV GP in comparison to naïve serum makes any interpretation about the kinetics of the response to these antigens unwise (Figure 8B, C). These antigens are unlikely to be less immunogenic than EBOV GP as a more robust response was mounted by rabbit 378-1 to each of these immunogens. However, for the sequentially immunised rabbits, the strong response to EBOV GP and generally lower responses to the following immunogens may be due to the inclusion of Complete Freund's Adjuvant (CFA) with EBOV GP alone in the first immunisation. In contrast, rabbit 378-1 received all four immunogens in the dose adjuvanted with CFA, which may have contributed to the more robust responses to BDBV GP and TAFV GP across time points.

The intended effect of sequential immunisation is to boost only antibodies that recognise both the previous and current immunogens, because their epitopes are conserved between the antigens. This will only be a limited portion of the previous response. Together, the lower initial response to any antigen not in the first immunisation, and the reduction in the number of B cells boosted each time, likely accounts for the lower serum IgG responses. Additionally, in this case, divergence of the GPs may be too great to sustain many germinal centres resulting in essentially a primary response to the antigen each time and lower titres of antibody. Although the dose of each individual antigen is lower, the mixed antigen cocktail at each subsequent immunisation provided a potential boost to all B cells that recognised any antigen in the first immunisation, all of which were adjuvanted with CFA. This higher overall concentration of antibody may consist, however, of a lower proportion of cross-reactive mAbs as there is less selection directed towards conserved epitopes. An assessment of monoclonal quality, not simply polyclonal quantity is therefore obviously required; but also a consideration of absolute number, rather than simply proportion, as regards the practicality of recovering individual cross-reactive B cells. In hypothetical terms, if only 2% of B cells made by the animal receiving mixed antigen immunisations are cross-reactive, compared to 10% of B cells made by the sequentially

immunised animals, yet the overall quantity of anti-GP antibody made is 10 fold greater, then the mixed antigen immunised rabbit would be a potentially better source of cross-reactive mAbs, with an increased chance of recovery of cross-reactive V region sequences from the available samples.

Due to the n=1 nature of this experiment, the results cannot be interpreted as robust data to support one vaccination regime over another for generating cross-reactive antibodies in general, but these data do indicate that at the polyclonal level all three rabbits had a humoral response to all four antigens, with rabbit 378-1 alone having a robust serum response to all for antigens at the time of B cell harvest for antibody discovery. Although a detectable binding response, none of the polyclonal sera were able to neutralise pseudotyped S-FLU viruses *in vitro*, hence this could not be used as a discriminator between immune samples for progression towards antibody discovery. The lack of neutralisation by serum was disappointing as Keck *et al*. had reported achieving cross-neutralising antibody in the serum of vaccinated macaques which they then used for antibody discovery [210]. In addition, as discussed in Section 1.4.1, a range of other vaccine platforms have also been able to induce cross-neutralising antibody titres in a range of animals and assays [159-162]. However, none used the same vaccine platform or virus system as in this thesis.

The UCB Core Antibody Discovery high throughput B cell culture platform in tandem with secondary screening using an iQUE flow cytometer does provide the potential to interrogate in high numbers the cross-reactivity of B cells at essentially monoclonal level, allowing comparison between vaccination regimens and between B cells from different lymphoid tissues without having to make large panels of recombinant antibodies. The technical problems encountered during the experiment reported here means that only B cell cultures for rabbit 387-1 lymph nodes were assayed in high numbers so a comparison between B cell compartments and animals was not possible. However, the remaining material from the immunised rabbits could be utilised in this manner.

3.6.2 Screening strategy

High throughput screening assays are necessary to select a feasible number of B cell cultures for manual single B cell recovery (tens to hundreds) from the thousands of cultures prepared from immune samples. Screening assays are designed to identify culture wells containing IgG with desired

binding characteristics; in this case, breadth of binding to multiple *Ebolavirus* GPs without binding to other cell surface antigens.

As the immune samples were generated using a cell-based vaccine rather than recombinant protein or peptides, it was necessary first to identify IgGs made in response to the vaccine that recognise other mammalian cell-surface antigens so as to exclude them from selection. The New Zealand White rabbit fibroblast cell line (RAB-9) was chosen for immunisation to reduce the response to the cells by exploiting immune tolerance to rabbit antigens. However, as RAB-9 cells are maintained in continuous culture and subjected to transfection protocols they will most likely have antigenic differences to primary cells. In this case, polyclonal serum from immunised rabbits did not show a strong response to HEK cells, but did to MDCK SIAT-1 cells (data not shown). The high level of consistent GP expression on the surface of these robust MDCK SIAT-1 cells makes them useful reagents in screening assays. Counter-screening against the parental cell-line was therefore necessary to reduce false positives due to the potential cross-reactivity of MDCK SIAT-1 cell lines and the RAB-9 cell line used in immunisations. Using a mixture of EBOV GP and SUDV GP expressing cells for positive selection increases the chance of any GP binding mAb being detected in the primary screen, compared to using a single GP. There was also the possibility that a mAb that can recognise both EBOV GP and SUDV GP would result in a higher signal in the primary screen, however strength of signal in the primary assay is also highly dependent on the concentration of IgG in the B cell culture supernatant which was not measured.

Secondary screening, occurring after B cells were frozen, was utilised to identify culture supernatants containing likely broadly reactive IgG and provided another opportunity to confirm antigen-specific binding over background cell binding in assays in which internal positive and negative assay controls could be included.

Two different fluorescence-based assays using FMAT and iQUE flow cytometer machines were used to facilitate selection of top candidates for V region recovery. However, there was a large amount of disagreement between the two assays when the same culture supernatants were assayed. The lack of supernatants showing reactivity to the parental MDCK SIAT-1 cells in the FMAT assay may simply

reflect the fact that these samples had been pre-screened for negative binding to the parental cell line in a similar homogenous assay format (using the Mirrorball, TTP) in the primary screening step in (Figure 10). The FMAT result confirms that the gating of the primary screen results successfully excluded almost all MDCK SIAT-1 cell binding samples identified in that assay prior to secondary screening. However, a comparison of the FMAT and iQUE assays indicates a large discrepancy in the identification of cell-binding samples. The threshold for identifying binding to podoplanin transfected cells in the iQUE assay was very conservative, with the many of samples classified as positive in the analysis displaying relatively weak binding (Figure 11D). This likely accounts for the majority of the discrepancy between identification of cell binding samples between the iQUE and FMAT assays, making them more consistent than the applied thresholds suggested. In addition, the assays were conducted with different cell lines, and antigenic differences between the MDCK-SIAT 1 and HEK293F cell lines may account for some of the discrepancy when comparing non-GP binding samples, especially as MDCK SIAT-1 cells had to be lifted from their culture flask for use in the assay, whereas the HEK cells were maintained in suspension culture and therefore were not treated with any enzyme before the assay. In future, rather than assigning an arbitrary cut off value based on background fluorescence in the iQUE assay, a similar strategy to that applied to the primary screen could be used, essentially gating based on median and distribution of signals in the assay.

There were a large number of samples identified as containing GP-binding IgG in the FMAT assay that were not classed as binding to EBOV or SUDV GP antigens in the iQUE assay (Figure 11C). Whilst some of this is accounted for by the re-classification of EBOV/SUDV GP positive wells in the FMAT assay as GP binders in the iQUE assay, this may also be due to differences in antigen expression between the two systems; the iQUE used transiently transfected cell samples in which not all cells were expressing GP, whereas the FMAT system used a stably transduced cell line that has been previously sorted for high GP expression. In addition, the assay formats differ with the homogenous 'no-wash' assay more susceptible to the 'hook-effect' at high concentrations if IgG, and the flow cytometry assay with wash steps more stringent for removal of low affinity binders. Hence, samples with low affinity or low concentrations of GP binding IgG may be positive in the FMAT, but not the

iQUE. Ultimately, selecting the samples that were positive in both assays resulted in a panel enriched for mAbs that bind multiple *Ebolavirus* GPs.

3.6.3 V region recovery and confirmation of cross-reactivity

Some of the antibodies showed a lack of binding to GP once expressed as recombinant antibody despite having come from foci wells identified as containing IgG that could bind full length GP expressed on cells, and that successfully formed antigen-specific fluorescent foci during B cell recovery using soluble GP protein. This is likely to have occurred at the foci picking stage in which single B cells are manually retrieved from mixed cell samples based on their ability to form florescent foci with beads coated with antigen. The individual B cells were snap frozen then subjected to reverse transcription and nested PCRs to recover V region cDNAs. Prior to subcloning and expression of recombinant antibody from plasmid expression vectors, small scale expression of the recovered sequences was tested by incorporating V regions into linear expression cassettes by PCR. Expression of IgG was confirmed in the majority of these supernatants, however when tested for binding to the same soluble protein antigen used for fluorescent foci formation, many samples no longer showed reactivity to GP. These antibodies generally also showed no binding to GP once expressed from plasmids. Likely this was the result of simply retrieving ambient contaminating RNA or a non-antigen specific B cell adjacent to a foci-forming B cell during foci picking. Even if sampled along with the foci-forming B cell, the mRNA from a non-antigen specific B cell could happen to be amplified preferentially. The B cell culture wells are seeded with multiple B cells, but at a density expected to achieve one antigen specific clone per well, based on previous experience at UCB and titres of antigen-specific IgG in the serum [216, 218]. In addition, the seeding density used could have been too high and as multiple B cells are seeded in every well, it is possible that multiple GP binding B cells were on occasion cultured in the same well. In this scenario, where cross-reactivity of B cell culture supernatant was the combined effect of >1 B cell, but only one V region was recovered, this would result in a decrease in cross-reactivity between secondary screening and recombinant IgG. However, due to the hit rate in primary screening (~4%), it is anticipated this would be the case in very few culture wells.

Despite the difficulties in rationalising the two secondary screening approaches and loss of antigen specific B cells at foci picking, a panel of GP-binding mAbs was generated which was enriched for highly cross-reactive antibodies. In a previous collaboration between UCB and the research groups of Professors Draper and Townsend, a larger panel of 82 human monoclonals was made from samples from vaccinees who had received ChAd3 EBOV and MVA-BN Filo (encoding EBOV GP, SUDV GP and Marburgvirus GP and TAFV nucleoprotein) vaccinations [3]. This panel was generated by a combination of screening and sorting strategies focusing only on binding to EBOV GP. Of these 82 mAbs, 75 have been tested for binding to EBOV, SUDV, BDBV [3] and TAFV GPs (Figure 14); 53% bound only EBOV GP, and 33% bound three or more GPs, 16% binding all four. For the panel of rabbit mAbs described in this Chapter, only 31% bound a single GP, with 58% binding three or more GPs, with 42% able to bind all four GPs. In addition, as SUDV GP was included in the primary screen, the rabbit panel includes two SUDV GP-specific mAbs.

3.6.4 Antibody sequences are reflective of reported rabbit Ig repertoires

The heavy and light gene usage across the panel is entirely reflective of reported germline gene usage for NZW rabbits in Ig repertoire deep-sequencing studies [241, 242]. V_H gene segment usage in rabbits is highly restricted compared to humans and mice [243], with ~40% of B cells using the IGHV1S69*01, and ~50% of remaining sequences using either IGHV1S45 or IGHV1S40. J_H germline usage is consistent with that reported by Kondagattil *et al.* with J4 contributing to >50% of sequences, and J2 and J3 contributing every other sequence.

Vκ germline usage in B cells in the rabbit bone marrow has been shown to be more diverse, yet still dominated by IGKV1S10 and IGKV1S36 (both individually accounting for ~10-20% of germline gene usage), with other germline genes each accounting for ~1-10% of gene usage [241, 242]. This is consistent with the usage seen in this panel, with IGKV1S10 and IGKV1S36 ~40% sequences total, and other gene segments occurring only once or twice. Jκ germline usage has been shown to be vastly dominated by the IGKJ-2 germline (~90%) in rabbits, which is reflected in the lack of variety in light chain Jκ segment usage in this panel [242]. Heavy and light chain CDR3 lengths of 6-19 and 9-13 amino acids respectively are also in keeping with reported analyses of rabbit Ig repertoire [241, 242].

There is no apparent relationship between the degree of mutation from germline and the degree of cross-reactivity of the antibodies generated here. In general, there is a higher degree of somatic mutation in both V_H and Vκ sequences (on average ~11 for heavy chain and ~14 for light chain) in comparison to anti-GP human monoclonal antibodies generated from vaccinees by Rijal *et al*., where only ~4-6 amino acid changes from germline on average were identified in V_H sequences [3]. Degree of SHM is similar between rabbits and humans [242], suggesting this greater degree of SHM reflects the three booster vaccinations the rabbits received in this study compared to single booster vaccination received by the human donors.

This panel of new mAbs, in particular those identified as broadly reactive to GP from all *Ebolaviruses* known to cause human disease, were taken forward and are the focus of epitope mapping and *in vitro* functional characterisation in Chapters 4 and 5.

4 *In vitro* characterisation of mAb panel

4.1 Introduction

In vitro characterisation of the new panel of rabbit monoclonal antibodies against *Ebolavirus* Glycoproteins (GPs) isolated in Chapter 3 is required to determine the best candidates to advance to *in vivo* testing in animal models. As described in Section 1.3.2, the entry of *Ebolaviruses* into host cells is a multistep process mediated by the GP from cell attachment to entry into the cytoplasm. After uptake of the virus into endosomes, processing of the GP by host proteases, host receptor binding, and global rearrangements of the GP to mediate fusion with host endosomal membranes are all necessary steps for infection. Monoclonal antibodies can act at many stages to neutralise viral infection of cells.

4.1.1 S-FLU pseudotype viruses

Ebolaviruses are BSL-4 pathogens due to the high morbidity and mortality of ED and their transmissibility between humans. As such, the use of safer surrogate viruses that can be handled at lower laboratory containment facilitates research such as screening of viral inhibitors, including mAbs, in more convenient laboratory settings. Multiple pseudotype or single-cycle *Ebolavirus* systems are used in the literature to study candidate viral inhibitors before progressing to studies in BSL-4 laboratories with wild type *Ebolaviruses* [244-246].

The Townsend lab has developed the S-FLU pseudotype system based on a disabled influenza virus core which is replication incompetent unless a viral attachment and fusion protein is provided in *trans* [247, 248]. The S-FLU core can be coated with *Ebolavirus* GP by passage through a GP expressing cell line. The resultant virus is infection competent, but replication incompetent in non-GP expressing cells and contains no viral RNA or DNA encoding the GP, allowing it to be handled at BSL-2. Infection is dependent on host cell NPC1 as for wild type *Ebolaviruses* [219].

A fluorescent reporter gene in the virus (e.g. GFP) replaces the hemagglutinin coding sequence, making infected MDCK SIAT-1 cells detectable by fluorescence. Loss of fluorescent signal indicates neutralisation of the virus and prevention of infection. The S-FLU microneutralisation (MN)

assay measures the ability of an antibody to prevent the infection of MDCK SIAT-1 cells by an S-FLU pseudotype virus coated with GP from an *Ebolavirus* and has previously been used to screen for small-molecule inhibitors and neutralising mAbs [3, 219].

4.1.2 Mechanisms of neutralisation by EBOV GP monoclonal antibodies

4.1.2.1 Inhibition of GP cleavage

Stepwise cleavage of GP1 by endosomal cysteine proteases Cathepsins B and L is necessary for viral entry into cells [111, 130]. In early *in vitro* experiments using GP coated Vesicular stomatitis virus (VSV) pseudovirions, the ~130kDa GP1 was cleaved to ~50kDa and then to ~20kDa, and finally to a ~19kDa fragment, in a processive time-dependent manner with the loss of the larger products as the smaller products appeared; this cathepsin-mediated proteolysis was faithfully mimicked by the protease thermolysin, suggesting a specific cleavage site in the GP structure [110, 112, 130].

The predicted cleavage site that results in the final cleaved GP (GP_{CL}, consisting of GP2 with the ~19kDa GP1 fragment) is close to the β13-β14 loop (residues 190-213) which is unresolved in many crystal structures, but bridges over the stem of the FL and is likely exposed in the pre-fusion conformation [1, 147].

Structural studies of published mAb100 with the GP, show it occludes this loop, and the same mAb has been demonstrated to inhibit thermolysin mediated cleavage of GP1 [186]. This mAb was able to neutralise lentivirus GP pseudotypes potently *in vitro* and was protective as part of a cocktail in an NHP model of Ebola virus infection; it is therefore an example of a protective mAb that utilises this mechanism of neutralisation [103].

4.1.2.2 Inhibition of receptor binding

Niemann-Pick C1 (NPC1), a lysosomal cholesterol transporter, is the essential host receptor for *Filovirus* infection; expression of NPC1 in non-permissive cells confers susceptibility, and knocking down expression in permissive cells abrogates infection [116, 249]. The receptor binding region (RBR) core is defined across studies as residues 54-201 of GP1 [1, 11, 106, 112, 250].

However, NPC1 binding only occurs after enzymatic processing of the GP1 to produce GP_{CL} in the endosome. Purified human NPC1-C luminal domain has been shown to interact directly and specifically with GP_{CL}, and to be necessary and sufficient for viral entry [249]. Therefore inhibition of either the processing required to produce GP_{CL} or NPC-C1 binding to GP_{CL} represent possible mechanisms of inhibition of viral entry by monoclonal antibodies; either of which could occur via steric hindrance or via allosteric effects. Published examples of mAbs that inhibit receptor binding are FVM04 and Ansuvimab (mAb114). The former is broadly neutralising *in vitro* although poorly protective *in vivo* [190, 210], and the latter is potently neutralising of EBOV, protective *in vivo,* showed efficacy in a human clinical trial and has been shown to act by direct receptor competition and not inhibition of GP processing [45, 186].

4.1.2.3 Inhibition of fusion

After the cleavage and receptor binding events that trigger fusion, the GP rearranges into a typical class I viral fusion protein six-helix bundle tipped by the fusion loop (FL) (GP2 511-556) which is inserted into the host endosomal membrane, resulting in membrane fusion and entry of the virus into the cytoplasm [136]. Highly conserved across *Ebolaviruses*, the function and location of the FL in members of another *Filovirus* genus, *Marburgvirus*, is highly similar although sequence identity and chemistry is not conserved outside of the tip of the loop (Figure 21A). As discussed in Section 1.3.1, the FL consists of a continuous portion of the primary protein sequence containing base and stem sequences that form an antiparallel β-sheet anchoring a central fusion peptide at the tip of the loop (FP) (Figure 21B). The FP of each GP1-GP2 monomer is cradled against the neighbouring GP1-GP2 monomer in the pre-fusion trimer and is the target of several broadly reactive mAbs (Section 1.4.3.2.1). Broadly neutralising mAb 6D6 which binds across the tip of the FL has been shown to inhibit entry of virus into cells at this fusion step, rather than cell attachment or internalisation [209]. Intuitively, this could occur by clamping the FL to the pre-fusion conformation of the GP preventing rearrangements required to generate the six-helix-bundle or other conformational intermediates required for membrane fusion, or by binding to the FP and preventing its insertion into the endosomal membrane even if not preventing global GP rearrangements. The FP of HIV-1 Envelope Glycoprotein

has also been shown to be the target of neutralising mAbs suggesting the FL is a site of vulnerability across class I viral fusion proteins [251].

This Chapter details *in vitro* neutralisation of GP pseudotyped viruses by the mAbs identified in Chapter 3, and begins to investigate the mechanisms of neutralisation employed by these mAbs in biochemical assays of inhibition of thermolysin cleavage, inhibition of receptor binding and binding to the fusion loop.

4.2 *In vitro* neutralisation of S-FLU pseudotype viruses

Purified recombinant antibodies were tested for their ability to neutralise S-FLU pseudotypes coated in GP from EBOV, SUDV and BDBV (NP_066246.1, YP_138523.1, YP_003815435.1 respectively) in a microneutralisation (MN) assay as summarised in Figure 16A.

All antibodies were tested in the MN assay at a limited range of concentrations, then all neutralising antibodies were tested again to completely titrate neutralisation as shown in Figure 16B. Additionally, CA45, a known broadly neutralising mAb [190], and a cocktail of human mAbs from the Oxford/UCB panel selected to provide broad neutralisation in combination and that protected guinea pigs against EBOV challenge (6541, 040, 6662 and 66-3-9C) [3] were tested as positive controls and comparators.

11886 could neutralise all three pseudotyped viruses tested (Figure 16B). To date this is the first mAb made in the Oxford/UCB collaboration that has been able to neutralise all three viruses. This assay suggests that, compared to CA45, 11886 neutralisation of the EBOV S-FLU pseudotype is less potent, but neutralisation of SUDV S-FLU pseudotype is more potent. However, additional repeats are required to see how significant the effect may be. The BDBV S-FLU neutralisation profile for both mAbs is highly similar.

Additionally, despite showing cross-reactivity to all three GPs, mAbs 11883 and 11889 showed neutralising activity against EBOV and SUDV S-FLU viruses only. Broadly reactive mAb 11892 (and virtually identical mAb 11881) could only partially neutralise BDBV S-FLU; and 11897 and 11878 were

In vitro neutralisation of *Ebolavirus* GP pseudotyped S-FLU viruses

A.

	EBOV			BDBV			SUDV			
		IC_{50} (µg/mL)	IC_{90} (µg/mL)		IC_{50} (µg/mL)	IC_{90} (µg/mL)		IC_{50} (µg/mL)	IC_{90} (µg/mL)	Binding
11897	nn			nn			nn			
11878	nn			nn			nn			
11883	Partial	0.549	n/a	nn			Strong	0.254	0.699	
11886	Strong	1.794	9.379	Strong	0.629	1.988	Strong	0.984	3.410	E, B, S, T
11881	nn			Partial	0.286	n/a	nn			
11892	nn			Partial	0.480	n/a	nn			
11889	Neut	2.474	*	nn			Partial	15.501	nn	
11888	nn			nn			nn			E, B, T, (S)
11895	nn			nn			nn			S, B, T
11887	nn			Neut	17.000	n/a	nn			E, B, T
11891	nn			nn			Strong	0.785	2.933	S, (E)
11896	nn			nn			nn			(E), (S)
13171	nn			nn			nn			(E), (T)
11893	nn			nn			nn			E
11894	Strong	2.581	8.249	nn			nn			E
11885	nn			nn			nn			E
11890	nn			nn			nn			E
11882	nn			nn			nn			S
Cocktail	Strong	1.604	2.955	Partial	2.571	n/a	nn			E, S, B, T
CA45	Strong	0.847	2.679	Strong	0.497	1.502	Strong	2.504	8.454	E, S, B, T

B.

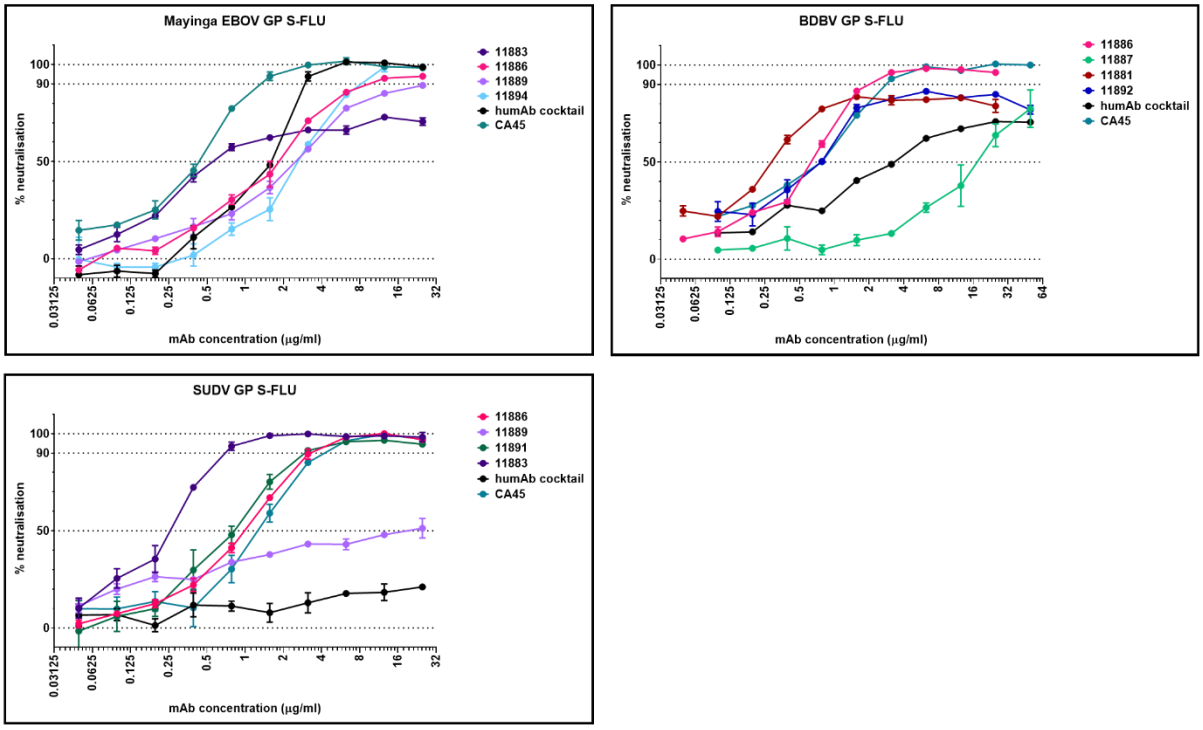


Figure 16. In vitro neutralisation of Ebolavirus GP pseudotyped S-FLU viruses. A. Table of inhibitory concentration at 50% (IC_{50}) and 90% (IC_{90}) values for in vitro neutralisation of S-FLU viruses pseudotyped with GP from EBOV, SUDV and BDBV calculated from data in B. *Strong: neutralises >90% of virus at tested concentrations; neut: neutralises virus to 90% at tested concentrations; partial: cannot neutralise 90% of virus at tested concentrations; nn: non-neutralising. *reached 90% at highest concentration tested. Cocktail of humAbs consists of an equimolar mixture of 6662, 6541, 040 and 66-3-9C with concentration of total mAbs reported above [3]. CA45 is a published broadly neutralising monoclonal antibody [190]. Binding summarised from Figure 13. Brackets indicate where binding was only detected in one binding assay.* ***B. Percent neutralisation of S-FLU viruses by titrated mAbs*** *calculated using maximal (no antibody added) and minimal fluorescence (no virus added) signals within assay. Mean and range of technical duplicates within assay shown.*

non-neutralising in this assay. Despite only partial neutralisation of EBOV S-FLU, 11883 showed very strong neutralisation of SUDV S-FLU with a ten-fold improvement on CA45 for IC_{50} and IC_{90}.

Consistent with their binding profiles, mAbs 11894 and 11891 showed strong neutralisation of EBOV and SUDV S-FLU virus respectively. Additionally, partially cross-reactive mAb 11887 was able to neutralise BDBV S-FLU alone.

The cocktail of human mAbs was able to fully neutralise EBOV S-FLU, only achieved ~70% neutralisation of BDBV S-FLU and could not neutralise the SUDV S-FLU pseudotype at the concentration tested. As the total concentration of antibody was consistent with the other mAbs tested, each component of the cocktail was at a quarter of the concentration of any single other antibody tested.

4.3 Inhibition of thermolysin cleavage

4.3.1 Inhibition of thermolysin cleavage in immunofluorescence assay

In this assay, a confluent monolayer of EBOV GP expressing MDCK SIAT-1 cells was incubated with thermolysin (THL) in HM buffer (pH7.4) to digest GP expressed on the surface to produce GP_{CL}. Cells were incubated with mAbs prior to or after digestion and the amount of antibody remaining bound to GP was compared with cells that had been treated with HM buffer only (Figure 17). Cells treated with THL become less adherent and are more likely to be lost during washing steps in the assay than untreated cells. All cells were stained with wheat germ agglutinin Alexa Fluor 488 conjugate to normalise for number of cells in each well at the end of assay, and the signal from each well was expressed as a ratio of two fluorescence signals.

R5.014 is a negative control antibody against a malaria antigen. P6 is a glycan cap antibody known to inhibit thermolysin-mediated processing of GP (personal communication, Pramila Rijal and Lisa Schimanski). ADI-15946 is a mAb from literature reported to inhibit proteolytic processing of the GP and was included as a comparator [211]. MR78 is a mAb derived from a survivor of Marburg virus infection and only cross-reacts with EBOV GP_{CL} not unprocessed GP [4, 195]. MR78 was included as an internal control for cleavage of GP by thermolysin, i.e. for each part of the assay, MR78 was added

after treatment with THL or HM buffer. The MR78 control indicates a consistent level of cleavage in both groups treated with THL and, as expected, does not show binding to cells that were not digested with THL.

Of the mAbs tested, 11894, 11897, 11883, 11886 and 11889 all behave like P6 and can inhibit thermolysin cleavage in this assay: they show almost no binding to processed GP_{CL}, but retain their epitopes if allowed to bind to GP before addition of THL.

11892 and ADI-15946 appear to have epitopes that are present after thermolysin cleavage, i.e., they can bind both to GP and GP_{CL}. Therefore this assay cannot determine whether they are able to inhibit THL processing of GP or not.

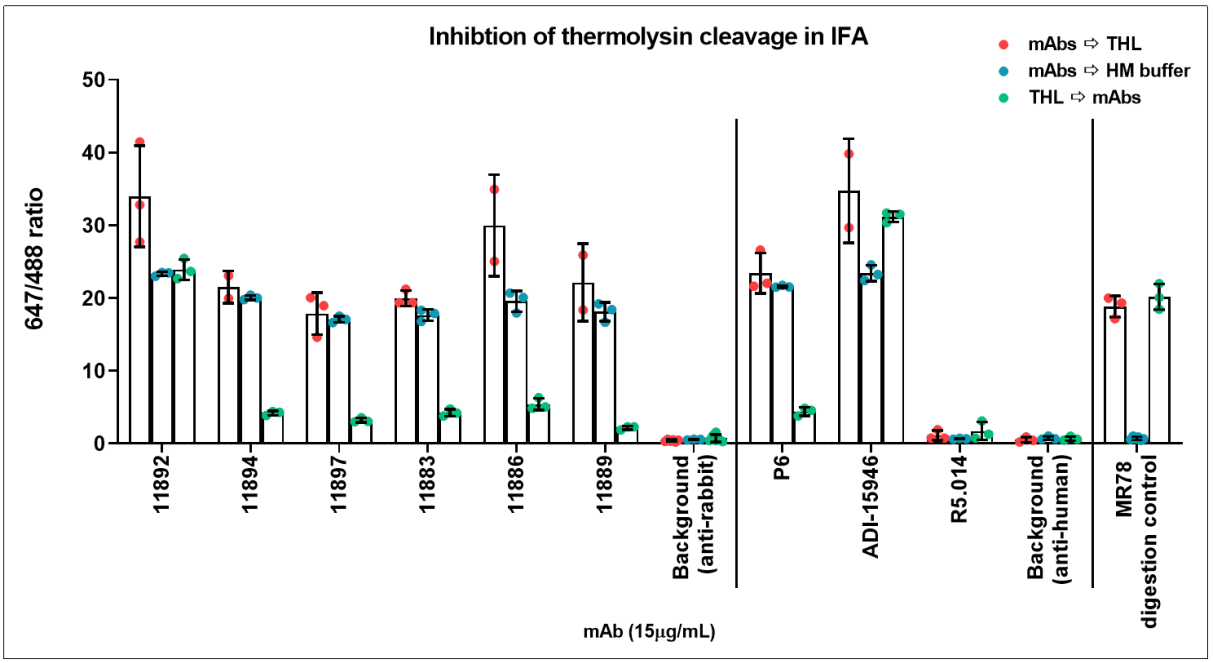


Figure 17. Inhibition of thermolysin (THL) cleavage of EBOV GP by mAbs in immunofluorescence assay. *For each mAb, binding to GP was tested under three conditions: mAbs were pre-incubated with GP then GP treated with THL (red); mAbs were pre-incubated with GP and cells were treated with buffer only without enzyme (blue); and mAbs were incubated with GP_{CL} (GP already treated with THL (green)). Rabbit antibodies were detected using an anti-rabbit IgG Alexa Fluor 647 conjugate and human mAbs P6, ADI-15946, R5.014 and MR78 were detected using an anti-human IgG Alexa Fluor 647 conjugate. Background indicates signals from duplicate wells containing cells stained with only secondary antibodies. All cells were stained with wheat germ agglutinin Alexa Fluor 488 (Invitrogen, W11261). Fluorescence intensity was read at 625-30/680-30 and 488-14/535-30 using a Clariostar plate reader (BMG Labtech). Gain was adjusted to give a ratio of approximately 1 in assay plate wells containing cells that had been stained with both 488 and 647 conjugates only. Mean and range of triplicate wells within assay shown, with wells excluded that gave overflow measurements or had too few cells remaining.*

4.3.2 Inhibition of thermolysin cleavage in immunoprecipitation assay

To confirm the ability of the mAbs to inhibit thermolysin cleavage of the GP, to determine whether mAbs that have epitopes that are not sensitive to THL cleavage are capable of inhibiting THL processing, and to begin investigating which part of the processing the mAbs are inhibiting, I next conducted immunoprecipitation experiments using a protocol optimised by Lisa Schimanski. Briefly, surface proteins of MDCK SIAT-1 cells expressing GP from EBOV, SUDV or BDBV were biotinylated, then incubated with Fab versions of the mAbs of interest. Cells were then treated with different concentrations of THL to process the biotinylated GP. Cells were lysed in the presence of protease inhibitors and lysate incubated with Protein A Sepharose and cross-reactive GP core/base mAb 6541. Protein A Sepharose was washed, and bound protein eluted and separated using SDS-PAGE electrophoresis. A western blot was conducted using a streptavidin AlexaFluor 647 conjugate and fluorescent bands imaged. The different products present in each sample compared to GP that was not pre-incubated with a Fab indicate the degree to which each Fab was able to protect certain bands from processive digestion with increasing concentrations of thermolysin. The same digestion pattern was seen by Lisa Schimanski when incubating the cells with the same concentration of THL sampled over a time course (*personal communication*) and is consistent with observations in similar experiments in the literature [130, 186].

In our assay, without the addition of THL the major band is at ~150kDa, and represents the full length EBOV GP1 with MLD intact. Even without addition of exogenous protease, there is spontaneous cleavage to form the ~60kDa product via another defined intermediate. The ~60kDa band is consistent with GP1 with the MLD removed (GPΔmuc), as it is a similar size to GP1 from GPΔmuc constructs created by genetic deletions in the literature [186]. The ~25kDa band is consistent with the predicted size of GP2 alone.

Upon the addition of THL the ~60kDa intermediate is further processed via a 40-50kDa intermediate. At higher concentrations of THL, the loss of the ~60kDa band is accompanied by the appearance of a 40-50kDa band and an increase in density in the ~20-25kDa region, which sometimes resolves into a doublet. This increase in density has been determined experimentally by Lisa

Schimanski, although these data are not shown here. The latter is approximately the predicted size of both the GP2 and the GP1 core. An ~19-20kDa GP1 band is seen in a variety of experiments in the literature using VSV-pseudovirions coated with GP and digested with THL or cathepsins, and VSV pseudoviruses coated in a ~19kDa GP1 with GP2 are still infection competent, hence it represents the GP1 core, containing the receptor binding region [130]. Therefore loss of the MLD appears to be somewhat independent of THL in this assay, although it happens more rapidly with the addition of THL. The loss of the GC and processing of the GPΔmuc by THL is likely represented by loss of the ~60kDa band in these experiments. The 40-50kDa band intermediate only appears in samples treated with THL, suggesting it is a true cleavage intermediate. Hence the ability of pre-incubation with a Fab to protect against loss of the ~60kDa band and formation of the 40-50kDa intermediate can be interpreted as inhibition of thermolysin catalysed processing of the GP to GP_{CL}. We plan to assess the content of each band in the assay via mass spectrometry.

Pre-incubation of EBOV GP with all Fabs except 11878 led to the ~60kDa band being retained after treatment with 250µg/mL THL. Additionally, for all Fabs tested except 11897 and 11878, the development of the 40-50kDa cleavage intermediate appears to have been largely prevented when treated with 50 or 250µg/mL THL. None of the mAbs can inhibit the initial processing of the full length GP, suggesting they cannot prevent removal of the mucin-like domain, but do affect the processing and removal of the glycan cap by THL. This pattern was reproducible (n=2) and representative blots are shown in Figure 18.

Additionally, whilst the digestion products vary slightly in size for BDBV and SUDV GP, a similar pattern is seen across species for those Fabs tested.

Pre-incubation of BDBV GP with 11883, 11886, 11889 and 11892 results in the strong retention of the ~80-90kDa GP1 bands, and prevention of the development of the bands that represent further processed, smaller GP1 cleavage products at ~40-50kDa.

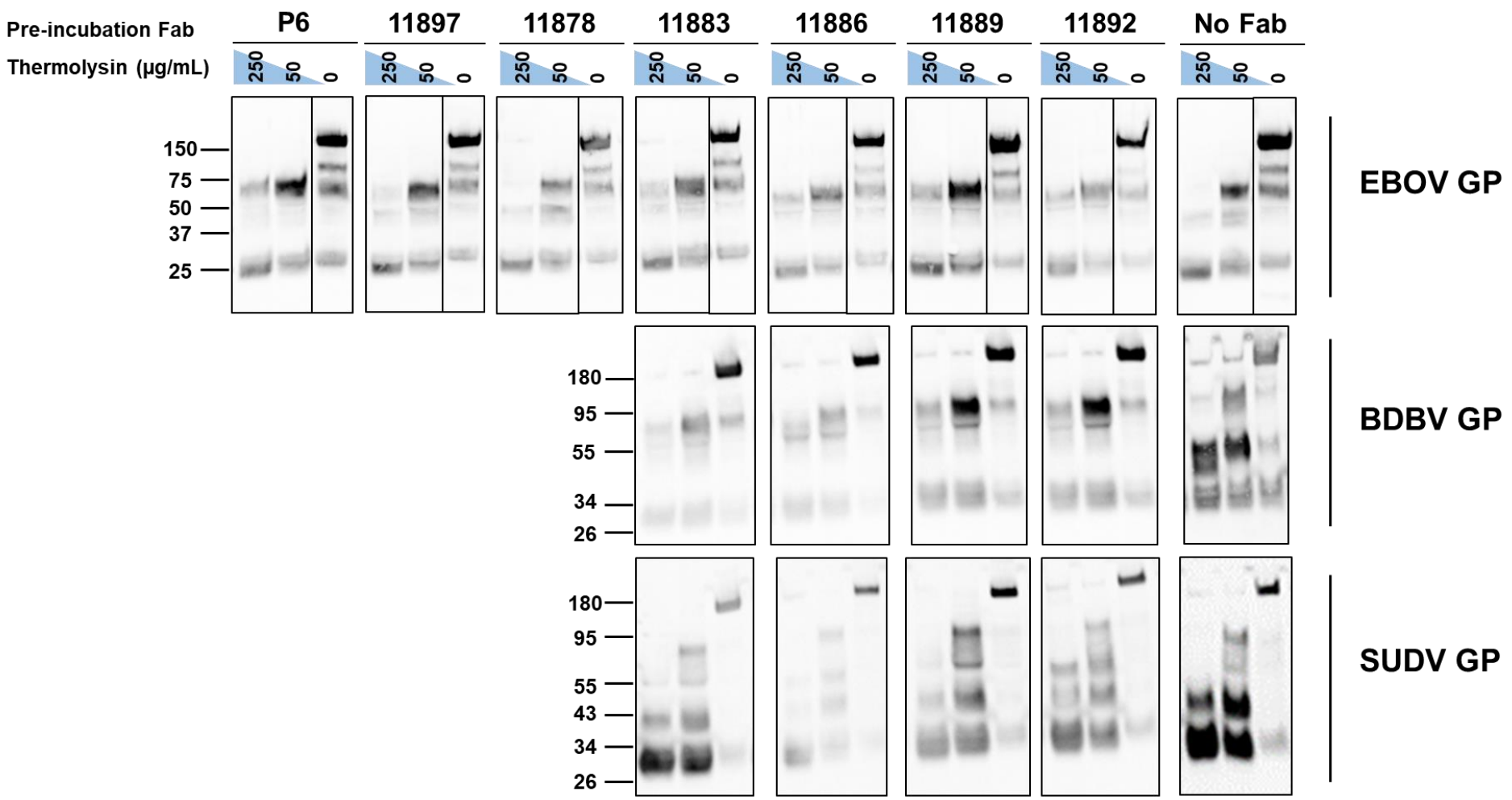


Figure 18. Immunoprecipitation of thermolysin cleaved GPs pre-incubated with broadly reactive Fabs. *Representative blots for EBOV GP (n=2), BDBV GP (n=1) and SUDV GP (n=1). Biotinylated surface proteins were treated with increasing concentrations of THL, then immunoprecipitated using Protein A Sepharose and anti-GP mAb 6541. Bands revealed using Streptavidin AlexaFluor-647 conjugate and imaged using iBright FL100.*

The effect is present but less pronounced in the blots using SUDV GP, with the retention of the ~55kDa band less strong, and no Fab tested prevented the processing of this ~55kDa fragment into the ~40kDa fragment for GP treated with either 50 or 250 µg/mL of THL.

11878 was not tested in the experiment described in Section 4.3.1 due to issues with limited supply of the mAb. However, the Fab appeared more stable and was able to be tested in this assay.

4.4 Inhibition of NPC1 binding to GP_{CL}

4.4.1 Inhibition of NPC1-C binding to GP_{CL}

Soluble NPC1 C domain (NPC1-C) protein was produced by the Townsend lab, based on previously published constructs in which the C domain (residues 372-622) is flanked by sequences that form an antiparallel coiled coil in order to maintain the overall structure of the luminal C domain [249, 252] with additional C terminal tags for purification and labelling. Interaction of biotinylated NPC1-C and antibody MR78 with EBOV GP_{CL} and with intact EBOV GP expressed on MDCK-SIAT 1 cells was confirmed and titrated prior to competition assay (Figure 19A). As previously, EBOV GP_{CL} was generated by treating EBOV GP expressing MDCK SIAT-1 cells with thermolysin. MR78 is a pan-filovirus monoclonal antibody with an epitope in the receptor binding region that prevents NPC1 binding [4].

The broadly reactive rabbit monoclonal antibodies of interest were tested to determine if they reduced NPC1-C binding to GP_{CL} (Figure 19B). Self-competition with unbiotinylated NPC1-C was used to determine a threshold for competition. Binding of biotinylated NPC1-C to GP and GP_{CL} without competing antibody (PBS only) defined upper threshold indicating no competition. MR78 successfully competed biotinylated NPC1-C, reducing signal to background levels. Negative control antibody R5.014 (mAb that binds a malaria antigen) did not compete biotinylated NPC1-C. None of the rabbit monoclonal antibodies tested reduced biotinylated NPC-C1 binding.

Within the assay, MR78 binding to the digested GP in the absence of NPC1-C confirmed that thermolysin digestion of GP was successful. Consistent with published literature, without thermolysin treatment, little or no NPC1-C bound to GP.

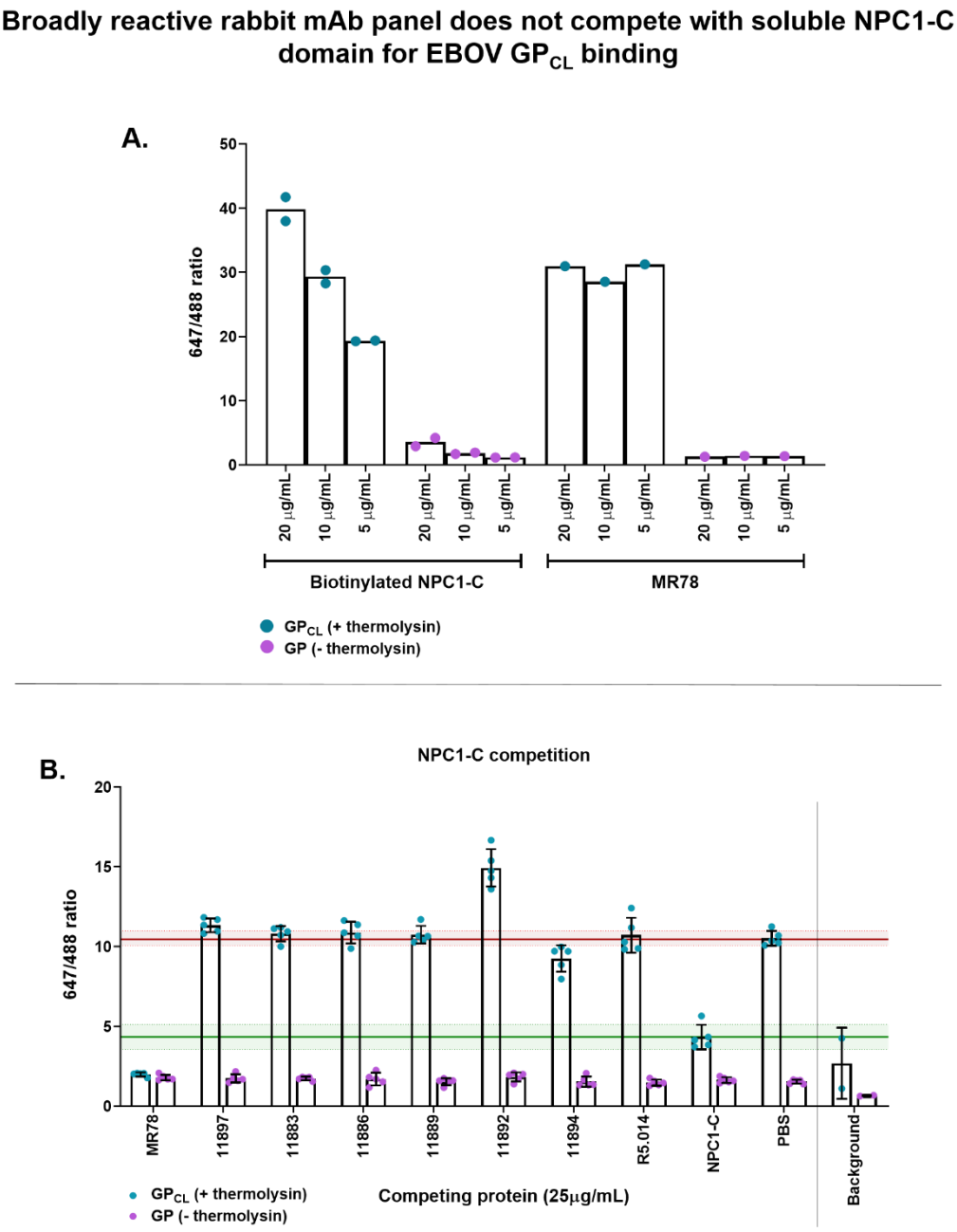


Figure 19. Inhibition of NPC1-C binding in immunofluorescence assay. A. Validation of NPC1-C domain protein binding to EBOV GP with and without thermolysin treatment. *Biotinylated NPC1-C and mAb MR78 titrated against EBOV GP$_{CL}$ and intact EBOV GP expressed on MDCK SIAT-1 cells. Bound NPC1-C and MR78 revealed with streptavidin Alexa Fluor 647 conjugate (Invitrogen, S21374) and anti-human IgG Alexa Fluor 647 conjugate (Invitrogen, A21445) respectively.* ***B. Competition of broadly reactive rabbit mAb panel with NPC1-C.*** *A pre-mixed solution of 5µg/mL biotinylated NPC1-C and 25µg/mL unbiotinylated mAb or unbiotinylated NPC1-C was incubated with cells that had been treated with 250µg/mL thermolysin in HM buffer (GP$_{CL}$), or HM buffer alone (GP). Bound biotinylated NPC1-C detected using streptavidin Alexa Fluor 647 conjugate. Bars represent mean and standard deviation of five replicates within assay. Red line with shaded area and green line with shaded area indicate mean and one standard deviation of signal from biotinylated NPC1-C binding to EBOV GP$_{CL}$ with no competition and when competed against itself. Background indicates signals from duplicate wells containing cells stained with secondary antibodies only.*

For all experiments shown, cells were stained with wheat germ agglutinin Alexa Fluor 488 (Invitrogen, W11261). Fluorescence intensity was read at 625-30/680-30 and 488-14/535-30 using a Clariostar plate reader (BMG Labtech). Gain was adjusted to give a ratio of approximately 1 in wells containing cells that had been stained with both 488 and 647 conjugates only (i.e. no NPC1-C or antibody added).

4.4.2 Prevention of NPC1-C binding by inhibition of GP processing

Assays to determine whether the ability of the mAbs to inhibit THL processing of the GP led to a reduction in receptor binding were also attempted; EBOV GP expressing cells were pre-incubated with antibodies of interest, treated with thermolysin, then incubated with biotinylated NPC1-C, and finally the amount of NPC1-C bound revealed using a streptavidin Alexa Fluor 647 conjugate with fluorescence measured using a Clariostar Plate reader. However, the overall signal in the assay remained low with much variability in signal between repeats. Whilst not consistently statistically significant, there did appear to be a reproducible trend towards the mAbs that had shown the ability to inhibit thermolysin cleavage to reduce the amount of NPC-C1 bound compared to controls without mAb treatment or samples treated with 11878 Fab or R5.014 mAb, neither of which are expected to prevent either THL cleavage or receptor binding (data not shown). The assay requires optimisation and may be improved by using flow cytometry.

In summary, whilst none of the mAbs directly prevented NPC1-C binding to pre-processed GP_{CL}, it is likely that, based on their ability to prevent processing of the GP, the mAbs may impede receptor binding by preventing the NPC1-C binding site being revealed, as demonstrated by other published mAbs [211].

4.5 Effect of I260R mutation on mAb binding and neutralisation

As discussed in Section 4.3, P6 is a glycan cap mAb that has been shown to inhibit thermolysin cleavage of the GP. As part of ongoing work to characterise this mAb by the Townsend lab, EBOV GP mutation I260R was selected for *in vitro* by passage of EBOVΔVP30 viruses [246] with mAb P6 by Professor Peter Halfmann, Kawaoka Lab, University of Wisconsin-Madison. Residue I260 is located in the glycan cap (Figure 20C), in the centre of a fragment of the GP identified as the P6 binding site by yeast peptide display [3].

Introduction of the I260R mutation into EBOV GP expressed on MDCK SIAT-1 cells by Pramila Rijal resulted in a 10 fold reduction in P6 binding to cells in a flow cytometry assay, but no significant

change in binding in IFA experiments. Rijal also reports that viruses pseudotyped with I260R EBOV GP were able to infect cells, and 10/11 GC mAbs tested, including P6, lost the ability to neutralise the virus, despite still being able to bind the GP (*Pramila Rijal, personal communication*). It was therefore of interest to see if these new rabbit mAbs are also susceptible to loss of neutralisation by this point mutation in the GP.

Consistent with data from Pramila Rijal (personal communication), mAbs against the GC (P6, 040), RBR (6662) and base (6541) are unaffected by the I260R mutation in a cell-based immunofluorescence assay, with MR78 acting as a negative control in this assay (Figure 20Ai); the partial decrease in 6662 binding to I260R GP was not reproduced across n=2 repeats. Similarly none of the new rabbit mAbs substantially lost binding to the I260R cells, with 11882 (a SUDV GP mAb) acting as a negative control (Figure 20Aii). Rabbit mAbs that showed EBOV S-FLU neutralisation in Section 4.2 were tested within the same assay for neutralisation of S-FLU pseudotype viruses coated in WT EBOV GP and I260R EBOV GP (Figure 20B). A reduction in fluorescence intensity indicates inhibition of infection. R5.015 is a negative control mAb against a malaria protein. P6 in this assay, consistent with Pramila Rijal's data, shows strong neutralisation of WT EBOV GP S-FLU, but cannot neutralise the I260R EBOV GP S-FLU. Rabbit mAbs 11889, 11894 and 11883 also lose neutralisation of the I260R variant. These mAbs have THL sensitive epitopes indicating a GC or MLD dependency for binding to GP (Section 4.3). Antibodies 6541 and ADI-15946 with known base epitopes that are retained after thermolysin treatment of the GP also retained neutralisation of the mutant I260R EBOV GP S-FLU. 11886, despite the fact it has a thermolysin sensitive epitope (Section 4.3), also retains neutralisation of the mutant I260R EBOV GP S-FLU.

4.6 Identification of fusion loop (FL) dependent mAbs

Having explored the relationship between the broadly reactive mAbs and GP processing and receptor binding, I next investigated the relationship between the panel of mAbs and the FL. The FL represents an important epitope distinct from the receptor binding region that is also retained on the GP_{CL}. It is a target for neutralising and broadly reactive mAbs with many examples of broadly-neutralising and/or broadly-protective mAbs with overlapping epitopes along the length of the FL (Figure 21).

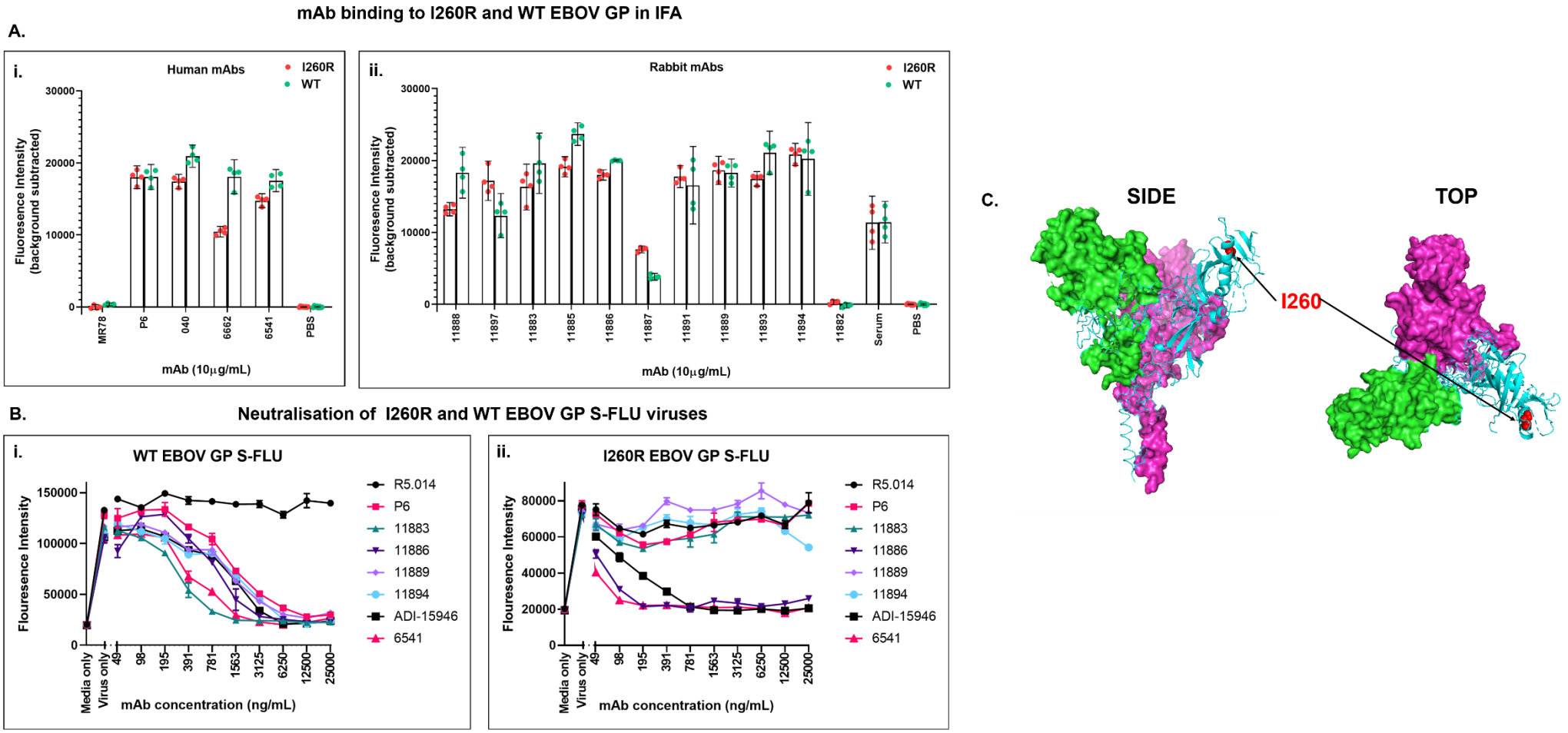


Figure 20. Effect of I260R mutation on mAb binding and neutralisation A. Binding of mAbs to WT EBOV GP and I260R EBOV GP expressed on MDCK SIAT-1 cells in immunofluorescence assay. *Rabbit serum used at 1:10 dilution in PBS. Mean and 95% CI on quadruplicates shown. Representative of n=2 repeats. mAb binding detected using anti-human IgG and anti-rabbit IgG AlexaFluor 488 conjugates respectively. Background fluorescence (average of 8 wells incubated with PBS instead of primary antibody) subtracted from signals.* ***B. Neutralisation of S-FLU viruses pseudotyped with (i) WT EBOV GP and (ii) I260R EBOV GP infecting MDCK SIAT-1 cells****. Mean and range of duplicates within assay shown. All fluorescence intensities measured using Clariostar (BMG Labtech).* ***C. Location of I260 in EBOV GP from side and top (rotated 90°).*** *PDB: 5JQ7 [6] recoloured by GP1-GP2 monomer (green, magenta and cyan). I260 (red sphere) highlighted on one monomer shown in ribbon representation.*

A high-throughput method for identifying mAbs that depend on the presence of the FL sequences to bind to the GP and/or neutralise the virus could provide a rapid method for screening for broadly protective mAbs as well as to investigate mechanism of neutralisation. This section describes the use of two different approaches to the identification of FL binding mAbs that avoid having to generate alanine scanning libraries, yeast/phage display GP peptide libraries and structural techniques.

4.6.1 ELISA using synthetic fusion loop peptide

ELISAs represent a quick method for testing binding of many samples to a peptide. A cyclic, biotinylated peptide with sequence from the fusion peptide (FP) mimicking the very tip of the FL was designed by Terry Baker (UCB Pharma) (Figure 21).

ELISA plates were coated in streptavidin then biotinylated peptide. A selection of human mAbs (humAbs) from the literature and the Oxford/UCB panel with known epitopes were tested by ELISA for binding to the synthetic FP (Figure 22). Included were GC mAb 040, and RBR mAbs 6662 and MR78 as negative controls, along with FL stem/base mAbs CA45 and ADI-15946. FVM02, a known FL antibody was able to bind to the peptide in the ELISA, consistent with published data showing it can bind linear peptides representing residues 526-535 in the FL [210]. All mAbs from the Oxford/UCB panel that had previously been determined to be base or GP core binding mAbs were also tested [3]; of these only 66-3-2B showed binding to the peptide after a 2 hour development incubation. Unlike FVM02, 66-3-2B is only partially cross-reactive as it is reported to bind EBOV and BDBV GP, but not SUDV GP [3]. Alongside the humAbs, the panel of new rabbit mAbs were tested, however none showed any reactivity to the peptide in this assay, even when the plate was left to develop for 24 hours (data not shown). Serum from rabbit 378-1 two weeks after final vaccination (Section 3.2.3.1) was included as a positive control for the anti-rabbit IgG alkaline phosphatase conjugate used to develop the ELISA.

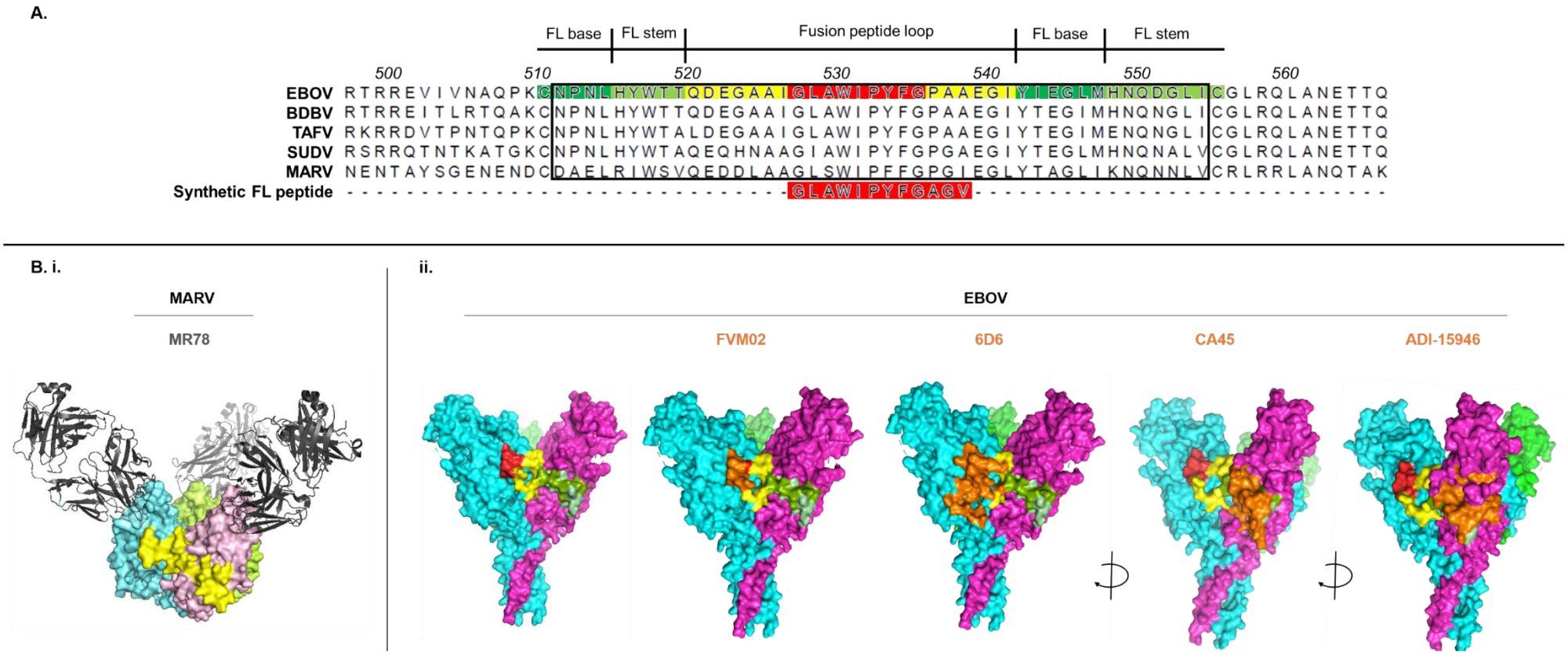


Figure 21. Fusion loop (FL) of Filovirus GPs and design of synthetic FL peptide and chimeric FL GPs. A. Fusion loop sequences from Ebolaviruses and Marburgvirus GPs Sequences obtained from NCBI Reference databas*e. EBOV (Ebola virus/H.sapiens-wt/GIN/2014/Makona-C15; ATY51135), SUDV (Sudan virus/H.sapiens-tc/UGA/2000/Gulu-808892; YP_138523.1), BDBV (Bundibugyo virus/H.sapiens-tc/UGA/2007/Butalya-811250; YP_003815435.1), TAFV (Tai Forest virus/H.sapiens-tc/CIV/1994/Pauleoula-CI; YP_003815426.1), MARV (Marburg virus/H.sapiens-tc/KEN/1980/Mt.Elgon-Musoke; YP_001531156.1). Numbering associated with EBOV sequence. Alignment created in MegAlign Pro 17 (Lasergene DNASTAR). Fusion loop base (dark green), stem (light green) and fusion peptide loop (yellow) are highly conserved between Ebolaviruses. Residues appearing in the Synthetic FL peptide (Terry Baker, UCB) represent the tip of the fusion loop (red). Region in box indicates sequences that are swapped to produce chimeric fusion loop GPs described in 4.6.2.* ***B. Location of fusion loop on Marburg virus (Lake-Victoria Marburgvirus, Kenya, 1987, RAVN-87) and EBOV GP with footprints of published broadly reactive FL mAbs. i****. MARV ΔMUC GP_{CL} ectodomain in complex with receptor binding region MR78 Fab (grey) ([4], PDB: 5UQY), coloured by GP1-GP2 monomers, with FL (yellow) of pink monomer unit highlighted****. ii.*** *EBOV ΔMUC GP ectodomain ([6], PDB: 5JQ7), coloured by GP1-GP2 monomer. FL of magenta monomer unit highlighted as in part A of this figure. Footprints of mAbs coloured in orange elucidate FL tip (FVM02, 6D6) and base/stem (CA45, ADI-15946) FL epitopes. 6D6 epitope also overlaps with epitope of ADI-15878. Adapted and redrawn from [10]. Images created in PyMOL.*

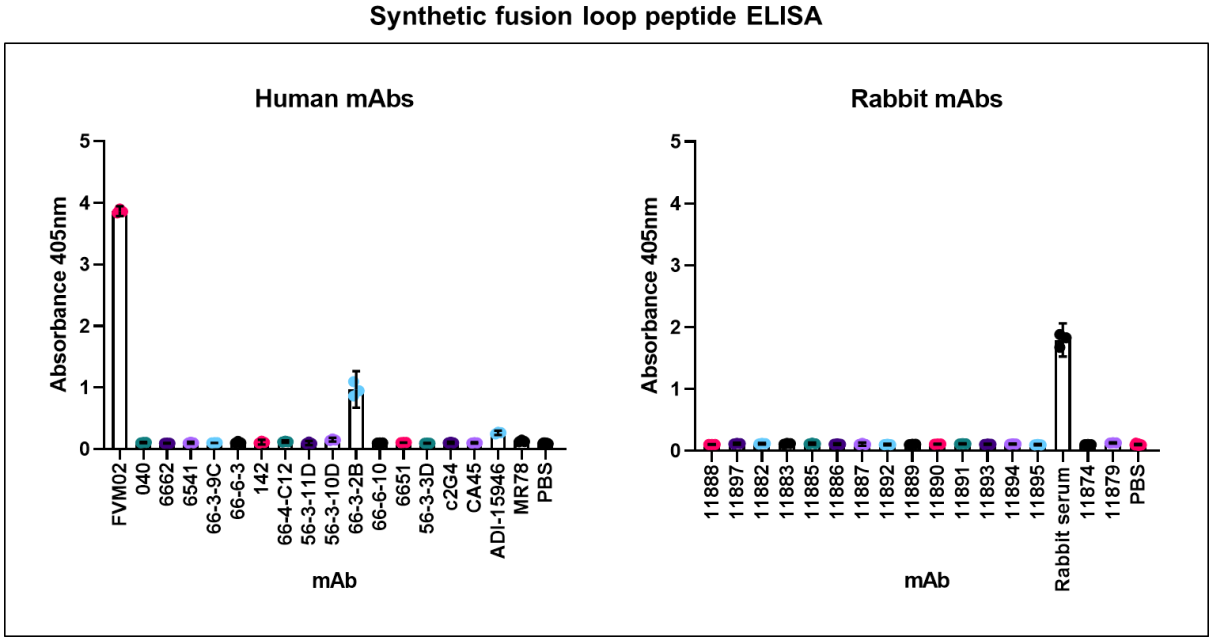


Figure 22. Synthetic fusion loop peptide ELISA to identify mAbs binding to the tip of Ebolavirus fusion loop. *Rabbit serum used at 1:10 dilution in PBS. Human mAb and rabbit mAb samples revealed using anti-human IgG and anti-rabbit alkaline phosphatase conjugates respectively. Developed for 2 h. Absorbance at 405nm read using BioTek Reader 808. Mean and 95% confidence intervals of triplicate wells shown, except for PBS (background) where eight wells were measured.*

4.6.2 Chimeric *Filovirus* GPs for identification of fusion loop (FL) binding mAbs

Filovirus GP FLs have functional and structural similarities, and the amino acids in the FP itself are conserved. However, across the GPs there is low sequence conservation between *Marburgvirus* and *Ebolavirus*es outside of the FP and receptor binding region [250], with very few pan-ebolavirus mAbs able to recognise other *Filoviruses*. A BLAST of EBOV GP (ATY51135.1) and MARV GP (YP_001531156.1) gives a percentage identity of just 50.30% (NCBI Protein Blast tool; accessed via https://blast.ncbi.nlm.nih.gov/Blast.cgi). Therefore chimeric *Filovirus* GPs with heterologous FLs could be used to identify mAbs that are reliant on the EBOV GP FL for binding, and that may interfere with membrane fusion as a mechanism of neutralisation.

An EBOV mAb that binds predominantly to the FL and does not bind MARV GP, would lose binding to a chimeric GP consisting of the EBOV GP with the MARV FL, and potentially gain binding to a chimeric GP consisting of the MARV GP with the EBOV FL. This section describes production and validation of cell lines expressing chimeric FL *Filovirus* GPs and their use to identify mAbs with FL dependency.

4.6.2.1 Design, production and validation of cell lines expressing fusion loop chimeric GPs

DNA constructs comprising the full length EBOV GP (Ebola virus/H.sapiens-wt/GIN/2014/Makona-C15; ATY51135) with residues 512-555 replaced with residues 513-556 from MARV GP (YP_001531156.1, Marburgvirus/H.sapiens-tc/KEN/1980/Mt. Elgon-Musoke) and the reciprocal chimeric GP comprising of full length MARV GP with residues 513-556 replaced with 512-555 from EBOV GP were designed by Alain Townsend and designated E/M and M/E chimeric GPs respectively. Sequences were cloned into vectors for producing lentivirus and MDCK SIAT-1 cell lines transduced to express full length chimeric GPs at the surface. Transduced cell lines were stained using mAbs 6662 for E/M cells and FVM02 for M/E cells. 6662 is an EBOV GP core/RBR binding mAb and so was not expected to lose binding with loss of the EBOV GP FL in the E/M chimeric GP. As previously discussed, FVM02 has been shown to be able to bind just to the FP [210] and as such should be able to bind M/E chimeric GP as the EBOV FP is retained in this construct. ~100,000-250,000 cells per sample were sorted into media and cells expanded for continued culture. Cell sorting was conducted by Craig Waugh, WIMM FACS Facility (Figure 23A).

Sorted cell lines expressing MARV GP, EBOV GP or chimeric GPs were interrogated using a variety of protein reagents including NPC1-C (Figure 23B) and antibodies against a range of epitopes (Figure 23C) to determine whether the chimeric GPs were correctly folded and expressed at levels on the surface of cells to make them useful reagents for identifying mAbs with FL epitopes.

NPC-C1 (endosomal receptor for *Filovirus* GP) and MR78 (pan-filovirus mAb) both bind to GP_{CL} from EBOV or MARV [11]. Both were able to bind to the E/M chimeric GP_{CL}, but binding to the M/E chimeric GP_{CL} was not significantly above background (Mann-Whitney two tailed t test, $P<0.05$) (Figure 23B). In the absence of THL cleavage, neither bound significantly above background at 10 µg/mL (data not shown). This was as expected for NPC1-C, but not for MR78 which is reported by Flyak *et al.* to bind to uncleaved MARV GP in ELISA, however in that study it did not bind as well to the GP from the Musoke isolate used in our studies which may explain this result [195].

040 and 6662 have previously been determined to bind the GC and core/RBR of EBOV GP respectively, and as expected do not bind the MARV or M/E chimeric GPs and retain binding to the E/M chimeric GP, confirming the expression of E/M GP on the surface of cells.

CA45 and ADI-15946 have epitopes with contacts across the base and stem of the EBOV FL (Figure 21Bii.). The predicted loss of binding by these mAbs to E/M chimeric GP compared to the EBOV GP (Figure 23C) suggests these cells can be used to identify FL binding mAbs, however, in this experiment, no base-binding mAbs were tested that retained binding to the E/M chimeric GP, so the possibility that swapping the FL has disrupted the base of the GP more substantially needed to be ruled out. ADI-15946 also appeared to show some cross-reactivity to the MARV GP although binding was much reduced in comparison to EBOV GP binding.

FVM02 showed binding to MARV, EBOV and the E/M chimeric GP, however, in this assay binding to M/E chimeric GP was not convincing. In later experiments, with a higher concentration of FVM02, more substantive binding to the M/E chimera was seen (Figure 23C).

6541 has been previously defined as a base-binding mAb and it competed mAb 6D6 for binding to the GP. Loss of binding in this assay prompted the testing of more antibodies in the Oxford/UCB panel to determine if 6541-like mAbs and competing mAbs similarly lost binding to the E/M chimeric GP (see next section 4.6.2.3).

E/M and M/E GP cell lines were used to produce S-FLU viruses pseudotyped with the chimeric GPs. In parallel, and using the same seed virus, batches of S-FLU were generated using the EBOV GP and MARV GP cell lines. When these virus preparations were titred to determine infectivity, neither chimeric GP virus appeared to be able to infect cells (Figure 23D).

4.6.2.2 Binding of mAb panels to FL chimeric GPs

During the previous characterisation of Oxford/UCB humAb panel by Pramila Rijal, a number were defined as base-binding mAbs by competition with mAb 3CSY/KZ52; within this group some were directly tested for competition with 6541, and assigned to related epitope bins within the base [3]. These base-binding mAbs were therefore all tested for binding to the new chimeric GPs along with

Production and initial validation of chimeric GPs for identifying fusion loop binding mAbs

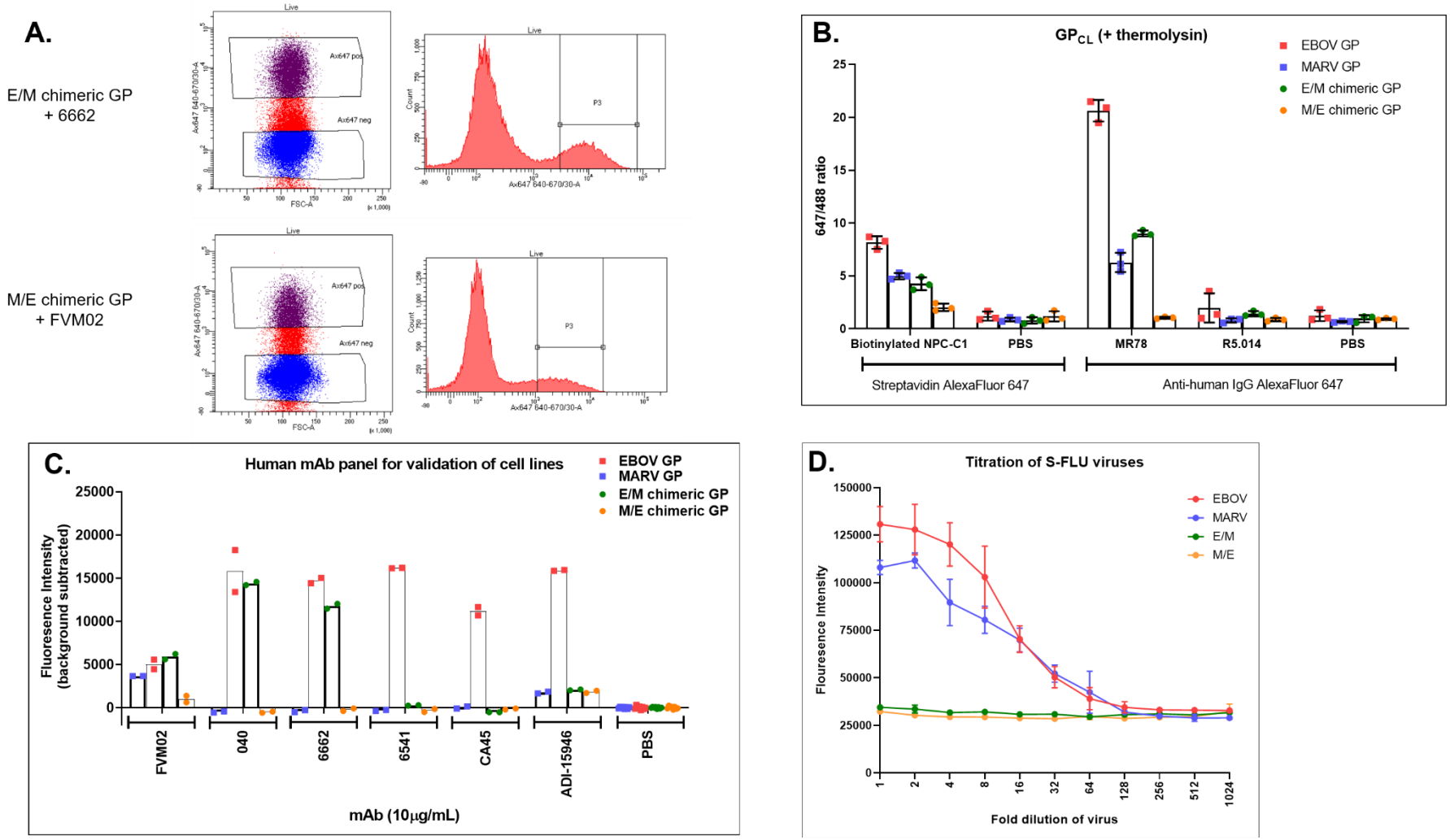


Figure 23. Production and initial validation of chimeric GPs for identifying fusion loop binding mAbs. Caption on next page.

Figure 23 caption. Production and initial validation of chimeric GPs for identifying fusion loop binding mAbs. A. Staining and sorting of E/M and M/E GP transduced MDCK SIAT-1 cells*. Cells were stained with 6662 or FVM02 and anti-human IgG Alexa Fluor 647 conjugate along with live/dead stain. Gates were applied to select singlets and live cells (not shown), then antigen-positive cells by comparison to an isotype control. Cells in gate P3 were sorted by Craig Waugh, WIMM FACS facility, for continued culture. Gating and analysis conducted in BD FACSDiva 8.0.1 by Craig Waugh.* ***B. NPC1-C binding and thermolysin digestion of chimeric GP expressing cell lines.*** *Confluent cell monolayers were treated with thermolysin or buffer alone then incubated with 10µg/mL biotinylated NPC-C1, MR78 or R5.014 in triplicate. NPC1-C was detected using a streptavidin AlexaFluor647 conjugate. MR78 (pan-filovirus, receptor–binding region mAb) and R5.014 (a malaria antigen antibody, negative control) were detected using anti-human IgG AlexaFluor647 conjugate. Cells were stained with Wheat germ agglutinin AlexaFluor-488 conjugate to normalise for number of cells remaining in well after thermolysin treatment. Mean and range of ratio between fluorescence intensities of 647 and 488 conjugates shown.* ***C. Binding of selected humAbs to chimeric GP expressing cell lines.*** *Confluent cell monolayers were incubated with mAbs at 10 µg/mL which was detected using anti-human IgG AlexaFluor 488 conjugate. Mean of duplicates in assay shown. Representative data for 040, 6662, 6541 and FVM02 which were repeated in n=3 experiments.* ***D. Titration of S-FLU viruses prepared by propagation in permissive EBOV GP, MARV GP and chimeric GP cell lines.*** *Viruses in viral growth media were diluted 1:2 in PBS then serially titrated in quadruplicate before addition of MDCK SIAT-1 cells. Cells were fixed after 24 hours and fluorescence intensity measured. For all data shown in B-D fluorescence intensity measured using Clariostar (BMG Labtech).*

c2G4 (Figure 24A). c2G4 is a component of the ZMapp therapeutic cocktail and has contacts in the base of FL [147].

4.6.2.3 Binding of mAb panels to FL chimeric GPs

During the previous characterisation of Oxford/UCB humAb panel by Pramila Rijal, a number were defined as base-binding mAbs by competition with mAb 3CSY/KZ52; within this group some were directly tested for competition with 6541, and assigned to related epitope bins within the base [3]. These base-binding mAbs were therefore all tested for binding to the new chimeric GPs along with c2G4 (Figure 24A). c2G4 is a component of the ZMapp therapeutic cocktail and has contacts in the base of FL [147].

Loss of binding to the E/M chimeric GP was observed for a subset of mAbs tested (56-3-11D, c2G4, 66-3-2B, 142, 6651, 66-6-10), but not all, which increases confidence that replacing the EBOV FL with the MARV FL has not destroyed all base epitopes, and that loss of binding is more likely directly due to loss of FL contacts.

The mAbs tested for which there is evidence of direct competition with 6541 (c2G4, 56-3-11D), both showed loss of binding to the E/M chimera. Of the remaining mAbs which exhibit a range of overlapping epitopes, the majority of mAbs that have been shown to directly compete c2G4 lost binding to the chimeric GP, and those shown not compete with either c2G4 or 6541, were more likely to retain binding to the chimera. Competition groups of base mAbs based on data reported in Rijal *et al* have been combined with FL sensitivity in Figure 24C.

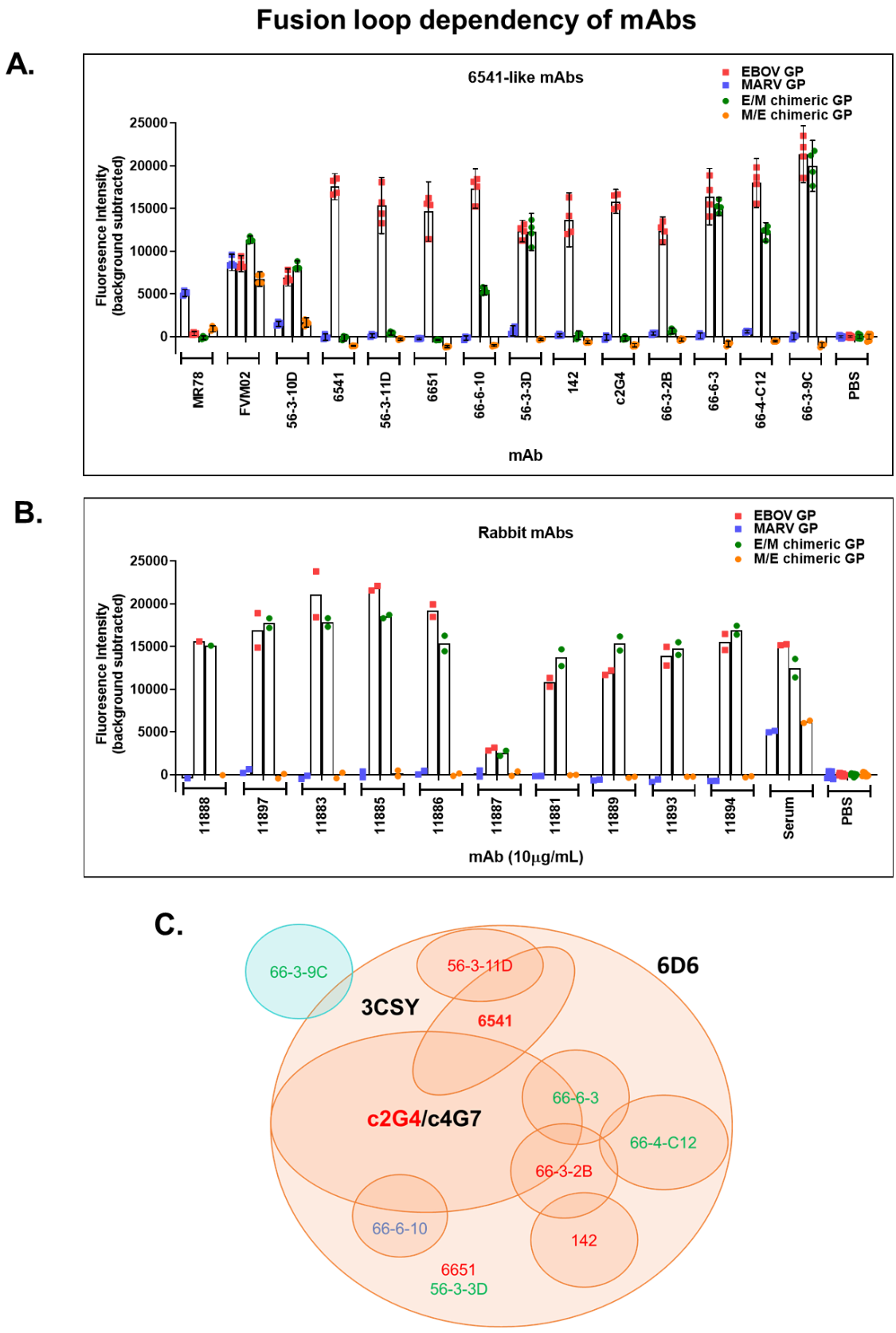


Figure 24. mAb binding to MARV GP, EBOV GP, E/M chimeric GP and M/E chimeric GP expressing cell lines*. Confluent cell monolayers were incubated with mAbs at 10 μg/mL except MR78 and FVM02 which were tested at 50μg/mL. Human mAbs and rabbit mAbs detected using anti-human IgG and anti-human Alexa Fluor 488 conjugates respectively. Fluorescence intensity measured using Clariostar (BMG Labtech).* ***A. GP base-binding humAbs panel.*** *Mean and 95% confidence intervals for 4 replicates in assay shown.* ***B. Rabbit GP mAb panel.*** *Mean of duplicates within assay shown, representative data shown from n=2 repeats of experiment.* ***C. Competing epitopes of GP base-binding humAbs defined by 3CSY/KZ52 competition, based on Rijal et al.*** *Overlapping shapes indicate competition. No data regarding 53-11D and c2G4 competition. Colour of text indicates change in binding to E/M chimera compared to EBOV GP in binding assay in A.: Black, not tested; Green, no loss of binding; Red, complete loss of binding; Blue, partial loss of binding.*

There is no reported competition group for 56-3-10D, however its epitope is not lost after thermolysin digestion suggesting it binds the base or core of the GP [3]. It showed a similar binding profile to FVM02 in this assay: binding to all the cell lines, but less strongly to the MARV GP and M/E GP, and without substantial loss of binding to the E/M chimeric GP. Apart form 56-3-10D and FVM02, none of the mAbs showed any binding to MARV and M/E chimeric GP cell lines.

The rabbit mAbs were also tested alongside serum from rabbit 378-1 two weeks after final vaccination as a positive control or detection of rabbit IgG in the assay. None of the rabbit mAbs showed loss of binding to the E/M chimeric GP in comparison to the EBOV GP (Figure 24B).

4.7 Discussion

4.7.1 Identification of broadly neutralising mAbs

Direct virus neutralisation is not the only mechanism of protection of mAbs against *Ebolavirus* GPs; however, the world-wide systematic analysis of EBOV GP mAbs conducted by the Viral Hemorrhagic Fever Immunotherapeutic Consortium (VIC) shows there is still a strong association between the ability to neutralise virus *in vitro* and protection in animal models, although certain pseudotype virus systems may be more easily neutralised than wild type viruses [173]. In addition, Ansuvimab (mAb114) and one of the components of Inmazeb (REGN-EB3) can neutralise pseudotype and wild type EBOV viruses *in vitro* and have shown efficacy in human clinical trials [45, 93, 182]. Therefore virus neutralisation is an important indicator for identifying potentially protective mAbs.

In vitro neutralisation of the EBOV S-FLU pseudotype virus used here has shown strong though incomplete correlation with protection against EBOV in a mouse challenge model [3]. No comparison has been conducted for neutralisation of the SUDV or BDBV S-FLU pseudotype, but CA45 is protective in SUDV challenge models [190] and was also neutralising when I tested it in the SUDV S-FLU *in vitro* neutralisation assay.

Previously, none of the human mAbs made as part of the Oxford/UCB collaboration were able to neutralise all three viruses tested, and the human mAb cocktail that showed protection of guinea pigs against EBOV is unable to fully neutralise SUDV and BDBV S-FLU *in vitro* at the concentrations I

tested, despite containing cross-reactive mAbs (Figure 9) [3]. Therefore the cross-neutralising rabbit mAbs identified in this study, especially 11886 for its breadth of neutralisation and 11883 for its strength of SUDV neutralisation, may represent alternative candidates for future pan-ebolavirus cocktail development in combination with the previously identified mAbs from vaccinees.

For the rabbit mAb panel reported here, the broadly reactive mAbs were also more likely to be neutralising. This supports the hypothesis that epitopes conserved across all GPs are likely to be conserved because they are involved in functions that are necessary for attachment and invasion of cells and so there is a selective pressure to maintain the sequences. Hence mAbs targeting these epitopes are more likely to be targeting sites of vulnerability on the GP that make them more likely to be able to neutralise the virus and prevent infection of cells. Of the cocktails of mAbs tested *in vivo* by Rijal *et al.*, the most cross-reactive cocktail was also the most protective [3].

A collaboration with Dr Thomas Strekker, Institut für Virolgie, Philipps-Universität Marburg, was initiated to test the ability of the rabbit mAbs to neutralise wild type *Ebolaviruses in vitro*, however, due to the COVID-19 pandemic, this work was paused and no data have yet been generated. Other planned future work includes testing neutralisation of other *Ebolaviruses* that have not yet caused large outbreaks of human disease; TAFV, RESTV and BOMV. Initial attempts to produce TAFV GP S-FLU resulted in extremely low titres of infectious virus, which may be rectified with further pseudotyping developments (see 4.7.4). BOMV GP and RESTV GP genes have been successfully cloned and will be used to transduce MDCK-SIAT-1 cell lines for both binding assays and S-FLU production. Whilst these viruses have not caused significant human disease yet, they represent the wider genus of emerging *Ebolaviruses* with potential to cause human disease and knowing whether current candidate mAbs are likely to be effective against these viruses is important for preparedness for the emergence of a new *Ebolavirus* into the human population.

4.7.2 Inhibition of GP processing for receptor binding as a mechanism of neutralisation

None of the rabbit mAbs were able to compete receptor binding directly, which suggests that binding to the cleaved GP and preventing receptor binding is not a mechanism of neutralisation by any of these mAbs – either by direct competition or via allosteric effects. This is not surprising as the receptor

binding region is largely occluded by the presence of the MLD, which was present in the full length GP used as the immunogen in the antibody discovery campaign (Section 3.2.1). In contrast, antibodies in the literature such as FVM04 and MR78, that do bind in the GP receptor binding site and compete NPC1 binding, were generated from vaccinations including MLD-deleted *Ebolavirus* GP proteins and a MARV survivor respectively [94, 213]. The removal of the MLD from *Ebolavirus* GPs exposes the RBR, and in MARV GP the MLD is more equatorial than in EBOV and is predicted to leave different areas of the GP exposed including the RBR [4].

That said, it may be possible for NPC1-C blocking mAbs to be generated in systems where full-length *Ebolavirus* GP is presented to the immune system. Antibody 6662 from the Oxford/UCB panel (generated from vaccinees who received a viral vectored vaccine that expressed full length GP) was able to compete NPC1-C binding to GP in a similar assay to that reported here (*Pramila Rijal, personal communication*). In addition, Ansuvimab (mAb114) was isolated from an ED survivor and binds to the GP at a vertical angle giving it access to the RBR despite the MLD, thereby allowing it to directly compete NPC1 for GP binding [186].

However, aside from direct or allosteric competition with NPC1-C, receptor binding can be inhibited by interfering with GP processing thereby preventing the NPC1-C binding site from being revealed. In assays assessing inhibition of thermolysin cleavage, antibodies 11894, 11897, 11883, 11886, 11889, 11892, but not 11878, were able to inhibit cleavage of EBOV GP to some degree. For mAbs 11883, 11886, 11889 and 11892 this extended to BDBV GP and to some extent SUDV GP. However, inhibition of thermolysin cleavage in these assays was not predictive of neutralisation in the pseudotype virus assays.

11886 could inhibit thermolysin cleavage and neutralise all three pseudotypes; however mAbs 11889, 11892 and 11883 showed inhibition of cleavage of all three GPs, but only neutralised one or two of the S-FLU viruses, and 11897 is non-neutralising and yet could inhibit EBOV GP cleavage. 11878, the only mAb that was not able to inhibit cleavage, was also non-neutralising.

Therefore, either these assays are not reflective of the conditions of cathepsin cleavage in the context of virus or pseudovirus infection, or inhibition of THL in these assays is a contributory mechanism, but not always sufficient to lead to direct viral neutralisation.

One explanation for this may be that inhibition of THL cleavage could be achieved in different ways, of which some are more efficient at preventing infection. Antibodies may interfere with the cleavage event itself – competing with proteases for the cleavage sites or causing conformational changes to the cleavage sites that make them less likely to be processed. Alternatively, antibodies may not prevent the cleavage event, but may prevent dissociation of the cleaved portions by retaining a binding footprint across the interface between cleaved domains. In the immunofluorescence assays these two scenarios would not be distinguishable as the antibody epitope would be retained in both cases when the mAb is bound to GP before THL treatment. However, in the immunoprecipitation experiments, one might expect to see that a mAb that only acts by holding the cleaved portions together would not necessarily retain the larger bands that represent uncleaved GP1, but may retain fragments of the size equivalent to the cleaved mucin-like domain and/or glycan cap in addition to the smaller bands representing the processed core GP1. There does not appear to be evidence of this in the immunoprecipitation experiments conducted by me or Lisa Schimanski in other experiments with a variety of antibodies. However, it would also be possible for a mAb to prevent THL cleavage of the GP by both mechanisms, which would also not be easily distinguishable from the immunoprecipitation blots alone.

A key factor that may make the assay more reflective of viral infection is pH. These assays were conducted using thermolysin at pH7.4, whilst during infection endosomal cathepsins are acting in a more acidic environment. The kinetics of the different proteases and their activities at different pHs may influence the ability of the mAbs to inhibit proteolytic cleavage of the GP in these different contexts, but this remains to be tested. It may also be that these assays do not distinguish which mAbs have the necessary kinetics and affinities to inhibit proteolytic cleavage in the context of infection, as mAbs were incubated with the GP for an hour before addition of thermolysin.

Of interest is the observation that mAb 11892 was able to bind cleaved GP at least as well as uncleaved GP in these assays, yet inhibited thermolysin cleavage as determined by the immunoprecipitation method. It is unlikely to have significant contacts in the GC or cleaved area itself and so how this mAb is able to stop thermolysin cleavage (which in this case did not translate to significant neutralisation potency) requires more exploration.

4.7.3 I260R as a mutation of concern

The I260R mutation has been identified by the Townsend Lab as of concern due to the loss of virus neutralisation by the majority of antibodies targeting the glycan cap, yet the virus containing the mutation remains viable and infection competent. Experiments by Lisa Schimanski show that the I260R mutation leads to much more rapid processing of the GP by thermolysin and more spontaneous GP cleavage. The I260 residue is located in the glycan cap itself, but is not proximal to the predicted thermolysin cleavage site that produces the primed GP_{CL}. In immunoprecipitation experiments using the I260R GP, pre-incubation with P6 fab does not prevent thermolysin cleavage of the GP (Lisa Schimanski, personal communication). This contrasts the strong retention of larger GP1 cleavage products when P6 fab is used to protect WT EBOV GP in the same experiments.

Faster processing of the GP may lead to the complete loss of GC epitopes before mAbs have had a chance to bind, although the fact that there is no significant change in binding to cells expressing I260R EBOV GP compared to WT EBOV GP does not support this. A more sensitive approach using flow cytometry may reveal changes in binding that are not detected by plate-based IFA. The relationship between the stability of the GP and the kinetics of each mAb may be elucidated by expressing soluble GPs for SPR analysis. Even if not leading to a complete loss of epitope in the first instance, the I260R mutation may reduce the kinetic window in which mAbs may act to prevent cleavage resulting in a loss of neutralisation.

In silico, I260R has previously been identified as leading to functional GP by combining molecular dynamics simulation snapshots with FoldX software to estimate the effect of point mutations on the stability of the GP by calculating ΔΔG (change in change in Gibbs free energy upon

folding) between the mutated and original GP sequences, [253]. This is supported experimentally by the infectivity of the I260R EBOV GP S-FLU pseudotype produced by Pramila Rijal.

To determine whether the I260R mutation has already occurred in isolates sequenced during outbreaks of EBOV, I interrogated the NCBI Virus Variation Ebolavirus resource (https://www.ncbi.nlm.nih.gov/genome/viruses/variation/, [254, 255]. 1802 protein sequences were retrieved (Protein, Species = *Zaire ebolavirus*, Host = any, Region/Country = any, Genome region = Spike Glycoprotein, Isolation source = any [Accessed: 2nd March 2021]) of which 1 contained the I260R mutation (AAC57989). This isolate was identified by Volchkov *et al*. as the causative agent of an outbreak of ED in Gabon in 1994 with a 60% mortality rate; the I260R mutation was identified in the Gabon-94 isolate, but was not present in sequences from isolates from the 1976 or Kikwit 1995 outbreaks [256]. This shows that the I260R mutant is transmissible between humans, causes severe disease and may still be circulating in animal reservoirs.

My search of the Nextstrain database of 1493 EBOV genome sequences sampled between March 2014 and October 2015 in West Africa in the largest outbreak of Ebola disease (ED) to date, does not indicate the emergence of the variant in the field during that time of high viral transmission, [https://nextstrain.org/ebola, Accessed: 2nd March 2021] [257] nor did a search of papers reporting on the emergence of several different viral lineages during the outbreak [258-261].

A key question surrounding this mutation is whether currently approved antibody-based therapeutics for EBOV infection are affected. Ansuvimab (mAb114) would likely be unaffected as, although it contacts the GC along with the GP core, the GC contacts are dispensable for binding [186]. However, one of the components of the Inmazeb mAb cocktail binds to the GC. This mAb cannot neutralise WT *Ebolavirus in vitro,* but was partially protective in guinea pig *in vivo* studies as a monotherapy and fully protective as part of the cocktail [182]. It remains to be tested whether the I260R mutation affects *in vivo* protection offered by non-neutralising GC mAbs.

Also of concern is whether people receiving current vaccines which are largely based on the Mayinga 1976 or Makona 2014 isolates that do not contain the I260R mutant and likely have not been

tested against the Gabon-94 isolate, or those who have naturally acquired immunity (outside of survivors of the Gabon 1994 outbreak) that have a neutralising serum response, retain neutralisation of this variant. This would be of concern especially if there was any indication of the GC being immunodominant in terms of antibody responses to vaccines or infection, as then a large portion of humoral immunity could be lost towards the I260R mutant if it was transmitted into the human population again. Access to vaccinee serum from recent or ongoing Ebola vaccine trials could be possible through collaboration with the Jenner Institute or EBOVAC trial consortium and could be readily tested in S-FLU neutralisation assays.

Worryingly, I260 is conserved across EBOV, SUDV, BDBV and TAFV; thus this mutation could potentially affect strategies to protect against all disease causing *Ebolavirus* species. We have not yet tested the effect of the same mutation in different species, but it could be achieved simply with the experimental systems to hand.

This mutation could therefore have large implications for both future therapeutic mAb development and rational vaccine design as we continue to prepare for *Ebolavirus* variants and species.

As 11886 retains neutralisation of the I260R EBOV GP S-FLU, this supports further progression of this mAb in *in vivo* experiments. The fact that it appears to have a thermolysin sensitive epitope, yet retains neutralisation of the I260R mutation may mean that elucidation of the epitope and mechanism of neutralisation by this mAb could help explain why the I260R mutation causes the loss of neutralisation by other mAbs with thermolysin sensitive epitopes, but not this mAb.

4.7.4 Use of chimeric GPs and synthetic fusion loop peptide to identify fusion loop mAbs

The aim of this section was to develop assays to help quickly screen for mAbs that may impact viral infection by interfering with the functioning of the GP fusion loop.

Whilst easily generated and incorporated into a high throughput assay set up, the synthetic fusion loop peptide (FP) ELISA has distinct limitations. Namely, that a short peptide is neither representative either of the whole FL outside of the tip, which we know contains the epitopes of multiple broadly-neutralising mAbs, nor of the quaternary nature of the FL location on the trimeric

protein. 6D6 for example whilst having been shown to bind to the tip of the FL, was reported not to bind to a synthetic FP [209], and has many contacts on the neighbouring GP1-GP2 monomer [10]. This is clearly reflected in the much greater number of mAbs that were detected as having an epitope involving the FL using the chimeric GPs than in the synthetic FP ELISA in Section 4.6.

However, the synthetic FP ELISA was positive for FVM02. FVM02 has been shown to bind linear overlapping peptides corresponding to the FL tip residues 526-535 [210]. Therefore the assay does work to detect mAbs that can bind the tip of the FL without requiring contacts on the rest of the GP. However, are these even smaller subset of mAbs worth identifying? Unlike many of the other mAbs that bind the FL that have conformational epitopes spanning other parts of the GP, FVM02 is not neutralising *in vitro*, but when given to mice in a 2 dose regimen 0 and 3 days post infection at 25mg/kg, it was partially protective against death [210]. Therefore FVM02-like mAbs are of interest as potentially protective.

The chimeric GPs overcome many of the limitations of the synthetic FP by retaining the structures around the FL and including the entire FL. However, swapping the FL may have altered the conformation of the GP, potentially by affecting packing of the FL against the rest of the GP. The experiments conducted with MR78, NPC1-C and non-FL mAbs suggest that the global fold of the E/M chimera is correct as both thermolysin processing of the GP was able to occur and the receptor binding region, GC and other base epitopes are in an appropriate conformation for receptor and mAb binding.

From the sorting performed with FVM02 (Figure 23A) a low level of M/E chimeric GP expression was not anticipated, but neither MR78 nor NPC1-C could bind the thermolysin cleaved M/E chimera although they both bound MARV GP in the same assay (Figure 23B). The lack of signal from M/E chimeric GP expressing cells suggests that the protein is in fact not expressed at the surface or is incorrectly folded in such a way that prevents processing by thermolysin and/or receptor and MR78 binding. The latter is consistent with successful FVM02 staining of the cells for the sort and in IFA experiments, as FVM02 can bind a linear epitope and denatured protein [210]. A repeat of the transduction or cell sorting with a conformational non-FL MARV GP mAb may resolve the issue with

this chimeric GP. Such a mAb could also act as a better internal positive control for future screening experiments using MARV GP or M/E chimeric GPs as MR78 repeatedly performed poorly in my IFA experiments with uncleaved proteins. FVM02 cross-reacts with MARV GP due to the conservation of the FP, but as discussed, the linear nature of its epitope makes it a poor control for the correct expression of these full length GPs.

Loss of binding to an EBOV GP with MARV FL is reported for mAb 6D6, although details of the design of both chimeric GP and binding assay are not published [209]. Apart from this, I am unaware of other reports of attempts to make similar chimeric GPs.

66-3-2B was the only mAb identified in either the Oxford/UCB humAb panel or new rabbit mAb panel that bound the synthetic FP; it also lost binding to the E/M chimeric GP – adding credence to its identification as a FL binding mAb. This also suggests that compared to FVM02, 66-3-2B binding is more dependent on the residues in the FP that are not conserved between MARV and the *Ebolaviruses*. 66-3-2B is a partially cross-reactive and neutralising mAb [3], which supports the idea that identification of these FL tip mAbs is a strategy to find candidate therapeutic mAbs. However, the much larger number of mAbs identified using the chimeric GP assay supports the use of these as complementary assays for finding FL mAbs.

The robust nature of the adherent MDCK SIAT-1 cells means that they are useful reagents for high throughput screening assays. An IFA-type process could be conducted in high throughput set ups using systems like those used in Section 3.3.2. Another advantage of making the chimeric GP cell lines is their use in producing pseudotype S-FLU viruses to study the effect on antibody-mediated neutralisation due to the presence or absence of the EBOV GP FL; however the attempt to pseudotype S-FLU with these chimeric GPs failed. This is unsurprising for the M/E GP cell line considering the lack of conformational GP on the surface of cells, however given that the E/M GP appeared to express well and be processed correctly up to the point of receptor binding, this result was disappointing. It is possible that the next steps in viral infection, namely the rearrangement of the GP to insert the FP into the host endosomal membrane and fusion with the viral membrane, were impeded by the MARV FL in the context of the EBOV GP. However, we have previously seen that some *Ebolavirus* GPs, such as

TAFV GP, do not produce high enough titres of infectious S-FLU virions to conduct neutralisation assays despite having functional GP. In addition, in comparison to hemagglutinin (HA) pseudotyped viruses, *Filovirus* GP S-FLUs are generally lower in titre and form diffuse plaques [219]. Recently, as part of her doctoral work, Holly Sadler in the Townsend lab has developed a next generation S-FLU system named CLEARFLU. CLEARFLU HA is designed with a series of mutations that prevent it being competent at fusion, but still allow it to mediate cell attachment via sialic acid binding. When grown in cells transduced with both the GP and the CLEARFLU HA version 3, EBOV GP S-FLU titres were higher and plaques more defined than when grown in GP expressing cells alone, suggesting the sialic acid binding provided by the CLEARFLU HA improved the cell attachment of progeny virions in the permissive cell line and hence titres of the GP pseudotyped virus (*Holly Sadler, personal communication).* Repeating the chimeric GP pseudotyping experiments in this new system may improve titres to levels that would be useful for neutralisation assays.

4.7.5 Summary

In this Chapter I began to characterise *in vitro* the panel of antibodies generated in Chapter 3 with a focus on the broadly reactive antibodies that could bind all four GPs tested. The assays in this chapter concentrated on neutralisation of GP S-FLU pseudotype viruses and the mechanisms by which this neutralisation might be occurring. Assays to determine whether the mAbs could inhibit thermolysin processing of the GP suggested that this is a mechanism by which many of the mAbs were able to affect the GP, even if this was not always sufficient for neutralisation. None of the mAbs directly competed for receptor binding but, by inhibiting processing of the GP required to reveal the NPC1-C binding site, they may be preventing receptor binding indirectly. As inhibition of GP processing seemed to be a feature of many of the mAbs tested, the loss of neutralisation by mAbs with thermolysin sensitive epitopes to a variant EBOV GP S-FLU containing the I260R mutation is concerning, especially as this variant has been detected in the 1994 Gabon outbreak isolate. None of the new rabbit mAbs, but some of the base-binding human mAbs were identified as having a fusion loop dependency by binding to a synthetic fusion peptide and loss of binding to a chimeric EBOV GP with a MARV fusion loop. Both of these assays could be used in a high throughput setting to screen for fusion loop

antibodies during antibody discovery. Antibodies that bind the fusion loop may act to inhibit infection by preventing insertion of the fusion peptide into the host endosomal membrane. Assays to look more directly at inhibition of membrane fusion are required to determine if the mAbs are acting at this step of infection. Other mechanisms of neutralisation that could be investigated include inhibition of attachment to host cells before uptake into endosomes, and holding the GP in a pre-fusion conformation preventing the global GP rearrangements necessary for infection once in the endosome. The latter could be assessed in FRET assays as described by Durham *et al*. [262].

Aside from direct neutralisation, it is known that mAbs can be protective against EBOV infection by engaging a range of Fc effector functions [173, 174]. It is hoped that a project to assess the ability of the mAbs isolated in this thesis to engage complement and to recruit NK cells and phagocytic cells will be initiated in the near future.

Alongside insights into mechanisms of neutralisation, the assays conducted in this chapter provide evidence for initial epitope mapping of the mAb panel and this is continued in more detail in Chapter 5, with a particular focus on mAb 11886. The epitope of this mAb is of particular interest as in this work it was identified as a rare broadly-neutralising mAb and also retained neutralisation of the I260R variant despite having an epitope that was sensitive to thermolysin cleavage.

5 Defining broadly reactive epitopes

5.1 Introduction

As reviewed in Chapter 1 (Section 1.4.3.2) several sites on the GP have been identified as epitopes for broadly neutralising and/or protective mAbs: the NPC1-C binding site; the base, length and tip of the GP2 fusion loop; the cryptic 3_{10} pocket on GP1; and the GC [99, 100, 115, 190, 191, 193, 209, 210, 212-214]. It is important to define the epitopes of the broadly reactive mAbs of interest generated in Chapter 3 for many reasons: to define conserved sites of vulnerability on the GP that might be the cross-protective target of inhibitors or vaccines; to inform hypotheses about mechanism of function of neutralising mAbs; to allow comparison with other mAbs to inform the design of therapeutic mAb cocktails for *in vivo* studies; to inform hypotheses for rational mutation of mAbs to improve their efficacy; to define potential escape mutations of concern; or conversely, to find alternative mAbs in the event of viral escape.

From the data presented in Chapter 4, some information on the epitopes of the mAb panel has already been gained. The cell-based immunofluorescence assay (IFA) used to determine if the mAbs could inhibit thermolysin (THL) cleavage of the GP, also provided information on the THL-sensitivity of the mAb epitopes on EBOV GP (Section 4.3): mAbs that lost binding to the cleaved GP (GP_{CL}) likely bind the GC or MLD removed by cleavage, whereas those that retained binding likely have significant contacts in the GP1 core and/or the base including GP2. None of the mAbs were able to directly compete NPC1-C for binding to GP_{CL} in a competition IFA (Section 4.4.1) which suggests that none of the mAbs bind to the receptor binding region in the GP1 core. The loss of neutralisation of the I260R EBOV GP S-FLU virus is a feature of almost all GC mAbs previously tested by Dr Pramila Rijal (*personal communication*); for those mAbs I tested in Section 4.5, that were able to neutralise EBOV GP S-FLU, but not I260R EBOV GP S-FLU, this is a strong indicator that they are GC mAbs. Lastly, the assays used to identify fusion loop (FL) targeting mAbs (Section 4.6) showed that none of the rabbit mAbs tested were able to recognise a peptide representing the FL tip nor lost binding to the chimeric EBOV GP with residues 512-555 swapped for the FL of MARV GP. As the majority of the swapped

residues are unconserved, this suggests that the mAbs do not rely on contacts in the FL for binding to GP.

The epitopes of all the broadly reactive mAbs are of interest as non-neutralising mAbs may still be protective *in vivo*, and antibody engineering technologies may improve the efficacy of mAbs that are not currently broadly neutralising. However, with the data to hand, the mAbs of most interest within the panel are currently 11886, which was defined as broadly neutralising, and 11883 which was cross-neutralising and especially potent against SUDV GP S-FLU (Section 4.2). In Chapter 4, these mAbs showed similar binding profiles as both are THL-sensitive, but if their epitopes are distinct they offer a sensible start for rational cocktail design within the panel. In this Chapter, epitope mapping studies were conducted for the entire panel with higher resolution mapping studies focusing on the epitope of mAb 11886.

5.2 Defining epitopes on EBOV GP by competition group

An immunofluorescence experiment was used to assign a GP antibody of interest to a broad epitope bin by competing antibodies of known epitope for binding to EBOV GP expressed on the surface of adherent MDCK-SIAT 1 cells. Panels of antibodies including >1 mAb for each epitope bin were used to confirm competition and place new antibodies of interest in one of four epitope bins; base, GC, RBR or FL (Figure 25). In addition, the rabbit mAbs were competed against each other. Each antibody was chemically biotinylated and then mixed with a 10-fold excess of unbiotinylated antibody before incubation with GP. Amount of biotinylated antibody remaining after washing was determined using a streptavidin Alexa Fluor 647 conjugate and compared to controls to determine degree of competition.

Antibodies c13C6, c4G7, c2G4, CA45, 6D6, FVM02, ADI-15946, ADI-15878 have epitopes published in the literature by other groups. Antibodies 6541, 66-4-C12, 6660, 6662, 66-3-9C, 66-3-2C and 040 are fully human antibodies from the Oxford/UCB humAb panel with epitopes determined via multiple assays as previously described [3]. Antibodies 21-D8-5A (influenza neuraminidase-specific mAb) and R5.014 (malaria RH5-specific mAb) were used as negative controls.

Epitope binning of EBOV GP mAbs using competition immunofluorescence assay

A.i.

BIOTINYLATED ANTIBODY	UNBIOTINYLATED ANTIBODY																			
							Fusion Loop					Base				RBR			GC	
	11897	11883	11889	11886	11892	ADI-15946	ADI-15787	CA45	6D6	FVM02	6541	c4G7	c2G4	66-4-C12	6660	6662	13C6	66-3-2C	040	66-3-9C
11897		0.15	0.36	1.81	1.70	1.93	1.29	1.20	0.64	1.03	1.53	1.43	1.27	0.64	0.07	0.08	0.37	0.09	-0.2	1.08
11883	0.08		0.83	1.41	1.47	1.55	1.07	0.91	0.77	0.90	1.54	1.29	1.35	0.99	-0.07	-0.01	0.17	0.57	0.012	1.14
11889	-0.08	0.05		0.87	1.04	1.80	0.28	0.97	1.05	0.96	2.61	2.10	2.12	1.33	-0.08	0.01	0.50	-0.14	-0.16	1.59
11894							0.14	1.09	1.08	1.02	1.91	1.27	1.60	1.47	1.12	1.54	0.41	-0.19	-0.22	0.97
11886	3.62	2.49	1.79		-0.09	-0.14	0.47	0.77	0.90	1.16	-0.25	-0.27	-0.15	0.60	0.63	0.63	0.01	0.76	0.983	2.00
11892	1.33	1.53	0.83	0.35		-0.24	0.65	0.65	0.72	0.98	-0.09	-0.10	0.01	0.72	0.89	0.91	0.09	0.93	1.346	1.78
ADI-15946	1.66	1.71	1.28	0.33	0.30															

ii.

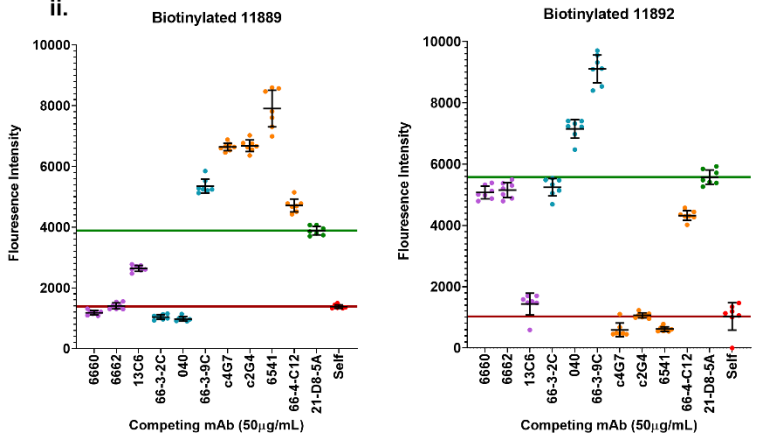


Figure 25. Competition of mAbs for binding to EBOV GP (Mayinga, NP_066246.1) expressed on cells. A.i. Summary table. *1 = signal when biotinylate body incubated with non-GP mAb or PBS (max), 0 = signal when competed against self (min). Values calculated using formula: (max-X)/(max-min), where X = mean signal of technical replicates for a given competition pair. Table coloured on scale from red (0, strong competition) to green (≥1 no competition). Blue; self-competition. Grey; no data. RBR; receptor binding region. GC; glycan cap. Background was subtracted from all data before calculation of competition.* ***ii. Example data used to generate table in part i. showing contrasting competition patterns.*** *Competition of biotinylated 11889 and 11892 with the panel of humAbs used to define epitope bins. Mean and 95% CI of 6 technical replicates within assay shown, background (mean of 12 wells per assay plate) subtracted. Data points coloured by epitope bin of competing mAb. RBR; purple. GC; blue. FL; yellow, Base; orange. Irrelevant mAb for no competition control, 12-D8-5A; green. Self-competition; red. Green and red lines represent mean of no competition and self-competition respectively. Biotinylated antibody detected using streptavidin Alexa Fluor 647 conjugate. Fluorescence intensity measured using Clariostar, BMG Labtech.* ***Figure continued on next page.***

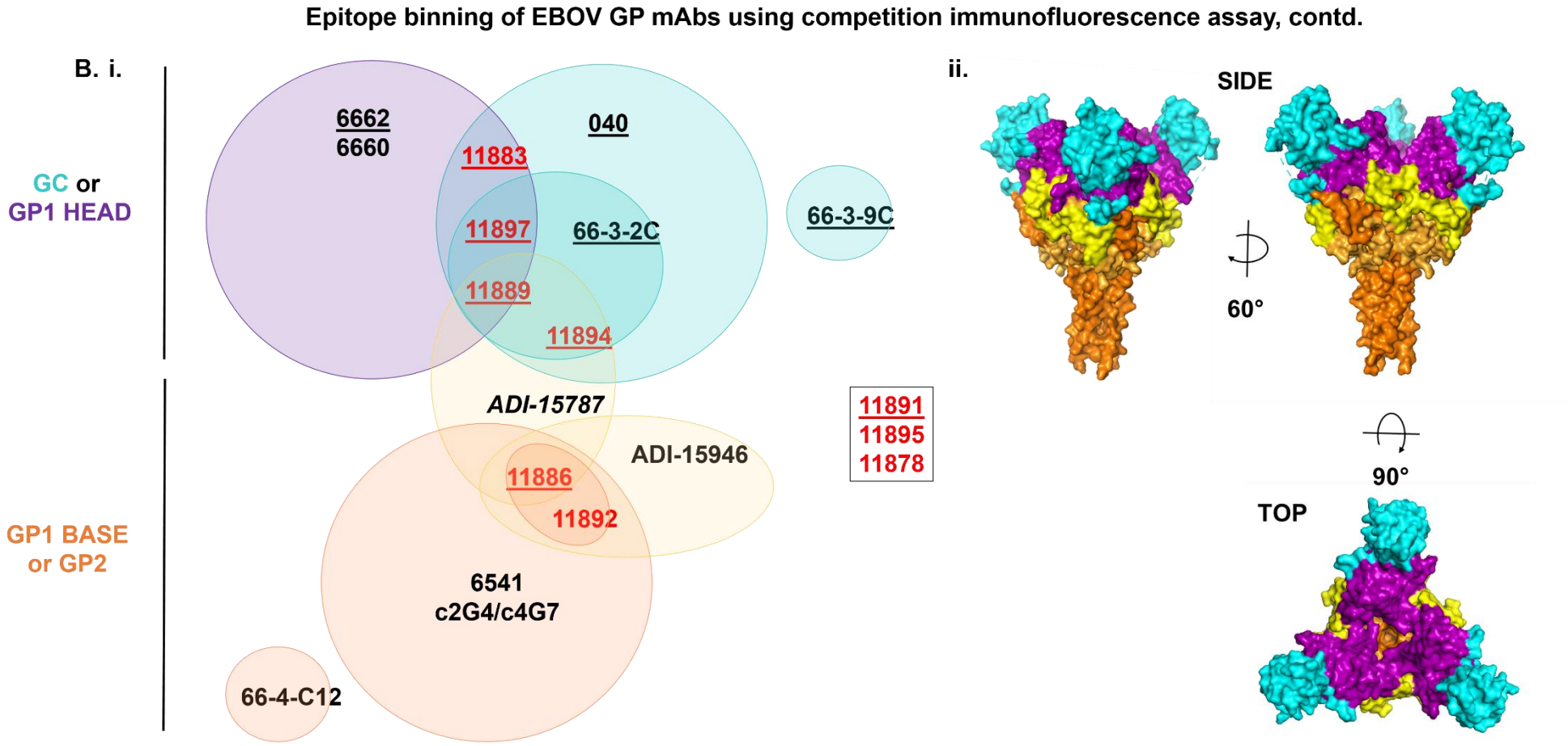


Figure 25 contd. B.i. Competition bins of mAbs based on summary table in A. *mAbs were defined as competing if have a completion value of <0.5. i.e. incubation with competing mAb reduced binding of biotinylated mAb by at least half. mAbs enclosed in the same shape are competing. Shapes are coloured by broad epitope bin as defined in part A. Red; rabbit mAb. Black; human mAb. Underlined; Epitope lost if GP treated with thermolysin (THL). Italics; THL sensitivity not tested. Box; not tested in competition assay. ADI-15878 and ADI-15946 do not compete each other [2], but no data to determine if they compete the other human mAbs. Competition between humAbs and THL sensitivity for humAbs 66-3-2C, 66-3-9C, 6662, 6660, 66-4-C12, c2G4, and c4G7 determined previously by Pramila Rijal [3].* ***ii. EM structure of EBOV GP (without mucin-like domain) coloured by epitope bin.*** *PDB: 5JQ7 [6]. Recoloured in Pymol using domains as defined by Lee et al. [1]. RBR and GP1 head; purple. GC; blue. FL; yellow, GP1 base; light orange. GP2; dark orange.*

Antibodies 6541, 66-4-C12, c2G4 and c4G7 represent different overlapping base epitopes, including some contacts with the FL. 6D6, FVM02, and ADI-15878 bind to tip of the conserved FL. CA45 and ADI-15946 binds to the stem and base of the FL with overlapping footprints. Antibodies 6660 and 6662 are designated as RBR mAbs. 66-3-9C, 66-3-2C, 040 and c13C6 represent different GC epitopes. 66-3-9C binds the β17-18 loop, an epitope also described for the mAb FVM09 [192].

It is noted all antibodies tested competed with c13C6, irrespective of the other mAbs they competed in the panel, hence any apparent competition with c13C6 is likely not meaningful. Similarly none of the mAbs competed 66-3-9C in this assay, suggesting none directly bind to the β17-18 loop.

11897 and 11883 show highly similar competition patterns and compete each other in both orientations. They strongly compete the RBR and GC mAbs, but not the base or FL mAbs.

11889 also competes the RBR and GC mAbs including 11897 in both orientations, and 11883 in one orientation. It does not compete any base or FL mAb except for ADI-15878. 11883, 11897 and 11889 may not directly contact the RBR/GP1 head, but at the very least their angle of binding hinders access of mAbs 6660 and 6662 to the GP.

11894 is not broadly reactive, but was assessed in this assay as it was used as a control mAb in other epitope mapping experiments (see Section 5.5). It competes only the GC mAbs and not the RBR, FL or other base mAbs, except ADI-15878. This suggests it is a GC mAb that approaches the GP from an angle that does not block access to RBR/ GP1 head epitopes. Competition with the other rabbit mAbs is yet to be tested.

11886 competes only base mAbs 6541, c2G4, c4G7 and ADI-15946. It partially competes FL tip mAb ADI-15878. This would suggest a base epitope. However, 11886 binding to GP appears to be THL sensitive. This suggests an intermediate epitope or angle of binding spanning the epitope bins; one that covers the base, but not the tip, of the fusion loop. ADI-15787 and ADI-15946 do not compete each other, but have epitopes on different sides of each GP1,2 monomer on the base of the GP chalice across the tip and base of the fusion loop respectively. It is therefore feasible for a mAb such as 11886

to bind between the two; to have an epitope adjacent to the base of the fusion loop, thereby competing base and FL base mAbs such as ADI-15946, but which also allows it to hinder access of a mAb such as ADI-15787 to the tip of the FL which is part of the neighbouring GP1,2 monomer depending on its angle of approach. This is consistent with the lack of competition with CA45, which binds further along the fusion loop base, i.e. on the other side of ADI-15946. The fact that GC mAbs 11894 and 11889 also compete ADI-15878, but not other FL tip binding mAbs, suggests that the angle of ADI-15878 binding to the GP is different to the other FL tip mAbs and that, whilst the footprint of the mAb on the GP is across the FL tip, the rest of the mAb is indeed angled in such a way to affect mAbs higher up the GP.

11892 competes 11886, the base mAbs 6541, c2G4, c4G7 and ADI-15946 strongly, but none of the other fusion loop, RBR or GC mAbs. Consistent with it having a true base epitope, 11892 can bind to GP after THL cleavage.

11892 and 11886 epitopes appear to be distinguished by the slightly stronger competition of ADI-15878 by 11886, although to properly dissect how their epitopes overlap requires repeats including the opposite orientation of biotinylated antibody.

11891 and 11895 do not bind to EBOV GP and limited availability of 11878 due to precipitation problems during expression prevented assessment in this assay.

5.3 Binding to thermolysin-cleaved GP in immunofluorescence assay

As described in Section 4.1.2.1, processing of the GP by endosomal cathepsins to remove the MLD and GC can be mimicked by cleavage catalysed by the protease thermolysin (THL). As an extension of the assay in Section 4.3.1, the binding of each mAb to cleaved GP from SUDV, BDBV and TAFV was tested (Figure 26), and repeated for EBOV (data not shown).

This allowed mapping of additional mAbs of interest outside of the broadly reactive panel: 11891 (SUDV GP specific) and 11895 (does not bind to EBOV GP). 11894 is an EBOV GP specific mAb and was included as a negative control. 11891 loses binding to the cleaved SUDV GP, suggesting it has

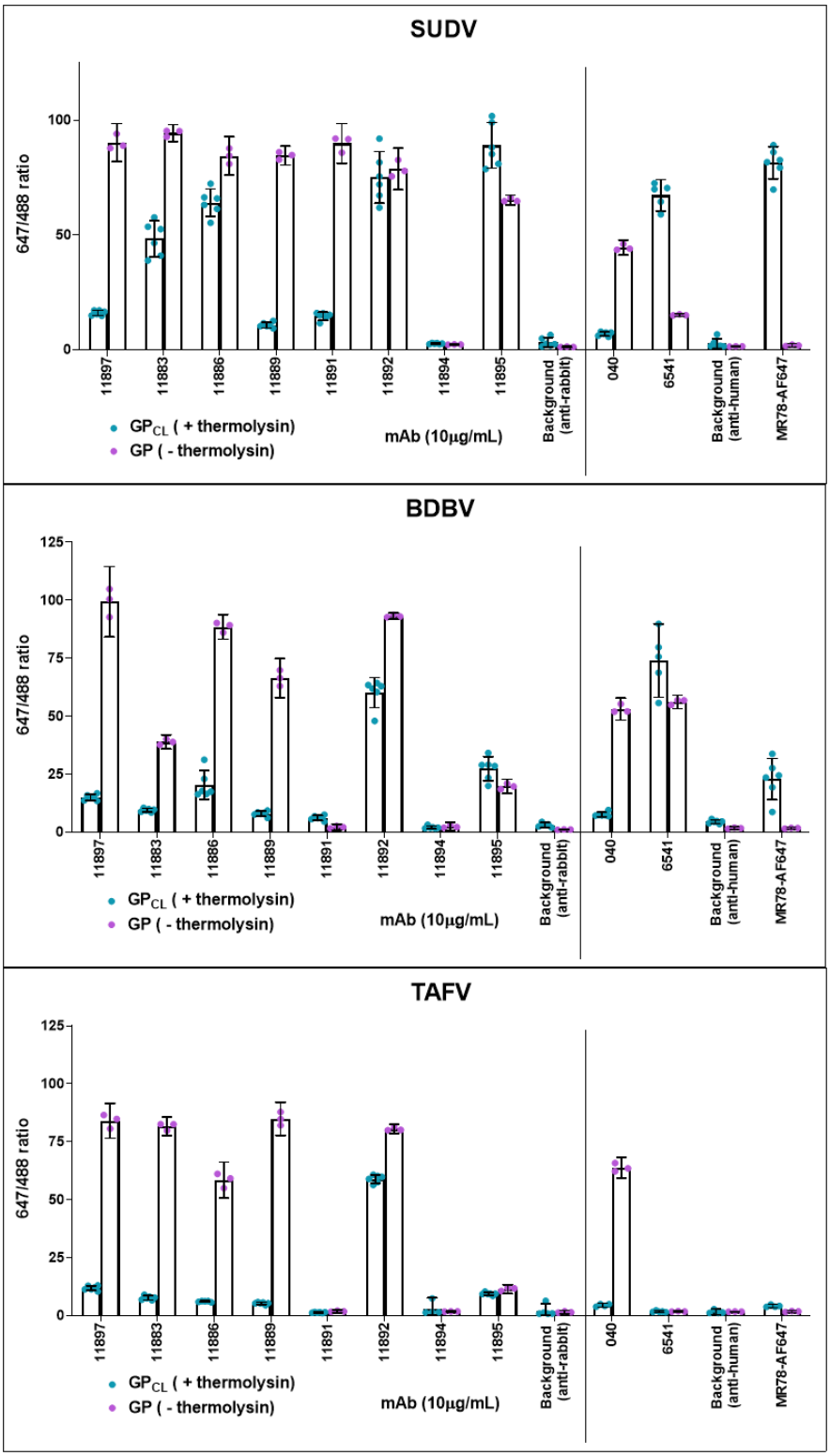


Figure 26. Binding of rabbit mAbs to GP and GP$_{CL}$ from SUDV (YP_138523.1), BDBV (YP_003815435.1) and TAFV (YP_003815426.1) expressed on MDCK SIAT-1 cells. *n=1 for BDBV and TAFV experiments, representative experiment shown for SUDV GP (n=2). Mean and 95% CI for six technical replicates for GP$_{CL}$ and three for GP. All cells were stained with wheat germ agglutinin Alexa Fluor 488 (Invitrogen, W11261). Fluorescence intensity was read at 625-30/680-30 and 488-14/535-30 using a Clariostar plate reader (BMG Labtech). Gain was adjusted to give a ratio of approximately 1 in assay plate wells containing cells that had been stained with both 488 and 647 conjugates only. Cells were excluded from analysis if 488 signal indicated too few cells in remaining in well. MR78 directly conjugated to Alexa Fluor 647 was used to confirm digestion by THL.*

an MLD or GC epitope. 11895 retains binding to GP_{CL} across SUDV, BDBV and TAFV, suggesting a base epitope in GP2 or the GP1 core.

Consistent with their binding patterns to EBOV GP_{CL}, the broadly reactive mAbs 11897 and 11889 do not bind GP_{CL}, whereas 11892 retains binding to GP_{CL} for all GPs tested. However, despite having an epitope conserved across GPs and showing essentially no binding to GP_{CL} from EBOV, BDBV and TAFV species, binding of 11886 to SUDV GP_{CL} was detectable and reproduced in a repeat of the experiment. Although still reduced compared to binding the uncleaved SUDV GP, 11886 binding to SUDV GP appears to be less reliant on the presence of the domains removed by THL cleavage. The effect seen of 11883 binding to cleaved SUDV GP_{CL} was less clear in repeated experiments.

040 (GC) and 6541 (base) were included as broadly reactive controls, but as previously shown (see Section 3.4.2), 6541 does not bind to TAFV GP. MR78 directly conjugated to Alexa Fluor 647 was used to confirm digestion by THL as previously discussed in Section 4.3.1.

5.4 Immunoprecipitation blots to identify epitope bins

The THL cleavage products of GP arise processively over time or with increasing concentrations of THL. These digestion products can be immunoprecipitated and analysed via western blot. In Section 4.3.2 this was used to assess how pre-incubation of the GP with Fabs of interest protected from THL cleavage when immunoprecipitated with a core/base mAb. In this section, GP was instead digested then precipitated using the mAbs of interest in order to determine which cleavage products are recognised by each mAb (Figure 27A).

Briefly, surface proteins of MDCK SIAT-1 cells expressing GP from EBOV were biotinylated, cells were then treated with increasing concentrations of THL to process the biotinylated GP. Cells were lysed in the presence of protease inhibitors and lysate incubated with Protein A Sepharose and mAb of interest. Protein A Sepharose was washed and bound protein eluted and separated using SDS-PAGE electrophoresis. A western blot was conducted using a streptavidin AlexaFluor 647 conjugate and fluorescent bands imaged. The digestion products seen in this assay are discussed in detail in Section 4.3.2, and here summarised in Figure 27B.

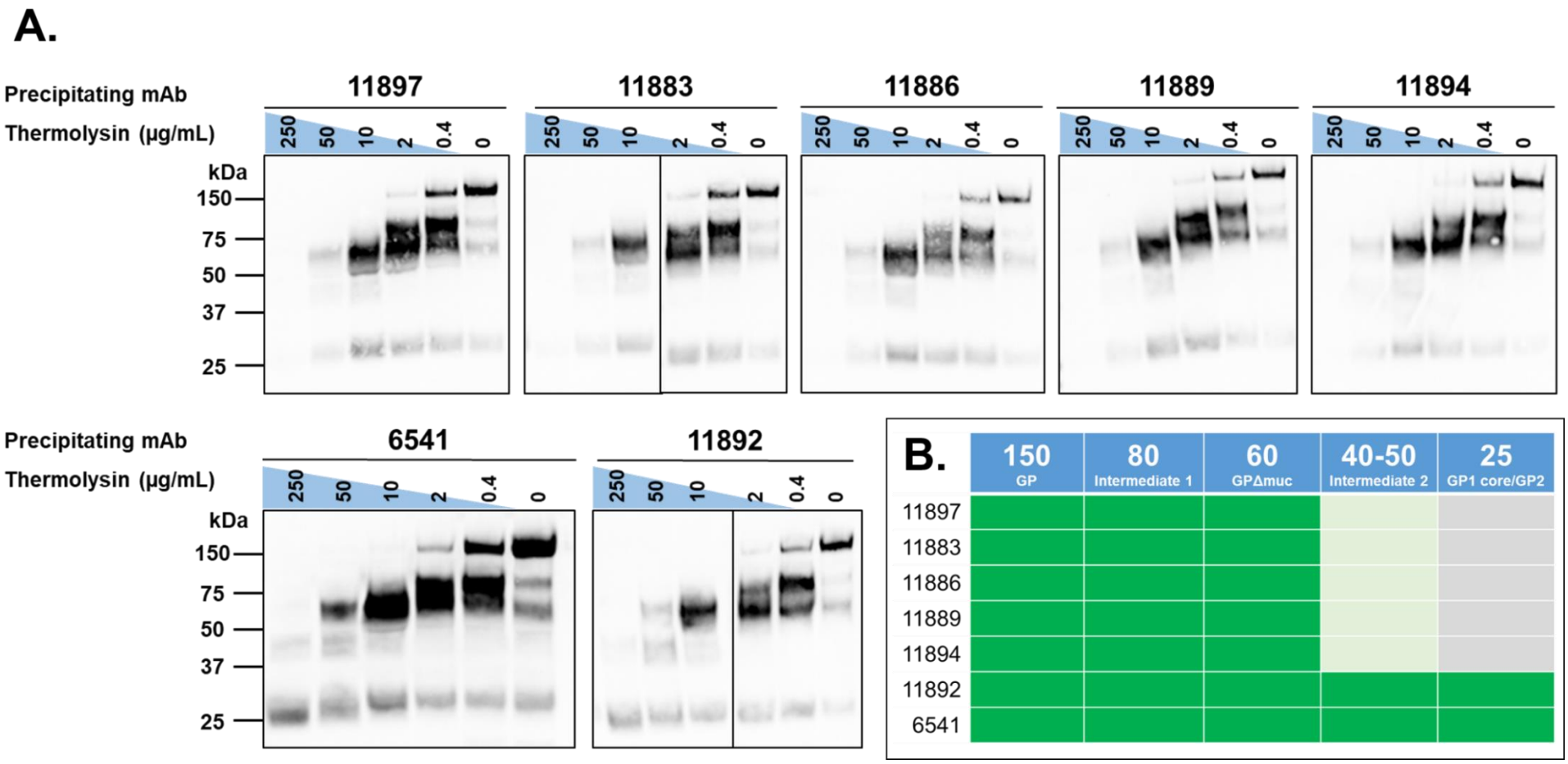


Figure 27. Immunoprecipitation of thermolysin cleaved EBOV GP by broadly reactive rabbit mAbs. A. Immunoprecipitation blots. *Biotinylated surface proteins of EBOV GP expressing MDCK SIAT-1 cells were treated with increasing concentrations of THL, then immunoprecipitated using Protein A Sepharose and anti-GP mAb of interest. Bands revealed using Streptavidin AlexaFluor-647 conjugate and imaged using iBright FL1000. 6541 blot was produced as part of a separate experiment and is included here to show comparison with a base binding mAb. All other blots were conducted as part of the same experiment with digested cells treated with each concentration of THL pooled before precipitating with individual mAbs.* ***B. Table summarising whether mAbs can still precipitate GP fragment after the loss of the previous fragment****. Sizes of indicated fragments are approximate kDa as estimated from blots. Green; strong binding. Light green; weaker binding. Grey; unable to precipitate band after loss of previous fragment.*

None of the mAbs lose the ability to precipitate the GP after the loss of the MLD (loss of ~150kDa and ~80kDa bands) suggesting that none rely on the MLD for binding to GP. 11897, 11883, 11886, 11889 and 11894 can no longer precipitate the GP after the complete loss of the ~60kDa band suggesting they have a large GC component to their epitopes. They can all precipitate the ~40-50kDa intermediate to some extent, suggesting that their epitopes span the two parts of the GP cleaved successively by THL to remove the GC. Like the base mAb 6541, 11892 can precipitate the ~25kDa band after the loss of the GC and other GP intermediates. This suggests an epitope in the GP1 core and/or GP2. These results are consistent with the immunofluorescence assays showing loss of binding to cleaved EBOV GP for all mAbs tested here except 11892 (see Section 4.3.1).

5.5 Epitope mapping using random combinatorial linear peptide phage display and next generation sequencing

Libraries of peptides displayed on the surface of bacteriophages can be used to agnostically identify partial or linear epitopes of certain antibodies via cycles of panning, elution and library expansion (Figure 28). Use of random combinatorial peptide libraries representing a high proportion of total possible sequence diversity facilitates the identification of epitopes of mAbs that can bind partial or linear epitopes without the need to generate libraries of peptides representing a specific antigen. Multiple antibodies can be panned in parallel making it a useful method for rapidly gaining insight into

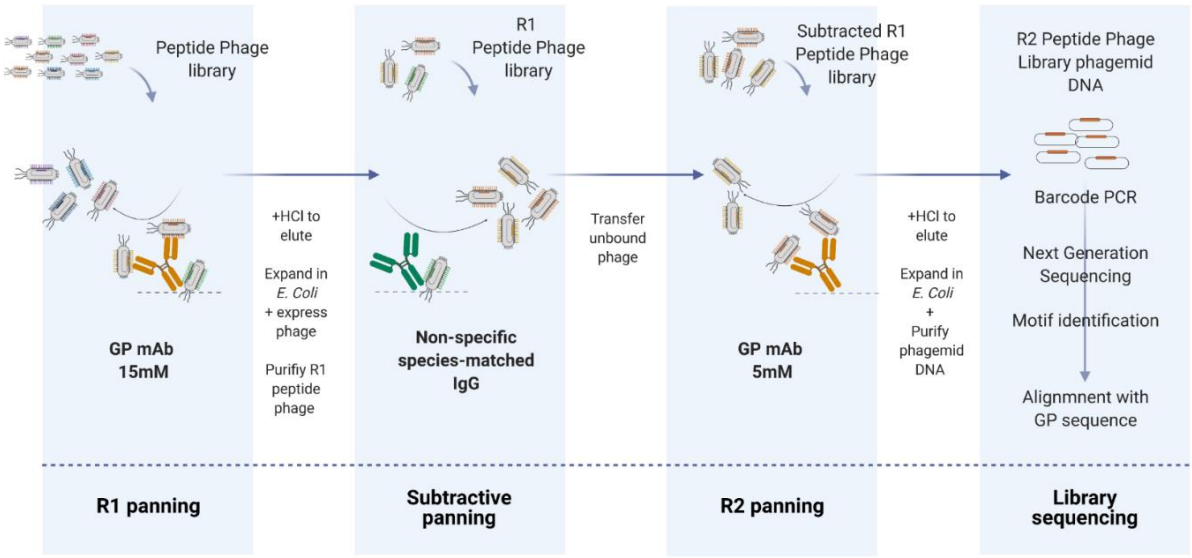


Figure 28. Peptide phage display strategy for identifying linear or partial epitopes of GP mAbs. *Figure drawn using BioRender.com.*

epitopes of small-medium panels of mAbs before more expensive and time-consuming structural methods are utilised. However, antibodies with largely conformational epitopes or epitopes where important contact residues are not consecutive within a linear stretch of the antigen sequence cannot be identified using this method.

These experiments utilised two independent random combinatorial linear peptide p8 phage libraries available at UCB created by Anastasios Spiliotopoulos consisting of either 9mer or 13mer sequences. Peptide sequences enriched by each mAb were analysed and motifs aligned with *Ebolavirus* GP sequences to identify potential linear or partial epitopes. In independent experiments using each library, biopanning was carried out with both libraries against each mAb with two rounds of selection using decreasing amounts of antibody and an additional subtractive panning against species relevant IgG. The latter was included in order to reduce amplification of phage bound to portions of the antibodies other than the CDRs. All panning steps occurred with mAbs adhered to a plastic surface. After biopanning, phagemid DNA sequences encoding peptide sequences were barcoded via PCR and subjected to Ion Torrent Next Generation Sequencing (NGS). Z score analysis was applied to compare sequences enriched by each mAb and a control antibody to identify sequences specifically enriched by the antibody of interest. For the 13mer library experiment the control antibody was FDRb6, a rabbit mAb generated at the same time as the rest of the panel, but that does not show binding to any *Ebolavirus* GP. For the 9mer library experiment, 11886 enriched peptides from an independent panning experiment were used as the set of control antibody peptides for all sets of enriched peptides except those enriched by 11886, where 66-3-9C was instead used as the control antibody. For experimental summary and barcoding see 7.3 Appendix 3. The top one hundred sequences specifically enriched by each mAb as determined by Z score (see 7.4 Appendix 4) were analysed for motifs using the MEME tool (http://meme-suite.org) [5]. Peptide motifs were aligned with GP sequences from EBOV (ATY51135), SUDV (YP_138523.1), BDBV (YP_003815435.1) and TAFV (YP_003815426.1).

A panel consisting of all the broadly-reactive mAbs (11897, 11878, 11883, 11886, 11889, 11891, 11892, 11881) and two mAbs of limited *Ebolavirus* species cross-reactivity (11894, 11895) was

panned against the 9mer and 13mer peptide libraries, with 11886 panned against the 9mer library twice in independent experiments. 66-3-9C, a human antibody with known epitope on EBOV GP that was previously identified via yeast display technologies [3], was used to pan phage libraries in parallel as a positive control to validate the method. Peptide motifs that aligned with GP sequences were identified for mAbs 11886, 11894 and 66-3-9C (Figure 29). All other mAbs enriched for peptide motifs that could not be aligned to the linear GP sequences (7.5 Appendix 5).

11886 enriched peptide sequences across both peptide libraries and experiments generated motifs containing the sequence arginine-cysteine-arginine-[serine] (RCR) often with another cysteine residue 3 to 4 residues away from the first (Figure 29Ai). The RCR sequence is in 12/15 of the most top most enriched sequences as ordered by Z score. The RCR motif aligned with a conserved sequences in the GP centred on R134, C135 and R136 of the GP sequences (Figure 29Aii). The conservation of these sequences across the GP is consistent with the broad reactivity of the mAb. Residues 134-136 form part of the GP1 head on the exterior of the chalice that is not part of the glycan cap, but sits between the GP1 GC and the GP2 fusion loop (Figure 30A). The second cysteine residue present in many of the motifs does not directly align with the GP sequences, but does suggests the possibility that enriched peptides are naturally cyclised via an internal disulphide bond. This pinch in the peptides possibly presents the arginines in a way that reflects the positioning of the R134 and R136 residues on a turn structure present in the GP. Therefore these constrained peptides may be reflective of a conformational rather than completely linear portion of a partial epitope. The arginine residues appear to be solvent accessible, but access to them by a mAb may be impeded by the disordered β13-β14 loop of the GP1 core that connects the GP1 head and base and is likely to be the site of cathepsin cleavage.

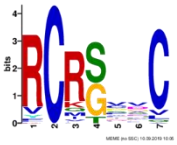
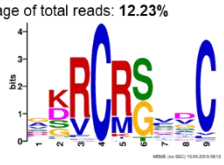
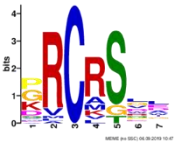
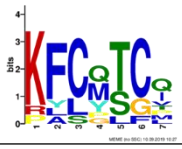
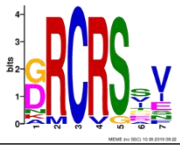
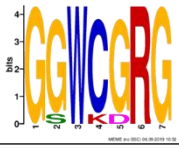
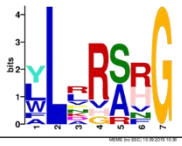
11894 strongly enriched peptides containing an NP[E/D] motif (Figure 29Bi). All of the top one hundred most enriched sequences in both the 9mer and 13mer libraries contain this three amino acid sequence. The motifs generated and the top five most enriched sequences in each library all align with N278, P279 and E280 in the EBOV GP (Figure 29Bii). The NPE sequence in the EBOV GP is not conserved in the other GP sequences; the charged E280 is replaced with uncharged threonine or

asparagine in other GPs, suggesting that an essential salt bridge may be lost in the interaction between these GPs and the mAb. This is supported by the fact that peptides with an alternative aspartic acid (NPD) were also strongly enriched as this residue could maintain this electrostatic interaction. This is consistent with the EBOV GP specificity of 11894 and lack of cross-reactivity to other *Ebolavirus* species. Residues 278-280 are part of the GC of the GP; the residues are solvent accessible on the outside of the GC and trimer chalice (Figure 30B). Their location is consistent with the THL sensitivity of 11894 binding and its ability to compete only GC mAbs for binding to the GP.

66-3-9C enriched peptides in both libraries that generated a similar motif, and both these motifs aligned with the GP within the GC β17-β18 loop GEWAFWET peptide previously identified by yeast display methods using a GP peptide library (Figure 29C) [3]. The motifs generated here suggest that interactions with the tryptophan and glutamic acid residues may be key within the 66-3-9C epitope, which is information that could not be derived using the yeast display GP peptide library. This result validates the phage display method used here and increases confidence in the results from panning with 11894 and 11886.

In addition, the NGS results were corroborated by Sanger sequencing of sequences from individual bacterial colonies generated from the R2 phage library (representing single peptide phage sequences). Sequences from 8-12 colonies sampled from each R2 9mer library reflected the motifs seen in the 11886 and 11894 NGS datasets with peptide sequences including the RCR and NP[E/D] sequences. The sequencing coverage of each R2 library was good with ~50,000 – 120,000 reads per library. Different mAbs were clearly enriching different sets of peptides even if the motifs generated could not be aligned with the GP. However the subtractive panning step did not completely avoid the enrichment of some parasitic sequences in some libraries as the enrichment of peptides from which the motif C[A/G]TC was derived was common across different mAbs (7.5 Appendix 5). The 9mer library resulted in individual peptides less strongly enriched than the 13mer library; as reflected in the percentage of total reads involved in the top motif for each library. However, the motifs derived from each library were highly similar. Taken together this suggests the method worked well to enrich mAb-

A.i.

9mer NGS Exp1 (BC11vs12)	9mer NGS Exp2 (BC04vs11)	13mer NGS (BC16vs23)
Motif 1 E value: **4.0e-113** Unique enriched sequences in motif: **69/100** Number of reads in motif: **11,810** Percentage of top ranked Z score reads: **54.40%** Percentage of total reads: **10.42%** 	**Motif 1** E value: **2.5e-174** Unique enriched sequences in motif: **68/100** Number of reads in motif: **14,140** Percentage of top ranked Z score reads: **71.77%** Percentage of total reads: **12.23%** 	**Motif 1** E value: **1.1e-31** Unique enriched sequences in motif: **35/100** Number of reads in motif: **77,104** Percentage of top ranked Z score reads: **68.79%** Percentage of total reads: **62.35%** 
Motif 2 E value: **2.8e-07** Unique enriched sequences in motif: **10/100** Number of reads in motif: **1992** Percentage of top ranked Z score reads: **9.18%** Percentage of total reads: **1.76%** 	**Motif 2** E value: **2.9e-02** Unique enriched sequences in motif: **13/100** Number of reads in motif: **1931** Percentage of top ranked Z score reads: **9.80%** Percentage of total reads: **1.67%** 	**Motif 2** E value: **9.2e-05** Unique enriched sequences in motif: **10/100** Number of reads in motif: **1733** Percentage of top ranked Z score reads: **1.55%** Percentage of total reads: **1.4%** 
Motif 3 E value: **2.5e-04** Unique enriched sequences in motif: **10/100** Number of reads in motif: **6050** Percentage of top ranked Z score reads: **27.87%** Percentage of total reads: **5.34%** 		

ii.

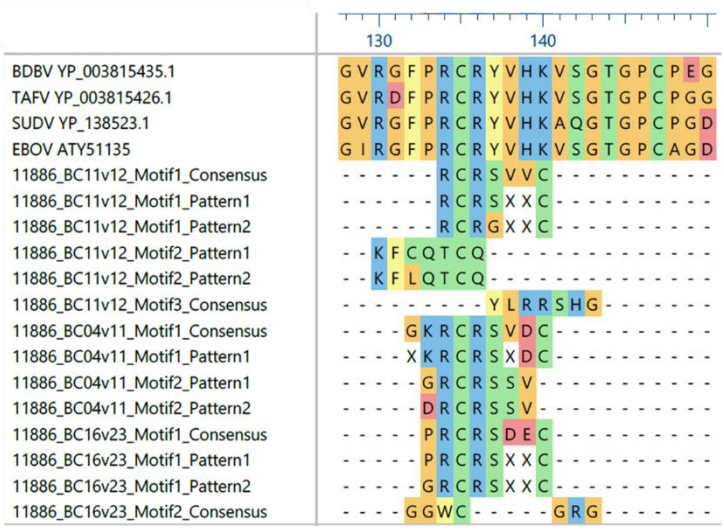
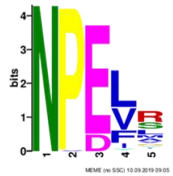
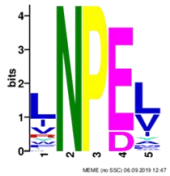
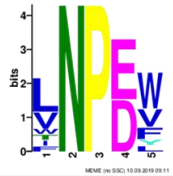
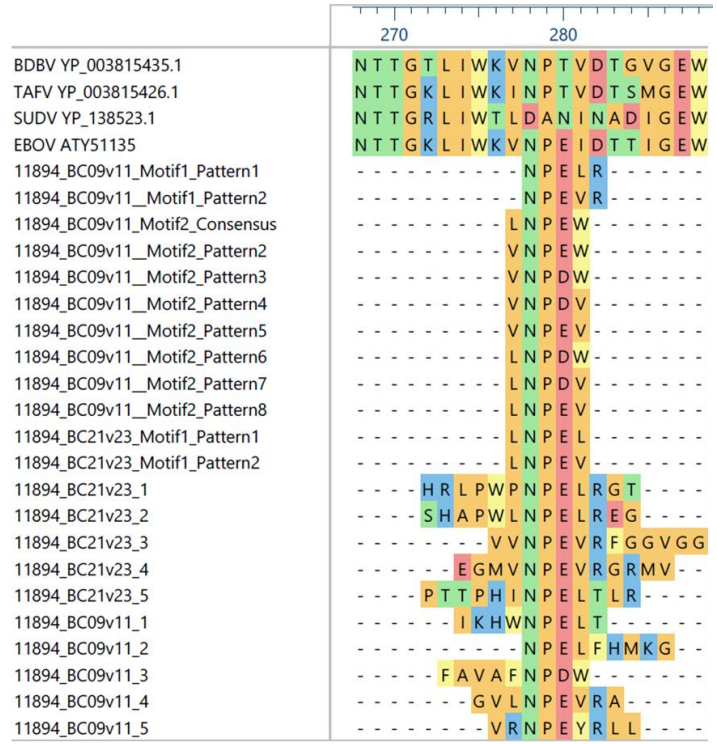


Figure 29. Motifs generated from sequences enriched from 9mer and 13mer linear peptide libraries by A. 11886 B. 11894 and C. 66-3-9C. *Motif logo plots generated using MEME tool (http://meme-suite.org) [5]. Alignments generated using MegaAlignPro with residues coloured by chemistry. GP sequences from EBOV (ATY51135), SUDV (YP_138523.1), BDBV (YP_003815435.1) and TAFV (YP_003815426.1). Percentages given to 2dp. The peptide sequences in BC04vs11 Motif 2 represent a subset of the sequences included in BC04vs11 Motif 1.* ***Figure continued on next page.***

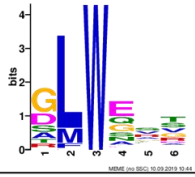
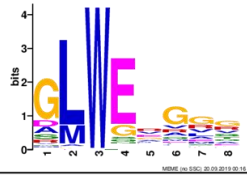
B.i.

9mer NGS (BC09vs11)	13mer NGS (BC21vs23)
Motif 1 E value: **1.3e-172** Unique enriched sequences in motif: **87/100** Number of peptide reads in motif: **9521** Percentage of top ranked Z score reads: **85.54%** Percentage of total peptide reads: **9.15%** 	**Motif 1** E value: **3.2e-189** Unique enriched sequences in motif: **91/100** Number of peptide reads in motif: **52,347** Percentage of top ranked Z score reads: **94.39%** Percentage of total peptide reads: **44.37%** 
Motif 2 E value: **3.7e-002** Unique enriched sequences in motif: **13/100** Number of peptide reads in motif: **1610** Percentage of top ranked Z score reads: **14.46%** Percentage of total peptide reads: **1.55%** 	

ii.



C.i.

9mer NGS (BC12vs11)	13mer NGS (BC24vs23)
Motif 1 E value: **1.4e-099** Unique enriched sequences in motif: **84/100** Number of peptide reads in motif: **782** Percentage of top ranked Z score reads: **76.52%** Percentage of total peptide reads: **0.80%** 	**Motif 1** E value: **7.6e-115** Unique enriched sequences in motif: **93/100** Number of peptide reads in motif: **8707** Percentage of top ranked Z score reads: **96.27%** Percentage of total peptide reads: **9.24%** 

ii.

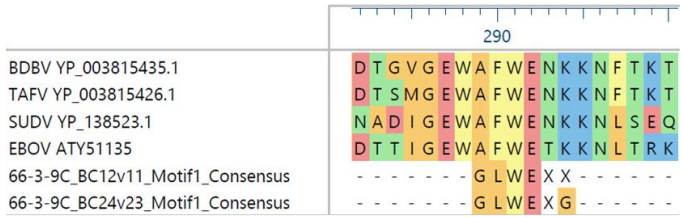


Figure 29 continued.

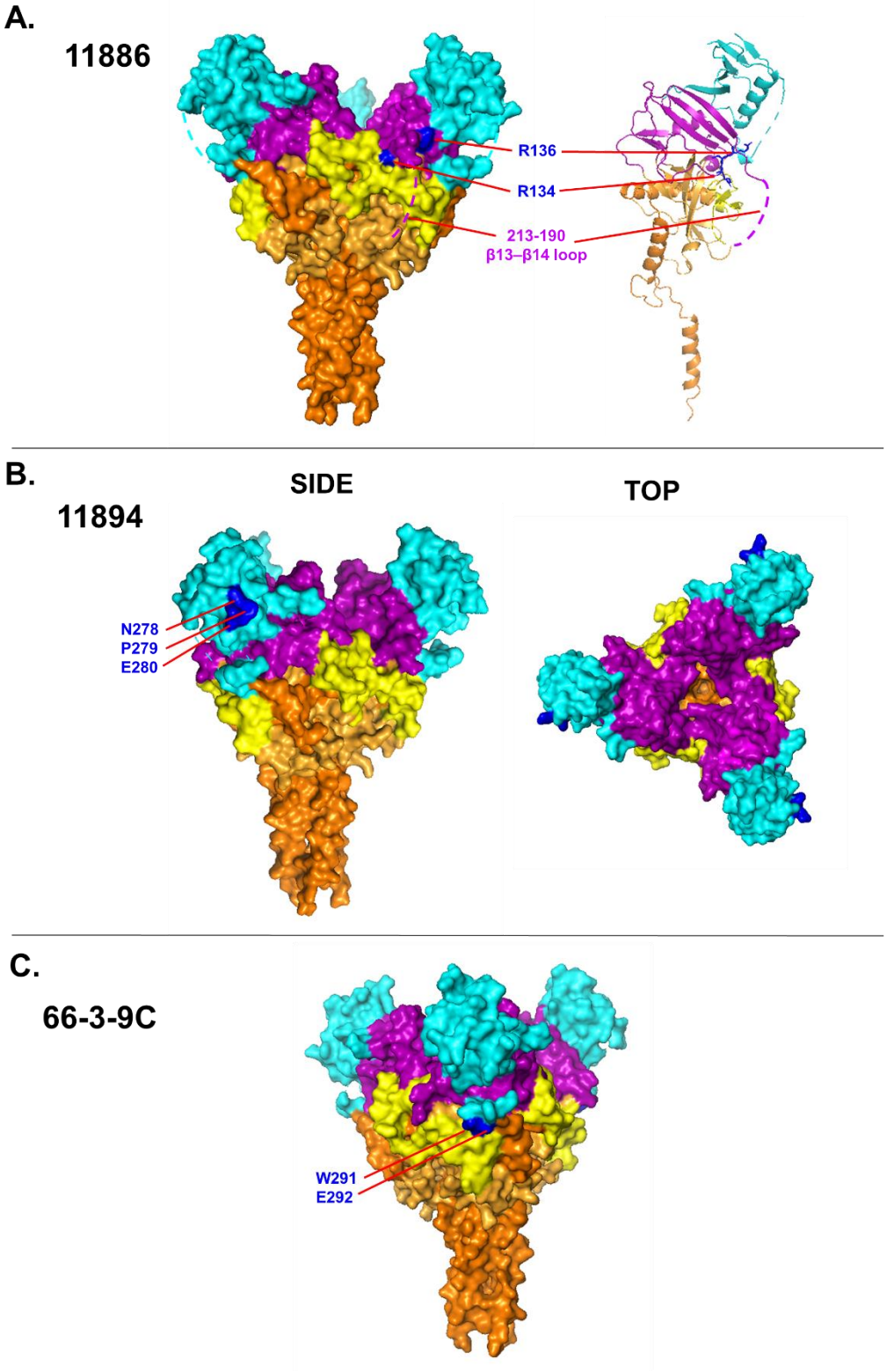


Figure 30. Location on EBOV GP of putative partial epitopes derived from peptide phage display for A. 11886, B. 11894 and C. 66-3-9C*. PDB: 5JQ7 [6]. Surface rendered and recoloured in Pymol using domains as defined by Lee et al. [1]. RBR and GP1 head; purple. GC; blue. FL; yellow, GP1 base; light orange. GP2; dark orange. Residues identified as putative partial epitopes for each mAb; Dark blue. In A. additionally annotated are GC β17-β18 loop (cyan, dashed) and GP1 head β13-β14 loop (purple, dashed) which are not resolved in crystal structure, and single EBOV GP monomer ribbon representation with R134, C 135 and R136 represented as sticks.*

specific peptides and that the NGS sequencing reflected a real collapse in diversity in the libraries generated by panning with mAbs that resulted in motifs that could align with the GP.

To corroborate the peptide phage display findings and to further investigate the epitope of the broadly neutralising mAb 11886, additional epitope mapping methods were next investigated. Linear and cyclic peptides representing the 11886 epitope across the 134-136 region were designed and then synthesised by GeneCust. Also designed were peptides matching the most enriched sequences in each library and control peptides in which the order of the residues was randomised. The intention was to use these peptides in ELISA, then additional SPR-based experiments to confirm binding of 11886 to the peptides enriched during phage display. However, preliminary attempts at the ELISA shows that optimisation or additional modification of the peptides is required to ensure sufficient coating of assay plates.

An alternative approach to corroborate the importance of these residues to 11886 binding is to mutate them in the context of the entire GP rather than short peptides. The enrichment of the RCR motif by 11886 guided a targeted approach to mutating just these solvent exposed arginine residues as discussed next (Section 0).

5.6 Effect of R134A, R136A and R134A/R136A mutations on 11886 interactions with EBOV GP

Highly conserved R134 and R136 in the *Ebolavirus* GP were identified as potentially important residues in a putative, partial epitope for 11886 by peptide phage display (Section 5.5). To investigate the contribution of these residues to the 11886 epitope via an alternative method, R134A and R136A mutations individually and together were introduced by site-directed mutagenesis into the full length EBOV GP. These variants were then used to investigate binding and neutralisation of EBOV by 11886.

5.6.1 11886 binding to R134A, R136A and R134A/R136A EBOV GP variants is unaffected

HEK293F cells were transiently transfected to express the variant GPs. To ensure that the variant GPs expressed on the surface of cells and were properly folded, binding of a broad panel of antibodies targeting multiple different epitopes across GC, RBR/head and base was assessed in a preliminary

experiment by flow cytometry. In parallel, HEK cells were transiently transfected with influenza H7 and specificity of H7 control antibody and EBOV GP mAbs confirmed. All eight EBOV GP mAbs tested were able to stain the variant GPs to a similar degree to the WT GP, confirming their conformational expression (data not shown). Due to the proximity of the region of interest to the glycan cap and cleavage site, suspension cells were used rather than the adherent MDCK SIAT-1 cells so as to avoid the use of any enzymes required to lift cells for flow cytometry in case proteases such as trypsin affect the GP. 11886 binding to WT, R134A, R136A and R134A/R136A EBOV GP variants was then assessed across a range of concentrations alongside c13C6 which was selected as a positive control for GP expression and R5.007 (malaria mAb) as a negative control for staining (Figure 31A). The latter was not an ideal isotype control due to the difference in species, however similar levels of mAb binding were seen to the mock transfected cells suggesting differences in non-specific staining between the anti-rabbit IgG and anti-human IgG fluorophore conjugated antibodies was not significant. EM structures of 13C6 Fab in complex with GP show it approaches the GP at an angle perpendicular to the membrane, binding to the GC on the face leading inside the GP chalice; alanine scanning identified residues T270 and K272 as key for 13C6 binding and so it was chosen as its binding should not be specifically affected by the R134 and R136 mutations unless they globally disrupt the fold of GP1 [179, 263]. Binding to each variant GP was highly similar between 13C6 and 11886 across concentrations. The variant GPs resulted in lower mean fluorescence that the WT GP, but the similarity of the 13C6 and 11886 binding across each experiment suggests this is due to transfection efficiency and expression levels rather than the mutations.

5.6.2 11886 neutralisation *of* R134A, R136A and R134A/R136A EBOV GP S-FLU pseudotype viruses is unaffected

As previously seen when investigating the effect of the I260R mutation on P6 and other GC antibodies (Section 4.5), the loss of residues in a mAb epitope may not affect binding to the extent that can be detected in IFA or flow cytometry, but has an effect that can be seen in changes in the ability of mAbs to neutralise the virus. Therefore the effect of the R134A, R136A and R134A/R136A mutations was further assessed by generating MDCK SIAT-1 cell lines expressing the variant GPs in order to produce

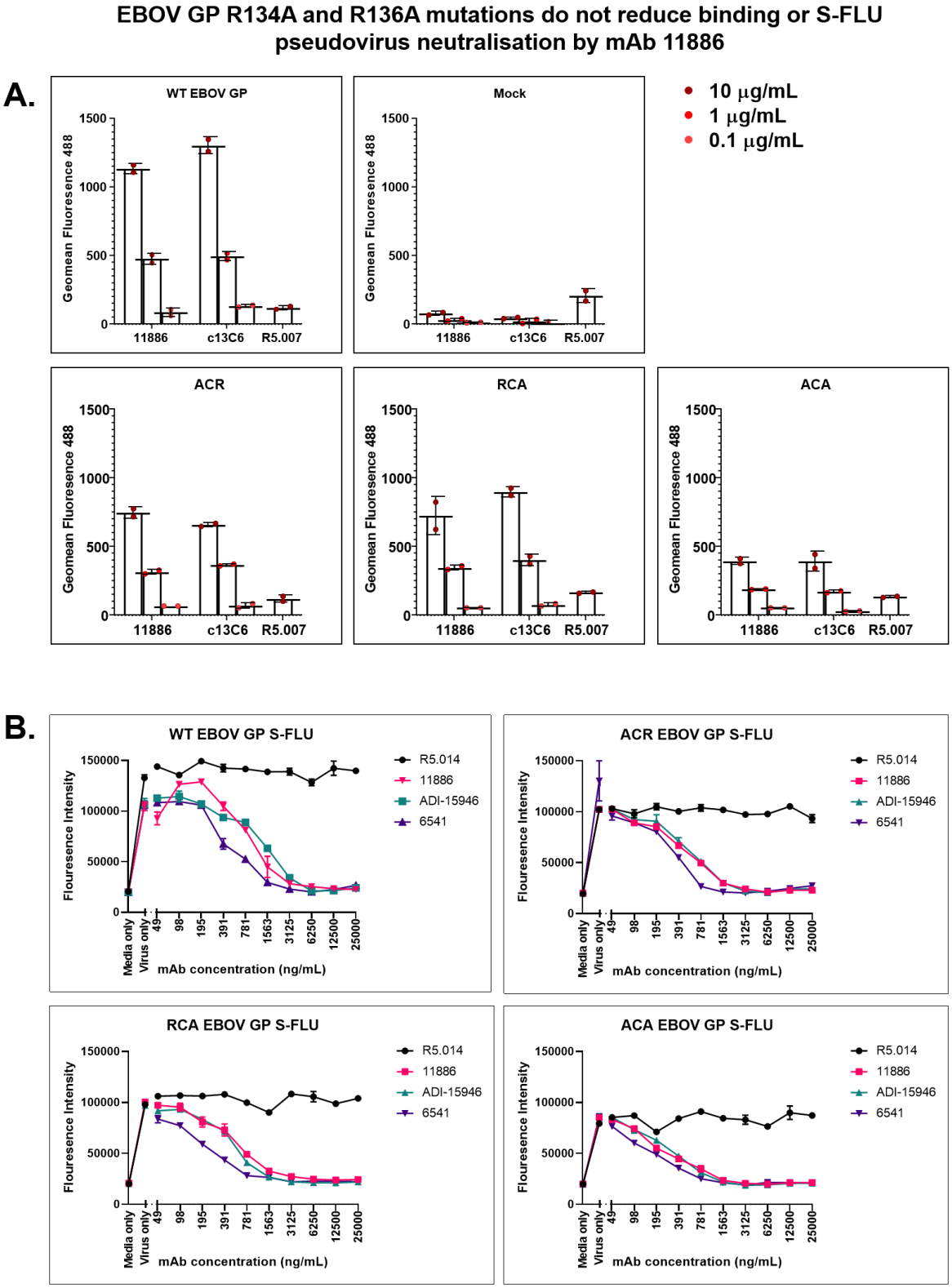


Figure 31. Effect of R134A (ACR), R136A (RCA) and R134A/R136A (ACA) EBOV GP mutations on 11886 binding and neutralisation of S-FLU pseudotype viruses. A. Binding of 11886 and control antibodies to HEK293F cells transiently transfected with variant GPs assessed by flow cytometry*. Mean and standard deviation of duplicates shown of background subtracted data (Background determined by mean of duplicates of cells stained with secondary antibody only). 11886 binding determined using anti-rabbit secondary Alexa Fluor 488 Goat Anti-Rabbit IgG (H L) Antibody (A-11008), c13C6 and R5.007 binding determined using Alexa Fluor 488 Goat Anti-Human IgG (H L) Antibody (Invitrogen, A-11013). >10,000 events collected for each sample after gating for singlets. Data collected using an LSR Fortessa X20 flow cytometer and FACS Diva software (BD Biosciences). Data analysis conducted using FlowJo software.* ***B. Neutralisation of S-FLU viruses coated with variant GPs infecting MDCK SIAT-1 cells.*** *Mean and range of duplicates within assay shown. All fluorescence intensities measured using Clariostar (BMG Labtech).*

S-FLU viruses coated with the variant GPs. Cells were transduced, sorted and expanded as previously described. All three variants cell lines were stained with base mAb 6541 for sorting via FACS. Pseudotype viruses were produced and titrated; all three variants were still able to infect MDCK SIAT-1 cells. Neutralisation of the variant viruses by 11886, ADI-15945, 6541 and negative control malaria mAb R5.014 was tested in the same experiment as neutralisation of EBOV GP S-FLU (WT). The infectivity of the variant GP S-FLU viruses was lower in this experiment so comparison of IC_{50} values should be approached cautiously. However, unlike the effect of the I260R mutation on GC mAbs, 11886 retains strong neutralisation of the variant GPs (Figure 31B). The pattern of neutralisation by the three antibodies against each virus is highly similar, with 6541 being more potent, and 11886 and ADI-15946 displaying highly similar neutralisation profiles to each other across the variant viruses. The R134 and R136 residues are not part of the ADI-15946 epitope which sits very close to R134 and R136, but primarily contacts the 3_{10} pocket (71-75) and GC β17-β18 loop [115]. This suggests that these mutations either alone or in combination are not sufficient to ameliorate binding or neutralisation by these mAbs. This suggests that if they are involved in the 11886 epitope, they are not key contacts of 11886.

5.7 Electron microscopy structure of 11886 Fab in complex with EBOV GPΔmuc

A structural biology approach was next taken to further define the 11886 epitope. I produced 11886 Fab which was used by Dr Amar Parvate (Saphire Lab, La Jolla Institute for Immunology, La Jolla, California) for an initial low resolution cryo-electron microscopy study of the Fab in complex with the EBOV GPΔmuc trimer. GP and complexes were produced by the Saphire lab.

Briefly, Fab was incubated at molar excess with trimeric GP and complexes were purified by size exclusion chromatography. Complex was loaded onto grids which were plunge frozen. Data was then collected using a FEI Titan Krios electron microscope. Single particle analyses excluded classes with suboptimal Fab occupancy. 3D refinement using 72,000 particles were used for final refinement with C3 symmetry imposed; a map of 5.9Å was achieved. A homology model for 11886 Fab was built using PDB map of ADI-15946 [115], and for the GP using PDB: 5JQ7 [6] which contains the GC (Figure 32).

11886 binds to the GP at a shallow angle from the perpendicular, with the constant region angled down towards the viral membrane. The Fab appears to engage an interface composed of regions of the GC, GP1 core and possibly GP2, sitting in the space between the FL base and the GC predominantly occupied by the β16-β17 loop of the GC binding the GP at the vertex of each GP1,2 protomer (Figure 32). This is broadly consistent with the other epitope mapping approaches.

Docking of crystal structure 5JQ7 shows that density of a portion of the GC (which is not resolved in cryo-EM due to high flexibility) is potentially clashing or very close to density of the 11886 Fab heavy chain CDRs, in particular residue W77 of the Fab (Figure 32B). This indicates an interaction with the helix of the GC or even that 11886 Fab partly displaces the GC to accommodate itself binding at an angle approaching from underneath the GC. The positioning of the Fab HC CDRs also does not exclude the possibility of the contacts with the residues forming the top of the sequence leading into and out of the β17-β18 loop.

A significantly larger dataset on the GP-Fab 11886 complex needs to be collected to obtain a map between 3-4 Å resolution and model contacts between the Fab and GP in more detail. However, the Fab footprint appears to be covering the β17-β18 loop of the GC (Figure 32B). The footprint also runs along the right hand edge of the predicted location of the β13-β14 loop; it is possible that it could contact this loop or at least occlude access of cathepsins to this cleavage site. In addition there appears to be contact from the CDRs L1 and L3 with GP2 in the approximate region of the N terminus (~502-511) that immediately precedes the fusion loop (512-555). This current model suggests the Fab binds in the space above the FL base/stem rather than contacting these residues directly. It is possible that it could inhibit the rearrangement necessary for the fusion peptide to be inserted in the host cell membrane by steric hindrance and preventing cleavage of the loop pinning the FL to the GP.

Fab density as modelled here does not strongly support a direct interaction with R134 or R136 although the latter is directly proximal to the Fab footprint.

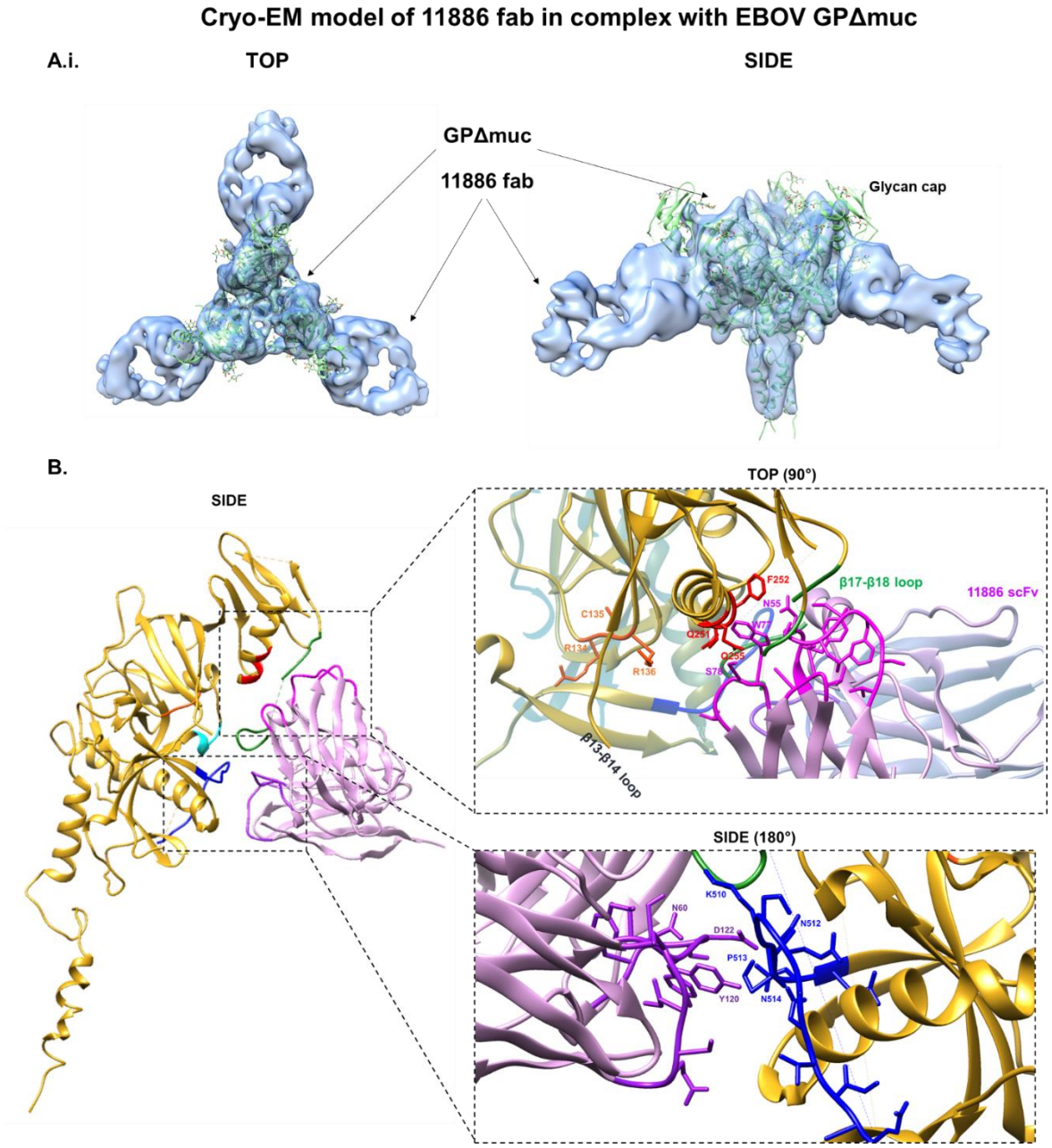


Figure 32. Cryo-EM model of 11886-EBOV GP Cryo-EM model of EBOV GPΔmuc-11886 Fab complex at 5.9Å. *A. Cryo-EM density (grey) map docked with crystal structure of EBOV GP with GC (green, cartoon) (PDB: 5JQ7, [6] B. Homology models of 11886 scFv (pink, ribbon) and single GP1,2, monomer (yellow, ribbon) (PDB: 5JQ7) fitted to EM density which has been hidden. Top box: Detail showing area where electron density of 11886 Fab HC CDR clashes with glycan cap in crystal structure. Side box: Detail of interface between 11886 scFV LC with GP2 N terminus. Side chains of residues that may clash or interact are shown. Red: possible area of GC alpha helix clashing with 11886 Fab. Cyan: 3_{10} pocket residues 71-75. Green: β17-β18 loop. Dark blue: GP2 N terminus. Purple: LC Fab CDR L1 and L3. Hot pink: HC Fab CDR H1, H2 and H3. Orange: R134, C135, R136. All images created in UCSF Chimera. Images in A. created by Dr Amar Parvate.*

5.8 Discussion

5.8.1 Vaccination strategy resulted in a panel of broadly reactive mAbs against distinct regions of the GP

The decision to vaccinate with full length GPs expressed on the surface of cells was taken in order to take an unbiased approach to finding broadly reactive mAbs against a range of epitopes on the pre-fusion potent broadly reactive mAbs when the project started. The epitope mapping in this Chapter combined with data in Chapter 4 has shown that, among the broadly reactive rabbit mAb panel derived in Chapter 3, there are antibodies that bind to distinct GC and base epitopes, with a range of overlapping epitopes within these competition groups, and none appear to predominantly contact the FL. This validates the approach taken to antibody discovery, and opens the possibility of designing cocktails of non-competing mAbs to make a potent broadly reactive mAb cocktail for *in vivo* study, including testing combinations with published potent fusion loop mAbs to supplement their efficacy and reduce chance of viral escape. It also shows, as is the growing consensus in the literature, that whilst the fusion loop is an important site of vulnerability on the GP, there are other distinct sites on the GP that are conserved and can be the target of neutralising mAbs.

In this Chapter, multiple complementary approaches were taken to identify the epitope bins of each broadly reactive mAb. The different assays were consistent with each other and with the data from Chapter 4. None of the mAbs were able to compete with NPC1-C domain binding to the GP_{CL} and so are unlikely to be binding directly to the same region in the RBR necessary for NPC1-C binding in the absence of the GC. In addition, none of the mAbs lost the ability to precipitate GP after the removal of just the MLD, suggesting that none have MLD epitopes; this is unsurprising given the fact that these mAbs are broadly reactive, and the MLD is the least conserved portion of the GP between *Ebolavirus* species. Peptide phage display methods did not generate any motifs for the broadly reactive rabbit mAbs that aligned with the GP except from 11886 suggesting they all have conformational epitopes. The epitope of 11886 was explored in more detail and is discussed separately in Sections 5.8.3 and 5.8.40.

11883 and 11897 are assigned to the same epitope bin as they compete each other and show similar characteristics across the range of assays used. They cannot bind to GP_{CL} in IFA or immunoprecipitation experiments, compete GC and GP1 core/RBR mAbs, do not compete FL mAbs or lose binding to chimeric EBOV GP with MARV FL. Together this suggests both mAbs predominantly bind the GC, but possibly at an angle that allows them to compete RBR/GP1 core mAbs. The two mAbs however are distinguished by their neutralisation patterns: 11883 is cross-neutralising (EBOV GP S-FLU, partial; SUDV GP S-FLU, strong), whereas 11878 is unable to neutralise any of the pseudotypes tested. This, along with the indication the 11883 may compete 66-3-2C less strongly that 11878, shows that the two mAbs have different epitopes within the GC. Apart from 11886, 11883 was the only mAb to show any kind of species dependency when binding GP vs GP_{CL}. 11883 in one repeat showed much stronger binding to SUDV GP_{CL} than to EBOV, BDBV or TAFV GP_{CL} to which its binding was almost completely lost. This was not convincingly reproducible in the repeated experiment undertaken, but could be of interest to investigate further as it may elucidate why this mAb can neutralise SUDV GP S-FLU so potently in comparison to the other pseudotype viruses. As is typical of almost all other mAbs assigned to the GC epitope bin, 11883 lost neutralisation of the I260R EBOV GP S-FLU.

11889 is also broadly in the GC epitope bin due to its competition of GC and GP1 core mAbs, and THL sensitivity of its epitope. However, its epitope is distinct from both 11883 and 11897 as it competes mAb ADI-15878 and competes 11883 in one orientation only. 11889 competes mAb ADI-15878, but not 6D6 or FVM02 which also bind the FL tip: as 6D6 and ADI-15787 share an almost entirely overlapping footprint it is unlikely 11889 shares the same footprint but does not compete them both. Therefore 11889 may compete ADI-15878 by steric hindrance rather than direct competition for the surface of the GP; this is consistent with fact that 11889 binding to the chimeric EBOV GP with MARV FL was unaffected. In addition, 11889 also showed loss of neutralisation of the I260R EBOV GP S-FLU. Therefore 11889 is likely a GC mAb that binds at an angle allowing it to compete a mAb binding the FL tip just below it. As 11889 does not compete with 11886 this is consistent with it binding on the side of the GC above the tip of the FL of the neighbouring protomer rather than the top

or side vertex of the GP1,2 protomer. This suggests an overlapping series of GC epitopes for 11883, 11897 and 11889 vertically down the GP along the side GC where the crest of the RBR/ GP1 head is exposed towards the base (Figure 25B).

11892 retains its epitope after thermolysin digestion in both IFA and immunoprecipitation experiments and competes base binding mAbs, but not GC or RBR mAbs. It strongly competes ADI-15946 and 11886, but not the mAbs that bind to the tip or further up the stem of the FL, and retains binding to the chimeric EBOV GP with MARV FL. This is consistent with it having a base epitope in the vicinity of the FL base, that does not rely heavily on contacts within the FL itself or that vary between the EBOV and MARV FL sequences. Its epitope is distinguished from 11886 and ADI-15946 by its neutralisation profile: it cannot neutralise EBOV GP or SUDV GP S-FLU viruses, and only partially inhibits infection by BDBV GP S-FLU.

11878 heavily precipitated when expressed and purified, leading to issues with testing it in the full range of assays. However, in two repeats of an IFA assay to determine if it could bind to EBOV GP_{CL}, it appeared to retain its epitope (data not shown), possibly to bind much more strongly to GP_{CL} than uncleaved GP. This is consistent with the inability of 11878 Fab to inhibit THL cleavage in the immunoprecipitation assay (Section 4.3.2) and suggests a base epitope for the mAb.

EBOV GP-specific 11894 cannot bind cleaved EBOV GP in IFA, competes only GC mAbs, loses neutralisation of the I260R EBOV GP S-FLU and completely collapsed the diversity of two peptide phage display libraries to sequences containing an NP[E/D] motif which aligns with a EBOV GP specific sequence in the GC close to the β17-β18 loop. The fact that it competes only GC mAbs and not the GP1 core mAbs in the competition IFA puts it in a separate GC epitope bin to 11883, 11897 and 11889. The binding of 11894 to the NPE motif is consistent with it competing ADI-15878, which binds to the FL tip of the neighbouring protomer just below and to the right of the NPE motif location on the GP. The consistency of results for this mAb across all of the assays increases confidence in the multiple approaches used to identify the epitopes of the broadly reactive mAbs of greater interest.

5.8.2 Peptide phage display as a rapid method of generating epitope data without having to have an antigen-specific peptide library.

In the absence of structural methods, alanine scanning libraries or antigen-specific peptide libraries, random combinatorial peptide libraries provide a strategy to gain information on the epitopes of mAbs against any antigen [220, 264]. They require libraries that cover a high proportion of the total possible combinatorial diversity of amino acids in peptides of a given length and a way to increase the confidence in the specificity of any enrichment seen, such as the subtractive panning and z-score analyses applied here.

For the libraries used in the experiments reported here, their high coverage of combinatorial diversity had previously been determined by Dr Anastasios Spiliotopoulos. For example for the 9mer library, > 98% of ~1 million sampled sequences were unique, and the library size was $1.5x10^{10}$ colony forming units: thus the starting library contains a large proportion of the theroteically possible $5x10^{11}$ 9mer peptide sequences. As with other non-structural approaches, peptide phage display of random linear peptides can only provide partial epitopes, and as with linear peptide display methods, generally only linear portions of epitopes that can be aligned with the primary sequence of the antigen. This makes the success of such an approach highly mAb dependent. However, the ability to use pre-prepared non-antigen specific libraries allows the investigation of mAb epitopes (or other protein-peptide interactions) without access to, or before the generation of, antigen-specific peptide libraries. The libraries used here performed well: peptide sequences were retrieved from the majority of sequence reads, different mAbs enriched different sets of peptides, there was strong agreement in motifs enriched by the same mAb using two independent libraries, and the motifs derived for the control antibody 66-3-9C aligned with the known epitope of the mAb. The combining of phage display techniques and next generation sequencing methods allows the interrogation of many mAbs in parallel via tens of thousands of sequence reads (replacing picking of individual clones) and here proved its ability to identify epitopes consistent with other epitope mapping data. In addition, the use of random peptides did not bias epitope mapping towards one GP sequence, which was a consideration when attempting to define the epitopes of mAbs that bind multiple GPs. This strategy

therefore also allows for the parallel mapping of mAbs that recognise different *Ebolavirus* GPs at the same time using the same starting peptide phage libraries; an approach that could be applied to any system where antibody responses to multiple serotypes or variants of an antigen may need interrogating [220, 265].

5.8.3 Epitope of mAb 11886

The epitope mapping data of 11886 is consistent with a mAb that binds at the interface between previously defined GC and base epitope bins close to the base of the GP2 fusion loop and N-terminus.

The location of the Fab in complex with the EBOV GPΔmuc shows the antibody as binding across the GC, GP1 and potentially GP2, sitting between the fusion loop base waist and GC rim on the outside of the GP chalice. EM density of the Fab suggests 11886 interacts with the bottom of the GC alpha helix and β17-β18 loop, although whether it displaces these portions of the GP is yet to be determined. The observation that the epitope is THL sensitive indicates that there are direct contacts with the GP that are lost in the cleaved GP (GP_{CL}). Unlike the other mAbs that are THL sensitive, 11886 does not appear to compete other GC mAbs. In addition, the binding of other rabbit mAbs that are in the GC epitope bin (11883, 11887) may have even improved binding of biotinylated 11886 in the competition assay (Figure 25). This apparent inconsistency is resolved by the angle of 11886 fab binding to the GP in the EM structure; it approaches the GC from beneath suggesting its Fc would be angled towards the viral membrane in a full IgG molecule. Its footprint appears highly similar to mAb ADI-15946, explaining its ability to most strongly compete ADI-15946 and other base mAbs such as c2G4, but largely not affect the binding of the GC or GP1 core mAbs tested.

5.8.3.1 Relation of 11886 epitope to the fusion loop

11886 did not lose binding to the chimeric EBOV GP with MARV fusion loop (FL) (E/M), which suggests that although it sits close to the residues in the very base and stem of the FL, it does not rely on them for binding. The EM structure indicates that there may be contacts with the residues in the N-terminus of GP2 (502-514) immediately preceding and overlapping into those changed in the chimeric GP construct (512-555). The residues that looked most likely to interact with 11886 in the docked homology models were N512 and P513. However as both of these were lost in the E/M chimeric GP

either they are not necessary for 11886 binding, or higher resolution structural data will show that different GP2 N-terminus residues are more likely to interact.

The base of the FL on one GP1,2 protomer is separated only by the N-terminus of GP2 from the tip of the FL of the neighbouring GP1,2, protomer. The positioning of 11886 footprint with potential contacts in this GP2 N-terminus and not the FL base, explains the partial competition of ADI-15878 and possible partial competition with 6D6. These mAbs bind across the FL tip, but also contact the GP outside of the FL [10, 193]. It is therefore consistent that a mAb such as 11886 that sits just to the left of this epitope would compete binding of these mAbs by steric hindrance or competition with the non-FL contacts of the ADI-15878 and 6D6 with the GP. In addition to the position of the Fab in the EM structure, lack of competition with FVM02 which can bind a linear epitope consisting of just tip of the FL, the inability of 11886 to recognise the synthetic FL peptide and the ability of 11886 to retain binding to the EBOV GP with MARV FL, suggests that 11886 does not in fact contact the tip of the FL itself.

5.8.3.2 Importance of RCR motif

The R134A and R136A mutations made no detectable difference to 11886 binding to EBOV GP via FACS nor to neutralisation of S-FLU pseudotype viruses *in vitro*. These residues were identified via alignments of phage display derived peptide motifs generated from peptides enriched by binding 11886. These residues may be sufficient for a mAb to bind a peptide (confirmatory experiments are planned) but do not contribute largely to the overall binding to GP or functioning of the mAb which can therefore tolerate their loss. However, the EM structure suggests that it is unlikely that both R134 and R136 contact the 11886 Fab unless there is a conformational change in the GP that is not detectable in the current resolution of EM density.

The peptide phage display panning and enriching method with Next Generation Sequencing for motif generation worked well for mAbs 11894 and 66-3-9C, and the same motif was derived from three independent panning experiments with two different peptide libraries; therefore the strong enrichment of the RCR motif is unlikely to be an artefact. However, the alignment with the GP may be misleading. It could be that the motif represents a conformational epitope which co-incidentally arises

in the linear sequence of the GP leading to a misattribution of the RCR motif to 134-136 of the GP sequence. The indication of cyclisation of many of the enriched peptides could be used as an argument for the R134 and R136 residues, but could also be applied to other conformational epitopes. There are other arginine residues close to the interface between the Fab and GP the docked homology structures, but no other pair seem likely to both interact with the Fab. Ultimately, the peptide phage display and EM data disagree on the likely importance of this binding motif, but both place 11886 epitope in roughly the same area on the GP and higher resolution structural data may explain this inconsistency.

5.8.3.3 Thermolysin sensitivity of 11886 epitope

Rather than residues 134-136, it instead seems that residues lost after thermolysin cleavage of the GP are more important to 11886 binding as the mAb cannot precipitate EBOV GP after the glycan cap is completely removed (GP_{CL}). These residues may be completely lost after cleavage or conformational changes (induced by cleavage) of remaining residues may impede 11886 binding. The thermolysin sensitivity of its epitope on EBOV GP was determined by IFA and immunoprecipitation, and the thermolysin sensitivity of its epitope on BDBV and TAFV GPs in IFA. Intriguingly it appeared in IFA experiments that the THL sensitivity of the 11886 epitope may be less pronounced for SUDV GP binding.

The EM data suggests that Fab residues N55, W77 and S78 are positioned closest to the GC alpha helix resides Q251 and Q255. In the SUDV GP, residue 255 is instead a Leucine. The β17-18 loop which may have multiple contacts with the Fab is largely conserved between all GPs, but the preceding sequence is less conserved in SUDV than other GPs. However it is hard to see how differences in the contacts with the portion of the GP which is lost would make the mAb better at binding cleaved SUDV GP. Alternatively, the GP2 N-terminus preceding the FL is less conserved between SUDV and EBOV than for BDBV and TAFV. In particular, there are more charged residues in the SUDV GP2 N-terminus. This may introduce additional interactions with the SUDV GP2 that allows the mAb to continue binding in the absence of the GC.

With the assays so far conducted it cannot be determined if 11886 binds or neutralises SUDV significantly better or worse than EBOV, but it can fully neutralise S-FLU pseudotypes. SPR with recombinant purified GP proteins would show if 11886 has higher affinity for one GP over another; affinities to GP_{CL} could also be tested. The immunoblot assay could also be used to corroborate the result seen in the IFA by testing if 11886 can precipitate SUDV GP_{CL}.

The ability of 11886 to inhibit thermolysin cleavage, and its ability to continue to neutralise the I260R EBOV GP variant is consistent with its epitope close to the cleavage site in the β13-β14 loop and its interactions with both the GC and base near the GP2 N-terminus, although exactly how it functions requires further structural study. It seems highly likely it acts via multiple mechanisms to hold the GP in a pre-fusion conformation thereby preventing receptor binding and membrane fusion.

5.8.4 Distinction of 11886 epitope from ADI-15946 and EBOV-520

ADI-15946 and EBOV-520 are broadly reactive mAbs that have footprints on the GP that largely overlap each other and that of 11886 [115] and involve contacts in both the GP1 and GP2 (Table 1). Like 11886, ADI-15946 and EBOV-520 are reported to be able to inhibit proteolytic cleavage of the GP [99, 211], and like 11886, ADI-15946 retained neutralisation against the I260R EBOV GP S-FLU (EBOV-520 was not tested in this assay). However, the direct contacts made by 11886 seem to be distinct from those of both ADI-15956 and EBOV-520.

Both ADI-15946 and EBOV-520 bind to cleaved EBOV GP (GP_{CL}) with higher affinity than to uncleaved GP and can neutralise viruses bearing GP_{CL} more potently [99, 211]. The crystal structure of ADI-15946 in complex with EBOV GP_{CL} showed that HC CDR3 competes with and mimics the β17-β18 loop (residues 287-291) by binding to a cryptic epitope in the 3_{10} pocket on the GP1 core (residues 71-75) [115]. In addition pairing ADI-15946 with FVM09, a mAb that binds to the β17-β18 loop, also enhanced binding, which the authors hypothesised was due to FVM09 improving displacement of the loop from the side of the GP chalice. The crystal structure of EBOV-520 in complex with GP_{CL} has also been determined (PDB: 6OZ9) [191] showing that it too contacts the 3_{10} pocket. When the authors increased access to the 3_{10} pocket by mutating the β17-β18 loop so that it was easier to displace from the 3_{10} pocket, they enhanced binding of EBOV-520. EBOV-520 also behaves co-operatively when

paired with a GC mAb, EBOV-548. However the authors hypothesise that this is not due EBOV-548 directly binding the loop and holding it away from the GP core, but instead EBOV-548 contacts the β17 strand pulling back the GC and the β17-β18 loop at the same time. This effectively achieves the same result, decreasing occupancy of the 3_{10} pocket by the GC loop and exposing the pocket for partner mAb binding [191].

ADI-15946 and EBOV-520 both also contact the GP2 N-terminus, with the conserved K510 residue identified as necessary for binding of ADI-15946 [115] and N512 necessary for EBOV-520 binding [99]. The decreased activity of ADI-15946 against SUDV was attributed in large part to the contact of ADI-15946 CDR H3 with EBOV GP N506, which in SUDV is instead an arginine. Antibody ADI-23774 was developed by specificity maturation of ADI-15946 to increase binding affinity for SUDV GP [2]. This antibody gained the ability to neutralise SUDV *in vitro* and offered protection against SUDV infection in a mouse model [2]. Although the sequence of this antibody and changes to its epitope were not included in published reports of that study, a parallel study using a structure-guided approach reduced steric clashes with GP N506 and increased activity of ADI-15946 against SUDV GP [115].

Whilst 11886 broadly has a similar footprint on the GP, its binding is distinct from the 3_{10} pocket mAbs in several key ways:

Firstly, as determined by both IFA and immunoblot, 11886 cannot bind EBOV, BDBV or TAFV GP_{CL}, and whilst it seems to still recognise SUDV GP_{CL}, binding is still decreased in comparison to uncleaved SUDV GP. This directly contrasts the enhanced binding of the 3_{10} pocket mAbs to GP_{CL} and is perhaps the clearest indication that these mAbs function differently. A comparison of EBOV-520 and 11886 binding (Figure 33B) shows an EBOV-520 CDR occupies an area of electron density by the 3_{10} pocket of the GP that is unoccupied by 11886 CDRs but is occupied by the bottom of the β17-β18 loop in the uncleaved GP structure shown for comparison.

Secondly, the angle of 11886 binding to the GP is different to both ADI-15946 and EBOV-520. Compared to ADI-15946 (Figure 33A) 11886 binds at a shallower angle to the perpendicular, and ADI-

15946 binds offset from the vertex of the GP1,2 protomer. This difference in binding angle is consistent with the fact that whilst 11886 partially competed fusion loop tip mAb ADI-15878 in the

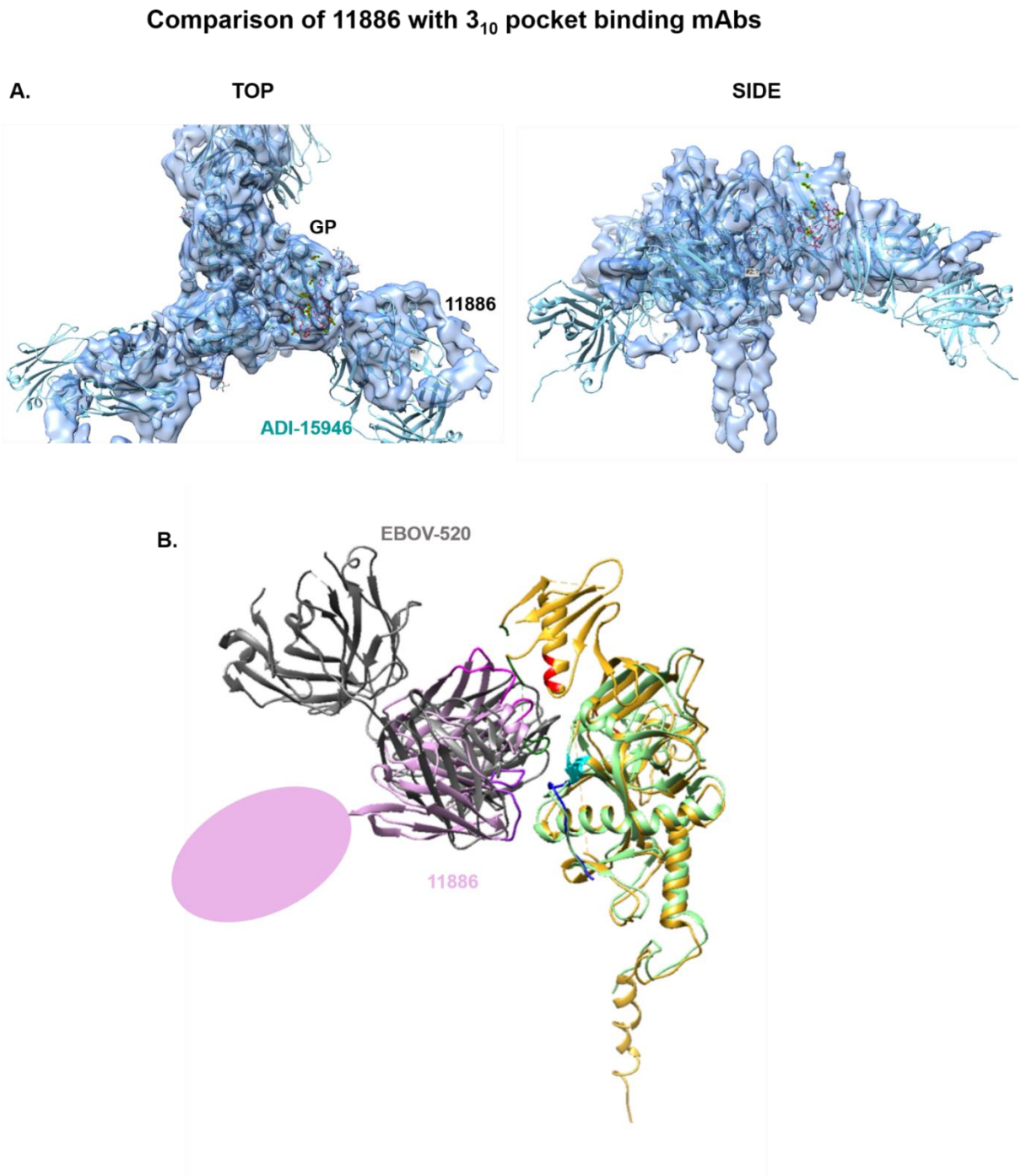


Figure 33. Comparison of 11886 and 310 pocket binding mAbs interacting with EBOV GP. A. Comparison with ADI-15946 *X-ray structure of ADI-15946 in complex with GP_{CL} (teal, ribbon) (PDB: 6MAM [115] fitted to EM density map of 11886 Fab in complex with EBOV GPΔmuc (grey surface). Images created in UCSF Chimera by Dr Amar Parvate.* ***B. Comparison with EBOV-520.*** *X-ray structure of EBOV-520 (grey, ribbon) in complex with GP_{CL} (pale green, ribbon) (PDB: 6OZ9 [191]), homology model of 11886 scFv (pink, ribbon) and X-ray structure of GPΔmuc (yellow, ribbon) (PDB: 5JQ7) fitted to EM density map of 11886 Fab in complex with EBOV GPΔmuc (hidden). Location of rest of 11886 Fab shown by pink oval. Red: possible area of GC alpha helix clashing with 11886 Fab. Cyan: 3_{10} pocket residues 71-75. Green: β17-β18 loop. Dark blue: GP2 N-terminus. Purple: 11886 LC Fab CDR L1 and L3. Hot pink: 11886 HC Fab CDR H1, H2 and H3. All images created in UCSF Chimera. Images in A. created by Dr Amar Parvate.*

competition IFA, ADI-14946 is reported not to compete this antibody [2]. The angle of 11886 binding is even more distinct from EBOV-520 (Figure 33B); the EBOV-520 Fab approaches the GP from a top down angle, whereas the 11886 Fab approaches the GP from underneath the GC.

Thirdly, 11886 whilst appearing to contact the GP2 N-terminus like the 3_{10} pocket mAbs, does not rely on the same contacts within this area. ADI-15946 lost almost all binding to the chimeric EBOV GP with MARV FL (E/M), whereas 11886 did not (Figure 24). Whilst EBOV-520 was not tested in that assay, the residue N512 which is necessary for EBOV-520 binding was swapped from an uncharged asparagine to a charged aspartic acid in the E/M chimera without affecting binding of 11886 in a cell-based immunofluorescence assay (IFA). EBOV-520 has multiple contacts in the base of the FL, but none of these appear necessary for 11886 binding. As 11886 can neutralise the SUDV GP S-FLU pseudotype this suggests that it also avoids substantial interaction with N506 increasing its breadth of activity compared to the original version of ADI-15946.

Due to the contrasting thermolysin sensitivity of 11886, it is more likely that it contacts the GC rather than displacing the β17-β18 loop like the 3_{10} pocket mAbs. To elucidate the relationship between 11886 and the β17-18 loop, similar experiments to those conducted with ADI-15946 and EBOV-520 could be carried out using loop binding mAbs. Initial data from the competition IFA in Section 5.2 suggests that 11886 does not compete 66-3-9C, a mAb known to bind the tip of the β17-β18 loop [3] – however there was no positive control for binding of 66-3-9C in that experiment and more thorough competition studies are required. Similarly, 11894 was identified via peptide phage display as having an epitope involving residues at the top of the loop. Binding of 11886 in the presence of 11894 is yet to be investigated, but they are likely to affect each other in some way.

5.8.5 Summary

In this Chapter a series of epitope mapping experiments were conducted which were highly consistent with each other and the data presented in Chapter 4. The panel of broadly reactive rabbit mAbs includes mAbs from different epitope bins including the GC and base, thereby validating the approach

taken using full length antigens to generate a panel of antibodies that bind a range of epitopes outside of the FL.

The lack of FL dependency seen for any rabbit mAb with the chimeric EBOV GP with MARV FL and synthetic FL peptide was supported by the minimal competition with known FL-binding mAbs in competition IFA. The lack of NPC1-C competition was supported by the absence of any mAbs in the panel that both competed GP1 core or GC mAbs and retained binding the GP_{CL}. Consistent with observations of other GC mAbs previously made by Dr Pramila Rijal, those mAbs that lost neutralisation of I260R EBOV GP S-FLU also competed GC mAbs and lost binding to GP_{CL} in two assays.

In addition, the epitope of 11886 was investigated in more detail and was determined by a combination of biochemical and structural methods to bind to an epitope spanning domains of the GP. The footprint of the 11886 on the GP is highly similar to that of 3_{10} pocket antibodies ADI-15946 and EBOV-520, but specific contacts are distinguished by differences in their interactions with the fusion loop, GP_{CL} and GP from different species of *Ebolavirus*. Therefore, whilst the 11886 footprint is similar, it likely does not constitute another 3_{10} pocket mAb, but acts via a different mechanism that also involves the β17-β18 loop. Like these ADI-15946 and EBOV-520, 11886 may utilise multiple mechanisms to mediate neutralisation – utilising a highly conformational, domain-spanning epitope to impact the multiple changes necessary for the GP to transition from its pre-fusion conformation.

Determining epitopes and competition between mAbs provides important information for rational design cocktails of antibodies: this and other future experiments are discussed in detail in Chapter 6.

6 Discussion

6.1 Thesis summary

Ebolaviruses cause devastating disease with high mortality. Their complex and, as yet, poorly defined ecology, viral persistence in survivors of infection, the risk of emergence of new species, and the challenging social and economic circumstances of communities experiencing outbreaks, means that despite the great strides forward in vaccine and therapeutic development in the last eight years, we continue to see regular outbreaks of ED. Whilst mAb therapies REGN-EB3 and mAb114 show laudable efficacy and undoubtedly have saved and will continue to save lives, there are still many challenges to overcome with regards to treating *Ebolavirus* infections. This thesis focused on broadly reactive antibodies targeting the GP that may form the basis of therapies for infections caused by multiple known and yet to emerge *Ebolaviruses*. The mAbs presented here contribute to a field of broadly reactive mAbs generated over the last five years and form part of a robust pipeline of mAbs to be evaluated as the next generation of improved ED therapies.

In Chapter 3, repeated immunisation of rabbits with RAB-9 cells expressing full length GPs from EBOV, SUDV, BDBV and TAFV was used to generate immune material for antibody discovery. A high throughput screening strategy was devised to enrich for broadly-reactive mAbs using the UCB Core Antibody Discovery B cell culture platform. A panel of 30 variable regions were recovered and recombinantly expressed as rabbit IgG, of which 19 were GP specific and six were unique sequences that could bind GP from all four species of *Ebolavirus* tested. The predicted germline gene usage of the panel of antibodies reflects reported rabbit Ig repertoires in the literature. To my knowledge, this panel of mAbs represents the first broadly reactive panel of antibodies against full length *Ebolavirus* GP utilising the rabbit immune repertoire.

In Chapter 4 the ability of mAbs to neutralise S-FLU pseudotype viruses coated with GP from EBOV, SUDV and BDBV was assessed; mAb 11886 was identified as the first mAb generated as part of the University of Oxford and UCB collaboration to be able to neutralise all three pseudotypes. The number of broadly reactive and neutralising mAbs generated in the project, reflects similar high throughput screening or single cell sorting approaches to discovery of other broadly reactive mAbs

[99, 190]. The mechanism of neutralisation by the mAbs was investigated using a combination of cell-based immunofluorescence assays, immunoprecipitation experiments, and competition assays. The expression of chimeric EBOV GP with the MARV GP fusion loop (FL) was established and investigated as a method of identifying mAbs with key contacts in the EBOV FL. Inhibition of thermolysin cleavage was identified as a feature of six of seven mAbs tested.

In Chapter 5, epitope mapping studies using competition for binding to GP and binding to thermolysin cleaved GP, combined with data from Chapter 4, showed that the panel contained broadly reactive mAbs against a range of overlapping epitopes in the GC and base regions of the GP, that were neither reliant on contacts in the FL, nor required for NPC1-C (host receptor) binding. In addition, peptide phage display and an initial EM structure of 11886 fab in complex with EBOV GP, generated by Dr Amar Parvate, elucidated the epitope of 11886 further. 11886 appears to bind a portion of the GP between the FL base and GC that has previously been identified as the target of other broadly reactive mAbs, EBOV-520 and ADI-15946. Importantly, these mAbs (or their derivatives) have already been evaluated in animal models and are broadly protective. Intriguingly, 11886 is distinguished from these mAbs by its lack of binding to cleaved GP, and its lack of sensitivity to the replacement of the EBOV GP FL with that of MARV GP. This suggests EBOV-520 and ADI-15946, whilst targeting similar footprints have different modes of binding and possibly of neutralisation to 11886; higher resolution structural data is required to determine if 11886 uses molecular mimicry or allosteric effects like EBOV-520 and ADI-15946 [115, 191]. From the current data presented here, 11886 has emerged as the leading candidate for further evaluation in *in vivo* models and in higher resolution structural studies.

This Chapter is primarily a discussion of the future directions of the work presented in this thesis.

6.2 Authentic, wild type virus neutralisation

As established by the Viral Hemorrhagic Fever Immunotherapeutic Consortium (VIC), not all virus neutralisation assays are equivalent; all neutralisation assays in that study correlated with mAb-mediated protection *in vivo*, but not necessarily with each other [173]. Different pseudotype viruses have different stringencies compared to authentic virus, as illustrated in the development of the

REGN-EB3 cocktail, where mAb REGN3470 strongly neutralised EBOV GP coated VSV pseudotype virus, but not authentic virus [182]. Differences in (i) particle morphology, (ii) amount and arrangement of GP on the virus particle, (iii) presence of sGP or additional proteins involved in attachment and invasion, and (iv) cell type and incubation times in assay set up, between pseudotyped and authentic virus assays likely account for these discrepancies. *In vitro* neutralisation is still a good predictor of *in vivo* efficacy, but using a single assay may be misleading for some mAbs [173] . The S-FLU assay used in this thesis is not widely used by other groups nor included in the VIC, but has also been shown to correlate well with *in vivo* efficacy [3]. It is a simple, safe, and easily adapted system for initial screening of mAbs [219]. However, it is clear that testing in more than one assay would be beneficial for candidate selection. Authentic virus assays require BSL-4 facilities, and, as discussed in Section 4.7.1, a collaboration with Dr Thomas Strekker, Institut für Virolgie, Philipps-Universität Marburg, was initiated for *in vitro* studies of authentic wild type *Ebolavirus* neutralisation; it is hoped this work will resume after the COVID-19 pandemic.

6.3 Humanisation

The rabbit origin of the antibodies generated in this thesis needs addressing if any are to progress towards the clinic. Non-human origin products are at greater risk of causing hypersensitivity and adverse events, and have shorter half-lives than human antibody products [266]. In addition, the ability of the mAbs to engage Fc effector functions in NHP models of infection and humans may be compromised without a human Fc. The V region sequences of the broadly reactive rabbit mAbs have been humanised *in silico* by Dr Kerry Tyson, UCB. These sequences will be expressed with human constant regions and recombinant antibody assessed for expression, binding and *in vitro* neutralisation to ensure they are comparable to the original rabbit mAbs.

6.4 Fc effector functions

Neutralisation is not the only mechanism by which mAbs can offer protection against *Ebolaviruses* [174]. Antibody dependent complement deposition (ADCD), cellular cytotoxicity (ADCC), cellular phagocytosis (ADCP), neutrophil phagocytosis (ADNP), and NK cell activation and degranulation engagement by *Ebolavirus* GP mAbs can all be assayed *in vitro*. Whilst neutralisation correlates with

protection, across a large number of mAbs assayed, non- or partially neutralising mAbs can offer protection *in vivo* by engaging Fc effector functions, and conversely neutralising mAbs that cannot engage phagocytosis can by ineffective *in vivo* [174].

However engaging Fcγ Receptors (FcγR) on immune cells to trigger these functions can be a double edged sword, especially if recruiting the cells that are permissive to infection by the virus [267]. Increased efficacy must be balanced against the risk of antibody-dependent enhancement (ADE) of infection. Many broadly protective mAbs have not been assessed *in vitro* for their ability to trigger Fc effector functions, others have been engineered to improve Fc receptor (FcR) engagement and others to ablate it (Table 1).

EBOV-520 is able to engage ADCP, ADCC and activate NK cells, although the latter only poorly [99]. However a LALA mutant variant was taken forward into an NHP challenge study on the basis of increased efficacy compared to the original mAb when given at low doses in an EBOV challenge study in mice [191, 207]. LALA (L234A/L235A) mutations in the hinge of the Fc decrease the affinity of the mAb for FcγRIIa and FcγRIIIa, and therefore decrease ADCC, ADCP and complement recruitment [268]. The EBOV-520-LALA partner mAb, EBOV-548, was assessed with intact Fc in order that the combined cocktail was able to engage some Fc effector functions [191]. The increased protection of EBOV-520 LALA *in vivo*, was hypothesised to be due to the decrease in ADE of infection, leading to an overall net gain in efficacy [207].

Concerns over antibody dependent enhancement (ADE) should not be dismissed. ADE has been shown to be a driver of severe disease in Dengue virus infections, and like Dengue viruses, *Ebolaviruses* infect some of the very cells that express FcγRs and enact the Fc-mediated phagocytosis and cell killing [269]. Evidence of ADE of *Ebolavirus* infection *in vitro* by sub-neutralising concentrations of antibody has been demonstrated with sera and mAbs against a range of GP epitopes [207, 270, 271].

In counterbalance to this, non-neutralising mAbs have shown efficacy in animal challenge studies and have a net protective effect as part of cocktails in humans [173, 174, 177, 182]. In addition,

and in contrast to the strategy taken with EBOV-520, the components of the MP134 cocktail (ADI-15878 and ADI-27734) were poor at activating NK cells in their original form, but glyco-engineered to improve this particular activity [2]. Afucosylation of IgG1 Fc increases the affinity of interactions with FcγRIIIa, promoting ADCC and also ADCP, in part by NK cells [272]. An afucosylated version of the MP134 cocktail ($MP134^{AF}$) was more efficient at activating NK cells *in vitro*. NK activation has been identified as a key protective mechanism of *Ebolavirus* mAbs [174]. The $MP134^{AF}$ cocktail was more potent than MP134 and did not appear to enhance disease in guinea pigs when given at lower doses [2]. When tested in NHPs, $MP134^{AF}$ was fully protective when given as a single 25mg/kg dose four days post infection [214] (Table 2). The potency of other EBOV GP mAbs has also been enhanced by production in cell lines that decrease fucosylation [174, 273]. Sufficient dosing may be all that is required to minimise the threat of sub-protective and infection-enhancing concentrations of antibody in patients. Careful consideration should perhaps therefore be taken regarding the ability of broadly reactive mAbs used against different *Ebolaviruses* to equally protect against all; ensuring mAb concentrations that are protective against one virus do not act to enhance infection by a different virus towards which they are less potent inhibitors.

Whilst outside the initial scope of this thesis, assessing the ability of the mAbs presented here is of interest to ascertain whether (i) non-neutralising mAbs may effectively engage FcγRs and may be of interest to advance into *in vivo* studies, (ii) whether current leading candidate 11886 is able to also engage Fc effector functions. The latter is important when considering if any Fc engineering is necessary to either abolish or increase FcγR engagement, or when selecting suitable candidates for inclusion in a cocktail with 11886. The assays conducted by Gunn *et al*. primarily use antigen coupled beads [174]. Assays such as these could be complemented with assays using pseudotype viruses and transfected cells to more closely mimic full length GP in the context of the virus and infected cells. For example, assessing the ability of mAbs to fix complement in the presence of GP expressing cells, or comparison of *in vitro* neutralisation with and without complement proteins. ADNP and ADCP assays using transfected cells or pseudotype viruses and effector cells labelled with different dyes could be established.

6.5 *In vivo* efficacy

The *in vitro* characterisation of these mAbs described in this thesis has identified 11886 as a definite candidate for *in vivo* assessment. Initial assessment in mouse challenge models will provide evidence as to whether to continue *to in vivo* experiments with larger animals. To prove that mAbs or cocktails are broadly protective will require multiple experiments with different challenge viruses. Evaluation of the breadth of protection in mice is limited as a BDBV model does not exist, and SUDV infection requires interferon receptor knock out animals in which the disease is not uniformly lethal (Table 2) [82]. It will therefore be necessary to use guinea pig or ferret models to assess breadth of efficacy [274, 275]. Any *in vivo* challenge experiment with replication competent virus will require collaboration with teams with appropriate BSL-4 facilities and expertise and potentially significant additional resources.

6.6 Monoclonal antibody cocktails

Combining mAbs into cocktails for treatment has two major advantages; increased potency, and reduced chance of viral escape. Increasing the potency of any drug is desirable as it reduces the overall dose necessary to achieve protection, reducing both cost of each treatment, and the side effects of the drug. Both are particularly important in the context of mAb therapies for ED, as mAb therapies are high cost and will be delivered in a context where single doses will be more practicable than repeat dosing and adverse reactions more challenging to treat. A cocktail of mAbs can be more effective than its components by: (i) targeting multiple sites and functions of the GP at once; (ii) increasing the polyfunctionality of the treatment by including mAbs that are individually effective at neutralisation or engaging particular Fc effector functions; (iii) increasing breadth of efficacy against multiple species of virus; and (iv) co-operative binding or other synergistic effects leading to a result that is more potent than the sum of its individual parts. By targeting multiple epitopes with a single treatment the chance of selecting viable escape mutants that can evade all components of the cocktail at the same time also becomes highly unlikely. Cocktails are already an established feature of mAb therapies for ED, with ZMapp and REGN-EB3 the most obvious examples, however multiple cocktails of broadly reactive mAbs have been assessed *in vivo*: EBOV-520-LALA and EBOV-548; BDBV289 and BDBV 223; MP134

(ADI-15878 and ADI-23774); 040, 66-6-3, 6662 and 6541 (Table 2) [2, 3, 94, 190, 191, 214, 215]. Rational cocktail design based on *in vitro* data is essential as the cost and ethical considerations of *in vivo* challenge studies limits the iterative experimentation that might be necessary to develop the most effective combinations.

6.6.1 Synergy

Synergy can be achieved by pairing antibodies acting on non-independent pathways targeting different antigens, or by mAbs acting on the same antigen. The latter can be achieved by binding of the first mAb expanding the kinetic window for action of the second mAb. This can be critical in systems where antigens are exposed to humoral immunity for very short periods of time such as recently observed for mAbs against RH5, a malaria antigen, where one non-neutralising mAb slows the invasion of the parasite, thereby increasing the time that the second neutralising mAb has to act on the parasite before it invades a red blood cell [276]. Alternatively, synergy may be achieved by co-operative binding; the binding of the first mAb increases the affinity of the interaction of the second mAb and the antigen. Examples of cooperative binding have been observed in pairs of GP mAbs. The EBOV-520 and EBOV-548 cocktail was developed by deliberately screening for partners that enhanced EBOV-520 binding after an observation that several GC mAbs had this property [99, 191]. As discussed in Section 5.8.4, EBOV-548 increases availability of the EBOV-520 epitope in the 3_{10} pocket on uncleaved GP, but does not improve EBOV-520 binding to GP to the same extent as removing the MLD and GC [191]. Similarly, FVM09, a non-neutralising mAb, displaced the same portion of the GC (β17-β18 loop) in order to improve binding of both another 3_{10} pocket binding mAb, ADI-15496, and a similar, though less broadly reactive, murine mAb m8C4 [115, 192]).The *in vitro* effect of cooperativity on FVM09 and m8C4 clearly translates into *in vivo* protection in mouse challenge: 25mg/kg of FVM09 or m8C4 alone offered 0% and 30% survival respectively, but when given together at just 10mg/kg of each mAb, 100% survival was achieved [192].

Due to the relationship of the epitope of 11886 to EBOV-520 and ADI-15946, it will be worth investigating whether the same type of synergy is observed with mAbs that can displace the β17-β18 loop. However, due to the difference observed between the thermolysin sensitivity of 11886 binding

compared to the 3_{10} pocket mAbs, it may be that a different interaction with the GC improves the action of 11886.

I have adapted the S-FLU neutralisation assay format to be able to observe synergy as defined using Bliss additivity (which compares an experimental value to a predicted value of independent interactions only) in a way analogous to the growth inhibition assays used to study synergy between malaria mAbs [276]. Using human mAbs from the Oxford/UCB panel I have established that competing mAbs show sub-additive neutralisation, and that non-competing mAbs have independent effects across a range of concentrations (data not shown). However, no pair of mAbs yet tested has shown synergy in this assay. The rabbit mAbs are yet to be assessed, but combinations of 11886 with non-competing GC/RBR mAbs 11883, 11897 and 11889 which potentially showed enhanced binding in the competition IFA will be tested.

In the literature, therapeutic mAb cocktails for *Ebolavirus* infection, and hence investigations of synergy, have largely focused on GP. Toxicity of host-acting treatments raise many concerns due to the vital role of cellular receptors for the virus and the broad tropism of the virus. However, *in vitro*, antibodies targeting host attachment receptor Siglec-1 (sialic acid-binding Ig-like lectin 1) were shown to offer broad protection from *Ebolavirus* infection to dendritic cells, and many small molecules act as inhibitors by altering the endosomal environment necessary for GP processing [219, 277]. Combining host acting and virus acting drugs, that act synergistically could possibly allow a reduction in dose that brings the toxicity of the host-acting components into an acceptable range for short term treatment of acute infection, without reducing efficacy. Alternatively, combinations of mAbs and small molecule inhibitors that bind directly to the GP or immune modulators could also be investigated.

6.6.2 Escape mutants

Generating viral escape mutants by passaging viruses in the presence of mAbs can identify residues critical for binding or function of the mAbs, and to identity mutations that, if they arise in the wild, may affect the efficacy of treatment. One hypothetical advantage of broadly reactive mAbs is that the conserved nature of their epitopes across *Ebolavirus* species, may make them less amenable to viral escape without a fitness cost to the virus, as conserved epitopes are more likely to have functional

importance to the virus. However, some broadly reactive mAbs in the literature have been able to select for escape mutants that affected neutralisation (Table 1); possibly because their broad reactivity is reliant on key interactions with a few highly conserved contacts in a larger surface of unconserved residues, which, if lost, abrogate activity. For example, the escape mutant P634H selected for by passaging BDBV GP chimeric EBOV-eGFP in the presence of BDBV223, could not be neutralised by BDBV223 or the other MPER/HR2 mAbs in the study [100]. Similarly the conserved residues of the GP2 N-terminus appear to be amenable to viral escape; EBOV-520 neutralisation was abrogated by the N514Y escape mutant and ADI-15946 neutralisation was evaded by a K510E escape mutant [191, 211]. EBOV-548 and ADI-15878 also selected for escape mutations in their footprints in the GC and FL respectively [191, 211]. In addition, the I260R escape mutation discussed in Chapter 4 is another example of a residue that is conserved across species, but has arisen as an EBOV variant that causes human disease, and can cause the loss of neutralisation of GC mAbs (section 4.5).

Escape mutants may be more successfully avoided by the use of mAb cocktails. This has been effective for mAbs targeting other infectious diseases including Zika, SARS and Rabies [278-280]. *In vitro* assessment of broadly reactive *Ebolavirus* GP mAbs suggests that cocktails are better at reducing the chance of escape mutants arising. Compared to the generation of complete escape mutants when passaged virus was passaged with the individual mAbs, no escape mutants were detected for the EBOV-520 and EBOV-548 cocktail, and the the MP134 cocktail (ADI-15878 and ADI-24773) generated only one partial escape mutant [2, 191]. However, it was also noted that the rVSV-SUDV GP N514Y EBOV-520 escape mutant was not fully neutralised by EBOV-548 nor the cocktail [191]. Both of these cocktails contain only two components and it may be that inclusion of additional mAbs would abolish even these partial escape mutants.

In an extension of the collaboration that identified the P6 escape mutant, I260R, Professor Peter Halfmann, University of Wisconsin-Madison had agreed to generate escape mutants for 11886 *in vitro* by passage of EBOVΔVP30 viruses [246]. However, due to the COVID19 pandemic, this project was paused before data could be generated. It is hoped that this collaboration will be resumed in the near future and will provide information on key interactions of 11886 with the GP, which will help elucidate

its mechanism of neutralisation, help derive hypotheses for identifying partner mAbs to test for synergy, and allow the rational assessment of coverage of potential escape mutants by any cocktails including 11886.

6.7 Concluding remarks

Monoclonal antibodies are attractive therapeutic candidates in a broad range of diseases due to their specificity, their ability to engage the patient's immune system, and the increasing range of ways they can be engineered to improve their efficacy, safety and pharmacokinetic properties. Advances in high-throughput antibody discovery techniques in the last few decades have accelerated antibodies against infectious disease targets into the clinic in ever-growing numbers [281].

The development of therapeutic mAbs to treat ED in response to the 2014 EBOV outbreak, laid the foundations for therapeutic mAb discovery becoming a part of the rapid response to pandemic viruses. Antibodies and new antibody discovery methods now provide an express route to novel therapies against emerging pathogens as evidenced by the clinical assessment of mAbs targeting the Spike protein of SARS-CoV-2 within a year of the discovery of the virus [47].

However, this success should not lead to complacency. In the context of *Ebolaviruses*, new virus species continue to emerge, current viruses continue to evolve, and approximately one third of people treated with the current best therapies still did not survive EBOV infection [45]. In addition, therapies can only be tested for their efficacy in humans during an outbreak. Lessons from the delays to establishing clinical trials of potential therapies during the 2014 outbreak that led to an inconclusive result for ZMapp, helped rapidly initiate the trial in 2019 that provided evidence for the efficacy of two mAb therapies for ED [282]. Reducing delays in having candidate therapies ready to use in an outbreak requires preparation, not simply reaction. To be prepared for an outbreak caused by a different known, or yet to emerge, *Ebolavirus*, we require a robust pipeline of mAbs that are likely to be effective against multiple *Ebolaviruses*. Creating an arsenal with breadth and depth will provide the best defence. Having multiple mAbs that target each of a range of epitopes on the GP means we can immediately substitute individual components of cocktails if found to be less effective; either due to differences between animal models and human infection, or differences between viruses. In addition

having a range of antibodies to select from enables rarer co-operative and synergistic interactions to be identified and maintained as cocktails are adapted. The mAbs generated and characterised in this thesis contribute to this pipeline. For example, 11886, if ultimately protective in *in vivo* models, could be used as an alternative to ADI-23774 or EBOV-520, in the event of viral escape from these mAbs, or if their efficacy in NHPs does not translate to humans, as was the case for ZMapp. These three antibodies have a similar footprint on the virus, but likely rely on different key residues for interaction with the GP which may mean that mutations in the shared footprint that reduce the efficacy of one may not impact another. In contrast, other mutations such as I260R seem to globally impact a whole class of antibodies against a whole domain; these types of variants can be combatted by ensuring a range GP functions and domains are targeted within a cocktail. The panel of broadly reactive mAbs generated and characterised covers a range of epitopes and could provide other cocktail partners.

The mAb therapy field needs to address the issue of equitable access of these types of drugs when developing therapies for infectious diseases that predominantly affect countries without their own manufacturing capacity and limited health care resources. Monoclonal antibody therapies require intravenous transfusions, often in high doses, and are high cost [283]. More potent therapies could reduce the cost per dose, as could improvements in manufacturing techniques that improve yields. Alternatively, delivery of broadly reactive mAbs by viral vector platforms could provide long lived prophylaxis for communities most at risk of exposure to *Filoviruses* for which there are no current vaccines at low cost [284, 285].

Even if the mAbs presented here are ultimately unsuccessful *in vivo*, they may find alternative use in point of care diagnostic assays to identify *Ebolavirus* infections [286] or inform rational vaccine design to improve vaccine efficacy and generate more broadly protective responses [100, 287].

Therefore the isolation and initial *in vitro* characterisation of rabbit mAbs against *Ebolavirus* GP presented here contributes to the field of broadly reactive mAbs that are undergoing evaluation in preparation for an inevitable outbreak of ED caused by an *Ebolavirus* other than EBOV. 11886 in particular represents a lead therapeutic candidate for assessment *in vivo*.

7 Appendices

7.1 Appendix 1: Additional materials and methods

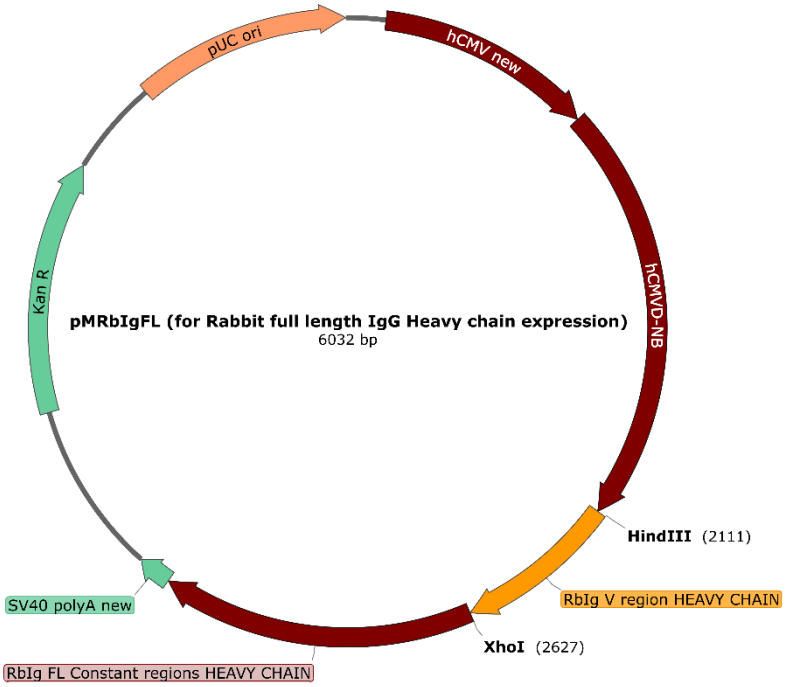
7.1.1 List of buffers

Table 6. Components of commonly used buffers

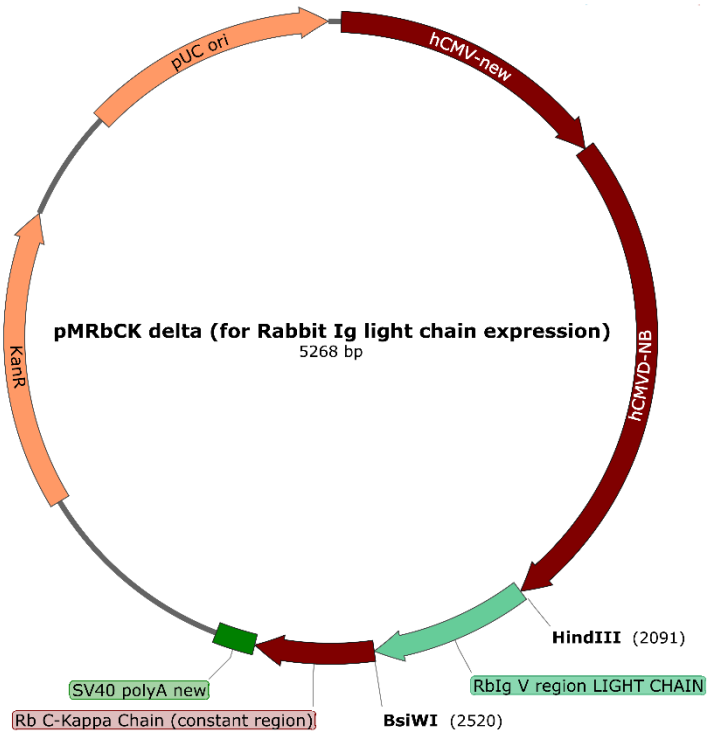
Buffer	**Components**
B cell culture media	RPMI, 10% FBS, 1% Penicillin/Streptomycin, 2% HEPES, 1% glutamine, 0.1% 2-mercaptopehtanol, 1% Normocin
Carbonate coating buffer	dH_2O, 15 mM Na_2CO_3, 35 mM $NaHCO_3$
CTAG elution buffer pH 7.4	dH_2O, 2 M $MgCl_2$, 20 mM Tris-HCl,
D10 media	DMEM (Sigma-Aldrich, D5796), 10% FBS (Sigma-Aldrich, F9665), 100U Pencillin/Streptomycin, 2mM L-Glutamine
FACS/iQUE buffer	PBS, 1% FBS, 0.02% Sodium azide
HM buffer pH 7.5	dH_2O, 20 mM MES, 20 mM HEPES, 130 mM NaCl, 10mM $CaCl_2$
RAB-9 cell culture media	RPMI, 10% FBS, 1% Penicillin/Streptomycin
Viral growth media	DMEM, 0.1% BSA (Sigma-Aldrich, A0336), 10mM HEPES (Sigma-Aldrich, H0887), 100U Penicillin/Streptomycin, 2mM L-Glutamine

7.1.2 Plasmid maps

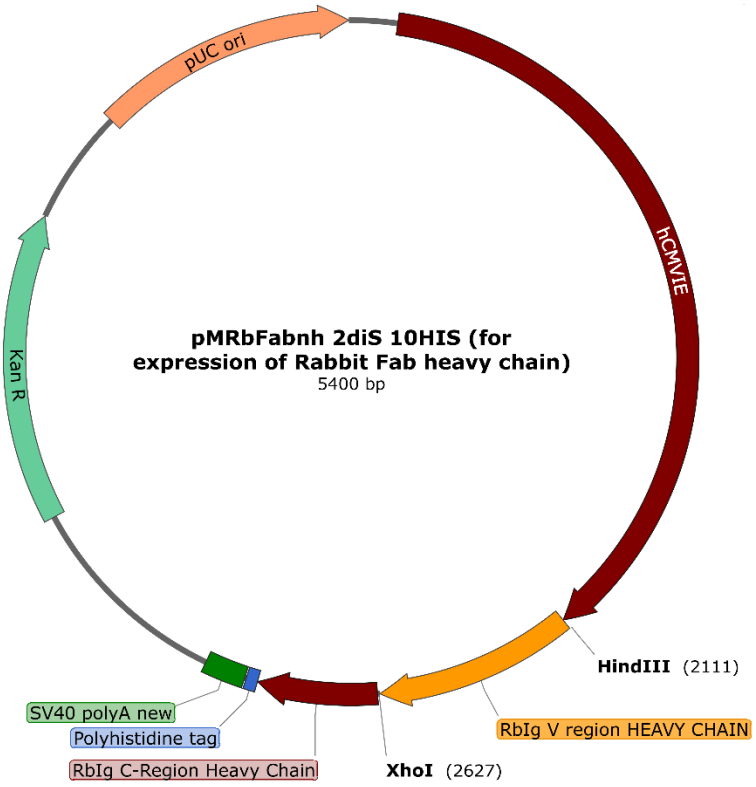
7.1.2.1 Vector for expression of Rabbit full length IgG Heavy chain



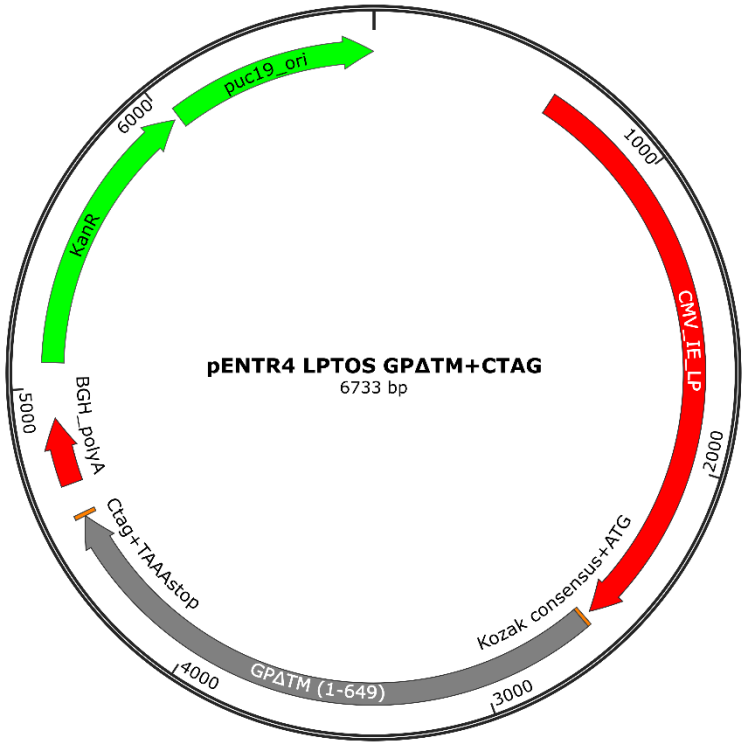
7.1.2.2 Vector for expression of Rabbit IgG Light chain



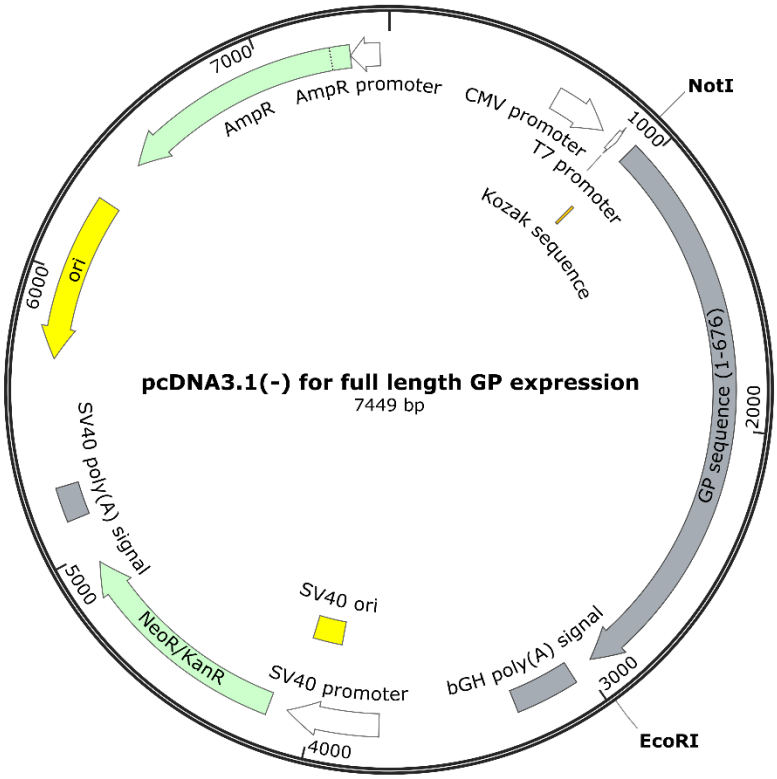
7.1.2.3 Vector for expression of Rabbit Fab Heavy chain



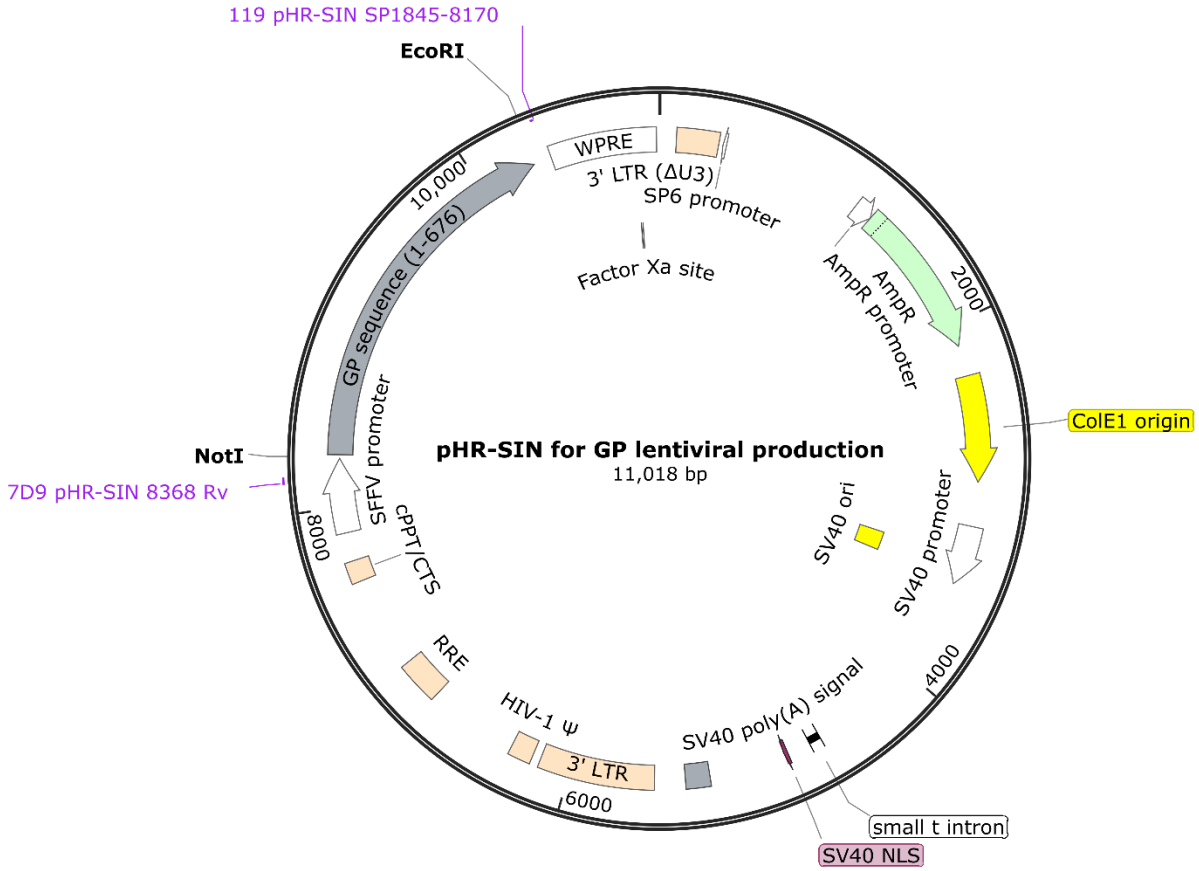
7.1.2.4 Vector for transient expression of soluble GP (GPΔTM)



7.1.2.5 pcDNA3.1(-) vector for transient expression of full length GPs on cell surface



7.1.2.6 pHR-SIN transfer vector for lentiviral production



7.1.3 Cell lines

Table 7. Cell lines and sources

Cell line	Source	Code	Publication
ExpiHEK 293F	Thermo Fisher	A14527	
HEK 293T	William Dunn School of Pathology, University of Oxford		
RAB-9	ATCC	CRL-1414	
MDCK SIAT-1	Townsend Lab		[288]
EBOV GP MDCK SIAT-1 (E-SIAT)	In house, Townsend Lab	n/a	[3]
SUDV GP MDCK SIAT-1	In house, Townsend Lab	n/a	[3]
BDBV GP MDCK SIAT-1	In house, Townsend Lab	n/a	[3]
MARV GP MDCK SIAT-1	In house, Townsend Lab	n/a	
I260R EBOV GP MDCK SIAT-1	In house, Townsend Lab	n/a	
TAFV GP MDCK SIAT-1	This study	n/a	
M/E GP MDCK SIAT-1	This study	n/a	
E/M GP MDCK SIAT-1	This study	n/a	
RCA MDCK SIAT-1	This study	n/a	
ACR MDCK SIAT-1	This study	n/a	
ACA MDCK SIAT-1	This study	n/a	

7.1.4 Oligonucleotide primers

Table 8. Primers used in GP subcloning

#	Primer	Sequence (5'-3')	Source
1	EBOVGP_ACR_F	GGGCTTCCCCGCATGCAGATACGTGCACAAGG	This study
2	EBOVGP_RCA_F	GGGCTTCCCCAGATGCGCATACGTGCACAAGG	This study
3	EBOVGP_ACA_F	GGGCTTCCCCGCATGCGCATACGTGCACAAGGTGTCCG	This study
4	EBOVGP_ACR_R	CCTTGTGCACGTATCTGCATGCGGGGAAGCCC	This study
5	EBOVGP_RCA_R	CCTTGTGCACGTATGCGCATCTGGGGAAGCCC	This study
6	EBOVGP_ACA_R	CGGACACCTTGTGCACGTATGCGCATGCGGGGAAGCCC	This study
7	pcDNA3.1(-)seqFOR	CTGCTTACTGGCTTATCG	This study
8	pcDNA3.1(-)seqREV	GTCGAGGCTGATCAG	This study
9	119 pHR-SIN SP1845-8170	CTTGCATGCCTGCAGGTC	Townsend Lab
10	7D9 pHR-SIN 8368 Rv	GAATTAACCAATCAGCCTGCTTCTCG	Townsend Lab
11	Seq primer for GPdTM (Note: does not anneal in Makona GP)	CCAGCAGCTACTAC	This study
12	200207.1 conserved GP seq	CCAGATGCAGATACGTGCACAA	This study
13	200218.1 MARV GP seq	CCACCACCATGTACAGAGGCAAAGTG	This study
14	200219.1 Makona GP seq	CCACCCTCTGCGCGAG	This study

7.2 Appendix 2: Rabbit monoclonal antibody sequences

ID	Heavy chain	Light chain
11886	AAGCTTCGAAGCCACCATGGAGACTGGGCTGCGCTGGCTTCTCCTGGTCGCTGTG CTCAAAGGTGTCCAGTGTCAGTCGGTGGAGGAGTCCGGGGGTCGCCTGGTCACG CCTGGGACACCCCTGACACTCACCTGCATAGTCTCTGGATTCTCCCTCAGTAACTA CTATATGAGCTGGGTCCGCCAGGTTCCAGGGAAGGGGCTGGAATGGATCGGAGT CATTGGTTGGAGTGGTACCTCATCCTACGCGAGCTGGGCGAAAGGCCGATTCACC ATCTCCAAAACCGCGTCGACCACGGTGGATCTGAAAATCACCAGTCCGACAACCG AGGACACGGCCACCTATTTCTGTGCCAGAGTCCTATATATTGGTGGTGGTTTTAAT TACTATGATGCTTTTGATCCCTGGGGCCCAGGCACCCTGGTCACCGTCTCGAG	AAGCTTCGAAGCCACCATGAACATGAGGGCCCCCACTCAGTTGCTGGGGCTCCTGCTGC TCTGGCTCCCAGGTGCCAGATGTGATGTTGTGATGACCCAGACTCCAGCCTCCGTGTCTG AACCTGTGGGAGGCACAGTCACCATCAAGTGCCAGGCCAGTCAGAGTATTGGTAGTAAT TTAGCCTGGTATCAGCAGAAACCAGGGCAGCCTCCCAAGCTCCTGATCTATGCTGCATCC ACTCTGGCATCTGGGGTCCCATCGCGGTTCAAAGGCAGTGGATCTGGGACAGAGTTCAC TCTCACCATCAGCGACCTGGAGTGTGCCGATGCTGCCACTTACTACTGTCAATGTACTTAT TATGATAGTAGTTATGTTTATAATAATTTTGGCGGAGGGACCGAGGTGGTGGTCAAACG TACG
11880	AAGCTTCGAAGCCACCATGGAGACTGGGCTGCGCTGGCTTCTCCTGGTCGCTGTG CTCAAAGGTGTCCAGTGTCAGTCGGTGGAGGAGTCCGGGGGTCGCCTGGTCACG CCTGGGACACCCCTGACACTCACCTGCATAGTCTCTGGATTCTCCCTCAGTAACTA CTATATGAGCTGGGTCCGCCAGGTTCCAGGGAAGGGGCTGGAATGGATCGGAGT CATTGGTTGGAGTGGTACCTCATCCTACGCGAGCTGGGCGAAAGGCCGATTCACC ATCTCCAAAACCGCGTCGACCACGGTGGATCTGAAAATCACCAGTCCGACAACCG AGGACACGGCCACCTATTTCTGTGCCAGAGTCCTATATATTGGTGGTGGTTTTAAT TACTATGATGCTTTTGATCCCTGGGGCCCAGGCACCCTGGTCACCGTCTCGAG	AAGCTTCGAAGCCACCATGGACACGAGGGCCCCCACTCAGTTGCTGGGGCTCCTGCTGC TCTGGCTCCCAGGTGCCAGATGTGATGTTGCGATGACCCAGACTCCAGCCTCCGTGTCTG AACCTGTGGGAGGCACAGTCACCATCAAGTGCCAGGCCAGTCAGAGTATTGGTAGTAAT TTAGCCTGGTATCAGCAGAAACCAGGGCAGCCTCCCAAGCTCCTGATCTATGCTGCATCC ACTCTGGCATCTGGGGTCCCATCGCGGTTCAAAGGCAGTGGATCTGGGACAGAGTTCAC TCTCACCATCAGCGACCTGGAGTGTGCCGATGCTGCCACTTACTACTGTCAATGTACTTAT TATGATAGTAGTTATGTTTATAATAATTTTGGCGGAGGGACCGAGGTGGTGGTCAAACG TACG
11889	AAGCTTCGAAGCCACCATGGAGACTGGGCTGCGCTGGCTTCTCCTGGTCGCTGTG CTCAAAGGTGTCCAGTGTCAGTCGTTGGAGGAGTCCGGGGGAGACCTGGTCAAG CCTGGGGCATCCCTGACACTCACCTGCACAGCCTCTGGATTCTCCTTCAGTAGCAG CTACTGGATATGCTGGGTCCGCCAGGCTCCAGGGAAGGGGCTGGAGTGGATCGC ATGCATTTATGCTGGTAGTAGTGGTAGCACTTACTACGCGAGCTGGGCGAAAGGC CGATTCACCATCTCCAAAACCTCGTCGACCACGGTGACTCTGCAAATGACCAGTCT GACAGCCGCGGACACGGCCACCTATTTCTGTGCGACCGCTTATGCTGTTTCTGCTC CTTTTGGTTATACTCTTTTTAAATACTTTGAATTGTGGGGCCCAGGCACCCTGGTCA CCGTCTCGAG	AAGCTTCGAAGCCACCATGGACATGACCCCCACTCAGCTGCTGGGGCTCCTGCTGCTCTG GCTCCCAGGTGCCAGATGTGCCGTCGTGCTGACCCAGACTGCATCCCCCGTGTCTGCACC TGTGGGAGGCACAGTCACCATCAAGTGCCAGGCCAGTCAGAACATTTACAGCGATTTAG CCTGGTATCAGCAGAAACCAGGGCAGCCTCCCAAGCTCCTGATCTATGATACATCCAATC TGGCATCTGGGGTCTCATCGCGGTTCAAAGGCAGTAGATCTGGGACAGAGTTCACTCTC ACCATCAGCGACCTGGAGTGTGCCGATGCTGCCACTTACTACTGTCAAGCCTATTATTAT AGTAGTAGTAGTGGTGATACTACTTTCGGCGGAGGGACCGAGGTGGTGGTCAAACGTA CG

11883	AAGCTTCGAAGCCACCATGGAGACTGGGCTGCGCTGGCTTCTCCTGGTCGCTGTG CTCAAAGGTGTCCAGTGTCAGGAGCAGCTGGTGGAGTCCGGGGGAGGCCTGGTC CAGCCTGAGGGATCCCTGACACTCACCTGCACAGCCTCTGGATTCTCCTTCAGTAG CAACTGCTGGAGATGCTGGGTCCGCCAGGCTCCAGGGAAGGGGCTGGAGTGGAT CGCATGCGTTTGTGCTGGTAGGAGTGGTGGTACCACTTACTACGCGAGCTGGGCG AAAGGCCGATTCACCATCTCCAAAACCTCGTCGCCCACGGTGACTCTTCAAATGAC CAGTCTGACAGCCGCGGACACGGCCACCTATTTCTGTGCGAGAGCTGGTTATGAT GATTATGGTGACGCTTCCTTCTTTAACTTGTGGGGCCCAGGCACCCTGGTCACCGT CTCGAG	AAGCTTCGAAGCCACCATGAACATGAGGGCCCCCACTCAGCTGCTGGGGCTCCTGCTGC TCTGGCTCCCAGGTGCCAGATGTGCCGTCGTGCTGACCCAGACTGCATCCCCCGTGTCTA CACCTGTGGGAGGCACAGTCACCATCAAGTGCCAGGCCAGTCAGAACATTTACAGTAAT TTAGCCTGGTATCAGCAGAAACCAGGGCAGCCTCCCAAGCTCCTGATCTATTCTGCATCC ACTCTGGCATCTGGGGTCCCATCGCGGTTCAAAGGCAGTGGATCTGGGACAGAGTACAC TCTCACCATCAGCGACCTGGAGTGTGCCGATGCTGCCACTTACTACTGTCAAAGGTATTA TTATCTTAGTGGTAGTGCTGATAATACTTTCGGCGGAGGGACCGAGGTGGTGGTCAAAC GTACG
11892	AAGCTTCGAAGCCACCATGGAGACTGGGCTGCGCTGGCTTCTCCTGGTCGCTGTG CTCAAAGGTGTCCAGTGTCAGTCAATGGAGGAGTCCGGGGGTCGCCTGGTCACA CCTGGGACACCCCTGACACTCACCTGCACAGTCTCTGGATTCTCCCTCAGTATCTAT TCAATGAGCTGGGTCCGCCAGGCTCCAGGGAAGGGGCTGGAATGGATCGGAATT ATTTATCTTGGTGATAGGGCATACTACGCGAGCTGGGCGAAAGGCCGATTCACCA TCTCCAAAACCTCGTCGACCACGGTGGATCTGAAAATCACCAGTCCGACAACCGA GGACACGGCCACCTATTTCTGTGCCAGAGTTGCTGGTTATGCTGGTTATGGTTATG CGTTCTATGATGCTTTTGATCCCTGGGGCCCAGGCACCCTGGTCACCGTCTCGAG	AAGCTTCGAAGCCACCATGAACACGAGGGCCCCCACTCAGCTGCTGGGGCTCCTGCTGC TCTGGCTCCCAGGTGCCAGATGTGATGTTGTGATGACCCAGACTCCAGCCTCCGTGTCTG AACCTGTGGGAGGCACAGTCACCATCAAGTGCCAGGCCAGTCAGAGCATTGGTAGTAAT TTAGCCTGGTATCAGCAGAAACCAGGGCAGCCTCCCAAGCTCCTGATCTATGCTGCATCC ACTCTGGCATCTGGGGTCCCATCGCGGTTCAAAGGCAGTGGATCTGGGACAGAGTTCAC TCTCACCATCAACGGCGTGCAGTGTGACGATACTGCCACTTACTACTGTCAATGTACTTAT TATGCTAGTACTTATGTGGCTGAGACTTTCGGCGGAGGGACCGAGGTGGTGGTCAAACG TACG
11897	AAGCTTCGAAGCCACCATGGAGACTGGGCTGCGCTGGCTTCTCCTGGTCGCTGTG CTCAAAGGTGTCCAGTGTCAGTCGGTGGAGGAGTCCGGGGGTCGCCTGGTCACG CCTGGGACACCCCTGACACTCACCTGCACAGTCTCTGGGTTCGACCTCAGTAGCTA TGCAATGGGCTGGGTCCGCCAGGCTCCAGGGAAGGGGCTGGAATACATCGGAAT CATTAATTATAGTGGTAACAGATATTACGCGAGCTGGGCGAAAGGCCGATTCACC ATCTCCAGAACCTCGACCACGGTGGATCTGTCAATGACCAGTCTGACAACCGAGG ACACGGCCACCTATTTCTGTGCCAGAGGGGGTTATGATGATTATGGTTATGTGTCC TACTTTGACATCTGGGGCCCAGGCACCCTGGTCACCGTCTCGAG	AAGCTTCGAAGCCACCATGGACATGAGGGCCCCCACTCAGCTGCTGGGGCTCCTGCTGC TCTGGCTCCCAGGTGCCAGATGTGCCGTCGTGTTGACCCAGACTCCAGCCTCTGTGTCTG CACCTGTGGGAGGCACAGTCACCATCAATTGCCAGGCCAGTCAGAATGTTTATGGTTTAT TGGCCTGGTATCAACAGAAACCAGGGCAGCCTCCCAAGCTCCTGATCTATTCTGCATCCA CTCTGGCATCTGGGGTCCCATCGCGATTCAAAGGCAGTGGATCTGGGACACAGTTCACT CTCACCATCAGCGACCTGGAGTGTGCCGATGCTGCCACTTACTACTGTCAAAGGTATTAT TATAGTAGTGGTACTACTGAGACTACTTTTGGCGGAGGGACCGAGGTGGTGGTCAAACG TACG
11878	AAGCTTCGAAGCCACCATGGAGACTGGGCTGCGCTGGCTTCTCCTGGTCGCTGTG CTCAAAGGTGTCCAATGTCAGTCGGTGGAGGAGTCGGGGGGTCGCCTGGTCACG CCTGGGACACCCCTGACACTCACCTGCACAGTCTCTGGATTCTCCCTCAGCAGCAA CGACATGAACTGGGTCCGCCAGGCTCCAGGGAAGGGGCTGGAGTGGATCGGACA CATTTGGAGTGGTGGTAGTACATACTACCCGAGCTGGGCGAGAGGCCGATTCACC ATCTCCAAAACCTCGACCACGGTGGATCTGAAAATCACCAGTCCGACAAGCGAGG	AAGCTTCGAAGCCACCATGGACACGAGGGCCCCCACTCAGCTGCTGGGGCTCCTGCTGC TCTGGCTCCCAGGTGCCAGATGTGATGTTGTGATGACCCAGACTGCATCCCCCGTGTCTG CACCTGTGGGAGGCACAGTCACCATCAAGTGCCAGGCCAGTGAGAACATTGATAACTTA TTGGCCTGGTATCAGCTGAAACCAGGGCAGCCTCCCAAGCTCCTGATCTACAGGGCATC CACTCTGGCATCTGGGGTCCCATCGCGGTTCAAAGGCAGTGGATCTGGGACAGAGTTCA CTCTCACCATCAGCGGCGTGCAGTGTGACGATGCTGCCACTTACTACTGTCAAAGCAATT

	ACACGGCCACCTATTTCTGTGCCAGAGGGCCTGTTAGTGACATCTGGGGCCCAGG CACCCTGGTCACCCGTCTCGAG	ATTATGGCTTTTATTATGGTATGACTTTCGGCGGAGGGACCGAGGTGGTGGTCAAACGT ACG
11888	AAGCTTCGAAGCCACCATGGAGACTGGGCTGCGCTGGCTTCTCCTGGTCGCTGTG CTCAAAGGTGTCCAGTGTCAGCAGCTGGAGGAGTCCGGGGGAGACCTGGTCAAG CCTGAGGGATCCCTGACACTCACCTGCACAGCCTCTGGATTCTCCTTCAGTAGCAG CTACTGGATATGCTGGGTCCGCCAGGCTCCAGGGAAGGGGCTGGAGTGGATCGC ATGCATTTATGCTGGTAGTAGTGGTAGCACTTACTACGCGAGCTGGGCGAAAGGC CGATTCACCATCTCCAAAACCTCGTCGACCACGGTGACTCTGCAAATGACCAGTCT GACAGCCGCGGACACGGCCACCTATTTCTGTGGGAGAGCTGGTTGGGATGATTAT GGTGATGCTTCCTACTTTAACTTGTGGGGCCCAGGCACCCTGGTCACCGTCTCGAG	AAGCTTCGAAGCCACCATGGACACGAGGGCCCCCACTCAGCTGCTGGGGCTCCTGCTGC TCTGGCTCCCAGGTGCCAGATGTGCCGTCGTGCTGACCCAGACTGCATCCCCCGTGTCTA CACCTGTGGGAGGCACAGTCACCATCAAGTGCCAGGCCAGTCAGAACATTTACAGTAAT TTAGCCTGGTATCAGCAGAAACCAGGGCGGCCTCCCAAGCTCCTGATCTATGGTGCATCC ACTCTGGCATCTGGGGTCCCCTCGCGGTTCAGAGGCAGTGGATCTGGGACAGAGTACAC TCTCACCATCAGCGGCGTGCAGTGTGCCGATGCTGCCACTTACTACTGTCAAAGGTATTA TTATCTTAGTGGTAGTGCTGATAATACTTTCGGCGGAGGGACCGAGGTGGTGGTCAAAC GTACG
11887	AAGCTTCGAAGCCACCATGGAGACTGGGCTGCGCTGGCTTCTCCTGGTCGCTGTG CTCAAAGGAGTCCAGTGTCAGGAGCTGGTGGAGTCTGGAGGGGGTCTGGTCCAG CCGGGGGAATCCCTGAAACTCTCCTGCAAAGCCTCTGGAATCGACTTCAGTAGAT ATGGCATTAGCTGGGTCCGCCAGGCTCCAGGGAAGGGGCTGGAGTGGGTCGCAT ACATTTATCCTCGTTTTGGTATCAGAAAGTACGCGAACTCTGTGAGGGGCCGATTC ACCATCTCCAGCGACAACGCCCAGAACACGGTGTTTCTGCAGATGACCAGTCTGA CAGCCTCGGACACGGCCACCTATTTCTGTGCAAGAGGGGATATAATAACCACGAC GCCTAGTATATTTAGAATATGGAACTACTTTAACTTGTGGGGCCCAGGCACCCTGG TCACCGTCTCGAG	AAGCTTCGAAGCCACCATGGACATGAGGGCCCCCACTCAGCTGCTGGGGCTCCTGCTGC TCTGGCTCCCAGGTGCCACATTTGCAGCCGTGCTGACCCAGACACCATCGCCCGTGTCTG CAGCTGTGGGAGGCACAGTCACCATCAAGTGCCAGTCTAGTCAGAATATTTATAATAGT GACTACTTAGCCTGGTTTCAGCAGAAACCAGGGCAGCCTCCCAACCTCCTGATCTATGAT GCATCCAGTCTGGCATCTGGGGTCCCAGATAGGTTCAGCGGCAGTGGATCTGGGACACA GTTCACTCTCACCATCAGCGGCGTGCAGTGTGACGATGCCGCCACTTACTACTGTCTAGG CGCTTATGATGATGCTGATACGTTTACTTTCGGCGGAGGGACCGAGGTGGTGGTCAAAC GTACG
11879	AAGCTTCGAAGCCACCATGGAGACTGGGCTGCGCTGGCTTCTCCTGGTCGCTGTG CTCAAAGGTGTCCAGTGTCAGTCGGTGGAGGAGTCCGGGGGTCGCCTGGTCACG CCTGGGACACCCCTGACACTCACCTGCACAGTCTCTGGAATCGACCTCAGTAACTA TGCAGTGGGCTGGGTCCGCCAGGCTCCAGGGAAGGGGCTGGAATGGATCGGAA TCATTGGCGGTGATAATACCACATACTGCGCAACCTGGGCGAAAGGCCGATTCAC CATCTCCAAAACCTCGTCGACCACGGTGGATCTGAAAATCACCAGTCCGACGACC GAGGACACGGCCACCTATTTCTGTGCCAGAGACACCTATCCTGGTGATGGTACTAT TGCTTTTAACATTTGGGGCCCAGGCACCCTGGTCACCGTCTCGAG	AAGCTTCGAAGCCACCATGGACATGAGGGCCCCCACTCAGCTGCTGGGGCTCCTGCTGC TCTGGCTCCCAGGTGCCAGATGTGCATTCGAATTGACCCAGACTCCATCCTCCGTGGAGG CAGCTGTGGGAGGCACAGTCACCATCAATTGCCAGGCCAGTCAGAGCATTAGCACGCAC TTATCCTGGTATCAGCAGAAACCAGGGCAGCCTCCCAAGCTCCTGATCTATGGTGTATCC ATTCTGGCATCTGGGGTCCCATCGCGGTTCAGCGGCAGTGGATCTGGGACAGAGTTCAC TCTCACCATCAGCGGCGTGCAGTGTGACGATGCTGCCACTTACTACTGTCAATATGCTTG GTATGATAGTAGCCAAAATACTTTCGGCGGAGGGACCGAGGTGGTGGTCAAACGTACG

11895	AAGCTTCGAAGCCACCATGGAGACTGGGCTGCGCTGGCTTCTCCTGGTCGCTGTG CTCAAAGGTGTCCAGTGTCAGGAGCAGCTGGAGGAGTCCGGGGGAGGCCTGGTC CAGCCTGAGGGATCCCTGACACTCACCTGCACAGCCTCTGGATTCTCCCTCAGTAC CAGCTACTGCATATGCTGGGTCCGCCAGGCTCCAGGGAAGGGGCTGGAGTGGTT CGCGTGTATTTATGTTGGTGGTAGTGGTAGCACTTACTACGCGAGCTGGGCGAAA GGCCGATTCACCATCTCCAAAGCCTCGTCGACCACGGTGACTCTGCAAATGACCA GTCTGACAGCCGCGGACACGGCCACCTATTTCTGTGCGCGAGATTGGACTCATTA TGTTGGTGGTGGTAATTATGGCTATGTGACTCGGTTGGATCTCTGGGGCCAGGGC ACCCTGGTCACCGTCTCGAG	AAGCTTCGAAGCCACCATGGACATGAGGGCCCCCACTCAGCTGCTGGGGCTCCTGCTGC TCTGGCTCCCAGCAACCACATTTGCCCAAGTGCTGACCCAGACTGGATCCCCCGTGTCTG CACCTGTGGGAGGCACAGTCAGCATCGATTGCCAGTCCAGTCAGAGTGTTTATGGTAAC TTCTTATCGTGGTATCAGCAGAAACCAGGGCAGCCTCCCAAGCTCCTGATCTATTATGCA GCCACTCTGACATCTGGGGTCCCATCGCGGTTCAAAGGCAATGGATCTGGGACACAGTT CACTCTCACCATCAGCGACCTGCAGTGTGACGATGCTGCCACTTACTACTGTGCAGGCGG TGGCAGTTCTGATATTGATAATATTTTCGGCGGAGGGACCGAGGTGGTGGTCAAACGTA CG
11894	AAGCTTCGAAGCCACCATGGAGACTGGGCTGCGCTGGCTTCTCCTGGTCGCTGTG CTCAAAGGTGTCCAGTGTCAGTCGGTGGAGGAGTCCGGGGGTCGCCTGGTCACG CCTGGGACACCCCTGACACTCACCTGCACAGTCTCTGGATTCTCCCTCAGTATCTAT ACAATGACCTGGGTCCGCCAGGCTCCAGGGAAGGGGCTGGAATGGATCGGAGTC ATTAGTGGTGGTGGTTCGACATACTACGCGACCTGGGCGAAAGGCCGATTCACCA TCTCCAAAACCTCGTCGACCACGGTGGGTCTGAAAATCACCAGTCCGACAACCGA GGACACGGCCACCTATTTCTGTGCCAGAGGGGGCTACGATGACTATGCTGATTAT CTATTAGGTTTCGATGGGTTTGATCCCTGGGGCCCAGGCACCCTGGTCACCGTCTC GAG	AAGCTTCGAAGCCACCATGAACATGAGGGCCCCCACTCAGCTGCTGGGGCTCCTGCTGC TCTGGCTCCCAGGTGCCAGATGTGATGTTGTGATGGCCCAGACTCCAGCCTCCGTGTCTG AACCTGTGGGAGGCACAGTCACCATCAAGTGCCAAACCAGTCAGAGCATTGGTAGTAAT TTAGCCTGGTATCAGCAGAAACCAGGGCAGCCTCCCAAACTCCTGATCTACAGGGCGTC CACTCTGGCATCTGGGGTCTCATCGCGGTTCAAAGGCAGTGGATCTGGGACACAGTTCA CTCTCACCATCAGCGACCTGGAGTGTGCCGATGCTGCCACTTACTACTGTCAGTGTACTT ATTATGGTAATACTTATTATAATACTTTCGGCGGAGGGACCGAGGTGGTGGTCAAACGT ACG
11893	AAGCTTCGAAGCCACCATGGAGACTGGGCTGCGCTGGCTTCTCCTGGTCGCTGTG CTCAAAGGTGTCCAGTGTCAGTCATTGGAGGAGTCCGGGGGAGACCTGGTCAAG CCTGGGGCATCCCTGACACTCACCTGCAAAGCCTCTGGACTCTCCTTCAGTGTCAT CTACTACATGTGCTGGGTCCGCCAGGCTCCAGGGAAGGGGCCTGAGTGGATCGC ATGCATTGATGCTGGTAGTAGTGGTACCACTTACTACGCGACCTGGGCGAAAGGC CGATTCACCATCTCCAAAACCTCGTCGACCACGGTGACTCTGCAAATGACCAGTCT GACAGCCGCGGACACGGCCACCTATTTCTGTGCGAGAGATGATGCTTATGTTGCT GGTTTTTACTTTGACTTGTGGGGCCCAGGCACCCTGGTCACCGTCTCGAG	AAGCTTCGAAGCCACCATGGACATGAGGGCCCCCACTCAGCTGCTGGGGCTCCTGCTGC TCTGGCTCCCAGGTGCCAGATGTGCATTCGAATTGACCCAGACTCCAGCCTCCGTGTCTG CAGCTGTGGGAGGCACAGTCACCATCAAGTGCCAGGCCAGTCAGAGCATTAGCAGTTG GTTTGCCTGGTATCAGCAGAAACCAGGGCAGCCTCCCAAGCTCCTGATTTACCAGGCATC CACTCTGGCATCTGGGGTCCCATCGCGGTTCAAAGGCAGTGGATCTGGGACAGAGTTCA CTCTCGCCATCAGCGGCGTGCAGTGTGACGATGCTGCCACTTACTACTGTCAAAGTTATT CTGGTAGTGATCGTACTACTTTCGGCGGAGGGACCGAGGTGGTGGTCAAACGTACG
11890	AAGCTTCGAAGCCACCATGGAGACTGGGCTGCGCTGGCTTCTCCTGGTCACTGTG CTCAAAGGTGTCCAGTGTCAGGAGCAGCTGGTGGAGTCCGGGGGAGGCCTGGTC CAGCCTGAGGGATCCCTGACACTCACCTGCAAAGCCTCTGGATTCGACTTCAGTAA CAATGCAATGTGCTGGGTCCGCCAGGCTCCAGGGAAGGGGCCTGAGTGGATCGC ATGCATTGGTAGTGGTGATGCCAGCGCGTACTACGCGAACTGGGTGAATGGCCG ATTCACCGTCTCCAGAAGCACCAGCCTAAGCACGGTGACTCTGCAAATGACCAGT	AAGCTTCGAAGCCACCATGGACATGAGGGCCCCCACTCAGCTGCTGGGGCTCCTGCTGC TCTGGCTCCCAGGTGCCAGATGTGCCGTCGTGCTGACCCAGACTCCAGCCTCCGTGTCTG CAGCTGTGGGAGGCACAGTCACCATCAAGTGCCAGGCCAGTCAGAACATTTACACAAAT TTAGCCTGGTATCAGCAGAAACCAGGGCAGCCTCCCAAGGTGTTGATCTATGGTGCATC CACTCTGGCATCTGGGGTCCCATCGCGGTTCAAAGGCAGTGGATCTGGGACAGAGTTCA CTCTCACCGTCAGCGACCTGGAGTGTGCCGATGCTGCCACTTACTACTGTCAAAGCTATG

	CTGACAGCCGCGGACACGGCCACCTATTTCTGTGCGAGTAAAGGGTTTTATGTTATTGCTGGTGTTGCTTATGCCTATGAATCTTACTTTAACTTCTGGGGCCCAGGCACCCTGGTCACCGTCTCGAG	TTGATAGTCGTAGTAGTACTGATTCTTTCGGCGGAGGGACCGAGGTGGTGGTCAAACGTACG
13171	AAGCTTCGAAGCCACCATGGAGACTGGGCTGCGCTGGCTTCTCCTGGTCGCTGTGCTCAAAGGTGTCCAGTGTCAGGAGCAGCTGGAGGAGTCCGGGGGAGACCTGGTCCAGCCTGAGGGATCCCTGACACTCACCTGCACAGCCTCTGGACTCGACTTCAGTAGCAACTACTGGATATGTTGGGTCCGCCAGGCTCCAGGGAAGGGGCTGGAGTGGATCGCATGCATTTATGTTGGTAGTTATGGTAGCACCCACTACGCGAGCTGGGCGAAAGGCCGATTCACCATCTCCAAAACCTCGTCGACCACGGTGACTCTGCAGATGACCAGTCTGACAGCCGCGGACACGGCCACCTATTTCTGTGCGAGAAATGGTGTTTATGTTGGTAGGAATTATGCCCTCTTTAACTTGTGGGGCCCAGGCACCCTGGTCACCGTCTCGAGT	AAGCTTCGAAGCCACCATGAACATGAGGGCCCCCACTCAGCTGCTGGGGCTCCTGCTGCTCTGGCTCCCAGGTGCCAGATGTGCCGTCGTGCTGACCCAGACTCCATCCCCCGTGTCTGCACCTGTGGGAGGCACAGTCACCATCAGGTGCCAGGCCAGTGAGAACATTTACAGCGATTTAGCCTGGTATCAGCAGAAACCAGGGCAGCCTCCCAAGCTCCTTGTCTATTATGCATCTGATCTGGCATCTGGGGTCCCATCGCGGTTCAAAGGCAGTAGATCTGGGACAGAGTTCACTCTCACCATCAGCGACCTGGAGTGTGCCGATGCTGCCACTTACTACTGTCAGGCCTATGATTGGATTGGTGATAGTGCTGCTAGTACTTTCGGCGGAGGGACCGAGGTGGTGGTCAAACGTACG
11885	AAGCTTCGAAGCCACCATGGAGACTGGGCTGCGCTGGCTTCTCCTGGTCGCTGTGCTCAAAGGTGTCCAGTGTCAGTCGTTGGAGGAGTCCGGGGGAGACCTGGTCAAGCCTGGGGCATCCCTGACACTCACCTGCACAGCCTCTGGATTCTCCTTCAGTGTAATCTACTACATGTGCTGGGTCCGCCAGGCTCCAGGGAAGGGGCCTGAGTGGATCGCATGCATTGATGCTGGTAGTAGTGGTACCACTTACTACGCGAGCTGGGCGAAAGGCCGATTCACCATCTCCAAAACCTCGTCGACCTCGGTGACTCTGCAAATGACCAGTCTGACAGCCGCGGACACGGCCACCTATTTCTGTGCGAGAGATGATGCTTATGTGGCTGGTTTTTACTTTGACTTGTGGGGCCCAGGCACCCTGGTCACCGTCTCGAG	AAGCTTCGAAGCCACCATGGACATGAGGGCCCCCACTCAGCTGCTGGGGCTCCTGCTGCTCTGGCTCCCAGGTGCCAGATGTGCATTCGAATTGACCCAGACTCCAGCCTCCGTGTCTGCAGCTGTGGGAGGCACAGTCACCATCAACTGCCAGGCCAGTCAGAGCATTAGTAGTTGGTTAGCCTGGTATCAGCAGAGCCCAGGGCAGCCTCCCAAGCTCCTGATCTACGAAGCATCCAAACTGGCATCTGGGGTCCCATCGCGGTTCAGCGGCAGTGGATCTGGGACAGAGTTCACTCTCACCATCAGCGGCGTGCAGTGTGACGATGCTGCCACTTACTACTGTCAAAGTTATACTGGTAGTGATCGTACTACTTTCGGCGGAGGGACCGAGGTGGTGGTCAAACGTACG
11891	AAGCTTCGAAGCCACCATGGAGACTGGGCTGCGCTGGCTTCTCCTGGTCGCTGTGCTCAAAGGTGTCCAGTGTCAGTCGTTGGAGGAGTCCGGGGGAGGCCTGGTCCAGCCTGAGGGATCCCTGACACTCACCTGCACAGCCTCTGGATTCTCCTTCAGTAGCAACTATTGGATATGCTGGGTCCGCCAGGCTCCAGGGAAGGGGCTGGAGTGGATCGGATGCATTGATACTGGTGATGGTAGCACTTACTTCGCGAGTTGGGCGACAGGCCGATTCACCATCTCCAAAACCTCGAACACGGTGGATCTGAAAATGACCAGTCTGACAGCCGCGGACACGGCCACCTATTTCTGTGCGAGGGGGTATTATTACCGTTATGGTTGGCTTTATACTAGGGATTCCTACTTTACCTTGTGGGGCCCAGGCACCCTGGTCACCGTCTCGAG	AAGCTTCGAAGCCACCATGGACATGAGGGCCCCCACTCAGCTGCTGGGGCTCCTGCTGCTCTGGCTCCCAGGTGCCAGATGTGCCGTCGTGCTGACCCAGACTGCATCCCCCGTGTCTGCACCTGTGGGAGGCACAGTCACCATCAAGTGCCAGGCCAGTCAGAACATTTACAACAATTTAGCCTGGTATCAGCAGAAACCAGGGCAGCCTCCCAAGCTCCTGATCTATTCTGCATCCACTCTGGCATCTGGGGTCCCATCGCGGTTCAAAGGCAGTGGATCTGGGACAGAGTACACTCTCACCATCAGCGACCTGGAATGTGCCGATGCTGCCACTTACTACTGTCAAAGCTATTATGATAGTAGTAGTAGTGGAGATAATACTTTCGGCGGAGGGACCGAGGTGGTGGTCAAACGTACG

11882	AAGCTTCGAAGCCACCATGGAGACTGGGCTGCGCTGGCTTCTCCTGGTCGCTGTG CTCAAAGGTGTCCAGTGTCAGTCGCTGGAGGAGTCCGGGGGAGACCTGGTCAAG CCTGGAGGAACCCTGACACTCACCTGCAAAGCCTCTGGATTCTCCTTCAGTAGCGG CTACGACATATGTTGGGTCCGCCAGGCTCCAGGGAAGGGGCTGGAGTGGATCGG ATGTATTGCTAGTGGTGATAGTAGCACTTACTACGCGAACTGGGCGAAAGGCCGA TTCACCATCACCAGAAGCACCAGCCTAAACACGGTGACTCTGCAACTGAACAGTCT GACAGCCGCGGACACGGCCACGTATTTCTGTGCGAGGGGTGATTATGTTGGGAA TGGTTATGGAGACTACTTTACCTTGTGGGGCCCTGGCACCCTGGTCACCGTCTCGA G	AAGCTTCGAAGCCACCATGGACACGAGGGCCCCCACTCAGCTGCTGGGGCTCCTGCTGC TCTGGCTCCCAGGTGCCAGATGTGCCCTTGTGATGACCCAGACTCCAGCCTCCGTGTCTG CGGCTGTGGGAGGCACAGTCACCATCAATTGCCAGGCCAGTCAGAGTATTAGTGCTGCA TTAGCCTGGTATCAGCAGAAACCAGGGCAGCCTCCCAAGCTCCTGATCTACGATGCATCG AAATTGGCATCTGGGGTCCCATCGCGGTTCAAAGGCAGTGGATCTGGGACAGAGTTCAC TCTCACCATCAGCGGTGTGCAGTGTGACGATGCTGCCACTTACTATTGTCAACAGGCTTA TAGTGATAGTAATACTTTCGGCGGAGGGACCGAGGTGGTGGTCAAACGTACG

7.3 Appendix 3: Peptide phage display experiment summary

ID	Species	Cross-reactivity	Exp	Library	R1 mAb concentration (µg/mL)	R2 subtraction	R2 mAb concentration (µg/mL)	Barcode	PCR	Total reads	Total reads in top 100 enriched seqs	Barcode compared in Z score analysis
11897	Rabbit	Pan-GP	2	9mer, linear	15	Chromopure Rabbit IgG, 011-000-003	5	Akey-BC1-p8	1 step	111288	4965	11
11878	Rabbit	Pan-GP	2	9mer, linear	4.5	Chromopure Rabbit IgG, 011-000-003	5	Akey-BC2-p8	1 step	114064	13701	11
11883	Rabbit	Pan-GP	2	9mer, linear	15	Chromopure Rabbit IgG, 011-000-003	5	Akey-BC3-p8	1 step	97911	5562	11
11886	Rabbit	Pan-GP	2	9mer, linear	15	Chromopure Rabbit IgG, 011-000-003	5	Akey-BC4-p8	1 step	115621	19703	11
11889	Rabbit	Pan-GP	2	9mer, linear	15	Chromopure Rabbit IgG, 011-000-003	5	Akey-BC5-p8	1 step	78886	1224	11
11891	Rabbit	SUDV	2	9mer, linear	15	Chromopure Rabbit IgG, 011-000-003	5	Akey-BC6-p8	1 step	111405	6099	11
11892	Rabbit	Pan-GP	2	9mer, linear	15	Chromopure Rabbit IgG, 011-000-003	5	Akey-BC7-p8	1 step	90682	3921	11
11881	Rabbit	Pan-GP	2	9mer, linear	15	Chromopure Rabbit IgG, 011-000-003	5	Akey-BC8-p8	1 step	112110	3921	11
11894	Rabbit	EBOV	2	9mer, linear	15	Chromopure Rabbit IgG, 011-000-003	5	Akey-BC9-p8	1 step	104098	7556	11
11895	Rabbit	SUDV, BDBV	2	9mer, linear	15	Chromopure Rabbit IgG, 011-000-003	5	Akey-BC10-p8	1 step	100319	28415	11
11886	Rabbit	Pan-GP	1	9mer, linear	15	Chromopure Rabbit IgG, 011-000-003	5	Akey-BC11- p8	1 step	113337	21708	12

66-3-9C	Human	Pan-GP	2	9mer, linear	15	Chromopure Human IgG, 009-000-003	5	Akey-BC12- p8	1 step	97947	1022	11
11897	Rabbit	Pan-GP	3	13mer, linear	15	Chromopure Rabbit IgG, 011-000-003	5	BC13_adpt	2 step	87693	52538	23
11878	Rabbit	Pan-GP	3	13mer, linear	15	Chromopure Rabbit IgG, 011-000-003	5	BC14_adpt	2 step	52130	40278	23
11883	Rabbit	Pan-GP	3	13mer, linear	15	Chromopure Rabbit IgG, 011-000-003	5	BC15_adpt	2 step	89763	73341	23
11886	Rabbit	Pan-GP	3	13mer, linear	15	Chromopure Rabbit IgG, 011-000-003	5	BC16_adpt	2 step	123666	112093	23
11889	Rabbit	Pan-GP	3	13mer, linear	15	Chromopure Rabbit IgG, 011-000-003	5	BC17_adpt	2 step	83101	10808	23
11891	Rabbit	SUDV	3	13mer, linear	15	Chromopure Rabbit IgG, 011-000-003	5	BC18_adpt	2 step	92825	28642	23
11892	Rabbit	Pan-GP	3	13mer, linear	15	Chromopure Rabbit IgG, 011-000-003	5	BC19_adpt	2 step	113621	96486	23
11881	Rabbit	Pan-GP	3	13mer, linear	15	Chromopure Rabbit IgG, 011-000-003	5	BC20_adpt	2 step	162587	157027	23
11894	Rabbit	EBOV	3	13mer, linear	15	Chromopure Rabbit IgG, 011-000-003	5	BC21_adpt	2 step	117986	55457	23
11895	Rabbit	SUDV, BDBV	3	13mer, linear	15	Chromopure Rabbit IgG, 011-000-003	5	BC22_adpt	2 step	92552	64184	23
FDRb6	Rabbit	None	3	13mer, linear	15	Chromopure Rabbit IgG, 011-000-003	5	BC23_adpt	2 step	n/a		
66-3-9C	Human	Pan-GP	3	13mer, linear	15	Chromopure Human IgG, 009-000-003	5	BC24_adpt	2 step	94278	9044	23

7.4 Appendix 4: Sequences enriched by mAb panel during peptide phage display

7.4.1 11897

13mer				9mer			
Sequence	**BC13**	**BC23**	**Z-score**	**Sequence**	**BC01**	**BC11**	**Z-score**
DGATLGRLVGGGG	6028	0	80.45	VAFLRGTGV	348	0	18.68
GADGLVGLVASGG	4500	0	68.87	FVSLIANGA	156	0	12.5
AGQVLGIHSLVIL	3106	0	56.75	VVQCATCAK	130	0	11.41
SGWRLAGVVESGG	2531	0	51.05	LGIYELVMT	108	0	10.4
GVGEVLGVVMGGG	2518	0	50.92	AMDCATCSV	106	0	10.3
EAACLMIPGACWD	1471	0	38.68	HACGTVACV	104	0	10.2
DRDRFWQMTMGGG	1429	0	38.11	KCGTAPCqL	99	1	9.81
AVNEVVGAYVSGG	1428	0	38.1	ACATCGVGG	94	0	9.7
AWEEVVGVVVNGG	1324	0	36.66	LGVYGDIVV	93	0	9.65
SGREMVGLVGGGG	1134	0	33.89	DMMVGRHVR	86	0	9.28
SGQVLALVVRqGG	1029	0	32.27	GGPWYLYGD	84	0	9.17
DARRFGEVVRQGG	961	0	31.17	LVGVLRNGG	84	0	9.17
qGLDAFWMNWFGG	909	0	30.31	MMRAIVNGA	80	0	8.95
ESMDMRAFVLNGG	888	0	29.95	REWIVGTGI	75	0	8.66
EELAREVRQGWVG	780	0	28.05	VIAGMGTVT	69	0	8.31
VYPYVVLVDGFGF	758	0	27.65	VPRWALFGD	68	0	8.25
THSPFLIYVLNSG	729	0	27.11	LGVYETIVL	62	0	7.88
GVGVEGGMVGWMV	703	0	26.62	YASLAANGL	59	0	7.68
GPDIARILLINGG	670	0	25.98	VPIFGqVVL	58	0	7.62
EGWRLAQVVGGGG	574	0	24.04	FVDMLFSGA	56	0	7.49
WAEEMVSLVGGGG	573	0	24.02	AYWLERSPE	56	0	7.49
EMARALETLDLTS	569	0	23.93	VRMLVVNGA	56	0	7.49
RGDGFVGAVVSGL	548	0	23.48	ATCGTNCVE	56	0	7.49
GVQELVGVVGGGG	493	0	22.27	LGIYSFALV	56	0	7.49
HESFTRLPMGWTD	442	0	21.08	GHRCGTVLC	53	0	7.28
RGSEMVRVVGGGG	392	0	19.84	ASYLLGTGL	50	0	7.07
DQSSLLVVLLNGG	388	0	19.74	IGIYSqVVL	49	0	7
DVDTIVAAVRAGG	380	0	19.54	ACGTNICGT	48	0	6.93
DGARIALVVLqGG	365	0	19.14	qKGCATCAL	48	0	6.93
SQPPWHQLSGYSA	363	0	19.09	SVGERLYGY	48	0	6.93
GEKISAVLqAASP	354	0	18.85	FAGLVLNGA	46	0	6.78
GGGMVVDPIRGPW	342	0	18.53	AAFISGTGL	46	0	6.78
AELRRLWGGGGGV	335	0	18.34	WSTLVANGA	45	0	6.71
GGREMVVLVGSGG	320	0	17.92	ASYLLGTGM	45	0	6.71
GGQWLAGVVLSGG	318	0	17.86	YGWLALNGA	45	0	6.71
VIPPSRLLSVISP	313	0	17.72	FGGCATCGW	45	0	6.71
GSSEVVGIIVSGG	307	0	17.55	GIAASRWTV	44	0	6.63
RVSDVVVAYINGL	307	0	17.55	DWTCGWTPC	44	0	6.63
GGMALGRVVQAGG	301	0	17.38	CATCDWGIT	43	0	6.56
MDVGFVGRVVGGW	296	0	17.23	ALPWYMWSD	43	0	6.56

PPGWREGWTELRA	294	0	17.18	FVAWMTNGL	43	0	6.56
GSREILAVVMGGG	294	0	17.18	LPVYSFVMG	43	0	6.56
RAWDVEFWFANGV	280	0	16.76	CATCGLSAT	43	0	6.56
DSWGPVFCVVNGC	278	0	16.7	YSGLLDSGV	41	0	6.4
VMWDLGGVAFRAM	278	0	16.7	RCATCSIGD	40	0	6.33
AWQELVGVVGGGG	274	0	16.58	LGIYSqIML	39	0	6.25
ESLYSRQVFGWGL	264	0	16.27	FTGLLVSGG	39	0	6.25
RGWELAAVVGSGG	262	0	16.21	LGRILLQGG	39	0	6.25
EHERLVGAIHSGA	261	0	16.18	RCGTGACSA	38	0	6.17
TPQAFIGAVQAGG	257	0	16.05	FVDMMFSGV	37	0	6.08
RGWVDSVQSGLWL	254	0	15.96	ADWIRGTGV	37	0	6.08
SVEDVIGVVGGGG	246	0	15.71	CATCVAGGV	36	0	6
AGEVLVGVVRNGG	234	0	15.32	VSIYGDIVI	36	0	6
SYAPSLLNGVTFS	229	0	15.15	DRLVGWFGM	36	0	6
IPLDLVRVVGGGG	224	0	14.99	FLGRYMSGW	36	0	6
FVPPSLLQSLYFS	219	0	14.82	MSRLVVNGA	35	0	5.92
DLCGWGFCGLGGV	210	0	14.51	LVGLIASGG	35	0	5.92
DSASMVGLVGGGG	203	0	14.26	LCGTNGCST	35	0	5.92
RGSGWHFQLPHGG	201	0	14.19	CATCAQTAE	34	0	5.83
DHFGLFSAFMNGV	193	0	13.91	IATRMSAGS	34	0	5.83
GGSVLAGVVGSGG	190	0	13.8	WVGLITNGV	34	0	5.83
GGEELVRLVGGGG	189	0	13.76	IAWLSATGI	33	0	5.75
SGQELLSVVGGGG	188	0	13.73	LMALVDSGV	33	0	5.75
GLWLHGVRRADGW	188	0	13.73	ASFIIGTGV	33	0	5.75
PRLASLLDSMRAG	182	0	13.5	VATSRLWWD	33	0	5.75
RHADFMMCALNGC	182	0	13.5	MAGIMLSGG	33	0	5.75
GAWVMGRRRVDGF	181	0	13.47	WRRLAVNGA	32	0	5.66
NKHAGLCRIENGC	178	0	13.36	MYSLVLNGA	32	0	5.66
RTFDGCWAGRWVC	177	0	13.32	LAMLKGTGI	32	0	5.66
GAYDLVRLVGGGG	175	0	13.24	FGGLIFNGG	32	0	5.66
GGGRLVALVGSGG	174	0	13.2	LATRERLSL	32	0	5.66
GGRELGRVVGSGG	174	0	13.2	GDLVGWTSV	31	0	5.57
VYFGGSVRRWDGW	174	0	13.2	VAGLVMSGV	31	0	5.57
GPDTVVGVVAGGG	171	0	13.09	FVGLAANGL	31	0	5.57
QARALLEAWLDRE	167	0	12.94	LGVVSSVAL	31	0	5.57
RGAELVGLVGGGG	167	0	12.94	AHFMLGTGI	31	0	5.57
HPTQLHLLMLNGG	164	0	12.82	AWLVINTGM	31	0	5.57
TSAGYLRAILNGG	164	0	12.82	LCGTNVCER	31	0	5.57
SEGDFWVRVMNGA	162	0	12.74	WYALSVNGA	31	0	5.57
PWMRAVGWWGSGG	158	0	12.58	ALWLLGTGV	31	0	5.57
DATRALVLVRNGG	158	0	12.58	ACGTGGCLG	31	0	5.57
NAGDMVVLVGGGG	158	0	12.58	MGIYGLTVV	31	0	5.57
SGEVMIGLVGGGG	151	0	12.3	GAFIVGTGL	30	0	5.48
RQAVRFLPVGWTD	151	0	12.3	VYGLVMNGA	30	0	5.48
RGEVLGLVVARGG	151	0	12.3	IGVVSSVMV	30	0	5.48
GCLVVDWGqVICV	150	0	12.26	LSVMSFAML	30	0	5.48
GWWYGRFGGGGRV	149	0	12.22	EACATCVNS	30	0	5.48
DVGSLVGLVGGGG	149	0	12.22	LAVAFLNGA	30	0	5.48

GGWDLVRVVGGGG	148	0	12.18	INLVGRFST	29	0	5.39
GGERLGDVVRRGG	147	0	12.13	ASYIAGTGL	29	0	5.39
GGMTLAGVVGSGG	144	0	12.01	LKGLVDSGV	29	0	5.39
GWMDGGFWVGTGL	143	0	11.97	GWQCGTSGC	29	0	5.39
DPARPMPFSRLWD	142	0	11.93	FYGLMANGV	29	0	5.39
SAARQCLLVRECR	142	0	11.93	LSIYGqVVV	29	0	5.39
GVEDVMRVVGGGG	140	0	11.84	WTALFVNGV	29	0	5.39
VVGDLVGVVGGGG	138	0	11.76	AAWLMGTGV	29	0	5.39
TGVEMVGVVGGGG	136	0	11.67	VGRWWMYSD	28	0	5.29
DSTKFVSAWLSGA	135	0	11.63	VAYVSTVST	28	0	5.29
EKRWRHEMEWGKP	135	0	11.63	TFGCGTAPC	28	0	5.29
RGHELGLVVAGGG	135	0	11.63	DLLRRLGTN	28	0	5.29

7.4.2 11878

13mer				9mer			
Sequence	**BC14**	**BC23**	**Z-score**	**Sequence**	**BC02**	**BC11**	**Z-score**
ARDYPRLPDGRYE	6395	0	85.38	qVWLDRVGG	725	0	27.01
GVGGRGGLLEVSL	3926	0	65.16	GYLVARSGI	687	0	26.29
DRGGSLVVFSRWG	3531	0	61.54	GMWLqRVGG	422	0	20.58
VGLEWEVQQARWG	3118	0	57.59	GYEWNRYVD	415	0	20.41
NDAGALMHFSRGV	2435	0	50.54	WRRNVESGA	410	0	20.28
EGEWGLqLARWGG	2006	0	45.68	GYLVDRYMA	265	0	16.3
GWLGGCFGGACWG	1216	0	35.29	EMNVARWGG	252	0	15.89
CVGGVCRGSGGWV	1122	0	33.86	AMWRDRVGG	227	0	15.08
NESQQFIYNRTVQ	908	0	30.4	DVWIRRTGG	220	0	14.85
GMVEYEIGRLRGW	690	0	26.44	GYMVNRVAS	191	0	13.83
RGGEWVVRGVKVG	613	0	24.91	qMWLQRMGG	189	0	13.76
ALWGVKQE	547	0	23.51	GLWIERMGG	189	0	13.76
DKLGFLVHFPRAD	523	0	22.98	FYLVNRVAG	183	0	13.54
HHFARHPDGTPIS	523	0	22.98	GFLVDRYGL	179	0	13.39
GFCRTGSGGWVGV	510	0	22.69	DYRWERVVG	178	0	13.35
GEDWqVMVGRWGG	481	0	22.03	AYMWNRNVD	172	0	13.12
AWLGGCGVGGCWG	439	0	21.04	GYLVNRSGT	172	0	13.12
AWMGGCWGGRCWG	375	0	19.43	GYVYNRAVA	170	0	13.05
VDCPVWLRRWGCG	372	0	19.36	GYMLDRGMV	169	0	13.01
EEKNREYKTWKME	358	0	18.99	GMWLVRMGG	164	0	12.82
HGMAWEYWqARQM	331	0	18.25	GYLVSRYGS	164	0	12.82
QQERRWPDGSRLV	316	0	17.83	GYVMNRVMG	152	0	12.34
AMGGRKV	313	0	17.75	GYLFDRINL	148	0	12.17
GTQWDVSLSRWGG	310	0	17.66	SYIWNRIVD	147	0	12.13
GDDWAVVLGRWGG	276	0	16.66	GYLVNRYGS	144	0	12.01
ASAEWRYHQDRMT	262	0	16.23	SMWLSRVGG	144	0	12.01
YLGGGVKG	258	0	16.1	GMWLGRMGG	143	0	11.97
KEVALNGEKVRVA	250	0	15.85	SYLVNRNMT	143	0	11.97
DAAEWGVAVMRWG	224	0	15	GVWLRRVGG	141	0	11.88
VGRGGWVGEVVGL	221	0	14.9	AYRWDRVVV	138	0	11.75

NEDMKYNLDRWME	218	0	14.8	GYAVNRVFL	131	0	11.45
HPHYLRHPDGHPI	218	0	14.8	GYWLDRYMA	130	0	11.41
GGGRWEWRLPSGE	218	0	14.8	GYLLARSSV	129	0	11.36
LPAHALLLYPFSL	201	0	14.2	AYHFNRHVV	125	0	11.19
GIDYSYQLARYAV	198	0	14.1	KVWMDRVGG	125	0	11.19
EQKNEWAKKWPTQ	195	0	13.99	GYLVARNGS	124	0	11.14
RNPDGSATSAVLL	187	0	13.7	AYNWDRIVD	124	0	11.14
VVRGVKF	184	0	13.59	LLKNILAGA	123	0	11.1
VYGGRKME	183	0	13.55	ALWVARTGG	123	0	11.1
ASGWqVWLERWGG	164	0	12.83	AYAISRYGF	122	0	11.05
ERDFQYRLNRSML	164	0	12.83	GYVLDRVWM	122	0	11.05
RGWGGEWWKGRWG	161	0	12.71	AYLVNRSFS	121	0	11.01
GEMMKYRLNRSMV	160	0	12.67	ARLPDGSYR	121	0	11.01
RDGWSVWLGRWGG	160	0	12.67	GVQLSRWGG	121	0	11.01
LDPASRYAVLRSV	156	0	12.51	DYLMSRNAT	118	0	10.87
LQWDSVVLARWGG	154	0	12.43	DYLVqRFGF	118	0	10.87
GSGGCTARDGCWG	151	0	12.31	GYLVNRWGV	115	0	10.73
GSPFSVWLGRWGG	144	0	12.02	GYLWNRSVA	111	0	10.54
VGRGGWMGMVVGL	136	0	11.68	EVqVLRWGG	107	0	10.35
GGREGVWVLRWGG	133	0	11.55	GYLVERYGS	103	0	10.15
AWRGRVYGRENVG	129	0	11.37	GYLVLRYGV	103	0	10.15
GWLGGCGLGGCWG	125	0	11.19	GYIVDRYGT	102	0	10.1
GGWYGCWGGGCWG	115	0	10.74	GVWLDRVGG	102	0	10.1
GDCVVGYGCWGRV	114	0	10.69	GLWVSRMGG	101	0	10.05
RNTYEYRLNRVMG	109	0	10.45	MYLVSRYSS	101	0	10.05
MCWSVGGGDRRCG	105	0	10.26	AVWVSRVGG	100	0	10
VVRGGFVGTVVGL	104	0	10.21	GYWVTRYSV	98	0	9.9
SLISKNLAAGWGG	101	0	10.06	VVQCATCAK	98	0	9.9
ALGGVRM	101	0	10.06	RGGLYGVML	98	0	9.9
KGAWEVQVMRWGG	97	0	9.86	MYALDRVWL	96	0	9.8
GGVGGCVLGSCWG	95	0	9.76	GYWMTRYAT	95	0	9.75
HELARYLASGREH	93	0	9.65	GYYVNRSFL	93	0	9.65
GGLGGMVWDVAWS	89	0	9.44	LLKNINAGA	93	0	9.65
REMAVGGVRVEVA	89	0	9.44	GRNPDGSIA	91	0	9.54
GRRGAENVFGGWG	87	0	9.34	AFLLDRYMT	89	0	9.44
VEGAWILKGVKYE	86	0	9.28	GYRWNRLVT	88	0	9.38
GPVGGFVGGTSWG	86	0	9.28	DLWLRRMGG	87	0	9.33
SGYYARRPDGQVL	82	0	9.06	AYNWNRAVT	87	0	9.33
WFGRGSMVYASYF	82	0	9.06	EYALARYMG	87	0	9.33
GDVGGCWAGRCWG	80	0	8.95	EYLVRRYGV	87	0	9.33
PPSLQYLQARYGQ	78	0	8.84	EYVWNRTVM	87	0	9.33
EGEWGLqLARWG	78	0	8.84	AFQWDRLVD	86	0	9.28
CRMPSGSWASSSR	76	0	8.72	AYLVDRVFG	86	0	9.28
GTSWQVHLVRWGG	76	0	8.72	GLSLLRWGG	86	0	9.28
LDNHAMFLLRSGD	74	0	8.61	GYMVSRFAT	84	0	9.17
CREAGGGWRRTGG	73	0	8.55	GFLLDRYME	84	0	9.17
FLGGGAKG	72	0	8.49	GIEVMRWGG	83	0	9.11

MGRNPDGSERGMS	72	0	8.49	GYLLNRSSV	83	0	9.11
SWVGGCRSSGCWG	71	0	8.43	VATARRVVE	83	0	9.11
AWKGRAV	69	0	8.31	AYLVTRWGT	83	0	9.11
AHWGGCVGGTCWG	68	0	8.25	AYRMNRVVG	83	0	9.11
SVIGGCVGGSCWG	67	0	8.19	GYLMSRYGS	82	0	9.06
ETQWWGRARGWWP	67	0	8.19	GMWLRRTGG	81	0	9
ERTYSYWMARYSE	67	0	8.19	ASWYRDRFG	79	0	8.89
EDFGLVMARWGGV	65	0	8.07	CWHSNSCFV	78	0	8.83
AWGGEKW	64	0	8	GYILDRYMG	77	0	8.78
AVGGEKK	64	0	8	GFLVDRYAS	77	0	8.78
CRLPTGLAVVEEG	64	0	8	GFLINRWGT	77	0	8.78
SGWRLAGVVESGG	62	0	7.88	SFLMDRYMT	77	0	8.78
CRGGMCMGSGSWG	62	0	7.88	GYLVDRVSM	75	0	8.66
CESYSRLTSYCMD	61	0	7.81	ACATCGVGG	74	0	8.61
DLALWMMRTGGGD	61	0	7.81	GYVWNRTVG	73	0	8.55
GTRWELVVGRWGG	60	0	7.75	NYMVSRWST	72	0	8.49
SSVGGCGLSGCWG	60	0	7.75	GYVWNRLVD	72	0	8.49
CEVAMKGRKFQCG	59	0	7.69	RYFSGGSIS	72	0	8.49
AQSWALQLGRWGG	58	0	7.62	GYLVNRHSS	72	0	8.49
DNSYSYSLARNAV	57	0	7.55	DYMMDRYAV	72	0	8.49
DDEKTKRLREMME	57	0	7.55	GYVFNRAVS	71	0	8.43
ARKGEKV	57	0	7.55	QVqVARWGG	70	0	8.37
THSWTVWLGRWGG	56	0	7.49	SFLVNRYGI	70	0	8.37

7.4.3 11883

13mer				9mer			
Sequence	**BC15**	**BC23**	**Z-score**	**Sequence**	**BC03**	**BC11**	**Z-score**
SGWRLAGVVESGG	12814	0	122.26	LGIYELVMT	680	0	26.17
AWEEVVGVVVNGG	7165	0	88.24	GKWDVRNVG	159	0	12.62
RGGQPGRTDVGLD	3897	0	63.83	VVQCATCAK	154	0	12.42
DGPLALRLWQTGG	3672	0	61.88	GAVLQDVFS	130	0	11.41
DGRWNVFFVRSGG	3550	0	60.8	VGMYETVVI	127	0	11.28
VqERILGLVLNGG	2749	0	53.25	AMDCATCSV	127	0	11.28
VTDGTRVRGVDGT	2702	0	52.78	VCLGNPHCE	117	0	10.82
GVHDVVRWVGGGG	2534	0	51.06	LGVYETIVL	113	0	10.64
DASRCYHRGRqVC	2310	0	48.69	MGVFEVIMI	106	0	10.3
RGRVCGGGVAECG	1910	0	44.18	HACGTVACV	102	0	10.1
RNASVERLQATGQ	1846	0	43.41	MGVYEVVLI	93	0	9.65
NADRWVGLGMGWG	1831	0	43.23	KCGTAPCqL	92	1	9.47
GGQPGKPGVSTWG	1687	0	41.46	LGMYELVMM	88	0	9.39
SREDVFRWVGAGG	1396	0	37.66	VIAGMGTVT	87	0	9.33
RGGQMGRPESWRW	1336	0	36.83	IGIFETIAV	86	0	9.28
PRACWGMLGGCGQ	1031	0	32.3	VGVYELVVV	75	0	8.66
AGEVLVGVVRNGG	874	0	29.71	VGAYEMLVL	72	0	8.49
DGATLGRLVGGGG	769	0	27.85	ACATCGVGG	72	0	8.49
GGQPGRAEGSWGP	674	0	26.06	CATCDWGIT	69	0	8.31

MGGWDGMRWAEGV	666	0	25.9	LGVYELVVT	68	0	8.25
GAWFDRFGGRGGG	656	0	25.71	FTGLLVSGG	68	0	8.25
GWLAVYDKVMAGG	645	0	25.49	VGVYEVVML	57	0	7.55
VRSCADHALGVCR	641	0	25.41	FGGCATCGW	54	0	7.35
GPSLAVRLVATGG	607	0	24.72	ATCGTNCVE	52	0	7.21
EGWRLAQVVGGGG	577	0	24.1	CFASGRqVC	52	0	7.21
AGQVLGIHSLVIL	562	0	23.78	LCGTNGCST	51	0	7.14
RDSLWARLVGGGV	538	0	23.26	GAQVQDVMG	51	0	7.14
GGSWAVRVLGGGG	498	0	22.38	RCATCSIGD	50	0	7.07
ESLGLVGCVVNGC	492	0	22.24	LGVYEYVMA	50	0	7.07
GPDWALKVVGGKG	463	0	21.57	SWLVSSGRI	49	0	7
MGGQPGRPGGEGE	450	0	21.27	WVGLAMNGA	49	0	7
GSSYWLRVVGGGG	440	0	21.03	YSGLVINGG	48	0	6.93
EDSVWIRIAAGGG	413	0	20.37	GIAASRWTV	48	0	6.93
GSEDWVRIVGGGG	383	0	19.61	GAFHNANVG	47	0	6.86
LPPLRRNADGSPF	373	0	19.35	GAIAqDVMI	47	0	6.86
EVPPHLRLAALYG	367	0	19.2	LGIYqLVVL	46	0	6.78
MGGQPGRTNVHEG	353	0	18.83	CATCGSISV	45	0	6.71
QGGQRGRPETVIS	329	0	18.17	GAVMQDTML	45	0	6.71
AGGQMGRPEGWEG	316	0	17.81	FTDLYMSGR	45	0	6.71
TGGQVGRPEHGIE	307	0	17.55	IRGLVLNGA	44	0	6.63
GAERLVSVVLGGG	285	0	16.91	MVGFLINGG	43	0	6.56
ASGWAVRIMGGGG	283	0	16.85	FGGLIFNGG	43	0	6.56
RGGQTGRPAGEAV	248	0	15.77	CATCAQTAE	42	0	6.48
RGGQPGRPEGQEY	242	0	15.58	VGIYSLVVI	41	0	6.4
QGGQPGRPGGGGD	234	0	15.32	GAHFQDVMV	40	0	6.33
VGWGTVVRGVDGR	198	0	14.09	CATCASHLG	40	0	6.33
QEPRILRYWMNTG	194	0	13.94	DLCLDRGMC	40	0	6.33
GDAWSVRVVGGGG	192	0	13.87	LGVYEAMLI	39	0	6.25
GVEDVMRVVGGGG	187	0	13.69	DILRYFSNT	38	0	6.17
KDGVFLRVLGGGG	186	0	13.65	VGAYELVVV	37	0	6.08
GGEWMVRVVGGGG	183	0	13.54	IGAYELVML	37	0	6.08
SSRNPDGSWIAPS	183	0	13.54	AWLVASGRI	37	0	6.08
SAYSYMRLAVNGG	180	0	13.43	VGAYEFVVL	36	0	6
WGERLVGVVAGGG	175	0	13.24	FAGLVLNGA	36	0	6
AAEGGWRRIGSLS	172	0	13.13	WVGLMTNGA	35	0	5.92
AVAGPINAFAVFG	171	0	13.09	WVGLIqNGG	34	0	5.83
RRDLCRLPSGASC	170	0	13.05	YTGLVINGS	34	0	5.83
DWQNVGRLLGGGG	170	0	13.05	YSGLVMqGL	34	0	5.83
RCVSYDCGVMVWE	170	0	13.05	MSRYAVATK	33	0	5.75
GTEEVLRWVGGGG	168	0	12.97	FSGLLHNGG	33	0	5.75
SGAGFGRFMGGGG	165	0	12.86	IGIYEMIVT	33	0	5.75
GQGWAVRVLGGGG	158	0	12.58	GHRCGTVLC	33	0	5.75
MGRDMVGWVVNGG	153	0	12.38	EACATCVNS	33	0	5.75
VDSPYVRFIGGGG	152	0	12.34	VGAYELLVL	33	0	5.75
AGPDVLRWLGGGG	139	0	11.8	LCGTNVCER	32	0	5.66

EPPRLFLCLMNGC	134	0	11.58	KPGGHEATW	32	0	5.66
ARPRGWDGERWGG	130	0	11.41	VGVYEVVMI	32	0	5.66
DGLVRFWRFSDLG	126	0	11.23	LCGTGGCMA	32	0	5.66
RGGQTGRPEGHWS	124	0	11.14	LGVFEAISI	32	0	5.66
VDPFALRVMGGGG	123	0	11.1	IGNAFSVWG	32	0	5.66
GLVRGWDGVRWGG	123	0	11.1	WSRIVGGGG	31	0	5.57
GESWAVRVVGGKG	122	0	11.05	YANLVLqRL	31	0	5.57
GGWDVFGVVVNGG	121	0	11.01	LATRERLSL	31	0	5.57
TGGQIGRPGVEEG	121	0	11.01	VGVYEGLLI	31	0	5.57
GWRDMVRVVGGGG	118	0	10.87	LGIYqVVAL	31	0	5.57
GPEHYAHFASTGA	116	0	10.78	GSVYQDVMT	31	0	5.57
RGGQMGRPAGGEW	116	0	10.78	RLGHMFGTY	31	0	5.57
ASAWWHKGERFGG	114	0	10.68	ACGTDGCVV	30	0	5.48
MDVGFVGRVVGGW	113	0	10.64	LGIYqLVLL	30	0	5.48
CARKEWVRAGGQC	113	0	10.64	CATCGLSAT	29	0	5.39
GCVRSGRGVVECG	111	0	10.54	ACGTNICGT	29	0	5.39
DWVGRVCGVGGKC	111	0	10.54	GALIqDVMG	29	0	5.39
GLEYWLRVVGGGG	109	0	10.45	MGCATCGGS	29	0	5.39
RAAEAYRLALSGQ	109	0	10.45	FTGLFVNGS	29	0	5.39
GAEWWLRVVGGGG	108	0	10.4	VGVYEVVLL	29	0	5.39
GCRFDGVVFMCGV	105	0	10.25	GWVATRASV	29	0	5.39
qESRLYWLLTTGQ	104	0	10.2	IGVYESLIM	29	0	5.39
DEGYNVCTWEGQC	102	0	10.11	VWWqKLWGF	29	0	5.39
AGFSNQARMWTMT	102	0	10.11	MLRLLGGGV	29	0	5.39
PREWIMQIAANRG	99	0	9.96	RGGWRDVGT	29	0	5.39
RGQDVVRWVGGGG	97	0	9.85	AYCATCDAS	28	0	5.29
QGGQIGRPGGAVE	95	0	9.75	VMGKLLSGG	28	0	5.29
REGVFERLVGGGV	95	0	9.75	RSFHVRDIG	28	0	5.29
AGEGMVRFVGGGG	89	0	9.44	YWLCGTSLC	28	0	5.29
GPEQVEQQMKRGM	89	0	9.44	ACGTGGCLG	28	0	5.29
VGGQPGRPGEGEG	87	0	9.33	ERFLNTLGE	27	0	5.2
GGETLARVVLGGG	86	0	9.28	FAGLLFNGA	27	0	5.2
GGEGLVGVVLNGG	84	0	9.17	GAILQDVTS	27	0	5.2
AGGQKGRPGVTGE	83	0	9.11	CATCQSqVV	27	0	5.2
GQEYWVRVVGGQG	81	0	9	SCGTAGCLS	27	0	5.2

7.4.4 11886

9mer Experiment 1			
Sequence	**BC11**	**BC12**	**Z-score**
qDYLLRSHG	4375	0	67.46
NEYLLRARG	609	0	24.74
GKRCISWHC	577	0	24.08
SGRCRSGWC	525	0	22.97
GKFCqTCML	494	1	22.19
LRCRGPVCL	476	0	21.86
GCKVGKFGC	412	0	20.33
SDRCRGMLC	384	0	19.63

EKLCRSCqV	355	0	18.87
GRCRSSACS	354	0	18.84
qKVCRSSQC	346	0	18.63
GKFCQTCqV	343	0	18.55
ARRCMSWTC	334	0	18.3
DRCRSSTCN	333	0	18.28
GKVCKSLGC	323	0	18
SARCVSPVC	318	0	17.86
AGRCRSWTC	315	0	17.77
IVRGLGFFP	312	0	17.69
EGRCKGIVC	306	0	17.52
ADRCISFSC	300	0	17.34
GGRCRSVHC	291	0	17.08
DKVCRSWTC	255	0	15.99
SFPFGGGFP	242	0	15.57
STRCRTCVL	206	0	14.37
qKFCMTCLS	204	0	14.3
GSRCRSLVC	191	0	13.83
AELLRRSRG	181	0	13.46
ALRCRTCVV	179	0	13.39
RDFLNRSRG	179	0	13.39
AGRCRGFDC	177	0	13.31
GKFCYTCqV	170	0	13.05
GNRCRTCLV	169	0	13.01
RLLPSGILD	169	0	13.01
ESRCRGVDC	161	0	12.7
GYGTVAFVM	159	0	12.62
RILLDRARG	158	0	12.58
GGLGRDIMS	157	0	12.54
AERCRGFGC	150	0	12.26
GKVCRSCEq	149	0	12.21
VTPFSGSFP	149	0	12.21
DTWLPGRVG	146	0	12.09
EDYLRRSHG	145	0	12.05
NDRCRGVDC	144	0	12.01
DGKALYLGT	139	0	11.8
ERCRGPLCF	137	0	11.71
SGRCRGVMC	136	0	11.67
GGRCRGPEC	136	0	11.67
SARCRSFDC	134	0	11.58
GRCqSDWCR	128	0	11.32
MGRCRSAFC	127	0	11.28
NGRCKGIGC	126	0	11.23
GKFCRSCEq	119	0	10.91
RKYCMTCqQ	119	0	10.91
ARCRGPVCL	118	0	10.87

FFDDWARVG	118	0	10.87
NRCRSSMCV	118	0	10.87
PSRCISAVCRGqV	117	0	10.82
NGRCRSDAC	113	0	10.64
GRCRSDFCR	113	0	10.64
DRCRGIGCL	112	0	10.59
LGRLVDRqM	109	0	10.45
qKFCMTCVM	108	0	10.4
VGRCRGVDC	105	0	10.25
RERCRSLMC	104	0	10.2
VGRCRGVEC	103	0	10.15
VGRCRSGFC	103	0	10.15
VSRCKTCSL	102	0	10.1
GRCRGVDCH	102	0	10.1
GKYCqSCVL	100	0	10
NKKCLGVGC	99	0	9.95
LERCRGIGC	97	0	9.85
SKFCqSCIL	97	0	9.85
RCRSLACVN	96	0	9.8
DSRCKGEAC	96	0	9.8
AERCRSFSC	95	0	9.75
GSRCRGPVC	95	0	9.75
GLRCRGSDC	94	0	9.7
VGRSVDRVM	92	0	9.6
RLRCVSVVC	91	0	9.54
RCRGPDCLG	94	1	9.51
GALVHAFGV	90	0	9.49
ARCRGDQCS	89	0	9.44
VDRCRGTDC	89	0	9.44
GSWLKVSNG	87	0	9.33
FGRCRGPEC	87	0	9.33
GKVCMSGSC	87	0	9.33
CGLRRCLTF	87	0	9.33
YGRCRSVSC	87	0	9.33
GKYCASGNC	87	0	9.33
IGRSVNRSF	86	0	9.28
DKFCqSCVA	85	0	9.22
GKVCKGTMC	84	0	9.17
GKRCMTVWC	84	0	9.17
GDRCRGVAC	83	0	9.11
VGRLVDRqM	82	0	9.06
YALLARAHG	80	0	8.95
YERCRSADC	80	0	8.95
AARCMGFGC	80	0	8.95
SKYCqTCAV	80	0	8.95
FNRCRSSVC	80	0	8.95

13mer Experiment 2				9mer Experiment 2			
Sequence	**BC16**	**BC23**	**Z-score**	**Sequence**	**BC04**	**BC11**	**Z-score**
GKVCKSqACLEME	52646	0	302.77	NDRCRGVDC	748	144	20.12
GRCASGDPLGCGG	13040	0	120.74	VDRCRGTDC	427	89	14.77
PLNREMLAGVTFG	6752	0	84.51	ARCRGPVCL	479	118	14.64
GNTYALLLARAHG	5225	0	73.86	DRCRSVVCK	346	65	13.77
GPRGKLGVGAGLG	2835	0	53.87	GKVCMGIGC	198	3	13.76
GRCRSGILSGCGG	2757	0	53.1	RDRCRSSAC	300	49	13.36
GWqKVCKGEGACS	1833	0	43.13	ESRCRGVDC	497	161	12.93
CLVSGqYVCYDLF	1288	0	36.08	VGRCRGVDC	378	105	12.29
PLNREMLAGVTLG	1261	0	35.69	FGRCRGPEC	338	87	12.06
DLSGPRCASVECR	1230	0	35.25	EGRCKGIVC	691	306	11.95
VPTFAEYLERGRG	1156	0	34.16	GKVCRSAAC	216	32	11.62
MGRCRSELCRGEG	886	0	29.87	AGRCRGVDC	237	45	11.35
MSFRMTVADMMGK	845	0	29.17	RDRCRSDMC	227	42	11.2
GVTRGGKCGLECR	796	0	28.3	DKVCMSVQC	231	48	10.87
LPVPPGLLTRATI	737	0	27.23	YARCRGVDC	204	36	10.77
GRCVSWRPDGCGG	667	0	25.9	NEYLLRARG	1059	609	10.68
GLPVGGGWCGRGP	657	0	25.7	GKYCMSCVE	158	20	10.3
RLADPALKWERIG	634	0	25.24	RDRCRSALC	121	6	10.19
ERYQEYVQRMKLA	622	0	25	GKVCMSGSC	283	87	10.06
SGKACTMARGCGD	599	0	24.53	GRCRSDFCR	326	113	10.02
RLPFCGRLLCAAG	556	0	23.63	GKRCISWHC	978	577	9.84
EEARCRSPECQAY	503	0	22.47	AGRCRSFSC	240	71	9.47
SGVGGGWIAGRVG	478	0	21.91	KCGPECRSL	174	36	9.45
EPLWPRCKTRDLC	476	0	21.86	LDRCRSSVC	188	43	9.45
SRPGIRLYPHDMA	468	0	21.67	GDRCRGVAC	253	83	9.15
AASSPRCRSLACS	432	0	20.82	ADRCRGSDC	188	50	8.85
AHGPAWLQTPALP	431	0	20.8	DKVCRSWVC	99	8	8.77
GRCRSMqATGCAM	430	0	20.77	GKVCMGLGC	208	63	8.7
GLWGGAWGGKAVG	413	0	20.36	KMCRSIICM	121	18	8.69
RVGKDGVGWLVGV	401	0	20.06	AGRCRGFDC	388	177	8.68
GRGNGCGWGKWGC	394	0	19.88	AKKCMSGMC	154	34	8.68
RNPNPRCRSLECV	381	0	19.55	SGKCGPECR	120	19	8.51
NPGVNKVCLSGLC	376	0	19.42	DKVCRSWTC	492	255	8.44
MRCRSGLPDGCGL	359	0	18.97	FNRCRSSVC	231	80	8.43
VQSGKRCRTLWVC	341	0	18.49	DKRCRGTGC	111	16	8.38
LGVEGARVGKAVG	302	0	17.4	DKVCRSCQL	106	14	8.36
DWLRSKLVPQVSL	289	0	17.02	RKVCKSDWC	92	10	8.09
GLPSPRCRSDACV	259	0	16.11	SRCRGPDCG	102	15	8
ACFGVSSKGKCIR	247	0	15.73	GKRCMSSSC	83	7	7.99
KDRCMSRECKLEQ	241	0	15.54	WLRCRGGDC	119	23	7.99
LLPFPSAILDLGA	234	0	15.31	GVWLKSANG	182	58	7.9
GWMLQPTTGKRVG	230	0	15.18	DRCRSTVCL	140	35	7.86
WGLEVGGSWVRVG	229	0	15.15	AGRCVSFVC	208	74	7.85
QPAPLGKCDLRCR	226	0	15.05	GKVCRSEVC	110	20	7.84

RGEALGGWCGRGS	219	0	14.81	YDRCRSSVC	147	39	7.83
YTWSGGYVAKSGG	207	0	14.4	LSRCRGVGC	176	56	7.77
GWKPSDFAGGRLG	204	0	14.29	SSRCMSFVC	158	46	7.75
DVQWYRVAVAQGQ	198	0	14.08	SSRCRSFEC	63	1	7.75
DPAAKLCRSLAVC	197	0	14.05	VKVCRSGLC	156	47	7.55
RGGVGAGYSAGKA	188	0	13.72	KKVCMSVSC	94	15	7.52
EWQGRCRGVGCGL	186	0	13.65	SDRCRGFWC	89	13	7.48
EHMPRPKPFHLWD	183	0	13.54	ADLLMRAHG	144	41	7.48
RDLLNHDYLRHAG	182	0	13.5	GRCVSIECR	72	6	7.45
WAGQSPRCKTCVL	177	0	13.31	LHRCRGSDC	101	19	7.43
QWPRALGPTHAWN	159	0	12.62	qKFCMTCLS	389	204	7.39
GREQGWWDWRRVG	158	0	12.58	VGRSVDRVM	224	92	7.28
SRCRSVFCRWWEG	157	0	12.54	LMRCRGLDC	161	53	7.28
EGFWRGKDGKVLG	156	0	12.5	VDRCRGPGC	53	0	7.28
SSVAELCALCGYG	156	0	12.5	VSRCRGDQC	139	41	7.21
RPDGYWPLGRAAT	154	0	12.42	RVLPGFLID	84	13	7.17
GLRGEKVCMSLWC	151	0	12.3	GKRCMSLTC	69	7	7.08
GDGVSGGWCGRGS	149	0	12.21	DRCRSSSCL	76	10	7.08
SLPSFVFSSFILP	142	0	11.92	QGRCKSGLC	69	7	7.08
GVGGFGGWCGRGA	142	0	11.92	ADRCRSAGC	120	32	7.06
ELSFADFLARGRG	135	0	11.63	WVMLDRARG	69	8	6.92
GPPWPPSVITAAL	131	0	11.45	DRCRSVVCR	47	0	6.86
FPPPAVFPGLALR	131	0	11.45	DKACRGIGC	110	29	6.79
GGRKACRGQGCGT	131	0	11.45	PDWLYSYGM	78	13	6.77
RDPSHYQHVHYMI	124	0	11.14	VSRCRGVDC	154	55	6.74
DRGVGGGWCGRGA	117	0	10.82	NGRCRSDAC	243	113	6.73
NHYEEYLMRSKVK	115	0	10.73	RSRCRGDWC	105	27	6.72
GKVCKGEGACS	114	0	10.68	SSRCRGPVC	61	6	6.69
DRGVMGVVFGRSG	114	0	10.68	GCKVGKFGC	638	412	6.68
AGPRCRGPGACqD	113	0	10.64	SKFCMSCSV	53	3	6.67
GGWEMVLLDGRAG	109	0	10.44	GRCRSEVCR	53	3	6.67
GSGGVGGWCGRGV	101	0	10.05	SCVPGKYCP	115	33	6.66
TSFPLFIFTLSS	100	0	10	KKCMSTACA	47	1	6.64
GWVGGIYMGKAMH	99	0	9.95	AEYLLRSRG	107	29	6.61
SRGKVCMPGGDCD	98	0	9.9	ANRCRSVVC	156	58	6.59
EHRRLGSWCGRGE	95	0	9.75	FGRCRGVDC	110	31	6.57
GLGPSGGWCGRGS	94	0	9.7	AKRCRSLEC	168	66	6.55
GRCRSGLASGCGG	93	0	9.65	TLRCLGPDC	63	8	6.49
QPRCMSPDCRGAG	93	0	9.65	VCLRDGRCG	63	8	6.49
KFPQGLLNNAIGG	91	0	9.54	ASRCRSFAC	162	63	6.48
EGWRVGGWCGRGV	90	0	9.49	GKVCRSFGC	82	17	6.48
GRCRSDSCQLRAG	89	0	9.44	VDRCRTCVL	174	71	6.46
VGQDMLCGGKVCW	88	0	9.38	LFSDWSRVG	146	54	6.4
RIGSAGWDFARIG	87	0	9.33	GDLLARSHG	163	65	6.37
ERCRSTDCLRHAA	87	0	9.33	GRCRSHECL	61	8	6.35
PRCVSEICLAVER	84	0	9.17	SDRCMTCHV	40	0	6.33
LALPPLMIDLTSR	84	0	9.17	ADRCISFSC	484	300	6.32
RFGDKMCASQWCL	80	0	8.95	GSRCRGADC	175	74	6.27

GRCRSWLAGGCGG	76	0	8.72	NRCRSAVCQ	45	2	6.26
DVRCLSRECRLqV	75	0	8.66	RSWPRFVVE	42	1	6.25
QPSPDRCRSDNCR	74	0	8.6	ASRCRGIGC	68	12	6.22
VGWGGDRCRGVMC	73	0	8.55	DRCRSSICM	41	1	6.17
DRCRSWLGDGCGG	72	0	8.49	GRCRSSNCL	124	43	6.17
VSVTSRCASDICR	72	0	8.49	GKVCMTVHC	170	72	6.17
WTWEGGWKDRGGG	69	0	8.31	FDYLARSRG	75	16	6.13
VYAGSCIFDFRCG	68	0	8.25	SDRCRSATC	76	17	6.06

7.4.5 11889

13mer				9mer			
Sequence	**BC17**	**BC23**	**Z-score**	**Sequence**	**BC05**	**BC11**	**Z-score**
GGREVWVYSDRGE	549	0	23.51	RGWVLYESV	32	0	5.66
ESRGRDVWVVAGS	468	0	21.69	RSGNWIFSS	29	0	5.39
RRGGqVWVMEVGG	441	0	21.06	LRGGWVLAG	28	0	5.29
VAPCSGFAGLCVV	368	0	19.23	IVRRGWVVV	25	0	5
WIPWDGAMPVRHG	316	0	17.81	GSGRQWVFT	23	0	4.8
ERRDGAVWVVLGM	276	0	16.64	RGWVLSGSP	23	0	4.8
GRQDMRGWVVEVG	263	0	16.24	GHTYRWVLE	22	0	4.69
GQGGWVVLGDRRA	231	0	15.22	GSRWVWLSS	22	0	4.69
RDINGLVWVVGGD	227	0	15.09	VRRGWVFVE	22	0	4.69
PIAPSGRLVWLSD	223	0	14.95	AGWVLYDSA	20	0	4.47
GGFGCGGRGWVCV	221	0	14.89	VNVRGWVWT	19	0	4.36
DICPKGWVCVGGG	220	0	14.85	GGLSGWVIV	19	0	4.36
DRVVGGWVVRGGG	197	0	14.05	LGRHIWIGS	19	0	4.36
GREVWVVGYGGGG	191	0	13.84	GSRVIWEIT	18	0	4.24
VGPPGREWVYYDV	173	0	13.17	NAMRGWVLY	18	0	4.24
TRAEKERWLWVRE	162	0	12.74	GYHSRWVFF	18	0	4.24
RDGGAVWVLAGGD	138	0	11.76	QSVRGWVIV	17	0	4.12
KGRDVWLIDVSSN	133	0	11.54	RMGGNTWVG	17	0	4.12
ALPRSTWLPVPSD	133	0	11.54	FTSRGWVVE	16	0	4
REWMGqKWVMEGV	131	0	11.45	RLRGNVWVG	15	0	3.87
RSLNGDVWVMAEG	129	0	11.37	GYGRGWVIV	15	0	3.87
GCVGRGWVCVGDG	127	0	11.28	LARSGWVLF	15	0	3.87
GDRIqGGWVVYGG	123	0	11.1	GRVRWVMVD	15	0	3.87
ERRGRDVWVHEQG	120	0	10.96	VWGERWVVA	14	0	3.74
RTGERGWRQWVIE	113	0	10.64	GRGWVLTSA	14	0	3.74
RTGGQVWAVPHTN	105	0	10.25	GSMRGWVLF	14	0	3.74
RSGGQVWVMGGGG	103	0	10.16	MCGRGWVCG	13	0	3.61
GCLGRGWVCLSGE	100	0	10.01	VDARMWVRW	13	0	3.61
QGSGWVWLVGGDG	97	0	9.85	GRGWVMIGE	13	0	3.61
NALKSRWVAYEDA	91	0	9.54	GVLRGWVVT	13	0	3.61
RPAVERWVVLDTG	89	0	9.44	FVWNGYVIE	12	0	3.46
VVMRLQGqEWVAF	89	0	9.44	GVMRGWVFT	12	0	3.46
RGSDGYMWVLGGG	88	0	9.39	AGWVVSHGG	12	0	3.46

RDGDGVWWVGGND	86	0	9.28	VLGRGWILD	12	0	3.46
VRFGNDVWISSGF	86	0	9.28	MRHWVLDRS	12	0	3.46
PQSRFHWLSNPDR	85	0	9.22	RWGGSVYIG	12	0	3.46
GVGSCGRGWVCLN	82	0	9.06	SRRFVWIVD	12	0	3.46
IVSVPRMGWIVTq	82	0	9.06	SRSSWVYWE	12	0	3.46
WLTDPGLLSNAMT	80	0	8.95	GGRVRWIVD	12	0	3.46
PRERWVWLTGGDG	80	0	8.95	MRGWVLEAA	12	0	3.46
DRGREVWLVGGGV	80	0	8.95	VSSSGWIFF	12	0	3.46
AVREGGQVWLRSV	79	0	8.89	MVINGWVVS	12	0	3.46
EGRAKERWSWVYY	77	0	8.78	RPYSGWVIT	11	0	3.32
TDERFTWVMASKE	77	0	8.78	ECERGWICT	11	0	3.32
RDGEGLGWVWIIA	77	0	8.78	RTWVMYGSD	11	0	3.32
TTVQTVSGGWVVT	75	0	8.66	THSGWVLYR	11	0	3.32
RVGNSVWVVGGGD	74	0	8.61	TGRWLWIRE	10	0	3.16
WCDRSqGECWVVL	72	0	8.49	EVYRGWVLY	10	0	3.16
EERISERWKWVQA	72	0	8.49	MSGRSWVIF	10	0	3.16
SEHRWLWVQEELN	71	0	8.43	GAWRSWVVS	10	0	3.16
NDYRLLGPWVLNE	71	0	8.43	TRSATWVFY	10	0	3.16
RWLGSWWVVGEGV	70	0	8.37	DMVRGWVVV	10	0	3.16
MGAVEGLRGWVVV	70	0	8.37	VGVRGWVLT	10	0	3.16
VHRSGGFVWVTQD	70	0	8.37	GLLRGWVVE	10	0	3.16
GQGGQGWRGWEMT	68	0	8.25	GYRTVWVVT	10	0	3.16
VRRGDGVWVVSNV	68	0	8.25	RYGRDVWIG	10	0	3.16
RHSGSVWVVGGGV	68	0	8.25	FRGWVLGNG	10	0	3.16
KDAGRVVWVLKDG	68	0	8.25	GSQRGWIVT	10	0	3.16
DGSGQRWFWIRHA	67	0	8.19	ATRFIWIID	10	0	3.16
DVVRGGWVVHRDV	67	0	8.19	GGWVVSHEG	10	0	3.16
ERPRVTWVLNVQP	66	0	8.13	MRGWVIQVS	9	0	3
VETGGQWRGWVLE	65	0	8.07	GGRYVWLRH	9	0	3
GLAAGqMRGWVLF	65	0	8.07	GGRAIWVYV	9	0	3
ARYGSEVWVLGGG	64	0	8	GRLVWLYSS	9	0	3
ARGTGDFRGWVRW	64	0	8	RTGWVVEHG	9	0	3
EGVGMWAREWVFV	64	0	8	ANRYTWLVE	9	0	3
GMGREVWILHAEK	64	0	8	GRVRWVLFD	9	0	3
ERLGRQVWVMTAE	63	0	7.94	IGGRGWVLL	9	0	3
REGWWWLPRGDEG	63	0	7.94	GMQRGWVAV	9	0	3
RWGGSWWVTEGLG	63	0	7.94	VMSRHWVMV	9	0	3
GVMQVRGVAWETS	62	0	7.88	RGWVIASSA	9	0	3
RQDAGRVTWIAHE	62	0	7.88	LRGWVVIGE	9	0	3
PCRRGWICVGGET	62	0	7.88	GAQARWVYY	9	0	3
ARLDGSWWVVGGA	61	0	7.81	RRNQWVFAE	9	0	3
CAGGMCWFVGGAV	60	0	7.75	LSGHWVLSS	9	0	3
ETVRIGGSKWVVE	60	0	7.75	AGRVIWVMD	8	0	2.83
ERYTWVLADGGGV	60	0	7.75	GAFRGWVVD	8	0	2.83
RLLGSTWVFDGSN	60	0	7.75	GVMRNWVIL	8	0	2.83
PWRSWVFVDVGGG	60	0	7.75	GIRHAWVFD	8	0	2.83
GGRNVWVLDVGDG	58	0	7.62	RGWVLAGPA	8	0	2.83

VVRDGVWLNVRYG	58	0	7.62	SGWVFHSGW	8	0	2.83
GSGVVLWRGWVVE	57	0	7.55	SGPSYYALP	8	0	2.83
VRSGREVWVVREG	57	0	7.55	RGWIPWEQD	8	0	2.83
GTRYAGQAWVLVG	57	0	7.55	GIMQGWVVT	8	0	2.83
EHRGGGWVVDVGG	57	0	7.55	GGWRSWVIE	8	0	2.83
EAGRSVWELYRMD	57	0	7.55	GGLRGWVYY	8	0	2.83
IMSGSTGERWVVR	56	0	7.49	SGWVVMHDS	8	0	2.83
RDAPAYWVVDHSS	56	0	7.49	VRGGWIFVS	8	0	2.83
SWGGSTWVTMP	55	0	7.42	NSLRGWVLV	8	0	2.83
GVERGERWVWVWE	55	0	7.42	GRLRGWVIE	8	0	2.83
DGGRVSWVSWGGE	55	0	7.42	RGWVMGPVD	8	0	2.83
QVGRWVVFGGEGG	55	0	7.42	GLGREWVVT	8	0	2.83
NRVGEGWVVEHGG	55	0	7.42	RRGWILYER	8	0	2.83
RGGDGVVWVLGGG	54	0	7.35	VMAGGWVVM	8	0	2.83
ETVGSGLRGWVTV	54	0	7.35	DGRGWVMYG	8	0	2.83
EARGRDGRVWVVL	54	0	7.35	DRRLVWIVE	8	0	2.83
RRGGQVWVSVGEE	54	0	7.35	SqRWAWVYV	8	0	2.83
GGWVVDHGGSSEV	54	0	7.35	GRFRWISVE	8	0	2.83
VWRERGAVWEASG	53	0	7.28	NSLRGWVWT	8	0	2.83
GLMSFGGqVWVVQ	52	0	7.21	GLERGWVLH	8	0	2.83

7.4.6 11891

13mer				9mer			
Sequence	**BC18**	**BC23**	**Z-score**	**Sequence**	**BC06**	**BC11**	**Z-score**
LPLRSLLAFASFP	3422	0	59.61	MSSFHPVLP	1928	0	44.29
QPLSTSLWSLPFP	2959	0	55.28	VDLKALYLP	505	0	22.52
MTSTANQLSLVSL	1787	0	42.69	NGALRTVVL	498	0	22.37
GSRTVGEAVGGGM	1158	0	34.24	GARLLLAGK	250	0	15.83
RPTPSILASASMP	669	0	25.96	GLVLAAGIP	116	0	10.78
RSIqMAISDAALG	534	0	23.18	IALLISqGK	115	0	10.73
RLGSLHqAVNALP	521	0	22.89	AVEMIMAGK	87	0	9.33
RVMGVGEAVRGGW	518	0	22.82	DKLSLILGR	72	0	8.49
RQSHSQMFTAVVD	511	0	22.67	VGIWGLGLP	70	0	8.37
APLYSHIQLLSLP	491	0	22.22	VGYMLASGK	61	0	7.81
GGYSVGTAVLRGL	467	0	21.66	GGWCGHGPC	49	0	7
ARPVGqVVGAVSV	445	0	21.15	VRMLIAqGK	49	0	7
KPSMTIYEAVKMS	420	0	20.54	KSVFRAMSL	48	0	6.93
EAGGMDLRSVWFP	402	0	20.09	RSVLSGVSL	46	0	6.78
GWGTNFAVRSVAL	381	0	19.56	KTVFRqVSW	46	0	6.78
GVYSVGqAVAAGF	380	0	19.53	RSVFAAVSV	40	0	6.33
ESSSVGEAIGRGF	372	0	19.33	RSVFRAVNL	37	0	6.08
ERGRSFWEVTGMG	356	0	18.9	RYTMREMSI	36	0	6
RSLTqMLGEVGVP	354	0	18.85	GqIVRLASL	36	0	6
RGRTMLVDGALLS	352	0	18.8	VSIWRVFLP	36	0	6
RMVGLGEAVGKAG	336	0	18.36	LGSRGVVLP	36	0	6

DFASFPLRMFSNP	331	0	18.23	KSVLGSVTV	35	0	5.92
RPAGIMTAWAQLP	330	0	18.2	RSAVMAVTL	34	0	5.83
RDLGSEMRKVLNP	320	0	17.92	LGGRGVTFP	34	0	5.83
GGGSVAWGVGLGM	315	0	17.78	MMRLVRDGK	34	0	5.83
RSVGREMEQYTLG	271	0	16.49	VLSLIqNGK	33	0	5.75
GVLCqVFCCGCDS	257	0	16.05	VADFFGRGK	33	0	5.75
QSLPALVSTLSYP	237	0	15.41	SVqLMLFGK	33	0	5.75
NGLPAAMRAVSNP	231	0	15.22	RAVMAGVVM	32	0	5.66
QSTGSAMSYALGA	227	0	15.08	RRNFAWFLD	32	0	5.66
FPLSSVLASVVSP	223	0	14.95	RRVLAMYSL	32	0	5.66
HHVATVLPMLTYP	221	0	14.88	GSVLRSVGL	31	0	5.57
ALRSVGqVVQEGF	215	0	14.68	VRMSQVFFP	31	0	5.57
GMALGEqMQVVML	201	0	14.19	VVDLIRAGK	31	0	5.57
RVAGMGRAMGEVM	200	0	14.16	VNMWMLSMP	30	0	5.48
RLIDMKMMMTQQM	198	0	14.09	KGTMGLVTL	30	0	5.48
SGEGVGKAMGAVL	198	0	14.09	GVTRMVALP	30	0	5.48
WPVRTTLLSTMLP	195	0	13.98	MGVRGPVFP	30	0	5.48
RGVGGAVHEVAVG	189	0	13.76	MAYLMTKGK	30	0	5.48
GAVSLRGEVSRVM	179	0	13.39	DSLLKVLGA	30	0	5.48
VTSIGDVSANLWR	178	0	13.35	VRAWYRSGM	30	0	5.48
KGSVAEVVERVAV	178	0	13.35	VGWSHNLAM	29	0	5.39
LPSASVLRTLTLP	171	0	13.09	VADLVRMGK	29	0	5.39
QQVRWAVLSAEAM	170	0	13.05	GRVMGVVLP	29	0	5.39
RGTAVGVAVGGGF	170	0	13.05	RSTMSDVTL	29	0	5.39
RGRSVAEALTMVT	167	0	12.93	IVqMIKGGK	28	0	5.29
MSLGRGVREMTVP	163	0	12.78	RSMLSQVSL	28	0	5.29
SQSMREAVQQVTA	162	0	12.74	GQVVAMVSL	28	0	5.29
VRPVGAFLEqVSL	160	0	12.66	RNFRDLYLP	27	0	5.2
RYFSARSFIGDVM	157	0	12.54	MGRSSGPIP	27	0	5.2
VPLQTAIANATGR	156	0	12.5	ASVLSQVSL	26	0	5.1
LPLHSHLLTLSMP	154	0	12.42	REYFRNVSW	26	0	5.1
VDLKGVMRGGASV	153	0	12.38	GNLKAMAMT	26	0	5.1
LSLPSVVRSIAAP	152	0	12.34	VNMMLRqGK	26	0	5.1
KATREMLWEVGL	146	0	12.09	YVARSPWLP	25	0	5
RRGVGDMVGAVSV	145	0	12.05	GLVRAASFG	25	0	5
QPTHSLLRFISYP	143	0	11.97	AHDMLRAGK	24	0	4.9
GAGVRGVMDQVAL	141	0	11.88	VALMLqMGK	24	0	4.9
GMGVGEAVAKVAG	139	0	11.8	NNLKSIGLP	24	0	4.9
WAVPMPWPVGLG	136	0	11.67	SGVLRSVSL	24	0	4.9
HGMGLGEGIRKQF	133	0	11.54	YGSLSPAIP	23	0	4.8
QPLGVAVRTAFGL	132	0	11.5	ATVLRqMSL	23	0	4.8
DVSVSqAVSSVTV	128	0	11.32	YGVDLKTLG	23	0	4.8
GMSLGqAILSVMV	128	0	11.32	MSFWDAYRW	23	0	4.8
QDMRVWFWIDRGV	126	0	11.23	RGVMTSVVL	22	0	4.69
RSLSDAVSGVAVA	124	0	11.14	RqVFREVSL	22	0	4.69
QGLGSVVEAMAAG	121	0	11.01	VSLDVNRLG	22	0	4.69

APNSVWAAVGqGL	121	0	11.01	GGALRMVSW	22	0	4.69
LKPAYMTLqEVTL	118	0	10.87	VREVLSGRL	22	0	4.69
RNASSVIAGISFP	116	0	10.78	VNLGALWFP	22	0	4.69
STSLFSKQVFSFP	115	0	10.73	ATLRLVAMP	22	0	4.69
RRMAMGEAMMDVM	111	0	10.54	KqTMASVSL	22	0	4.69
ATGLDAGIRNAMW	107	0	10.35	REVWRNVAS	21	0	4.58
RAQSAALAEAMGR	106	0	10.3	DLRAFSWGR	21	0	4.58
GQIGLGQAMARQF	106	0	10.3	ASEVSSGRL	21	0	4.58
GGSLTTAMQAVGS	106	0	10.3	RHSLVDGAL	21	0	4.58
RqAVGNVVTGVSL	106	0	10.3	DVLGRILGA	20	0	4.47
qNLVQVVSMMSFG	104	0	10.2	RqVLREISW	20	0	4.47
MVGDGGGGWFVLV	106	1	10.17	AADVSAGRV	20	0	4.47
PQPYFNVWSLASP	102	0	10.11	RTVLSSVSV	20	0	4.47
GTDLGGAMRAVLS	100	0	10.01	GAWLVSRGK	20	0	4.47
RTMAVTQAVQVVM	99	0	9.96	VVGLLRqGK	19	0	4.36
SPISFHVSALAFG	99	0	9.96	MGHAIARGM	19	0	4.36
RSVGQEVGFVVFG	94	0	9.7	HAMMSKWTS	19	0	4.36
HPVGSQLLIMSVP	94	0	9.7	VAGLLMSGK	19	0	4.36
QSLSDMVNRATSP	92	0	9.6	ILELVRSGK	19	0	4.36
PRGLSEVVMGASI	92	0	9.6	RSVLHSIMS	18	0	4.24
RDLQTVVSDFGLP	91	0	9.54	RSVLRMVSF	18	0	4.24
QRTGVWNAVQRAG	90	0	9.49	MGKSVGPLP	18	0	4.24
MPLRSVVLSVVSP	90	0	9.49	AGVLLSAGK	18	0	4.24
QqSMAEVMSMVMR	90	0	9.49	ASRLVSRGI	18	0	4.24
NAMDWTLESMKKS	89	0	9.44	LKVFWGQGQ	18	0	4.24
GGGVGQQVGMVSW	89	0	9.44	GLLHMVSAP	18	0	4.24
GRGGVGAAVASGL	89	0	9.44	MAELILRGK	18	0	4.24
GRAMGAAVSSVAI	87	0	9.33	SAWSALSMP	18	0	4.24
GALSAVMGRSTLG	86	0	9.28	ASVMRLASL	18	0	4.24
GLGGLATAVSGSM	86	0	9.28	INPRVWSMP	18	0	4.24
RSLSATILLLSFP	85	0	9.22	GIWGAMRMS	18	0	4.24
DSVGEAMSRALGR	85	0	9.22	FRSGWVLGA	18	0	4.24
RSLVGEVGSLAVP	85	0	9.22	KSELSRVTW	18	0	4.24

7.4.7 11892

13mer				**9mer**			
Sequence	**BC19**	**BC23**	**Z-score**	**Sequence**	**BC07**	**BC11**	**Z-score**
MVGDGGGGWFVLV	48971	1	293.36	IERLVSWqT	186	0	13.65
THYMRGLDYAWGL	7088	0	86.95	GWLVAGVTI	138	0	11.76
GGSWVVRGIASGQ	3860	0	63.21	VVQCATCAK	138	0	11.76
GPYRFSELVERRV	3006	0	55.57	AMDCATCSV	96	0	9.8
GVDLFWRGMHSGW	2387	0	49.38	GFLVDGVMG	83	0	9.11
RVWDYVGGGWWGG	2263	0	48.05	HACGTVACV	81	0	9
GHWCTVWAGGGEC	2252	0	47.93	GYLIAGVAS	79	0	8.89
QSLPSLRFSPHGP	1666	0	41.12	VPFRIVGGM	75	0	8.66
RRAEPGGPAGLWG	1385	0	37.44	YRRFVASGD	72	0	8.49

VRSRDGLQVVGAG	1366	0	37.18	YWSLLSAGD	71	0	8.43
KGVEWERLMRGYG	988	0	31.57	FRRLSLSGY	68	0	8.25
NHSPPHKPDAPYN	961	0	31.13	GFLITGVST	60	0	7.75
TEIERWMETMRqT	906	0	30.22	WRRLIDSHS	60	0	7.75
RDGREWMARWGGG	812	0	28.6	GYLVDGVTR	57	0	7.55
RPDWGVRLVKSGE	715	0	26.82	YTRYVMSGW	53	0	7.28
SSEMAWRMMASKE	676	0	26.08	FERLVAWqL	53	0	7.28
GVAELWRRMGGAA	628	0	25.13	RYLVDGSMS	53	0	7.28
RPASERLIDVFGY	563	0	23.79	YNRLRASDY	49	0	7
GILELWGRFGGSG	522	0	22.9	VIAGMGTVT	49	0	7
DGAQDGWRAFRGA	502	0	22.46	FIRLqASGI	48	0	6.93
RQTPPRCPTGDCG	477	0	21.89	GWLVVGVMS	47	0	6.86
GWGDLVEGVMRGH	467	0	21.65	GWLVSGVTV	45	0	6.71
KYGQEWLSRMGGG	462	0	21.54	YSRFRDSGF	43	0	6.56
SGRGAFMDVMTGV	445	0	21.14	GWLAQGVMS	43	0	6.56
AKGGMGWAVFQGA	438	0	20.97	RWLIqGVMS	42	0	6.48
EFQRWRDSIVMGV	374	0	19.37	YERFVRSGE	41	0	6.4
MGGVEWLRRWGGG	369	0	19.24	KCGTAPCqL	43	1	6.39
QPSVSLRFAPPFP	367	0	19.19	FGGCATCGW	40	0	6.33
GNRDLGHVLKDLN	352	0	18.79	GMRFVPSGP	40	0	6.33
DQSRFFHPGWHSG	347	0	18.66	MRRMVDSGV	40	0	6.33
WGERRLSWEALVT	347	0	18.66	WAMDRIYGY	39	0	6.25
TRPQYQRFVDSGW	341	0	18.49	EWLVSGVSL	38	0	6.17
VGEGSGWAVFGGA	341	0	18.49	LDYERLFGY	38	0	6.17
YPGPPHKPDRPQV	313	0	17.72	ACATCGVGG	38	0	6.17
WLRPGADWVAPGM	288	0	16.99	MEWERLFGY	37	0	6.08
GGAEDWLRRWGGG	265	0	16.3	WDVRRWGVM	37	0	6.08
PVPPSFRMSPYGP	256	0	16.02	SASLYHMSL	37	0	6.08
DLCRWGVGCGVVG	251	0	15.86	GYLVLGAMA	36	0	6
RNSPPHKPDAPWH	249	0	15.8	ATCGTNCVE	36	0	6
GWDERWRRLRGVE	245	0	15.67	GWLVSGVSV	34	0	5.83
PDAWALRYIRSGA	244	0	15.64	GYLVSGVAS	34	0	5.83
CLRSGVCVFLTqD	236	0	15.38	GIAASRWTV	34	0	5.83
ALGRDGLVVLGGG	226	0	15.05	DSqVFHLAW	34	0	5.83
VRVDDKWFDRWGG	211	0	14.54	SMLVAGIMS	33	0	5.75
GSQGNGWRVFGGA	210	0	14.5	MSRYAVATK	33	0	5.75
qERWLNVEQLLGG	206	0	14.37	GFLVTGVMA	33	0	5.75
VRSKGGWVVVGGG	191	0	13.83	SWLVQGSMA	32	0	5.66
LAPTWRVPSPHGP	187	0	13.69	CATCDWGIT	32	0	5.66
GNGYEWLDRWGDK	186	0	13.65	GHRCGTVLC	31	0	5.57
RVWDYVGGGWFVLV	183	0	13.54	ARLDRWLGY	31	0	5.57
GQGGAGWGYFGGA	177	0	13.31	HARLVDSGW	31	0	5.57
FPGPPHKPAYPGF	175	0	13.24	CATCAAGTT	30	0	5.48
GCLGRLSCVAGGF	175	0	13.24	GFRVPRMGP	30	0	5.48
GWLTEVGVGRRTG	170	0	13.05	ACGTNICGT	29	0	5.39
PSLWSMFPGLQGE	166	0	12.89	MYLTRGVSG	29	0	5.39

GGMEMKELLQRGV	162	0	12.74	AWLVSGVTL	29	0	5.39
PPQTTIRFARHGD	161	0	12.7	GFLVRGVTS	29	0	5.39
RGGWDWLARWGGD	160	0	12.66	FARFLqSGW	29	0	5.39
YRPTNYRISPHGP	160	0	12.66	CATCGLSAT	28	0	5.29
GRSVGGWRVFEGA	154	0	12.42	CATCSFVTN	27	0	5.2
APDSWVRLTESGW	152	0	12.34	RCATCSIGD	27	0	5.2
GSREVRYRVEGVM	147	0	12.13	GGVDRWLGY	27	0	5.2
GRPWWVWAGRVGE	147	0	12.13	GFLMEGVMT	27	0	5.2
VGWGEWGLRMWGE	147	0	12.13	FYLVqGVTS	26	0	5.1
PPRPTVRFPPHGP	144	0	12.01	FDRFRSSGQ	26	0	5.1
RANYEELVMMRKW	142	0	11.92	GYLVAGVFS	26	0	5.1
PPMHKPDAPPGYE	138	0	11.75	WAKFVGVGE	26	0	5.1
VDRPCLLKGDCVG	128	0	11.32	GFLVAGVSV	25	0	5
WWATRYAERVWGQ	128	0	11.32	LCGTNVCER	25	0	5
SVLVqERGFFAYT	127	0	11.28	LYRQTEATW	25	0	5
GSGLAWLARWGDG	125	0	11.19	LCGTNGCST	25	0	5
GSREWWGRWGPGV	118	0	10.87	FqRLLASDY	25	0	5
RLRHPHDPSLAIP	116	0	10.78	YWSMLGRGD	25	0	5
MPRWTMLEAYTGS	112	0	10.59	GLLMSGILM	25	0	5
WDAPWWIQGGKGE	112	0	10.59	YARSIASGV	24	0	4.9
DHSLPFRLSPLSP	109	0	10.45	CATCAQTAE	24	0	4.9
GMVELWLRSGGMF	107	0	10.35	CATCASGLD	24	0	4.9
PLPPLFLRCYPSC	103	0	10.15	CATCQSqVV	23	0	4.8
GPAQTIRYARHGD	102	0	10.1	EYLVSGVMG	23	0	4.8
GRGVGGWAIFGGA	101	0	10.05	GLVATRFTA	23	0	4.8
RLPRCGHTFCLPG	101	0	10.05	CEFGCLEWG	23	0	4.8
FCYGVVCWDVFGE	99	0	9.95	RDVRAWLGY	23	0	4.8
GSGYALDFGRWGG	99	0	9.95	YRRFSASGD	23	0	4.8
YSSCVLFPFVGSC	99	0	9.95	RLGHMFGTY	23	0	4.8
DCVSWWPRCVSLV	98	0	9.9	FRRMVDSRL	22	0	4.69
EEGIVERLMRGHG	98	0	9.9	TYLTAGVMA	22	0	4.69
LPFRVVGSYDGTL	97	0	9.85	GGGSRWATW	22	0	4.69
GGGWWRVEGWRGG	96	0	9.8	ACGTGGCLG	22	0	4.69
EMRFNRGLQYYSV	95	0	9.75	GYLVTGVSL	22	0	4.69
DYVFLPYARFGYG	94	0	9.7	IATRVASSG	22	0	4.69
MVGDGGGGWWGG	91	0	9.54	FYLIqGVGV	22	0	4.69
GERLPLRIAPNGP	89	0	9.44	YNRLVAWqM	22	0	4.69
RPLSPDGPFVLTG	88	0	9.38	LCGTGGCMA	22	0	4.69
PGPTYHLAVRQGE	87	0	9.33	YqRLILSGF	22	0	4.69
VGTRGGVVFVGEG	84	0	9.17	GSLERYLGY	22	0	4.69
LVAPYMFPLALVF	83	0	9.11	MGRLLASGW	21	0	4.58
GRSDRWCSPVGCV	80	0	8.95	GYLVEGVSR	21	0	4.58
RLPPPFP	79	0	8.89	FTRSLASGM	21	0	4.58
GPFSLVRPSSYML	77	0	8.78	YLRGLVSGW	21	0	4.58
RLAASRCPSLQCG	75	0	8.66	GFRVSPFGP	21	0	4.58

7.4.8 11881

13mer				9mer			
Sequence	**BC17**	**BC23**	**Z-score**	**Sequence**	**BC08**	**BC11**	**Z-score**
MVGDGGGGWFVLV	91286	1	456.22	IERLVSWqT	225	0	15.02
THYMRGLDYAWGL	34322	0	208.58	GSLERYLGY	217	0	14.75
GHWCTVWAGGGEC	14348	0	125.45	GYLIAGVAS	214	0	14.64
GGSWVVRGIASGQ	3066	0	55.9	GYLVDGVTR	196	0	14.01
GVDLFWRGMHSGW	3007	0	55.35	GFLVDGVMG	187	0	13.69
KGVEWERLMRGYG	1814	0	42.83	YWSLLSAGD	185	0	13.61
RVWDYVGGGWWGG	984	0	31.46	GWLVAGVTI	174	0	13.2
GPYRFSELVERRV	688	0	26.29	AMDCATCSV	149	0	12.21
QSLPSLRFSPHGP	454	0	21.34	YRRFVASGD	145	0	12.05
RRAEPGGPAGLWG	449	0	21.22	GFLITGVST	134	0	11.58
GILELWGRFGGSG	339	0	18.43	YNRLRASDY	133	0	11.54
ALGRDGLVVLGGG	300	0	17.34	GWLVSGVTV	130	0	11.41
EFQRWRDSIVMGV	222	0	14.91	GWLVVGVMS	129	0	11.36
WLRPGADWVAPGM	190	0	13.79	VVQCATCAK	129	0	11.36
RLPRCGHTFCLPG	166	0	12.89	YTRYVMSGW	112	0	10.59
GWGDLVEGVMRGH	165	0	12.85	YSRFRDSGF	109	0	10.45
VDRPCLLKGDCVG	165	0	12.85	FRRLSLSGY	107	0	10.35
KYGQEWLSRMGGG	151	0	12.29	RYLVDGSMS	104	0	10.2
MVGDGGEC	150	0	12.25	FIRLqASGI	102	0	10.1
RPDWGVRLVKSGE	149	0	12.21	RHVSAYLGY	99	0	9.95
LCWSWYEGVGCRG	141	0	11.88	WRRLIDSHS	97	0	9.85
QPSVSLRFAPPFP	138	0	11.75	FERLVAWqL	95	0	9.75
AKGGMGWAVFQGA	137	0	11.71	FqRLLASDY	94	0	9.7
DGAQDGWRAFRGA	131	0	11.45	WAMDRIYGY	93	0	9.65
RVWDYVGGGWFVLV	119	0	10.91	GWLAQGVMS	90	0	9.49
RQTPPRCPTGDCG	109	0	10.44	DSqVFHLAW	90	0	9.49
WWATRYAERVWGQ	109	0	10.44	VPFRIVGGM	85	0	9.22
DYVFLPYARFGYG	106	0	10.3	HACGTVACV	82	0	9.06
FCYGVVCWDVFGE	106	0	10.3	GYLVLGAMA	81	0	9
TEIERWMETMRqT	100	0	10	NFLIEGIMR	78	0	8.83
CLRSGVCVFLTqD	97	0	9.85	GYLVSGVAS	77	0	8.78
EMRFNRGLQYYSV	94	0	9.7	SASLYHMSL	76	0	8.72
GRSVGGWRVFEGA	93	0	9.65	YERFVRSGE	74	0	8.61
MPRWTMLEAYTGS	92	0	9.59	GRIDRWLGF	72	0	8.49
VRSKGGWVVVGGG	90	0	9.49	YLRGLVSGW	72	0	8.49
GQGGAGWGYFGGA	89	0	9.44	RWLIqGVMS	69	0	8.31
RLPPHALDPLVLA	86	0	9.28	VIAGMGTVT	68	0	8.25
CGAVWGDFGWRCG	86	0	9.28	GWLVSGVSV	68	0	8.25
VCLWPVGVCVGDG	78	0	8.83	GGVDRWLGY	67	0	8.19
GGAEDWLRRWGGG	77	0	8.78	YWSMLGRGD	67	0	8.19
DCAHPDITVGCTK	77	0	8.78	MYLTRGVSG	67	0	8.19
MGGVEWLRRWGGG	77	0	8.78	GYLVSGIMA	66	0	8.13

MVGDGGGGWWGG	76	0	8.72	LCGTNGCST	65	0	8.06
GCAGRPPFVCAVV	71	0	8.43	FERYVASGF	65	0	8.06
LVAPYMFPLALVF	70	0	8.37	GYLVEGSMR	63	0	7.94
VGEGSGWAVFGGA	70	0	8.37	LDYERLFGY	63	0	7.94
GGGWWRVEGWRGG	69	0	8.31	CATCGLSAT	62	0	7.88
TNHEKPMPPMWKW	68	0	8.25	YARSIASGV	61	0	7.81
MVGDGGGGWFVLL	62	0	7.88	SYLIEGLMT	60	0	7.75
GCLGRLSCVAGGF	56	0	7.48	FDRFRSSGQ	59	0	7.68
TRPQYQRFVDSGW	54	0	7.35	FTRYVSSGW	59	0	7.68
RqVTSGWAWSGTV	54	0	7.35	YqRLILSGF	59	0	7.68
PDAWALRYIRSGA	54	0	7.35	SMLVAGIMS	58	0	7.62
GVRPCHSAGDCWG	53	0	7.28	GFLVRGVTS	57	0	7.55
RGGWDWLARWGGD	52	0	7.21	MEWERLFGY	57	0	7.55
FDCVALWPVCYPV	50	0	7.07	RDVRAWLGY	56	0	7.49
RANYEELVMMRKW	50	0	7.07	FARFLqSGW	56	0	7.49
DQSRFFHPGWHSG	49	0	7	SWLVQGSMA	56	0	7.49
RLPVGVGWVPGKG	48	0	6.93	GYLVYGVMA	55	0	7.42
GTLELFGRFGGAE	47	0	6.86	FGGCATCGW	55	0	7.42
HWLATLGVGRCCG	47	0	6.86	GYLVAGVFS	53	0	7.28
VDCVSNWPVCAIG	46	0	6.78	GYLMAGVSS	53	0	7.28
GSQGNGWRVFGGA	43	0	6.56	ARLDRWLGY	52	0	7.21
VPGYQAVEKGLLK	42	0	6.48	HARLVDSGW	51	0	7.14
RDGREWMARWGGG	42	0	6.48	GYLMSGIMG	51	0	7.14
TSVICRGLDYAWGL	41	0	6.4	CATCDWGIT	49	0	7
FLPSNLFGSFYSV	40	0	6.33	DAKVHLVAI	48	0	6.93
GNGYEWLDRWGDK	38	0	6.17	ATLVDRFGG	48	0	6.93
RLAASRCPSLQCG	38	0	6.17	FRRMVDSRL	48	0	6.93
PVPPSFRMSPYGP	38	0	6.17	EWLVSGVSL	47	0	6.86
DCVSWWPRCVSLV	37	0	6.08	GSRYFDYSR	46	0	6.78
YRPTNYRISPHGP	37	0	6.08	YYLVqGAAM	46	0	6.78
DLRFLEWARFMGF	36	0	6	CATCAQTAE	45	0	6.71
NHSPPHKPDAPYN	35	0	5.92	LCGTGGCMA	45	0	6.71
GRGVGGWAIFGGA	34	0	5.83	RYLVAGVTL	45	0	6.71
PPRPTVRFPPHGP	34	0	5.83	GFLVAGVSV	45	0	6.71
MVGEWGGGWFVLV	33	0	5.75	SYLVAGVGT	45	0	6.71
qERWLNVEQLLGG	33	0	5.75	IGMDRWLGF	44	0	6.63
DPRDTQLRNPSMG	33	0	5.75	EYLVRGLVN	44	0	6.63
SGRGAFMDVMTGV	33	0	5.75	SASTFHVSV	44	0	6.63
ATSCAVFSGCWLL	32	0	5.66	RPISAFLGY	44	0	6.63
THYMRGLDYASGQ	32	0	5.66	GYLVqGIMS	44	0	6.63
GRPWWVWAGRVGE	32	0	5.66	GLLMSGILM	43	0	6.56
VRSRDGLQVVGAG	32	0	5.66	GIAASRWTV	43	0	6.56
TNIPSGRSFPPGF	32	0	5.66	ERYARWYMS	43	0	6.56
RVPINLWSTYAAV	31	0	5.57	GYLVEGVSR	43	0	6.56
LDWHRFFSPALLS	30	0	5.48	MRRMVDSGV	42	0	6.48
RPASERLIDVFGY	29	0	5.39	YRRSVDSGL	42	0	6.48
THYMRGLDYAWGV	28	0	5.29	GLVATRFTA	42	0	6.48
PVPYPLLDVLRSL	28	0	5.29	YSRMVDSGF	42	0	6.48

EGRMWDVRVWGVL	28	0	5.29	FSIERFLGN	41	0	6.4
LAPTWRVPSPHGP	28	0	5.29	CATCASHLG	41	0	6.4
MVGDGGGGEC	27	0	5.2	FRRYVDSGF	41	0	6.4
RVWDYVGGGWWGGGEC	27	0	5.2	GSIERYIGY	40	0	6.33
PSLWSMFPGLQGE	27	0	5.2	HGAERWLGY	40	0	6.33
GVAELWRRMGGAA	26	0	5.1	YRRFIDSGE	40	0	6.33
GSREWWGRWGPGV	26	0	5.1	GFLIDGVSR	40	0	6.33
SYADLLERFSFSS	25	0	5	FSRGLSSGM	40	0	6.33
MVGDGGGGWFVLA	24	0	4.9	RCATCSIGD	40	0	6.33
GERLPLRIAPNGP	23	0	4.8	KCGTAPCqL	42	1	6.26

7.4.9 11894

13mer				9mer			
Sequence	**BC21**	**BC23**	**Z-score**	**Sequence**	**BC09**	**BC11**	**Z-score**
HRLPWPNPELRGT	6000	0	79.51	IKHWNPELT	613	0	24.83
SHAPWLNPELREG	4184	0	65.86	NPELFHMKG	425	0	20.66
VVNPEVRFGGVGG	3119	1	56.57	FAVAFNPDW	333	0	18.28
EGMVNPEVRGRMV	2734	0	52.9	GVLNPEVRA	333	0	18.28
PTTPHINPELTLR	2057	0	45.75	VRNPEYRLL	308	0	17.58
VNPEARWPAGGGE	1740	0	42.02	VRRSLNPEV	263	0	16.24
PRIPLTNPELTAH	1569	0	39.88	RSWYNPELS	246	0	15.7
GHAPKLNPDALIK	1247	0	35.5	qRTVLNPEF	236	0	15.38
RPARNPEASLS	1049	0	32.53	NPDLMAIGW	227	0	15.08
TINPEYRGLQSGV	1043	0	32.44	MVNPEIANR	216	0	14.71
PLAATRNPELSST	1030	0	32.23	GWFVNPELS	204	0	14.3
HPPLARHVNPEVT	877	0	29.72	VINPEILRT	203	0	14.26
RQLRTLNPDLSWH	862	0	29.47	RGINPEVMG	196	0	14.01
QQGPYLNPEVTAQ	854	0	29.33	MWNPELLRT	179	0	13.39
DHMRFGRAVNPEV	819	0	28.72	FVNPEFLSL	178	0	13.35
QLSPAINPELRSV	800	0	28.38	WGLNTNPDW	166	0	12.89
AGPRAYNPDALLT	700	0	26.54	NPEVSRGEW	154	0	12.42
LPIPHINPELRHT	692	0	26.38	MFNPELRGL	144	0	12.01
TMNPELRYPGNVN	677	0	26.09	NPDLSVIGT	139	0	11.8
NTHRTWNPEVIKG	648	0	25.53	VRNVEVRLS	138	0	11.76
VINPELRNLCTMq	572	0	23.97	ALNPELVVP	133	0	11.54
NPEWMKQQKMMME	559	0	23.7	LNPEISGLS	128	0	11.32
ACRGLVCGVNPEL	551	0	23.53	FFVNPEVRL	128	0	11.32
KDARIVWNPELVM	498	0	22.36	RRLNPEIGQ	125	0	11.19
NPEWRQQMRVLQE	498	0	22.36	GGMWNPELR	124	0	11.14
QLPLSFNPDLVSR	497	0	22.34	NPEVRVGMS	121	0	11.01
LNPEVRTGKVSVP	490	0	22.18	ERTLNPEVS	118	0	10.87
NPELTRLMVRMQD	466	0	21.63	RqINPDVRS	112	0	10.59
RPPPQLNPDLLFL	464	0	21.58	VFNPEVRSW	111	0	10.54
PLIGTLNPELRNH	448	0	21.21	WLNPEIIET	110	0	10.49
RGWNPELFVGGGL	436	0	20.92	GRGFNPELS	109	0	10.45

YPPRTLNPDLLSH	414	0	20.38	VNPEFLRGL	108	0	10.4
GPAPYLNPEIQYR	403	0	20.11	VNPELRFTV	108	0	10.4
NPELRLLQNQPWP	392	0	19.83	RLNPEVVqG	107	0	10.35
IGTRqLNPEIVQR	384	0	19.63	GRAFNPELS	105	0	10.25
RHALWRNPELVqA	373	0	19.34	LNPELAIGR	98	0	9.9
QASALNPEIRARM	362	0	19.06	ALNPELRSS	95	0	9.75
NLNPDYYASRMWT	359	0	18.98	LLNPELRGL	91	0	9.54
WPIVYNPEVLATR	356	0	18.9	ARLNPEVYG	90	0	9.49
YqGINPEVTAKTM	346	0	18.63	NPEVMANGT	89	0	9.44
NPELRAGLEGLK	346	0	18.63	VRVFNPEVA	88	0	9.38
GPTSSLNPEYRSL	339	0	18.44	RRHVNPDFS	88	0	9.38
PQARTLNPEWTKA	331	0	18.22	NPEVVRFGL	85	0	9.22
HPAISQHLNPEVT	320	0	17.91	VNPELRSLV	84	0	9.17
GGRGAQLNPELYQ	318	0	17.86	NPDFRSVGM	84	0	9.17
EQGRGSWVNPEMT	316	0	17.8	MNPEFLVSR	82	0	9.06
PTNPELRLYRLQD	313	0	17.72	RGVTVNPDW	82	0	9.06
NPEVRYVQLGTME	309	0	17.6	KFFGVNPEW	81	0	9
RRSAEPHLNPEFG	305	0	17.49	VWSVNPELR	80	0	8.95
SPGLHRNPEVGEA	299	0	17.31	RGFNPELRM	79	0	8.89
RGWNPEVRLDKEW	296	0	17.23	RRINPEVSG	78	0	8.84
VNPELHQHFRGFG	280	0	16.75	LLNPDVYSR	78	0	8.84
SGGMLNPEVAPWR	278	0	16.69	NPELISMGV	78	0	8.84
DQKVQHRLNPELR	277	0	16.66	MFNPDLLLS	76	0	8.72
LNPELRLAGSGQL	277	0	16.66	RLNPEVWqV	74	0	8.61
EGRRGMNPELVPA	271	0	16.48	GVINPEVGR	74	0	8.61
RRSSQLNPELVSF	261	0	16.17	LGINPEVRS	74	0	8.61
SSLPKPGPVNPDF	259	0	16.11	RTRVLNPDW	73	0	8.55
RHTAVLNPDLRSA	255	0	15.99	FRRGLNPEL	72	0	8.49
VINPECRLARLGq	255	0	15.99	RGLNPELSA	71	0	8.43
RNDRTINPEALGG	253	0	15.92	NAFTNPEVG	71	0	8.43
RLPTKQVLNPEIQ	251	0	15.86	GGLNPEIRG	71	0	8.43
NRSPSLNPEFRAV	237	0	15.41	YRSFINPEV	68	0	8.25
PqTPSHHLNPELL	234	0	15.31	HLNPDLLGA	68	0	8.25
TPHLNPELDNHWR	228	0	15.11	ARNPDVVAG	67	0	8.19
GARHVLNPDVRDR	226	0	15.05	INPELSRGL	67	0	8.19
RNLPREFLNPEIN	221	0	14.88	NPEVYRSGV	66	0	8.13
KHSLNPEVTAMGT	219	0	14.81	VLNPELAQW	66	0	8.13
AQIPTYNPELRVH	216	0	14.71	VINPEFRLY	65	0	8.06
LPSSPRFAFNPDW	214	0	14.64	NPEYRAAFS	65	0	8.06
LPPPFLNPDLSSL	210	0	14.5	NPELLSSGF	65	0	8.06
NDPRWVNPEVLMS	210	0	14.5	VGGFLNPDW	65	0	8.06
RNSPALNPDLLGA	208	0	14.43	NPELYSLGL	64	0	8
RLPREYMHFNPEL	208	0	14.43	GAVNPELRG	63	0	7.94
VNPEIAVLGARTE	206	0	14.37	FSMRWNPEV	63	0	7.94
VHSPSLNPELYLG	205	0	14.33	RYVNPEVMG	63	0	7.94
YPTPHLNPELMSR	200	0	14.15	RGLNPDFFA	63	0	7.94
RHHLAFNPEVLAG	197	0	14.05	LSLNPELMR	62	0	7.88

TINPELAGSAAAA	196	0	14.01	LLALNPELM	62	0	7.88
GPTSWRNPEVAqG	196	0	14.01	NPEFLRIGL	62	0	7.88
PLNPELRTFVRGG	195	0	13.98	NPEFLSFSV	62	0	7.88
PAPRTLNPDYLMQ	191	0	13.83	NPDFRFAMG	61	0	7.81
DRATAYNPEVRGI	190	0	13.8	CNPELCRSq	61	0	7.81
GVRPSFHLNPDIR	190	0	13.8	NPDLRSISV	61	0	7.81
ALHQQLNPELHSL	188	0	13.72	WLNPEFSSS	60	0	7.75
NRYTTINPELEKL	188	0	13.72	NPEFGALWR	58	0	7.62
PRPIFPNPELLSR	188	0	13.72	FLNPELSRY	57	0	7.55
GqRWINPELREGM	187	0	13.69	NPELRFFGA	57	0	7.55
NPEVRLFGTPHDF	184	0	13.58	GGRYNPELA	56	0	7.49
SKTARGHWNPEVq	184	0	13.58	VLNPELASL	56	0	7.49
RPPRSFNPELVYG	183	0	13.54	FVNPEFAST	56	0	7.49
SQSQHYNPEMIAL	181	0	13.46	GWNPELISL	56	0	7.49
NHPRAQWLNPEYK	180	0	13.43	VYQRNPELL	56	0	7.49
NPEVFATPAWLVG	178	0	13.35	VQLNPELGG	55	0	7.42
QINPEMKAMMLSS	178	0	13.35	NPEVRMRGS	54	0	7.35
DRLHSRNPELRFY	178	0	13.35	FGRVVNPDF	54	0	7.35
NPEVRKRLEHTEV	178	0	13.35	RAVTLNPEY	54	0	7.35
GRSALNPEVLEGQ	169	0	13.01	FLNPEMVLS	53	0	7.28
GRPARHTLNPEVS	169	0	13.01	FFLNPEFMS	53	0	7.28
RGAQWLNPEVYGR	169	0	13.01	GFYVNPELS	53	0	7.28

7.4.10 11895

13mer				9mer			
Sequence	**BC22**	**BC23**	**Z-score**	**Sequence**	**BC10**	**BC11**	**Z-score**
ERIASGVLGGRED	15638	0	137.18	WVqIVRSGS	7881	0	92.48
VGWDCGEGCVGVV	9345	0	101.95	WAqIVRSGD	2636	0	52.03
YTWSGGYVAKSGG	5165	0	73.96	WVqIASSGS	1823	0	43.09
GLREVCRWGRCEG	3497	0	60.29	WVqVVKAGS	748	0	27.45
WTWEGGWKDRGGG	2822	0	53.95	WVqVVRSGS	698	0	26.51
ERCEWLVGGGCKS	1874	1	43.7	WYqIVLSGS	685	0	26.26
CMWDEGLGWCAWE	1750	0	42.23	WAqIVRSGS	638	0	25.34
VTWDGGWLSMSRG	1662	0	41.14	RTWDGGYIS	462	0	21.54
WMWECGLGCVGVV	1593	0	40.26	GYRTWTGGF	458	7	21.03
WRMSYDASWGGAG	1342	0	36.9	DRVYFDSVT	434	0	20.88
GQAWVTWEGGYKI	1256	1	35.64	GYVTWAGGF	409	1	20.2
CFWNGVDCAVVVV	1180	0	34.57	NRAYFVTVT	375	0	19.4
TYTEVGRVVWAVR	1046	0	32.53	WVqIARSGS	375	0	19.4
VCWSGGLEVDCVG	943	0	30.87	YTRWFYAVD	328	0	18.14
SCLWDDVAGSCVV	846	6	28.94	GFITWDGGF	289	0	17.02
GLWVGGRGWQWFE	682	0	26.21	GYVTWAGGW	280	0	16.76
CLWSDVLGMCVGS	651	0	25.6	GFVTWDGGY	263	1	16.16
TYWDSGGRVYMME	604	0	24.66	DATLFPFII	258	0	16.08
RYWEYGGGVWWWL	565	0	23.84	YARLYREGF	242	0	15.58

WFPSVSGEWWVRI	435	0	20.91	WVQVLRAGS	234	0	15.31
ESWQGGWLDRGKS	428	0	20.74	WRSWTGGYE	229	0	15.15
FVCVDGWPCYLYF	408	0	20.24	FTWEGGYRW	231	8	14.55
GPGWVSWAVGEEV	379	0	19.51	WVqISQSGS	211	0	14.54
CWWDGLGCLVARV	361	0	19.04	RILRNFDAM	206	0	14.37
SEGDGWVSWSGGW	344	0	18.58	GYITWAGAY	214	4	14.3
RYFSYGGTVLAWR	330	0	18.2	FRLGMMSAV	202	0	14.23
CFWDGVECVGEGG	315	0	17.78	WMqIVRSGS	199	0	14.12
YTWDGSWTNRRSL	298	0	17.29	ISAMRIVSS	196	0	14.01
VSWSGAWYLGPGG	299	1	17.24	VTWSGGYVR	192	0	13.87
VCLWDGVGCLGGV	273	0	16.55	DSRLFVYTV	182	0	13.5
VCEWDGFGCVYRV	272	0	16.52	WMqIVRAGD	170	0	13.05
ADRCFWDGFVCVV	270	0	16.46	SLGWYLRVI	173	2	12.97
GRMAACGMGVCWG	269	0	16.43	AqVIRFYGF	165	0	12.86
RDWACGIGCVGVV	251	0	15.86	GYISWNGGY	170	2	12.85
DHCVWGCVGVGRD	249	0	15.8	FTWDGGFRY	191	11	12.84
VDLGRDVSWFSLG	232	0	15.25	WIqVIRSGS	164	0	12.82
CLWSASLGGCLDD	217	0	14.75	FRGLFFYTM	154	0	12.42
GWRECGVGCVGVV	216	0	14.71	ITWSGPWRP	155	1	12.36
GGGWVSWSGGYSG	203	0	14.26	GWMSWTGGF	152	0	12.34
GFVSWAGCWTCGG	186	0	13.65	ARFMQDVGL	150	0	12.26
GYWDWGGVVVVWD	185	0	13.62	WMQVVKSGS	146	0	12.09
GSLHCWWGDGGCV	184	0	13.58	SDARVAWVM	142	0	11.92
LFPWqLLPFLDWG	177	0	13.32	CVSNLFCPV	135	1	11.52
DHCVWVDGGGCGV	176	0	13.28	FRADIGYWG	125	0	11.19
GHCVWVDRGFCAV	174	0	13.2	FRRLMLEGF	120	0	10.96
WCWDCFFHVMGGG	171	0	13.09	GGLLRRVID	119	0	10.92
PCVSVGWTMVCWA	165	0	12.86	WGHFVGYFP	118	0	10.87
GWRTWEGPWFGGG	152	0	12.34	GWLSWSGGY	117	0	10.82
RGESEWRTWAGGY	144	0	12.01	GRLYSVAGR	113	0	10.64
GDGWVSWSGGGSV	141	0	11.88	VTWEGGYRY	112	0	10.59
TGSRCFWGYGGCV	139	0	11.8	FTWVGGYRG	112	0	10.59
HCVWSVDLSMCVL	136	0	11.67	FTWTGGYDK	114	1	10.56
RGWYTWSGGWREV	136	0	11.67	GHSLRSAMD	111	0	10.54
VQVDLHGRILGPV	134	0	11.58	VTWDGGYSS	107	0	10.35
PVCAFDLPVCYWV	133	0	11.54	GWTTWVGGF	104	0	10.2
WGWDCGYGCVGVV	128	0	11.32	EYYVWVGGY	106	2	10.05
LQRHCWWGEGGCV	118	0	10.87	YTWDGGFTG	112	5	10
VTWEGGWRGGGGG	113	0	10.64	YSWNGGYRF	99	1	9.83
CWWVDELGVCGVV	112	0	10.59	NQVMRYYSV	95	0	9.75
GGDWRDNWRGAVE	105	0	10.25	GYVSWTGPF	99	2	9.7
VVWDGGWLDRMSF	104	0	10.2	FTWVGGYRA	93	0	9.65
GGTWYADWVRGGP	103	0	10.15	VATRIVGGV	91	0	9.54
VSWTGGWRAGAVV	103	0	10.15	GGLLRRVFD	88	0	9.38
PAYVGHYGMSLGL	99	0	9.96	YRMVVWYGL	87	0	9.33
qICFWEVFGGCRV	96	0	9.8	LRWALVSSG	87	0	9.33

CWWDGVGCARGGV	92	0	9.6	GYLSWVGGY	86	1	9.14
LTWTGSWREGRGL	88	0	9.39	GWMSWNGGF	82	0	9.06
VFPPRVSHFLFAL	86	0	9.28	VTWSGGWVR	86	2	9
RCSDIGWTFVCGL	86	0	9.28	VTWDGGYRV	83	1	8.97
DWVVCVGGFGACG	85	0	9.22	YTWAGGYVA	83	2	8.84
GVCLWVDGVGCSG	84	0	9.17	HRWVATAAN	77	0	8.78
GRCFWVDGLGCGV	83	0	9.11	YTWTGGWKM	80	2	8.67
VLPWQHVIWFYSV	78	0	8.84	VTWSGGYVE	74	0	8.61
DICYWVSAVGCqV	75	0	8.66	FKSTMVVAL	73	0	8.55
GVVCWWDGVGCAG	73	0	8.55	GSWVQVASR	72	0	8.49
GYVSWVGGWLGGG	72	0	8.49	ETWTGSWLR	76	2	8.43
HCWWMGDGCVFqV	70	0	8.37	YTWSGGMHR	73	1	8.4
CGWDVGLGVCVDV	69	0	8.31	GLGWYLRVI	73	1	8.4
GSWCWWVGDGCWV	66	0	8.13	GDVFRRVFD	70	0	8.37
DVqVWSGGWLAGG	58	0	7.62	RTWSGGYLS	70	0	8.37
GWCRWEVGGGCAG	54	0	7.35	GWMSWSGGF	72	1	8.34
RGHCWWVMGLGCV	54	0	7.35	DASLWVYSV	69	0	8.31
GHVWWGDYGGEVG	53	0	7.28	KDqVVVWIM	68	0	8.25
ERIASGVLGGREE	52	0	7.21	GGFLRRTMD	68	0	8.25
VCLWDGFECKGSG	51	0	7.14	RNAVSVATW	67	0	8.19
DSWNKWVLFGGLT	51	0	7.14	RTWAGGYAG	65	0	8.06
CWWVEGLGVCREQ	49	0	7	GLNWYMTVI	67	1	8.03
GWKVWEGGWLEGG	47	0	6.86	EWFSWTGGY	66	1	7.97
STWSGGWGDGPGV	46	0	6.78	YTWSGGYTL	68	2	7.94
RMLVWGGFLREGV	45	0	6.71	WVqVIKAGS	62	0	7.88
CWWDELKAMCGWS	45	0	6.71	CASNIFCSV	62	0	7.88
CSWSSYHGTCLSV	44	0	6.63	FVRLSREGF	64	1	7.84
AHCFWDEPLMqCV	44	0	6.63	GWLSWDGGF	61	0	7.81
VDCYWLEGGDCWV	42	0	6.48	FRSVMMVSF	61	0	7.81
YTWTGGPQAMKVQ	40	0	6.33	AVPVRVVMG	61	0	7.81
GSCYRVCDGGYGE	39	0	6.25	HWTGNWLAR	69	4	7.71
WMGDGWVSWSGPR	38	0	6.17	WRLAHDSSF	59	0	7.68
CFWDVVLGSCVGV	35	0	5.92	FRLGMIYAV	58	0	7.62
GRGAGWVEWSGGY	34	0	5.83	NRVYFEVVT	58	0	7.62
WGPWCMVTSAMQq	32	0	5.66	VTWVGGYTA	58	0	7.62

7.4.11 66-3-9C

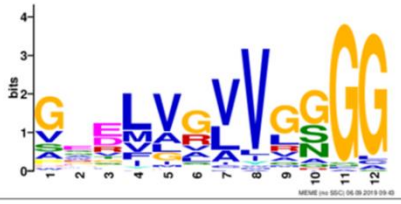
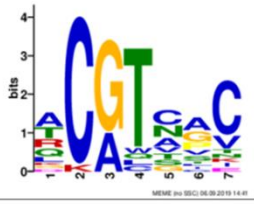
13mer				**9mer**			
Sequence	**BC24**	**BC23**	**Z-score**	**Sequence**	**BC12**	**BC11**	**Z-score**
SMWESAKTIGEGG	2134	8	46.48	qIGRVVDRV	38	0	6.17
GLWEERARGMRYM	448	2	21.09	GLWAATGVG	38	0	6.17
RLWENqLLSEAPS	278	1	16.61	SLWqVSQST	30	0	5.48
GLWEGRGVVSGSV	267	0	16.36	WSGGSRTME	26	0	5.1
GAWSMWGVGQDGA	241	2	15.37	LRIGEGVMG	22	0	4.69
GLWEGEVVGTMQS	231	0	15.22	GLWEINLGG	21	0	4.58
GLWqVGVSGGGGI	220	0	14.85	GNWSMWAKE	21	0	4.58

SLWEAGRVEMKqG	200	0	14.16	DFWNARLKV	21	0	4.58
GLWGVEGGGIVGG	191	0	13.83	TLWqVTAVS	20	0	4.47
GPGGAWGPGASQV	159	0	12.62	GLWEPSGRH	19	0	4.36
GLWERGASRSVLL	124	1	11.02	SFWERqVHD	18	0	4.24
GLWEGGLRGRGLD	107	0	10.35	LRIGEGVFN	17	0	4.12
GPFALWSAGLQRA	103	0	10.15	GLWESTNRG	17	0	4.12
GLWEGGLVGGGGD	100	0	10.01	ALWSRSSGT	16	0	4
AMWEQMSERREVM	96	0	9.8	DLWESVSGN	15	0	3.87
GLWESVRAVGGDP	89	0	9.44	WKGGTPVHL	15	0	3.87
RLWEALPTKGATG	88	1	9.24	ALWSHTMVA	14	0	3.74
GGWGLWEVGVGRV	85	0	9.22	WLQAYANRA	14	0	3.74
DLWGAGARQEEVV	83	1	8.96	ALWETAAGS	13	0	3.61
WPAQGWQLWEEER	78	0	8.84	GLWFDPKRS	12	0	3.46
SLWEKVqESPSAA	77	0	8.78	WYGGVVGAV	12	0	3.46
GLWDTSTSKVERV	77	0	8.78	DLWGGRRAV	12	0	3.46
GLPWGPTEKKMPE	79	1	8.74	DLWSLTKGA	12	0	3.46
ELWNRGMRQGEGV	74	0	8.61	TLWETGTVA	12	0	3.46
GLWENTMPRTAqT	74	0	8.61	SMWAGHASF	11	0	3.32
GEWGLWERGGSGG	74	0	8.61	NLWSTVSSA	11	0	3.32
GLWGAGDDGGGGS	74	0	8.61	DLWSRRGAP	11	0	3.32
GLWEIGRSVPGVP	70	0	8.37	AMWEDLRAR	11	0	3.32
GLWEMASKWGSMD	68	0	8.25	GLWNMNADA	11	0	3.32
DLWEKGMGRQEQA	67	0	8.19	TLWqDTVLD	10	0	3.16
TSEVGRIGGEGLN	66	0	8.13	VLWELSRTA	10	0	3.16
DMWEAERLEQRMG	65	0	8.07	GLWNVVVPG	10	0	3.16
GTWGLWKEREEMG	64	0	8	DLWGEDYRR	10	0	3.16
GLWEEVAQRAGMA	64	0	8	GLWERYEVS	10	0	3.16
ALWEEQRYLHQqG	61	0	7.81	GLWAVGAGL	10	0	3.16
GVWDLWGRGVGGV	68	3	7.76	RMWEILTAq	9	0	3
GLWERAGqPGDKG	57	0	7.55	DMWERTSLV	9	0	3
GEWAMWARE	53	0	7.28	DLWqARSKV	9	0	3
GEWVLWqGVESGI	52	0	7.21	RGRIGSEMA	9	0	3
GLWERqMPEMGRE	52	0	7.21	GLWEARVES	9	0	3
GLWEQGGRGGGQG	52	0	7.21	ALWASTLGI	9	0	3
GMWEHVGVVGGGE	51	0	7.14	SMWEPTASA	9	0	3
GLWEEGRQVVGKG	51	0	7.14	WLAPAMSNG	8	0	2.83
GLWEARGRVGVVV	53	1	7.09	GLWNGSLSS	8	0	2.83
GPWRLWGEGGGRL	49	0	7	KAFNLWSPV	8	0	2.83
TEWGLWSVARAEE	49	0	7	GLWqTANRV	8	0	2.83
MLWEDVGGGLESG	48	0	6.93	GLWqAYQWE	8	0	2.83
GLWEVGGRVSGGE	48	0	6.93	SLWqIPVRD	8	0	2.83
GGWRMWERRSVTQ	48	0	6.93	AMWGGRGLQ	8	0	2.83
DMWELRGTAGPAG	48	0	6.93	MALWGSGVG	8	0	2.83
ACPPAWMSPRCSS	47	0	6.86	WSGGELGID	8	0	2.83
GLWELGLGGRSEG	45	0	6.71	TLWNAARVG	8	0	2.83
GGPRWMLWEETAK	45	0	6.71	GLWHRDLGL	7	0	2.65

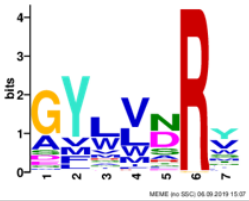
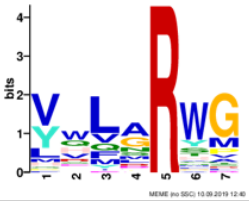
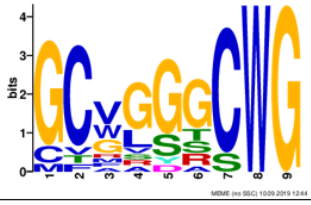
GLWEFGLGRGVAM	45	0	6.71	GGRIGEGSQ	7	0	2.65
GLWEMGRGIEAGG	44	0	6.63	GLWEqVLAQ	7	0	2.65
GLWEVGEGLGAKG	43	0	6.56	GLWGSSLGD	7	0	2.65
GLWEqGLGGEGRV	43	0	6.56	RLWESDWGA	7	0	2.65
GLWGEGMKGAGPG	43	0	6.56	GLWSVGSVG	7	0	2.65
GLWEKQSALITRQ	42	0	6.48	GLWGRSEVL	7	0	2.65
DMWQQNRQVRTWE	44	1	6.43	GLWESRVPV	7	0	2.65
DMWERMGRMGDRE	41	0	6.4	MRIGENVTV	7	0	2.65
SQSqGLWFALDAY	41	0	6.4	GLWQKGDWT	7	0	2.65
RGPIGSERLMTQG	41	0	6.4	ALWqSVGVE	7	0	2.65
GMWESGVPGVGEG	46	2	6.38	RLWENDWGR	7	0	2.65
NLWEGVNGSGAGG	40	0	6.33	RLWDAGIVE	7	0	2.65
GLWNIGVGVGGGGG	40	0	6.33	DMWAESRPV	7	0	2.65
AGAEALWGRGGEG	40	0	6.33	ALWqSVGFL	7	0	2.65
GLWEMGVVGAKAV	39	0	6.25	DFWVHSGAH	7	0	2.65
KEQLRWDMWELRW	41	1	6.19	GLWGRGDQW	7	0	2.65
TEGQIGRMKERME	38	0	6.17	GMWEGGSGI	7	0	2.65
GMWEETGTGSVAR	38	0	6.17	WTGGERGVL	7	0	2.65
GLWEGRSAAASSS	38	0	6.17	WLFPRGVGH	7	0	2.65
ALWEVGVGGGIGV	37	0	6.08	GLWQIGVPV	7	0	2.65
GLWEGGLVGGGVV	37	0	6.08	GLWELVYER	6	0	2.45
SMWGPGDVGGRLG	37	0	6.08	GLWNHVKVG	6	0	2.45
WQGGEASKV	37	0	6.08	GLWESSSSI	6	0	2.45
ALWEHAEAAAART	36	0	6	DLWTVAVVq	6	0	2.45
GSGWGLWTEGKEV	36	0	6	GLWGRGDEF	6	0	2.45
GSWGMWSADGGGG	36	0	6	GLWqVTGSD	6	0	2.45
ALWERGVTEGGGG	36	0	6	RLWESTNVS	6	0	2.45
GLWEEVRTLEKTq	38	1	5.94	TLWGMGGQG	6	0	2.45
VLWEVGGGIARDV	35	0	5.92	DLWSGTGRF	6	0	2.45
GLWEGAGGGMDRE	35	0	5.92	GLWNVSGAE	6	0	2.45
GLWEDARTGGGGL	34	0	5.83	DFWGRRELH	6	0	2.45
GLWEVMGRGEVQV	34	0	5.83	DLWNSVRDD	6	0	2.45
DLWERGLAGGDGI	36	1	5.77	SLWGAGDKV	6	0	2.45
WSGGQVGDGG	36	1	5.77	SLWAKDVTG	6	0	2.45
ALWEPQGRSVRVG	33	0	5.75	SLWDQRMWR	6	0	2.45
NLWEGAGTSRVAG	33	0	5.75	GLWALVRPE	6	0	2.45
GLWERLGGGGGGR	33	0	5.75	DMWSLSDLG	6	0	2.45
AMWERYESQKEWL	33	0	5.75	GLWRTVDTG	6	0	2.45
TLWEEGLGGVSRV	33	0	5.75	SLWGRSNGW	6	0	2.45
GLWERHAQAGGLQ	33	0	5.75	TMWEPRGTI	6	0	2.45
GVWELWGSVDGAV	33	0	5.75	DLWHRFPAD	6	0	2.45
SMWGVGESGQGGG	32	0	5.66	GLWEVVKAD	6	0	2.45
DLWEMGVVSRRGV	32	0	5.66	QCWGPMCTL	6	0	2.45
GLWEMEVRGGDGV	32	0	5.66	GMWESAVPV	6	0	2.45
REFSMWGVQGERL	32	0	5.66	RLWEVTRqE	6	0	2.45
GPWGLWEGREqRG	32	0	5.66	GLWEVTGRV	6	0	2.45
RLWEKMqKSNEVL	34	1	5.6	SLWqVTSRG	6	0	2.45

7.5 Appendix 5: Motifs derived from sequences enriched by peptide phage display

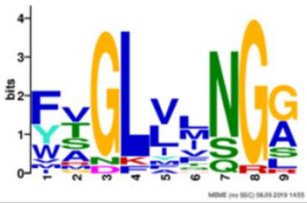
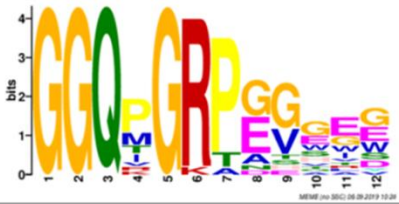
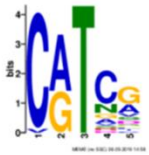
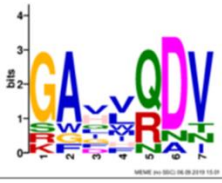
7.5.1 11897

9mer NGS (BC01vs11)	13mer NGS (BC13vs23)
Motif 1 E value: **5.9e-059** Unique enriched sequences in motif: **45/100** Number of reads in motif: **2237** Percentage of top ranked Z score reads: **45.06%** Percentage of total reads: **2.01%** 	**Motif 1** E value: **4.9e-105** Percentage of enriched sequences in motif: **61/100** Number of reads in motif: **36,479** Percentage of top ranked Z score reads: **69.43%** Percentage of total reads: **41.59%** 
Motif 2 E value: **9.4e-018** Unique enriched sequences in motif: **14/100** Number of reads in motif: **695** Percentage of top ranked Z score reads: **13.00%** Percentage of total reads: **0.62%** 	
Motif 3 E value: **5.6e-015** Unique enriched sequences in motif: **20/100** Number of reads in motif: **1121** Percentage of top ranked Z score reads: **22.58%** Percentage of total reads: **1.01%** 	

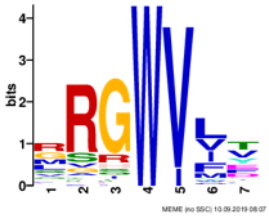
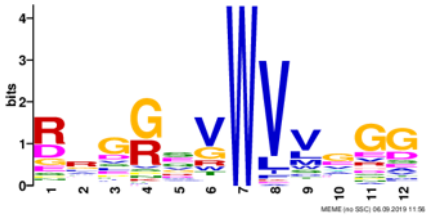
7.5.2 11878

9mer NGS (BC02vs11)	13mer NGS (BC14vs23)
Motif 1 E value: **3.6e-155** Unique enriched sequences in motif: **90/100** Number of reads in motif: **12,481** Percentage of top ranked Z score reads: **91.10%** Percentage of total reads: **10.94%** 	**Motif 1** E value: **9.8e-027** Unique enriched sequences in motif: **39/88** Number of reads in motif: **18,970** Percentage of top ranked Z score reads: **47.10%** Percentage of total reads: **36.40%** 
	Motif 2 E value: **4.2e-014** Unique enriched sequences in motif: **15/88** Number of reads in motif: **3124** Percentage of top ranked Z score reads: **7.76%** Percentage of total reads: **5.99%** 

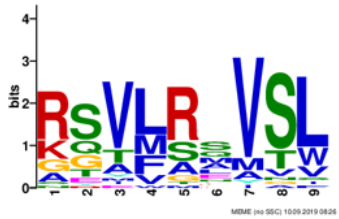
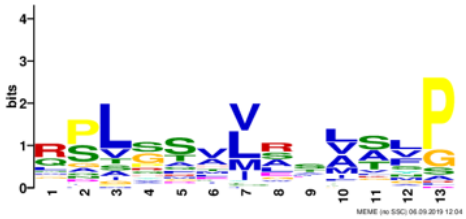
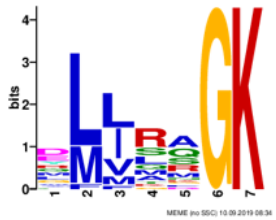
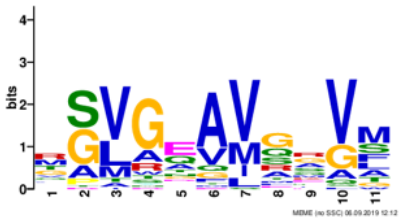
7.5.3 11883

9mer NGS (BC03vs11) Total reads: 97911	13mer NGS (BC15vs23)
Motif 1 E value: **8.3e-073** Unique enriched sequences in motif: **30/100** Number of reads in motif: **2240** Percentage of top ranked Z score reads: **40.27%** Percentage of total reads: **2.29%** 	**Motif 1** E value: **1.3e-046** Unique enriched sequences in motif: **54/100** Number of reads in motif: **49,506** Percentage of top ranked Z score reads: **67.50%** Percentage of total reads: **55.15%** 
Motif 2 E value: **6.8e-025** Unique enriched sequences in motif: **17/100** Number of reads in motif: **661** Percentage of top ranked Z score reads: **11.88%** Percentage of total reads: **0.68%** 	**Motif 2** E value: **2.6e-044** Unique enriched sequences in motif: **18/100** Number of reads in motif: **10,699** Percentage of top ranked Z score reads: **14.59%** Percentage of total reads: **11.92%** 
Motif 3 E value: **3.2e-022** Unique enriched sequences in motif: **26/100** Number of reads in motif: **1331** Percentage of top ranked Z score reads: **23.93%** Percentage of total reads: **1.36%** 	
Motif 4 E value: **1.9e-003** Unique enriched sequences in motif: **10/100** Number of reads in motif: **585** Percentage of top ranked Z score reads: **10.52%** Percentage of total reads: **0.60%** 	

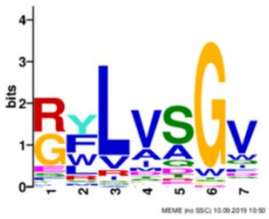
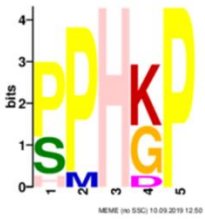
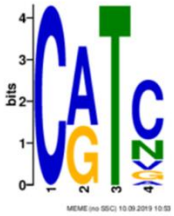
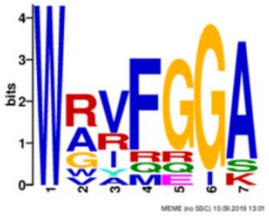
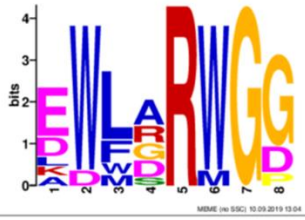
7.5.4 11889

9mer NGS (BC05vs11)	13mer NGS (BC17vs23)
Motif 1 E value: **1.1e-078** Unique enriched sequences in motif: **61/100** Number of peptide reads in motif: **732** Percentage of top ranked Z score reads: **59.80%** Percentage of total peptide reads: **0.93%** 	**Motif 1** E value: **2.2e-057** Unique enriched sequences in motif: **55/100** Number of peptide reads in motif: **6143** Percentage of top ranked Z score reads: **56.84%** Percentage of total peptide reads: **7.39%** 

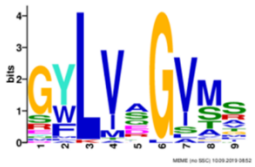
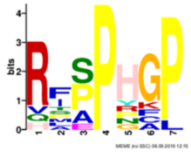
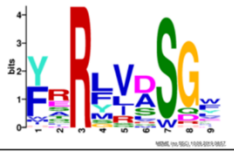
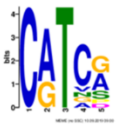
7.5.5 11891

9mer NGS (BC06vs11)	13mer NGS (BC18vs23)
Motif 1 E value: **1.5e-042** Unique enriched sequences in motif: **30/100** Number of peptide reads in motif: **1336** Percentage of top ranked Z score reads: **21.91%** Percentage of total peptide reads: **1.20%** 	**Motif 1** E value: **3.4e-038** Unique enriched sequences in motif: **43/100** Number of peptide reads in motif: **15,764** Percentage of top ranked Z score reads: **55.04%** Percentage of total peptide reads: **16.98%** 
Motif 2 E value: **4.7e-021** Unique enriched sequences in motif: **22/100** Number of peptide reads in motif: **1000** Percentage of top ranked Z score reads: **16.40%** Percentage of total peptide reads: **0.90%** 	**Motif 2** E value: **1.5e-024** Unique enriched sequences in motif: **44/100** Number of peptide reads in motif: **10,735** Percentage of top ranked Z score reads: **37.48%** Percentage of total peptide reads: **11.56%** 

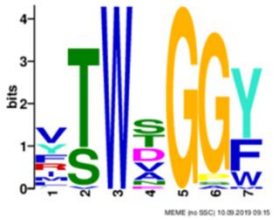
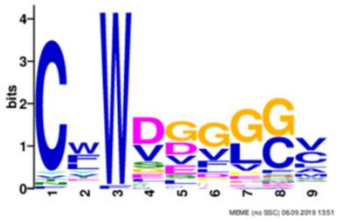
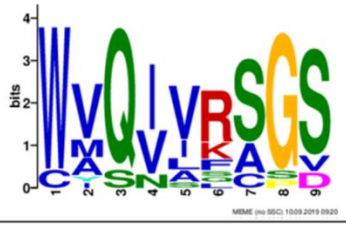
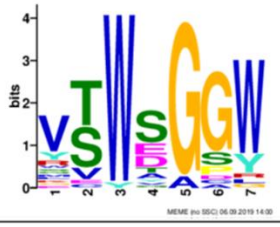
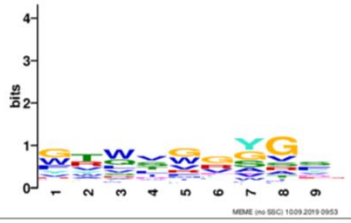
7.5.6 11892

9mer NGS (BC07vs11)	13mer NGS (BC19vs23)
Motif 1 E value: **1.5e-063** Unique enriched sequences in motif: **54/100** Number of peptide reads in motif: **2239** Percentage of top ranked Z score reads: **57.10%** Percentage of total peptide reads: **2.47%** 	**Motif 1** E value: **1.4e-011** Unique enriched sequences in motif: **10/99** Number of peptide reads in motif: **4109** Percentage of top ranked Z score reads: **4.26%** Percentage of total peptide reads: **3.62%** 
Motif 2 E value: 3.7e-022 Unique enriched sequences in motif: **21/100** Number of peptide reads in motif: **841** Percentage of top ranked Z score reads: **21.45%** Percentage of total peptide reads: **0.93%** 	**Motif 2** E value: 1.8e-003 Unique enriched sequences in motif: **11/99** Number of peptide reads in motif: **7045** Percentage of top ranked Z score reads: **7.30%** Percentage of total peptide reads: **6.20%** 
	Motif 3 E value: **3.8e-004** Unique enriched sequences in motif: **10/99** Number of peptide reads in motif: **2807** Percentage of top ranked Z score reads: **2.91%** Percentage of total peptide reads: **2.47%** 

7.5.7 11881

9mer NGS (BC08vs11)	13mer NGS (BC20vs23)
Motif 1 E value: **3.9e-087** Unique enriched sequences in motif: **36/100** Number of peptide reads in motif: **2860** Percentage of top ranked Z score reads: **37.85%** Percentage of total peptide reads: **2.55%** 	**Motif 1** E value: **1.0e-004** Unique enriched sequences in motif: **11/100** Number of peptide reads in motif: **1431** Percentage of top ranked Z score reads: **0.91%** Percentage of total peptide reads: **0.88%** 
Motif 2 E value: **6.0e-052** Unique enriched sequences in motif: **26/100** Number of peptide reads in motif: **2070** Percentage of top ranked Z score reads: **27.40%** Percentage of total peptide reads: **1.85%** 	**Motif 2** E value: **2.1e-004** Unique enriched sequences in motif: **11/100** Number of peptide reads in motif: **1036** Percentage of top ranked Z score reads: **0.66%** Percentage of total peptide reads: **0.64%** 
Motif 3 E value: **1.3e-015** Unique enriched sequences in motif: **12/100** Number of peptide reads in motif: **804** Percentage of top ranked Z score reads: **10.64%** Percentage of total peptide reads: **0.72%** 	**Motif 3** E value: **2.0e-002** Unique enriched sequences in motif: **10/100** Number of peptide reads in motif: **37,793** Percentage of top ranked Z score reads: **24.07%** Percentage of total peptide reads: **23.24%** 
Motif 4 E value: **1.5e-014** Unique enriched sequences in motif: **14/100** Number of peptide reads in motif: **985** Percentage of top ranked Z score reads: **13.04%** Percentage of total peptide reads: **0.88%** 	**Motif 4** E value: **2.4e-003** Unique enriched sequences in motif: **10/100** Number of peptide reads in motif: **92,693** Percentage of top ranked Z score reads: **59.03%** Percentage of total peptide reads: **57.01%** 

7.5.8 11895

9mer NGS (BC10vs11)	13mer NGS (BC22vs23)
Motif 1 E value: **5.1e-089** Unique enriched sequences in motif: **41/100** Number of peptide reads in motif: **6053** Percentage of top ranked Z score reads: **21.30%** Percentage of total peptide reads: **6.03%** 	**Motif 1** E value: **3.0e-028** Unique enriched sequences in motif: **42/100** Number of peptide reads in motif: **12,435** Percentage of top ranked Z score reads: **19.37%** Percentage of total peptide reads: **13.44%** 
Motif 2 E value: **7.3e-053** Unique enriched sequences in motif: **17/100** Number of peptide reads in motif: **16,867** Percentage of top ranked Z score reads: **59.36%** Percentage of total peptide reads: **16.81%** 	**Motif 2** E value: **3.8e-014** Unique enriched sequences in motif: **26/100** Number of peptide reads in motif: **15,071** Percentage of top ranked Z score reads: **23.48%** Percentage of total peptide reads: **16.28%** 
Motif 3 E value: **1.8e-016** Unique enriched sequences in motif: **100/100** Number of peptide reads in motif: **28,415** Percentage of top ranked Z score reads: **100%** Percentage of total peptide reads: **28.32%** 	

8 References

1. Lee, J.E., et al., *Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor.* Nature, 2008. **454**(7201): p. 177-82.
2. Wec, A.Z., et al., *Development of a Human Antibody Cocktail that Deploys Multiple Functions to Confer Pan-Ebolavirus Protection.* Cell Host Microbe, 2019. **25**(1): p. 39-48 e5.
3. Rijal, P., et al., *Therapeutic Monoclonal Antibodies for Ebola Virus Infection Derived from Vaccinated Humans.* Cell Rep, 2019. **27**(1): p. 172-186.
4. Hashiguchi, T., et al., *Structural basis for Marburg virus neutralization by a cross-reactive human antibody.* Cell, 2015. **160**(5): p. 904-912.
5. Bailey, T.L., et al., *The MEME Suite.* Nucleic Acids Research, 2015. **43**(W1): p. W39-W49.
6. Zhao, Y., et al., *Toremifene interacts with and destabilizes the Ebola virus glycoprotein.* Nature, 2016. **535**(7610): p. 169-172.
7. Bharat, T.A.M., et al., *Structural dissection of Ebola virus and its assembly determinants using cryo-electron tomography.* Proceedings of the National Academy of Sciences, 2012. **109**(11): p. 4275-4280.
8. Collar, A.L., et al., *Comparison of N - and O -linked glycosylation patterns of ebolavirus glycoproteins.* Virology, 2017. **502**: p. 39-47.
9. Forbes, K.M., et al., *Bombali Virus in Mops condylurus Bat, Kenya.* Emerg Infect Dis, 2019. **25**(5): p. 955-957.
10. Milligan, J.C., et al., *Structural Characterization of Pan-Ebolavirus Antibody 6D6 Targeting the Fusion Peptide of the Surface Glycoprotein.* J Infect Dis, 2019. **219**(3): p. 415-419.
11. Bornholdt, Z.A., et al., *Host-Primed Ebola Virus GP Exposes a Hydrophobic NPC1 Receptor-Binding Pocket, Revealing a Target for Broadly Neutralizing Antibodies.* mBio, 2016. **7**(1): p. e02154-15.
12. Wang, H., et al., *Ebola Viral Glycoprotein Bound to Its Endosomal Receptor Niemann-Pick C1.* Cell, 2016. **164**(1-2): p. 258-268.
13. Mehedi, M., et al., *A new Ebola virus nonstructural glycoprotein expressed through RNA editing.* J Virol, 2011. **85**(11): p. 5406-14.
14. Sanchez, A., et al., *The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing.* Proc Natl Acad Sci U S A, 1996.
15. Volchkova, V.A., et al., *Ebola Virus GP Gene Polyadenylation Versus RNA Editing.* Journal of Infectious Diseases, 2015. **212**(suppl 2): p. S191-S198.
16. Burk, R., et al., *Neglected filoviruses.* FEMS Microbiol Rev, 2016. **40**(4): p. 494-519.
17. Edenborough, K.M., et al., *Dendritic Cells Generated From Mops condylurus, a Likely Filovirus Reservoir Host, Are Susceptible to and Activated by Zaire Ebolavirus Infection.* Front Immunol, 2019. **10**: p. 2414.
18. Leroy, E.M., et al., *Fruit bats as reservoirs of Ebola virus.* Nature, 2005. **438**(7068): p. 575-576.
19. Schuh, A.J., et al., *Comparative analysis of serologic cross-reactivity using convalescent sera from filovirus-experimentally infected fruit bats.* Sci Rep, 2019. **9**(1): p. 6707.
20. Wittmann, T., et al., *Isolates of Zaire ebolavirus from wild apes reveal genetic lineage and recombinants.* 2007.
21. Leroy, E.M., et al., *Multiple Ebola virus transmission events and rapid decline of central African wildlife.* Science, 2004. **303**(5656): p. 387-90.
22. Bausch, D.G., et al., *Assessment of the risk of Ebola virus transmission from bodily fluids and fomites.* J Infect Dis, 2007. **196 Suppl 2**: p. S142-7.
23. Marí Saéz, A., et al., *Investigating the zoonotic origin of the West African Ebola epidemic.* EMBO Molecular Medicine, 2015. **7**(1): p. 17-23.
24. Leroy, E.M., et al., *Human Ebola outbreak resulting from direct exposure to fruit bats in Luebo, Democratic Republic of Congo, 2007.* Vector Borne Zoonotic Dis, 2009. **9**(6): p. 723-8.
25. Formenty, P., et al., *Human Infection Due to Ebola Virus, Subtype Côte d'Ivoire Clinical and Biologic Presentation.* The Journal of Infectious Diseases, 1999.

26. Barrette, R.W., et al., *Discovery of swine as a host for the Reston ebolavirus.* Science, 2009. **325**(5937): p. 204-6.
27. Jahrling, P.B., et al., *Preliminary report: isolation of Ebola virus from monkeys imported to USA.* Lancet, 1990. **335**(8688): p. 502-5.
28. Marsh, G.A., et al., *Ebola Reston virus infection of pigs: clinical significance and transmission potential.* J Infect Dis, 2011. **204 Suppl 3**: p. S804-9.
29. Penas, J.A., et al., *Risk assessment of Ebola Reston virus in humans in the Philippines.* Western Pac Surveill Response J, 2019. **10**(3): p. 1-8.
30. Goldstein, T., et al., *The discovery of Bombali virus adds further support for bats as hosts of ebolaviruses.* Nat Microbiol, 2018. **3**(10): p. 1084-1089.
31. Karan, L.S., et al., *Bombali Virus in Mops condylurus Bats, Guinea.* Emerg Infect Dis, 2019. **25**(9).
32. Aruna, A., et al., *Ebola Virus Disease Outbreak — Democratic Republic of the Congo, August 2018–November 2019.* MMWR. Morbidity and Mortality Weekly Report, 2019. **68**(50): p. 1162-1165.
33. Prevention, C.f.D.C.a. *2021 Democratic Republic of the Congo, North Kivu Province*. 2021 23rd February 2021 [cited 2021 27th March 2021]; Available from: https://www.cdc.gov/vhf/ebola/outbreaks/drc/2021-february.html.
34. Prevention, C.f.D.C.a. *2021 Guinea, N'Zérékoré Prefecture*. 2021 15th March 2021 [cited 2021 27th March 2021]; Available from: https://www.cdc.gov/vhf/ebola/outbreaks/guinea/2021-february.html.
35. Jacobs, M., et al., *Late Ebola virus relapse causing meningoencephalitis: a case report.* The Lancet, 2016. **388**(10043): p. 498-503.
36. Thorson, A., et al., *Systematic review of the literature on viral persistence and sexual transmission from recovered Ebola survivors: evidence and recommendations.* BMJ Open, 2016. **6**(1): p. e008859.
37. Varkey, J.B., et al., *Persistence of Ebola Virus in Ocular Fluid during Convalescence.* N Engl J Med, 2015. **372**(25): p. 2423-7.
38. Subissi, L., et al., *Ebola Virus Transmission Caused by Persistently Infected Survivors of the 2014–2016 Outbreak in West Africa.* The Journal of Infectious Diseases, 2018. **218**(suppl_5): p. S287-S291.
39. Mulangu, S., et al., *Serologic Evidence of Ebolavirus Infection in a Population With No History of Outbreaks in the Democratic Republic of the Congo.* The Journal of Infectious Diseases, 2018. **217**(4): p. 529-537.
40. Mulangu, S., et al., *High prevalence of IgG antibodies to Ebola virus in the Efé pygmy population in the Watsa region, Democratic Republic of the Congo.* BMC Infectious Diseases, 2016. **16**(1).
41. Arcos Gonzalez, P., et al., *The Epidemiological Presentation Pattern of Ebola Virus Disease Outbreaks: Changes from 1976 to 2019.* Prehosp Disaster Med, 2020. **35**(3): p. 247-253.
42. Henao-Restrepo, A.M., et al., *Efficacy and effectiveness of an rVSV-vectored vaccine in preventing Ebola virus disease: final results from the Guinea ring vaccination, open-label, cluster-randomised trial (Ebola Ça Suffit!).* The Lancet, 2017. **389**(10068): p. 505-518.
43. Inungu, J., K. Iheduru-Anderson, and O.J. Odio, *Recurrent Ebolavirus disease in the Democratic Republic of Congo: update and challenges.* AIMS Public Health, 2019. **6**(4): p. 502-513.
44. Kennedy, S.B., et al., *Phase 2 Placebo-Controlled Trial of Two Vaccines to Prevent Ebola in Liberia.* N Engl J Med, 2017. **377**(15): p. 1438-1447.
45. Mulangu, S., et al., *A Randomized, Controlled Trial of Ebola Virus Disease Therapeutics.* New England Journal of Medicine, 2019. **381**(24): p. 2293-2303.
46. Kaplon, H., et al., *Antibodies to watch in 2020.* mAbs, 2020. **12**(1): p. 1703531.
47. Kaplon, H. and J.M. Reichert, *Antibodies to watch in 2021.* mAbs, 2021. **13**(1): p. 1860476.
48. Qin, E., et al., *Clinical Features of Patients With Ebola Virus Disease in Sierra Leone.* Clinical Infectious Diseases, 2015. **61**(4): p. 491-495.
49. Hunt, L., et al., *Clinical presentation, biochemical, and haematological parameters and their association with outcome in patients with Ebola virus disease: an observational cohort study.* Lancet Infect Dis, 2015. **15**(11): p. 1292-9.

50. Chertow, D.S., et al., *Ebola virus disease in West Africa--clinical manifestations and management.* N Engl J Med, 2014. **371**(22): p. 2054-7.
51. Malvy, D., et al., *Ebola virus disease.* The Lancet, 2019. **393**(10174): p. 936-948.
52. Bah, E.I., et al., *Clinical presentation of patients with Ebola virus disease in Conakry, Guinea.* N Engl J Med, 2015. **372**(1): p. 40-7.
53. Schieffelin, J.S., et al., *Clinical illness and outcomes in patients with Ebola in Sierra Leone.* N Engl J Med, 2014. **371**(22): p. 2092-100.
54. Muehlenbachs, A., et al., *Ebola Virus Disease in Pregnancy: Clinical, Histopathologic, and Immunohistochemical Findings.* Journal of Infectious Diseases, 2017. **215**(1): p. 64-69.
55. Macneil, A., et al., *Proportion of Deaths and Clinical Features in Bundibugyo Ebola Virus Infection, Uganda.* Emerging Infectious Diseases, 2010. **16**(12): p. 1969-1972.
56. Okware, S.I., et al., *An outbreak of Ebola in Uganda.* Tropical Medicine and International Health, 2002. **7**(12): p. 1068-1075.
57. Mattia, J.G., et al., *Early clinical sequelae of Ebola virus disease in Sierra Leone: a cross-sectional study.* The Lancet Infectious Diseases, 2016. **16**(3): p. 331-338.
58. Steptoe, P.J., et al., *Evolving Longitudinal Retinal Observations in a Cohort of Survivors of Ebola Virus Disease.* JAMA Ophthalmol, 2020. **138**(4): p. 395-403.
59. Clark, D.V., et al., *Long-term sequelae after Ebola virus disease in Bundibugyo, Uganda: a retrospective cohort study.* The Lancet Infectious Diseases, 2015. **15**(8): p. 905-912.
60. Munoz-Fontela, C. and A.K. McElroy, *Ebola Virus Disease in Humans: Pathophysiology and Immunity.* Curr Top Microbiol Immunol, 2017. **411**: p. 141-169.
61. Commission, W.I., *Ebola haemorrhagic fever in Zaire, 1976.* Bull World Health Organ, 1978. **56**(2): p. 271-93.
62. Geisbert, T.W., et al., *Pathogenesis of Ebola Hemorrhagic Fever in Primate Models.* The American Journal of Pathology, 2003. **163**(6): p. 2371-2382.
63. Mahanty, S., et al., *Cutting edge: impairment of dendritic cells and adaptive immunity by Ebola and Lassa viruses.* J Immunol, 2003. **170**(6): p. 2797-801.
64. Stroher, U., et al., *Infection and activation of monocytes by Marburg and Ebola viruses.* J Virol, 2001. **75**(22): p. 11025-33.
65. Geisbert, T.W., et al., *Pathogenesis of Ebola Hemorrhagic Fever in Cynomolgus Macaques.* The American Journal of Pathology, 2003. **163**(6): p. 2347-2370.
66. McElroy, A., et al., *Biomarker Correlates of Survival in Pediatric Patients with Ebola Virus Disease.* Emerging Infectious Disease journal, 2014. **20**(10): p. 1683.
67. Mcelroy, A.K., et al., *Ebola Hemorrhagic Fever: Novel Biomarker Correlates of Clinical Outcome.* Journal of Infectious Diseases, 2014. **210**(4): p. 558-566.
68. Lüdtke, A., et al., *Ebola Virus Disease Is Characterized by Poor Activation and Reduced Levels of Circulating CD16*

+

Monocytes. Journal of Infectious Diseases, 2016. **214**(suppl 3): p. S275-S280.
69. Wauquier, N., et al., *Human Fatal Zaire Ebola Virus Infection Is Associated with an Aberrant Innate Immunity and with Massive Lymphocyte Apoptosis.* PLoS Neglected Tropical Diseases, 2010. **4**(10): p. e837.
70. Messaoudi, I., G.K. Amarasinghe, and C.F. Basler, *Filovirus pathogenesis and immune evasion: insights from Ebola virus and Marburg virus.* Nature Reviews Microbiology, 2015. **13**(11): p. 663-676.
71. Gupta, M., et al., *Monocyte-Derived Human Macrophages and Peripheral Blood Mononuclear Cells Infected with Ebola Virus Secrete MIP-1α and TNF-α and Inhibit Poly-IC-Induced IFN-α in Vitro.* Virology, 2001. **284**(1): p. 20-25.
72. Mcelroy, A.K., et al., *Kinetic Analysis of Biomarkers in a Cohort of US Patients With Ebola Virus*

Disease. Clinical Infectious Diseases, 2016. **63**(4): p. 460-467.
73. Martines, R.B., et al., *Tissue and cellular tropism, pathology and pathogenesis of Ebola and Marburg viruses.* J Pathol, 2015. **235**(2): p. 153-74.

74. Zaki, R., Sherif, et al., *A Novel Immunohistochemical Assay for the Detection of Ebola Virus in Skin: Implications for Diagnosis, Spread, and Surveillance of Ebola Hemorrhagic Fever.* The Journal of Infectious Diseases, 1999. **179**(s1): p. S36-S47.
75. Baskerville, A., et al., *Ultrastructural pathology of experimental Ebola haemorrhagic fever virus infection.* J Pathol, 1985. **147**(3): p. 199-209.
76. Hunt, L., et al., *Clinical presentation, biochemical, and haematological parameters and their association with outcome in patients with Ebola virus disease: an observational cohort study.* The Lancet Infectious Diseases, 2015. **15**(11): p. 1292-1299.
77. Mcelroy, A.K., et al., *Von Willebrand Factor Is Elevated in Individuals Infected with Sudan Virus and Is Associated with Adverse Clinical Outcomes.* Viral Immunology, 2015. **28**(1): p. 71-73.
78. Prescott, J.B., et al., *Immunobiology of Ebola and Lassa virus infections.* Nature Reviews Immunology, 2017. **17**(3): p. 195-207.
79. Baize, S., et al., *Defective humoral responses and extensive intravascular apoptosis are associated with fatal outcome in Ebola virus-infected patients.* Nature Medicine, 1999. **5**(4): p. 423-426.
80. Mcelroy, A.K., et al., *Human Ebola virus infection results in substantial immune activation.* Proceedings of the National Academy of Sciences, 2015. **112**(15): p. 4719-4724.
81. Leroy, E., et al., *Human asymptomatic Ebola infection and strong inflammatory response.* The Lancet, 2000. **355**(9222): p. 2210-2215.
82. Brannan, J.M., et al., *Interferon alpha/beta Receptor-Deficient Mice as a Model for Ebola Virus Disease.* J Infect Dis, 2015. **212 Suppl 2**: p. S282-94.
83. Chan, Y.K. and M.U. Gack, *RIG-I-like receptor regulation in virus infection and immunity.* Current Opinion in Virology, 2015. **12**: p. 7-14.
84. Fanunza, E., et al., *Insights into Ebola Virus VP35 and VP24 Interferon Inhibitory Functions and their Initial Exploitation as Drug Targets.* Infectious Disorders - Drug Targets, 2019. **19**(4): p. 362-374.
85. Bhattacharyya, S., et al., *Enveloped Viruses Disable Innate Immune Responses in Dendritic Cells by Direct Activation of TAM Receptors.* Cell Host & Microbe, 2013. **14**(2): p. 136-147.
86. Kotsias, F., I. Cebrian, and A. Alloatti, *Chapter Two - Antigen processing and presentation*, in *International Review of Cell and Molecular Biology*, C. Lhuillier and L. Galluzzi, Editors. 2019, Academic Press. p. 69-121.
87. Ruibal, P., et al., *Unique human immune signature of Ebola virus disease in Guinea.* Nature, 2016. **533**(7601): p. 100-104.
88. Tipton, T.R.W., et al., *Characterisation of the T-cell response to Ebola virus glycoprotein amongst survivors of the 2013–16 West Africa epidemic.* Nature Communications, 2021. **12**(1).
89. Villinger, F., et al., *Markedly Elevated Levels of Interferon (IFN)-σ, IFN-α, Interleukin (IL)-2, IL-10, and Tumor Necrosis Factor-α Associated with Fatal Ebola Virus Infection.* The Journal of Infectious Diseases, 1999. **179**(s1): p. S188-S191.
90. Hutchinson, L., Karen and E. Rollin, Pierre, *Cytokine and Chemokine Expression in Humans Infected with Sudan Ebola Virus.* The Journal of Infectious Diseases, 2007. **196**(s2): p. S357-S363.
91. Rowe, K., Alexander, et al., *Clinical, Virologic, and Immunologic Follow-Up of Convalescent Ebola Hemorrhagic Fever Patients and Their Household Contacts, Kikwit, Democratic Republic of the Congo.* The Journal of Infectious Diseases, 1999. **179**(s1): p. S28-S35.
92. Ksiazek, G., T., et al., *Clinical Virology of Ebola Hemorrhagic Fever (EHF): Virus, Virus Antigen, and IgG and IgM Antibody Findings among EHF Patients in Kikwit, Democratic Republic of the Congo, 1995.* The Journal of Infectious Diseases, 1999. **179**(s1): p. S177-S187.
93. Corti, D., et al., *Protective monotherapy against lethal Ebola virus infection by a potently neutralizing antibody.* Science, 2016. **351**(6279): p. 1339-42.
94. Flyak, A.I., et al., *Cross-Reactive and Potent Neutralizing Antibody Responses in Human Survivors of Natural Ebolavirus Infection.* Cell, 2016. **164**(3): p. 392-405.
95. Sobarzo, A., et al., *Profiling the Native Specific Human Humoral Immune Response to Sudan Ebola Virus Strain Gulu by Chemiluminescence Enzyme-Linked Immunosorbent Assay.* Clinical and Vaccine Immunology, 2012. **19**(11): p. 1844-1852.

96. Kreuels, B., et al., *A Case of Severe Ebola Virus Infection Complicated by Gram-Negative Septicemia.* New England Journal of Medicine, 2014. **371**(25): p. 2394-2401.
97. Onyango, O., Clayton, et al., *Laboratory Diagnosis of Ebola Hemorrhagic Fever during an Outbreak in Yambio, Sudan, 2004.* The Journal of Infectious Diseases, 2007. **196**(s2): p. S193-S198.
98. Towner, J.S., et al., *Rapid Diagnosis of Ebola Hemorrhagic Fever by Reverse Transcription-PCR in an Outbreak Setting and Assessment of Patient Viral Load as a Predictor of Outcome.* Journal of Virology, 2004. **78**(8): p. 4330-4341.
99. Gilchuk, P., et al., *Multifunctional Pan-ebolavirus Antibody Recognizes a Site of Broad Vulnerability on the Ebolavirus Glycoprotein.* Immunity, 2018. **49**(2): p. 363-374 e10.
100. Flyak, A.I., et al., *Broadly neutralizing antibodies from human survivors target a conserved site in the Ebola virus glycoprotein HR2-MPER region.* Nat Microbiol, 2018. **3**(6): p. 670-677.
101. Bornholdt, Z.A., et al., *Isolation of potent neutralizing antibodies from a survivor of the 2014 Ebola virus outbreak.* Science, 2016. **351**(6277): p. 1078-1083.
102. Maruyama, T., et al., *Ebola Virus Can Be Effectively Neutralized by Antibody Produced in Natural Human Infection.* Journal of Virology, 1999. **73**(7): p. 6024-6030.
103. Corti, D., et al., *Protective monotherapy against lethal Ebola virus infection by a potently neutralizing antibody.* Science, 2016. **351**(6279): p. 1339-1342.
104. Luczkowiak, J., et al., *Specific neutralizing response in plasma from convalescent patients of Ebola Virus Disease against the West Africa Makona variant of Ebola virus.* Virus Research, 2016. **213**: p. 224-229.
105. Volchkov, V.E., et al., *Processing of the Ebola virus glycoprotein by the proprotein convertase furin.* Proceedings of the National Academy of Sciences, 1998. **95**(10): p. 5762-5767.
106. Brindley, M.A., et al., *Ebola Virus Glycoprotein 1: Identification of Residues Important for Binding and Postbinding Events.* Journal of Virology, 2007. **81**(14): p. 7702-7709.
107. Lee, J., et al., *Structure of the Ebola virus envelope protein MPER/TM domain and its interaction with the fusion loop explains their fusion activity.* Proc Natl Acad Sci U S A, 2017. **114**(38): p. E7987-E7996.
108. Malashkevich, V.N., et al., *Core structure of the envelope glycoprotein GP2 from Ebola virus at 1.9-A resolution.* Proceedings of the National Academy of Sciences, 1999. **96**(6): p. 2662-2667.
109. Lee, J., et al., *The Roles of Histidines and Charged Residues as Potential Triggers of a Conformational Change in the Fusion Loop of Ebola Virus Glycoprotein.* PLoS One, 2016. **11**(3): p. e0152527.
110. Brecher, M., et al., *Cathepsin Cleavage Potentiates the Ebola Virus Glycoprotein To Undergo a Subsequent Fusion-Relevant Conformational Change.* Journal of Virology, 2012. **86**(1): p. 364-372.
111. Chandran, K., et al., *Endosomal Proteolysis of the Ebola Virus Glycoprotein Is Necessary for Infection.* Science, 2005. **308**(5728): p. 1643-1645.
112. Dube, D., et al., *The Primed Ebolavirus Glycoprotein (19-Kilodalton GP1,2): Sequence and Residues Critical for Host Cell Binding.* Journal of Virology, 2009. **83**(7): p. 2883-2891.
113. Weissenhorn, W., et al., *Crystal Structure of the Ebola Virus Membrane Fusion Subunit, GP2, from the Envelope Glycoprotein Ectodomain.* Molecular Cell, 1998. **2**(5): p. 605-616.
114. Beniac, D.R. and T.F. Booth, *Structure of the Ebola virus glycoprotein spike within the virion envelope at 11 Å resolution.* Scientific Reports, 2017. **7**(1): p. 46374.
115. West, B.R., et al., *Structural basis of broad ebolavirus neutralization by a human survivor antibody.* Nat Struct Mol Biol, 2019. **26**(3): p. 204-212.
116. Carette, J.E., et al., *Ebola virus entry requires the cholesterol transporter Niemann–Pick C1.* Nature, 2011. **477**(7364): p. 340-343.
117. Côté, M., et al., *Small molecule inhibitors reveal Niemann–Pick C1 is essential for Ebola virus infection.* Nature, 2011. **477**(7364): p. 344-348.
118. Miller, E.H. and K. Chandran, *Filovirus entry into cells - new insights.* Curr Opin Virol, 2012. **2**(2): p. 206-14.
119. Wheeler, S. and D.J. Sillence, *Niemann–Pick type C disease: cellular pathology and pharmacotherapy.* Journal of Neurochemistry, 2020. **153**(6): p. 674-692.

120. Gregory, S.M., et al., *Ebolavirus Entry Requires a Compact Hydrophobic Fist at the Tip of the Fusion Loop.* Journal of Virology, 2014. **88**(12): p. 6636-6649.
121. Alvarez, C.P., et al., *C-Type Lectins DC-SIGN and L-SIGN Mediate Cellular Entry by Ebola Virus in cis and in trans.* Journal of Virology, 2002. **76**(13): p. 6841-6844.
122. Jemielity, S., et al., *TIM-family Proteins Promote Infection of Multiple Enveloped Viruses through Virion-associated Phosphatidylserine.* PLoS Pathogens, 2013. **9**(3): p. e1003232.
123. Richard, A.S., et al., *Virion-associated phosphatidylethanolamine promotes TIM1-mediated infection by Ebola, dengue, and West Nile viruses.* Proceedings of the National Academy of Sciences, 2015. **112**(47): p. 14682-14687.
124. Shimojima, M., et al., *Tyro3 Family-Mediated Cell Entry of Ebola and Marburg Viruses.* Journal of Virology, 2006. **80**(20): p. 10109-10116.
125. Takada, A., et al., *Downregulation of beta1 integrins by Ebola virus glycoprotein: implication for virus entry.* Virology, 2000. **278**(1): p. 20-6.
126. Aleksandrowicz, P., et al., *Ebola Virus Enters Host Cells by Macropinocytosis and Clathrin-Mediated Endocytosis.* The Journal of Infectious Diseases, 2011. **204**(suppl_3): p. S957-S967.
127. Nanbo, A., et al., *Ebolavirus Is Internalized into Host Cells via Macropinocytosis in a Viral Glycoprotein-Dependent Manner.* PLoS Pathogens, 2010. **6**(9): p. e1001121.
128. Saeed, M.F., et al., *Cellular Entry of Ebola Virus Involves Uptake by a Macropinocytosis-Like Mechanism and Subsequent Trafficking through Early and Late Endosomes.* PLoS Pathogens, 2010. **6**(9): p. e1001110.
129. Spence, J.S., et al., *Direct Visualization of Ebola Virus Fusion Triggering in the Endocytic Pathway.* mBio, 2016. **7**(1): p. e01857-15.
130. Schornberg, K., et al., *Role of Endosomal Cathepsins in Entry Mediated by the Ebola Virus Glycoprotein.* Journal of Virology, 2006. **80**(8): p. 4174-4178.
131. Misasi, J., et al., *Filoviruses Require Endosomal Cysteine Proteases for Entry but Exhibit Distinct Protease Preferences.* Journal of Virology, 2012. **86**(6): p. 3284-3292.
132. Hood, C.L., et al., *Biochemical and Structural Characterization of Cathepsin L-Processed Ebola Virus Glycoprotein: Implications for Viral Entry and Immunogenicity.* Journal of Virology, 2010. **84**(6): p. 2972-2982.
133. Bruchez, A., et al., *MHC class II transactivator CIITA induces cell resistance to Ebola virus and SARS-like coronaviruses.* Science, 2020. **370**(6513): p. 241-247.
134. Simmons, J.A., et al., *Ebolavirus Glycoprotein Directs Fusion through NPC1*

+

Endolysosomes. Journal of Virology, 2016. **90**(1): p. 605-610.
135. Wang, M.K., et al., *Biochemical Basis for Increased Activity of Ebola Glycoprotein in the 2013-16 Epidemic.* Cell Host Microbe, 2017. **21**(3): p. 367-375.
136. Gregory, S.M., et al., *Structure and function of the complete internal fusion loop from Ebolavirus glycoprotein 2.* Proceedings of the National Academy of Sciences, 2011. **108**(27): p. 11211-11216.
137. Harrison, J.S., et al., *Designed protein mimics of the Ebola virus glycoprotein GP2 α-helical bundle: Stability and pH effects.* Protein Science, 2011. **20**(9): p. 1587-1596.
138. Hoenen, T., et al., *Inclusion Bodies Are a Site of Ebolavirus Replication.* Journal of Virology, 2012. **86**(21): p. 11779-11788.
139. Nanbo, A., et al., *The spatio-temporal distribution dynamics of Ebola virus proteins and RNA in infected cells.* Scientific Reports, 2013. **3**(1).
140. Takamatsu, Y., L. Kolesnikova, and S. Becker, *Ebola virus proteins NP, VP35, and VP24 are essential and sufficient to mediate nucleocapsid transport.* Proceedings of the National Academy of Sciences, 2018. **115**(5): p. 1075-1080.
141. Adu-Gyamfi, E., et al., *Host Cell Plasma Membrane Phosphatidylserine Regulates the Assembly and Budding of Ebola Virus.* Journal of Virology, 2015. **89**(18): p. 9440-9453.
142. Harty, R.N., *No exit: Targeting the budding process to inhibit filovirus replication.* Antiviral Research, 2009. **81**(3): p. 189-197.
143. Misasi, J. and N.J. Sullivan, *Immunotherapeutic strategies to target vulnerabilities in the Ebolavirus glycoprotein.* Immunity, 2021. **54**(3): p. 412-436.

144. Noda, T., et al., *Ebola Virus VP40 Drives the Formation of Virus-Like Filamentous Particles Along with GP.* Journal of Virology, 2002. **76**(10): p. 4855-4865.
145. Kaletsky, R.L., et al., *Tetherin-mediated restriction of filovirus budding is antagonized by the Ebola glycoprotein.* Proceedings of the National Academy of Sciences, 2009. **106**(8): p. 2886-2891.
146. Vande Burgt, N., R. Kaletsky, and P. Bates, *Requirements within the Ebola Viral Glycoprotein for Tetherin Antagonism.* Viruses, 2015. **7**(10): p. 5587-5602.
147. Pallesen, J., et al., *Structures of Ebola virus GP and sGP in complex with therapeutic antibodies.* Nat Microbiol, 2016. **1**(9): p. 16128.
148. de La Vega, M.A., et al., *The multiple roles of sGP in Ebola pathogenesis.* Viral Immunol, 2015. **28**(1): p. 3-9.
149. Mohan, G.S., et al., *Antigenic subversion: a novel mechanism of host immune evasion by Ebola virus.* PLoS Pathog, 2012. **8**(12): p. e1003065.
150. Volchkova, V.A., H.-D. Klenk, and V.E. Volchkov, *Delta-Peptide Is the Carboxy-Terminal Cleavage Fragment of the Nonstructural Small Glycoprotein sGP of Ebola Virus.* Virology, 1999. **265**(1): p. 164-171.
151. Radoshitzky, S.R., et al., *Ebolavirus -Peptide Immunoadhesins Inhibit Marburgvirus and Ebolavirus Cell Entry.* Journal of Virology, 2011. **85**(17): p. 8502-8513.
152. Wells, C.R., et al., *Ebola vaccination in the Democratic Republic of the Congo.* Proceedings of the National Academy of Sciences, 2019. **116**(20): p. 10178-10183.
153. Marzi, A., et al., *Antibodies are necessary for rVSV/ZEBOV-GP-mediated protection against lethal Ebola virus challenge in nonhuman primates.* Proceedings of the National Academy of Sciences, 2013. **110**(5): p. 1893-1898.
154. Lambe, T., G. Bowyer, and K.J. Ewer, *A review of Phase I trials of Ebola virus vaccines: what can we learn from the race to develop novel vaccines?* Philosophical Transactions of the Royal Society B: Biological Sciences, 2017. **372**(1721): p. 20160295.
155. Agnandji, S.T., et al., *Phase 1 Trials of rVSV Ebola Vaccine in Africa and Europe.* New England Journal of Medicine, 2016. **374**(17): p. 1647-1660.
156. Suschak, J.J. and C.S. Schmaljohn, *Vaccines against Ebola virus and Marburg virus: recent advances and promising candidates.* Human Vaccines & Immunotherapeutics, 2019. **15**(10): p. 2359-2377.
157. Ehrhardt, S.A., et al., *Polyclonal and convergent antibody response to Ebola virus vaccine rVSV-ZEBOV.* Nature Medicine, 2019. **25**(10): p. 1589-1600.
158. Cagigi, A., et al., *Vaccine Generation of Protective Ebola Antibodies and Identification of Conserved B-Cell Signatures.* J Infect Dis, 2018. **218**(suppl_5): p. S528-S536.
159. Wang, Y., Z. Liu, and Q. Dai, *A highly immunogenic fragment derived from Zaire Ebola virus glycoprotein elicits effective neutralizing antibody.* Virus Res, 2014. **189**: p. 254-61.
160. Sebastian, S., et al., *A Multi-Filovirus Vaccine Candidate: Co-Expression of Ebola, Sudan, and Marburg Antigens in a Single Vector.* Vaccines, 2020. **8**(2): p. 241.
161. Geisbert, T.W., et al., *Single-Injection Vaccine Protects Nonhuman Primates against Infection with Marburg Virus and Three Species of Ebola Virus.* Journal of Virology, 2009. **83**(14): p. 7296-7304.
162. Matassov, D., et al., *Single-Dose Trivalent VesiculoVax Vaccine Protects Macaques from Lethal Ebolavirus and Marburgvirus Challenge.* Journal of Virology, 2017. **92**(3).
163. Mupapa, K., et al., *Treatment of Ebola Hemorrhagic Fever with Blood Transfusions from Convalescent Patients.* The Journal of Infectious Diseases, 1999. **179**(s1): p. S18-S23.
164. van Griensven, J., et al., *Evaluation of Convalescent Plasma for Ebola Virus Disease in Guinea.* N Engl J Med, 2016. **374**(1): p. 33-42.
165. Dye, J.M., et al., *Postexposure antibody prophylaxis protects nonhuman primates from filovirus disease.* Proceedings of the National Academy of Sciences, 2012. **109**(13): p. 5034-5039.
166. Jahrling, B., Peter, et al., *Ebola Hemorrhagic Fever: Evaluation of Passive Immunotherapy in Nonhuman Primates.* The Journal of Infectious Diseases, 2007. **196**(s2): p. S400-S403.
167. Sullivan, N.J., et al., *CD8+ cellular immunity mediates rAd5 vaccine protection against Ebola virus infection of nonhuman primates.* Nature Medicine, 2011. **17**(9): p. 1128-1131.

168. Wang, H., et al., *Equine-Origin Immunoglobulin Fragments Protect Nonhuman Primates from Ebola Virus Disease.* Journal of Virology, 2018. **93**(5).
169. Dowall, S.D., et al., *Post-exposure treatment of Ebola virus disease in guinea pigs using EBOTAb, an ovine antibody-based therapeutic.* Sci Rep, 2016. **6**: p. 30497.
170. Ulitzka, M., et al., *Engineering therapeutic antibodies for patient safety: tackling the immunogenicity problem.* Protein Engineering, Design and Selection, 2020. **33**.
171. The PREVAIL II Writing Group, f.t.M.-N.P.I.S.T., *A Randomized, Controlled Trial of ZMapp for Ebola Virus Infection.* New England Journal of Medicine, 2016. **375**(15): p. 1448-1456.
172. Saphire, E.O., et al., *Antibody-mediated protection against Ebola virus.* Nat Immunol, 2018. **19**(11): p. 1169-1178.
173. Saphire, E.O., et al., *Systematic Analysis of Monoclonal Antibodies against Ebola Virus GP Defines Features that Contribute to Protection.* Cell, 2018. **174**(4): p. 938-952.e13.
174. Gunn, B.M., et al., *A Role for Fc Function in Therapeutic Monoclonal Antibody-Mediated Protection against Ebola Virus.* Cell Host Microbe, 2018. **24**(2): p. 221-233 e5.
175. Qiu, X., et al., *Characterization of Zaire ebolavirus glycoprotein-specific monoclonal antibodies.* Clinical Immunology, 2011. **141**(2): p. 218-227.
176. Wilson, J.A., et al., *Epitopes Involved in Antibody-Mediated Protection from Ebola Virus.* Science, 2000. **287**(5458): p. 1664-1666.
177. Qiu, X., et al., *Reversion of advanced Ebola virus disease in nonhuman primates with ZMapp.* Nature, 2014. **514**(7520): p. 47-53.
178. Tran, E.E., et al., *Mapping of Ebolavirus Neutralization by Monoclonal Antibodies in the ZMapp Cocktail Using Cryo-Electron Tomography and Studies of Cellular Entry.* J Virol, 2016. **90**(17): p. 7618-27.
179. Murin, C.D., et al., *Structures of protective antibodies reveal sites of vulnerability on Ebola virus.* Proc Natl Acad Sci U S A, 2014. **111**(48): p. 17182-7.
180. Pettitt, J., et al., *Therapeutic Intervention of Ebola Virus Infection in Rhesus Macaques with the MB-003 Monoclonal Antibody Cocktail.* Science Translational Medicine, 2013. **5**(199): p. 199ra113-199ra1.
181. Qiu, X., et al., *Two-mAb cocktail protects macaques against the Makona variant of Ebola virus.* Science Translational Medicine, 2016. **8**(329): p. 329ra33-329ra33.
182. Pascal, K.E., et al., *Development of Clinical-Stage Human Monoclonal Antibodies That Treat Advanced Ebola Virus Disease in Nonhuman Primates.* J Infect Dis, 2018. **218**(suppl_5): p. S612-S626.
183. Macdonald, L.E., et al., *Precise and in situ genetic humanization of 6 Mb of mouse immunoglobulin genes.* Proceedings of the National Academy of Sciences, 2014. **111**(14): p. 5147-5152.
184. Sivapalasingam, S., et al., *Safety, pharmacokinetics, and immunogenicity of a co-formulated cocktail of three human monoclonal antibodies targeting Ebola virus glycoprotein in healthy adults: a randomised, first-in-human phase 1 study.* The Lancet Infectious Diseases, 2018. **18**(8): p. 884-893.
185. Administration, U.F.a.D., *Drugs@FDA: FDA-Approved Drugs*.
186. Misasi, J., et al., *Structural and molecular basis for Ebola virus neutralization by protective human antibodies.* Science, 2016. **351**(6279): p. 1343-1346.
187. Gaudinski, M.R., et al., *Safety, tolerability, pharmacokinetics, and immunogenicity of the therapeutic monoclonal antibody mAb114 targeting Ebola virus glycoprotein (VRC 608): an open-label phase 1 study.* The Lancet, 2019. **393**(10174): p. 889-898.
188. Hensley, L.E., et al., *Demonstration of Cross-Protective Vaccine Immunity against an Emerging Pathogenic Ebolavirus Species.* PLoS Pathogens, 2010. **6**(5): p. e1000904.
189. Macneil, A., Z. Reed, and P.E. Rollin, *Serologic cross-reactivity of human IgM and IgG antibodies to five species of Ebola virus.* PLoS Negl Trop Dis, 2011. **5**(6): p. e1175.
190. Zhao, X., et al., *Immunization-Elicited Broadly Protective Antibody Reveals Ebolavirus Fusion Loop as a Site of Vulnerability.* Cell, 2017. **169**(5): p. 891-904 e15.
191. Gilchuk, P., et al., *Analysis of a Therapeutic Antibody Cocktail Reveals Determinants for Cooperative and Broad Ebolavirus Neutralization.* Immunity, 2020. **52**(2): p. 388-403 e12.

192. Howell, K.A., et al., *Cooperativity Enables Non-neutralizing Antibodies to Neutralize Ebolavirus.* Cell Rep, 2017. **19**(2): p. 413-424.
193. West, B.R., et al., *Structural Basis of Pan-Ebolavirus Neutralization by a Human Antibody against a Conserved, yet Cryptic Epitope.* mBio, 2018. **9**(5).
194. Gilchuk, P., et al., *Efficacy of Human Monoclonal Antibody Monotherapy Against Bundibugyo Virus Infection in Nonhuman Primates.* The Journal of Infectious Diseases, 2018. **218**(suppl_5): p. S565-S573.
195. Flyak, A.I., et al., *Mechanism of human antibody-mediated neutralization of Marburg virus.* Cell, 2015. **160**(5): p. 893-903.
196. Burke, M.J., P.G. Stockley, and J. Boyes, *Broadly Neutralizing Bovine Antibodies: Highly Effective New Tools against Evasive Pathogens?* Viruses, 2020. **12**(4): p. 473.
197. Doria-Rose, N.A., et al., *New Member of the V1V2-Directed CAP256-VRC26 Lineage That Shows Increased Breadth and Exceptional Potency.* Journal of Virology, 2016. **90**(1): p. 76-91.
198. Stanfield, R.L., et al., *The Unusual Genetics and Biochemistry of Bovine Immunoglobulins*. 2018, Elsevier. p. 135-164.
199. Sok, D., et al., *Rapid elicitation of broadly neutralizing antibodies to HIV by immunization in cows.* Nature, 2017. **548**(7665): p. 108-111.
200. Koch, K., et al., *Selection of nanobodies with broad neutralizing potential against primary HIV-1 strains using soluble subtype C gp140 envelope trimers.* Scientific Reports, 2017. **7**(1): p. 8390.
201. Diebolder, C.A., et al., *Complement is activated by IgG hexamers assembled at the cell surface.* Science, 2014. **343**(6176): p. 1260-3.
202. De Jong, R.N., et al., *A Novel Platform for the Potentiation of Therapeutic Antibodies Based on Antigen-Dependent Formation of IgG Hexamers at the Cell Surface.* PLOS Biology, 2016. **14**(1): p. e1002344.
203. Cook, E.M., et al., *Antibodies That Efficiently Form Hexamers upon Antigen Binding Can Induce Complement-Dependent Cytotoxicity under Complement-Limiting Conditions.* J Immunol, 2016. **197**(5): p. 1762-75.
204. Smith, R.I., M.J. Coloma, and S.L. Morrison, *Addition of a mu-tailpiece to IgG results in polymeric antibodies with enhanced effector functions including complement-mediated cytolysis by IgG4.* The Journal of Immunology, 1995. **154**(5): p. 2226.
205. Blundell, P.A., et al., *Engineering the fragment crystallizable (Fc) region of human IgG1 multimers and monomers to fine-tune interactions with sialic acid-dependent receptors.* J Biol Chem, 2017. **292**(31): p. 12994-13007.
206. Wec, A.Z., et al., *A "Trojan horse" bispecific-antibody strategy for broad protection against ebolaviruses.* Science (New York, N.Y.), 2016. **354**(6310): p. 350-354.
207. Kuzmina, N.A., et al., *Antibody-Dependent Enhancement of Ebola Virus Infection by Human Antibodies Isolated from Survivors.* Cell Rep, 2018. **24**(7): p. 1802-1815 e5.
208. Fan, P., et al., *Potent neutralizing monoclonal antibodies against Ebola virus isolated from vaccinated donors.* MAbs, 2020. **12**(1): p. 1742457.
209. Furuyama, W., et al., *Discovery of an antibody for pan-ebolavirus therapy.* Sci Rep, 2016. **6**: p. 20514.
210. Keck, Z.Y., et al., *Macaque Monoclonal Antibodies Targeting Novel Conserved Epitopes within Filovirus Glycoprotein.* J Virol, 2015. **90**(1): p. 279-91.
211. Wec, A.Z., et al., *Antibodies from a Human Survivor Define Sites of Vulnerability for Broad Protection against Ebolaviruses.* Cell, 2017. **169**(5): p. 878-890 e15.
212. Holtsberg, F.W., et al., *Pan-ebolavirus and Pan-filovirus Mouse Monoclonal Antibodies: Protection against Ebola and Sudan Viruses.* J Virol, 2016. **90**(1): p. 266-78.
213. Howell, K.A., et al., *Antibody Treatment of Ebola and Sudan Virus Infection via a Uniquely Exposed Epitope within the Glycoprotein Receptor-Binding Site.* Cell Rep, 2016. **15**(7): p. 1514-1526.
214. Bornholdt, Z.A., et al., *A Two-Antibody Pan-Ebolavirus Cocktail Confers Broad Therapeutic Protection in Ferrets and Nonhuman Primates.* Cell Host Microbe, 2019. **25**(1): p. 49-58 e5.
215. Brannan, J.M., et al., *Post-exposure immunotherapy for two ebolaviruses and Marburg virus in nonhuman primates.* Nature Communications, 2019. **10**(1).

216. Tickle, S., et al., *A Fully Automated Primary Screening System for the Discovery of Therapeutic Antibodies Directly from B Cells.* Journal of Biomolecular Screening, 2015. **20**(4): p. 492-497.
217. Lightwood, D.J., et al., *Antibody generation through B cell panning on antigen followed by in situ culture and direct RT-PCR on cells harvested en masse from antigen-positive wells.* Journal of Immunological Methods, 2006. **316**(1-2): p. 133-143.
218. Clargo, A.M., et al., *The rapid generation of recombinant functional monoclonal antibodies from individual, antigen-specific bone marrow-derived plasma cells isolated using a novel fluorescence-based method.* mAbs, 2014. **6**(1): p. 143-159.
219. Xiao, J.H., et al., *Characterization of Influenza Virus Pseudotyped with Ebolavirus Glycoprotein.* J Virol, 2018. **92**(4).
220. Naqid, I.A., et al., *Mapping polyclonal antibody responses to bacterial infection using next generation phage display.* Scientific Reports, 2016. **6**(1): p. 24232.
221. Goddard, T.D., C.C. Huang, and T.E. Ferrin, *Visualizing density maps with UCSF Chimera.* Journal of Structural Biology, 2007. **157**(1): p. 281-287.
222. Waterhouse, A., et al., *SWISS-MODEL: homology modelling of protein structures and complexes.* Nucleic Acids Research, 2018. **46**(W1): p. W296-W303.
223. Francica, J.R., et al., *Steric Shielding of Surface Epitopes and Impaired Immune Recognition Induced by the Ebola Virus Glycoprotein.* PLoS Pathogens, 2010. **6**(9): p. e1001098.
224. Brisse, M., et al., *Emerging Concepts and Technologies in Vaccine Development.* Frontiers in Immunology, 2020. **11**.
225. Wang, F., et al., *Reshaping Antibody Diversity.* Cell, 2013. **153**(6): p. 1379-1393.
226. Deiss, T.C., et al., *Immunogenetic factors driving formation of ultralong VH CDR3 in Bos taurus antibodies.* Cellular & Molecular Immunology, 2019. **16**(1): p. 53-64.
227. Dong, J., et al., *Structural Diversity of Ultralong CDRH3s in Seven Bovine Antibody Heavy Chains.* Frontiers in immunology, 2019. **10**: p. 558-558.
228. Muyldermans, S., *Nanobodies: Natural Single-Domain Antibodies.* Annual Review of Biochemistry, 2013. **82**(1): p. 775-797.
229. Blanc, M.R., et al., *A one-step exclusion-binding procedure for the purification of functional heavy-chain and mammalian-type gamma-globulins from camelid sera.* Biotechnol Appl Biochem, 2009. **54**(4): p. 207-12.
230. Desmyter, A., et al., *Antigen Specificity and High Affinity Binding Provided by One Single Loop of a Camel Single-domain Antibody.* Journal of Biological Chemistry, 2001. **276**(28): p. 26285-26290.
231. Stanfield, R.L., *Crystal Structure of a Shark Single-Domain Antibody V Region in Complex with Lysozyme.* Science, 2004. **305**(5691): p. 1770-1773.
232. Sellmann, C., et al., *A One-Step Process for the Construction of Phage Display scFv and VHH Libraries.* Molecular Biotechnology, 2020. **62**(4): p. 228-239.
233. Feng, M., et al., *Construction and next-generation sequencing analysis of a large phage-displayed V(NAR) single-domain antibody library from six naïve nurse sharks.* Antib Ther, 2019. **2**(1): p. 1-11.
234. Weiss, R.A. and C.T. Verrips, *Nanobodies that Neutralize HIV.* Vaccines (Basel), 2019. **7**(3).
235. Hanke, L., et al., *An alpaca nanobody neutralizes SARS-CoV-2 by blocking receptor interaction.* Nature Communications, 2020. **11**(1).
236. Yu, J.-S., et al., *Detection of Ebola virus envelope using monoclonal and polyclonal antibodies in ELISA, surface plasmon resonance and a quartz crystal microbalance immunosensor.* Journal of Virological Methods, 2006. **137**(2): p. 219-228.
237. Ducloux, D., et al., *Antithymocytes globulins: Time to revisit its use in kidney transplantation?* International Reviews of Immunology, 2018. **37**(4): p. 183-191.
238. Weber, J., H. Peng, and C. Rader, *From rabbit antibody repertoires to rabbit monoclonal antibodies.* Exp Mol Med, 2017. **49**(3): p. e305.
239. Rief, N., et al., *Production and characterization of a rabbit monoclonal antibody against human CDC25C phosphatase.* Hybridoma, 1998. **17**(4): p. 389-94.
240. Becker, R.S. and K.L. Knight, *Somatic diversification of immunoglobulin heavy chain VDJ genes: evidence for somatic gene conversion in rabbits.* Cell, 1990. **63**(5): p. 987-97.

241. Kodangattil, S., et al., *The functional repertoire of rabbit antibodies and antibody discovery via next-generation sequencing.* mAbs, 2014. **6**(3): p. 628-636.
242. Lavinder, J.J., et al., *Systematic characterization and comparative analysis of the rabbit immunoglobulin repertoire.* PLoS One, 2014. **9**(6): p. e101322.
243. Knight, K.L. and C.R. Winstead, *B lymphocyte development in the rabbit.* Int Rev Immunol, 1997. **15**(3-4): p. 129-63.
244. Brouillette, R.B. and W. Maury, *Production of Filovirus Glycoprotein-Pseudotyped Vesicular Stomatitis Virus for Study of Filovirus Entry Mechanisms.* Methods in molecular biology (Clifton, N.J.), 2017. **1628**: p. 53-63.
245. Beck, S., et al., *Identification of entry inhibitors of Ebola virus pseudotyped vectors from a myxobacterial compound library.* Antiviral Research, 2016. **132**: p. 85-91.
246. Halfmann, P., et al., *Generation of biologically contained Ebola viruses.* Proceedings of the National Academy of Sciences, 2008. **105**(4): p. 1129-1133.
247. Baz, M., et al., *Nonreplicating influenza A virus vaccines confer broad protection against lethal challenge.* mBio, 2015. **6**(5): p. e01487-15.
248. Powell, T.J., et al., *Pseudotyped influenza A virus as a vaccine for the induction of heterotypic immunity.* J Virol, 2012. **86**(24): p. 13397-406.
249. Miller, E.H., et al., *Ebola virus entry requires the host-programmed recognition of an intracellular receptor.* The EMBO Journal, 2012. **31**(8): p. 1947-1960.
250. Kuhn, J.H., et al., *Conserved Receptor-binding Domains of Lake Victoria Marburgvirus and Zaire Ebolavirus Bind a Common Receptor.* Journal of Biological Chemistry, 2006. **281**(23): p. 15951-15958.
251. Kong, R., et al., *Fusion peptide of HIV-1 as a site of vulnerability to neutralizing antibody.* Science, 2016. **352**(6287): p. 828-833.
252. Deffieu, M.S. and S.R. Pfeffer, *Niemann-Pick type C 1 function requires lumenal domain residues that mediate cholesterol-dependent NPC2 binding.* Proceedings of the National Academy of Sciences, 2011. **108**(47): p. 18932-18936.
253. Miller, C.R., et al., *Initiating a watch list for Ebola virus antibody escape mutations.* PeerJ, 2016. **4**: p. e1674.
254. Hatcher, E.L., et al., *Virus Variation Resource – improved response to emergent viral outbreaks.* Nucleic Acids Research, 2017. **45**(D1): p. D482-D490.
255. Brister, J.R., et al., *NCBI Viral Genomes Resource.* Nucleic Acids Research, 2015. **43**(D1): p. D571-D577.
256. Volchkov, V., et al., *Emergence of Subtype Zaire Ebola Virus in Gabon.* Virology, 1997. **232**(1): p. 139-144.
257. Hadfield, J., et al., *Nextstrain: real-time tracking of pathogen evolution.* Bioinformatics, 2018. **34**(23): p. 4121-4123.
258. Simon-Loriere, E., et al., *Distinct lineages of Ebola virus in Guinea during the 2014 West African epidemic.* Nature, 2015. **524**(7563): p. 102-104.
259. Ladner, T., Jason, et al., *Evolution and Spread of Ebola Virus in Liberia, 2014–2015.* Cell Host & Microbe, 2015. **18**(6): p. 659-669.
260. Gire, S.K., et al., *Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak.* Science, 2014. **345**(6202): p. 1369-1372.
261. Carroll, M.W., et al., *Temporal and spatial analysis of the 2014–2015 Ebola virus outbreak in West Africa.* Nature, 2015. **524**(7563): p. 97-101.
262. Durham, N.D., et al., *Real-Time Analysis of Individual Ebola Virus Glycoproteins Reveals Pre-Fusion, Entry-Relevant Conformational Dynamics.* Viruses, 2020. **12**(1).
263. Davidson, E., et al., *Mechanism of Binding to Ebola Virus Glycoprotein by the ZMapp, ZMAb, and MB-003 Cocktail Antibodies.* J Virol, 2015. **89**(21): p. 10982-92.
264. Spiliotopoulos, A., et al., *Discovery of peptide ligands targeting a specific ubiquitin-like domain-binding site in the deubiquitinase USP11.* J Biol Chem, 2019. **294**(2): p. 424-436.
265. Naqid, I.A., et al., *Mapping B-cell responses to Salmonella enterica serovars Typhimurium and Enteritidis in chickens for the discrimination of infected from vaccinated animals.* Scientific Reports, 2016. **6**(1): p. 31186.

266. Dowall, S.D., et al., *Development of a Cost-effective Ovine Polyclonal Antibody-Based Product, EBOTAb, to Treat Ebola Virus Infection.* J Infect Dis, 2016. **213**(7): p. 1124-33.
267. Lu, L.L., et al., *Beyond binding: antibody effector functions in infectious diseases.* Nature Reviews Immunology, 2018. **18**(1): p. 46-61.
268. Hessell, A.J., et al., *Fc receptor but not complement binding is important in antibody protection against HIV.* Nature, 2007. **449**(7158): p. 101-104.
269. Bhatt, P., et al., *Current Understanding of the Pathogenesis of Dengue Virus Infection.* Current Microbiology, 2021. **78**(1): p. 17-32.
270. Takada, A., et al., *Infectivity-Enhancing Antibodies to Ebola Virus Glycoprotein.* Journal of Virology, 2001. **75**(5): p. 2324-2330.
271. Takada, A., et al., *Epitopes Required for Antibody-Dependent Enhancement of Ebola Virus Infection.* The Journal of Infectious Diseases, 2007. **196**(s2): p. S347-S356.
272. Reusch, D. and M.L. Tejada, *Fc glycans of therapeutic antibodies as critical quality attributes.* Glycobiology, 2015. **25**(12): p. 1325-1334.
273. Zeitlin, L., et al., *Enhanced potency of a fucose-free monoclonal antibody being developed as an Ebola virus immunoprotectant.* Proceedings of the National Academy of Sciences, 2011. **108**(51): p. 20690-20694.
274. Cross, R.W., et al., *The Domestic Ferret (Mustela putorius furo) as a Lethal Infection Model for 3 Species of Ebolavirus.* J Infect Dis, 2016. **214**(4): p. 565-9.
275. Wong, G., et al., *Development and Characterization of a Guinea Pig-Adapted Sudan Virus.* J Virol, 2016. **90**(1): p. 392-9.
276. Alanine, D.G.W., et al., *Human Antibodies that Slow Erythrocyte Invasion Potentiate Malaria-Neutralizing Antibodies.* Cell, 2019. **178**(1): p. 216-228.e21.
277. Perez-Zsolt, D., et al., *Anti-Siglec-1 antibodies block Ebola viral uptake and decrease cytoplasmic viral entry.* Nat Microbiol, 2019. **4**(9): p. 1558-1570.
278. Keeffe, J.R., et al., *A Combination of Two Human Monoclonal Antibodies Prevents Zika Virus Escape Mutations in Non-human Primates.* Cell Rep, 2018. **25**(6): p. 1385-1394 e7.
279. ter Meulen, J., et al., *Human monoclonal antibody combination against SARS coronavirus: synergy and coverage of escape mutants.* PLoS Med, 2006. **3**(7): p. e237.
280. Bakker, A.B.H., et al., *Novel Human Monoclonal Antibody Combination Effectively Neutralizing Natural Rabies Virus Variants and Individual In Vitro Escape Mutants.* Journal of Virology, 2005. **79**(14): p. 9062-9068.
281. Gilchuk, P., et al., *Integrated pipeline for the accelerated discovery of antiviral antibody therapeutics.* Nature Biomedical Engineering, 2020. **4**(11): p. 1030-1043.
282. Piot, P., M.J. Soka, and J. Spencer, *Emergent threats: lessons learnt from Ebola.* International Health, 2019. **11**(5): p. 334-337.
283. Elgundi, Z., et al., *The state-of-play and future of antibody therapeutics.* Advanced Drug Delivery Reviews, 2017. **122**: p. 2-19.
284. Rghei, A.D., et al., *AAV Vectored Immunoprophylaxis for Filovirus Infections.* Tropical Medicine and Infectious Disease, 2020. **5**(4): p. 169.
285. Van Lieshout, L.P., et al., *Intramuscular Adeno-Associated Virus–Mediated Expression of Monoclonal Antibodies Provides 100% Protection Against Ebola Virus Infection in Mice.* The Journal of Infectious Diseases, 2018. **217**(6): p. 916-925.
286. Yoshida, R., et al., *Development of an Immunochromatography Assay (QuickNavi-Ebola) to Detect Multiple Species of Ebolaviruses.* J Infect Dis, 2016. **214**(suppl 3): p. S185-S191.
287. Crowe, J.E., *Principles of Broad and Potent Antiviral Human Antibodies: Insights for Vaccine Design.* Cell Host & Microbe, 2017. **22**(2): p. 193-206.
288. Matrosovich, M., et al., *Overexpression of the α-2,6-Sialyltransferase in MDCK Cells Increases Influenza Virus Sensitivity to Neuraminidase Inhibitors.* Journal of Virology, 2003. **77**(15): p. 8418-8425.