

STRUCTURAL BASIS FOR DISTINCTIVE RECEPTOR TROPISM OF EMERGING PARAMYXOVIRUSES

Alice Stelfox, Pembroke college, Trinity term 2019

A thesis submitted in partial fulfilment of the requirements for the degree of *Doctor of Philosophy*

ABSTRACT

A diverse array of viruses capable of infecting both humans and animals constitute the family *Paramyxoviridae* (Amarasinghe et al. 2019; Fields, Knipe, and Howley 2013). Concerns regarding the interspecies transmission potential paramyxoviruses harboured by wildlife reservoirs has motivated subsequent concerted viral surveillance efforts (Thibault et al. 2017; Drexler et al. 2012). Our understanding of many of these newly identified paramyxoviruses is primarily limited to genome information. However, paramyxovirus genomic variation (Holmes 2010; Kitchen, Shackelton, and Holmes 2011) often impedes our ability to predict the pathogenic potential based upon study of primary sequence alone. Therefore, investigations into paramyxovirus structure and function are vital to elucidating the potential biomedical and economic significance of emerging paramyxoviruses (Zeltina, Bowden, and Lee 2016).

The six-bladed β -propeller receptor binding protein (RBP), responsible for host-cell attachment and initiation of the merger of the viral and host-cell membranes (Fields, Knipe, and Howley 2013; Jardetzky and Lamb 2014), is therefore a central determinant of tissue and species tropism (Thibault et al. 2017; Zeltina, Bowden, and Lee 2016). Characterisation of the RBP, through structural and functional approaches, may enhance our ability to rationalise host range and interspecies transmission potential. Within this thesis, the RBPs of a number of recently classified paramyxoviruses are studied: the pararubulavirus, Sosuga virus (SosV); the narmoviruses, Mossman virus (MosV) and Nariva virus (NarV); the jeilongviruses, J virus (JPV) and Beilong virus (BeiV) (Amarasinghe et al. 2019).

In my investigations, I found that despite presentation of several stringently conserved residues associated with sialic acid activity, SosV-RBP lacks hemadsorption and neuraminidase functionality characteristic of closely related viruses, such as mumps virus (Kubota et al. 2016). Absence of activity is rationalised by discovery of structural rearrangements to key sialic acid interacting motifs, such as the triarginyl motif (Crennell et al. 2000), and physicochemical incompatibilities at the expected receptor binding site. Similarly, the jeilongviruses do not display hemadsorption or neuraminidase activity despite conservation of canonical sialic acid interacting motifs. Structural study of the jeilongviral RBPs, has enabled the characterisation of a C-terminal extension that forms a domain that occludes the putative sialic acid site. This provides evidence that jeilongviral RBPs employ an auto-inhibitory function analogous to the one utilised by the RBP of Newcastle disease virus Ulster strain (Yuan et al. 2012). Contrasting pararubulavirus and jeilongvirus RBPs, narmovirus RBPs display a β -propeller architecture that bears no motifs associated with known modes of paramyxovirus-receptor interactions. However, sequence conservation mapping reveals an extended footprint of elevated amino acid sequence identity on the surface of the protein. Efforts utilising site-directed mutagenesis are being implemented to explore the importance of this region in dictating cell interactions.

Through crystallographic and biochemical analysis, this thesis provides a structure-based rationale for understanding the pathobiological independence of the newly defined (Amarasinghe et al. 2019) *Pararubula*-, *Narmo*-, and *Jeilongviruses* from other genera within the *Paramyxoviridae*, where the RBPs of these viruses are likely incapable of utilising currently known modes of paramyxoviral host-cell recognition. Finally, this work provides a platform to identify the receptors utilised by these unique groups of pathogens.

ACKNOWLEDGEMENTS

There are a number of people without whom this work would have been impossible. Firstly, I would like to thank my supervisor, Thomas Bowden, who has provided great supervision and support throughout the course of my D.Phil. Secondly, I want to thank Ilona Rissanen, who introduced and guided me through the wet lab here in STRUBI, and continued to provide advice throughout my project. I am also grateful to Thomas, Ilona and David Stuart for proofreading this thesis, and for providing references for my upcoming postdoctoral position. For their supporting role in STRUBI and their invaluable experience, I am grateful to Margaret Jones, Oliver Hayward, Tom Walter, Weixian Lu and Karl Harlos. Additionally, the smooth running of STRUBI would not be possible without the work of Agata Krupa, Melissa Matthews, Claire Morris and Jun Dong. I'd also like to thank the members of the laboratory of Benhur Lee, for their collaboration, advice and discussions. Additionally, there are a number of others within STRUBI and Diamond who have provided instrumental scientific expertise relating specifically to this project: Kamel El-Omari, Rhys Pryce, Liz Allen, Max Renner and Antra Zeltina. Overall, my experience in STRUBI has been wonderful, and I must thank everyone for contributing to the collaborative atmosphere here.

Importantly, I need to thank a number of postdocs and fellow D.Phil. students in STRUBI, with whom I shared lunchtime chats and coffee break discussions. Here, whilst we spent much time having stimulating scientific discussions, we also used this time to vent, chat about anything but lab work and have fun. This includes STRUBI members both past and present, Rhys, Liz, Eamon, the Helens, Max, Rachel, Rob, Sam, Michael, Steinar, Serban, Vicky, Suzi, Vojta, Lindsay, Yasa, Mattia, Nayab, Joy, Christina, Christiane, Emily, Ben, Claire, Ruben, Ilona and Weng. I must additionally thank Robert Stass, Liz Allen and Vicky Avanzato for the many cat and dog pictures they have shared, these have been vital during the writing of this thesis.

Whilst not doing anything relating directly to the creation of this thesis, I must recognise the importance of the support I have had beyond the laboratory. It has been wonderful living with my three housemates for the past four years. Specifically, Rudy Bunel for his baking (which is excellent) and our chats about sourdough, Sam Watson for the ‘tinnies’, and Alban Desmaison for being a wonderful climbing partner (in addition to Liz, Eamon and Steinar). However, undoubtedly I would not have been able to make it through these final months without the support and generous opportunity to stay with my boyfriend, Tariq Hussain. My family have also been wonderful in supporting me throughout my D.Phil., but especially during the writing of this thesis. Ta everyone!

TABLE OF CONTENTS

CONTENTS

STRUCTURAL BASIS FOR DISTINCTIVE RECEPTOR TROPISM OF EMERGING

PARAMYXOVIRUSES	i
1 Introduction	2
1.1 Aims	2
1.2 The global burden of paramyxoviruses	3
1.3 Viral surveillance and reform of classification	7
1.4 Paramyxovirus molecular biology.....	9
1.4.1 Genome organisation and the paramyxoviral proteins	9
1.4.2 The paramyxovirus lifecycle	12
1.5 The envelope proteins of paramyxoviruses.....	13
1.5.1 Sialic acid mediated entry (and exit)	17
1.5.2 Protein-binding paramyxoviral receptor binding proteins	20
1.5.4 Trends in dimer angle organisation	29
2 Materials and Methods	34
2.1 Molecular Biology.....	34
2.1.1 cDNA templates	34
2.1.2 Cloning strategy and expression vectors	34
2.1.3 Primer specifications	38
2.1.4 Polymerase chain reaction (PCR).....	38
2.1.5 Agarose gel electrophoresis.....	43
2.1.6 Restriction enzyme digestion	43
2.1.7 DNA ligation.....	43
2.1.8 In-Fusion®	43
2.1.9 Transformation	44
2.1.10 Plasmid amplification and purification	44
2.1.11 Mutagenesis using splicing by overlap PCR.....	45
2.2 Mammalian cell culture.....	48
2.2.1 Maintenance of cell lines.....	48
2.2.2 Storage and revival of cell lines	49

2.3	Recombinant protein expression and purification.....	49
2.3.1	Rationale and workflow	49
2.3.2	Small-scale transient transfection.....	50
2.3.3	SDS-PAGE.....	51
2.3.4	Western blot analysis	51
2.3.5	Large-scale transient transfection.....	53
2.3.6	Production of stable cell lines	54
2.3.7	Large-scale stable cell line culture	56
2.3.8	Cell harvesting and diafiltration	57
2.3.9	Protein purification.....	57
2.3.10	Affinity chromatography	57
2.3.11	Enzymatic deglycosylation.....	58
2.3.12	3C cleavage	59
2.3.13	Size-exclusion chromatography	59
2.3.14	Protein concentration determination	60
2.4	Protein crystallisation.....	60
2.5	Structure determination	61
2.5.1	Crystal harvesting and cryo-protection	61
2.5.2	X-ray data collection and processing	61
2.5.3	Phase determination	62
2.5.4	Model building and refinement	63
2.6	Structure analysis	64
2.6.1	Molecular visualisation and presentation	64
2.6.2	Structural phylogenetic analysis.....	64
2.7	Biophysical characterisation	64
2.7.1	High performance liquid chromatography small-angle X-ray scattering.....	64
2.8	Sialic-acid interaction analysis.....	65
2.8.1	Hemadsorption and neuraminidase activity analysis	65
2.8.2	Warren test	66
2.9	Novel receptor identification.....	68
2.9.1	Glycan array analysis	68
2.9.2	Pull-down experiments.....	69
2.9.3	SosV-RBP binding analysis for potential haploid screen.....	71
3	The sosuga virus receptor binding protein	74
3.1	Summary of findings.....	74

3.2	Introduction	74
3.3	Protein production and structure determination of SosV-RBP	77
3.3.1	Construct design and small-scale expression testing.....	77
3.3.2	Large-scale expression and purification.....	78
3.3.3	Crystallisation and X-ray diffraction data collection	80
3.3.4	Structure determination	82
3.4	Structure analysis	82
3.4.1	Structure overview	82
3.4.2	Comparison of SosV-RBP to characterised paramyxoviral RBPs	83
3.5	Assessment of sialic acid interactions.....	92
3.6	Discussion	95
3.7	Conclusion.....	97
4	The Mossman virus and Nariva virus receptor binding proteins	99
4.1	Summary of findings.....	99
4.2	Introduction.....	99
4.3	Protein production and structure determination of MosV-RBP and NarV-RBP.....	101
4.3.1	Construct design and small-scale expression testing.....	101
4.3.2	Large-scale expression and purification of MosV-RBP and NarV-RBP	104
4.3.3	Crystallisation and X-ray diffraction data collection	106
4.3.4	Structure determination	107
4.4	Structure analysis	110
4.4.1	Structure overview of the MosV and NarV RBPs.....	110
4.4.2	Small-angle X-ray scattering analysis.....	112
4.4.3	Comparison of MosV-RBP and NarV-RBP to different paramyxoviral RBPs.....	115
4.5	Assessment of MosV-RBP and NarV-RBP sialic acid interactions.....	121
4.6	Comparisons within the Narmovirus genus	122
4.6.1	Production of mutant MosV-RBPs.....	124
4.6.2	Comparison of MosV-RBP and NarV-RBP to the recently discovered bank vole virus RBP	126
4.7	Discussion	126
4.8	Conclusions	128
5	The J paramyxovirus and Beilong virus receptor binding proteins.....	131
5.1	Summary of findings.....	131
5.2	Introduction	132
5.3	Protein production and structure determination of BeiV-RBP and JPV-RBP	141

5.3.1	Construct design and small-scale expression testing.....	141
5.3.2	Large-scale expression and purification.....	142
5.3.3	Crystallisation and X-ray diffraction data collection	146
5.3.4	Structure determination	149
5.4	Structure analysis	149
5.4.1	Structure overview of the BeiV and JPV RBPs	149
5.4.2	Comparison of BeiV-RBP and JPV-RBP to different paramyxoviral RBPs	153
5.4.3	The extended C-terminus	158
5.5	Assessment of JPV-RBP and BeiV(HMC and rat)-RBP sialic acid associated activity	163
5.5.1	Design and small-scale screening of truncated JPV and BeiV RBP constructs	164
5.6	Discussion	166
5.7	Conclusion.....	169
6	Future perspectives and concluding remarks	171
6.1	Future perspectives.....	171
6.1.1	Novel receptor identification strategies.....	171
6.1.2	Characterisation of the determinants of receptor interactions	183
6.2	The paramyxoviral receptor binding proteins in context with the sialidase superfamily	188
6.3	Concluding remarks	189
Appendix A: List of publications		191
Appendix B: Bibliography		192
Appendix C: List of abbreviations		212
Appendix D: List of genbank accession numbers		215

CHAPTER 1

INTRODUCTION

1 INTRODUCTION

1.1 AIMS

The capability of a virus to undergo interspecies transmission and zoonotic transfer, the process by which pathogenic agents are transmitted between animals and humans, is contingent on the interactions and dynamics between reservoir species, domesticated species and humans (Morse et al. 2012). Drivers leading to disease emergence include the ecological, environmental and socio-economic factors that cause altered patterns in human and animal interactions (Daszak, Cunningham, and Hyatt 2000), potentially contributing to the statistically significant increase in outbreaks of emerging infectious disease (Jones et al. 2008; Smith et al. 2014). However, the pathobiological features of a virus also define its ability to spread into new animal populations, with the molecular determinants of virus-host-cell recognition constituting a key determinant of host range (Rosenberg 2015; Cunningham, Daszak, and Wood 2017; Zeltina, Bowden, and Lee 2016).

While virus identification initiatives have expanded our understanding of paramyxoviral diversity, the large genetic variation revealed by these efforts hinders our ability to grasp the distinct characteristics of these potentially dangerous viruses. This thesis is predicated on the supposition that an understanding of the true significance of these viruses can only come with a knowledge of structure and function. Here, I focus on the receptor binding protein (RBP) presented on the surface of emerging and recently identified paramyxoviruses. The RBP is a key tropism determinant, and as such, contributes to the ability of a viruses to jump the species barrier. Through the integration of crystallographic and biochemical analysis, I report functional insights into the RBPs of viruses from three newly defined genera, *Narmovirus*, *Jeilongvirus* and *Pararubulavirus* (Amarasinghe et al. 2019). This work provides a rationale for the divergent properties of these viruses and identifies potential novel receptor binding sites, paving the way for future receptor identification and validation studies and rationalising the diverse pathobiological features of the *Paramyxoviridae* family, more broadly.

1.2 THE GLOBAL BURDEN OF PARAMYXOVIRUSES

Viruses of the *Paramyxoviridae* family have a broad impact on the world through their extensive interactions with human and animal populations. Paramyxoviruses cause widespread and ubiquitous disease in many volant, terrestrial, and aquatic animals, with their role as pathogens of domestic and wild animals presenting a major economic and ecological burden (Roeder 2011; Brown and Bevins 2017; Kumar et al. 2014; Ellis 2010; Sutummaporn et al. 2019; Rendon-Marin et al. 2019; Glennon et al. 2018; Liu et al. 2015; Duignan et al. 2014; Van Bresseem et al. 2014; Vizoso and Thomas 1981). Members of this family are also causative agents of disease in humans, including the incredibly infectious measles virus (MeV) (Griffin, Lin, and Pan 2012; Fields, Knipe, and Howley 2013; Xu et al. 2008), the highly prevalent parainfluenza viruses (Branche and Falsey 2016; Connaris et al. 2002), and the highly pathogenic Nipah and Hendra viruses (NiV and HeV, respectively) (Marsh and Wang 2012; Eaton et al. 2006).

Viral surveillance has led to a continued increase in the number of known members of the *Paramyxoviridae* family (Amarasinghe et al. 2019). Prior to the emergence of the henipaviruses in the 1990s (Murray et al. 1995; Chua et al. 2000; Iorio et al. 2001), the *Paramyxoviridae* family was typified by morbilliviruses, sialic acid binding paramyxoviruses, and pneumoviruses (Enders 1996). Human infection with the archetypal morbillivirus MeV is characterised by fever, cough, conjunctivitis, and subsequent development of a maculopapular rash (Bester 2016). Despite development of an effective vaccine, MeV continues to be a major cause of mortality of children in developing nations (Kaufmann 2019). The hallmarks of human infection with mumps virus (MuV) are parotitis and orchitis, though central nervous complications, such as virus-induced meningitis and encephalitis, may sometimes lead to fatality (Rubin et al. 2015). The parainfluenza viruses (PIVs), initially named due to the influenza-like illness that results from infection, are often the causative agents of croup, pneumonia and bronchiolitis in children (Schomacker et al. 2012). The parainfluenza related, avian pathogenic, Newcastle disease virus (NDV) was discovered following outbreaks of deadly disease in poultry flocks in Indonesia and the UK (Fields, Knipe, and Howley 2013).

As demonstrated by the agriculturally relevant NDV, paramyxoviruses are important aetiological agents of animal disease. For example, whilst MeV host range is restricted to primates (Laksono et al. 2016), the related *Morbillivirus*, canine distemper virus (CDV) demonstrates a broad animal host range and varying degrees of pathogenicity in hosts, with dogs, lions and ferrets all exhibiting neurologic complications (Rendon-Marin et al. 2019). At the ecological level, occurrence of CDV in lions, the related phocid distemper virus (PDV) and cetacean morbillivirus (CeMV) within marine mammals, often result in outbreaks with devastating consequences to populations (Van Bresse et al. 2014; Duignan et al. 2014; Beineke, Baumgärtner, and Wohlsein 2015). These impacts are reflected agriculturally, for example, peste des petits ruminants (PPRV) morbillivirus infection in ruminant species is manifested by erosive stomatitis, pneumonitis, diarrhoea, fever and lymphoid cell depletion. PPRV is also an economically relevant pathogen of ruminant species, causing an estimated loss of \$1.45–\$2.1 billion (USD) per year (OIE 2015; Fields, Knipe, and Howley 2013; Kumar et al. 2014; Baron et al. 2017).

Many paramyxovirus can be readily transmitted between species and are responsible for emerging diseases in animals and humans (Thibault et al. 2017). Indeed, due to their low fidelity polymerase, paramyxoviruses may readily acquire mutations that allow host-switching (Holmes 2010; Kitchen, Shackelton, and Holmes 2011). This potential for paramyxoviruses to precipitate devastating zoonotic events, where disease transmits to humans from animal reservoirs, was demonstrated by the emergence of henipaviruses (Marsh and Wang 2012). In 1994, HeV was identified as the aetiological agent of acute respiratory disease in horses in Australia. Subsequent transmission of HeV to humans resulted in high associated fatality rates (Murray et al. 1995; Marsh and Wang 2012). Similarly, NiV infection of humans and pigs caused a devastating outbreak event throughout Malaysia in 1998–1999, where 1 million pigs were culled and of the 265 humans that contracted disease, 40% succumbed to the disease (Chua et al. 2000; Mohd Nor, Gan, and Ong 2000). NiV continues to emerge in South Asia, with case fatality rates of up to 100% (Marsh and Wang 2012). Both HeV and NiV cause severe respiratory disease and encephalitis (Wong and Ong 2011). Domesticated species, such as pigs and horses, are intermediate hosts for the henipaviruses (Glennon

et al. 2018; AbuBakar et al. 2004), with the pteropid species of bats being the likely host reservoir (Chua, Lek Koh, et al. 2002). Given the potential to cause a public health emergency and the current lack of effective drugs and vaccines, NiV has been listed for research and development prioritisation by the World Health Organization (WHO).

One factor defining the anthropological and ecological barriers to zoonosis is the geographic range of the reservoir host (Daszak, Cunningham, and Hyatt 2000; Drexler et al. 2012). For example, bats (*Chiroptera*) constitute a large and diverse number of mammalian species, and have a great spatial range across all continents excluding Antarctica (Han et al. 2015; Wibbelt et al. 2010). Recent studies have identified a number of novel paramyxoviruses in bat and rodent (*Rodentia*) species worldwide (Drexler et al. 2012). Similar to species of the order *Chiroptera*, rodents also have high species diversity, they exist in large, social, and population dense groups, and share habitats with numerous other bat and rodent species. These behaviours promote a high prevalence and diversity of viruses within these species (Thibault et al. 2017). Whilst such reservoir species enable dissemination of viruses, the infection of domestic species with potentially zoonotic paramyxoviruses can facilitate transmission of a virus from natural reservoirs, due to their close association with both human and wild-life populations (Glennon et al. 2018), as exemplified in the HeV and NiV outbreaks (Han et al. 2015). Beyond sporadic emergence of paramyxoviruses, such zoonotic events have been suggested to cause the establishment of MeV as a human pathogen. Rinderpest virus (RPV), an eradicated virus restricted to even-toed ungulates (Roeder 2011) has been postulated to represent the basal group of the *Morbillivirus* genus, and therefore is hypothesised to be a precursor to MeV (Fields, Knipe, and Howley 2013; Pearce-Duvet 2006). However, owing to limited evidence, the origin of MeV from a pathogen maintained in domestic or wild animals remains unclear (Pearce-Duvet 2006). Supportive of the hypothesis that morbilliviruses are capable of emerging into human populations, recent gain-of-function (GOF) and structural studies on CDV (Bieringer et al. 2013) and PPRV (Abdullah et al. 2018), respectively, identified the minimal molecular alterations to the RBP required for zoonotic transfer to be possible.

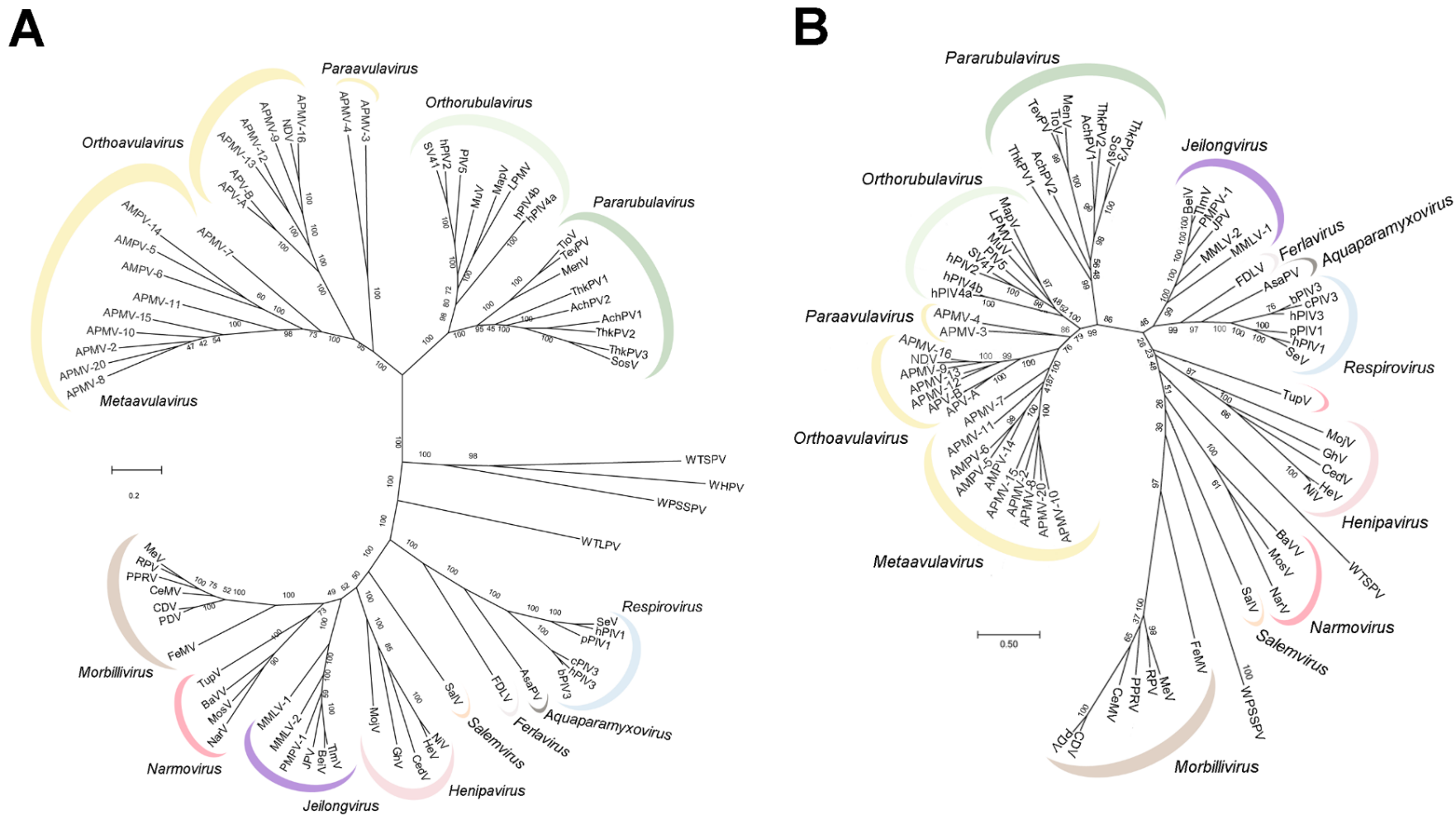


Figure 1.2: Maximum likelihood phylogenetic trees of (a) the polymerase (L) and (b) RBP protein sequences using MEGA-X (Molecular Evolutionary Genetic Analysis Version X 10.1). Genera are coloured and annotated with classification defined by the International Committee on Taxonomy of Viruses. Scale bars describe amino acid substitutions per site. Bootstrap values (100 replicates) annotated at nodes. Genbank accession numbers listed in Appendix D.

1.3 VIRAL SURVEILLANCE AND REFORM OF CLASSIFICATION

Increased viral surveillance efforts have aided the prediction of virus emergence, and have revealed an unprecedented diversity of paramyxoviruses worldwide (Thibault et al. 2017). Indeed, the increased number of known paramyxoviruses has led to redefinition of paramyxovirus taxonomy, for example the pneumoviruses, such as respiratory syncytial virus (RSV), have recently been classified into a distinct family, termed *Pneumoviridae* (Afonso et al. 2016). These efforts have been driven by a combination of factors, including improved molecular diagnostic and increased interest and funding (Lipkin 2013). A recent study identified 66 novel paramyxovirus species in rodent and bat populations across Africa, including a number of viruses which do not classify into established genera (Drexler et al. 2012). Further small-scale sporadic studies revealed unique paramyxoviruses in animal populations worldwide (Mortlock et al. 2019; Wu, Yang, et al. 2014; Vanmechelen, Bletsa, Laenen, Lopes, Vergote, Beller, Deboutte, Korva, Zupanc, et al. 2018; Jack et al. 2005; Li et al. 2006a; Woo et al. 2016a; Woo, Lau, Wong, Fan, et al. 2012; Lambeth et al. 2009; Miller et al. 2003; Anderson et al. 2012; Chua, Wang, et al. 2002; Bowden et al. 2001; Burroughs et al. 2015; Alkhovsky et al. 2018; Tidona et al. 1999; Woo et al. 2011a; Baker et al. 2013).

The identification of so many novel paramyxoviruses has precipitated recent restructuring of the *Paramyxoviridae* taxonomy (Amarasinghe et al. 2019; Rima et al. 2018), where the *Orthoparamyxovirinae* subfamily has been created and is defined by five existing genera (*Aquaparamyxovirus*, *Ferlavirus*, *Henipavirus*, *Morbillivirus* and *Respirovirus*) and three newly described genera (*Narmovirus*, *Jeilongvirus* and *Salemvirus*) (Fig. 1.2). The subfamily *Rubulavirinae* was created to account for the recently discovered distinct “rubula-like” viruses (Thibault et al. 2017), now termed *Pararubulavirus*, with established rubulaviruses being classified within the new *Orthorubulavirus* genus. In my thesis, I focus on emerging and newly classified paramyxoviruses belonging to the genera: *Narmovirus*, *Jeilongvirus*, and *Pararubulavirus*. These genera represent a growing, uncharacterised fraction of the paramyxovirus family, with potential to impact human and animal populations.

Narmoviruses were initially classified as ‘morbilli-like’ viruses due to sequence homology in the P protein, however, subsequent analysis has supported the re-categorisation of these viruses as a discrete genus (Rima et al. 2018). The *Narmovirus* genus is comprised of four rodent and treeshrew originating viruses (*Rodentia* and *Scandentia*) isolated in disparate locations worldwide. The type species is Nariva virus (NarV) and was isolated from short-tailed cane mouse (*Zygodontomys brevicauda*) trapped in the Nariva swamp, Trinidad, Trinidad and Tobago (Lambeth et al. 2009). Other members of the *Narmovirus* genus, include Mossman virus (MosV), which was isolated from *Rattus* sp. in Australia (Miller et al. 2003), bank vole virus 1 (BavV), which was detected in bank voles (*Myodes glareolus*) in Russia (Alkhovsky et al. 2018), and Tupaia paramyxovirus (TupV), which was isolated from northern treeshrews (*Tupaia belangeri*) captured in Thailand (Tidona et al. 1999).

The viruses of the similarly rodent-borne jeilongviruses have the largest genomes (~19 kb) of all viruses within the order, *Mononegavirales* (Fields, Knipe, and Howley 2013). The first jeilongvirus isolated was J paramyxovirus (JPV) from house mice (*Mus musculus*) presenting extensive haemorrhagic lung lesions in Australia (Jack et al. 2005). Beilong virus (BeiV) has been listed as the archetypal species by the International Committee on Taxonomy of Viruses (ICTV) (Amarasinghe et al. 2019), and was discovered in a human kidney mesangial cell line (Li et al. 2006a) and subsequently detected in black and brown rats (*R. rattus* and *R. norvegicus*, respectively) populations in China (Woo et al. 2016a). Other members of the *Jeilongvirus* genus include, Mount Mabu lophuromys viruses 1 and 2 (MMLV-1/2), which was identified in Rungwe brush furred rats (*Lophuromys machangui*) in Mozambique, Pohorje Myodes paramyxovirus 1 (PMPV-1), which was detected in bank voles (*Myodes glareolus*) in Slovenia (Vanmechelen, Bletsa, Laenen, Lopes, Vergote, Beller, Deboutte, Korva, Zupanc, et al. 2018), and Tailam virus (TaiV), which was discovered in Sikkim rats (*R. andamanensis*) in Hong Kong (Woo et al. 2011a).

Due to genetic similarities, many pararubulaviruses have been described to be “rubula-like” (Thibault et al. 2017). However, the recently identified structural and pathobiological differences

between pararubulaviruses and other rubulaviruses (e.g. MuV), has motivated the classification of pararubulaviruses into a distinct genus, *Pararubulavirus* (Amarasinghe et al. 2019). Menangle virus (MenV), was isolated in Australia in the 1990s and found to cause reproductive disease in pigs and severe influenza-like illness in humans (Chant et al. 1998; Philbey et al. 1998). Similarly, Sosuga virus (SosV), another recently classified member of the *Pararubulavirus* genus, infected a wildlife biologist working in South Sudan and Uganda, and is thought to be responsible for severe febrile illness (Albarino et al. 2014). Both MenV (*Pteropus Alecto*) and SosV (*Rousettus aegyptiacus*) are thought to utilise fruit bats as a reservoir (Barr et al. 2012; Amman et al. 2015). Further pararubulaviruses of unknown interspecies transmission potential and pathogenicity are related to these two viruses: Achimota viruses 1 and 2 (AchPV-1/2) were discovered in urine collected from straw-coloured bat (*Eidolon helvum*) urban roosts in Ghana (Baker et al. 2013), Teviot virus (TevPV) was isolated from pteropid bat (*P. alecto* and *P. poliocephalus*) urine samples in Australia (Burroughs et al. 2015), Tioman virus (TioV) was isolated from pteropid bat (*P. hypomelanus*) urine in Malaysia (Chua, Wang, et al. 2002), and Tuhoko viruses 1, 2 and 3 (ThkPV-1/2/3) were discovered in Leschenaults rousette (*Rousettus leschenaultii*) bats in China (Lau et al. 2010).

Our understanding of these viruses is mostly limited to the primary sequence information, and given the low fidelity of the paramyxovirus polymerase and high mutation acquisition rates across their genomes (Holmes 2010), predictions of virus zoonosis risk and pathobiology are often impeded. Conversely, virus function and structure remain much more conserved and provide a means to comparatively analyse new paramyxoviruses (Zeltina, Bowden, and Lee 2016).

1.4 PARAMYXOVIRUS MOLECULAR BIOLOGY

1.4.1 Genome organisation and the paramyxoviral proteins

Viruses within the family *Paramyxoviridae* have negative-sense single-stranded RNA genomes, typically 15–21 kb in length (Vanmechelen, Bletsa, Laenen, Lopes, Vergote, Beller, Deboutte, Korva, Zupanc, et al. 2018), encoding for at least six structural genes (Fields, Knipe, and Howley

2013). The nucleoprotein (N) helically encapsidates the RNA genome and associates with the phosphoprotein (P) and RNA-dependent RNA-polymerase (L) for RNA replication and transcription (Noton and Fearn 2015). The matrix protein (M) is involved in maintaining virion architecture beneath the viral envelope (Battisti et al. 2012), which is in turn decorated with the fusion protein (F), and RBP, both responsible for driving virus entry (Harrison, Sakaguchi, and Schmitt 2010). Depending on functionality contingent on cognate receptor binding, the RBP has been variously termed the hemagglutinin-neuraminidase (HN), the hemagglutinin (H), or attachment glycoprotein (G). Several paramyxoviruses also variously encode accessory proteins, including the small-hydrophobic protein (SH) (Abraham et al. 2018; Wilson et al. 2006), the transmembrane protein (TM) (Li et al. 2015), and the V, W and C proteins produced from mRNA editing of the P gene or initiation of an alternative reading frame (He et al. 2002; Poole et al. 2002; Shaw et al. 2005; Yamaguchi et al. 2014; Kolakofsky 2016; Uchida et al. 2018). Recently identified jeilongviruses have an additional open reading frame (ORF) encoded after the RBP, currently termed 'X' (Jack et al. 2005; Li et al. 2006; Woo et al. 2016). Paramyxovirus genes are transcribed sequentially from the 3' end of the genome initiating and terminating at the gene junctions between the viral genes (Noton and Fearn 2015). This is performed in the order 3'-N-P/(V)/(W)/(C)-M-F-(SH)- (TM)-RBP-(X)-L-5', resulting in an mRNA gradient, that correlates with the required abundance of the proteins in infected cells (Wignall-Fleming et al. 2019).

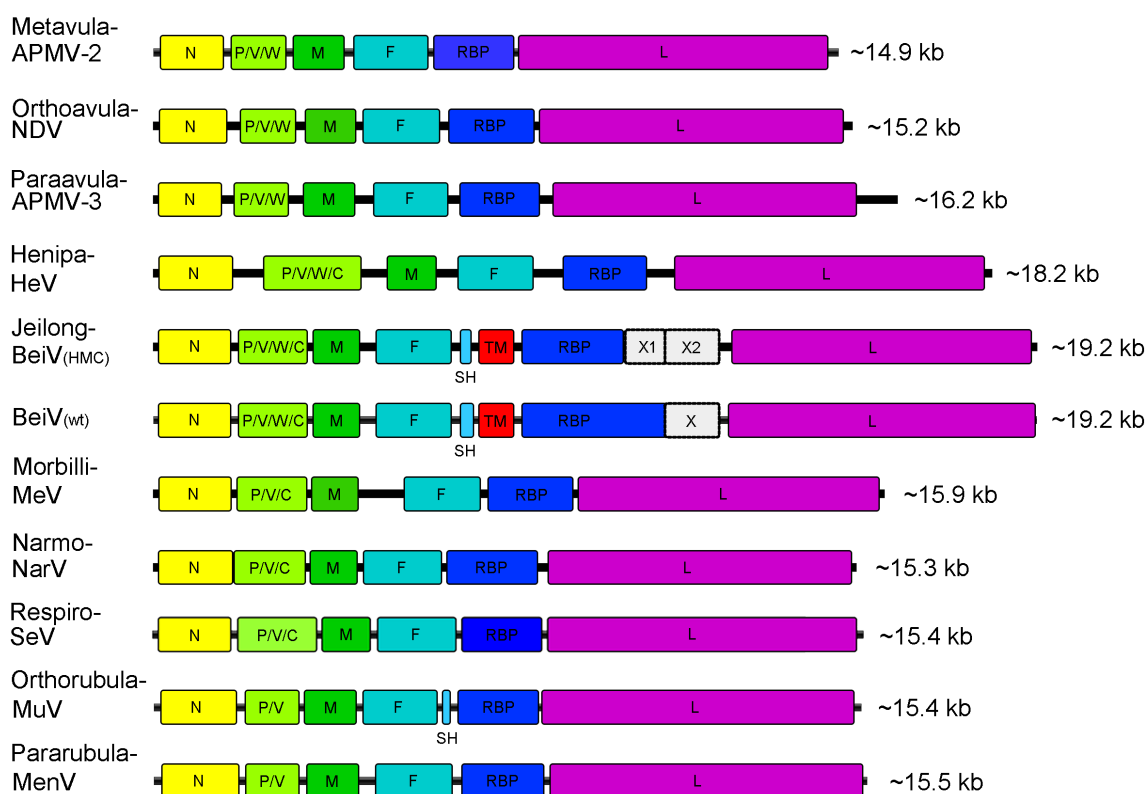


Figure 1.4.1: Genome organisation of representative species from the genera defined by the recent Mononegavirales taxonomy update (Amarasinghe et al. 2019). The representative species are as follows: *Metavulavirus*, avian paramyxovirus 2 (APMV-2); *Orthoavulavirus*, NDV; *Paraavulavirus*, APMV-3; *Henipavirus*, HeV; *Jeilongvirus*, beilong virus (BeiV) isolated from human mesangial cells (HMCs) and wildtype (wt) isolated from rats, *Morbillivirus*, MeV; *Narmovirus*, Nariva virus (NarV); *Respirovirus*, Sendai virus (SeV); *Orthorubulavirus*, MuV; *Pararubulavirus*, MenV. The standard genes N (yellow), P (green), M (dark green), F (teal), RBP (blue), L (purple), are encoded by all viruses. The accessory proteins C, V and W, are produced through mRNA editing or initiation from an alternative start codon of the P gene. The SH (light blue) gene is encoded by the Jeilong viruses and Rubulaviruses and the TM (red) gene only by the Jeilongviruses. The open reading frame (ORF) X (dashed grey) is considered part of the RBP of Jeilongviruses. Genbank accession codes are listed in Appendix D.

1.4.2 The paramyxovirus lifecycle

As access to the host-cell cytoplasm is required for paramyxovirus replication, paramyxoviruses utilise the combined functions of the RBP and F to bring about the merger of the host and viral membranes (Aguilar et al. 2016). This process is initiated through binding of the RBP to a cognate host-cell receptor, which regulates the triggering of the type I fusion protein, F. Fusion enables release of the viral ribonucleoprotein (vRNP), which is composed of genomic negative-sense RNA encapsidated by N protein and associated with the L and P proteins (Alayyoubi et al. 2015). Upon genome release into the host-cell, the viral polymerase machinery initiates transcription *de novo*, for production of protein required for subsequent rounds of replication and virus assembly. Replication proceeds through creation of antigenome templates for negative-sense genomic RNA production, which is subsequently encapsidated with N, and associated with L and P to form the vRNP (Noton and Fearn 2015). For viral particle assembly, the vRNP is transported to the membrane by the M protein, where M interacts with the surface expressed F and RBP and drives formation of the virion, and subsequent scission (Harrison, Sakaguchi, and Schmitt 2010; Fields, Knipe, and Howley 2013). In some paramyxoviruses, the RBP has dual function as a receptor destroyer (Crennell et al. 2000), so as to prevent reinfection and promote successful dissemination of progeny.

During the paramyxoviral lifecycle the different proteins likely have a varying contribution to dictating cell and species tropism and resulting pathogenicity. For example, the V protein of NiV was found to be a key determinant of pathogenesis in animal infection models (Yoneda et al. 2010; Satterfield et al. 2015), due to its role in suppression of RIG-I and MDA-5-dependent interferon signalling (Uchida et al. 2018). The RBP, on the other hand, is utilised to overcome the first molecular barrier to host-cell infection and interspecies transmission, as this is the initial point of contact within the host and therefore a major determinant of tropism (Thibault et al. 2017). Due to the plasticity of the paramyxoviral RBP, paramyxoviruses display a considerable capability to adapt to a diversity of receptors (Zeltina, Bowden, and Lee 2016; Hadac et al. 2004). To understand this propensity to bind a diversity of receptors, it is necessary to explore the structure of the RBP and its context within the organisation of the paramyxoviral envelope.

1.5 THE ENVELOPE PROTEINS OF PARAMYXOVIRUSES

All paramyxoviruses possess an RBP and F protein that decorate the surface of the pleomorphic paramyxovirus virion (Harrison, Sakaguchi, and Schmitt 2010). These proteins act in concert to bring about the pH-independent fusion of the viral and cellular membranes (Chang and Dutch 2012). A third protein, the SH protein, is an integral membrane protein found in viruses of the genera *Othorubulavirus* (Wilson et al. 2006) and *Jeilongvirus* (Li et al. 2006a; Jack et al. 2005) and plays a role in virulence and apoptosis inhibition, with mechanisms linked to modulation of tumour necrosis factor (TNF)- α production in the case of JPV (Abraham et al. 2018).

The RBP is a type II integral membrane protein consisting of a short N-terminal intraviral region, a transmembrane domain, α -helical stalk region and C-terminal six-bladed β -propeller head domain. Structural, biochemical and biophysical studies revealed that the RBP assembles as tetramers on the virion surface, where each tetramer is composed of a dimer-of-dimers, which form through a non-covalent dimeric interactions between head domains, and disulphide bonds between the helical stalk domains (Yuan et al. 2011; Welch et al. 2013) (**Fig. 1.5A**). The β -propeller head domain is implicated in receptor binding and is arranged as six β -sheets comprised of four antiparallel strands organised around a central axis (Jardetzky and Lamb 2014; Aguilar et al. 2016). Viruses of *Jeilongvirus* genus encode a lengthened C-terminus that extends beyond the predicted six-bladed β -propeller head domain and exhibits no structural or functional homology with characterised paramyxoviral proteins. RBPs from this genus range from 709–1,589 amino acids in length (JPV-RBP and PMPV-1-RBP, respectively), as opposed to the typical ~550 amino acids in length for typical paramyxoviral RBPs (Vanmechelen, Bletsa, Laenen, Lopes, Vergote, Beller, Deboutte, Korva, Zupanc, et al. 2018; Jack et al. 2005; Li et al. 2006a; Woo et al. 2016a). Whilst the overall fold of the RBP remains conserved throughout the paramyxovirus family, the primary sequence is highly variable, enabling the RBPs adapt to a diverse range of host-cell receptors (Zeltina, Bowden, and Lee 2016; Bowden, Crispin, Jones, et al. 2010).

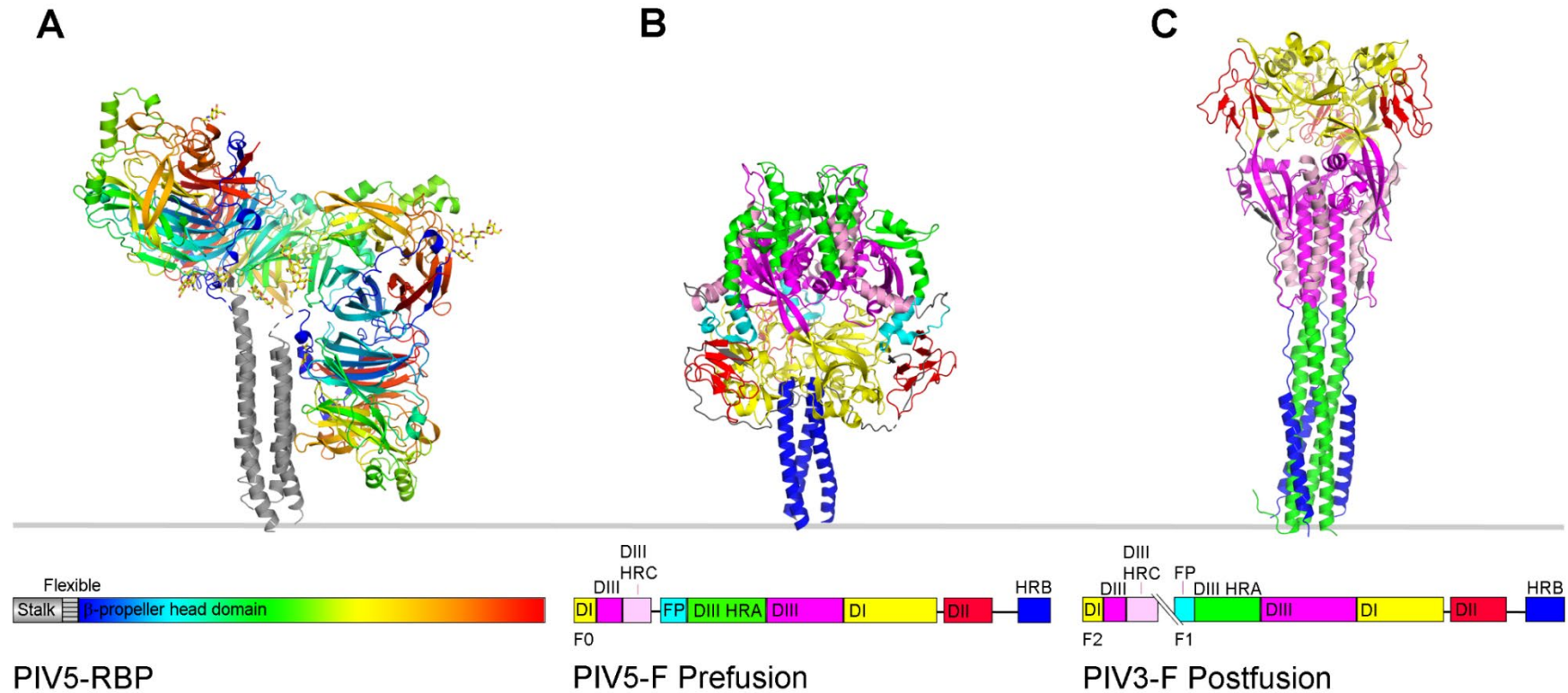


Figure 1.5: The RBP and fusion glycoproteins are presented on the virion surface. (a) Tetrameric 'heads-up' and 'heads-down' orientated PIV5-RBP ectodomain (PDB ID: 4JF7) shown in cartoon representation, with the six-bladed β -propeller coloured rainbow from N- to C-terminus (blue to red), and the stalk coloured grey. The gene diagram is similarly coloured. (b) The trimeric prefusion PIV5-F ectodomain (PDB ID: 2B9B) displayed in cartoon representation, with domain I (DI) coloured yellow, domain II (DII) coloured red, domain III (DIII) coloured magenta, heptad repeat A (HRA) coloured green, heptad repeat B (HRB) coloured blue, heptad repeat C (HRC) coloured light pink, and the fusion peptide (FP) coloured cyan. Both HRA and HRC are also contained within DIII. The uncleaved F₀ gene diagram is coloured as in the structure. (c) The hexameric postfusion PIV3-F ectodomain (1ZTM) and gene diagram coloured as in 1.5B, with cleavage location located between F₂ and F₁, and represented as two diagonal lines.

The F protein activated upon receptor attachment, directs the merger of the host and cellular membranes (Jardetzky and Lamb 2014). It is a type I integral membrane protein and a class I fusion protein, and like all proteins of this class, is presented on the virion surface as a prefusion homotrimer (**Fig 1.5B**) (Yin et al. 2006; Welch et al. 2012; Xu et al. 2015; Hashiguchi et al. 2018), which forms a six-helix bundle postfusion structure following activation (**Fig. 1.5C**) (Yin et al. 2005). The F protein is expressed as a fusion incompetent precursor (F0) and is activated by a cellular protease (e.g. cathepsin-L for NiV, subtilisin-like proteases for virulent strains of NDV), which cleaves F0 into the disulphide-linked F1 and F2 chains during transport to the cell surface (Chang and Dutch 2012; Pager et al. 2006; Fujii et al. 1999). Similar to the RBP, the F protein encodes a short cytoplasmic tail and a TM domain as a membrane anchor, however these features are located at the C-terminus. Structural analysis of PIV5, MeV and NiV F proteins reveal that in the prefusion state, domains I, II and III (DI, DII and DIII, respectively) of the typical F protein homotrimer form a globular head through extensive interactions, and are connected to a three-helix coiled-coil stalk consisting of heptad repeat B (HRB) projecting from the viral membrane (Xu et al. 2015; Welch et al. 2012; Yin et al. 2006; Hashiguchi et al. 2018). The well-conserved fusion peptide remains buried between DIII of the same monomer and DII of an adjacent monomer in the prefusion form. The PIV3-F structure shows that in the postfusion form, the F protein forms a hexamer and undergoes a conformational change to form an elongated structure with a compact head. These conformational changes include restructuring of DIII, enabling projection of the now extended α -helical heptad repeat A (HRA), resulting in the formation of a six-helix coiled-coil (Yin et al. 2005; Aguilar et al. 2016). In the context of the two membranes, these irreversible change of the F protein from the prefusion state to the stable postfusion state results in the merger of the virus and host membrane and the creation of a pore enabling entry of the viral genomic material (Chang and Dutch 2012).

Whilst several models of F activation by the RBP currently exist, data indicate a primary role for the stalk region of the RBP in triggering the F protein (Jardetzky and Lamb 2014; Aguilar et al. 2016). Arrangements of tetrameric RBPs have been observed in both ‘heads-down’ (Yuan et al.

2011) and ‘heads-up’ (Yuan et al. 2005; Welch et al. 2013) orientations of the receptor binding β -propeller head (**Fig. 1.5A**), with the ‘heads-up’ conformation consistent with receptor binding, revealing the stalk regions able to interact with the F protein (Jardetzky and Lamb 2014). Studies across different paramyxoviral species have confirmed that expression of the RBP stalk region alone can induce fusion driven by the F protein (Liu et al. 2013; Bose et al. 2012; Brindley et al. 2013), and limitation of the flexibility of the loop between the stalk and head domains using proline substitutions perturbs fusion activation (Adu-Gyamfi et al. 2016). Furthermore, as the RBP head domain is not subject to functional constraints associated with fusion triggering, it is more structurally plastic and able to accommodate the diverse assortment of receptors known to allow paramyxovirus entry into a host-cell (Jardetzky and Lamb 2014; Porotto et al. 2012b; Hadac et al. 2004; Bowden, Crispin, Jones, et al. 2010). Demonstrating this structural plasticity, engineering efforts have generated fusion competent vaccine strains of MeV that retargeted towards epidermal growth factor receptor (EGFR), for potential use in virus oncolytic therapy (Paraskevskou et al. 2007; Hadac et al. 2004).

The intricacies of the attachment and fusion molecular choreography are a major determinant of the ability of a paramyxovirus to be transmitted to a new species. The process of host-cell attachment is an important determinant of cellular and host tropism, and paramyxoviral RBPs display a great propensity to adapt to new binding partners and therefore hosts (Zeltina, Bowden, and Lee 2016; Kitchen, Shackelton, and Holmes 2011). Across the *Paramyxoviridae* family, members utilise a spectrum of carbohydrate and proteinaceous receptors. For example, viruses within the *Respirovirus* (Lawrence et al. 2004), *Orthoavulavirus*, *Metaavulavirus*, *Paraavulavirus* (Crennell et al. 2000), and *Orthorubulavirus* genera (Kubota et al. 2016) utilise sialic acid, viruses within the *Morbillivirus* genus utilise the Ig-like molecules Nectin-4 and signalling lymphocyte activation molecule F1 (SLAMF1) (Sato et al. 2012), and members of the *Henipavirus* genus differentially utilise EphrinB1, EphrinB2, and EphrinB3 (Xu, Broder, and Nikolov 2012; Negrete et al. 2006b; Negrete et al. 2005; Pryce et al. 2019). Furthermore, structural and functional studies have expanded

the known functional diversity within the family, revealing that recently identified paramyxoviruses may not utilise these established modes of entry (Zeltina, Bowden, and Lee 2016; Rissanen et al. 2017).

In the following sections, I review characterised modes of paramyxovirus–receptor interactions.

1.5.1 Sialic acid mediated entry (and exit)

Through variation of linkage and presentation, sialic acid provides diverse, yet simple structures permissive to viral attachment. Moreover, as sialic acid is almost ubiquitously distributed in different tissues and species, it provides an excellent means for viruses to interact with diverse host organisms (Wasik, Barnard, and Parrish 2016). Indeed, sialic acid is known to be utilised as an entry receptor by a number of viruses with interspecies transmission potential, including influenza virus, Middle East respiratory syndrome coronavirus (MERS-CoV) (Li et al. 2017), and many paramyxoviruses bearing a HN RBP.

Sialic acid is usually located at the termini of N- and O-linked glycan chains presented on host-cell glycoproteins and gangliosides. The linkage of sialic acid on these glycan chains is an important determinant of cellular and species tropism (Wasik, Barnard, and Parrish 2016). For example, human influenza viruses are efficiently transmitted due to the high expression of α 2,6-linked sialic acid in the upper respiratory tract. Furthermore, the ability of avian influenza viruses to productively interact with α 2,3-linked sialic acid presented in the lower respiratory tract is, in part, responsible for outbreaks of more severe disease (Kuiken et al. 2012). Similarly, likely due to the presentation of α 2,8-linked sialic acid on neurons (Mühlenhoff, Eckhardt, and Gerardy-Schahn 1998), adaption of variant Urabe AM9 mumps vaccine strain to α 2,8-linked sialic acid is associated with neurovirulence (Santos-Lopez et al. 2009). Sialic acid also has the capacity to shape virus dynamics beyond receptor attachment. Indeed, sialic acid presented on secreted mucin, within

mucus, has been shown to trap escaping virus, requiring sialic acid tropic viruses to develop means to sever these interactions in the form of neuraminidase activity (Zanin et al. 2016).

Paramyxoviruses of the genera, *Respirovirus*, *Orthoavulavirus*, *Metaavulavirus*, *Paraavulavirus*, and *Orthorubulavirus*, express RBPs that utilise a conserved means to bind and hydrolyse sialic acid (Kubota et al. 2016; Crennell et al. 2000; Lawrence et al. 2004; Yuan et al. 2005). These hemagglutinin-neuraminidase (HN) type RBPs present a six-bladed β -propeller fold common to sialidases presented by viruses, prokaryotes, archaea and eukaryotes, and have likewise preserved conserved mechanisms for binding and hydrolysing terminally presented sialic acid residues, resulting in the release of N-acetylneuramic acid (Neu5Ac) (Quistgaard and Thirup 2009) (**Fig 1.5.1**). Primarily, paramyxoviral HN-type RBPs retain the seven conserved sialidase residues Arg₁, Asp₁, Glu₄, Arg₄, Arg₅, Tyr₆, and Glu₆ (sub-script refers to location on blades 1–6 of the β -propeller fold), and likewise the cognate catalytic pocket located at the central axis of the top face of the propeller (Langedijk, Daus, and VanOirschot 1997). Structural studies of ligand-bound HN-type RBPs have enabled the mapping of the conserved paramyxoviral binding site (Crennell et al. 2000; Kubota et al. 2016; Lawrence et al. 2004; Yuan et al. 2005). Studies have revealed the triarginyl cluster (Arg₁, Arg₄, and Arg₅) binds the carboxy group of sialic acid, Tyr₆ is positioned near the C1–C2 bond of sialic acid, where it interacts with Glu₄ likely to stabilise a transition state intermediate (Vavricka et al. 2013b), Glu₆ stabilises Arg₁, and Asp₁ is primed over the glycosidic oxygen. The different contribution of these residues in different viruses has been confirmed using mutagenesis studies, which utilise techniques measuring the adherence of red blood cells (RBCs) presenting sialic acid (hemadsorption) and the capacity for hydrolysis of sialic acid from a glycoconjugate (neuraminidase activity) (Connaris et al. 2002; Iorio et al. 2001). In the context of NDV-RBP, site-directed mutagenesis of residues Asp₁, Arg₄ and Tyr₆ renders the protein incapable of both activities (Crennell et al. 2000; Connaris et al. 2002; Iorio et al. 2001), whereas Asp₁ is thought to have a lesser role in the activity of MuV (Kubota et al. 2016).

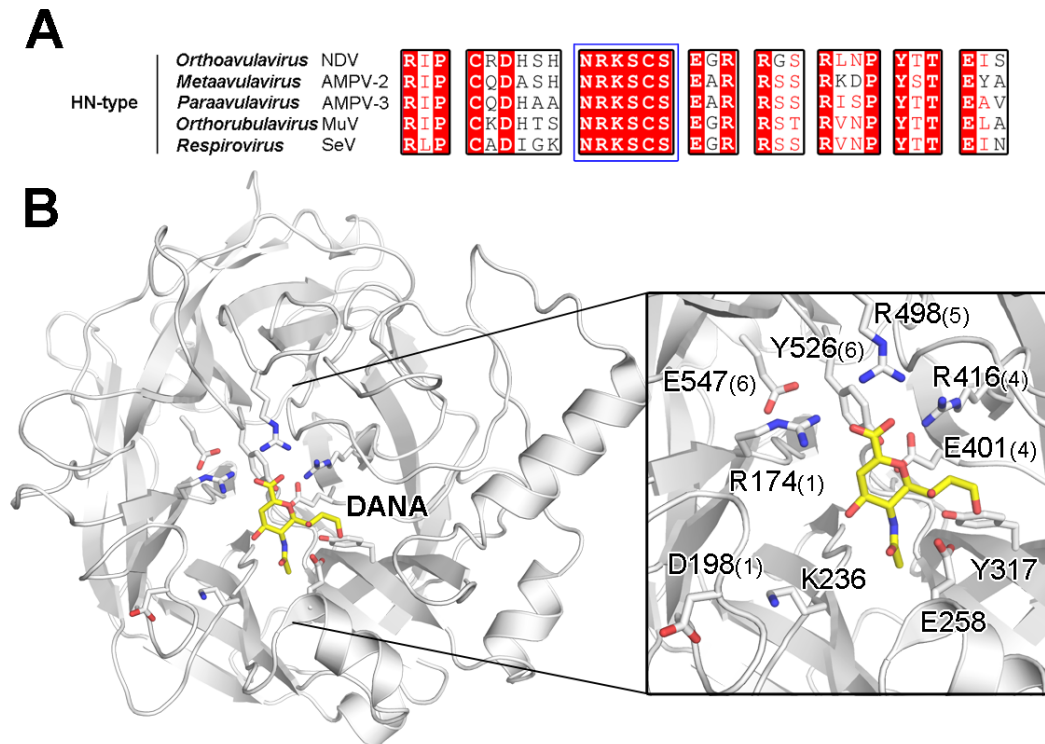


Figure 1.5.1: Conserved residues are implicated in the binding and hydrolysis of sialic acid by HN-type RBPs. (a) An amino acid sequence alignment of RBPs from archetypal sialic acid interacting paramyxoviruses, hPIV3, MuV and NDV. The conserved ‘Asn–Arg–Lys–Ser–Cys–Ser’ hexapeptide motif (Mirza, Deng, and Iorio 1994) and seven sialidase residues are annotated (Crennell et al. 2000; Iorio et al. 2001; Connaris et al. 2002), with sub-script referring to blade location on the propeller. (Langedijk, Daus, and VanOirschot 1997). Genbank accession codes are listed in Appendix D. (b) NDV-RBP (1E8V) is shown in cartoon representation and is coloured white, with the sialic acid binding site highlighted as a zoom-in panel. Within the site, functionally important residues are annotated and side chain sticks shown. The Neu5Ac analogue, N-Acetyl-2,3-dehydro-2-deoxyneuraminic acid (DANA), is coloured yellow and shown using stick representation. Bracketed sub-script refers to the blade location of the seven conserved sialidase residues.

Furthermore, a number of other residues are implicated in the interactions of paramyxoviral RBPs with sialic acid, despite not being as stringently conserved. In addition to Glu₄, it has been noted that Glu258^{NDV-RBP} and Tyr317^{NDV-RBP} (Crennell et al. 2000; Connaris et al. 2002), and homologous residues in related structures, interact with the glycerol group of sialic acid within the context of RBP structures. Moreover, site-directed mutagenesis provides evidence that alteration of these features impacts hydrolysis, likely due to their stabilising role (Connaris et al. 2002). Similarly, through site-directed mutagenesis, the hexapeptide ‘Asn–Arg–Lys–Ser–Cys–Ser’ motif has been discovered to be integral to the enzymatic capability of the paramyxoviral RBP. The hexapeptide

motif is conserved throughout the HN-type RBPs of the *Paramyxoviridae*, and has been structurally shown to interact closely with the seven sialidase residues. In particular, site-directed mutagenesis of the first four residues ‘Asn–Arg–Lys–Ser’ of NDV-RBP ablates neuraminidase activity (Mirza, Deng, and Iorio 1994), potentially due to the loss of choreographed interactions with Arg₁ (Crennell et al. 2000).

Many HN-type RBPs present this dual functionality binding site alone (Yuan et al. 2005; Kubota et al. 2016), however, a secondary sialic acid binding site lacking neuraminidase activity has been identified for NDV (Zaitsev et al. 2004; Porotto et al. 2012b; Bousse et al. 2004), and potentially PIV1 (Bousse and Takimoto 2006; Alymova et al. 2008) and PIV3 (Mishin et al. 2010a; Porotto et al. 2007; Porotto et al. 2012b). Indeed, structural studies of NDV-RBP, reveals binding of sialic acid analogues to a region at the dimer interface, with contributions to site formation by both monomers. A number of hydrophobic residues are implicated in binding the acetamido methyl group of sialic acid, including Gly169^{NDV-RBP}, Leu552^{NDV-RBP}, and Phe553^{NDV-RBP} from one monomer, and Phe156^{NDV-RBP}, Val517^{NDV-RBP}, and Leu561^{NDV-RBP} from the other. Interactions between Arg516^{NDV-RBP} and the sialic acid carboxy group, and the mainchain amide group of Ser519^{NDV-RBP} and O-8 also contribute (Zaitsev et al. 2004). The hydrophobic pocket is likewise conserved in the PIV3 structure, and whilst there is no residue homologous to Arg516^{NDV-RBP}, Arg173^{PIV3-RBP} may provide an alternative means to interaction (Mishin et al. 2010b). Functional data support a model by which this secondary sialic acid binding site promotes fusion through triggering the F protein, however, the exact mechanisms of this interaction are yet to be understood (Porotto et al. 2012b; Porotto et al. 2007; Yuan et al. 2012).

1.5.2 Protein-binding paramyxoviral receptor binding proteins

Morbilliviruses and henipaviruses have adapted to proteinaecous receptors, and in the process, have lost the sialic acid binding-site associated structural and sequence features, resulting in an inability to interact with sialic acid whilst maintaining the overall conserved six-bladed β -propeller fold (Sato et al. 2012; Pernet, Wang, and Lee 2012). It is not clear why viruses adapt to a sialic acid independent

paradigm, however, there is precedent for six-bladed β -propeller scaffolds to adapt to novel functionalities beyond sialidase functionality, as demonstrated by the recently discovered bat influenza viruses H17N10 and H18N11. Here, the neuraminidase-like proteins of the bat influenza viruses were found to have lost residues associated with the sialic acid binding site, in addition to presenting a wide cavity on their top surface incongruous to sialic acid interactions (Wu, Wu, et al. 2014; Li et al. 2012; Zhu et al. 2012; Karakus et al. 2019; Tong et al. 2013). These observations illustrate the potential for new RBP architectures to give way to unique attachment and entry paradigms.

1.5.2.1 Morbilliviral entry

The morbilliviruses present a diversified RBP, with the capacity to bind a number of proteinaceous receptors. The RBP of MeV was initially dubbed the H protein, due to its ability to agglutinate primate erythrocyte cells (Waterson, Rott, and Ruckle-Enders 1963). However, subsequent studies revealed that the targeted host-cell receptor was not the typical paramyxovirus target, sialic acid, and instead includes several proteinaceous receptors (Sato et al. 2012).

Both the Edmonston MeV vaccine strain and Vero-cell passaged strains of MeV infect cells that are CD46 positive. Interestingly, however, these cell types are not permissive to infection by virus isolated from B- and T-cell lines, therefore CD46 likely did not represent a naturally targeted receptor (Dorig et al. 1993). The cell-surface receptor, SLAMF1 (also known as CD150), is expressed on an array of different lymphocytes (Tatsuo et al. 2000) and was discovered to be the likely primary receptor for MeV. This interaction is thought to direct initial infection and therefore constitutes a major determinant of host-cell and species tropism. Interestingly, the usage of SLAMF1 correlates with the observed immunosuppression caused by MeV infection, due to its expression in lymphocytes, monocytes and dendritic cells (Dhiman, Jacobson, and Poland 2004; Lin and Richardson 2016). Nectin-4 constitutes the second known functional receptor of MeV. Nectin-4 is expressed on the basolateral surface of polarised epithelium and enables entry of the virus into

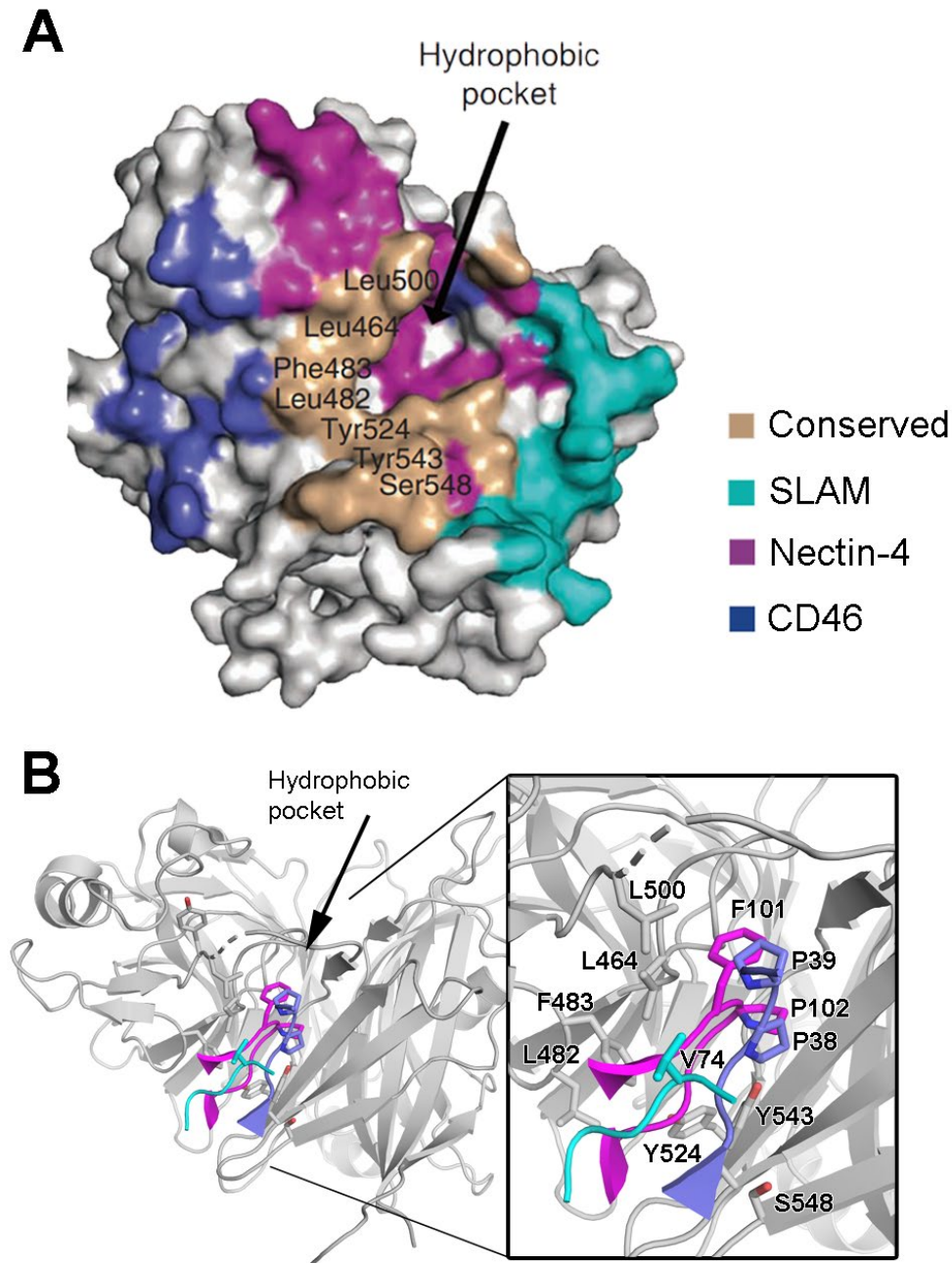


Figure 1.5.2.1: The modes of MeV-RBP receptor engagement. (a) The MeV-RBP shown in surface representation with the footprints of the three receptors engaged by MeV, SLAM (cyan) (Hashiguchi et al. 2011), nectin-4 (magenta) (Zhang et al. 2013), and CD46 (blue) (Santiago et al. 2010) coloured. Surface accessed by all three receptors is coloured orange (labelled ‘conserved’ in the adjacent legend) with residues annotated. This figure has been adapted from Zhang et al, 2013. (b) The MeV-RBP is shown in cartoon representation, with functionally important residues of the $\beta 4$ – $\beta 5$ groove annotated and shown in stick representation. The mode by which the MeV-RBP receptors engage the $\beta 4$ – $\beta 5$ groove is shown, with loops of SLAM (cyan) (3ALZ), nectin-4 (magenta) (4GJT), and CD46 (blue) (3INB) displayed using cartoon representation and sticks for important residues.

tissues, such as the airway epithelium, that enable dispersal of MeV between host organisms (Noyce et al. 2011).

The receptors bound by MeV, SLAMF1, nectin-4, and CD46, recognise the same hydrophobic binding region on MeV-RBP, which is located at the side of the $\beta 4$ – $\beta 6$ propeller face at a concave groove formed by loop $\beta 4L23$ and β -strands $\beta 5S2$ and $\beta 5S3$ (Hashiguchi et al. 2007; Santiago et al. 2010; Zhang et al. 2013; Hashiguchi et al. 2011) (**Fig 1.5.2.1**). In the case of the CD46–MeV-RBP interaction, the D'D- loop of the short consensus repeat (SCR) 1 in CD46 penetrates into the groove and is stabilised by hydrophobic interactions with $\text{Leu464}^{\text{MeV-RBP}}$, $\text{Leu500}^{\text{MeV-RBP}}$, $\text{Tyr541}^{\text{MeV-RBP}}$ and $\text{Tyr543}^{\text{MeV-RBP}}$ (Santiago et al. 2010). SLAM, on the other hand, utilises hydrophobic interactions between $\text{Val74}^{\text{SLAM}}$ and $\text{Asn72}^{\text{SLAM}}$ and residues including, $\text{Phe483}^{\text{MeV-RBP}}$, $\text{Tyr524}^{\text{MeV-RBP}}$, $\text{Tyr543}^{\text{MeV-RBP}}$ and $\text{Pro545}^{\text{MeV-RBP}}$. Aromatic stacking also contributes to interactions across this pocket, where residues $\text{Tyr524}^{\text{MeV-RBP}}$, $\text{Tyr541}^{\text{MeV-RBP}}$, $\text{Tyr543}^{\text{MeV-RBP}}$ and $\text{Phe552}^{\text{MeV-RBP}}$ interact with $\text{Phe119}^{\text{SLAM}}$ and the $\text{Glu75}^{\text{SLAM}}$ – $\text{Arg130}^{\text{SLAM}}$ salt bridge (Hashiguchi et al. 2011). Finally, the nectin-4–MeV-RBP interaction appears to resemble that of CD46–MeV-RBP, and is achieved through insertion of the F–G loop of nectin-4 into the concave $\beta 4$ – $\beta 5$ groove of MeV-RBP. This interaction is further strengthened by the formation of hydrogen bonds between $\text{Ala103}^{\text{Nectin-4}}$ and $\text{Gly465}^{\text{MeV-RBP}}$ and $\text{Tyr543}^{\text{MeV-RBP}}$ (Zhang et al. 2013).

Whilst all characterised MeV-RBP–receptor interactions involve the $\beta 4$ – $\beta 5$ groove, interactions beyond this site differ drastically. The CD46 SCR1 and SCR2 interact predominantly with blade $\beta 4$, and utilise a number of hydrogen bonds (Santiago et al. 2010). The SLAMF1 interaction appears to be centred across the concave $\beta 4$ – $\beta 5$ groove and blade $\beta 5$, with contributions from the neighbouring $\beta 4$ and $\beta 6$ blades. Importantly, salt bridges between MeV-RBP and SLAM are formed between the acidic $\text{Asp505}^{\text{MeV-RBP}}$ and $\text{Asp507}^{\text{MeV-RBP}}$ and basic $\text{Lys77}^{\text{SLAM}}$ and $\text{Arg90}^{\text{SLAM}}$ (Hashiguchi et al. 2011). The nectin-4 binding site exists predominantly over blades $\beta 4$ and $\beta 5$, with some contribution from $\beta 3$ (Zhang et al. 2013).

Studies have further confirmed that receptors homologous to SLAMF1 and nectin-4 are utilised by other members of the morbillivirus family in their respective target species, such as PPRV, CDV, PDV and CeMV (Sato et al. 2012; Birch et al. 2013; Rendon-Marin et al. 2019; Tatsuo, Ono, and Yanagi 2001; Noyce, Delpeut, and Richardson 2013). The intricacies of the interactions mapped on the MeV surface also provide insights into the binding between other morbilliviruses and their receptors (Langedijk et al. 2011). Indeed, similar to MeV, host-specific SLAMF1 and nectin-4 have been identified as the receptors for a number of morbilliviruses (Duignan et al. 2014; Van Bresseem et al. 2014; Sato et al. 2012; Birch et al. 2013; Noyce, Delpeut, and Richardson 2013; Tatsuo, Ono, and Yanagi 2001; Rendon-Marin et al. 2019). In fact, prior to discovery of nectin-4 as the epithelial receptor for MeV, structure-informed site-directed mutagenesis studies revealed the importance of the CDV-RBP $\beta 4$ – $\beta 5$ groove in regulating the interactions with both immune cells and SLAMF1 negative keratinocytes, indicating that CDV-RBP also specifically bound a then unidentified epithelial receptor enabling keratinocyte infection (Langedijk et al. 2011). While replication of morbilliviruses is restricted at the point of SLAMF1 and nectin-4-mediated entry into a host-cell (Ohishi et al. 2014), recent studies have revealed that minimal changes are required to drive binding of CDV-RBP and PPRV-RBP to human SLAM (Bieringer et al. 2013; Abdullah et al. 2018), and no changes are required for engagement of human nectin-4 receptor by CDV-RBP, raising concerns about the potential for these viruses to overcome the current species barrier.

1.5.2.2 Henipaviral entry

The henipaviruses, like morbilliviruses, utilise an RBP that recognises proteinaceous receptors (Weis et al. 2014; Lee et al. 2015; Marsh et al. 2012; Pernet, Wang, and Lee 2012; Pryce et al. 2019). However, while both RBP architectures have diversified from the sialic acid-dependent mode of viral entry, the henipaviral RBPs (HNV-RBPs) maintain a much closer structural relationship with the HN-type RBPs (Bowden, Crispin, et al. 2008; Bowden, Aricescu, et al. 2008). The ability of the pathogenic henipaviruses, HeV and NiV, to spillover from bat reservoirs into many host species and incite devastating disease is partly dependent on the specificity of their RBP for B-class ephrin cell-

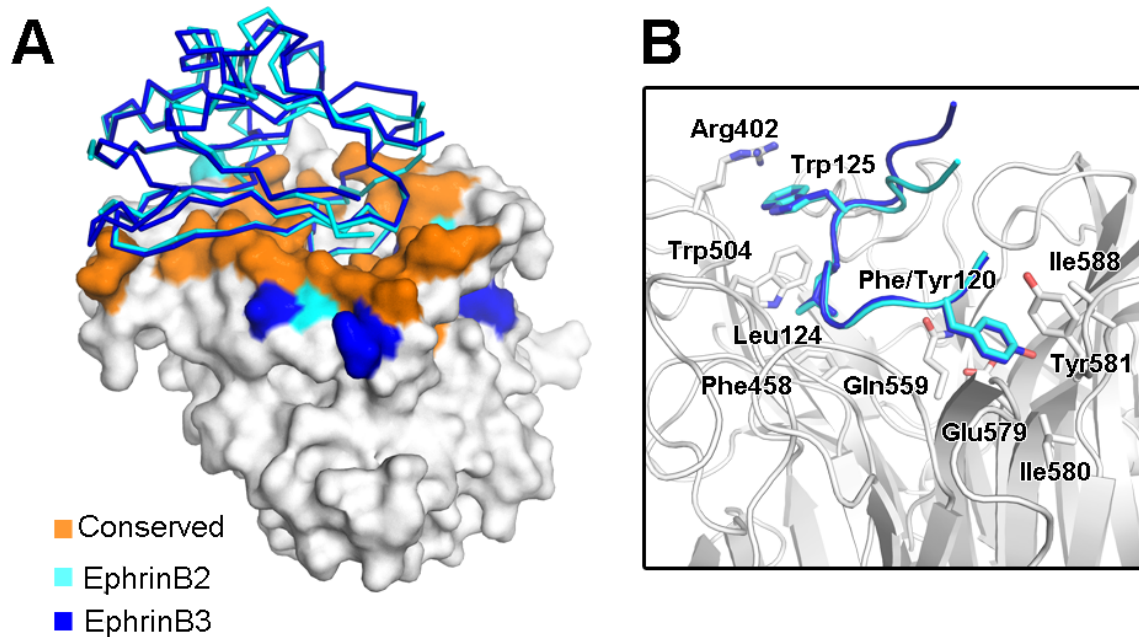


Figure 1.5.2.2: Interactions between NiV-RBP and ephrinB2 and ephrinB3. (a) EphrinB2 (2VSM) and ephrinB3 (3D12) (cyan and blue, respectively) are shown in ribbon representation bound to NiV-RBP (2VSM), displayed using surface representation (Bowden, Aricescu, et al. 2008; Xu et al. 2008). Residues bound by both ephrinB2 and ephrinB3 are coloured orange. (b) Cartoon representation of NiV-RBP bound by ephrinB2 and ephrinB3, showing the interactions between the G–H loop and the top face pocket presented by NiV-RBP. Residues forming important contributions to the virus–receptor interface are shown using stick representation and annotated.

surface receptors. The cell-surface presented ephrins are ligands to the Eph family of receptor tyrosine kinases (RTKs), and are involved in mediating numerous developmental and neurological processes (Negrete et al. 2005; Negrete et al. 2006b; Bonaparte et al. 2005; Lisabeth, Falivelli, and Pasquale 2013). Consistent with this diverse functionality, B-class ephrins (ephrinB1, ephrinB2, and ephrinB3) are widely presented across numerous tissues and cell-types and exhibit an expression pattern that relates to the associated pathology of henipaviral disease (Pernet, Wang, and Lee 2012). Furthermore, consistent with the broad tropism exhibited by henipaviruses (Rockx 2014) there is high level of sequence conservation between ephrins of different species (Pernet, Wang, and Lee 2012), resulting in a highly adaptable paramyxovirus entry mechanism.

Ephrins are differentially utilised by henipaviral RBPs. Indeed, while the RBPs of NiV and HeV recognise ephrinB2 and ephrinB3 (Negrete et al. 2005; Negrete et al. 2006a), the Ghana

henipavirus RBP (GhV) utilises only ephrinB2 (Lee et al. 2015), and the RBP of Cedar virus (CedV) binds ephrinB1 and ephrinB2 (Pryce et al. 2019). Interestingly, in contrast to NiV and HeV, CedV and GhV are thought to be non-pathogenic. Furthermore, it has been hypothesised that usage of neuron-specific ephrinB3 correlates with the dramatic clinical outcomes associated with NiV and HeV infection (Weis et al. 2014; Marsh et al. 2012; Lee et al. 2015; Pryce et al. 2019).

Similar to the mode of engagement between B-class ephrins and their cognate receptors, the mechanism of viral engagement relies upon an interaction between the highly flexible G–H loop of the B-class ephrins and a shallow cleft presented on the top face of the HNV-RBP (Pernet, Wang, and Lee 2012) (**Fig 1.5.2.2**). The interactions between the HeV and NiV RBPs with ephrinB2 and ephrinB3 have been extensively characterised through site-directed mutagenesis and structural studies (Bowden, Aricescu, et al. 2008; Lee et al. 2015; Xu et al. 2008). At the outer edge of blade $\beta 5$, both HeV-RBP and NiV-RBP present Glu533, which forms a salt bridge with the positively charged Lys60^{ephrinB2} or Arg57^{ephrinB3}, stabilising G–H loop engagement. Residues Leu124^{ephrinB2/B3}, Trp125^{ephrinB2/B3}, F120^{ephrinB2} and Y120^{ephrinB3} of the G–H contribute to these extensive protein–protein interfaces (Bowden, Aricescu, et al. 2008; Xu et al. 2008). Despite being unable to bind ephrinB3 and exhibiting a low sequence identity between GhV and the archetypal NiV and HeV-RBPs, GhV-RBP utilises a similar mechanism of ephrinB2 engagement (Lee et al. 2015). Similarly, the sequence diverse CedV-RBP has maintained its ability to bind ephrinB2 (Marsh et al. 2012), but has also adopted the ability to engage ephrinB1 (Pryce et al. 2019) through the presentation of the Tyr121^{ephrinB1} and Met122^{ephrinB1} (YM) motif on the G–H loop, in place of the Leu124^{ephrinB2/B3} and W125^{ephrinB2/B3} (LW) motif. Interestingly, a W504Y substitution to NiV-RBP creates a cavity similar to that observed on CedV-RBP, rendering the glycoprotein capable of accommodating the YM motif of ephrinB1 (Pryce et al. 2019). These experiments demonstrate that only single substitutions to the HNV-RBP are necessary to alter ephrin receptor tropism.

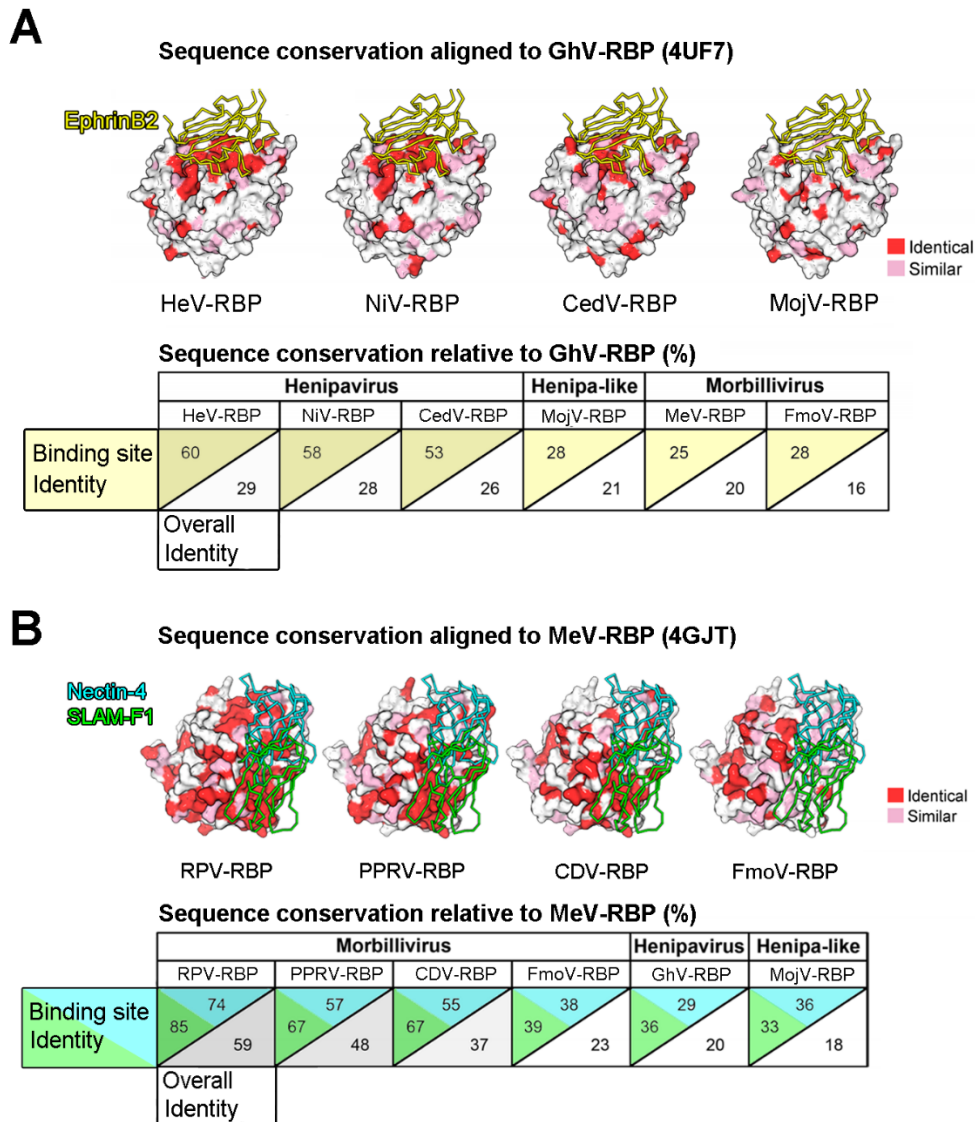


Fig. 1.5.4: Paramyxoviral RBPs present an elevated amino acid sequence similarity across receptor binding sites. (a) GhV-RBP is displayed using surface representation, with ephrinB2 (yellow) bound and presented using ribbon representation (4UF7) (Lee et al. 2015). Amino acid sequence similarity between GhV-RBP and HeV-RBP, NiV-RBP, CedV-RBP and MojV-RBP is mapped onto the GhV-RBP structure, with identical residues coloured red, and similar coloured pink. The table below summarises the overall amino acid similarity between GhV-RBP and the other RBPs, and the similarity across the ephrinB2 binding site (yellow). (b) MeV-RBP is displayed using surface representation, with SLAM (green) and nectin-4 (cyan) presented in a ribbon format (PDB IDs 3ALZ and 4GJT, respectively) (Zhang et al. 2013; Hashiguchi et al. 2011). Amino acid sequence similarity between RPV-RBP, PPRV-RBP, CDV-RBP and feline morbillivirus (FmoV)-RBP is mapped onto the MeV-RBP structure, with residues coloured as in 1.5.4A. The table below summarises the overall similarity between GhV-RBP and the other RBPs, and the similarity across the ephrinB2 binding site. The table below summarises the overall similarity between MeV-RBP and the other RBPs, and the similarity across the SLAMF1 (green) and nectin-4 (blue) binding sites. This figure has been adapted from Zeltina et al, 2016.

1.5.3 Conservation across paramyxoviral receptor binding sites

As described above, the molecular basis of paramyxoviral receptor engagement has been extensively characterised for a number of HN, H, and G-type RBPs. Importantly, the necessity to interact with a specific receptor(s) likely acts as a functional constraint during paramyxovirus evolution and has been observed to lead to elevated levels of sequence conservation at surfaces of the RBP required for receptor recognition (Zeltina, Bowden, and Lee 2016). (Langedijk, Daus, and VanOirschot 1997; Crennell et al. 1993; Burmeister, Ruigrok, and Cusack 1992; Grobe et al. 1998; Gut, King, and Walsh 2008; Luo et al. 1999; Roggentin et al. 1989; Stummeyer et al. 2005). For example, as described in **section 1.5.1**, the hexapeptide and residues that facilitate hemagglutinin and neuraminidase activity in HN RBPs are strictly conserved across respiro, orthoavula, metavula, paravula, and orthorubulaviruses. Similarly, the functional constraints imposed by receptor binding are exemplified by the maintenance of the conserved B-class ephrin binding site on CedV and GhV RBPs, despite high sequence variation (Zeltina, Bowden, and Lee 2016; Lee et al. 2015) (**Fig. 1.5.3**). These findings were contrasted by the Mojiang virus RBP (MojV-RBP), a divergent HNV which exhibits high sequence variability across the regions associated with ephrin binding, and is therefore unable to engage B-class ephrins (Zeltina, Bowden, and Lee 2016) (**Fig. 1.5.3A**). This conclusion was confirmed through structural and functional characterisation of MojV-RBP, which revealed a lack of the canonical B-class ephrin binding site and an inability to interact with known paramyxoviral receptors *in vitro* (Rissanen et al. 2017).

Whilst MeV-RBP represents the only structurally characterised morbilliviral RBP to date, sequence mapping analysis reveals a similar trend, where SLAMF1 and nectin-4 receptor binding sites exhibit elevated levels of sequence conservation amongst morbilliviruses (Zeltina, Bowden, and Lee 2016; Langedijk et al. 2011; Liang et al. 2016; Abdullah et al. 2018). For example, although CDV-RBP is only 37% identical with MeV-RBP across the full-length sequence, the proteins exhibit 67% and 55% sequence similarity at SLAM and nectin-4 binding sites, respectively (**Fig. 1.5.3B**) (Zeltina, Bowden, and Lee 2016). Despite a propensity to bind receptors that are not necessarily well

conserved between different host species, several important determinants of MeV infection remain conserved, such as the residues that form the $\beta 4$ – $\beta 5$ groove. Indeed, studies have confirmed that the $\beta 4$ – $\beta 5$ groove is likely to be similarly utilised in other morbilliviral RBPs, with mutagenesis of the CDV-RBP $\beta 4$ – $\beta 5$ groove resulting in abrogation of infection in nectin-4 expressing keratinocytes and SLAM-expressing cells (Abdullah et al. 2018; Liang et al. 2016; Langedijk et al. 2011). Based on this understanding of the molecular determinants of the establish paramyxoviral receptor binding mechanisms, as novel paramyxoviruses emerge from the environment, comparisons with homologous structures may provide insight into receptor usage. However, these analyses are limited to comparisons with characterised binding sites. In the case of MojV-RBP, while a lack of B-class ephrin interactions has been confirmed, receptor tropism is still yet to be determined (Rissanen et al. 2017).

1.5.4 Trends in dimer angle organisation

The paramyxovirus RBP is presented as a dimer-of-dimers on the virion surface, where the β -propeller receptor binding head domain forms homodimers (Hashiguchi et al. 2007; Lawrence et al. 2004; Crennell et al. 2000; Bowden, Crispin, Harvey, et al. 2010), and the β -helical stalk region facilitates tetramerisation through inter-protomer disulphide bond formation (Yuan et al. 2005; Yuan et al. 2011; Welch et al. 2013). A flexible linker region enables the movement of dimerically-associated β -propeller domains between ‘heads-down’ and ‘heads-up’ orientations (Welch et al. 2013), with such movements associated with fusion activation (Adu-Gyamfi et al. 2016; Herren et al. 2018; Jardetzky and Lamb 2014; Bose et al. 2014) (**Fig. 1.5A**). Whilst the dimeric interface is typically comprised of residues from blades $\beta 1$ and $\beta 6$ of the RBP propeller, structural characterisation of RBP architectures has revealed that these assemblies may adopt different modes of association, depending on whether the RBP exhibits H, HN, or G functionality (**Fig 1.5.4**) (Bowden, Crispin, et al. 2008; Bowden, Crispin, Harvey, et al. 2010; Crennell et al. 2000; Yuan et al. 2005; Lawrence et al. 2004; Hashiguchi et al. 2007).

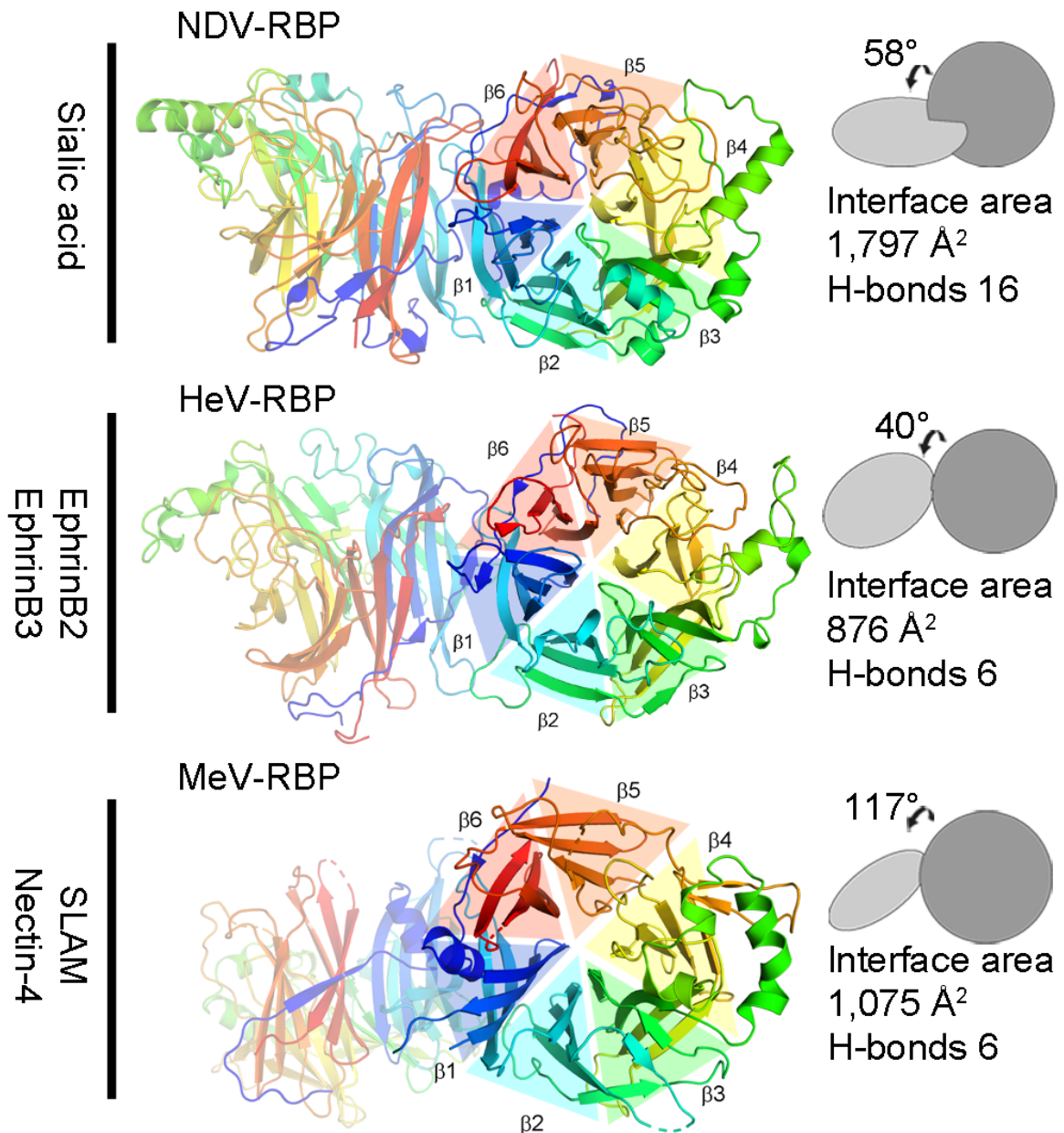


Fig 1.5.5: Known dimeric assemblies of paramyxoviral RBPs. Representative HN, G, and H RBP dimeric assemblies from NDV-RBP, HeV-RBP and MeV-RBP, respectively, are shown in cartoon representation with monomers coloured rainbow from the N-terminus to C-terminus (blue to red, respectively). Calculated angle of association, interface area and number of hydrogen bonds is displayed to the right-hand side of the relevant dimer. The known receptors bound by each RBP type is listed to the left-hand side. This figure has been adapted from Stelfox et al, 2019.

Structures of higher order oligomers of RBPs with sialic acid interacting capabilities demonstrate a common mode of organisation utilising extensive contacts between blades $\beta 1$ and $\beta 6$, resulting in an average buried surface area of approximately 1,790 Å² (**Fig 1.5.5**) (Stelfox and Bowden 2019). This dimer association sets the top faces of the protomers, and therefore the sialic acid binding site, at an approximate angle of 60° relative to each other (Kubota et al. 2016; Lawrence et al. 2004; Crennell et al. 2000; Yuan et al. 2005). This is contrasted by dimeric arrangements observed to be adopted by the protein binding morbilliviral and henipaviral RBPs. The observed HeV-RBP dimer is similarly formed through interactions involving the $\beta 1$ and $\beta 6$ blades, however, the contribution from $\beta 6$ is less dominant and overall the buried surface area is approximately 880 Å². It has been hypothesised that the dimeric association between two protomeric receptor binding domains is influenced by an evolutionary requirement to access sometimes bulky cell surface receptors, such as the B-class ephrins, SLAM or nectin-4 (Hashiguchi et al. 2007; Hashiguchi et al. 2011; Santiago et al. 2010; Zhang et al. 2013; Yuan et al. 2008; Stelfox and Bowden 2019; Bowden, Crispin, Harvey, et al. 2010). The association between the protomers of the dimeric HeV-RBP structure is 40° and results in the upward presentation of the B-class ephrin binding site, potentially facilitating receptor attachment without disruption of the overall oligomeric assembly (Bowden, Crispin, Harvey, et al. 2010; Yuan et al. 2008). Similarly, the crystallographically observed organisation of the MeV-RBP deviates from the paradigm originally presented by the HN-type RBPs. (**Fig 1.5.5**). Several structures of MeV-RBP dimers in complex with cognate SLAMF1, nectin-4, and CD46 receptors have been solved, revealing a conserved mechanism of association primarily mediated through the $\beta 1$ blade, and lacking involvement from the $\beta 6$ blade. The receptor bound structures, with binding sites orientated upwards, reveal that the unique oligomeric organisation could potentially facilitate the docking of multiple receptor molecules at the cell surface (Santiago et al. 2010; Zhang et al. 2013; Hashiguchi et al. 2011; Hashiguchi et al. 2007).

The trends in RBP association angle and level of buried surface across the paramyxovirus family, are consistent with the hypothesis that the overall presentation of RBP molecules is linked

to the type of receptor(s) utilised. Indeed, changes in the angle of association between protomers of a RBP dimer may enable engagement to a wider array of receptors, including larger proteinaceous receptors.

CHAPTER 2

MATERIALS AND METHODS

2 MATERIALS AND METHODS

2.1 MOLECULAR BIOLOGY

2.1.1 cDNA templates

Gene encoding nucleotide sequences from receptor-binding proteins (RBP) were codon optimised for use in human cell expression systems and synthetically produced by GeneArt™ (Life Technologies) as GeneStrings or within transfer plasmids. All genes were produced with an AgeI site at the 5' end and a KpnI site at the 3' end, enabling cloning of the constructs into appropriate vectors. Genes supplied as lyophilised fragments were resuspended in sterile filtered deionised water to a final concentration of 100 ng/μL and amplified using polymerase chain reaction (PCR) prior for cloning. Plasmids containing genes were resuspended as above, and subsequently subjected to PCR or template amplification in *E. coli*.

2.1.2 Cloning strategy and expression vectors

All constructs are listed (Table 2.1.2). Typically, standard restriction digest cloning protocols were utilised, but at times the In-Fusion® method was utilised due to high cloning efficiency, when reagents were available (Park, Throop, and LaBaer 2015). All plasmids used, encoded ampicillin resistance as a selection marker.

2.1.2.1 Full-length RBP constructs

Entire receptor-binding protein gene open reading frames (ORFs) encompassing the intraviral portion, transmembrane domain, stalk region and β-propeller were cloned into the high-copy number plasmid, pHLsec (Aricescu, Lu, and Jones 2006), resulting in a stock of template for further construct production and for the expression of full-length constructs for use in cell-based assays.

2.1.2.2 Constructs for soluble expression

Design of constructs for soluble RBP expression was guided by structure based sequence alignment with closely related and structurally characterised paramyxoviral RBPs.

2.1.2.2.1 pHLsec hexa-histidine tagged constructs

Typically, soluble constructs encoding the stalk and β -propeller region and/or the β -propeller only were cloned into the pHLsec vector, facilitating production of large quantities of recombinant protein. Typically, fragments were inserted between the AgeI and KpnI restriction sites, enabling incorporation of the vector secretion signal peptide (SP) and the C-terminal hexa-histidine tag (**Fig 2.1.2**). Incorporation at these sites resulted in the soluble protein retaining the amino acids ETG and GTK proximal to the signal sequence and hexa-histidine tag, respectively.

2.1.2.2.2 pHLsec Fc tag constructs

Soluble constructs were sub-cloned into a pHLsec vector containing a C-terminal Fc tag followed by a hexa-histidine tag. The addition of an Fc allowed efficient protein G based purification and subsequent analysis of the proteins using flow cytometry and pulldown studies.

2.1.2.2.3 pHLsec octa-histidine SUMO-tagged constructs

To aid in soluble expression, SosV-RBP and BeiV-RBP(wt) were cloned into a modified pHLsec vector encoding an octa-histidine tag, *Saccharomyces cerevisiae* small ubiquitin-like modifier (SUMO) (Panavas, Sanders, and Butt 2009) and 3C cleavage site at the N-terminus (Waugh 2011) and a 1D4 tag at the C-terminus (Chang et al. 2015; Molday and Molday 2014). Inserts were digested with AgeI and KpnI, and the plasmid with the restriction enzyme XmaI, which produces compatible cohesive ends. The SUMO-tag could be cleaved using human rhinovirus (HRV) 3C protease (Waugh 2011), if necessary.

2.1.2.2.4 pHLSEC AviTag™ constructs

The 8xHis-SUMO-**SosV-RBP** cassette was cloned into an AviTag™ encoding pHLsec using AgeI and KpnI (Fairhead and Howarth 2015). The addition of an AviTag™ enabled subsequent biotinylation with the BirA enzyme, making the proteins suitable for flow cytometry (**Section 2.9.3.1**).

2.1.2.2.5 pURD octa-histidine SUMO-tagged construct

The SP-8xHis-SUMO-**SosV-RBP**-1D4 cassette was sub-cloned into the pURD expression vector between the EcoRI and KpnI restriction sites (Seiradake 2015). Cloning into pURD enabled subsequent production of stable cell lines upon cotransfection with a plasmid containing the ϕ 31C integrase. Production of a stable cell line enabled constitutive production of protein and boosted yields. The method is described in **section 2.3.6.1** (Seiradake 2015).

2.1.2.2.6 pHR-CMV-TetO₂ construct for lentivirus production

The SP-8xHis-SUMO-**SosV-RBP**-AviTag™ was cloned into a lentivirus pHR-CMV-TetO₂ transfer plasmid using EcoRI and XhoI sites. Transfection of this plasmid into a Lenti-X HEK293T cell line alongside helper plasmids enabled the production of Lentivirus capable of infecting and transducing a target cell line, as described in **section 2.3.6.2** (Elegheert et al. 2018).

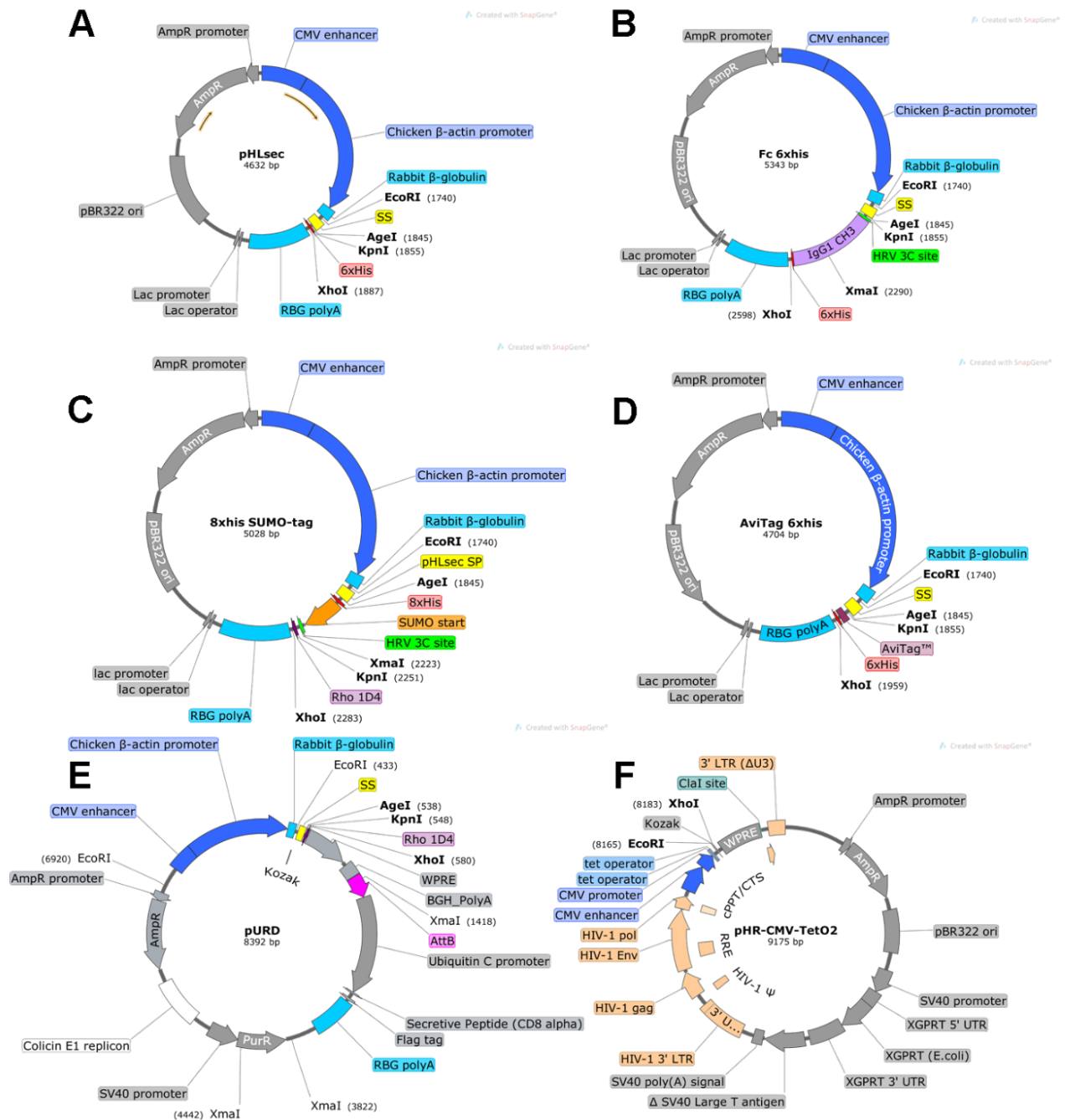


Figure 2.1.2: Plasmid maps annotated with features and relevant restriction enzyme sites. All inserts were cloned into plasmids at the multiple cloning site (MCS), where restriction sites are encoded. Plasmids A-D derived from pHLsec mammalian expression vector for transient transfection (Aricescu, Lu, and Jones 2006). (a) pHLsec encoding a hexa-histidine tag 3' to the MCS. (b) pHLsec encoding a Fc and hexa-histidine tag at the 3' to MCS. (c) pHLsec vector encoding a octa-histidine tag and SUMO tag 5' to the MCS and a ID4 tag 3' to the MCS. (d) pHLsec vector encoding an AviTag™ and a hexa-histidine tag 5' to the MCS. (e) pURD plasmid, capable of site specific integration, encoding a ID4 tag 5' to the MCS (Seiradake 2015). (f) pHR-CMV-TetO₂ transfer plasmid, which in combination with helper plasmids can produce lentivirus for stable cell line production (Elegheert et al. 2018).

2.1.3 Primer specifications

DNA primers were produced by Life Technologies or Eurofins. Custom primers were all designed to have restriction endonuclease sites, short overhangs to facilitate digestion and an overlapping region to the desired target sequence. To facilitate cloning with the Thermo Scientific® Phusion® High-Fidelity DNA Polymerase, the melting temperature (T_m) of the primer regions overlapping with desired ORFs was adjusted to between 71°C and 74°C using the online NEB T_m calculator (<https://tmcaculator.neb.com>). In addition, some primers were synthesised with a further minimum of fifteen bases at the 5' end homologous to target vector regions encompassing and proximal to the AgeI and KpnI restriction sequences to enable In-Fusion® cloning (Park, Throop, and LaBaer 2015).

2.1.4 Polymerase chain reaction (PCR)

Entire open reading frames or specific truncated constructs were amplified using the manufacturer's recommended PCR protocol for the Phusion® High-Fidelity PCR Mastermix with HF buffer (New England Biolabs). Reactions were assembled in a 50 µL volume, with 100 ng of template DNA, 0.5 µM concentration of each primer and the mastermix diluted to a 1x concentration. Reactions were heated to 98°C for an initial denaturation prior to immediate thermocycling. Thermocycling was then performed 30x with the following setup: 10 s at 98°C denaturation, 10–15 s at 55°C annealing, 30 s per kb at 72°C extension. Finally, elongation was allowed to proceed for a further 10 min at 72°C, before cooling to 4–12°C. The resulting DNA fragments were purified either using a QIAquick PCR Purification Kit (Qiagen) and PCR fragments analysed using agarose gel electrophoresis, or using a QIAquick Gel Extraction Kit (Qiagen) if fragments were to be applied to a subsequent splicing by overlap extension PCR or In-Fusion® reaction. In both cases, the manufacturers protocol was adhered to, unless fragments were to be used for a subsequent In-Fusion® reaction, where an extra ethanol wash was required.

VIRUS	GENBANK	VECTOR	TAG(S)	START	END	PRIMERS	USE
MOSV	NP_958054.1	pHLSEC	6xhis	1	632	N/A	Template, and cell based assays
	NP_958054.1	pHLSEC	6xhis	202	632	F: CGCACCGGTGAGCTGTACACCAGAGTGCGG R: GCGGTACCAGTGTGGCAATTGTC	Crystallisation
	NP_958054.1	pHLSEC	Fc-tag, 6xhis	157	632	F: CGTAGCTGAAACCGGTACCTGTCTGCGGATCGAGGAAA R: AACAGCACCTCCAGGGTACCAGTGTGGCACTTGTCTACTG	Flow cytometry
NARV	YP_006347588.1	pHLSEC	6xhis	1	657	N/A	Template, and cell based assays
	YP_006347588.1	pHLSEC	6xhis	184	657	F: CGCACCGGTGAAGCCAGCGAGATCCACAGC R: GCGGTACCGGGTGTTC TAGGGG	Crystallisation
	YP_006347588.1	pHLSEC	Fc-tag, 6xhis	184	657	F: CGTAGCTGAAACCGGTGAAGCCAGCGAGATCCACAGC R: AACAGCACCTCCAGGGTACCGGGTGTTC TAGGGGCG	Flow cytometry
JPV	YP_338084.1	pHLSEC	6xhis	1	709	N/A	Template, and cell based assays
	YP_338084.1	pHLSEC	6xhis	149	709	F: CGCACCGGTGAAGTGGGCATCGAGGCCATC	Crystallisation

JPV						R: GCGGTACCACCGTTGTGGCCGATGGG	
	YP_338084.1	pHLSEC	Fc-tag, 6xhis	123	709	F: CGTAGCTGAAACCGGT CAGTGCACCCCCAGCTTCATG R: AACAGCACCTCCAGGGTACC GTTGTGGCCGATGG	Flow cytometry
	YP_338084.1	pHLSEC	8xhis, SUMO, 1D3	149	709	N/A	Biochemical assays
	YP_338084.1	pHLSEC	8xhis, SUMO, 1D4	149	584	F: CGTAGCTGAAACCGGT GAAGTGGGCATCGAGGCCATC R: AACAGCACCTCCAGGGTACC GTGGGCATCCAGAGGTGTCAGG	Biochemical assays
BEIV	YP_512253.1	pHLSEC	6xhis	1	734	N/A	Template, and cell based assays
	YP_512253.1	pHLSEC	6xhis	145	734	F: CGCACCGGTAACCCACCCACAGCAGATACTTCG R: GCGGTACCCTCTTTGGGGTATTCC	Attempted expression
	YP_512253.1	pHLSEC	8xhis, SUMO, 1D4	145	734	N/A	Crystallisation
	AOV81769.1	pHLSEC	6xhis	1	1046	N/A	Template, and cell based assays
	AOV81769.1	pHLSEC	6x his	145	1046	F: CGTAGCTGAAACCGGT AATCCCACACACAGCCGGTACTT R: TGATGGTGGTGCTTGGTACCAGGTGTCACTCCCCG	Attempted expression

	AOV81769.1	pHLSEC	8xhis, SUMO, 1D4	145	1046	N/A	Attempted EM
	AOV81769.1	pHLSEC	8xhis, SUMO, 1D4	145	589	F: <u>CGTAGCTGAAACCGGT</u> AATCCCACACACAGCCGGTACTT R: <u>TGATGGTGGTGCTTGGTACC</u> TTTGCCGGTGAACCGCATG	Biochemical assays
SOSV	YP_009094033.1	pHLSEC	6xhis	1	582	N/A	Template, and cell based assays
	YP_009094033.1	pHLSEC	6xhis	125	582	F: <u>CGTAGCTGAAACCGGT</u> AACGCCGTGTGTAGAGGCCCT R: <u>TGATGGTGGTGCTTGGTACC</u> CTTGATCAGGTACAGAAAGGACCGGATG	Attempted expression
	YP_009094033.1	pHLSEC	8xhis, SUMO, 1D3	125	582	N/A	Attempted crystallisation
	YP_009094033.1	pURD	8xhis, SUMO, 1D4	125	582	N/A	Crystallisation
	YP_009094033.1	pHLSEC	6xhis, AviTag	125	582	N/A	Attempted purification for flow cytometry
	YP_009094033.1	pHR	6xhis, AviTag	125	582	N/A	Flow cytometry
NDV	Q9Q2W5.1	pHLSEC	6xhis	1	577	N/A	Template, and cell based assays
	Q9Q2W5.1	pHLSEC	6xhis	74	577	F: <u>CGTAGCTGAAACCGGT</u> AGCATGGGCGCCAGCA	Biochemical assays

					R: <u>TGATGGTGGTGCTT</u> GGTACCTCCGCTTCTGGCCTCGC	
--	--	--	--	--	--	--

Table 2.1.2 Constructs synthesised contributing to doctoral work.

^aRestriction site shown with red text

^bRegion overlapping with target plasmid for In-Fusion® shown with underline

2.1.5 Agarose gel electrophoresis

To monitor DNA purity and identity of fragments throughout cloning samples were separated by agarose gel electrophoresis using TAE buffer (Tris, Acetic acid, ethylenediaminetetraacetic acid (EDTA)). Typically, samples were run for 20 min at 100V on an 1% w/v agarose gel produced with a 1:10,000 (v/v) SYBRTM safe (Thermo Fisher scientific). 6x loading dye (NEB) was added to samples and 1 kb Plus DNA ladder at a 1:6 dilution. Visualisation of DNA was performed using a blue light transilluminator if required for gel excision, or imaged with UV light with a ChemiDocTM system (BioRad).

2.1.6 Restriction enzyme digestion

To produce fragments amenable to DNA ligation, insert and target vector DNA were digested with complimentary restriction enzymes (NEB). Reaction mixtures were set up as per the manufacturer's standard protocol and incubated for 30 min–12 h (if enzymes are known not to exhibit star activity). Subsequently, digested fragments were separated using agarose gel electrophoresis (**Section 2.1.5**) and purified using a QIAquick Gel Extraction kit (Qiagen). Typically, restriction digests were performed with AgeI-HF[®] (NEB), KpnI-HF[®] (NEB) in CutSmart[®] buffer (NEB). Inserts required for In-Fusion[®] cloning were not digested, however vector digests were performed and similarly treated.

2.1.7 DNA ligation

Where In-Fusion[®] was not used, digested DNA fragments were ligated using a Quick LigationTM kit (NEB), assembling reactions as per the manufacturer's protocol. Reactions were incubated for 5 min at 21°C prior to immediate transformation of competent *E. coli* or storage at -20°C.

2.1.8 In-Fusion[®]

PCR amplified insert with overlaps amenable to In-Fusion[®] cloning were mixed with the complimentary linearized vector as per the manufacturer's protocol. Briefly, the In-Fusion HD EcoDry Mix was resuspended in 10 µL volume reaction mixtures, which were set up with a 2:1 Insert:Vector ratio. Reactions were subsequently incubated at 37°C followed by 50°C for 15 min

each. Reactions were diluted with 40 μ l TE (Tris, EDTA) buffer prior to transformation or storage at -20°C.

2.1.9 Transformation

Following insertion of desired constructs into target plasmids via either ligation or In-Fusion® methods, plasmids were transformed into chemically competent DH5 α ™ subcloning efficiency *E. coli*. (ID: 18265017, Invitrogen) 30 μ L aliquots of cells were thawed gently on ice prior to adding ligation or In-Fusion® reactions to the cells at a volumetric ratio of 1:10. When transforming purified DNA of higher concentration (e.g. from miniprep), the volumetric ratio of DNA to cells was reduced to 0.3:10. Cells were subsequently incubated on ice prior to heat shock at 42°C for 20 s. Following heat shock, cells were cooled on ice for 2 min prior to addition of 200 μ L SOC (Super optimal broth, Sigma) and incubated at 37°C for 45 min in a shaking incubator (180 rpm), to aid recovery. Cultures were plated on Luria broth (LB)-Agar plates containing ampicillin at 100 μ g/mL and subsequently incubated at 37°C for a further 16–18 h to enable selection of colonies successfully transformed with circularised vector, as the pHLsec vector and its derivatives all confer ampicillin resistance.

2.1.10 Plasmid amplification and purification

Following transformation and overnight incubation, single colonies were inoculated into 5 mL LB supplemented with ampicillin at 100 μ g/mL. Cultures were incubated at 37°C for 16–18 h in a shaking incubator (180 rpm). Bacteria were harvested by centrifugation at 4,000 x g for 5 min, supernatant discarded and plasmids subsequently purified from the bacterial pellets using a QIAprep Spin Miniprep kit (Qiagen), following the manufacturer's protocol.

To determine the success of cloning, purified plasmid was digested and subject to agarose gel electrophoresis to assess the size of inserts. Inserts of expected length were sent for sequence analysis by Eurofins genomics (previously GATC Biotech). Mixtures were prepared prior to sending according to the service requirements: 5 μ l of plasmid DNA at 80–100 ng/ μ L and 5 μ L of primer at

a concentration of 5 μ M. Primers complementary to the regions adjacent to the multiple cloning sites in pHLsec and its derivatives (incl. pURD) were used to ensure coverage of the insert and tags (**Table**

Table 2.1.10: Sequencing primers

Primer	Sequence
pHLsec (pURD) Forward	TTCCTACAGCTCCTGGGCAAC
pHLsec Reverse	CACCAGCCACCACCTTCTGATAG
pURD Reverse	AAGCAATAGCATGATACAAAGGCATTAAAGCA
pHR Forward	CGTCGTCGTGCTCGTTTAGTGAAC
pHR Reverse	CCAGAGGTTGATTATCGATACGTACGATATCAAG

2.1.10). However, in some cases where ORFs were particularly long, intermediate primers had to be sent. Intermediate primers were designed with a target T_m between 71°C and 74°C, as determined by the online NEB T_m calculator. DNA concentration was determined by measuring absorbance at 260 nm (A^{260}) on a Nanodrop 1000 spectrophotometer (Thermo Fischer Scientific).

For production of plasmid DNA for large scale purposes colonies (selected and confirmed to be correct) were inoculated into 5 mL LB supplemented with ampicillin at 100 μ g/mL. Cultures were incubated for 6–8 h at 37°C in a shaking incubator (180 rpm), prior to dilution 1:1,000 in 3 L of LB culture (six 2 L flasks each containing 500 mL media) supplemented with ampicillin at 100 μ g/mL. Large scale cultures were similarly grown for 16–18 h prior to bacterial pellet harvest by centrifugation at 4,000 $\times$ g for 15 min. Plasmids were purified from bacterial pellets using QIAGEN plasmid giga kits.

2.1.11 Mutagenesis using splicing by overlap PCR

Splicing by overlap extension (SOE) PCR was utilised as a simple method of site-directed mutagenesis (Heckman and Pease 2007) (**Fig 2.1.11**). The initial step required production of overlapping gene segments with the desired mutation engineered into the internal primer. In addition to mutations, internal primers were designed with regions of approximately 20 bases complementary to one another at the 3' end and a region capable of annealing to the template DNA at the 5' end as per the standard primer specification (**Section 2.1.3**). Two overlapping gene segments were created by pairing internal primers with start and end primers including the necessary engineering for

subsequent cloning. The initial PCR proceeded as outlined in **section 2.1.4**. To ensure purity prior SOE PCR the fragments were gel purified (**Section 2.1.5**). Instead of 100 ng of template, the SOE

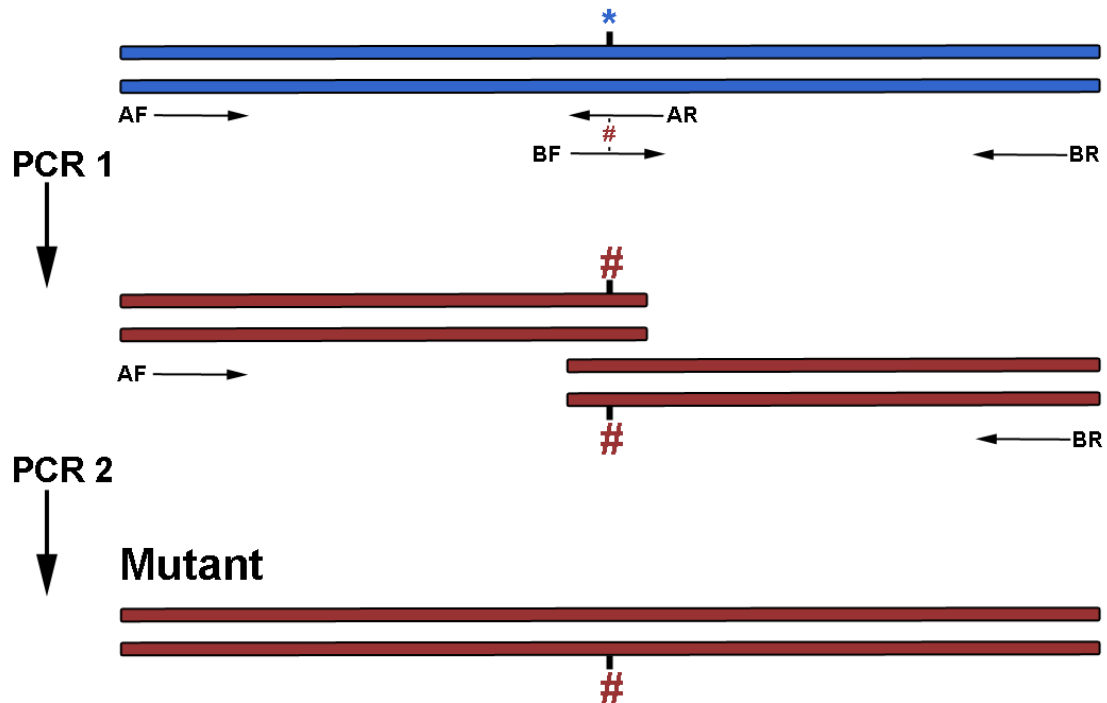


Figure 2.1.11: Splice-by-overlap PCR schematic. The blue template represents the gene of interest encoding the wildtype nucleotide (*). A first PCR produces two overlapping DNA fragments (red), each carrying a copy of the mutation (#), introduced by primers AR and BF. A final PCR step produces a copy of the original template (red) containing the mutation.

Virus	Mutant	Construct		Tag(s)	Primers
		Start	End		
MosV	N230R	157	632	Fc 6xHis	AF: CGTAGCTGAAACC GGTACCT TGTCTGCGGATCGAGGAAA AR: CAGGGATCAGG CG CAGGTCGGTT BF: AACCGACCTG CG CCTGATCCCTG BR: AACAGCACCTCCAG GGTACC AGTGTGGCACTTGTCTGTAC TG
	R268A				AF: CGTAGCTGAAACC GGTACCT TGTCTGCGGATCGAGGAAA AR: GACAGGTGCCC CC CATGGTCAG BF: CTGACCATG CG GGGCACCTGTC BR: AACAGCACCTCCAG GGTACC AGTGTGGCACTTGTCTGTAC TG
	R268N				AF: CGTAGCTGAAACC GGTACCT TGTCTGCGGATCGAGGAAA AR: TACCTGACCATG AAT GGCACCTGTCAG BF: GACAGGTGCC ATT CATGGTCAGGTATGT BR: AACAGCACCTCCAG GGTACC AGTGTGGCACTTGTCTGTAC TG
	Q272R				AF: CGTAGCTGAAACC GGTACCT TGTCTGCGGATCGAGGAAA AR: GTCTGTGGTCCC CG CACAGGTGC BF: GCACCTGTC CG GGACCACAGAC BR: AACAGCACCTCCAG GGTACC AGTGTGGCACTTGTCTGTAC TG
SosV	R242N L243R Y244K H245S	001	582	6xHis	AF: <u>CGTAGCTGAAACCGGT</u> TATGCACGCCAGAAATAGC AR: GTGTGGTCATTGTGTCACGCTGCA AGACTTCCGGTT TGTC BF: CAGCAGCTGAGCCAGACCGACA ACCGGAAGTCTTGC BR: <u>TGATGGTGGTGCTTGGTACC</u> CTTGATCAGGTACAGAAAG GACCGGATG
	R242N L243R Y244K H245S	125	582	8xHis SUMO ID4	AF: <u>CGTAGCTGAAACCGGT</u> AACGCCGTGTGTAGAGGCCCT AR: GTGTGGTCATTGTGTCACGCTGCA AGACTTCCGGTT TGTC BF: CAGCAGCTGAGCCAGACCGACA ACCGGAAGTCTTGC BR: <u>TGATGGTGGTGCTTGGTACC</u> CTTGATCAGGTACAGAAAG GACCGGATG

Table 2.1.11: Mutagenised constructs. Details of mutagenised constructs produced, including residue substitution, construct design and primers used. Restriction sites are highlighted in green, vector overlap for In-Fusion® cloning is underlined.

PCR mixture contained 25 ng of each overlapping gene segment. With the exception of alterations to template, the mixture and reaction proceeded as in **section 2.1.4**. Subsequent cloning, restriction enzyme dependent or In-Fusion® based, continued as outlined in **sections 2.1.5–2.1.10**. Specifics of mutagenised constructs are outlined further in table 2.1.11.

2.2 MAMMALIAN CELL CULTURE

2.2.1 Maintenance of cell lines

A number of mammalian cell lines were utilised for the required experiments (**Table 2.2.1**). All cell culture was performed in a laminar flow cabinet, using sterile reagents and standard aseptic technique. For maintenance cells were cultured in T175 flasks (175 cm² surface area, Corning) at 37°C and 5% CO₂, and were passaged upon reaching ~70–90% confluence, with the exception of HAP1 cells which were passaged at ~50% confluence as cells began to detach at higher densities.

Table 2.2.1 Details of cell culture growth media and maintenance.

CELLS	CELL TYPE	CELL CULTURE MEDIA	DILUTION
HEK293T	Human embryonic kidney	DMEM ^a (Sigma) 10% FCS ^b (Sigma) 1% NEAA ^c (Gibco) 1% L-glutamine (Gibco)	1:10
HAP1	Human chronic myeloid Leukaemia haploid	IMDM ^d (Sigma) 10% FCS (Sigma) 1% NEAA (Gibco) 1% L-glutamine (Gibco)	1:5
BHK21	Baby hamster kidney	GMEM ^e (Gibco) 5% FCS (Sigma) 1% NEAA (Gibco)	1:10

^aDulbecco's Modified Eagle's Medium

^bFetal calf serum

^cNon-essential amino acids

^dIscove's Modified Dulbecco's Medium

^eGlasgow's Minimum Essential Medium

To passage, media was removed from flasks and cells subsequently washed with 10 mL phosphate buffered saline pH 7.4 (PBS, Gibco), and the monolayer detached through 5 min incubation with 5 mL trypsin-EDTA (Gibco). Cells were subsequently resuspended in 20 mL of the appropriate cell culture medium, and a suitable dilution of resuspended cells added to a flask containing 25 mL of fresh culture media (**Table 2.2.1**).

2.2.2 Storage and revival of cell lines

To maintain a diverse set of cells, it was necessary to keep vials of cells at -80°C and/or in liquid nitrogen storage. Cells were prepared for cryo-storage by washing with PBS (Gibco) followed by detachment with 3ml of trypsin-EDTA (Gibco). Following incubation as per **section 2.2.1**, cells were resuspended with a further 5 mL of the appropriate media supplemented with FCS (**Table 2.21**). For storage, cell solution was distributed to cryo-vials, and the cryo-protectant dimethyl sulphoxide (DMSO) added to a final concentration of 10%. Vials were stored in a CoolCell® (Biocision), enabling a controlled reduction in temperature when stored in a -80°C freezer. Cells were then transferred to long-term storage either in a -80°C freezer or liquid nitrogen.

Cell revival required rapid heating to 37°C and removal of the high concentrations of trypsin-EDTA and DMSO. Heated cell solutions were diluted in 5 mL of the appropriate media (**Table 2.2.1**) prior to centrifugation at 1000 x g for 3 min. Media containing DMSO and trypsin-EDTA was removed, cells resuspended in 15 mL of media and transferred to T75 flasks (75 cm² surface area, Corning).

2.3 RECOMBINANT PROTEIN EXPRESSION AND PURIFICATION

2.3.1 Rationale and workflow

Both stable and transient recombinant protein production was performed in human embryonic kidney (HEK) 293T cells due to their fast growth rate, transfectability and ease of handling (Seiradake et al. 2015; Elegheert et al. 2018; Aricescu, Lu, and Jones 2006). Furthermore, as a human cell line, appropriate folding and processing of these proteins from mammalian targeting viruses would be

expected. Processing in HEK293T cells, such as N-glycosylation, was controlled (**Sections 2.3.5 and 2.3.11**), if necessary, to produce homogenous sample for crystallisation (Chang et al. 2007). Transfection of constructs contained within plasmids encoding a secretion signal sequence at the N-terminus enabled easy collection of protein from the cell culture medium. Media was subsequently buffer exchanged and concentrated before application to affinity and size-exclusion chromatography (SEC), typically yielding milligram quantities of pure protein.

2.3.2 Small-scale transient transfection

To determine the suitability of protein constructs for large-scale expression small-scale transient expression tests were performed. This method was also utilised for small scale expression of protein for sialic acid interaction assays (**Section 2.8**), with amounts scaled as necessary. HEK293T cells were inoculated into 24-well plates, 0.5 mL wells, to a density that would yield ~90% confluence following a day of growth in DMEM (supplemented with 10% FCS, 1% NEAA and 1% L-glutamine). Per well, 0.8 µg of purified DNA was diluted in 50 µL of DMEM FCS-free (1% NEAA, 1% L-glutamine) and in parallel 2 µL of Lipofectamine® 2000 (Invitrogen) was similarly diluted. To allow for proper dispersion of DNA and Lipofectamine® 2000 (Invitrogen), samples were incubated for 5 min prior to mixing 1:1 and complexation allowed to proceed for 20 min at room temperature. The Lipofectamine® 2000 cationic liposome and neutral co-lipid formulation, complexes with DNA and facilitates interaction and fusion with the negatively charged plasma membrane, enabling entry of the plasmid (Dalby et al. 2004). The media was exchanged with 400 µL fresh DMEM (10% FCS, 1% NEAA, 1% L-glutamine), and the 100 µL complexed DNA-Lipofectamine® mix added. Cells were incubated at 37°C, 5% CO₂ until ready to harvest following 48–72 h.

Cell culture supernatants and cells dislodged using 0.5 mL of NuPAGE ® MES SDS Running Buffer (1x) (Life Technologies) were collected for small-scale expression analysis. In all tests, both a mock well, where all of the above components were added in the absence of plasmid DNA, and a well transfected with a construct known to be produced and secreted successfully, were used as negative and positive controls, respectively. Western blotting was used to determine the

suitability of a construct for large-scale protein expression, through comparative analysis of secreted protein quantity.

2.3.3 SDS-PAGE

Protein samples were separated according to molecular weight using sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) using 4–12% Bis-Tris PAGE gels (Life Technologies). Prior to loading, samples were mixed with either reducing dye (250 mM Tris-HCl pH 6.8, 10% SDS, 50% glycerol, 25% β -mercaptoethanol, 0.01% bromophenol-blue) or non-reducing dye (125 mM Tris-HCl, pH 6.8, 20% glycerol, 4% SDS, 0.1% bromophenol-blue). Appropriate ladders were loaded in conjunction. Owing to its resolution, BenchMark™ protein ladder (Invitrogen) was used for SDS-PAGE analysis unless otherwise stated. Separation was performed by applying a 160–180 V current for 40–60 min, stopping when the dye front reached the end of the gels. To visualise protein gels were either stained using InstantBlue™ Protein Stain (Expedeon), or western blot performed. Images of stained gels were captured using the ChemiDoc™ system (Bio-Rad).

2.3.4 Western blot analysis

In all cases, the BenchMark™ pre-stained protein ladder (ID: 10748010, Invitrogen) was loaded onto SDS-PAGE gels prior to western blotting, to aid judgement of successful separation and transfer. Blotting was variously performed using dry and wet blotting methods.

2.3.4.1 Wet blotting

For wet blotting, SDS-PAGE gels were assembled within a Mini Blot Module (Life Technologies), with sandwiches consisting of sponge pads, Whatman® Filter Paper (Sigma) and Amersham 0.45 μ m nitrocellulose membrane (GE Healthcare) placed against the gel. When assembling, care was taken to remove air, keep the contents wet with Bolt™ Transfer Buffer (Life Technologies) (Supplemented with 10% ethanol), and to place the nitrocellulose on the anode orientated side of the gel. The Mini Blot Module (Life technologies) was placed in the Mini Gel Tank (Life Technologies) and transfer set up for 70 min at 10 V.

2.3.4.2 Dry blotting

The SDS-PAGE gel was assembled into commercial iBlot™ 2 transfer stacks (Life Technologies), and blotted onto the included nitrocellulose membrane using the associated iBlot™ 2 gel transfer system (Life Technologies). Dry blotting was performed by running the iBlot™ 2 gel transfer system 20 V for 1 min, 23 V for 4 min, and 25 V for 2 min.

2.3.4.3 Anti-histidine tag detection

Nitrocellulose membrane blotted with protein was blocked in 5% milk powder in PBS (w/v) for a minimum of 1 h at room temperature. The blocked nitrocellulose membrane was subsequently incubated with primary antibody. For detection of the histidine tag, the membrane was incubated for 1 h at 21°C with the primary antibody, a mouse-derived monoclonal anti-penta-His antibody (ID: 34660, Qiagen) diluted 1:1000 (v/v) in PBS containing 5% milk (w/v). The membrane was washed five times for three min with PBS containing 0.05% TWEEN20 (v/v), prior to incubation for 1 h at 21°C with the secondary antibody, a horseradish peroxidase (HRP)-conjugated goat-derived anti-mouse IgG antibody (Sigma), diluted 1:2000 (v/v) in PBS containing 5% milk (w/v). Membrane was washed as above before signal development and detection using the Clarity™ Western ECL substrate (Bio-Rad) and the ChemiDoc™ system (Bio-Rad). When using histidine-tag detection, samples were loaded alongside 1 µL His-tagged Benchmark™ protein ladder (ID: 10747012, Invitrogen) for confirmation of successful staining and to enable size determination.

2.3.4.4 Anti-biotin tag detection

Nitrocellulose membrane was incubated for a minimum of 1 h at room temperature in 2.5% Bovine Serum Albumin (BSA) (Sigma) in PBS. Blocked membranes were subsequently incubated for 1 h at 21°C with a HRP-conjugated goat derived anti-biotin antibody (ID: ab19221, Abcam), diluted 1:1000 (v/v) in PBS containing 2.5% BSA (w/v). The membrane was washed five times for three min with PBS 0.05% containing TWEEN20 (v/v), prior to signal development and detection using the Clarity™ Western ECL substrate (Bio-Rad) and the ChemiDoc™ system (Bio-Rad). When using

biotin detection, samples were loaded alongside 10 μ L a biotinylated protein ladder (ID: 7727, CST) for confirmation of successful staining and to enable size determination.

2.3.4.5 *Anti-ephrinB2 detection*

Nitrocellulose membrane blotted with protein was blocked in 5% milk powder in PBS (w/v) for a minimum of 1 h at room temperature. Blocked membranes were subsequently incubated for 1 h at 21°C with an anti-ephrinB2 antibody (ID: ab150411, Abcam), diluted 1:10,000 (v/v) in PBS containing 5% Milk. The membrane was washed five times for three min with PBS 0.05% (v/v) containing tween-20, prior to incubation with the secondary HRP-conjugated goat anti-rabbit IgG (ID: PI-1000, Vector laboratories). The membrane was washed five times for three min with PBS containing 0.05% TWEEN20 (v/v), prior to signal development and detection using the Clarity™ Western ECL substrate (BioRad) and the ChemiDoc™ system (BioRad).

2.3.5 Large-scale transient transfection

For large-scale transient transfection, HEK293T cells were grown in DMEM (10% FCS), in 2,125 cm² area roller bottles (Cellmaster™, Greiner BioOne) at 37°C 5% CO₂. Cells were subsequently transfected with purified plasmid DNA upon reaching ~90% confluence. Prior to preparation of a transfection mixture, 500 μ g of DNA per roller bottle was sterilised using chloroform. 10% by volume chloroform was added to the appropriate amount of DNA require for the number of roller bottles (500 μ g per bottle), mixed vigorously and centrifuged for 10 min at 13,000 x g. Sterilised DNA was collected from the upper-phase of liquid, mixed in a 1:2 ration with of polyethylenimine (PEI, Polyscience Inc.) in FCS-free DMEM, (50 mL per roller bottle), and allowed to complex at room temperature for a minimum of 10 min. PEI complexes with and condenses nucleic acid into positively charged particles capable of interacting with the negatively charged plasma membrane, leading to endocytosis by the cell (Boussif et al. 1995). Meanwhile, roller bottle media was replaced with 200 mL of fresh DMEM (2% FCS). If proteins were intended for crystallography the class I mannosidase inhibitor, kifunensine, was added at a concentration of 5 μ M to the DNA-PEI mixture following complexation (Elbein et al. 1990; Chang et al. 2007). Addition of kifunensine results in

the presentation of high-mannose glycoforms, typically Man₉GlcNAc₂, amenable to cleavage by endoglycosidase F₁ (Grueninger-Leitch et al. 1996; Maley et al. 1989). Cells were incubated in a rolling cabinet for a further 72–96 h at 37°C, 5% CO₂ to allow for recombinant protein expression.

2.3.6 Production of stable cell lines

Production of protein from stable cell lines has the advantage of reliability and cost effectiveness, and despite the time required for selection of successfully transduced cells lines, this method offers the opportunity for tuneable and time economic expression once established (Chaudhary et al. 2012). This method has proved particularly useful for the production of poorly expressing protein, with reported improvements in yield (Elegheert et al. 2018). To this end, two methods were used for the production of stable cell lines, site specific recombination (Seiradake 2015), a reliable yet time consuming method, and the newly established lentiviral method of transduction (Elegheert et al. 2018).

2.3.6.1 Site specific recombination

The mechanism by which *Streptomyces* bacteriophage ϕ C31 uses an integrase to insert into the host genome has been exploited for subsequent manipulation of mammalian genomes (Keravala and Calos 2008; Seiradake 2015; Büssow 2015). This integrase is capable of recombining attachment sites, attB and attP. Co-transfection of a plasmid encoding the integrase and the pURD plasmid with the gene of interest (GOI) and an attB site (**Fig 2.1.2**) enables successful gene integration as mammalian genomes encode endogenous pseudo attP sites (Seiradake 2015).

To generate a HEK293T based stable cell line, cells were grown to ~90% confluency in 2 mL of DMEM (10% FCS, 1% NEAA, 1% L-glutamine), prior to co-transfection with pURD encoding SP-8xHis-SUMO-**SosV-RBP**-1D4 and a plasmid encoding the ϕ C31 integrase (pgk- ϕ C31), as described previously (**Section 2.3.2**). A day following transfection, cells were dislodged using 300 μ L trypsin-EDTA and transferred to a T75 flask containing 15 mL DMEM (10% FCS, 1% NEAA, 1% L-glutamine) (**Table 2.2.1**). Cells were incubated at 37°C, 5% CO₂, until 100% confluency when puromycin was added at a concentration of 2 μ g/mL. The following day, the media

was replaced with 10 mL DMEM (10% FCS, 1% NEAA, 1% L-glutamine) and puromycin at a concentration of 1 µg/mL. When individual cell clumps were observed to be growing on the flask surface, cells were maintained as outlined in section 2.2.1, with media supplemented with 1 µg/mL puromycin. Following analysis of successful soluble protein expression using western blotting, stocks of cells were frozen (**Section 2.2.2**) and/or expanded for large scale protein production.

2.3.6.2 *Lentiviral transduction*

The nature of lentiviruses to integrate into host genomes has been manipulated for production of recombinant protein in mammalian cell lines (Bandaranayake et al. 2011). Utilising a lentivirus producer cell line (HEK293 Lenti-X), lentivirus can be produced through co-transfection of transfer (encoding GOI), envelope and packing plasmids. The transfer plasmid consisting of the cis-acting HIV components the polypurine tract (PPT), Rev response element (RRE) and the psi (ψ) packing signal, and GOI flanked by the 5' and 3' long terminal repeats (LTRs), is stably integrated into the host-cell genome as proviral DNA. Subsequently produced lentiviral particles containing the GOI are capable of infected and transducing a target cell line, resulting in stable cell line production.

HEK293T Lenti-X cells were cultured in a T75 flask at 37°C, 5% CO₂, and upon reaching >90% confluency were co-transfected with a mix consisting of the transfer (pHR-CMV-TetO₂ encoding SP-8xHis-SUMO-**SosV-RBP**-AviTag™), packing and envelope plasmids, mixed 1:1:1 to a total of 30 µg in 250 µL DMEM FCS-free (1% NEAA, 1% L-glutamine). In parallel, an appropriate amount of PEI (1:2.5 DNA:PEI ratio) was prepared in 250 µL DMEM FCS-free (1% NEAA, 1% L-glutamine). Solutions were incubated at 21°C for 5 min, prior to mixing and a further incubation for 20 min. Subsequently, the media was discarded from the HEK293T Lenti-X cells and replaced with 12 mL DMEM (2% FCS, 1% NEAA, 1% L-glutamine) containing the DNA-PEI mixture. Cells were incubated at 37°C, 5% CO₂ for 3 days prior to collection of the media.

For infection, a T75 flask of target HEK293T cells was grown to >90% confluency. Media from the HEK293T Lenti-X, lentivirus producing cell line was collected and 6 mL of DMEM (10% FCS, 1% NEAA, 1% L-glutamine) added. Media was filtered through a 0.45 µm filter unit

(Millipore) attached to a Luer-lock syringe prior to addition of 18 μL of polybrene infection reagent (Merck), which reduces the electrostatic repulsion between the viral particles and the negatively charged plasma membrane. The target cell line media was replaced with the lentivirus containing media, and incubated at 37°C, 5% CO_2 . Following incubation for 3 days, the T75 flask was washed 3 times with PBS prior to maintenance as outline in **section 2.2.1**. The target cell line was infected with the lentivirus and transgene successfully integrated, as confirmed through western blotting. Subsequently, cells were expanded for large scale protein production and/or frozen for long-term storage.

2.3.7 Large-scale stable cell line culture

Large-scale culture of stably transfected HEK293T cells, was performed in the same manner as large-scale culture of transiently transfected HEK293T cells. Roller bottles (Cellmaster™, Greiner BioOne) with a surface area of 2,125 cm^2 , were seeded with stably transfected HEK293T cells, cultured in DMEM (10% FCS) and incubated to ~90% confluence prior to media exchange. All media was retained for future purification due to constant expression of protein. Media was exchanged with DMEM (10% FCS), and cells cultured in a rolling incubator for a further 72–96 h at 37°C, 5% CO_2 . When necessary, kifunensine was added to a final concentration of 5 μM , making sure to retain concentration upon media exchange. In addition, to induce production of protein from stable cell lines produced using the lentiviral method, doxycycline was maintained in culture at a concentration of 1 $\mu\text{g/mL}$.

2.3.7.1 Large-scale stable cell culture with *in vivo* biotinylation

Biotinylation of SP-8xHis-SUMO-SosV-RBP-AviTag™ was performed during large-scale protein expression from a stable lentivirus-transduced cell line. Prior to preparation of a transfection mixture, 125 μg (per roller bottle) of *E. coli* biotin ligase (BirA) encoding pDisplay™ (Thermo) plasmid was chloroform purified (**Section 2.3.5**). Addition of a the 15 amino acid Avitag™ enables site-specific biotinylation of protein using the BirA enzyme, facilitating an effective method of protein detection for downstream applications (Fairhead and Howarth 2015). Sterilised DNA was collected from the

upper-phase of liquid, mixed 1:2 with of PEI in FCS-free DMEM, (50 mL per roller bottle), and allowed to complex at room temperature for a minimum of 10 min. D-biotin (ID: B4639, Sigma) was diluted in the DNA–PEI mixture so the final concentration in the rollers bottles (~250 mL) was 100 μ M.

2.3.8 Cell harvesting and diafiltration

To harvest secreted and soluble protein, cell supernatants were collected and centrifuged to remove cell debris, at 4,000 x g for 20 min, prior to filtration using 0.22 μ m vacuum filter units (Millipore). As components of DMEM are known to strip nickel ions from HisTrap™ HP (GE Healthcare) nickel affinity columns, supernatants were first buffer exchanged into 10 mM Tris pH 8.0, 150 mM NaCl (TBS), and concentrated using the ÄKTA flux tangential flow filtration system (GE Healthcare) or QuixStand™ benchtop system (GE Healthcare), fitted with 10 kDa molecular cut-off (MWCO) membranes.

2.3.9 Protein purification

Protein purification typically proceeded with a combination of affinity and size exclusion chromatography. Chromatography was variously performed using ÄKTApurifier, ÄKTABasic, ÄKTAFPLC and ÄKTApure liquid chromatography systems (GE Healthcare), utilising the UNICORN 6 and 7 software packages (GE Healthcare) to programme purification runs.

2.3.10 Affinity chromatography

Affinity chromatography was initially used in order to purify the majority of contaminants from samples prior to secondary purification using size exclusion chromatography (SEC).

2.3.10.1 Nickel-affinity chromatography

Protein constructs tagged with a hexa- or octa-histidine tag were purified using 5 mL HisTrap™ FF (GE Healthcare) nickel affinity columns and ÄKTA liquid chromatography systems. The HisTrap™ FF column was equilibrated in TBS, prior to loading of the concentrated supernatant. Proteins expressing an accessible oligo-histidine tag bound specifically to the immobilised nickel ions, and flowthrough contained proteins unable to bind the resin. As histidine has an imidazole side chain,

competition with this chemical can be used to increase stringency during purification and for elution. Subsequent controlled mixing with a 0–50 mM imidazole enabled the removal of non-specifically bound protein and an increase of imidazole concentration to 250 mM allowed competitive displacement of the specifically bound protein.

2.3.10.2 Protein G-affinity chromatography

Protein constructs conjugated to a Fc tag were purified using 5 mL and 1 mL HiTrap® Protein G HP (GE Healthcare) columns and ÄKTA liquid chromatography systems. Supernatant was concentrated as above for ease of application, and to ensure buffer components would not strip the columns. Concentrated supernatant was loaded onto a HiTrap® Protein G HP column, pre-equilibrated into 20 mM sodium phosphate pH 7.0. The protein G used in affinity purification derived from the group G streptococci bound specifically to proteins expressing an Fc tag. Protein was eluted by washing the column into 0.1 M glycine-HCl pH 2.7 (Akerstrom and Bjorck 1986). Eluted fractions were immediately neutralised using 60 µL of 1 M Tris-HCl pH 9.0 per mL of eluted fraction. Subsequently, protein was exchanged back into standard buffers using size exclusion chromatography (SEC) (**Section 2.3.13**).

2.3.11 Enzymatic deglycosylation

To reduce heterogeneity of protein samples for crystallisation, proteins deglycosylation was performed with the enzyme endoglycosidase F₁ (EndoF₁) (Grueninger-Leitch et al. 1996) at a mass ratio of 1:100 (EndoF₁:Protein), and incubated overnight at 21°C. EndoF₁ cleaves the high-mannose glycoforms resulting from protein expression conducted in the presence of kifunenzine, an α -mannosidase inhibitor (Chang et al. 2007). Cleavage occurs in the N-linked diacetylchitobiose core between the two N-acetylglucosamine (GlcNAc) residues (Maley et al. 1989), resulting in a truncated glycan consisting a single GlcNAc. Subsequent to deglycosylation, new protein constructs were subject to SEC (**Section 2.3.13**) and SDS-PAGE (**Section 2.3.3**) analysis.

2.3.12 3C cleavage

Constructs cloned into the modified pHLsec vector encoding a SUMO-tag were expressed with an N-terminal octa-histidine tag, SUMO-tag and Human Rhinovirus (HRV)-3C protease cleavage site (Chang et al. 2015). To remove the SUMO-tag, purified protein was incubated with glutathione S-Transferase (GST)-tagged HRV-3C protease at a 1:10 (3C:protein) molar ratio, overnight at 21°C. To obtain a homogeneous sample of successfully cleaved protein, sample was passed over a 5 mL HisTrap™ FF column, taking care to collect the unbound flowthrough. The specifically bound protein, with the N-terminal tags, was washed from the column using 250 mM imidazole. Comparison of the two peaks was performed using anti-histidine western blot and SDS-PAGE analysis. Subsequently, cleaved target protein was purified using SEC, to ensure a homogeneous crystallisation sample.

2.3.13 Size-exclusion chromatography

SEC was utilised to ensure purity of samples for crystallography and other applications, determine approximate size and therefore oligomerisation state, and for buffer exchange of samples. Depending on the expected size of constructs, both Superdex™ 200 Increase 10/300 GL and Superose 6 10/300 columns were used. Columns were loaded onto ÄKTA chromatography systems (GE Healthcare) so the UV absorbance (mAU) during elution could be tracked. All purifications were performed using TBS, unless otherwise stated. Samples were concentrated using Millipore™ Amicon™ centrifugal filter units with a 10 kDa MWCO in order to load them onto a 500 µL sample delivery loop. Prior to loading all samples were filtered using Millipore™ Ultrafree® 0.22 µm centrifugal filter units. SEC was performed at a flow rate of 0.5 mL/min. Fractions from SEC were assessed using SDS-PAGE and western blot analysis and subsequently pooled. SEC fraction analysis was often performed in parallel to analysis of previous steps of the purification, i.e. affinity purification, enzymatic digestion, etc.

2.3.14 Protein concentration determination

Throughout purification protein concentration was determined to keep track of yield and potential steps resulting in protein loss, e.g. concentration steps. Protein concentration was estimated following measurement of absorbance of light at 280 nm using a Nanodrop 1000 (Thermo) Spectrophotometer. Adhering to the Beer-Lambert law ($\text{Absorbance (A)} = \text{molar extinction coefficient } (\epsilon) \times \text{molar concentration (c)} \times \text{path length (l)}$) the concentration of individual protein samples could be determined upon calculation of their unique extinction coefficient (Grimsley and Pace 2004). The molar extinction coefficient describes the ability of the protein samples to absorb light at 280 nm is dependent on its tryptophan, tyrosine, phenylalanine and disulphide bond composition, and was calculated for individual protein constructs and their tags using the ‘Protparam’ tool (Wilkins et al. 1999).

2.4 PROTEIN CRYSTALLISATION

Protein structure solution using X-ray diffraction typically necessitates the production of good quality protein crystals, such crystals preferentially form from high concentration solutions of pure and homogeneous sample (McPherson and Gavira 2013). To achieve this, protein samples were deglycosylated with EndoF₁ (**Section 2.3.11**), purified by size exclusion chromatography (**Section 2.3.13**), concentrated and filtered prior to crystallisation. Following purification, the sample was concentrated with a Millipore™ Amicon® Ultra centrifugal filter, in all cases a 10 kDa MWCO filter was appropriate. Prior to crystallisation, proteins were filtered with Millipore™ Ultrafree® 0.22 µm centrifugal filter units to remove potential aggregates. To determine an optimum protein crystallisation concentration a pre-crystallisation test (PCT™, Hampton) was performed.

Utilising nanolitre scale sitting-drop vapour-diffusion, drops consisting of 100 nL protein and 100 nL reservoir solution were dispensed into Greiner CrystalQuick® plates by a crystallisation robot (Cartesian Technologies) (Walter et al. 2003). Plates were sealed, enabling the drop to equilibrate against the precipitant reservoir solution, increasing precipitant and protein concentration

(McPherson and Gavira 2013). Plates were subsequently stored in a Formulatrix® system at either 21°C or 4°C, and automatically imaged (Walter et al. 2005).

Proteins were initially screened for crystal formation using in-house conditions (Division of Structural Biology, University of Oxford), adapted from commercial screens available from Molecular Dimensions and Hampton Research (Walter et al. 2005). Crystals formed from initial screens were optimised by varying precipitant concentration, protein concentration, seeding and through the use of additives. The specifics of successful protein crystallisation experiments are outlined in the relevant result chapters.

2.5 STRUCTURE DETERMINATION

2.5.1 Crystal harvesting and cryo-protection

Crystals were harvested from Greiner CrystalQuick® plates using Litholoops™, and transferred into a cryo-protectant solution containing either glycerol, ethylene glycol or 2-Methyl-2,4-pentanediol (MPD). Crystals protected using glycerol or ethylene glycol were transferred to a solution consisting of 25% cryo-protectant in mother liquor. Crystal hits obtained from Morpheus® (Molecular Dimensions) screens containing precipitant mix 4 (25% v/v MPD; 25% PEG 1000; 25% w/v PEG 3350) were transferred to a mother liquor solution supplemented with 10% MPD. Many of the crystals obtained were harvested with the help of Dr. Karl Harlos.

2.5.2 X-ray data collection and processing

Data were collected at beamlines 102, I04, I04-1 and I24 at the Diamond Light Source (DLS), UK. Diffraction data were collected at 100 K using a range of Dectris detectors utilised by DLS beamlines between 2015 and 2018.

To achieve successful constructive interference resulting in measurable X-ray diffraction, experiments were performed in accordance with Bragg's law (Bragg William and Bragg William 1913):

$$n\lambda = 2d\sin\theta$$

Where n is an integer, λ refers to the wavelength of the incident X-ray, θ , to the scattering angle, and d , to the inter-plane spacing of the crystal lattice. Reflection positions (θ and d) and intensities were measured and processed using the xia2 suite of programmes (Winter 2010), which index, integrate and scale data. Utilising the measured reflection positions and knowledge of detector geometry, data were assigned Miller indices enabling definition of unit cell parameters. Defined crystal and instrument parameters were refined, and reflection intensities were subsequently converted to structure factor amplitudes.

Data were initially processed automatically by xia2 with the data resolution limit determined by standard criteria ($CC_{1/2} > 0.5$, merged $I/\sigma I > 1$, unmerged $I/\sigma I > 0.25$). By monitoring the values for resolution dependent statistics, such as correlation coefficient ($CC_{1/2}$) (Karplus and Diederichs 2012), signal to noise ratio ($I/\sigma I$), completeness, multiplicity and $R_{\text{merge}}/R_{\text{pim}}/R_{\text{meas}}$, it was determined whether automatic resolution limits imposed by xia2 were appropriate. Subsequently, data were often merged with multiple data sets and reprocessed with the application of manually defined parameters within the xia2 software. The details of processing procedures will be covered in the relevant results chapters.

2.5.3 Phase determination

Calculation of an electron density map from X-ray diffraction data requires the structure factor amplitudes and the phase information of each reflection (McPherson and Gavira 2013). The structure factor amplitudes are calculated from measured reflection intensities, whereas information about reflection phase is lost during an X-ray diffraction experiment. However, phases can be estimated, utilising both experimental methods, which take advantage of added or intrinsic heavy atoms in the crystal, or computationally using molecular replacement (MR), by applying the phases of a related structure (Taylor 2010; McPherson and Gavira 2013).

Initially, MR was performed with the programme Phaser (McCoy et al. 2007), in order to attempt to derive phases for novel data using those from similar structures. Potential MR models were selected from the protein data bank (PDB) based upon sequence identity, and subsequently

edited to isolate the protein core using COOT (Emsley et al. 2010), reducing the potential for clashes impacting structure solution. Solutions were assessed according to the generated F_o-F_c and $2F_o-F_c$ density maps, log likelihood gain (LLG) value and Z-score.

Structure solution of MosV-RBP using molecular replacement was unsuccessful, therefore experimental methods of phase determination were employed. MosV-RBP crystals were soaked with potassium tetrachloroplatinate (K_2PtCl_4) with the help of Dr. Karl Harlos. A solution comprising the precipitant and 5mM K_2PtCl_4 was added to the crystal containing droplet, and incubation allowed to proceed for 3 hr prior to cryo-cooling in a precipitant, K_2PtCl_4 solution containing 25% glycerol. Incorporation of the heavy metal platinum into the crystal enabled phase determination using single isomorphous replacement with anomalous scattering (SIRAS). The location of K_2PtCl_4 within the crystal was determined by exploiting the differences in the structure factor amplitudes of native and derivatised MosV-RBP data from isomorphous crystals. The resulting phase ambiguity was broken using anomalous scattering methods. The derivatised data were collected at the LIII absorption edge for platinum ($\lambda = 1.072 \text{ \AA}$), resulting in a breakdown of Friedel's law, which states that a reflection and its inverse have the same intensity and opposite phases (McPherson and Gavira 2013). Further details of phase determination are provided in the relevant chapters.

2.5.4 Model building and refinement

Iterative rounds of building and refinement were performed to improve electron density maps. Real-space model building and refinement was performed using maps generated by phenix.refine in COOT (Emsley et al. 2010). Refinement in reciprocal-space was performed using phenix.refine (Afonine et al. 2012), a component of the PHENIX suit of programmes (Adams et al. 2010). Where relevant, non-crystallographic symmetry (NCS) restraints were employed throughout, and translation-libration-screw (TLS) parameters employed for later rounds on refinement. Structure quality and the suitability of the refinement strategy employed were assessed using the statistics, R_{free} and R_{work} , which describe the agreement between the model and data.

2.6 STRUCTURE ANALYSIS

2.6.1 Molecular visualisation and presentation

Structures presented were obtained within this project or sourced from the PDB. Visualisation of molecules and subsequent graphics production was performed using the PyMOL Molecular Graphics System (Schrödinger, available online: www.schrodinger.com/pymol) and UCSF Chimera (Pettersen et al. 2004). Dimer association angle was calculated in UCSF Chimera using the ‘Structure Measurements’ functionality. Calculation of surface electrostatic potential was performed using the Adaptive Poisson-Boltzmann server, APBS (Jurrus et al. 2018).

2.6.2 Structural phylogenetic analysis

For structural phylogenetic analysis, RBP six-bladed β -propeller monomers were prepared by removal of water molecules, ligands and protein chains extending from the canonical fold. Structures were applied to the Structural Homology Program, SHP (Stuart et al. 1979), generating pairwise evolutionary distance matrices which were subsequently used to generate un-rooted phylogenetic trees in PHYLIP (Felsenstein 1989).

2.7 BIOPHYSICAL CHARACTERISATION

2.7.1 High performance liquid chromatography small-angle X-ray scattering

Small-angle X-ray scattering (SAXS) was used to characterise the properties of glycosylated MosV-RBP and NarV-RBP crystallisation constructs in solution. Here, the shape and mass of a protein was described by the angle of scattered intensity (Rambo and Tainer 2013b). Data were collected on beamline B21, at the DLS, UK, configured to measure across the scattering vector range $0.0032 \text{ \AA}^{-1} < q < 0.38 \text{ \AA}^{-1}$. For high performance liquid chromatography (HPLC) mode, a $55 \mu\text{L}$ sample at a concentration of 5 mg/mL was loaded onto a Superdex™ 200 PC 3.2/30 (GE healthcare). Buffer TBS, was washed over at a rate of 0.075 mL/min . The SAXS instrument was coupled directly inline enabling efficient measurement of the solution as it eluted (Rambo 2017). Buffer subtraction, peak range selection and data reduction were performed using the Software package ScÅtter

(www.bioisis.net). SAXS profiles were calculated for MosV-RBP and NarV-RBP structures using the fast X-Ray scattering (FoXS) server, and matched to SAXS experimental profiles. Determination of the success of the fit was demonstrated by assessment of the χ^2 statistic.

2.8 SIALIC-ACID INTERACTION ANALYSIS

2.8.1 Hemadsorption and neuraminidase activity analysis

To understand whether uncharacterised paramyxoviral RBPs were capable of binding and hydrolysing sialic acid from glycoconjugates, hemadsorption and neuraminidase experiments were performed. HEK293T cells were transfected with full-length RBP genes, encompassing the intraviral portion, transmembrane domain, stalk, head domain, and vector derived hexa-histidine tag as described previously (Section 2.3.2), except the amount of DNA and volume of Lipofectamine® 2000 were scaled to 1 mL wells. Experiments were performed 18 h following transfection. All experiments were normalised to a NDV-RBP positive control and the relative level of protein cell surface expression, as assessed by enzyme-linked immunosorbent assay (ELISA).

2.8.1.1 Hemadsorption assay

Binding of purified RBPs to terminal sialic acid residues presented on the surface of red blood cells (RBCs) was measured using a hemadsorption assay (Bousse et al. 1994). RBP expressing HEK293T monolayers were washed on ice with PBS pH 7.4 containing MgCl₂ and CaCl₂ (Sigma), prior to incubation with a 2% solution of sheep blood (Thermo Scientific Oxoid) diluted in PBS (MgCl₂, CaCl₂) at 4°C for 30 min. Cells were washed four times whilst on ice using PBS (MgCl₂, CaCl₂), careful to only remove unabsorbed erythrocytes and not to disturb the monolayer. The remaining cells were lysed using 50 mM Tris pH 7.4, 5 mM EDTA, 150 mM NaCl and 0.5% NP-40. Absorbance at 540 nm was measured using a CLARIOStar plate reader (BMG Labtech).

2.8.1.2 Neuraminidase activity assay

The ability of the paramyxoviral RBPs to hydrolyse sialic acid from glycoconjugates was determined using a neuraminidase activity assay (Potier et al. 1979). The sialic acid analogue, 2'-(4-

methylumbelliferyl)- α -D-*N*-acetylneuraminic acid (MU-Neu5Ac), releases the fluorescent 4-methylumbelliferyl upon hydrolysis. Monolayers of HEK293T cells expressing full-length RBPs were washed and dislodged with PBS pH 7.4 (Gibco), counted and seeded at a density of 25,000 cells per mL into a 96-well V-bottom plate (Greiner Bio One). Cells were pelleted with centrifugation at 1,500 rpm for 5 min, before aspiration into a black non-transparent 96-well flat-bottomed plate (Nunc) with 0.1M sodium acetate pH 6.0 containing 1 mM MU-Neu5Ac. The mixture was incubated for 1 h at 37°C and the reaction stopped by addition of 0.25 M glycine buffer pH 10.7. Free 4-methylumbellifrone levels could be fluorimetrically determined at 365 nm for excitation and 450 nm for emission using a CLARIOStar plate reader (BMG Labtech).

2.8.1.3 Cell-surface expression enzyme-linked immunosorbent assay

An ELISA was performed to determine relative levels of cell surface protein expression across HEK293T monolayers (Bousse et al. 1994). Transfected HEK293T cells were washed and dislodged with PBS pH 7.4 (Gibco), counted and seeded at a density of 25,000 cells per mL into 96-well Nunc-immuno™ PolySorp® plates (Sigma), and bound overnight at 4°C. Supernatant was removed and cells fixed in 4% paraformaldehyde for 15 min at 21°C. Cells were washed three times with PBS, prior to blocking with 5% milk powder in PBS (w/v) for a minimum of 1 h. The hexa-histidine tag at the C-terminus of the protein was bound with a rabbit anti-6xHis-tag antibody (ab9108) diluted 1:1000 in PBS containing 5% milk for 1 h at 21°C. Cells were repeatedly washed prior to addition of HRP-conjugated goat anti-rabbit IgG (ID: PI-1000, Vector laboratories), for 1 h at 21°C. Cells were repeatedly washed prior to addition of TMB substrate kit (Pierce™). Stop solution (2 M sulphuric acid) was added following a 15 min incubation at room temperature, and absorbance was read at 430 nm using a using an Infinite F200 plate reader (TECAN).

2.8.2 Warren test

The warren test is a thiobarbituric method to determine free sialic acid concentration (FSA) (Warren 1959), which takes advantage of the formation of β -formyl pyruvic acid following oxidation of sialic acid. β -formylpyruvic acid reacts with thiobarbituric acid to produce a pink coloured product, which

can be measured fluorimetrically at $\lambda_{\text{ex}} = 555/\lambda_{\text{em}} = 585$ nm. Two slightly differing methods were utilised, depending on the reagents available.

2.8.2.1 *Warren test using purified protein*

Free sialic acid (FSA) concentration was determined following incubation of purified protein of interest and the heavily sialidated protein, fetuin (Spiro 1960). Potential neuraminidases were incubated with 10 μM fetuin (Sigma) for 2.5 h at 37°C, 5% CO_2 in PBS, pH 7.4 (Gibco). Protein was precipitated using a 2% trichloroacetic acid solution followed by centrifugation at 14,000 x g. Supernatants were oxidized, and reacted with dye containing thiobarbituric acid, resulting in a pink coloured product if sialic acid was present. Samples were aliquoted into a black non-transparent 96-well flat-bottomed plate (Nunc) before reading using a CLARIOStar plate reader (BMG Labtech). Concentration could be determined by plotting signal against a standard curve determined by samples containing a known concentration of sialic acid.

2.8.2.2 *Warren test using small-scale expressed protein*

If purified recombinant protein was not available, a method only requiring the small-scale transient expression of protein was utilised. Histidine-tagged constructs to test were transfected and expressed in 1 mL volumes, as outlined in **section 2.3.2**. 96-well Nunc-immuno™ PolySorp® plates (Sigma) plates were incubated with a 50 μL PBS, pH 7.4, solution containing mouse-derived monoclonal anti-5xHis antibody (ID: 34660, Qiagen) diluted 1:200, one day prior to harvesting small-scale transiently expressed protein. Following 72 h incubation at 37°C, 5% CO_2 , supernatants were collected and 100 μL added to anti-His antibody coated plates, and incubated for 2 h at 21°C. Plates were washed prior to addition of 50 μL of 50 μM fetuin (Sigma) diluted in PBS. Samples were incubated for 18 h at 37°C, 5% CO_2 , prior to continuing with the protocol as above (**Section 2.8.2.1**). Successful expression of soluble protein was confirmed by applying supernatant to an SDS-PAGE gel, performing a western blot with anti-histidine detection, as detailed in **sections 2.3.3 and 2.3.4**.

2.9 NOVEL RECEPTOR IDENTIFICATION

A number of techniques were used in order to attempt to identify novel carbohydrate or protein receptors. Despite negative results thus far, these methods are detailed below.

2.9.1 Glycan array analysis

Interactions between paramyxovirus hemagglutinin-neuraminidase receptor binding proteins (HN-RBPs) and sialic acid have been well characterised (Connaris et al. 2002; Crennell et al. 2000; Langedijk, Daus, and VanOirschot 1997; Iorio et al. 2001; Mirza, Deng, and Iorio 1994). However, further RBP interactions with other carbohydrate species have not been identified, despite this being possible (Raman et al. 2016). Utilising a microarray based assay developed commercially by RayBiotech, MosV (Glu202–Thr632), NarV (Glu184–Pro657), JPV (Glu149–N709) RBPs were screened for binding to a panel of 100 glycans (<https://www.raybiotech.com/glycan-array-100/>). This was done in parallel to testing of the positive controls, NDV-RBP (Ser47–Gly577) and glycine max soybean lectin (ID: L1395, Sigma).

Prior to application to the glass slide based microarray, it was necessary to biotinylate samples for biotin detection downstream. Samples were dialysed into PBS, pH 8.0, for 6 h at 4°C with constant stirring, changing the 500 mL of buffer once during this time. Proteins were labelled non-specifically by incubation of 30 µg of dialysed sample at room temperature with 3.3 µL of the 1x labelling reagent provided in the Glycan Array 100 kit (RayBiotech). Subsequently, 3 µL of ‘Stop Solution’ was added to the mixture and samples dialysed as above.

Glycan sub-arrays were blocked prior to application of biotinylated sample by application of 400 µL of ‘Sample Diluent’ and subsequent incubation at room temperature for 30 min. Particulates were removed from biotinylated samples by centrifugation at 10,000 rpm for 5 min. Sample was diluted to 10 µg/mL in 400 µL of ‘Sample Diluent’, and applied to the blocked sub-arrays. Slides were incubated at room temperature for 3 h. Following incubation, sub-arrays were washed 5 times with 800 µL of 1x ‘Wash Buffer I’ and 2 times with 800 µL 1x ‘Wash Buffer II’, 5 min each, taking care to remove the buffer following each wash step. Meanwhile the 1x Dye-

conjugated Streptavidin was prepared by first adding 1 mL 'Sample Diluent' to the Cy3 equivalent dye-conjugated streptavidin tube, creating a stock solution. The stock solution was diluted 400 μ L in 1.4 mL of 'Sample Diluent'. Subsequently, 400 μ L of 1x Cy3 equivalent dye-conjugated streptavidin was added to each sub-array well, the slide covered with foil and incubated at room temperature for 1 h. Samples were decanted from each sub-array well, which were promptly washed 5 times with 800 μ L 1x 'Wash Buffer I', for 5 min at room temperature. The whole slide was then placed within the slide washer dryer and washed with 30 mL 1x 'Wash Buffer I' for 15 min, 30 mL 1x 'Wash Buffer II' for 5 min, and finally washed with 30 mL ultrapure water, before thorough, yet gentle drying. Slides were sent to Raybiotech Inc. (Peachtree Corners, GA, USA), where scanning and data extraction services were performed.

2.9.2 Pull-down experiments

Pull-down biochemical methods of receptor hunting utilise a 'bait' protein that can bind and isolate interacting partners from a host-cell. This method has been successful in identifying the SARS-CoV receptor angiotensin-converting enzyme 2 (ACE2) (Li et al. 2003) and ephrinB2 as the receptor utilised by NiV-RBP (Negrete et al. 2005).

2.9.2.1 Preparation of a biotinylated BHK21 membrane protein solution

BHK21 cells were grown to ~90% confluency in 4x T75 flasks before collection, and centrifugation in a 50 mL tube at 500 x g for 3 min. The cell pellet was gently lysed in 500 μ L buffer consisting of 25 mM Tris pH 7.4, 150 mM NaCl, 1% n-dodecyl-b-D-maltopyranoside (DDM) sol-grade (ID: D301S, Anatrace), 1x protease inhibitor cocktail (ID: P8340, Sigma). Solubilisation was aided by ultrasound based sonication using the Bioruptor® Pico (Diagenode). Cells were sonicated for 3 cycles, 30 sec ON/30 sec OFF, at 4°C, before solubilisation proceeded for a further 2 h at 4°C. Insoluble material was removed by centrifugation at 10,000 x g for 15 min. To maintain the solubility of membrane proteins without need of detergents, DDM was exchanged for the biotinylated A8-35 amphipol (ID BAM01, Anatrace). The solubilised solution was supplemented with 5 g of biotinylated A8-35 per g of protein, exceeding the typical binding capacity of membrane proteins

(Popot et al. 2003). The solution was incubated for 3 h at 4°C prior to removal of detergent by incubation of the solution overnight at 4°C with Bio-Beads® SM-2 (ID: 152-3920, Biorad) (Byrne et al. 2016), added at 5 g per 25 mL of solution. The amphipol exchanged solution was collected for application to the pulldown experiment.

2.9.2.2 *MosV-RBP and NiV-RBP Fc pulldown assay*

Fc-tagged MosV-RBP (Thr157–Thr632) and Nipah virus(NiV)-RBP (Gly183–Thr602) (the latter provided by Dr Rhys Pryce) were expressed in HEK293T cells, harvested and purified using protein G affinity purification as per **section 2.3.10.2**. Dynabeads™ Protein G (ID: 10004D, Thermo) magnetic beads were used for capture of Fc-tagged MosV-RBP and NiV-RBP pre-incubated with the solubilised membrane protein solution. However, prior to this the solubilised membrane protein was pre-cleared, with two rounds of incubation with unbound protein G beads, and one round with IgG coated protein G beads. Solubilised membrane protein was added to 80 µL beads, rotated for 30 min at 4°C prior to collection and application to further rounds of pre-clearing. Per 1 mg of beads, 8 µg of IgG was required to coat the beads. A final pre-clearing step against unbound protein G beads was performed prior to immunoprecipitation.

For binding to Fc tagged protein, 50 µL protein G beads per reaction was mixed with 15 µg MosV-RBP Fc diluted in 200 µL PBS. Similarly, Fc tagged NiV RBP was also added to protein G beads. Samples were incubated at room temperature for 10 min, prior to washing with PBS, followed by the application of the pre-cleared solubilised membrane protein. The mixture was incubated on ice for 30 min, with occasional gentle pipetting. The supernatant containing unbound membrane protein was collected and beads were washed 3 times with PBS, prior to elution in 40 µL of 0.1 M glycine, pH 2.7, and subsequent neutralisation with 1.6 µL 1 M Tris, pH 9.

The successful complexation of biotinylated amphipol and testing of bead eluate was performed using SDS-PAGE and anti-biotin dot-blot analysis. 2 µL volumes of sample (BHK21 lysate, supernatant following bead bind, washes 1, 2 and 3, and eluate) to test were applied to a

nitrocellulose membrane, subsequently treated as per **section 2.3.4.4**. A biotinylated molecular weight marker was used as a positive control, and PBS as a negative to confirm successful staining.

2.9.3 SosV-RBP binding analysis for potential haploid screen

2.9.3.1 Flow cytometry

Screening viral protein binding to mutagenised haploid cells enables the study of phenotypes that could be impactful to virus interactions at the early stages of infection (Pillay and Carette 2015). HAP1 cells, derived from the human myeloid leukemia cell line, KBM-7 (Andersson et al. 1995), have been used as tools to identify Neimann-Pick C1 (NPC1) as the receptor for Ebola virus and the lysosomal-associated membrane protein-1 (LAMP-1), as an intracellular receptor for Lassa virus (Jae et al. 2014).

Prior to pursuing haploid screening for the SosV-RBP target, it was necessary to confirm binding to the haploid cells, HAP1, using flow cytometry. HAP1 cells were grown to ~50% confluency in 4x T175 flasks prior to collection. Cells were dislodged using 5 mL PBS, pH 7.4, 1 mM EDTA, and counted prior to centrifugation at 1200 rpm for 5 min at 21°C. Cells were resuspended in an appropriate amount of PBS prior to transferring 1 million cells per well to a V-bottom plate (Greiner Bio One). Cells were centrifuged for 5 min at 1,500 rpm, 21°C, prior to removal of supernatant and resuspension in an Fc block solution consisting of 5 µL per well of Human TruStain FcX™ (ID: 422301, Biolegend), 10 µg/mL mouse IgG (ID:12-371, MerckMillipore), and 10 µg/mL human IgG (ID: I4506, Sigma), diluted in 50 µL fluorescence activated cell sorting (FACS) buffer (PBS, pH 7.4, 1% bovine serum albumin). Following a 5 min incubation and centrifugation for 5 min at 1,500 rpm, the relevant cells were resuspended in solutions of mouse antibody specific to the HAP1 cell marker CD59 (phycoerythrin (PE) tagged ID: 304707, Biolegend; Alexafluor®-647 tagged ID: ab187769, Abcam; Biotin tagged ID: 555762, BD Biosciences) and biotinylated SosV-RBP, with specifics regarding dilutions detailed in the ‘Future perspectives and concluding remarks’ chapter. Incubation was allowed to proceed for 30 min at 4°C, with samples covered with foil. HAP1 cells were washed 3 times with 150 µL FACS buffer, with

centrifugation for 5 min at 1,500 rpm, 4°C, between steps. A second incubation was performed, applying streptavidin tagged with PE diluted 1:3000 in FACS buffer for 20 min at 4°C. Cells were washed 3 times with FACS buffer, before being resuspended in 200 µL FACS buffer.

Samples were analysed using a FACSCelesta™ (BD Biosciences) operated with the FACSDiva software (BD Biosciences). The cytometer was calibrated with CS&T-beads (BD Biosciences) and compensation, where the overlap of the excitation and emission spectra of the fluorochromes corrected. Data were analysed in the software FlowJo (BD Biosciences).

2.9.3.2 *Enzyme-linked immunosorbent assay*

An ELISA was performed as a secondary technique to flow cytometry to analyse SosV-RBP binding to HAP1 cells. HAP1 cells were washed and dislodged with PBS pH 7.4 (Gibco) supplemented with 1mM EDTA, counted and seeded at a density of 25,000 cells per mL into 96-well Nunc-immuno™ PolySorp® plates (Sigma), and bound overnight at 4°C. In parallel a subset of wells was coated with recombinant His-tagged SosV-RBP, at a concentration of 1 µg/mL in PBS. Following overnight incubation, supernatant was removed and cells fixed in 4% paraformaldehyde for 15 min at 21°C. Cells were washed three times with PBS, prior to blocking with 2.5% BSA in PBS for 1 h. Subsets of HAP1 cells were either treated with SosV-RBP or biotinylated CD59 diluted in 2.5% BSA in PBS to a concentration of 5 µg/mL, or left untreated and incubated for a further hour in blocking buffer. For detection of SosV-RBP wells were incubated with a rabbit anti-6xHis-tag antibody (ab9108; Abcam) diluted 1:5000 in PBS containing 2.5% BSA for 1 h at 21°C, before incubation with HRP-conjugated goat anti-rabbit IgG (ID: PI-1000, Vector laboratories), diluted 1:5000 2.5% BSA in PBS, for 1 h at 21°C. For detection of biotinylated CD59 antibody bound to HAP1 cells, wells were incubated for 1 h at 21°C with HRP-conjugated goat derived anti-biotin antibody (ID: ab19221, Abcam) diluted 1:10,000 in 2.5% BSA in PBS. Wells were repeatedly washed prior to addition of TMB substrate kit (Pierce™). Stop solution (2 M sulphuric acid) was added following a 15 min incubation at room temperature, and absorbance was read at 430 nm using a using an Infinite F200 plate reader (TECAN).

CHAPTER 3

THE SOSUGA VIRUS RECEPTOR BINDING PROTEIN

3 THE SOSUGA VIRUS RECEPTOR BINDING PROTEIN

3.1 SUMMARY OF FINDINGS

Sosuga virus (SosV), a bat-borne paramyxovirus, is one of a host of recently identified viruses to be classified within the recently defined *Pararubulavirus* genus, family *Paramyxoviridae* (Amarasinghe et al. 2019). In 2012, a wildlife biologist presenting severe febrile illness was found to be infected with SosV, suggestive that the virus may be a human pathogen (Albarino et al. 2014; Iorio et al. 2001). The pathobiological characteristics of SosV can be, in part, studied by analysis of the SosV receptor binding protein (SosV-RBP), which facilitates host-cell attachment and is a major determinant of tissue and species tropism (Thibault et al. 2017; Zeltina, Bowden, and Lee 2016). Despite the close relation of SosV-RBP to the hemagglutinin neuraminidase (HN)-type RBPs of other paramyxoviruses, SosV-RBP lacks the stringently conserved hexapeptide motif integral to sialic acid binding and hydrolysis (Albarino et al. 2014). In this chapter, I show that while SosV-RBP structurally aligns closely with RBPs with HN functionality, the putative receptor-binding site contrasts those displayed by RBPs with HN functionality. Observed structural alterations across this rationalise hemadsorption and neuraminidase activity analyses, which demonstrate the inability of SosV-RBP to interact with sialic acid. Altogether, this combined structural and functional investigation provides data that indicate that SosV undergoes a non-classical route of host-cell entry, and provides a model by which pararubulaviruses only recently diverged from sialic acid binding functionality.

3.2 INTRODUCTION

SosV was identified following infection of a wildlife biologist, who presented with severe febrile illness upon returning to the US from Uganda and South Sudan (Albarino et al. 2014). Metagenomic studies revealed that this virus was closely related the recently identified ‘rubula-like’ virus, Tuhoko virus 3 (ThkV3), isolated in southern China from Leschenault's rousette bats (*Rousettus leschenaultia*) (Lau et al. 2010). Investigations of bat populations at multiple locations throughout

Uganda over a 3-year period (2009-2012), found a number of Egyptian rousette bats (*Rousettus aegyptiacus*) that were PCR positive for SosV-RBP, indicating this species as a possible reservoir host (Amman et al. 2015). Given the recent identification of putatively functionally similar viruses circulating in Egyptian rousette bat populations in South Africa, and the range of this species across sub-Saharan Africa (Mortlock et al. 2019), concern regarding the transmission of SosV or similar rubula-like viruses to human populations has increased.

‘Rubula-like’ viruses were defined as close genetic relatives to rubulaviruses, such as mumps virus (MuV), whilst having key differences in the amino acid sequences of their RBPs which dictate interactions with the host-cell receptor, sialic acid (Albarino et al. 2014; Baker et al. 2013; Bowden et al. 2001; Burroughs et al. 2015; Chua, Wang, et al. 2002; Lau et al. 2010). The rubulaviruses present a tetrameric HN-type RBP, composed of an intraviral region, transmembrane domain, α -helical stalk, and a HN-type six-bladed β -propeller head domain which utilises a conserved mechanism of sialic acid binding and hydrolysis to facilitate viral entry and escape. Paramyxoviral HN-type RBPs retain the seven sialidase residues Arg₁, Asp₁, Glu₄, Arg₄, Arg₅, Tyr₆, and Glu₆ (sub-script refers to location on blades 1–6 of the β -propeller fold) (Langedijk, Daus, and VanOirschot 1997; Connaris et al. 2002; Crennell et al. 2000; Iorio et al. 2001). Paramyxoviral HN-type RBPs also encode a stringently conserved ‘Asn–Arg–Lys–Ser–Cys–Ser’ hexapeptide motif. Mutagenesis of the hexapeptide motif NDV-RBP resulted in an ablated HN functionality (Mirza, Deng, and Iorio 1994). Whilst, the SosV-RBP encodes for the seven conserved sialidase residues, it only retains the final two residues of the hexapeptide motif (Cys–Ser) (Albarino et al. 2014) (**Fig 3.2**). The absence of this hexapeptide and variation across the conserved residues required for sialic acid binding and hydrolysis has been observed in a number of ‘Rubula-like’ viruses, such as Menangle virus (MenV) (Bowden et al. 2001), Tevot virus (TevPV) (Burroughs et al. 2015) and Tioman virus (TioV) (Chua, Wang, et al. 2002). Furthermore, recent studies have confirmed that sialidase treatment of permissive cells does not ablate infection (Johnson et al. 2019), supporting a sialic acid independent entry paradigm for the pararubulaviruses.

		R ₁	D ₁	Hexapeptide	E ₄	R ₄	R ₅	Y ₆	E ₆
<i>Pararubulavirus</i>	SosV	RFP	CDG SNP	RLYHCS	ESK	RSS	RQGP	YTD	ELD
	ThkPV-3	RFP	CSGSSP	RLYHCS	ESR	RGS	RTGP	YTD	ELD
	AchPV-1	RTP	CNGSSP	VTYQCS	ESQ	RSS	RKGP	YSI	ELD
	ThkPV-2	RNP	CNGSLG	VSRQCS	ESQ	RSA	RTYP	YSV	ELD
	AchPV-2	RFP	CSDISS	FRRGCS	EGR	RST	RQNP	YSE	ELT
	MenV	RFP	CELNQN	PVRTCS	KAQ	QSA	FRKG	DTT	EWG
	TevPV	RFP	CELDSD	QTRGCS	RAQ	QSA	FQRP	STT	EWG
	TioV	RFP	CELDSD	QARGCS	RSQ	QSA	FQKG	STT	EWG
	ThkPV-1	RFP	CSDHGP	WLRSCS	EGQ	RSS	RSNP	YST	ELM
HN-type	<i>Orthoavulavirus</i> NDV	RIP	CRD HSH	NRKSCS	EGR	RGS	RLNP	YTT	EIS
	<i>Metaavulavirus</i> AMPV-2	RIP	CQDASH	NRKSCS	EAR	RSS	RKDP	YST	EYA
	<i>Paraavulavirus</i> AMPV-3	RIP	CQDHAA	NRKSCS	EAR	RSS	RISP	YTT	EAV
	<i>Orthorubulavirus</i> MuV	RIP	CKDHTS	NRKSCS	EGR	RST	RVNP	YTT	ELA
	<i>Respirovirus</i> SeV	RLP	CADIGK	NRKSCS	EGR	RSS	RVNP	YTT	EIN

Figure 3.2: Amino acid sequence alignment. Alignment of the receptor binding protein (RBP) amino acid sequences from Sosuga virus (SosV), Tuhokoviruses 1, 2 and 3 (ThkPV-1/2/3), Achimota viruses 1 and 2 (AchPV-1/2), Menangle virus (MenV), Teviot paramyxovirus (TevPV), Tioman virus (TioV), Newcastle disease virus (NDV), avian paramyxoviruses 2 and 3 (AMPV-2/3), mumps virus (MuV) and Sendai virus (SeV). The seven conserved sialidase residues (Langedijk, Daus, and VanOirschot 1997) and hexapeptide motif (Mirza, Deng, and Iorio 1994) are labelled according to residue and blade location (Langedijk, Daus, and VanOirschot 1997) and annotated above alignments. Genbank accession codes are listed in Appendix D.

Given the variation across the hexapeptide motif and other residues integral to sialic acid interactions, and recent data which supports a sialic acid independent mode of virus entry (Johnson et al. 2019), the genus, *Rubulavirus*, has undergone a reorganisation. *Rubulavirinae*, now constitutes a subfamily containing two genera, *Pararubulavirus* and *Orthorubulavirus* within the *Paramyxoviridae*. The *Orthorubulavirus* genus contains well-characterised sialic acid-interacting viruses, including MuV and parainfluenza virus 5 (PIV5). The genus, *Pararubulavirus*, on the other hand, contains newly identified and divergent viruses, including SosV, MenV, TevPV (Amarasinghe et al. 2019).

Here, I sought to clarify the structure of the unique SosV-RBP, therefore providing initial characterisation of a newly defined genus of the paramyxovirus family with pathogenic potential. In accordance with the discovery of a binding site that is incompatible with known modes of HN-mediated sialic acid recognition, I found that SosV-RBP maintains only a limited capability to interact with sialic acid, *in vitro*. These data provide evidence that SosV utilises a non-classical mode of paramyxovirus entry, in addition to presenting a rationale for the recent distinction of the genus, *Pararubulavirus* (Amarasinghe et al. 2019).

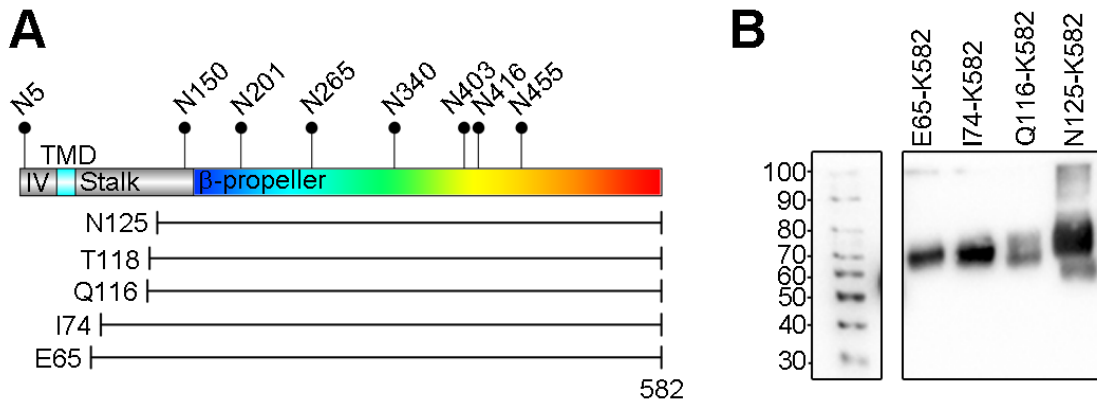


Figure 3.3.1 Construct design and small-scale expression test of SosV-RBP. (a) Gene diagram displaying features of the SosV-RBP, including the intraviral (IV) region, transmembrane domain (TMD), stalk and six-bladed β -propeller receptor binding head (rainbow). Predicted N-linked glycosylation sites (NXT/S, where X $\neq$ P) are marked with pins and relevant asparagine residue numbered. The designed constructs are illustrated below the schematic. (b) Western blot of small scale expression test of constructs tagged with an N-terminal SUMO tag, where an SDS-PAGE gel of HEK293T cell supernatants was performed under reducing conditions. Western blot analysis was performed alongside the Benchmark™ Protein Ladder (Thermo, left).

3.3 PROTEIN PRODUCTION AND STRUCTURE DETERMINATION OF SOSV-RBP

3.3.1 Construct design and small-scale expression testing

SosV constructs were generated from a human codon-optimised gene of the full-length RBP (GenBank NC_025343). In order to determine appropriate construct boundaries required to obtain soluble six-bladed β -propeller head domains for structural and functional characterisation, the SosV-RBP amino acid sequence was aligned against the sequences of the constructs used for structure determination of PIV3-RBP (Lawrence et al. 2004), the PIV5-RBP (Yuan et al. 2005; Welch et al. 2013) and the MuV-RBP (Kubota et al. 2016). Constructs were generated using PCR, with primers specific to unique N-terminal start sites (Glu65, Ile74, Gln116, Thr118, Asn125) and the C-terminal end site, Lys582, and cloned into the pHLsec expression vector (Aricescu, Lu, and Jones 2006).

The ability of the constructs to produce soluble protein was assessed by small-scale expression screening in HEK293T cells. Whilst almost all constructs expressed within the cell fraction (excluding Thr118–Lys582), no constructs were secreted, indicative that the protein was unstable. To aid in solubility and stability, a small ubiquitin-like modifier (SUMO) tag was

introduced to the N-terminus (Peroutka et al. 2008) by sub-cloning constructs into a modified pHLsec vector (Chang et al. 2015). The addition of a SUMO tag resulted in the production of secreted protein (**Fig 3.3.1**). Due to the observed equal expression of the SosV-RBP constructs, the shortest construct, Asn125–Lys582, was selected for large-scale expression, as the presence of an unstructured N-terminus protruding beyond the six-bladed β -propeller head could impede crystallogenesis (Savitsky et al. 2010).

3.3.2 Large-scale expression and purification

Large-scale transient transfection of the SosV-RBP construct Asn125–Lys582 encoding an N-terminal octa-histidine, and SUMO tag fusion was performed in the presence of α -mannosidase type I inhibitor, kifunensine (Chang et al. 2007). Cell culture supernatants were clarified by centrifugation, filtered and subject to buffer exchange into 10 mM Tris pH 8, 150 mM NaCl buffer (TBS). Concentrated and exchanged supernatants were purified using nickel-affinity chromatography using a HisTrap™ FF column (GE Healthcare), and partially deglycosylated with the enzyme endoglycosidase F₁ (EndoF₁), to yield protein bearing single N-acetylglucosamine (GlcNAc) residues (Maley et al. 1989).

Size exclusion chromatography (SEC) was performed on the deglycosylated sample using a Superdex™ 200 Increase 10/300 GL column (GE Healthcare), revealing an absorbance spectrum with two prominent peaks and multiple shoulders (**Fig 3.3.2A**). The peak was analysed by SDS-PAGE, revealing a species corresponding to the ~65 kilodalton (kDa) tagged SosV-RBP (**Fig 3.3.2C**). The SEC indicated that SosV-RBP likely exists as an oligomeric species in solution. Consistent with previous studies of paramyxovirus RBPs (Yuan et al. 2005; Yuan et al. 2008; Bowden, Crispin, Harvey, et al. 2010; Hashiguchi et al. 2007; Lawrence et al. 2004; Yuan et al. 2011), it is possible that the un-partnered cysteine in the N-terminal stalk domain potentially drives homoligomerisation and the formation of the native-like dimer-of-dimers. However, neither an expected homodimeric nor a homotetrameric species conjugated to a SUMO tag corresponds to the ~160 kDa band visualised when non-reduced sample was tested (**Fig 3.3.2C**). Further analyses are

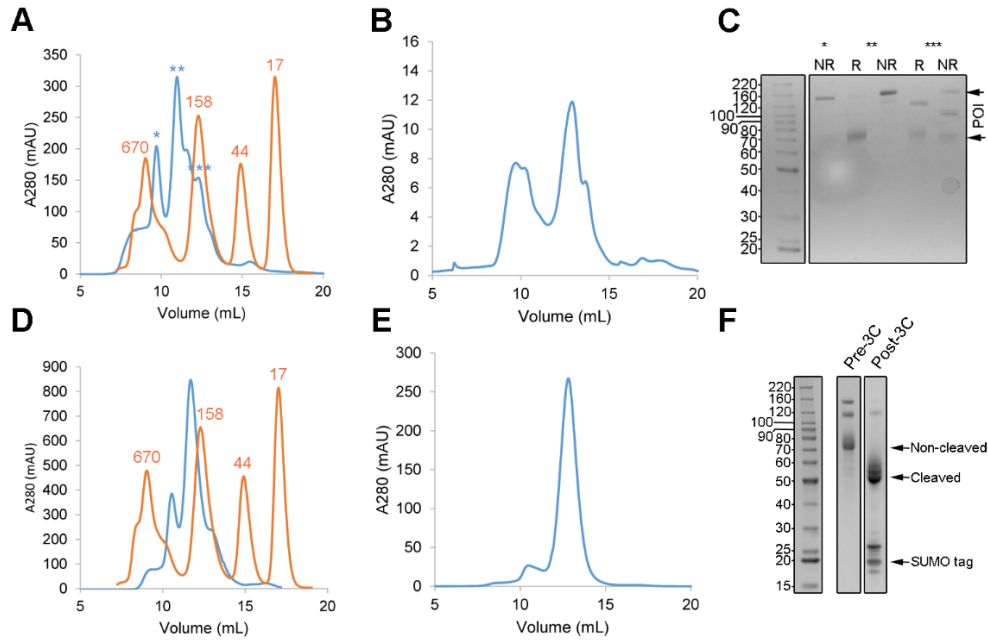


Figure 5.3.2: Purification of SosV-RBP. (a) Size exclusion chromatogram of SosV-RBP expressed transiently in the presence of kifunensine, and treated with endoglycosidase F₁ (EndoF₁), aligned against gel filtration standards (ID: 1511901, BioRad, orange). Blue stars refer to pooled peaks analysed using SDS-PAGE analysis. (b) Size exclusion chromatogram of SosV-RBP following 3C cleavage of the N-terminal 8xHis and SUMO tag. Poor yields resulted in a change of strategy. (c) SDS-PAGE analysis of SosV-RBP purification prior to cleavage of tags with protein of interest (POI) highlighted, indicating the presence of oligomeric species composed of ~65 kDa SosV-RBP monomers. Pooled peaks (*) run under reducing and non-reducing conditions. (d) Size exclusion chromatogram of SosV-RBP expressed stably in the presence of kifunensine, and treated with EndoF₁, aligned against gel filtration standards (BioRad, orange). (e) Size exclusion chromatogram following cleavage of tags. (f) SDS-PAGE analysis of SosV-RBP pre- and post-3C cleavage, revealing a final sample with some minor tag contamination. In both panel C and F, samples were run alongside Benchmark™ Protein Ladder (Thermo, left), and ran under reducing conditions.

required to determine the oligomeric status of the protein in solution (e.g. multi-angle light scattering (MALS)). The peak fractions were pooled, excluding sample within the void, and treated with human rhinovirus (HRV) 3C protease overnight. Successfully cleaved sample was separated from residually-tagged protein and subject to SEC (**Fig 3.3.2B**). The SosV-RBP peak was not monodisperse and neither was there sufficient protein yield required to perform SDS-PAGE analysis and crystallisation screening. Therefore, utilising site-specific recombination, a stable cell line was produced to enable more efficient production of larger quantities of SosV-RBP.

In order to produce a stable cell line, the SP-8xHis-SUMO-SosV-RBP(Asn125–Lys582)-ID4 cassette was sub-cloned into the pURD vector (Seiradake 2015). The plasmid was co-transfected with a pDisplay vector encoding the ϕ 31C integrase, in order to enable host-cell integration and constitutive expression. Stably expressing cells were expanded to large-scale production and protein purified from supernatant described above (**Fig 3.3.2D-F**). Whilst yields were greatly improved, the protein precipitated during concentration, resulting in limited initial crystallisation screening.

3.3.3 Crystallisation and X-ray diffraction data collection

An appropriate SosV-RBP crystallisation concentration of 2.1 mg/mL, was determined using a pre-crystallisation test (PCT™, Hampton). SosV-RBP purified protein crystallised following 25 d in 0.2 M magnesium acetate tetrahydrate, 0.1 M sodium cacodylate pH 6.5, 20% PEG 8000. The crystal was immersed in a cryo-protectant solution, comprising 20% glycerol diluted in reservoir solution. X-ray data collection was performed at a wavelength of 0.9795 Å on beamline I04, Diamond Light Source (DLS), Didcot, UK. Reflections were indexed, integrated and scaled using the xia2 package (Winter 2010). Data processing statistics are presented in **Table 3.3.3**.

Table 3.3.3: Data collection and refinement statistics for SosV-RBP

SosV-RBP	
Data collection statistics	
Beamline	DLS I04
Spacegroup	$P\ 1\ 2_1\ 1$
Cell dimensions	
a, b, c (Å)	75.7, 84.0, 81.7
α, β, γ (°)	90.0, 112.9, 90.0
Resolution range (Å)	51.60–2.50 (2.54–2.50) ^a
Wavelength (Å)	0.9795
R_{merge}	0.08 (0.94)
R_{meas}	0.08 (1.02)
R_{pim}	0.03 (0.39)
I/σ I	13.5 (1.5)
CC_{1/2}	1.00 (0.80)
Completeness (%)	99.8 (98.9)
Multiplicity	6.7 (6.5)
Refinement statistics	
Resolution (Å)	43.44–2.50 (2.59–2.50)
No. reflections	32,699
R_{work} / R_{free}^b (%)	19.5/23.3
No. atoms	
Protein	5,870
Ligand/ion	182
Water	39
B-factors	
Protein	77.6
Ligand/ion	91.2
Water	61.0
R.m.s deviations^c	
Bond lengths (Å)	0.003
Bond angles (°)	0.65
Ramachandran analysis^d	
Favoured region (%)	96.6
Allowed region (%)	3.3

^aNumbers in parentheses refer to the outer resolution shell.

^bR_{free} is calculated as for R_{work}, but using only 5% of the data which were separated prior to refinement.

^cr.m.s deviations: root mean square deviation from ideal geometry.

^dDetermined using the Molprobit server (Chen et al. 2010).

3.3.4 Structure determination

The structure of SosV-RBP was solved by molecular replacement in Phaser (McCoy et al. 2007) using the structure of unliganded MuV-RBP as a search model (PDB: 5B2C) (Kubota et al. 2016). The structure was improved with iterative rounds of building and refinement, using COOT (Emsley et al. 2010) and PHENIX (Adams et al. 2010) software packages, respectively. Refinement statistics are presented in **Table 5.3.3**.

3.4 STRUCTURE ANALYSIS

3.4.1 Structure overview

Crystallographic analysis confirmed that residues Arg158–Lys582 of SosV-RBP correspond to the canonical six-bladed β -propeller (**Fig 3.4.1**), with each blade ($\beta 1$ to $\beta 6$) comprised of four antiparallel β -strands. Within the asymmetric unit were two near-identical β -propeller head domains (root-mean-square deviation [RMSD] of 0.5 Å over 366 aligned Ca atoms) (**Fig 3.4.2i**), associated by a putative native dimeric interface formed through interaction of the first ($\beta 1$) and sixth ($\beta 6$) blades of the β -propeller. A portion of the N-terminal stalk region (residues Asn125–Tyr155), and loops $\beta 3L23$ (Arg344–Gly373) and $\beta 5L01$ (Val464–Thr479) could not be visualised and were directed towards the solvent channels in the crystal. This disorder likely reflects an intrinsic flexibility within these regions of the RBP, which may be stabilised by higher order RBP and fusion (F) protein assemblies formed on the native virion envelope. Eight disulphide bonds stabilise the six-bladed β -propeller fold, seven of which are conserved amongst the RBPs of the related orthorubulaviruses (Kubota et al. 2016; Yuan et al. 2005). A further disulphide, Cys522–Cys527, is formed between the 3rd and 4th strands of $\beta 5$ (**Fig 3.4.2.2iiiB**).

Paramyxoviral RBPs are typically post-translationally modified to display N-linked glycosylation, which has been shown to have roles in protein folding, host immune evasion, virulence and modulation of entry (Biering et al. 2012; Bradel-Tretheway et al. 2015; Sawatsky and von Messling 2010; Panda et al. 2004; Chu et al. 2013). Consistent with this general RBP feature,

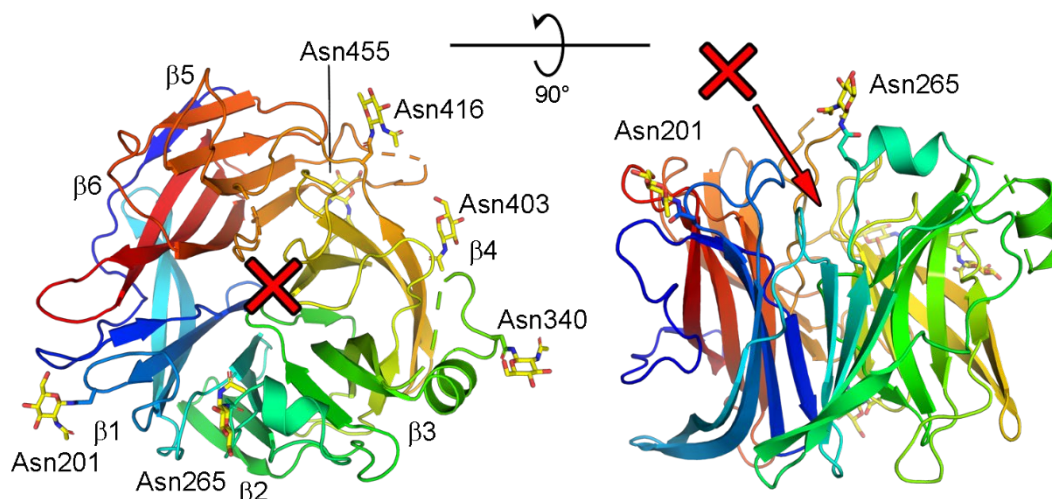


Figure 3.4.1: SosV-RBP β -propeller monomer. (a) Six-bladed β -propeller SosV-RBP monomer is shown in cartoon representation with blades labelled (coloured blue to red, from N- to C-terminus), viewing from both top down and the side. The N-linked glycosylation sites (Asn201, Asn265, Asn340, Asn403, Asn416, Asn455) observed within the crystal structure are labelled and represented as sticks (yellow). The putative sialic acid binding pocket is labelled with a red cross.

the crystallised SosV-RBP construct encodes six predicted N-linked glycosylation sequons (NXS/T, where $X \neq P$) (Asn201, Asn265, Asn340, Asn403, Asn416, and Asn455) (**Fig 3.4.1**). Electron density corresponding to GlcNAc moieties for each of these sites was observed in the crystal, consistent with the hypothesis that all the N-linked glycosylation sites of SosV-RBP are glycosylated in the context of the native virion.

3.4.2 Comparison of SosV-RBP to characterised paramyxoviral RBPs

3.4.2.1 Macro-comparisons

Structural comparison reveals that SosV-RBP exhibits a greater structural relatedness to the HN-type RBPs of the orthoavula-, orthorubula- and respiroviruses (1.7–2.1 Å root-mean-square deviation [RMSD] upon superposition of equivalent C α atoms), than the RBPs from the protein binding morbilliviruses (3.1 Å RMSD upon overlay with measles virus RBP; MV-RBP) and henipaviruses (2.2–2.3 Å RMSD) (**Fig 3.4.2.1iA**). This structural conservation is reflected upon

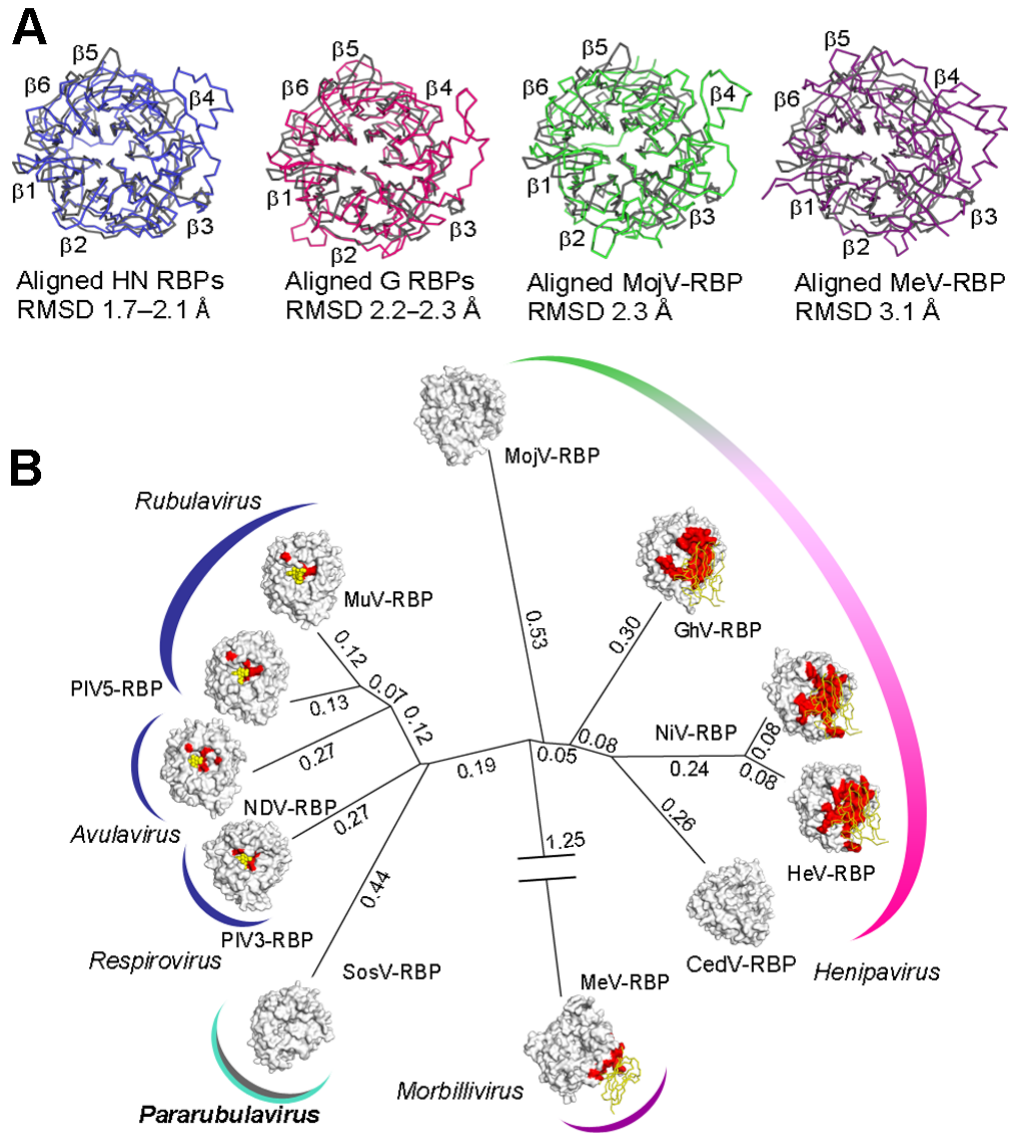


Figure 3.4.2.1i: SosV-RBP is structurally related to the HN-type RBPs. (a) Overlays of SosV-RBP (grey) with other paramyxoviral RBP structures: from left to right, NDV, Newcastle disease virus (blue, 1E8V) (Crennell et al. 2000), NiV, Nipah virus (pink, 2VWD) (Bowden, Crispin, et al. 2008), MojV, Mojiang virus (green, 5NOP) (Rissanen et al. 2017), MeV, measles virus (purple, 2ZB5) (Hashiguchi et al. 2007). C α trace rendered and root-mean-square deviation (RMSD) annotated. (b) Structure-based phylogenetic analysis of paramyxoviral RBP monomers, clockwise: SosV, Sosuga virus; PIV3, parainfluenzavirus 3 (1V2I) (Lawrence et al. 2004); NDV (1E8V); PIV5, parainfluenza virus 5 (4JF7) (Welch et al. 2013); MuV, mumps virus (5B2C) (Kubota et al. 2016); MojV (5NOP); GhV, Ghana virus (4UF7) (Lee et al. 2015); NiV (2VWD); HeV, Hendra virus (2X9M) (Bowden, Crispin, Harvey, et al. 2010); CedV, Cedar virus (PDB pending) (Pryce et al. 2019); MeV (2ZB5). Structural Homology Program (SHP (Stuart et al. 1979)), utilised to calculate evolutionary distance matrices by means of pairwise superposition of RBP structures. The resultant matrices were used to plot an unrooted tree in PHYLIP (Felsenstein 1989). RBP surfaces are represented with cognate receptor binding sites coloured (red), and receptors presented either as yellow ribbon (protein) or spheres (carbohydrate). Calculated evolutionary distances are indicated beside the branches. Figure adapted from Stelfox, et al (2019).

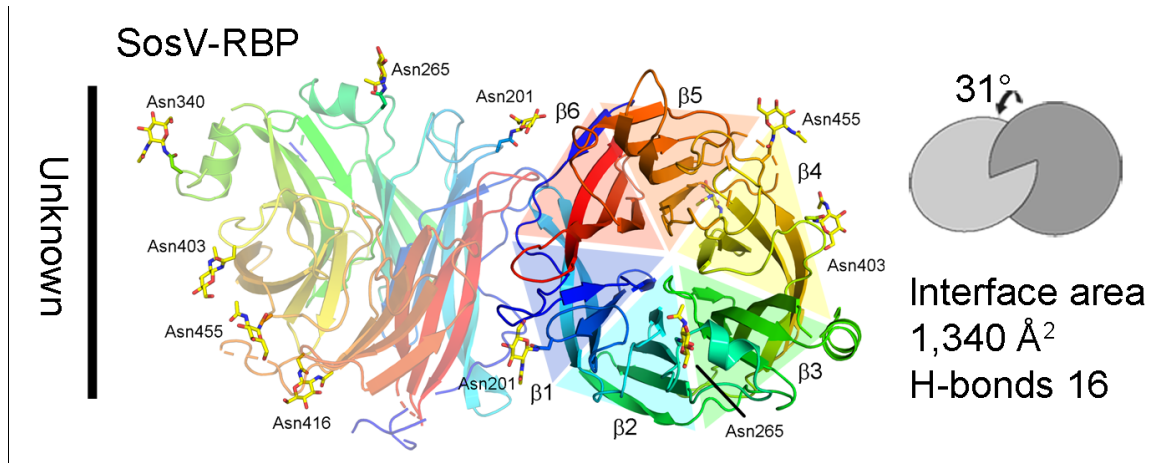


Figure 3.4.2.1ii: Dimeric assembly of SosV-RBP. The crystallographic SosV-RBP dimer shown in cartoon representation with monomers coloured rainbow from the N-terminus to C-terminus (blue to red, respectively). Calculated angle of association, interface area and number of hydrogen bonds is displayed to the right-hand side of the relevant dimer. This figure has been adapted from Stelfox et al, 2019.

structure-based phylogenetic analysis (**Fig 3.4.2.1iB**), which reveals that SosV-RBP peripherally clusters with the HN-type RBP structures.

Whilst the dimeric interface is typically comprised of residues from blades $\beta 1$ and $\beta 6$ of the RBP propeller, structural characterisation of RBP architectures has revealed that these assemblies may adopt different modes of association, depending on whether the RBP exhibits H, HN, or G functionality. The interaction between paramyxoviruses and their receptors is hypothesised to impact the overall oligomeric organisation of the different RBPs, i.e. an upward facing orientation of the ephrin binding sites on Hendra virus (HeV) accommodates this bulky receptor (Yuan et al. 2008; Bowden, Crispin, Harvey, et al. 2010; Zhang et al. 2013; Hashiguchi et al. 2007; Hashiguchi et al. 2011). The two molecules of SosV-RBP in the asymmetric unit form a putative homodimer comparable to other previously observed oligomeric paramyxoviral RBP organisations, including NDV-RBP (Crennell et al. 2000; Yuan et al. 2011), PIV3-RBP (Lawrence et al. 2004), PIV5-RBP (Yuan et al. 2005; Yuan et al. 2008; Welch et al. 2013), HeV-RBP (Bowden, Crispin, Harvey, et al. 2010), and MuV-RBP (Kubota et al. 2016). Furthermore, in-line with the crystallographically observed SosV-RBP dimer representing a native oligomeric assembly, the dimeric interface was

devoid of N-linked glycosylation (**Fig 3.4.2.1ii**) (Bowden, Crispin, Harvey, et al. 2010). Similar to the dimeric assemblies of NDV-RBP, PIV3-RBP, and MuV-RBP, which are stabilised by an average of 14 hydrogen bonds, the SosV-RBP dimer is stabilised by 16 hydrogen bonds (as calculated by the Proteins, Interfaces, Structures and Assemblies (PISA) server (Krissinel and Henrick 2007)).

The SosV-RBP dimer occludes approximately 1,340 Å² of solvent exposed surface area (**Fig 3.4.2.1ii**), which differs from the average area of approximately 1,790 Å² observed for archetypal HN-type NDV-RBP, PIV3-RBP, PIV5-RBP and MuV-RBP structures (**Fig 1.5.5**). Unlike the HN-type RBP dimers, blade one strand 4 (β1S4) does not form part of the interaction surface between the protomeric β-propeller units, and likely contributes to the reduced dimerisation surface. In addition, the angle of association between SosV-RBP protomers of 31° deviates from the average of ~60° observed across the sialic acid binding NDV-RBP (Crennell et al. 2000; Yuan et al. 2011), PIV3-RBP (Lawrence et al. 2004), PIV5-RBP (Yuan et al. 2005; Yuan et al. 2008), and MuV-RBP (Kubota et al. 2016) structures. This parallels the varied angles of association observed in HeV-RBP (40°) and MeV-RBP (117°) dimeric structures, which have been hypothesised to have arisen during the adaptation to binding bulky proteinaceous receptors, such as B-class ephrin receptors and SLAMF1/nectin-4, respectively (Bowden, Crispin, Harvey, et al. 2010; Yuan et al. 2008; Hashiguchi et al. 2011; Hashiguchi et al. 2007; Zhang et al. 2013).

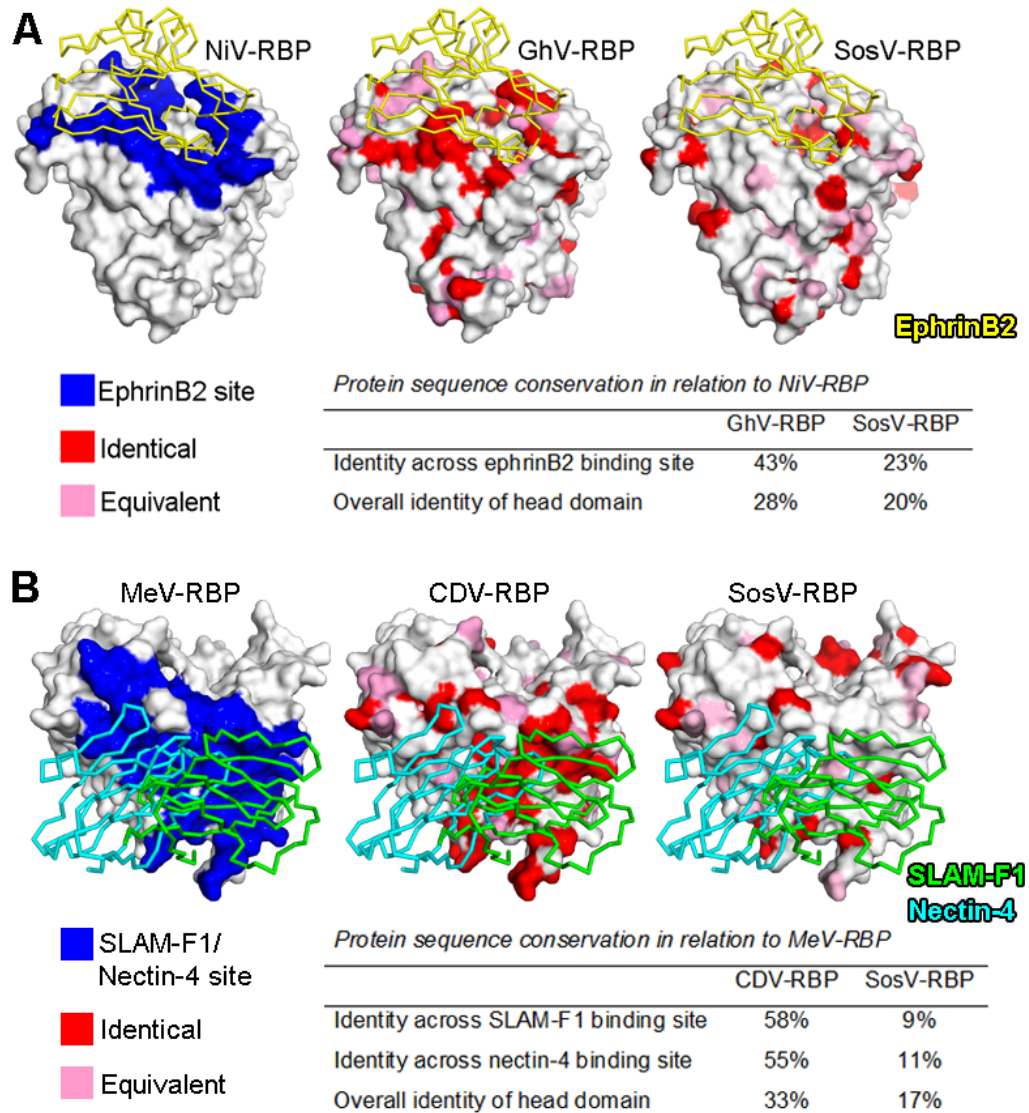


Figure 3.4.2.2i: Mapping of SosV-RBP sequence conservation onto henipaviral and morbilliviral RBPs. (a) NiV-RBP (PDB ID 2VSM) (Bowden, Aricescu, et al. 2008) is shown as a surface with ephrinB2 shown as a yellow ribbon and the cognate receptor-binding footprint coloured blue (left). Sequence conservation of GhV-RBP (middle) and SosV-RBP (right) with NiV-RBP mapped onto the surface of NiV-RBP. Consistent with the known shared usage of ephrinB2, GhV-RBP and NiV-RBP share an elevated level of sequence conservation at the receptor binding site. This feature of elevated sequence conservation was not observed using the same approach with NiV-RBP and SosV-RBP, suggestive that SosV-RBP does not utilise ephrin receptors. Residues coloured red are identical, those that are pink are equivalent residues. (b) MeV-RBP (PDB ID 3ALZ) (Hashiguchi et al. 2011) is shown in surface representation with SLAM-F1 (green) and nectin-4 (cyan) shown in ribbon representation and the cognate joint receptor-binding footprint coloured blue (left). Sequence conservation of CDV-RBP (middle) and SosV-RBP (right) with MeV-RBP is mapped onto the surface of MeV-RBP. Residues coloured red are identical, those coloured pink are equivalent. Consistent with the shared usage of SLAMF1/nectin-4, MeV-RBP and CDV-RBP share an elevated level of sequence conservation at the receptor binding sites. This feature of elevated sequence conservation was not observed using the same approach with MeV-RBP and SosV-RBP, suggestive that SosV-RBP does not utilise SLAMF1/nectin-4 receptors.

3.4.2.2 Comparisons to characterised receptor-binding sites

3.4.2.2.1 Proteinaceous receptors

Previous studies have demonstrated that genetically diverse paramyxoviruses exhibit elevated levels of sequence conservation at binding sites specific to common or similar receptors (Zeltina, Bowden, and Lee 2016). Sequence conservation mapping was performed in order to detect whether SosV-RBP may bind proteinaceous receptors utilised by henipaviral (B-class ephrins) or morbilliviral RBPs (SLAM-F1/nectin-4). SosV-RBP has an overall amino acid sequence identity with NiV-RBP of 20%, and a similarly low sequence identity of 23% across ephrinB2 binding residues, in contrast to the elevated identity demonstrated by GhV-RBP across this region (Bowden, Aricescu, et al. 2008; Lee et al. 2015) (**Fig 3.4.2.2iA**). Likewise, there is no detectable sequence conservation between SosV-RBP and MeV-RBP across nectin-4 and SLAM-F1 binding residues (Hashiguchi et al. 2011; Zhang et al. 2013) (**Fig 3.4.2.2iB**), nor does SosV-RBP display the integral β 4-5 groove required morbilliviral receptor recognition (Zhang et al. 2013; Hashiguchi et al. 2011; Santiago et al. 2010). These analyses suggest that SosV-RBP does not utilise characterised henipaviral or morbilliviral proteinaceous receptors.

3.4.2.2.2 Sialic acid

The well understood sialic acid interaction site presented by HN-type RBPs is composed of a specific set of seven conserved residues, which act collaboratively to bind and hydrolyse sialic acid (Connaris et al. 2002; Crennell et al. 2000; Iorio et al. 2001; Langedijk, Daus, and VanOirschot 1997), in addition to the paramyxoviral hexapeptide motif (Asn-Arg-Lys-Ser-Cys-Ser) (Mirza, Deng, and Iorio 1994), involved in completing the necessary receptor binding and destroying activities (Wasik, Barnard, and Parrish 2016). While the SosV-RBP structure retains some features consistent with the HN-type RBPs, including a site for binding the glycerol component of sialic acid (**Fig 3.4.2.2iiB**), and a conserved cation binding site (**Fig 3.4.2.2iiC**) (Crennell et al. 2000; Kubota et al. 2016; Lawrence et al. 2004; Yuan et al. 2005), there are a number of distinct differences across the putative sialic acid recognition site (**Fig 3.4.2.2iiA**). Firstly, residues corresponding to the SosV-RBP hexapeptide motif present a surface that is structurally incompatible to sialic acid recognition

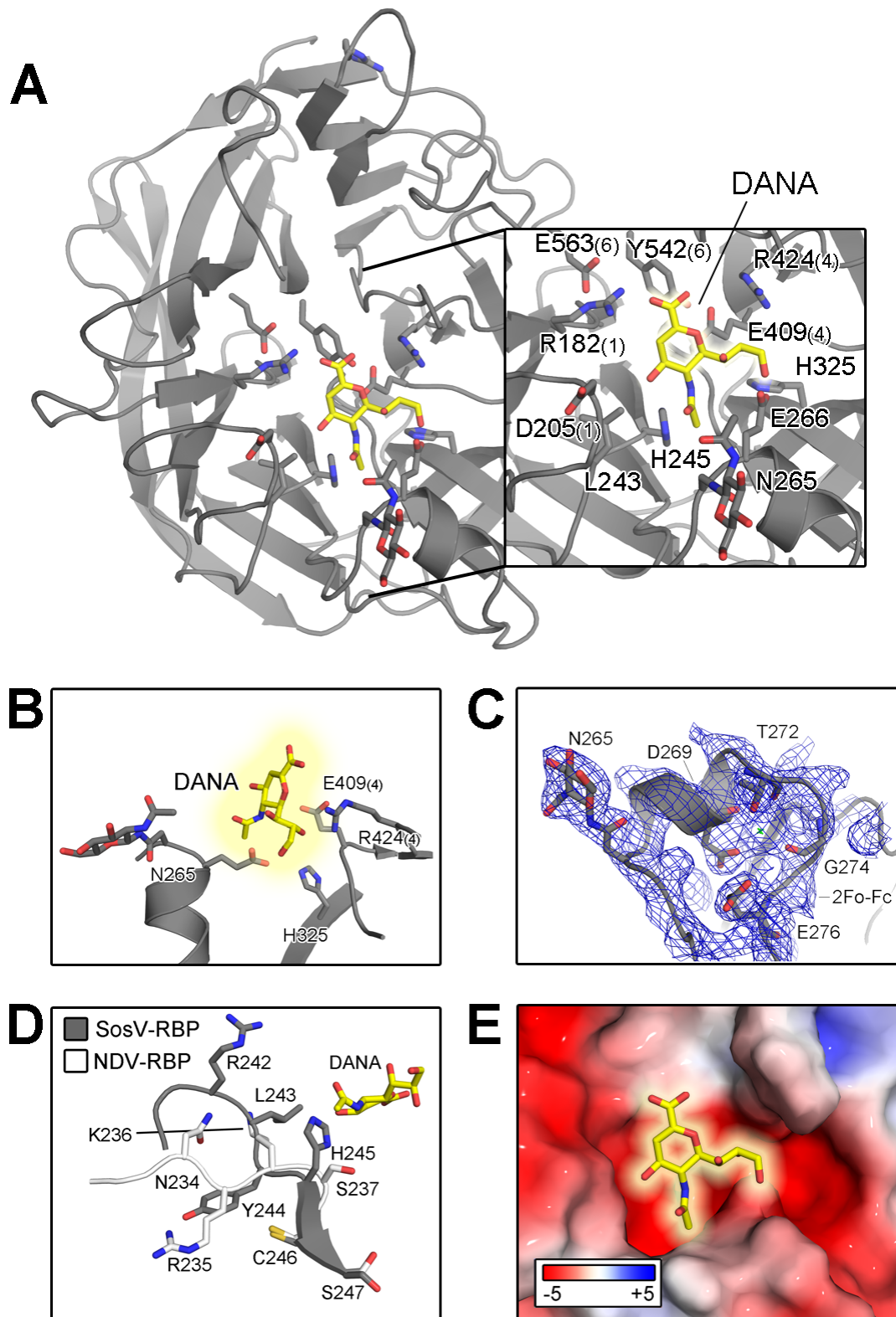


Figure 3.4.2.2ii: SosV-RBP displays similarities and differences to the HN-type across the region corresponding to the sialic acid binding site. (a) The SosV-RBP β -propeller, and close up of region corresponding to sialic acid binding site shown with the inhibitor N-Acetyl-2,3-dehydro-2-deoxyneuraminic acid (DANA) (yellow) superimposed (1E8V). Residues equivalent to those utilised by HN-type RBPs to bind and hydrolyse sialic acid shown using stick representation and annotated. (b) Residues and loops implicated in binding glycerol region of sialic acid and its analogues displayed, with important residues shown using stick representation and annotated, and DANA (yellow) superimposed. (c) Fitting of a calcium ion (green) into loop β 2L23 is consistent with the 2Fo–Fc difference map generated following refinement. A similar mechanism of stabilization of β 2L23 through coordination of a calcium cation has been observed in several HN RBP structures, and loss of calcium inhibits activity in NDV-RBP (Connaris et al. 2002). The 2Fo–Fc electron density is rendered at 2.0 sigma and displayed as mesh (blue). SosV-RBP is displayed in cartoon (grey) representation, with important residues and glycans displayed as sticks. (d) SosV-RBP (grey) and NDV-RBP (white) (1E8V) (Crennell et al. 2000) are shown in cartoon representation with the relevant side chains represented as sticks. N-Acetyl-2,3-dehydro-2-deoxyneuraminic acid (DANA) is shown as yellow sticks. Homologous residues are numbered according to SosV-RBP. In contrast to NDV-RBP, SosV-RBP residues Arg242, Leu243, and His245 protrude into the putative sialic acid binding site. (e) Electrostatic charges of SosV-RBP calculated using the Adaptive Poisson-Boltzmann Solver (Jurrus et al. 2018), represented using surface by colouring residues from red to blue (-5 kT e^{-1} to $+5 \text{ kT e}^{-1}$), demonstrating the electrostatic incompatibility of the SosV-RBP surface with binding of sialic acid and analogues. DANA is represented as above.

(Fig 3.4.2.2iiD). Residues Leu243^{SosV-RBP} and His245^{SosV-RBP} impede into the region where the O₄ of sialic acid is located in liganded HN-type structures (Crennell et al. 2000; Kubota et al. 2016; Lawrence et al. 2004; Yuan et al. 2005), altering both local structure but also the chemical composition of the pocket (Fig 3.4.2.2iiD). Analysis of surface electrostatics using the Adaptive Poisson-Boltzmann server (APBS) (Jurrus et al. 2018), confirms that SosV-RBP presents a considerably altered surface charge, resulting in an extended acidic patch incongruous to binding the negatively charged sialic acid (Fig 3.4.2.2iiE). The most compelling change, however, is the rearrangement of the β 5L23 loop of SosV-RBP, which presents Arg514 (R₅) of the triarginyl motif. In HN-type RBP structures, the β 5L23 loop folds inwards enabling stabilisation of the carboxyl group of an incoming sialic acid, however the homologous loop in the SosV-RBP structure peels outwards towards a small external cavity (Fig 3.4.2.2iii). This altered conformation causes a dramatic displacement of Arg514^{SosV-RBP}, resulting in an approximately 14 Å distance between the

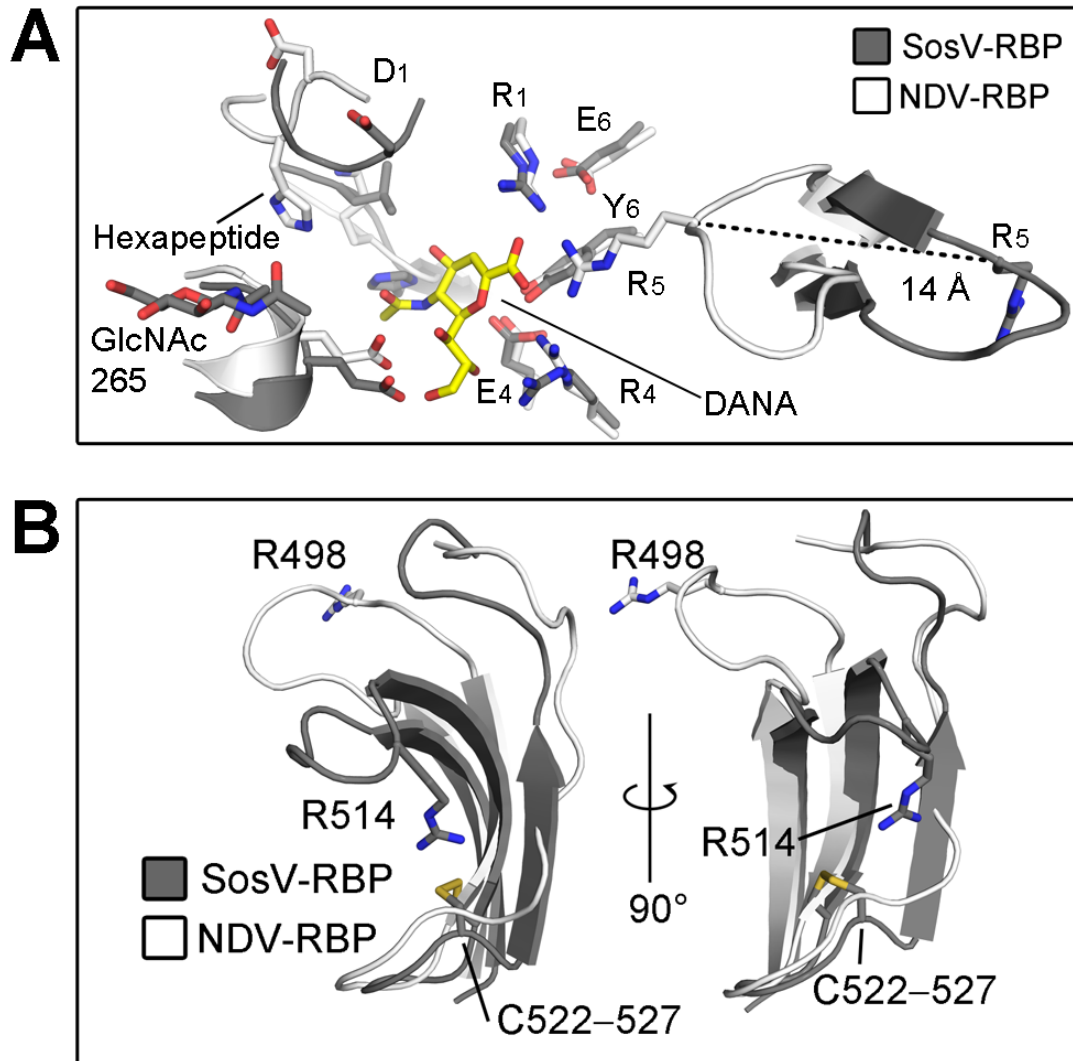


Figure 3.4.2.iii: Arg514₍₅₎ of SosV-RBP displays a unique orientation which may impact sialic acid interactions. (a) Active site of NDV-RBP (white) superposed with SosV-RBP (grey) structure, highlighting the differing locations of the Arg₅ residues. Residues are labelled as in Fig. 3.2. Loops are shown in cartoon representation, with functionally important residues and DANA (yellow) (1E8V) shown as sticks. (b) Both SosV-RBP (grey) and NDV-RBP (white) (1E8V) are shown in cartoon representation and the relevant residues represented using sticks. The conserved triarginyl motif of the SosV-RBP structure is disrupted, loop β 5L23 bends outwards, away from the putative active site. The disulphide (Cys522–Cys527) formed between β 5S3 and β 5S4 has not been observed in other paramyxovirus structures and may contribute to the structural displacement of residue Arg514₍₅₎ in SosV-RBP.

equivalent Arg5 C α atom in the homologous HN-type RBP structures (Crennell et al. 2000; Kubota et al. 2016; Yuan et al. 2005; Lawrence et al. 2004) (**Fig 3.4.2.2iiiA**). Stabilisation of this large local structural difference could be attributed to the introduction of the novel SosV-RBP disulphide bond between residues Cys522 and Cys527 of the third and fourth strands of blade five (β 5S3 and β 5S4, respectively) (**Fig 3.4.2.2iiiB**). Upon collection of crystallographic data following production of SosV-RBP crystals soaked with (3-sialyllactose, 30mM) and co-crystallised in the presence of sialyllactose (3- and 6- sialyllactose, 5x greater molar concentration), no changes to the β 5L23 loop conformation and no ligand density within the putative sialic acid binding site, or elsewhere within the structure, was observed. Similarly, there are no alternate confirmations of the β 5L23 loop which resemble the SosV-RBP mode in the liganded and un-liganded HN-type structures (Crennell et al. 2000; Kubota et al. 2016; Yuan et al. 2005; Lawrence et al. 2004), indicative that SosV-RBP may utilise a sialic acid independent mode of attachment. Furthermore, there is a lack of structural conservation at the region equivalent to the secondary binding site on NDV-RBP (Crennell et al. 2000), and likewise no evidence of glycan binding to this site following crystal soaks.

3.5 ASSESSMENT OF SIALIC ACID INTERACTIONS

Whilst these structural analyses provide evidence that SosV-RBP is distinct from the receptor binding architectures presented by characterised paramyxoviruses, given the close topological similarity to the HN-type RBPs and the maintenance of conserved residues implicated in sialic acid interactions, we assessed the hemadsorption (Morrison and McGinnes 1989) and neuraminidase activity (Potier et al. 1979) of SosV-RBP. Consistent with the recent finding that related pararubulaviruses utilise a sialic acid independent mechanism of host-cell entry (Johnson et al. 2019) and the distinctive structural organisation of SosV-RBP, human embryonic kidney (HEK) 293T cells presenting full-length SosV-RBP exhibited no neuraminidase (Potier et al. 1979) and minimal hemadsorption activity (Morrison and McGinnes 1989) with respect to an NDV-RBP control (**Fig. 3.5iA**). This apparent lack of sialic acid interactions was further corroborated by assessment of neuraminidase activity using the Warren method (Warren 1959), where a lack of free sialic acid was

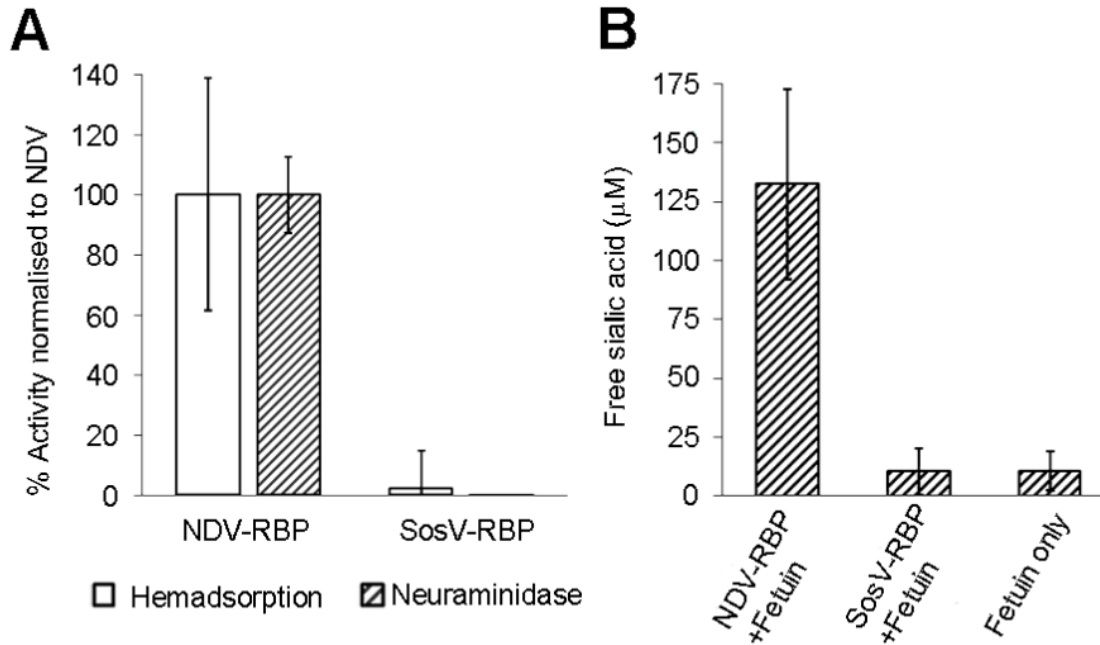


Figure 3.5i: SosV-RBP displays limited hemadsorption activity and a lack of neuraminidase activity. (a) SosV-RBP neuraminidase (Potier et al. 1979) and hemadsorption (Morrison and McGinnes 1989) activity normalized to cell surface expression of a NDV-RBP control. (b) Free sialic acid concentration detected following incubation of RBP with fetuin (Warren 1959). For panels A ($n = 10$) and B ($n = 13$) error bars represent the standard deviation across all replicates. Figure adapted from Stelfox, et al (2019).

measured following incubation of SosV-RBP with fetuin, a heavily sialylated protein presenting both $\alpha 2,3$ and $\alpha 2,6$ linked sialic acid (Spiro 1960; Cointe, Leroy, and Chirat 1998; Guttman and Lee 2016; Wu et al. 2016), when compared to a fetuin only control and NDV-RBP (**Fig 3.5iB**). Soluble NDV-RBP (Ser47–Gly577) was generated for this initial Warren test through transient transfection of HEK293T cells, followed by purification utilising nickel-affinity chromatography and SEC. Given the apparent lack of binding or enzymatic activity demonstrated by SosV-RBP, these data were indicative that sialic acid dependent host-cell engagement is not utilised by SosV as a primary mode of infection.

Subsequently, a SosV-RBP construct presenting the full hexapeptide motif ‘Asn–Arg–Lys–Ser–Cys–Ser’ was recombinantly engineered (termed ‘eSosV-RBP’) to assess whether neuraminidase activity is acquired when compared to wildtype SosV-RBP and an NDV-

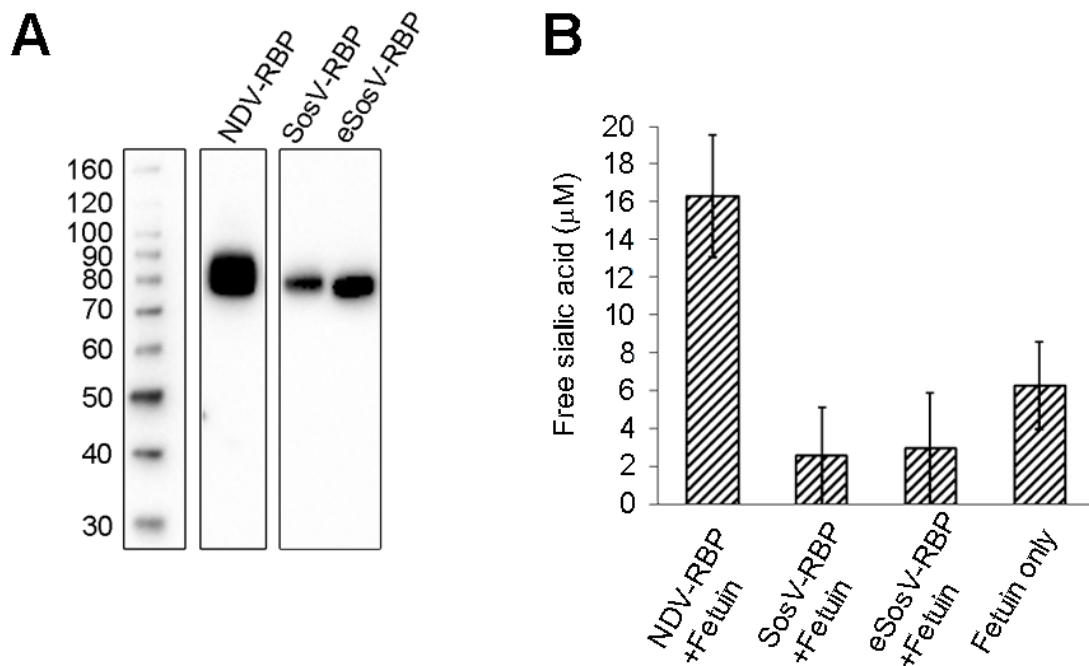


Figure 3.5ii: Engineered (e)SosV-RBP bearing full ‘NRKSCS’ motif lacks neuraminidase activity. (a) HEK293T cells were transiently transfected with NDV-RBP, wildtype SosV-RBP, and eSosV-RBP. All constructs encoded an N-terminal hexa-histidine tag and SUMO-tag. The expression of folded soluble protein harvested from the supernatants of transfected HEK293T cells was determined using western blotting with a mouse-derived monoclonal anti-penta-histidine antibody. Samples were run alongside Benchmark™ Protein Ladder (Thermo, left), on the same gel with the same exposure time. (b) Free sialic acid concentration detected following incubation of NDV-RBP, SosV-RBP, ‘NRKS’ mutant eSosV-RBP, and mock-transfected cell supernatant with fetuin (Warren 1959). For B ($n = 6$) error bars represent the standard deviation across all replicates. Figure adapted from Stelfox, et al (2019).

RBP control using the Warren method (Warren 1959). Whilst soluble eSosV-RBP was shown to be properly folded (**Fig 3.5iiA**), similar to the wildtype SosV-RBP no neuraminidase activity was detected when assessed against the NDV-RBP positive control and a mock transfected negative control (**Fig 3.5iiB**). These data supported a model by which the lack of the full hexapeptide motif did not exclusively dictate the likely sialic acid independence of SosV, but also the structural and chemical composition of the potential interaction site, in particular the displacement of Arg514₍₅₎.

3.6 DISCUSSION

A recent study in Uganda which reported that 62% of cases of severe febrile illness were misdiagnosed as malaria demonstrates that the impact of emerging viruses, such as SosV, in Africa is likely underestimated (Ghai et al. 2016; Stoler and Awandare 2016). The bat-borne SosV, was recently identified as the cause of a severe febrile illness suffered by a wildlife biologist (Albarino et al. 2014), and was subsequently identified in bat populations throughout Uganda (Amman et al. 2015). While the identification of new pararubulaviruses continues and the full biomedical and economic impact of these viruses is yet to be understood, other pararubulaviruses have demonstrated the ability to cross the species barrier. For example, MenV infects pigs, fruit bats and humans (Bowden et al. 2001; Chant et al. 1998), and the Achimota viruses (AchPV) are capable of causing respiratory disease in experimentally infected ferrets (Barr et al. 2018), and further serological evidence of AchPV2 human infection in Ghana and Tanzania indicates that the virus is potentially zoonotic (Baker et al. 2013). Given the importance of the paramyxoviral RBP in determining the zoonotic potential of paramyxoviruses (Thibault et al. 2017), I proceeded to determine the structure of the SosV-RBP and define the structural and functional relatedness to other paramyxoviral RBPs, with focus on key receptor binding motifs, and utilise this information to tailor subsequent functional analyses.

Our structural analyses demonstrate that whilst the overall SosV-RBP presents an architecture most similar to the HN-type RBPs (**Fig 3.4.2.1i**), observation of the regions homologous to receptor binding sites reveal molecular divergence. This work also reveals that SosV-RBP does not share structural features associated with SLAMF1/nelectin-4 and ephrin recognition, and interestingly displays a structurally and chemically distinct profile at the canonical sialic acid interaction site. Specifically, despite the site homologous to sialic acid binding presenting the residues that support the binding of the glycerol group of sialic acid (**Fig 3.4.2.2iiB**) and a conserved cation binding site (**Fig 3.4.2.2iiC**), non-favourable surface charge properties are exhibited (**Fig 3.4.2.2iiE**) and sialic acid binding is sterically disrupted by the protrusion of residues, Leu243 and His245 (**Fig 3.4.2.2iiD**) and the displacement of Arg514₍₅₎ (**Fig 3.4.2.2iii**).

Consistent with recent studies regarding related pararubulaviruses, our analysis shows that SosV-RBP has minimal ability to bind sialic acid and lacks neuraminidase activity (**Fig 3.5**). However, the observed low level hemadsorptive ability exhibited by SosV-RBP in the context of an NDV-RBP control could indicate low-affinity or secondary interactions, such as those which support MERS-CoV infection (Li et al. 2017). Similarly, whilst these assays do not assess binding to the neurotropic $\alpha 2,8$ linked sialic acid (Mühlenhoff, Eckhardt, and Gerardy-Schahn 1998), we note that a broadened receptor tropism exhibited by vaccine strains of MuV did not result in a departure from $\alpha 2,3$ and $\alpha 2,6$ specificity (Santos-Lopez et al. 2009), suggestive that if SosV-RBP interacted with $\alpha 2,8$ linked sialic acid, hemadsorption and neuraminidase activity would have been observed. In addition, our finding that SosV-RBP likely does not utilise sialic acid as a primary receptor are corroborated by recent studies which demonstrate that related pararubulaviruses, MenV, TioV and TevPV, employ a sialic acid independent mode of entry (Johnson et al. 2019). The hemadsorption and neuraminidase data are rationalised by our structural findings, as described, SosV-RBP presents a potential sialic acid binding site with a number of alterations likely incompatible with the typical HN-type activity. Interestingly, we were able to demonstrate that production of SosV-RBP bearing the full hexapeptide motif ‘Asn–Arg–Lys–Ser–Cys–Ser’ did not confer neuraminidase activity (**Fig 3.5C**), indicating that the principal contributor to the reduced sialic acid associated activity demonstrated by SosV-RBP, is the observed structural and chemical differences and not the alterations to the primary sequence.

However, consistent with genetic analysis, structural phylogenetic analysis and overlay with existing RBP structures, demonstrate that SosV-RBP displays a higher structural similarity to the HN-type RBPs rather than the protein tropic viruses. Given this closer relation to the HN-type RBPs, SosV-RBP likely diverged from a sialic acid dependent ancestral paramyxovirus, in an event that was distinct from the evolution of the morbilli- and henipaviruses. Furthermore, SosV-RBP exhibits the smallest known level of structural reorganisation required to gain sialic acid independence. These observed structural differences may have arisen following acquisition of binding motifs required for interaction with a unique protein or glycan based receptor. However, given the structural plasticity

of the paramyxoviral β -propeller, it is plausible that another novel binding site is utilised, and structural diversification across the sialic acid site is a product of the lack of functional constraints imposed across this region. In addition, alterations to the SosV-RBP architecture are not limited to the sialic acid binding site. Whilst SosV-RBP demonstrates a fold related to the HN-type RBPs, the organisation of the dimeric SosV-RBP deviates from the conserved $\sim 60^\circ$ association angle observed in HN-type RBP structures (MuV-RBP, PIV3-RBP, PIV5-RBP and NDV-RBP), similar to the associations observed for HeV-RBP and MeV-RBP dimers, where alterations to the high-order RBP assembly may facilitate interactions with bulky proteinaceous receptors (Bowden, Crispin, Harvey, et al. 2010; Yuan et al. 2008; Hashiguchi et al. 2011; Zhang et al. 2013), supportive of SosV-RBPs acquisition of a novel receptor-binding modality.

3.7 CONCLUSION

Our investigations, that utilise both structural and functional methods, support the sialic acid independent paradigm presented by pararubulaviruses (Johnson et al. 2019; Bowden et al. 2001; Chua, Wang, et al. 2002; Burroughs et al. 2015; Lau et al. 2010). Our structural studies rationalise a lack of observed neuraminidase and minimal hemadsorption activity, and enhance our understanding of the diversity of receptor binding architectures presented by paramyxoviruses. However, further characterisation of the receptor interactions mediated by SosV-RBP and other pararubulaviral RBPs is necessary to grasp the potential impact and spillover risk of these emerging paramyxoviruses.

CHAPTER 4

THE MOSSMAN VIRUS AND NARIVA VIRUS RECEPTOR BINDING PROTEINS

4 THE MOSSMAN VIRUS AND NARIVA VIRUS RECEPTOR BINDING PROTEINS

4.1 SUMMARY OF FINDINGS

The rodent-isolated paramyxoviruses Mossman and Nariva (MosV and NarV, respectively) (Miller et al. 2003; Lambeth et al. 2009), represent an uncharacterised arm of the *Paramyxoviridae*, the recently defined *Narmovirus* genus (Amarasinghe et al. 2019). The propensity of paramyxoviruses to undergo interspecies transmission is partially dictated by the molecular details of paramyxoviral receptor recognition, an interaction that impacts both cellular and host tropism, and therefore the resultant pathobiological characteristics (Thibault et al. 2017; Zeltina, Bowden, and Lee 2016). In this chapter, the structural characterisation of MosV and NarV RBPs (MosV-RBP and NarV-RBP) reveals that whilst the narmovirus RBPs exhibit the canonical paramyxoviral six-bladed β -propeller, they lack the necessary structural motifs for known paramyxoviral entry pathways. Structural-phylogenetic analysis reveals that MosV-RBP and NarV-RBP present a distinct structural class of paramyxovirus RBP. Furthermore, through sequence conservation mapping a potential binding site unique amongst the paramyxoviral RBPs has been identified, with efforts to characterise this region ongoing. Altogether, our data indicate that the narmoviruses, MosV and NarV, likely utilise a novel paramyxoviral entry pathway and therefore support the recent independent classification of narmoviruses from other members of the *Paramyxoviridae*.

4.2 INTRODUCTION

Viral surveillance and molecular characterisation efforts have led to the discovery and study of a number of novel paramyxoviruses with unknown pathogenic potential (Drexler et al. 2012; Thibault et al. 2017). MosV was isolated in Queensland, Australia, from the pooled organs of wild rats originating from two separate locations in the 1970s (Miller et al. 2003) and NarV was isolated in the early 1960s from forest rodents in Trinidad (Lambeth et al. 2009). Despite the disparate locations in which they were discovered, sequencing and phylogenetic analysis revealed these viruses

clustered closely within the *Paramyxoviridae* (Amarasinghe et al. 2019; Thibault et al. 2017; Zeltina, Bowden, and Lee 2016). However, as there is little functional and structural information regarding these viruses, our understanding of MosV and NarV is limited to sequence. Due to their tenuous phylogenetic relationship with paramyxoviruses from the established genera, the newly defined *Narmovirus* genus was formed to encompass MosV, NarV and the more recently sequenced bank vole virus (BaVV) (Amarasinghe et al. 2019).

The RBP is presented on the paramyxovirus surface as a dimer-of-dimers, with each protomer consisting of an intraviral (IV) region, transmembrane (TM) domain, α -helical stalk, and a six-bladed β -propeller head domain, which mediates the interaction with the cognate host-cell receptor (Yuan et al. 2008; Yuan et al. 2011; Yuan et al. 2005). Primary amino acid sequence of the paramyxoviral RBPs is highly variable despite conservation of the overall protein fold (Bowden, Crispin, Jones, et al. 2010; Zeltina, Bowden, and Lee 2016). This variable nature of paramyxoviral RBPs is illustrated upon BLAST alignment of full-length RBPs to NarV-RBP, where there is an overall poor query cover and low sequence identity (**Table 4.2**). Indeed, despite the distinct clustering of MosV-RBP and NarV-RBP together within the paramyxoviral RBP phylogenetic tree, the primary sequences exhibit only 31.4% identity (**Table 4.2**). Furthermore, NarV-RBP presents an extended C-terminus (Asn626–Pro657) bearing no sequence homology to regions of other paramyxoviral RBPs, or any known proteins. This diversity complicates predictions of viral tropism based on amino acid sequence alone. However, the structure-based methods proved useful in predicting the independence of Mojiang virus (MojV) RBP from the established receptor-specific groupings (Rissanen et al. 2017).

Here, I solved the structures of MosV-RBP and NarV-RBP glycoproteins using X-ray crystallography to 1.6 Å and 2.1 Å resolution, respectively. Utilising structure-based phylogenetic analysis and finer comparisons against characterised paramyxovirus binding surfaces, the recent distinction of MosV-RBP and NarV-RBP within the newly defined genus, *Narmovirus*, is rationalised. Furthermore, this work has led to the identification of a putative receptor binding site

Table 4.2: Percent sequence identity of amino acid sequences of paramyxoviral RBPs against NarV-RBP.

	Protein ^b	Identity ^c (Calculated with BLAST)	Query coverage ^c (Calculated with BLAST)
<i>Metaavulavirus</i>	APMV2-RBP ^a	20.0%	60.0%
<i>Orthoavulavirus</i>	NDV-RBP ^a	23.5%	24.0%
<i>Paraavulavirus</i>	APMV3-RBP ^a	19.0%	84.0%
<i>Respirovirus</i>	SeV-RBP ^a	22.1%	71.0%
<i>Orthorubulavirus</i>	MuV-RBP ^a	22.8%	37.0%
<i>Pararubulavirus</i>	MenV-RBP ^a	26.2%	42.0%
<i>Henipavirus</i>	HeV-RBP ^a	17.8%	81.0%
	MojV-RBP ^a	21.7%	27.0%
<i>Morbillivirus</i>	MeV-RBP ^a	25.9%	15.0%
<i>Narmovirus</i>	MosV-RBP ^a	31.4%	94.0%

^aReceptor binding protein (RBP)

^bRBPs of viruses representing each of the current genera of the *Paramyxoviridae*, including: avian paramyxovirus 2 (APMV2), Newcastle disease virus (NDV), avian paramyxovirus 3 (APMV3), Sendai virus (SeV), mumps virus (MuV), Menangle virus (MenV), Hendra virus (HeV), Mojiang virus (MojV), measles virus (MeV) and MosV.

^cIdentity and query cover values are coloured using a scale format, with the lowest values represented with yellow and the highest with green.

shared by MosV-RBP and NarV-RBP, which has allowed the prediction of residues that are potentially important for negotiating host-cell recognition, with these hypotheses being currently investigated.

4.3 PROTEIN PRODUCTION AND STRUCTURE DETERMINATION OF MOSV-RBP AND NARV-RBP

4.3.1 Construct design and small-scale expression testing

MosV and NarV RBP constructs for protein expression were generated from human codon-optimised genes of the full-length RBPs (Genbank AY286409.1 and FJ362497.2, respectively). Due to the low amino acid sequence similarity between MosV-RBP and NarV-RBP, and the other paramyxoviral RBPs, construct boundaries were determined by alignment of the amino acid sequences from a number of solved RBP structures, including Mojiang virus (MojV)-RBP (Rissanen et al. 2017), Nipah virus (NiV)-RBP (Bowden, Crispin, et al. 2008), measles virus (MeV)-RBP

(Hashiguchi et al. 2007) and Newcastle disease virus (NDV)-RBP (Crennell et al. 2000). Constructs containing the putative six-bladed β -propeller head and minimal stalk region were preferred as they were likely to be more amenable to crystallisation (Savitsky et al. 2010) (**Fig 4.3.1A**). Using a primer specific to the selected N-terminal start sites and the fixed C-terminal end site, PCR was performed to generate fragments for cloning into the pHLsec mammalian expression vector (Aricescu, Lu, and Jones 2006). For MosV-RBP, constructs Lys181–Thr632 and Glu202–Thr632 (**Fig 4.3.1A**), and for NarV-RBP, constructs Glu184–Pro657 and Ile202–Pro657 (**Fig 4.3.1C**), were initially designed for further testing. However, due to issues cloning MosV-RBP Lys181–Thr632 and NarV-RBP Ile202–Pro657, only the soluble expression of the MosV-RBP Glu202–Thr632 and NarV-RBP Glu184–Pro657 constructs were analysed by small-scale expression screening in human embryonic kidney (HEK)293T cells. The soluble protein fraction was analysed through anti-histidine western blotting of the supernatant, revealing successful expression of the folded MosV-RBP Glu202–Thr632 and NarV-RBP Glu184–Pro657 constructs (**Fig 4.3.1B and D**).

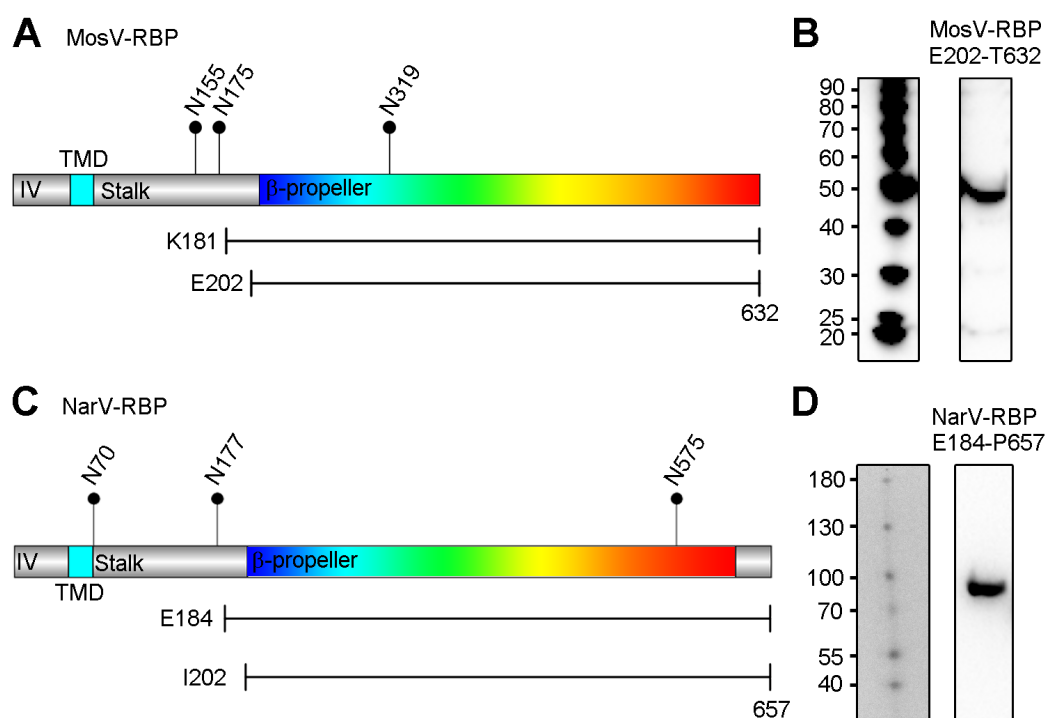


Figure 4.3.1 Construct design and small-scale expression test of MosV-RBP and NarV-RBP.

(a) Gene diagram displaying features of the MosV-RBP, including the intraviral (IV) region, transmembrane domain (TMD), stalk and six-bladed β -propeller receptor binding head (rainbow). Predicted N-linked glycosylation sites (NXS/T, where X $\neq$ P) are marked with pins and relevant asparagine residue numbered. The designed constructs are illustrated below the schematic. (b) Soluble expression of MosV-RBP construct Glu202–Thr632 tagged with a C-terminal histidine tag shown using western blot analysis of HEK293T cell supernatants from small-scale expression test. Western blot analysis was performed alongside the Benchmark™ Protein Ladder (Thermo, left), under reducing conditions. (c) Gene diagram displaying features of the NarV-RBP, including the intraviral (IV) region, transmembrane (TM) domain, stalk and six-bladed β -propeller receptor binding head (rainbow). Predicted N-linked glycosylation sites are marked with pins and relevant asparagine residue numbered. The designed constructs are illustrated below the schematic. (d) Soluble expression of NarV-RBP construct Glu184–Pro657 tagged with a C-terminal histidine tag shown using western blot analysis of HEK293T cell supernatants from small-scale expression test. Samples were run under reducing conditions alongside a PageRuler™ prestained protein ladder (Thermo, left), where the ladder was imaged using white light.

4.3.2 Large-scale expression and purification of MosV-RBP and NarV-RBP

Large-scale transient expression of both MosV-RBP (Glu202–Thr632) and NarV-RBP (Glu184–Pro657) was performed in the presence of the class I mannosidase inhibitor kifunesine (Chang et al. 2007). Cell culture supernatants were harvested by clarification through centrifugation and filtration, prior to buffer exchange into 10 mM Tris pH 8, 150 mM NaCl buffer (TBS). Subsequently, supernatants were purified using nickel-affinity chromatography using a HisTrap™ HP (GE Healthcare) column, and subjected to an initial round of SEC using a Superdex™ 200 Increase 10/300 GL column (GE Healthcare). Pooled peaks were deglycosylated using overnight treatment with EndoF₁, resulting in protein decorated with single GlcNAc residues. A further round of SEC was performed to remove EndoF₁ prior to crystallisation.

Given the calculated molecular weight of monomeric MosV-RBP is 49 kDa, the species observed through SEC analysis (44 kDa–158 kDa) would be consistent with a dimeric arrangement of MosV-RBP (**Fig 4.3.2A**). A shift in size between the pre- and post-EndoF₁ treated MosV-RBP observed following SDS-PAGE analysis confirmed successful deglycosylation (**Fig 4.3.2B**). Furthermore, both SEC and SDS-PAGE analysis indicated that a monodisperse sample of MosV-RBP had been obtained, and therefore was suitable for crystallisation.

The species observed during SEC of NarV-RBP, conversely, is larger, with a size consistent with the 158 kDa gel filtration size standard (**Fig 4.3.2C**). This is somewhat larger than the expected size of the monomer, consistent with hypothesis that NarV-RBP forms a stable higher order oligomer in solution, which purifies as a monodisperse peak. The solution state characteristics of both MosV-RBP and NarV-RBP are examined in greater detail in **section 4.4.2**. Following SDS-PAGE analysis (**Fig 4.3.2D**), which confirmed successful deglycosylation of NarV-RBP, the protein was applied into crystallisation screens.

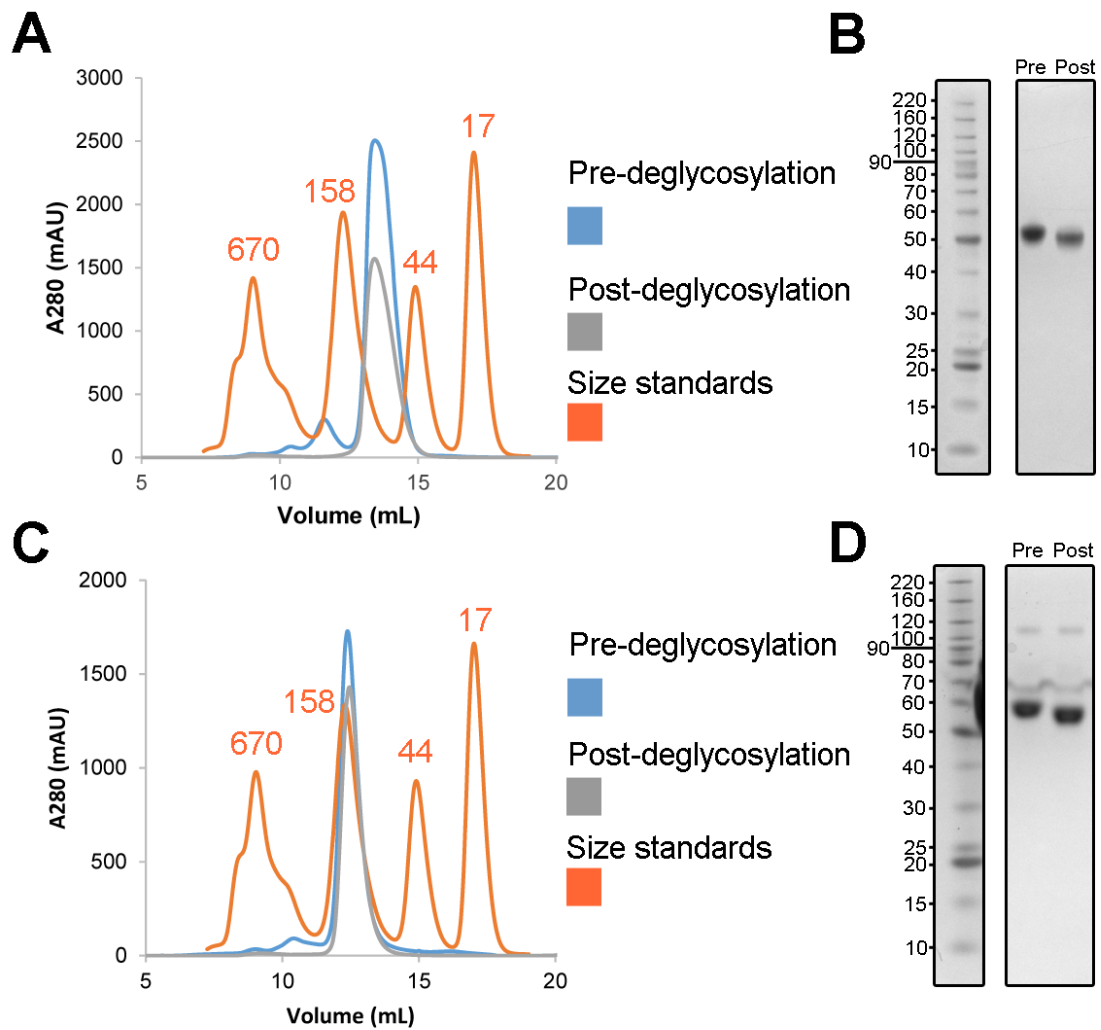


Figure 4.3.2: Purification of MosV-RBP and NarV-RBP. (a) Size exclusion chromatogram of MosV-RBP expressed transiently in the presence of kifunensine, pre- (blue) and post-deglycosylation (grey) with EndoF₁, aligned against gel filtration standards (BioRad, orange). (b) The SDS-PAGE analysis of MosV-RBP purification pre- and post-deglycosylation indicates a shift following treatment with EndoF₁. (c) Size exclusion chromatogram of NarV-RBP expressed stably in the presence of kifunensine, pre- (blue) and post-deglycosylation (grey) with EndoF₁, aligned against gel filtration standards (BioRad, orange). (d) The SDS-PAGE analysis of NarV-RBP pre- and post-deglycosylation reveals a shift following treatment with EndoF₁. In both panel B and D, samples were run under reducing conditions, alongside Benchmark™ Protein Ladder (Thermo, left).

4.3.3 Crystallisation and X-ray diffraction data collection

Utilising a pre-crystallisation test (PCTTM, Hampton), the optimal concentration for crystallisation of MosV-RBP was determined to be 5.5 mg/mL. MosV-RBP crystals were harvested 4 days following plate setup. Crystals were grown in the following precipitant mix: 0.2 M L-arginine, 0.1 M Tris pH 7.8, 8% poly- γ -glutamic acid (PGA)-LM, 6% dextran sulphate. Crystals were immersed in 20% glycerol diluted in reservoir solution prior to cryo-cooling. X-ray diffraction data were collected at a wavelength of 0.9795 Å on beamline I04, DLS. Reflections were processed to a resolution of 2.75 Å (Table 4.3.4i). With the help of Dr Karl Harlos, for experimental phasing, crystals from this same condition were soaked in a solution containing potassium tetrachloroplatinate (K₂PtCl₄) for 3 hr, prior to cryo-cooling with a 20% glycerol solution containing K₂PtCl₄ diluted in precipitant. At beamline I04, DLS, crystals were exposed to a wavelength consistent with the platinum LIII absorption edge ($\lambda = 1.072$ Å) to facilitate experimental structure determination. A further MosV-RBP dataset was collected on crystals which grew between 52 and 79 days post plate preparation, utilising the following precipitant mix: 0.2 M potassium bromide, 0.2 M potassium thiocyanate, 0.1 M sodium cacodylate pH 6.5, 3% PGA-LM, 20% w/v polyethylene glycol (PEG) 550 MME. Crystals were harvested and cryo-cooled in a 20% glycerol solution diluted with precipitant. Data were collected at a wavelength of 0.9795 Å, on beamline I04, DLS, and subsequently reflections were processed to 1.62 Å (Table 4.3.4i).

The optimal concentration for NarV-RBP crystallisation was determined to be 4.5 mg/mL. Fine needle-like crystals appeared following 4 days after plate preparation, in the following precipitant condition: 0.2 M magnesium chloride, 0.1 M HEPES pH 7.5, 25% PEG 3350, 10% PEG 400. Crystals were optimised by seeding into a lower concentration solution of NarV-RBP (Bergfors 2007; Walter et al. 2008). Crystals were pulverised, utilising a seed bead (Hampton Research, USA) (Luft and DeTitta 1999), and dispensed onto a plate prepared with the original precipitant mix and NarV-RBP at a concentration of 3 mg/mL (Walter et al. 2008). Fine fan-like crystals were harvested and cryo-cooled in a 20% glycerol, precipitant mix, 8 days following plate preparation. Data were collected at wavelength of 0.9795 Å, on beamline I04, DLS, and subsequently reflections were

processed to 2.07 Å (**Table 4.3.4ii**). All MosV-RBP and NarV-RBP reflections were indexed, integrated and scaled using the xia2 package (Winter 2010).

4.3.4 Structure determination

Phase determination of the original MosV-RBP dataset using molecular replacement was not successful, however, as the crystals were numerous and reproducible, heavy metal soaks using K₂PtCl₄ were performed. The isomorphous differences between the native dataset and this new derivatised dataset, in addition to the anomalous signal derived from platinum, enabled subsequent phase determination using single isomorphous replacement with anomalous scattering (SIRAS). SIRAS was performed using the Autosol wizard (Terwilliger et al. 2009), which identified heavy atom sites, calculated initial phases, performed density modification, and proceeded with the initial building and refinement of the MosV-RBP model. Further iterative phase improvement and model building was performed using the ‘Phase and Build’ programme (Adams et al. 2010). The partial MosV-RBP model obtained was used as a molecular replacement model to phase the NarV-RBP dataset using PHASER (McCoy et al. 2007). Iterative rounds of NarV-RBP model building and refinement were performed using the programmes COOT (Emsley and Cowtan 2004) and Phenix.refine (Afonine et al. 2012), with the refinement statistics presented in **Table 4.3.4ii**.

Despite multiple attempts to iteratively build and refine the MosV-RBP model, it remained partially built and the R-factors were unsatisfactory (>30%). Using PHASER (McCoy et al. 2007), molecular replacement was performed using the refined NarV-RBP structure, making sure to avoid potential model bias imposed through maintenance of a consistent set of MosV-RBP R_{free} flags. As above, the model underwent iterative rounds of building and refinement, and was considered complete. However, following collection of a higher resolution MosV-RBP dataset, molecular replacement was performed once again, this time utilising the lower resolution MosV-RBP as a model to enable phasing of the higher resolution MosV-RBP dataset. Model building and refinement was performed as above, with the refinement statistics presented in **Table 4.3.4i**.

Table 4.3.4i: Data collection and refinement statistics for MosV-RBP

	Native	Pt Peak	High resolution native
Data collection statistics			
Beamline	DLS I02	DLS I04	DLS I04
Spacegroup	<i>I</i> 4 ₁ 2 2	<i>I</i> 4 ₁ 2 2	<i>P</i> 2 ₁ 2 ₁ 2 ₁
Cell dimensions			
a, b, c (Å)	214.0, 214.0 138.4	214.0, 214.0 136.2	80.2, 83.9, 122.1
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution range (Å)	67.66–2.75 (2.82–2.75) ^a	63.22–3.24 (3.32–3.24)	69.17–1.62 (1.65–1.62)
Wavelength (Å)	0.9795	1.0721	0.9795
R_{merge}	0.12 (2.26)	0.18 (4.63)	0.08 (1.47)
R_{meas}	0.13 (2.46)	0.20 (4.76)	0.08 (1.62)
R_{pim}	0.04 (0.79)	0.03 (0.65)	0.03 (0.68)
I/σ I	15.3 (1.3)	22.1 (1.2)	13.0 (1.1)
CC_{1/2}	1.00 (0.51)	1.00 (55.9)	1.00 (0.50)
Completeness (%)	99.7 (99.7)	99.9 (100.0)	99.8 (99.4)
Multiplicity	11.7 (12.0)	53.6 (53.8)	6.5 (5.5)
Resolution (Å)	48.58–1.62 (1.68–1.62)		
No. reflections	10,324		
R_{work} / R_{free}^b (%)	20.1 (17.3)		
No. atoms			
Protein	6,766		
Ligand/ion	14		
Water	695		
B-factors			
Protein	32.9		
Ligand/ion	36.4		
Water	38.6		
R.m.s deviations^c			
Bond lengths (Å)	0.006		
Bond angles (°)	0.89		
Ramachandram analysis^d			
Favoured region (%)	96.3		
Allowed region	3.4		

^aNumbers in parentheses refer to the outer resolution shell.^bR_{free} is calculated as for R_{work}, but using only 5% of the data which were separated prior to refinement.^cr.m.s deviations: root mean square deviation from ideal geometry.^dDetermined using the Molprobit server (Chen et al. 2010).

Table 4.3.4ii: Data collection and refinement statistics for NarV-RBP

NarV-RBP	
Data collection statistics	
Beamline	DLS I04
Spacegroup	$P\ 1\ 2_1\ 1$
Cell dimensions	
a, b, c (Å)	55.4, 83.0, 93.0
α, β, γ (°)	90.0, 90.0, 90.0
Resolution range (Å)	82.96–2.07 (2.14–2.07) ^a
Wavelength (Å)	0.9795
R _{merge}	0.18 (0.81)
R _{meas}	0.24 (1.09)
R _{pim}	0.13 (0.59)
I/ σ I	4.2 (1.3)
CC _{1/2}	0.98 (0.62)
Completeness (%)	100 (99.9)
Multiplicity	3.3 (3.4)
Refinement statistics	
Resolution (Å)	55.38–2.07 (2.14–2.07)
No. reflections	50,856
R _{work} / R _{free} ^b (%)	23.0/27.0
No. atoms	
Protein	6,702
Ligand/ion	0
Water	721
B-factors	
Protein	23.3
Ligand/ion	N/A
Water	26.6
R.m.s deviations ^c	
Bond lengths (Å)	0.003
Bond angles (°)	0.63
Ramachandran analysis ^d	
Favoured region (%)	97.1
Allowed region (%)	2.8

^aNumbers in parentheses refer to the outer resolution shell.

^bR_{free} is calculated as for R_{work}, but using only 5% of the data which were separated prior to refinement.

^cr.m.s deviations: root mean square deviation from ideal geometry.

^dDetermined using the Molprobit server (Chen et al. 2010).

4.4 STRUCTURE ANALYSIS

4.4.1 Structure overview of the MosV and NarV RBPs

Crystallographic analysis confirmed that the six-bladed β -propellers of MosV-RBP and NarV-RBP consist of the residues Thr209–Thr632 and Arg203–Asn626, respectively (**Fig 4.4.1**). For both of the structures, the asymmetric unit consisted of two near-identical β -propeller head domains, with a putative native dimeric interface composed of interactions between the first ($\beta 1$) and sixth ($\beta 6$) blades of the β -propeller (**Fig 4.4.1**). The root-mean-square deviation [RMSD] between MosV-RBP monomers is 0.4 Å over 428 aligned C α atoms, with a buried surface area of 2,064 Å², as calculated by the Proteins, Interfaces, Structures and Assemblies (PISA) server (Krissinel and Henrick 2007). The RMSD between NarV-RBP monomers is 0.4 Å over 427 aligned C α atoms, with a calculated dimer interface area of 2,165 Å². The first residue, Glu202, of MosV-RBP and final residue, Thr632, including the C-terminal could not be resolved. More regions of NarV-RBP could not be resolved, including residues Glu184–Asn197 at the N-terminus, and the extended C-terminus Asn626–Pro657, which were disordered and directed towards the solvent channels in the crystal. Potentially, such disorder is indicative of flexibility, which may only be stabilised in the context of the native virion, where neighbouring RBPs or fusion (F) proteins may influence NarV-RBP structure.

Despite a sequence similarity of only 31.4%, MosV-RBP and NarV-RBP monomers appear to maintain a close structural relationship, demonstrating a RMSD of 1.6 Å over 401 aligned C α atoms. This is also reflected at the dimer level, where alignment of the two dimeric RBPs results in an RMSD of 1.8 Å over 801 aligned C α atoms. Similarly, both structures are stabilised by a set of seven conserved disulphide bonds, consistent with the NiV-RBP structure, which presents this same conserved set of seven disulphides (Bowden, Crispin, et al. 2008). This is in contrast to NDV-RBP (Crennell et al. 2000) and MeV-RBP (Hashiguchi et al. 2007), which both show a less related pattern of disulphide bonding to the narmoviral RBPs, with only five and three conserved disulphides across the β -propeller, respectively.

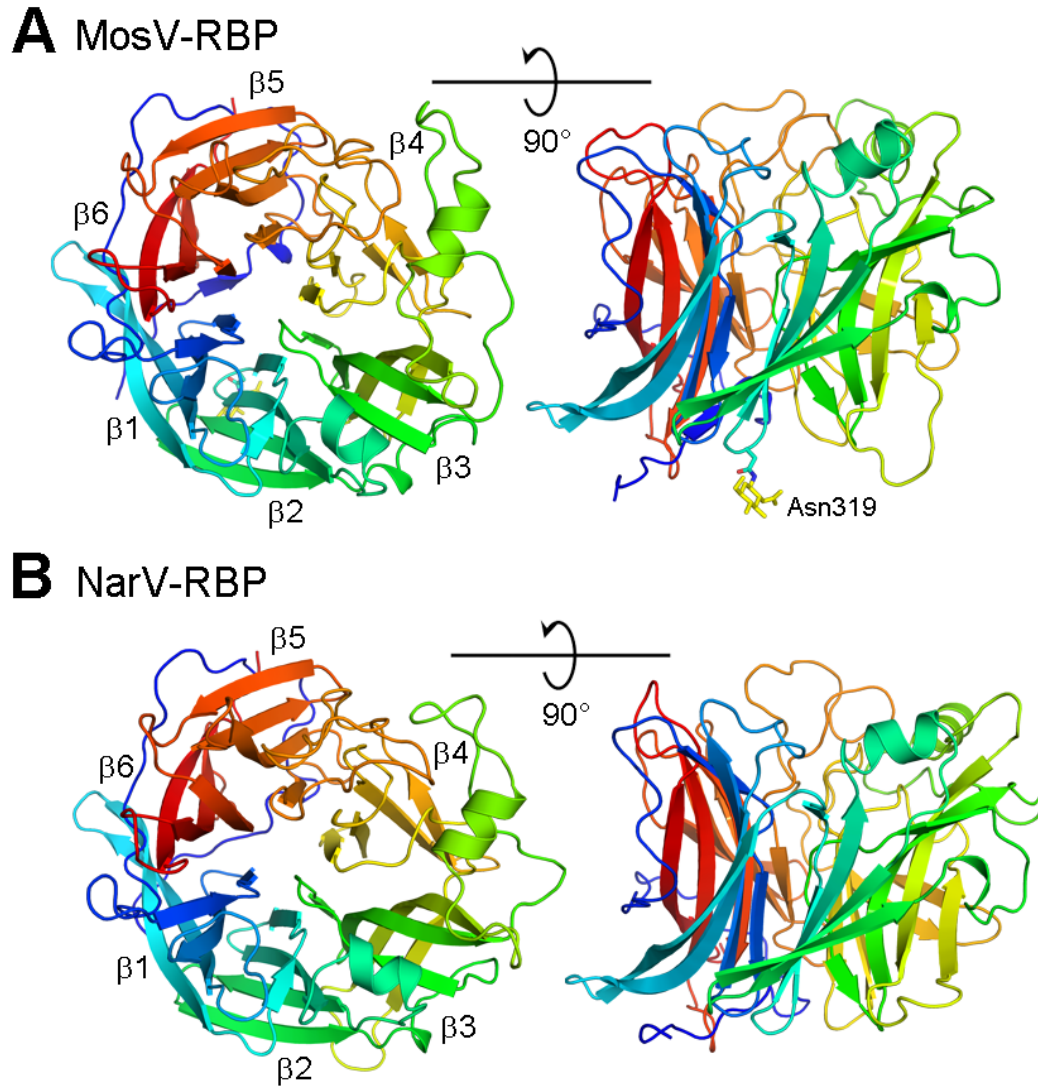


Figure 4.4.1: Crystal structures of MosV-RBP and NarV-RBP β -propellers. (a) The six-bladed β -propeller MosV-RBP monomer is shown in cartoon representation with blades labelled (coloured blue to red, from N- to C-terminus), viewing from both top down and the side. The N-linked glycosylation site (Asn319) and corresponding N-acetylglucosamine residue (yellow sticks) observed within the crystal structure is annotated. (b) The six-bladed β -propeller NarV-RBP monomer, viewed from both top down and the side, is shown in cartoon representation with blades labelled (coloured blue to red, from N- to C-terminus).

Unlike other paramyxoviruses, which typically present highly glycosylated surfaces (Bowden, Crispin, et al. 2008; Sawatsky and von Messling 2010; Hu et al. 1994; Segawa et al. 2003; Ng, Hiebert, and Lamb 1990; McGinnes and Morrison 1995; Panda et al. 2004; Chu et al. 2013), MosV-RBP and NarV-RBP, encode only three predicted N-linked glycosylation sequons (NXS/T where $X \neq P$), with only one within the β -propeller, Asn319 for MosV-RBP and Asn575 for NarV-RBP. In contrast to MosV-RBP where density observed at the Asn319 site is consistent with a single N-acetylglucosamine (GlcNAc), density for N-linked glycosylation at Asn575 was not observed in the NarV-RBP structure, either due to the site not being occupied (Jones, Krag, and Betenbaugh 2005) or the glycan being disordered.

4.4.2 Small-angle X-ray scattering analysis

To assess the solution state characteristics of glycosylated MosV-RBP and NarV-RBP β -propellers, SAXS was carried out at beamline B21, Diamond Light Source (DLS), Didcot, UK. Utilising the software package ‘ScÅtter’, the molecular mass was calculated following determination of the radius of gyration (R_g) using Guinier analysis, the distance pair ($P(r)$) distribution using real space analysis and the volume-of-correlation (V_c) (Putnam 2016; Rambo and Tainer 2013b, 2013a). The mass calculated for the recombinantly produced MosV-RBP and NarV-RBP was approximately 117 kDa and 112 kDa, respectively (Rambo and Tainer 2013a). Considering that MosV-RBP and NarV-RBP were glycosylated and accounting for the typical approximately 10% error observed when measuring protein mass using SAXS, these data were consistent with the expected molecular mass of MosV-RBP and NarV-RBP dimers (98 kDa and 106 kDa, respectively). Experimental scattering data were compared to a theoretical scattering profile calculated from the MosV-RBP and NarV-RBP crystal structures using the ‘Fast X-ray Scattering’ (FoXS) server (Schneidman-Duhovny, Hammel, and Sali 2010). Whilst the calculated χ^2 between the experimental SAXS profile and dimeric crystal structure data was not ideal in both cases (4.5 and 3.9, for MosV-RBP and NarV-RBP, respectively), when compared to fitting of a theoretical profile generated from MosV-RBP and NarV-RBP monomers alone ($\chi^2 = >100$), this analysis supported the dimeric organisation in solution (**Fig 4.4.2**). Furthermore, such discrepancies between the experimental and theoretical dimer profiles could arise

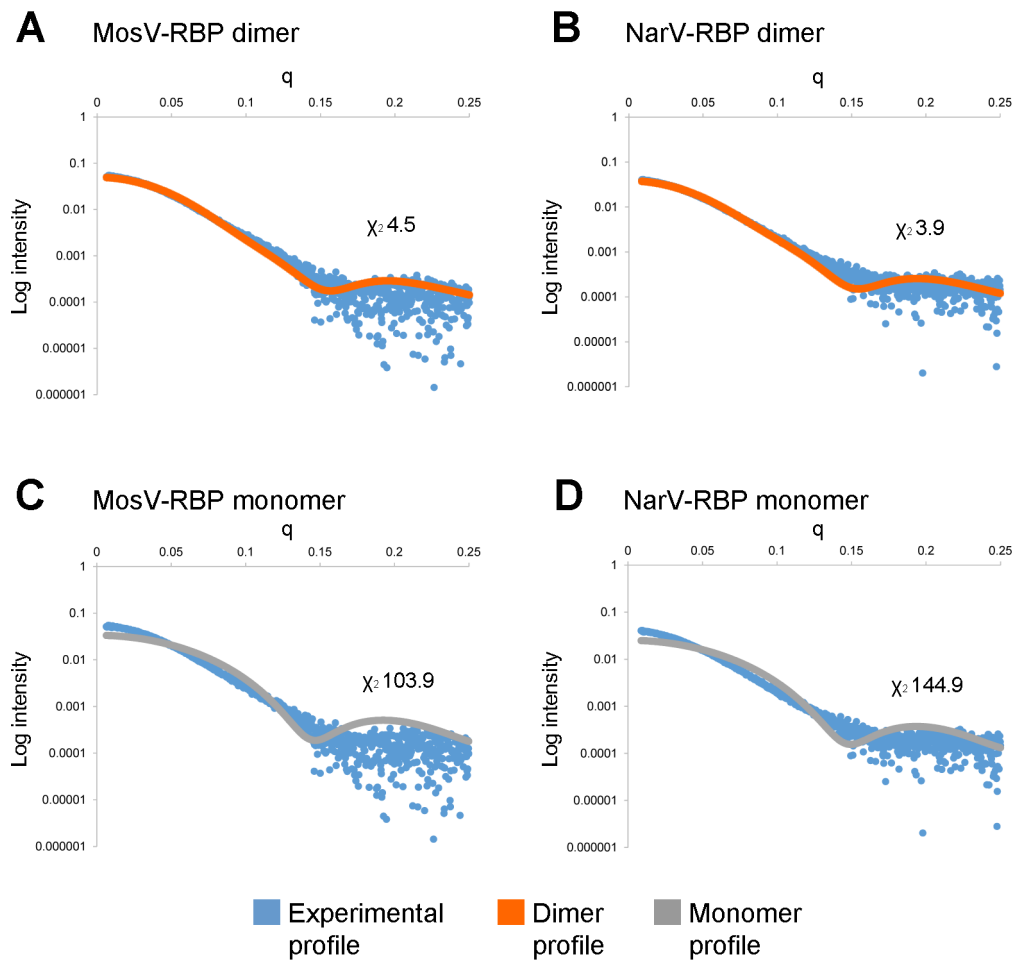


Figure 4.4.2: SAXS analysis indicates MosV-RBP and NarV-RBP β -propellers form dimers in solution. Experimental MosV-RBP and NarV-RBP SAXS data (blue) compared to theoretical data calculated from dimeric (orange) and monomeric (grey) crystal structures, plotted log intensity ($I(q)$) against the Porod invariant (q) (Schneidman-Duhovny, Hammel, and Sali 2010). (a) MosV-RBP SAXS data against dimeric MosV-RBP, (b) NarV-RBP SAXS data against dimeric NarV-RBP, (c) MosV-RBP SAXS data against monomeric MosV-RBP, (d) NarV-RBP SAXS data against monomeric NarV-RBP. The χ^2 for each fit is represented above the plotted experimental and theoretical profiles.

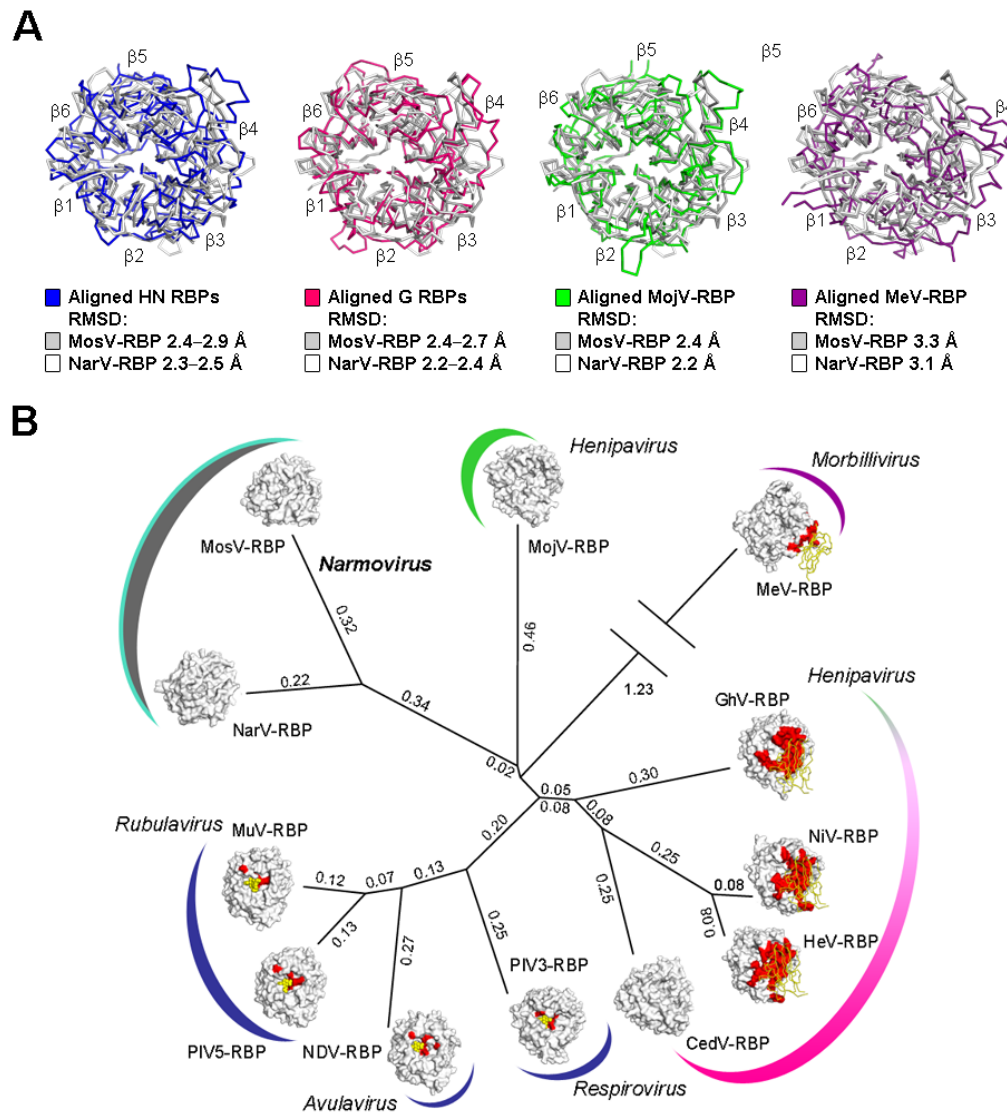


Figure 4.4.3.1i: MosV-RBP and NarV-RBP are structurally distinct from characterised paramyxoviral RBPs. (a) Overlays of MosV-RBP (white) and NarV-RBP (grey) with other paramyxoviral RBP structures: from left to right, NDV (blue, 1E8V) (Crennell et al. 2000), NiV (pink, 2VWD) (Bowden, Crispin, et al. 2008), MojV (green, 5NOP) (Rissanen et al. 2017), MeV (purple, 2ZB5) (Hashiguchi et al. 2007). α trace rendered and root-mean-square deviation (RMSD) annotated. (b) Structure-based phylogenetic analysis of paramyxoviral RBP monomers: MosV; NarV; PIV3, parainfluenzavirus 3 (1V2I) (Lawrence et al. 2004); NDV (1E8V); PIV5, parainfluenza virus 5 (4JF7) (Welch et al. 2013); MuV (5B2C) (Kubota et al. 2016); MojV (5NOP); MeV (2ZB5); GhV, Ghana virus (4UF7) (Lee et al. 2015); NiV (2VWD); HeV (2X9M) (Bowden, Crispin, Harvey, et al. 2010); CedV, cedar virus (PDB code N/A) (Pryce et al. 2019). Structural Homology Program (SHP (Stuart et al. 1979)), utilised to calculate evolutionary distance matrices by means of pairwise superposition of RBP structures. The resultant matrices were used to plot an unrooted tree in PHYLIP (Felsenstein 1989). RBP surfaces are represented with cognate receptor binding sites coloured (red), and receptors presented either as yellow ribbon (protein) or spheres (carbohydrate). Calculated evolutionary distances are indicated beside the branches.

from an inability to resolve regions of disorder in the crystal structure and differences in the glycosylation state of the samples utilised for each technique.

4.4.3 Comparison of MosV-RBP and NarV-RBP to different paramyxoviral RBPs

4.4.3.1 Macro-comparisons

Despite a conserved pattern of disulphide bonding across the β -propeller, pairwise analysis reveals that MosV-RBP and NarV-RBP deviate from the HNV-RBP structures (2.4–2.7 Å and 2.2–2.4 Å RMSD, respectively) (Bowden, Crispin, et al. 2008) (**Fig 4.4.3.1iA**). Indeed, a similar level of structural divergence is observed upon pairwise analysis of MosV-RBP and NarV-RBP against MeV-RBP (3.3 Å and 3.1 Å RMSD, respectively) and the HN-type RBPs of the *Respiro*-, *Orthoavula*-, and *Orthorubulavirus* genera (2.4–2.9 Å and 2.3–2.5 Å RMSD, respectively) (Yuan et al. 2005; Lawrence et al. 2004; Crennell et al. 2000; Hashiguchi et al. 2007; Bowden, Aricescu, et al. 2008; Rissanen et al. 2017; Kubota et al. 2016) (**Fig 4.4.3.1iA**). This observed independence is reflected upon structural phylogenetic analysis of the narmovirus RBPs and published RBP structures, where MosV-RBP and NarV-RBP form a distinct branch equidistant from the HN-type RBPs and the protein binding RBPs (**Fig 4.4.3.1iB**). The apparent distinction of the MosV-RBP and NarV-RBP is reflective of an evolutionary divergence from an ancestral paramyxoviral RBP.

This apparent structural independence of MosV-RBP and NarV-RBP from other paramyxoviral RBPs extends to the observed dimeric arrangement (**Fig 4.4.3.1ii**). As calculated by the PISA server (Krissinel and Henrick 2007), the association between MosV-RBP protomers occludes 2,066 Å² of solvent accessible surface area, and is stabilised by 16 hydrogen bonds and 3 salt bridges. The interface between NarV-RBP protomers is similarly substantial, demonstrating a surface area of 2,157 Å², stabilised by 19 hydrogen bonds and 3 salt bridges (**Fig 4.4.3.1ii**). MosV-RBP and NarV-RBP show a similar mode of dimerisation as that observed in HN-type oligomeric RBP structures (Crennell et al. 2000; Kubota et al. 2016; Yuan et al. 2005; Lawrence et al. 2004), where residues of both the 1st and 6th blades (β 1 and β 6, respectively) contribute to the interface. However, the *Narmovirus* RBP structures demonstrate a greater buried surface area, and utilise

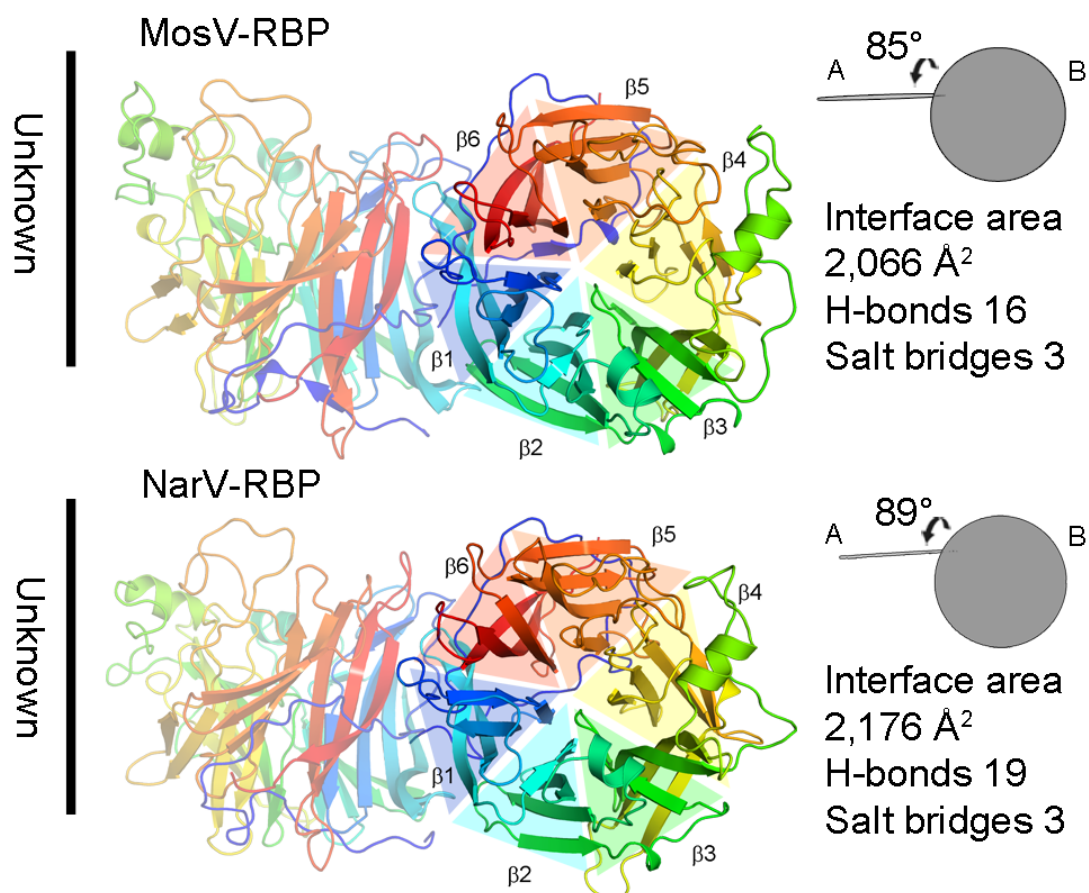


Figure 4.4.3.1ii: Dimeric assembly of MosV-RBP and NarV-RBP. The crystallographic MosV-RBP and NarV-RBP dimer shown in cartoon representation with monomers coloured rainbow from the N-terminus to C-terminus (blue to red, respectively). Calculated angle of association, interface area and number of hydrogen bonds is displayed to the right-hand side of the relevant dimer. The circles labelled 'A' and 'B' represent different monomers of the dimeric assembly.

salt bridges in the arrangement. Calculation of the angle of association of MosV-RBP and NarV-RBP protomers also reveals that their overall organisation is divergent from the established RBP dimerisation modes presented by the henipaviral RBPs, morbilliviral RBPs and HN-type RBPs. Indeed, whilst the HN-type RBP dimers typically assemble with a 60° of association between protomers, the MosV-RBP and NarV-RBP protomers are organised approximately 90° relative to each other.

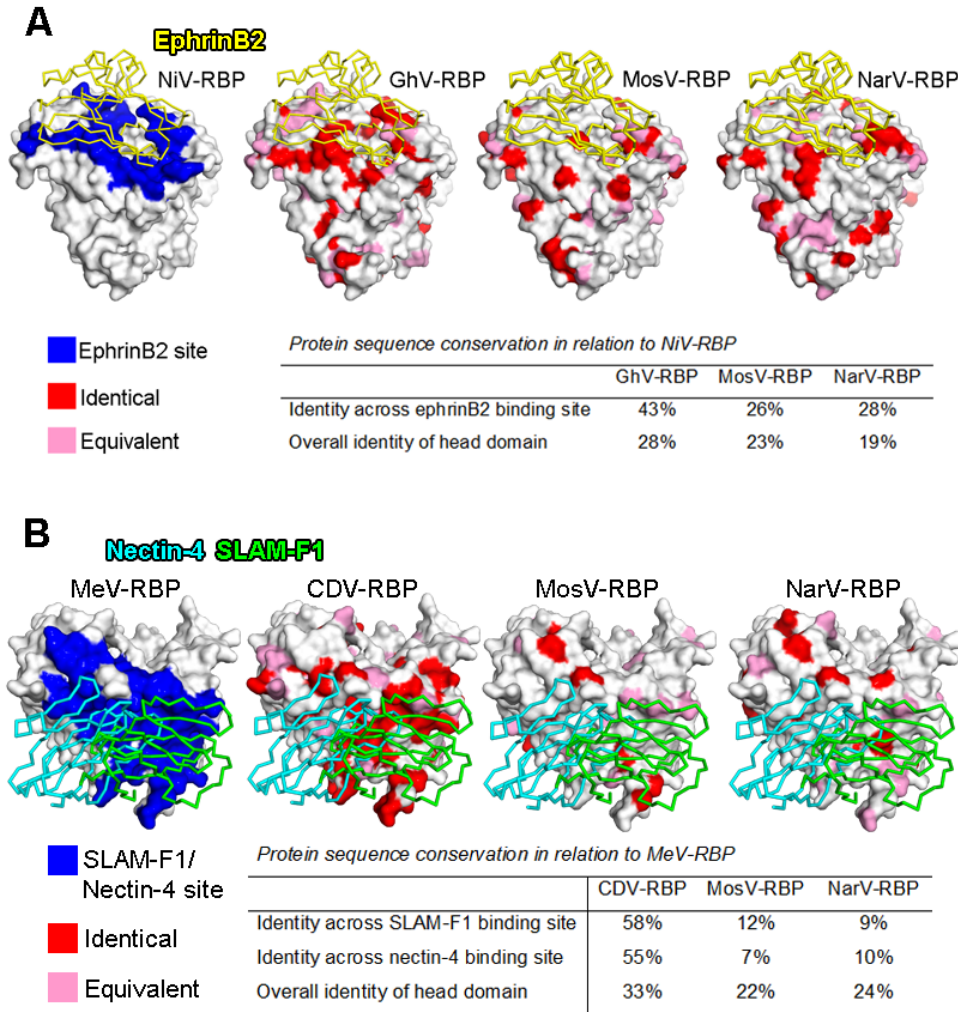


Figure 4.4.3.2i: Mapping of MosV-RBP and NarV-RBP sequence conservation onto henipaviral and morbilliviral RBPs. (a) NiV-RBP (2VSM) (Bowden, Aricescu, et al. 2008) is shown as a surface with ephrinB2 shown as a yellow ribbon and the cognate receptor-binding footprint coloured blue (far left). Sequence conservation of GhV-RBP (second left), MosV-RBP (second right), and NarV-RBP (far right), with NiV-RBP mapped onto the surface of NiV-RBP. Consistent with the known shared usage of ephrinB2, GhV-RBP and NiV-RBP share an elevated level of sequence conservation at the ephrin receptor binding site. This feature of elevated sequence conservation was not observed using the same approach with NiV-RBP and the narmoviral RBPs, suggestive that MosV-RBP and NarV-RBP do not utilise ephrin receptors. Residues coloured red are identical, those that are pink are equivalent residues. (b) MeV-RBP (3ALZ) (Hashiguchi et al. 2011) is shown in surface representation with SLAM-F1 (green) and nectin-4 (cyan) shown in ribbon representation and the cognate joint receptor-binding footprint coloured blue (far left). Sequence conservation of CDV-RBP (second left), MosV-RBP (second right), and NarV-RBP (far right), with MeV-RBP is mapped onto the surface of MeV-RBP. Residues coloured red are identical, those coloured pink are equivalent. Consistent with the shared usage of SLAMF1/nectin-4, MeV-RBP and CDV-RBP share an elevated level of sequence conservation at the receptor binding sites. This feature of elevated sequence conservation was not observed using the same approach with MeV-RBP and the narmoviral RBPs, suggestive that MosV-RBP and NarV-RBP do not utilise SLAMF1/nectin-4 receptors.

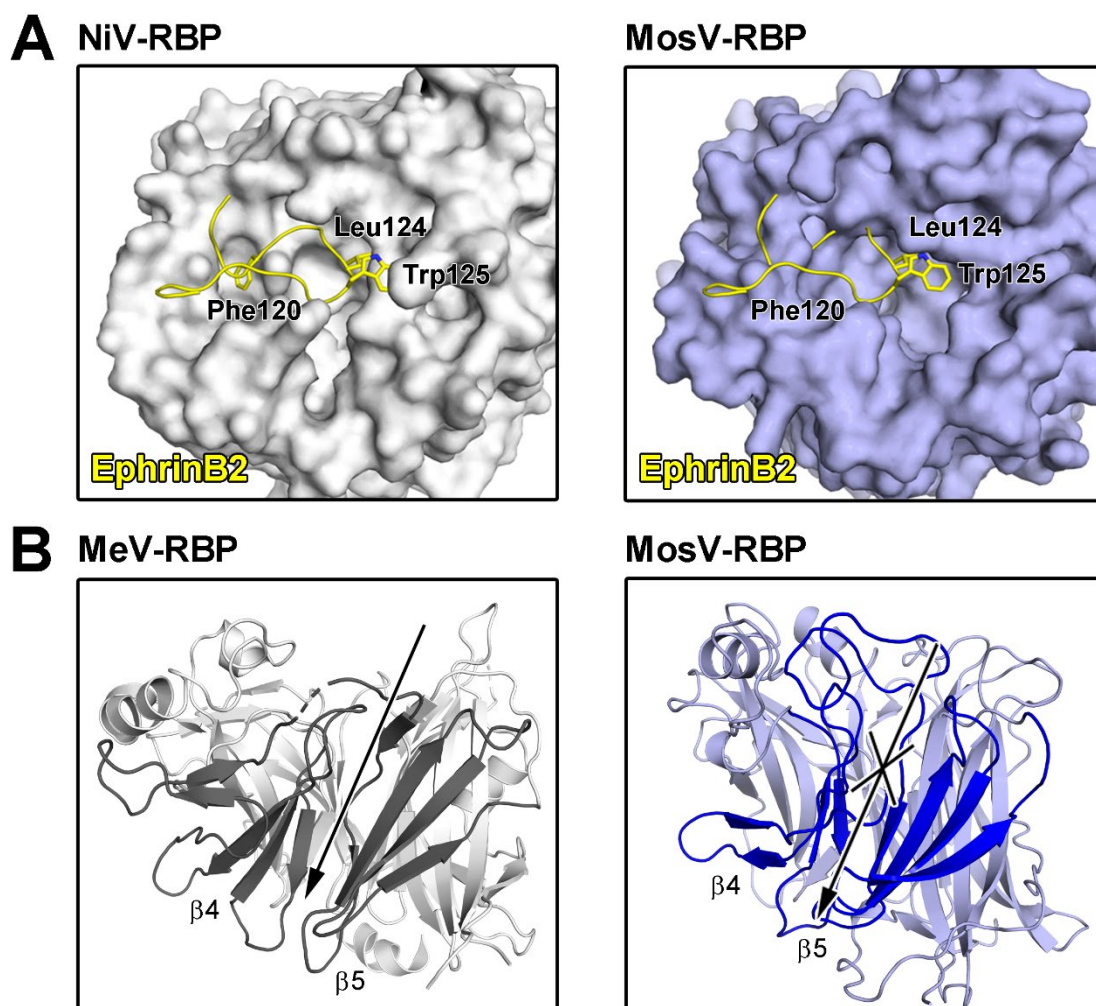


Figure 4.4.3.2ii: The surfaces of MosV-RBP and NarV-RBP appear incompatible to binding characterised paramyxoviral proteinaceous receptors. (a) The surface of NiV-RBP (left) (white), with the G-H loop of ephrinB2 shown in cartoon representation (yellow) (2VSM) (Bowden, Aricescu, et al. 2008). To the right MosV-RBP (blue) is shown in surface representation with the G-H loop docked into the putative binding site, revealing that the loop would be unlikely to bind due to the lack of the necessary pockets required to accommodate residues Phe120, Leu124 and Trp125 of ephrinB2. (b) MeV-RBP (left) shown as grey cartoon with blades $\beta 4$ and $\beta 5$ coloured darker grey (Hashiguchi et al. 2007). The arrow indicates the location of the deep $\beta 4$ - $\beta 5$ groove. MosV-RBP (right) shown as a blue cartoon with darker blue used to represent blades $\beta 4$ and $\beta 5$. The arrow indicates the lack of the characteristic morbilliviral groove.

4.4.3.2 Comparisons to characterised receptor-binding sites

To determine the level of conservation across characterised henipaviral and morbilliviral receptor binding sites, the identical and equivalent residues encoded by the narmoviral RBPs were mapped onto NiV-RBP (2VSM) (Bowden, Aricescu, et al. 2008) and MeV-RBP (3ALZ) (Hashiguchi et al. 2011) structures (**Fig 4.4.3.2i**). Conservation mapping of MosV-RBP and NarV-RBP sequences revealed that in addition to displaying a low sequence identity with the henipaviral RBP β -propeller (23% and 19%, respectively), the ephrinB2 binding site was not maintained, demonstrating only 26% and 28% identity, respectively (**Fig 4.4.3.2iA**). Furthermore, superposition of ephrinB2 (2VSM) (Bowden, Aricescu, et al. 2008) and ephrinB3 (3D12) (Xu et al. 2008) onto MosV-RBP and NarV-RBP structures reveals a non-complementary receptor-binding surface, with ephrinB2/ephrinB3 recognition likely sterically hindered by an extended β 5L01 loop, not present in the HNV-RBP structures (Pryce et al. 2019; Lee et al. 2015; Bowden, Aricescu, et al. 2008; Xu et al. 2008). Additionally, MosV-RBP and NarV-RBP lack the hydrophobic pockets required to accommodate residues Phe120^{EphrinB2}/Tyr120^{EphrinB3}, Leu124, and Trp125 of presented by the G-H loop of ephrinB2 and ephrinB3 (Bowden, Aricescu, et al. 2008; Xu et al. 2008) (**Fig 4.4.3.2iiA**). The presentation of a surface incompatible to B-class ephrin interactions is suggestive that MosV-RBP and NarV-RBP do not utilise the established mode of henipaviral entry, a finding corroborated by flow cytometry data collected by collaborators in the laboratory of Prof. Benhur Lee (Icahn School of Medicine at Mount Sinai, USA), where MosV-RBP and NarV-RBP did not bind Chinese hamster ovary (CHO) cells presenting ephrinB2 (unpublished data).

Similarly, MosV-RBP and NarV-RBP do not maintain identity with MeV-RBP (3ALZ) (Hashiguchi et al. 2011) over the β -propeller (22% and 24%, respectively), or across SLAM-F1 (12% and 9%, respectively) and nectin-4 (7% and 10%, respectively) binding sites (**Fig 4.4.3.2iB**). In addition, they lack the β 4-5 groove integral to wildtype morbilliviral receptor interactions (Hashiguchi et al. 2011; Zhang et al. 2013) (**Fig 4.4.3.2iiB**).

Of the seven residues commonly implicated in sialidase activity, Arg₁, Asp₁, Glu₄, Arg₄, Arg₅, Tyr₆ and Glu₆, both MosV-RBP and NarV-RBP display some identity at the level of primary

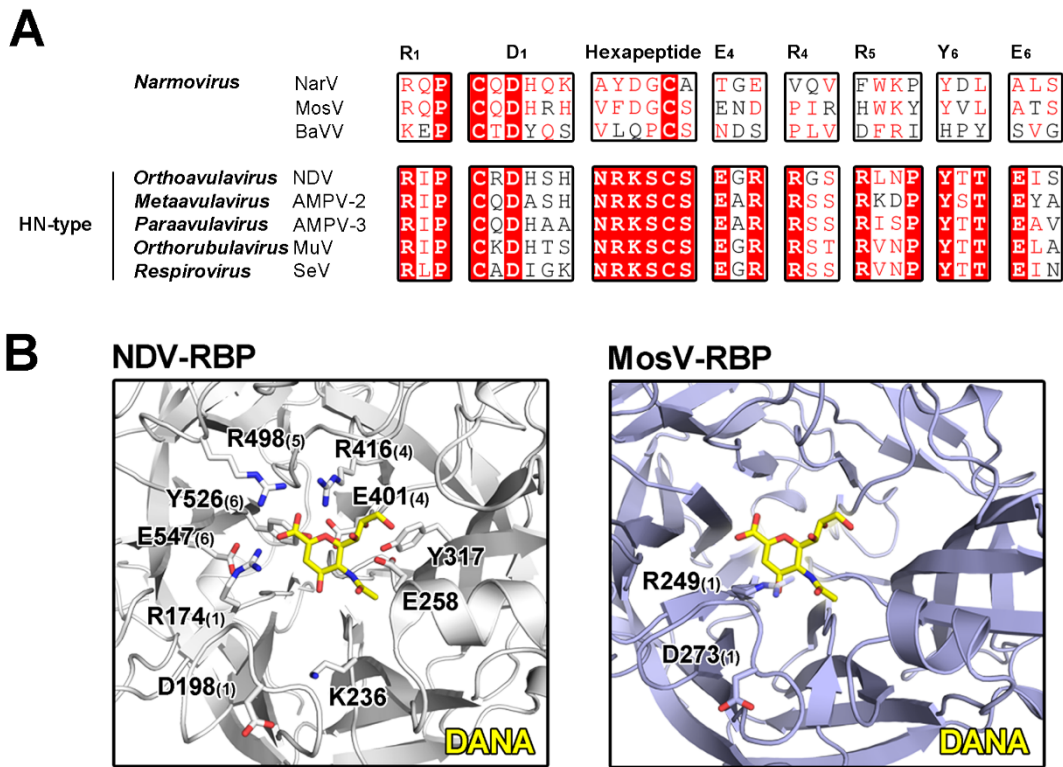


Figure 4.4.3.2iii: Conservation of MosV-RBP and NarV-RBP across the characterised sialic acid interacting site. (a) Alignment of the receptor binding protein (RBP) amino acid sequences from MosV, NarV, bank vole virus (BaVV), Newcastle disease virus NDV, AMPV-2, AMPV-3, MuV and SeV. The seven conserved sialidase residues (Langedijk, Daus, and VanOirschot 1997) and hexapeptide motif (Mirza, Deng, and Iorio 1994) are labelled according to residue and blade location (Langedijk, Daus, and VanOirschot 1997) and annotated above alignments. Genbank accession codes are listed in Appendix D. (b) NDV-RBP (left) (1E8V) shown using cartoon representation (grey) with important sialic acid interacting residues rendered as sticks and annotated (Crennell et al. 2000). Sialic acid analogue N-Acetyl-2,3-dehydro-2-deoxyneuraminic acid (DANA) is represented as sticks and coloured yellow. The equivalent region of MosV-RBP (right) (blue) is shown with DANA superimposed (yellow), also utilising cartoon representation and sticks to illuminate important residues. MosV-RBP lacks the majority of the residues involved in sialic acid interactions.

sequence (Crennell et al. 2000; Iorio et al. 2001; Langedijk, Daus, and VanOirschot 1997). MosV-RBP and NarV-RBP encode the residues, Arg₁ (Arg249 and Arg244, respectively), Asp₁ (Asp273 and Asp269, respectively), Tyr₆ (Tyr600 and Tyr595, respectively), in addition to MosV-RBP also encoding Glu₄ (Glu475) (**Fig 4.4.3.2iiiA**) (Miller et al. 2003; Lambeth et al. 2009). However, structural analysis reveals that both the Glu₄ and Tyr₆ residues of MosV-RBP and NarV-RBP are not presented within the putative sialic acid binding site (**Fig 4.4.3.2iiiB**), and are instead located at the underside of the propeller. Additionally, both proteins present an altered hexapeptide motif, ‘Asn–Arg–Lys–Ser–Cys–Ser’ (Mirza, Deng, and Iorio 1994), with MosV-RBP only presenting the final two residues, ‘Cys–Ser’, and NarV-RBP only retaining the cysteine residue (**Fig 4.4.3.2iiiA**). Nonetheless, despite lack of these integral sialic acid interacting residues, both MosV-RBP and NarV-RBP present a deep central pocket consistent with the site observed in the HN-type RBP structures (Crennell et al. 2000; Lawrence et al. 2004; Yuan et al. 2005; Kubota et al. 2016) (**Fig 4.4.3.2iiiB**).

Observation of such differences across these receptor binding sites, indicates both a structural and functional divergence from the characterised modes of paramyxoviral host-cell engagement.

4.5 ASSESSMENT OF MOSV-RBP AND NARV-RBP SIALIC ACID INTERACTIONS

Structural analysis of MosV-RBP and NarV-RBP indicate that both proteins are unlikely to utilise sialic acid as a host-cell receptor. To assess the ability of MosV-RBP and NarV-RBP to interact with sialic acid, the proteins were applied to hemadsorption and neuraminidase activity assays (Morrison and McGinnes 1989; Potier et al. 1979). Human embryonic kidney (HEK) 293T cells displaying either full-length MosV-RBP or full-length NarV-RBP exhibited no capacity for neuraminidase activity or hemadsorption when compared to an NDV-RBP control (**Fig 4.5**) (Potier et al. 1979; Morrison and McGinnes 1989). The lack of interactions with sialic acid was consistent with the

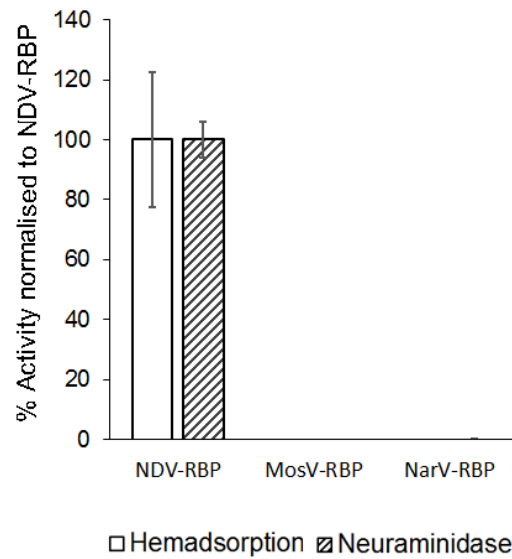


Figure 4.5: MosV-RBP and NarV-RBP display no hemadsorption and neuraminidase activity. (a) MosV-RBP and NarV-RBP neuraminidase (Potier et al. 1979) and hemadsorption (Morrison and McGinnes 1989) activity normalized to cell surface expression and a NDV-RBP control. For both hemadsorption (n = 12) and neuraminidase (n = 6) analysis error bars represent the standard deviation across all replicates.

absence of key sialidase residues, indicating that MosV and NarV likely utilise a sialic acid independent mode of entry.

4.6 COMPARISONS WITHIN THE NARMOVIRUS GENUS

Study of the MosV-RBP and NarV-RBP within the context of the characterised RBPs of the *Paramyxoviridae*, indicates that these viruses are unlikely to utilise any characterised paramyxoviral receptor. However, as observed through sequence conservation mapping of the henipaviral and morbilliviral RBPs, paramyxovirus RBPs tend to maintain conserved receptor binding sites despite high overall sequence variability (Zeltina, Bowden, and Lee 2016). Mapping the sequence conservation between MosV-RBP and NarV-RBP onto the MosV-RBP revealed an extensive elevated region of identity spanning across the two protomers (**Fig 4.6A**). Flow cytometry analysis performed by collaborators in the laboratory of Prof. Benhur Lee have identified that MosV-RBP interacts with both the mouse lung adenoma cell line, LA-4, and the African green monkey kidney

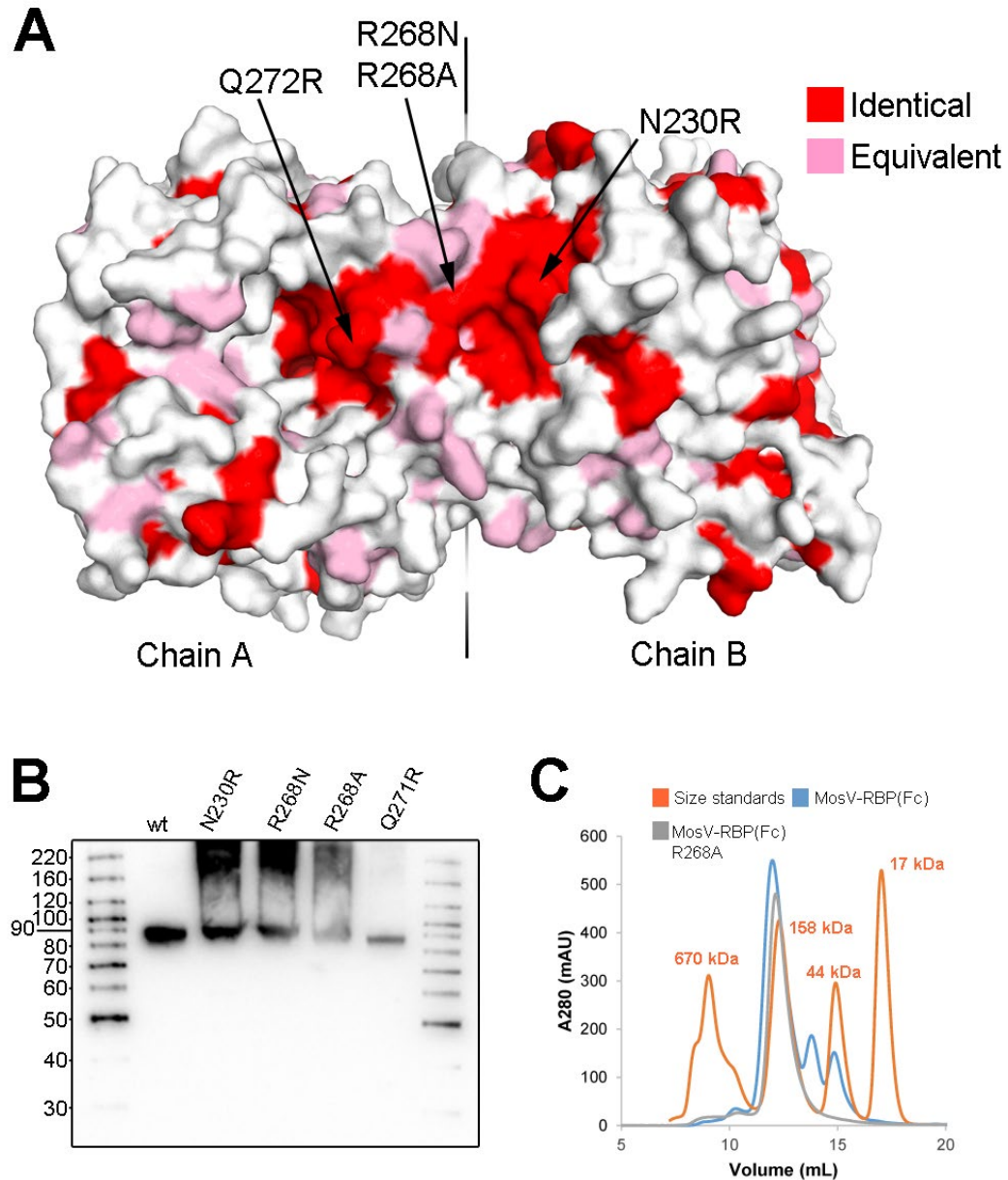


Figure 4.6: Mapping of sequence conservation between MosV-RBP and NarV-RBP onto the surface of MosV-RBP reveals a patch of elevated identity. (a) MosV-RBP is shown as surface with identical and equivalent residues shared with NarV-RBP, coloured red and pink respectively. Residues selected for mutation studies are annotated. (b) Western blot analysis following small-scale expression screening of wildtype (wt) and mutant MosV-RBP conjugated to an Fc in HEK293T cells. (c) Size exclusion analysis shows that MosV-RBP conjugated to an Fc tag, in both wt (blue) and mutant (grey) forms, elutes at a volume consistent with the expected size. The samples were aligned against a size exclusion standard (ID: 1511901, BioRad, orange), with markers at 670, 158, 44 and 17 kDa.

epithelial Vero cell line (unpublished data). Whilst this region of identity was only an indication of a receptor binding site shared by MosV-RBP and NarV-RBP, these data provided a mechanism to study *Narmovirus* receptor interactions further through site-directed mutagenesis and subsequent flow cytometry analysis.

4.6.1 Production of mutant MosV-RBPs

To assess the impact of the patch of identity shared by MosV-RBP and NarV-RBP in cell binding, site-directed mutagenesis of MosV-RBP was performed in order to introduce substitutions across this region. Three residues were selected for initial substitution, with those chosen predicted to cause either a chemical or structural change to the surface. The substitutions were as follows: Asn230Arg, Arg268Asn, Arg268Ala, and Gln271Arg (**Fig 4.6A**). Utilising the full-length human codon-optimised MosV-RBP gene, the substitutions were introduced by splicing by overlap extension (SOE) polymerase chain reaction (PCR) (Heckman and Pease 2007), and mutated fragments ligated into the pHLsec vector (Aricescu, Lu, and Jones 2006). Constructs for soluble production of wildtype (wt) and mutant MosV-RBPs were generated through PCR of Thr157–Thr632 from the full-length MosV-RBP backbones. These were subsequently inserted into a pHLsec vector with a C-terminal Fc tag and hexa-histidine tag encoded, enabling purification and analysis by flow cytometry. The ability of the constructs to produce soluble protein was assessed by small-scale expression screening in HEK293T cells (**Fig 4.6B**).

Large-scale production of wt and mutant MosV-RBP (Thr157–Thr632)-Fc-6xHis was performed through transient transfection of HEK293T cells. Cell culture supernatants were clarified by centrifugation, filtered and subject to buffer exchange into 20mM sodium phosphate pH 7.0. The concentrated and exchanged supernatants were purified using protein G affinity chromatography with a HiTrap® Protein G HP (GE Healthcare) column. Subsequently, fractions comprising the peak were concentrated and applied to a Superdex™ 200 Increase 10/300 GL column for gel filtration. Consistent with dimeric and Fc tagged MosV-RBP molecules, both the mutagenised and wt protein presented a molecular mass larger than that observed with the non-Fc tagged version (**Fig 4.6C**).

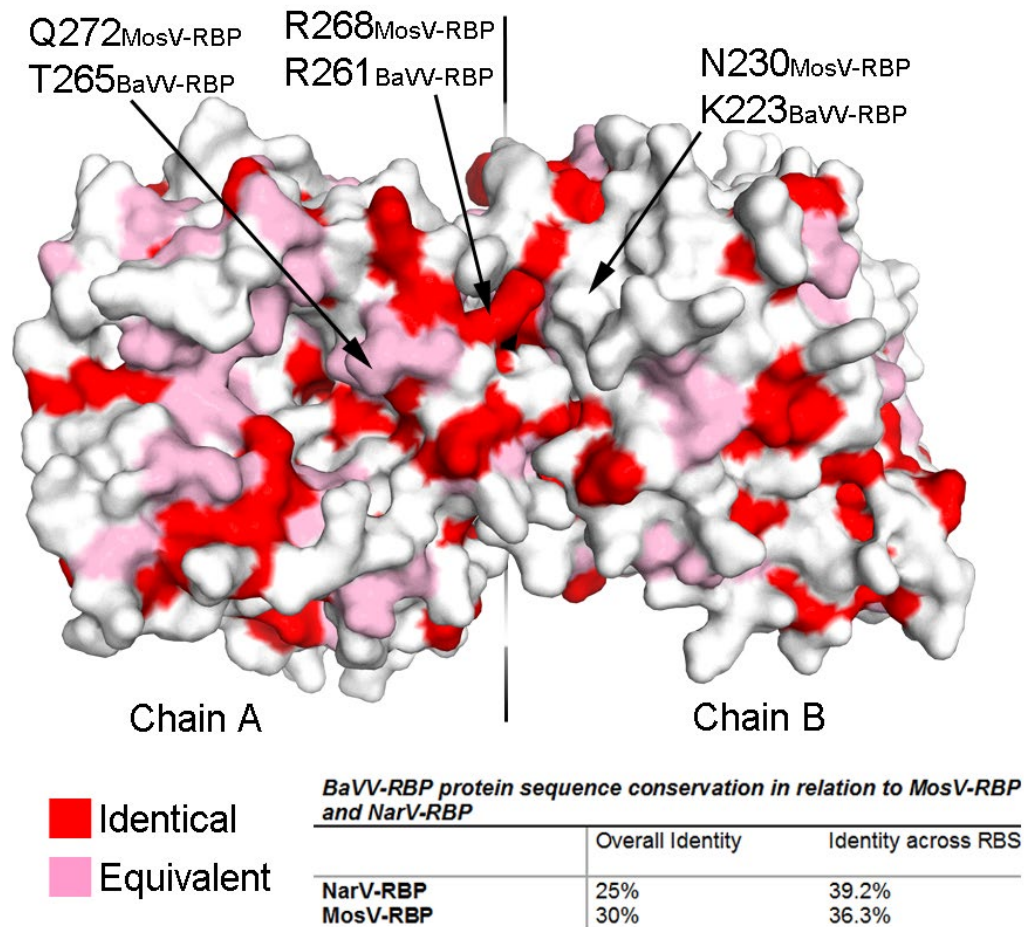


Figure 4.6.2: The potential receptor binding site (RBS) is as pronounced following mapping of sequence conservation between MosV-RBP and BaVV-RBP onto the surface of MosV-RBP. (a) MosV-RBP is shown as surface with identical and equivalent residues shared with BaVV-RBP, coloured red and pink, respectively. The MosV-RBP residues selected for site-directed mutagenesis studies and the corresponding BaVV residues are annotated. A table outlines the overall identity and identity across the putative RBS between BaVV-RBP and the structurally characterised MosV-RBP and NarV-RBP.

These samples were sent to be analysed by flow cytometry in the laboratory of Prof. Benhur Lee, where experiments are ongoing.

4.6.2 Comparison of MosV-RBP and NarV-RBP to the recently discovered bank vole virus RBP

Following generation of the MosV-RBP mutants, work was published concerning the genetic characterisation of a third *Narmovirus* from Russia, BaVV (Alkhovsky et al. 2018; Amarasinghe et al. 2019). Sequence conservation mapping revealed that the RBP of BaVV had a slight elevation of sequence identity across the putative *Narmovirus* RBP binding site (**Fig 4.6.2**), though not to the extent observed between MosV-RBP and NarV-RBP. A number of residues that form the region of elevated identity observed between MosV-RBP and NarV-RBP are not conserved, in particular, BaVV-RBP lacks a Asn at the position corresponding to 230^{MosV-RBP} (**Fig 4.6.2**), one of the mutants selected for flow cytometry analysis. Given this, future mutagenesis studies concerning the narmoviral RBPs may benefit from the consideration of BaVV-RBP identity in addition to MosV-RBP and NarV-RBP.

4.7 DISCUSSION

Explorations into viruses harboured by various wildlife reservoirs has begun to unveil the unprecedented diversity of paramyxoviruses that fall outside the existing genera (Thibault et al. 2017). One major reservoir of paramyxoviruses are populations of rodents, which due to their extensive inter- and intra-species behaviours are amplifiers of viral diversity (Thibault et al. 2017; Drexler et al. 2012). The existence of these reservoirs has motivated our work to model and understand such novel viruses isolated from rodent species (Rissanen et al. 2017). MosV and NarV represent the narmoviruses, a recently defined paramyxoviral genus discovered in numerous rodent species worldwide (Lambeth et al. 2009; Miller et al. 2003; Amarasinghe et al. 2019). The molecular interactions between the paramyxoviral RBP and their cognate receptors is an important determinant of a virus's propensity for interspecies transmission (Zeltina, Bowden, and Lee 2016; Thibault et al. 2017). Concerning the RBP, I proceeded to determine the structures for both MosV-RBP and NarV-

RBP, utilising existing knowledge of paramyxoviral receptor binding sites to determine the structural independence of these viruses, supporting the recent establishment of the genus, *Narmovirus* (Amarasinghe et al. 2019).

Structural study of MosV-RBP and NarV-RBP revealed that whilst the head domains of the RBPs retained the conserved six-bladed β -propeller fold, the RBPs lacked the necessary residues and complimentary surface required for characterised paramyxoviral receptor interactions. Both MosV-RBP and NarV-RBP did not present the $\beta 4$ - $\beta 5$ groove integral to the interaction between MeV and its receptors (Hashiguchi et al. 2011; Zhang et al. 2013; Santiago et al. 2010) and the characteristic B-class ephrin binding pocket presented by the henipaviral RBPs (Bowden, Aricescu, et al. 2008; Xu et al. 2008). This apparent inability to interact with B-class ephrin receptors is consistent with data collected by collaborators from the laboratory of Prof. Benhur Lee (Icahn School of Medicine at Mount Sinai), where binding of soluble MosV-RBP and NarV-RBP to ephrinB2 presenting CHO cells was not detected. Reflecting this independence of the narmoviral RBPs, an absence of extensive glycosylation was not characteristic of paramyxoviral RBPs (Bowden, Crispin, et al. 2008; Sawatsky and von Messling 2010; Hu et al. 1994; Segawa et al. 2003; Ng, Hiebert, and Lamb 1990; McGinnes and Morrison 1995; Panda et al. 2004; Chu et al. 2013). It was also found that the narmoviral RBPs also lacked a number of residues required for binding and hydrolysing sialic acid, with an inability to interact with sialic acid confirmed through hemadsorption and neuraminidase activity analysis. Additionally, the distinct RBP does not align closely with other structurally characterised RBPs, as demonstrated utilising pairwise overlay analysis and the constructed structure-based phylogeny. Furthermore, the oligomeric organisation of MosV-RBP and NarV-RBP is unlike dimeric associations observed for the HN-type RBPs, where the angle of association is approximately 60° (Yuan et al. 2005; Yuan et al. 2008; Lawrence et al. 2004; Kubota et al. 2016; Crennell et al. 2000). Divergence of the dimeric organisation has been observed in the HeV and MeV structures, most likely as a requirement to accommodate larger receptors (Bowden, Crispin, Harvey, et al. 2010; Hashiguchi et al. 2007; Bowden, Aricescu, et al. 2008; Hashiguchi et al. 2011; Zhang et al. 2013; Santiago et al. 2010), however, the narmoviral RBP dimer with an

approximate 90° angle of association is also distinct from these protein binding RBPs. Together these data outline a number of structural differences which support the recent distinction of the narmoviruses (Amarasinghe et al. 2019).

Sequence conservation between MosV-RBP and NarV-RBP, when mapped onto the MosV-RBP, surface reveal the existence of a patch of elevated identity stretching across the two protomers of the dimer despite sequence variability between the two. The potential usage of this site for receptor interactions would be consistent with the dimeric organisation of the narmoviral dimer, and furthermore resembles the binding sites utilised by other RBPs in that such sites are conserved between the different RBPs within a specific genus (Zeltina, Bowden, and Lee 2016). The discovery of this site, and the flow cytometry data provided by our collaborators in the laboratory of Prof. Benhur Lee offered a means to potentially study the viruses further through mutagenesis of residues across this region. Following production of mutant MosV-RBP, protein samples were sent to our collaborators for further testing. However, the recent molecular characterisation of BaVV (Alkhovsky et al. 2018), a member of the *Narmovirus* genus (Amarasinghe et al. 2019), revealed that some of the selected mutations were not conserved across the BaVV-RBP surface. Future efforts to characterise the receptor binding site utilised by this group of viruses will undoubtedly account for this newly discovered virus. Furthermore, given the structural plasticity of the paramyxoviral β -propellers (Bowden, Crispin, Jones, et al. 2010), another region on the RBP surface may form the receptor binding site, and it may be that MosV-RBP, NarV-RBP, and BaVV-RBP do not interact with the same receptor. Further work with a view to identifying the receptors utilised by these viruses is ongoing and covered in chapter 6.

4.8 CONCLUSIONS

The structural characterisation of MosV-RBP and NarV-RBP support the recent independent classification of these viruses within the newly defined genus, *Narmovirus* (Amarasinghe et al. 2019). Furthermore, the narmoviral RBP surface is distinct from those already characterised, and MosV-RBP and NarV-RBP present a potential common binding site requiring further study. Overall

this data expands our appreciation of the diversity of RBPs exhibited by paramyxoviruses. However, further work is certainly required to further characterise these viruses and assess their potential for interspecies transmission, with particular attention to receptor interactions.

CHAPTER 5

THE J PARAMYXOVIRUS AND BEILONG VIRUS RECEPTOR BINDING PROTEINS

5 THE J PARAMYXOVIRUS AND BEILONG VIRUS RECEPTOR BINDING PROTEINS

5.1 SUMMARY OF FINDINGS

The paramyxoviruses of the newly defined genus, *Jeilongvirus*, within the *Paramyxoviridae* family, have the longest characterised genome of all viruses of the order *Mononegavirales* (Amarasinghe et al. 2019). A number of features are distinct within this uniquely large genome, including an extended receptor binding protein (RBP) (Jack et al. 2005; Vanmechelen, Bletsa, Laenen, Lopes, Vergote, Beller, Deboutte, Korva, Avšič Županc, et al. 2018). Indeed, in addition to exhibiting the characteristic paramyxoviral RBP domain architecture, jeilongvirus RBPs also have an extended C-terminus of varying length (139–1,014 amino acids) and unknown function. Sequence analysis reveals that J paramyxovirus (JPV) and beilong virus (BeiV) RBPs present residues known to be integral for sialic acid interactions (Langedijk, Daus, and VanOirschot 1997; Crennell et al. 2000; Iorio et al. 2001), despite a previous study suggesting that JPV likely lacks neuraminidase activity (Jun, Karabatsos, and Johnson 1977). Consistent with a close functional relationship of jeilongvirus RBPs with sialic acid binding RBPs, my X-ray crystallographic analysis reveals that JPV-RBP and BeiV-RBP structures align closely with hemagglutinin neuraminidase (HN)-type RBPs. Furthermore, JPV-RBP and BeiV-RBP maintain the structurally conserved primary sialic acid binding pocket on the top face of the propeller (Crennell et al. 2000; Lawrence et al. 2004; Yuan et al. 2005; Kubota et al. 2016), along with a potential secondary binding site at the dimeric interface (Zaitsev et al. 2004). Interestingly, analogous to the RBP structure from the Ulster strain of NDV, which presents a C-terminus involved in inhibiting sialic acid associated activity (Yuan et al. 2012), structural characterisation of the BeiV-RBP C-terminal segment reveals a ‘hat-like’ domain that extends into the putative sialic acid interaction site at the centre of the six-bladed β -propeller fold. Additionally, despite the C-terminal segment of JPV-RBP being disordered in the crystal, electron density potentially corresponding to a C α chain within the putative sialic acid site suggests that the JPV-RBP C-terminal extension may be structurally employed for a similar auto-inhibitory role.

Hemadsorption and neuraminidase activity analysis confirmed that JPV-RBP and BeiV-RBP presenting the full C-terminal extension are unable to interact with sialic acid. Subsequent analysis of JPV-RBP and BeiV-RBP C-terminal truncations has not been possible due to the timing of this work within this thesis, therefore the auto-inhibition hypothesis was not tested. Nonetheless, these structural findings provide evidence for a close evolutionary relationship of jeilongvirus RBPs with HN-type RBPs, and support the hypothesis that the extended C-terminus facilitates auto-inhibition functionality.

5.2 INTRODUCTION

Isolated from feral mice in Queensland, Australia, in the 1970s (Jun, Karabatsos, and Johnson 1977; Mesina et al. 1974), JPV is the first known member of the recently defined genus, *Jeilongvirus* (Amarasinghe et al. 2019). At the time of molecular characterisation, JPV presented the greatest known divergence from the typical six-gene paramyxovirus genome structure, and had the longest known genome of the order, *Mononegavirales* (Jack et al. 2005). Subsequently, BeiV was isolated following the detection of cDNA clones with sequence homology to paramyxoviral genes during a subtractive hybridisation experiment focused on identifying genes upregulated by Angiotensin II in human kidney mesangial cells (HMCs) (Schomacker, Collins, and Schmidt 2004; Liang et al. 2003; Basler, García-Sastre, and Palese 2005; Li et al. 2006b).

The extensively long JPV genome encodes an additional small hydrophobic (SH) protein, TM protein, an elongated RBP, and an additional open reading frame (ORF), termed ‘X’. In addition to the SH and TM genes, the BeiV also encodes two further ORFs with unknown functionality, termed ‘X1’, a proline, serine and threonine (P/T/S)-rich protein, and X2’. Analysis of a variety of laboratory cell lines during the original gene expression study revealed that the likely origin of BeiV was a rat mesangial cell (RMC) line from 1998 (Li et al. 2006b). Consistent with this hypothesis, BeiV was identified in black and brown rats in Hong Kong, China, and is thought to be endemic in these populations (Woo, Lau, Wong, Wu, et al. 2012). Subsequent molecular characterisation of these rat isolated variants revealed the absence of a stop codon at position 2,205 of the BeiV RBP

gene, resulting in an RBP 1,046 amino acids in length, as opposed to the 734 observed from the laboratory isolated BeiV (Woo et al. 2016b). The absence of the stop codon results in the incorporation of what was previously termed ORF 'X1'. Subsequent investigation found that unlike JPV-RBP and BeiV(HMC)-RBP, a number of newly identified jeilongviruses include the P/T/S-rich 'X1' region (Vanmechelen, Bletsa, Laenen, Lopes, Vergote, Beller, Deboutte, Korva, Avšič Županc, et al. 2018). Herein, BeiV will be variously appended with 'HMC' or 'rat' to denote origin, when necessary.

The identification and isolation of further members of the *Jeilongvirus* genus has expanded our understanding of the genome structures presented by these viruses and increased our appreciation of their widespread presence in animal populations worldwide. This includes the isolation of Tailam virus (TlmV) from Sikkim rats in China (Woo et al. 2011b), Mount Mabu Lophuromys paramyxoviruses 1 and 2 (MMLPV-1 and -2) from Rungwe brush-furred rats in Mozambique and Pohorje Myodes paramyxovirus 1 (PMPV-1) from bank voles in Slovenia (Vanmechelen, Bletsa, Laenen, Lopes, Vergote, Beller, Deboutte, Korva, Avšič Županc, et al. 2018). Other studies have identified the presence of jeilongviruses in other bat, rodent and shrew populations (Noh et al. 2018; Kurth et al. 2012; Maganga et al. 2014; Sasaki et al. 2014; Mortlock et al. 2019), however, oftentimes full genetic characterisation of these viruses has not been performed. Furthermore, the potential for interspecies transmission is demonstrated by JPV. Indeed, initial characterisation of this virus led to the isolation of serum neutralising antibodies from mice, rats, pigs, cattle and humans, however with the exception of rodents JPV infection was not associated with disease (Jack et al. 2005; Jun, Karabatsos, and Johnson 1977). Therefore, despite the continued identification and isolation of jeilongviruses in numerous species, worldwide, the disease potential and evolutionary origin of these viruses remains unknown. Given the role of the unique jeilongviral RBP in dictating tissue and species tropism (Thibault et al. 2017), studies of this protein are essential for clarifying jeilongvirus pathobiology.

Similar to all paramyxoviruses, jeilongviruses utilise an RBP and fusion (F) protein to mediate viral entry (Jardetzky and Lamb 2014). The RBP is presented on the paramyxoviral surface

as a dimer-of-dimers (Yuan et al. 2008; Yuan et al. 2011), with each protomer formed of an intraviral (IV) portion, transmembrane (TM) domain, α -helical stalk, and a six-bladed β -propeller head domain, which interacts with the host-cell receptor. The primary amino acid sequence of the six-bladed β -propeller head domain of JPV-RBP and BeiV-RBP is most closely related to hemagglutinin neuraminidase (HN)-type β -propellers, with JPV-RBP and BeiV-RBP demonstrating 27% and 26% sequence identity to parainfluenza virus 3 (PIV3)-RBP, respectively. Importantly, the RBPs of jeilongviruses typically retain the majority of the residues implicated in the sialic acid associated interactions characteristic of HN-type RBPs. The seven conserved sialidase residues presented by HN-type RBPs are: Arg₁, Asp₁, Glu₄, Arg₄, Arg₅, Tyr₆, and Glu₆ (sub-script refers to location on blades 1–6 of the β -propeller fold) (Iorio et al. 2001; Langedijk, Daus, and VanOirschot 1997; Crennell et al. 2000). RBPs from JPV, BeiV (both HMC and rat), MMPV-2, PMPV-1, and TImV present residues Arg₁, Glu₄, Arg₄, Arg₅, Tyr₆, and Glu₆ (**Fig 5.2iA**). The typically conserved Asp₁ is much more variable, with only MMPV-2 maintaining this residue. However, in the context of the jeilongviral RBPs, Asp₁ is often conservatively substituted with the similar residue, Glu, as is the case for BeiV-RBP. In addition, the importance of Asp₁ in mediating sialic acid activity of the related mumps virus (MuV) has been debated (Kubota et al. 2016). MMPV-1 is an outlier within the family, in that it only retains the sialidase associated residues, Arg₁ and Tyr₆. Similarly, the hexapeptide motif (Asn–Arg–Lys–Ser–Cys–Ser), which is integral to the enzymatic activity of paramyxoviral HN-type RBPs (Mirza, Deng, and Iorio 1994), displays only a small level of divergence with jeilongviral RBPs, where Lys/Arg and Ser/Thr are variously presented at the 3rd and 6th positions of the hexapeptide, respectively (**Fig 5.2iA**).

In addition to the well conserved sialic acid binding site formed by the seven sialidase residues and the hexapeptide motif, some HN-type paramyxoviral RBPs present a secondary sialic acid binding site associated with triggering fusion (Zaitsev et al. 2004; Porotto et al. 2012b; Porotto et al. 2007; Yuan et al. 2012). Structural studies of NDV-RBP reveal that this site is located at the dimer interface and is composed of residues from both propellers, primarily including, Gly169^{NDV}-

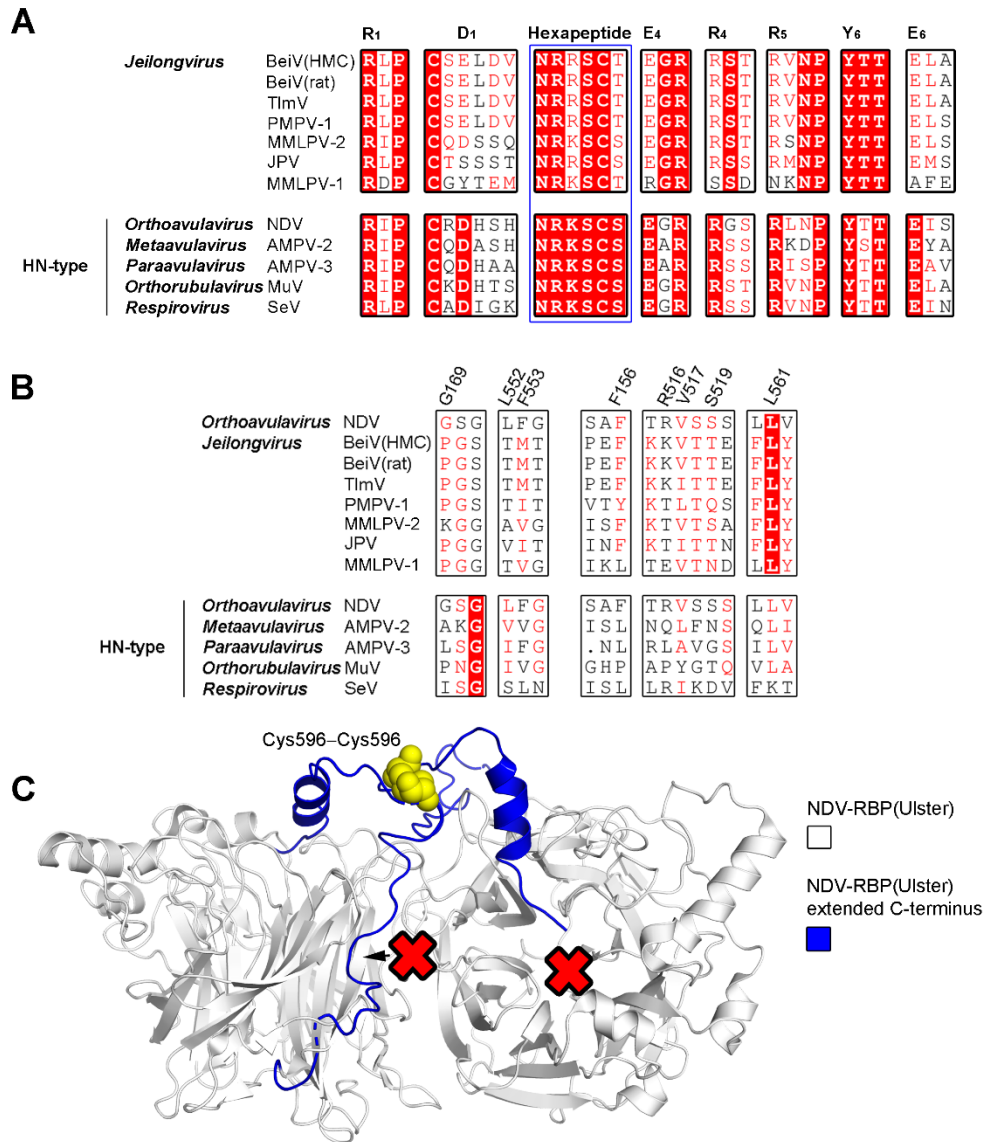


Figure 5.2i: Amino acid sequence alignment of regions known to confer sialic acid binding and hydrolysis functionality. (a) Alignment of the receptor binding protein (RBP) amino acid sequences from Newcastle disease virus (NDV), human kidney mesangial cell (HMC) and rat isolated beilong virus (BeiV), Tailam virus (TlmV), Pohorje Myodes paramyxovirus 1 (PMPV-1), Mount Mabu Lophuromys paramyxovirus 1 and 2 (MMLPV-1/2), J paramyxovirus (JPV), avian paramyxoviruses 1 and 2 (AMPV-2/3), mumps virus (MuV) and Sendai virus (SeV). Residues associated with primary sialic acid interactions, including the seven conserved sialidase residues (Langedijk, Daus, and VanOirschot 1997) and hexapeptide motif (Mirza, Deng, and Iorio 1994; Li et al. 2006b), are labelled according to residue and blade location (Langedijk, Daus, and VanOirschot 1997) and annotated above alignments. (b) Alignment of RBP sequences as above with residues associated with secondary sialic acid interactions annotated (Zaitsev et al. 2004). Genbank accession codes are listed in Appendix D. (c) Structure of the NDV-RBP Ulster strain (Yuan et al. 2012) (4ZFH), with the head-domain (white) and C-terminal extension (blue) shown in cartoon representation, and interchain disulphide bonds represented as spheres (yellow). The positions of the primary (right) and secondary (left) sialic acid recognition sites are labelled with a red cross, with the C-terminal extension blocking both these sites.

RBP, Leu552^{NDV-RBP}, Phe553^{NDV-RBP} from one monomer, and Phe156^{NDV-RBP}, Arg516^{NDV-RBP}, Val517^{NDV-RBP}, Ser519^{NDV-RBP} and Leu561^{NDV-RBP} from the other (Zaitsev et al. 2004). Whilst this sialic acid binding site has only been shown to be definitively functionally relevant in the case of the *Orthoavulavirus*, NDV (Zaitsev et al. 2004; Porotto et al. 2012b; Bousse et al. 2004), there is the potential for other paramyxoviruses, such as the *Respirovirus*, PIV3, to interact via this second sialic acid binding site (Mishin et al. 2010b; Porotto et al. 2007; Porotto et al. 2012b). It appears that the majority of jeilongviral RBPs retain either identical or equivalent residues at both the secondary and primary sialic acid binding sites, with the exception of the more divergent MMPV-1 (**Fig 5.2iB**). It is not clear how such small changes to these motifs impact the potential for jeilongviral RBPs to utilise sialic acid as a receptor. However, a preliminary investigation of JPV indicates that the virus lacks the expected neuraminidase activity associated with viral exit (Jack et al. 2005; Jun, Karabatsos, and Johnson 1977).

This apparent disconnect between the HN activity of jeilongviral RBPs and the presentation of sialic acid interacting residues would be consistent with an auto-inhibitory mechanism, such as that observed in the case of the RBP from Newcastle disease virus (NDV) Ulster strain (Yuan et al. 2012; Nagai and Klenk 1977) (**Fig 5.2iC**). For a subset of avirulent NDV strains, including NDV Ulster, proteolytic cleavage of a C-terminal segment of the RBP was required for the restoration of hemadsorption activity. Structural analysis reveals that this 45 amino acid C-terminal extension wraps around the NDV-RBP propeller, and is stabilised at the top of the dimeric interface by an interchain disulphide bond (chain A Cys596–chain B Cys596). The residues at the extremity of the C-terminus utilise both polar and non-polar contacts to insert into the sialic acid interaction site, thereby preventing receptor recognition (Yuan et al. 2012). In addition, residues Glu584–Lys587 of this extension occlude the secondary sialic acid binding site at the interface of the NDV dimers. Beyond the conserved six-bladed β -propeller presented by all known paramyxoviral RBPs, the jeilongvirus RBP encodes a similarly distinct extended C-terminus with uncharacterised structure or function. Whilst there is no sequence homology between the C-terminal extension of the NDV Ulster

A

Virus and protein	C-terminus length	C-terminus P/T/S content (%)			C-terminal O-linked glycosylation	Conservation residues 575-700 ^{BeiV(HMC)-RBP} (%)
		P	T	S		
BeiV(HMC)-RBP	159	8.8	6.3	6.3	0	-----
BeiV(wt)-RBP	471	8.9	12.1	12.1	93	99.2
JPV-RBP	139	5.8	7.2	7.2	0	18.4
MMLPV-1-RBP	287	6.6	8.4	17.4	49	33.3
MMLPV-2-RBP	234	2.6	9.8	19.2	35	22.0
TImV-RBP	477	5.9	9.6	14.3	107	82.5
PMPV-1-RBP	1,014	10.6	8.2	12.8	153	38.3

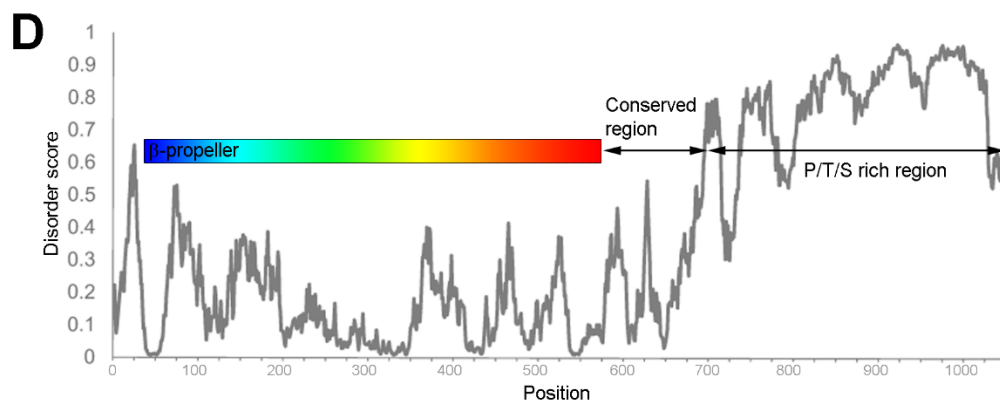
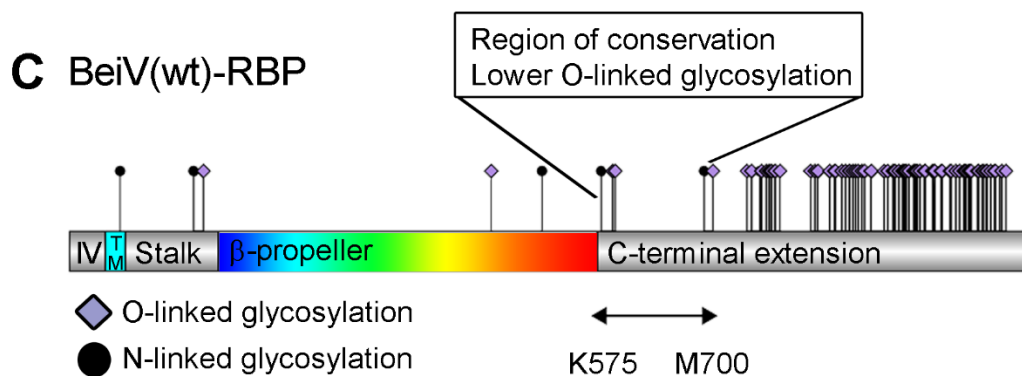
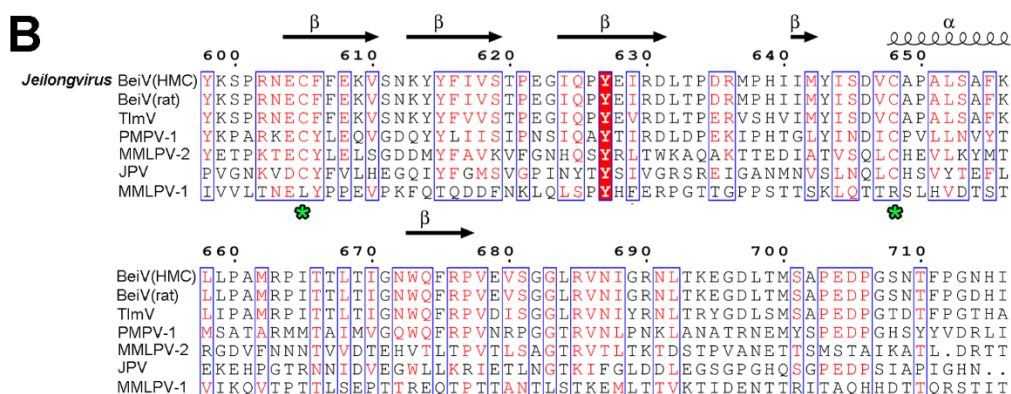


Figure 5.2ii: Sequence analysis of the C-terminal extension found on Jeilongvirus RBPs (a)

A table outlining the differences in length, proline/threonine/serine (P/T/S) content, and predicted O-linked glycosylation of the C-terminal extensions of jeilongviral RBPs. The final column details conservation between residues BeiV(HMC)-RBP K575–M700, and the analogous region in other jeilongviral RBPs. **(b)** Amino acid sequence alignment of the initial region of the jeilongviral RBP C-terminal extensions, revealing a conserved disulphide (green star). MMPLPV-1 is an exception, showing much more limited conservation across this region. **(c)** Gene diagram of BeiV(rat)-RBP with IV, TM, stalk and the six-bladed β -propeller (rainbow N- to C- terminus) annotated. In addition, the predicted O-linked and N-linked sequons are marked with purple diamonds and black circle pins, respectively. The initial region of the extended C-terminus aligned in the above panel B is annotated. **(d)** The BeiV(rat)-RBP disorder score for each amino acid position calculated using the ‘Prediction of Intrinsically Unstructured Proteins 2A’ (IUPred2A) server (Dosztányi 2018), shows an increase in disorder within the latter part of the C-terminal extension). The amino acids corresponding to the 6-bladed β -propeller are indicated with a box coloured as in panel C. The jeilongviral C-terminal region that displays amino acid conservation (K575-M700), and P/T/S rich region are indicated with arrows.

strain RBP and the jeilongviral RBPs (Jack et al. 2005; Li et al. 2006b; Vanmechelen, Bletsa, Laenen, Lopes, Vergote, Beller, Deboutte, Korva, Avšič Županc, et al. 2018), application of a C-terminal extension to the sialic acid site could be a common method of auto-inhibition to mediate pathogenicity, and therefore promote virus survival. However, the large size of the jeilongviral RBP indicates that these viruses potentially utilise the C-terminal extension for further diverse functionalities.

Within the *Jeilongvirus* genus, the C-terminal extensions of the RBPs display high levels of sequence diversity relative to the overall RBP, and vary considerably in length (139–1,014) (Vanmechelen, Bletsa, Laenen, Lopes, Vergote, Beller, Deboutte, Korva, Avšič Županc, et al. 2018) (**Fig 5.2iiA**). In addition to the high levels of sequence variation, the C-terminal extension is enriched with P/T/S residues (**Fig 5.2iiA**), a feature associated with extensive O-linked glycosylation (Carraway and Hull 1991) (**Fig 5.2iiC**). This feature has not been observed in paramyxoviruses beyond the *Jeilongvirus* genus, but is characteristic of the mucin-like domains of the attachment glycoprotein (G) from the related pneumoviruses, such as respiratory syncytial virus (RSV) and human metapneumovirus (hMPV) (Melero, Mas, and McLellan 2017; Leyrat et al. 2014; Wertz et al. 1985). These polypeptide chains with heavy O-linked glycosylation are structurally elongated, resulting in intrinsic disorder (Jentoft 1990; Leyrat et al. 2014) (**Fig 5.2iiD**). In combination with the

high level of sequence variability and heavy glycosylation, the steric hindrance caused by the elongated G protein presented by pneumoviruses, has been hypothesised to shield the pneumoviral F protein from the antibody-mediated immune response (Leyrat et al. 2014). Similarly, the Ebola virus glycoprotein (GP), utilises a mucin-like domain as a mechanism to mask epitopes (Francica et al. 2010; Reynard et al. 2009). Structural characterisation of these domains is inherently complicated, and similarly, these P/T/S enriched domains at the C-terminus of the jeilongviral RBPs are likely disordered (Leyrat et al. 2014). However, all the jeilongviral RBPs have a region prior to the P/T/S region, which whilst not as highly conserved as the propeller domain, demonstrates a conserved disulphide binding pattern (with the exception of MMLPV-1) (**Fig 5.2iiB**). Furthermore, this apparent P/T/S enrichment, and the associated predicted O-linked glycosylation, does not extend into this region (**Fig 5.2iiC**). Unlike other members of the jeilongvirus family, JPV-RBP and BeiV(HMC)-RBP altogether lack the variable P/T/S-rich region and only retain this initial conserved C-terminus with unknown functionality (Vanmechelen, Bletsa, Laenen, Lopes, Vergote, Beller, Deboutte, Korva, Avšič Županc, et al. 2018).

Considering the unique sequence of the jeilongviral RBPs, and the unknown role(s) of the C-terminal extension, I sought to understand the RBP structures of two archetypal, yet genetically disparate, jeilongviruses, JPV and BeiV. This investigation revealed that the six-bladed β -propellers of the RBPs are most closely related to the HN-type RBP propellers. Interestingly, the C-terminal extension of BeiV(HMC)-RBP was found to occupy the putative sialic acid binding site. Similarly, despite an inability to structurally characterise the C-terminal extension of JPV-RBP, density consistent with a C α chain was observable in the potential sialic acid interaction site. Additionally, functional analysis confirmed that JPV-RBP and BeiV(HMC and rat)-RBP bearing the C-terminal extension, did not exhibit neuraminidase or hemadsorption activity. Despite soluble expression of truncated JPV and BeiV RBPs, subsequent efforts to determine the potential role of the C-terminal extension were hindered by timing issues. However, despite incomplete functional analysis, these structural findings may rationalise the apparent lack of sialic acid associated activity demonstrated

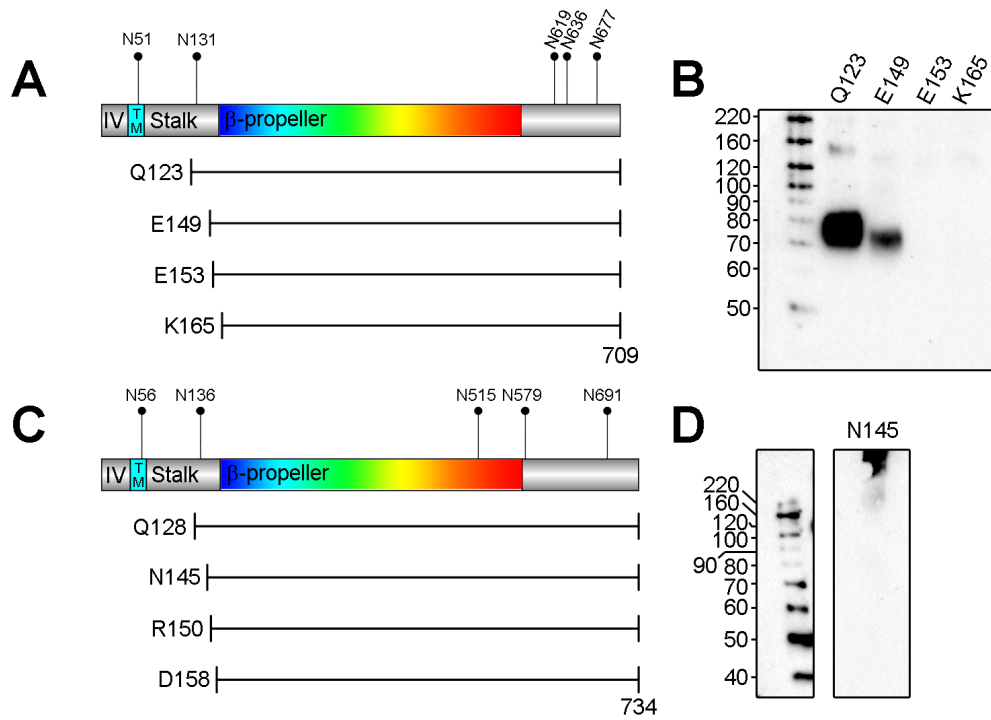


Figure 5.3.1i: Construct design and small-scale expression test of JPV-RBP and BeiV(HMC)-RBP. (a) Gene diagram displaying features of the JPV-RBP, including the IV region, TM domain, stalk, six-bladed β -propeller receptor binding head (rainbow), and C-terminal extension. Predicted N-linked glycosylation sites (NXT/S, where X $\neq$ P) are marked with pins and relevant asparagine residue numbered. The designed constructs are illustrated below the schematic. (b) Western blot of small scale expression test of constructs, where an SDS-PAGE gel of HEK293T cell supernatants was performed under reducing conditions. Western blot analysis was performed alongside the BenchmarkTM Protein Ladder (Thermo, left). (c) Gene diagram displaying features of the BeiV(HMC)-RBP, including the IV region, TM domain, stalk, six-bladed β -propeller receptor binding head (rainbow), and C-terminal extension. Predicted N-linked glycosylation sites (NXT/S, where X $\neq$ P) are marked with pins and relevant asparagine residue numbered. The designed constructs are illustrated below the schematic. (d) Western blot of small scale expression test of construct Asn145–Glu734, where an SDS-PAGE gel of HEK293T cell supernatants was performed under reducing conditions. Western blot analysis was performed alongside the BenchmarkTM Protein Ladder (Thermo, left).

by JPV-RBP and BeiV-RBP, where the initial segment of the jeilongviral RBP C-terminal extension performs a putative auto-inhibitory role.

5.3 PROTEIN PRODUCTION AND STRUCTURE DETERMINATION OF BEIV-RBP AND JPV-RBP

5.3.1 Construct design and small-scale expression testing

For production of soluble JPV-RBP and BeiV(HMC)-RBP for crystallisation, constructs were generated from full-length human codon-optimised RBP genes (Genbank NC_007454.1 and NC_007803.1, respectively). Similar to constructs generated for SosV-RBP, MosV-RBP and NarV-RBP, constructs containing minimal stalk region were preferred as they were likely more amenable to crystallisation (Savitsky et al. 2010), however the potentially flexible C-terminus was included as the structure of this was of significant interest. Appropriate construct boundaries were determined through amino acid sequence alignment to NDV-RBP (Crennell et al. 2000). The constructs were generated utilising primers specific to the selected N-terminal start site and the fixed C-terminal end site, and subsequently ligated into the pHLsec mammalian expression vector (Aricescu, Lu, and Jones 2006). For JPV-RBP, constructs Gln123–Asn709, Glu149–Asn709, Glu153–Asn709 and Lys165–Asn709 were generated (**Fig 5.3.1iA**), with only Gln123–Asn709 and Glu149–Asn709 exhibiting soluble expression upon small-scale expression screening in human embryonic kidney (HEK)293T cells (**Fig 5.3.1iB**). The BeiV(HMC)-RBP constructs Gln128–Glu734, Asn145–Glu734, Arg150–Glu734 and Asp158–Glu734 were generated (**Fig 5.3.1iC**), however, only soluble expression of Asn145–Glu734 was detected upon small-scale expression screening (**Fig 5.3.1iD**).

Following the discovery of a longer BeiV-RBP reading frame, encoded by BeiV found in rat populations in China (Woo et al. 2016b), a number of constructs were generated in order to attempt subsequent structural characterisation. Utilising a full-length human codon-optimised RBP gene (Genbank AOV81769.1) with 99.3% identity to BeiV(HMC)-RBP across amino acids Met1–Glu734, constructs Lys74–Pro1046, Arg85–Pro1046, Lys98–Pro1046, Gln113–Pro1046,

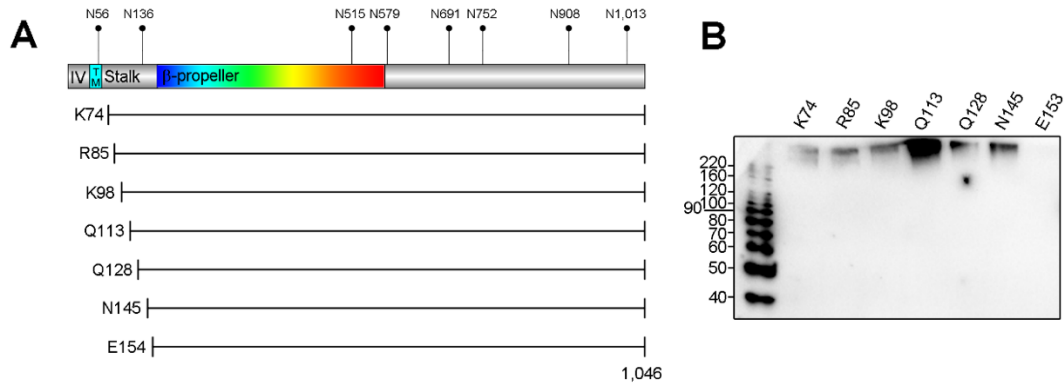


Figure 5.3.1ii: Construct design and small-scale expression test of BeiV(rat)-RBP. (a) Gene diagram displaying features of the BeiV(rat)-RBP, including the IV region, TM domain, stalk and six-bladed β -propeller receptor binding head (rainbow). Predicted N-linked glycosylation sites (NXT/S, where X $\neq$ P) are marked with pins and relevant asparagine residue numbered. The designed constructs are illustrated below the schematic. (b) Western blot of small scale expression test of constructs with an N-terminal SUMO tag, where an SDS-PAGE gel of HEK293T cell supernatants was performed under reducing conditions. Western blot analysis was performed alongside the Benchmark™ Protein Ladder (Thermo, left).

Gln128–Pro1046, Asn145–Pro1046, Glu154–Pro1046 were generated using PCR and cloned into the pHLsec vector (Aricescu, Lu, and Jones 2006) (**Fig 5.3.1iiA**). However, soluble expression of the constructs was not observed upon small-scale expression screening in HEK293T cells. Therefore, to aid in solubility constructs were sub-cloned into a pHLsec vector encoding an N-terminal octahistidine tag and small ubiquitin-like modifier (SUMO) tag (Chang et al. 2015). Expression of BeiV(rat)-RBP constructs conjugated to a SUMO tag was observable upon western blot analysis (**Fig 5.3.1iiB**). Given the success of the initial work with BeiV(HMC)-RBP construct Asn145–Glu734, construct BeiV(rat)-RBP Asn145–Pro1046 was selected for further structural work.

5.3.2 Large-scale expression and purification

The JPV and BeiV-RBP constructs were transiently expressed in HEK29T cells, in the presence of the α -mannosidase type I inhibitor, kifunensine (Chang et al. 2007). Samples were centrifuged and the supernatant filtered and buffer exchanged into 10 mM Tris pH 8, 150 mM NaCl buffer (TBS). Utilising nickel-affinity chromatography, RBP was purified from the supernatant using a HisTrap™ FF column (GE healthcare). For both JPV-RBP (Glu149–Asn709) and BeiV(HMC)-RBP

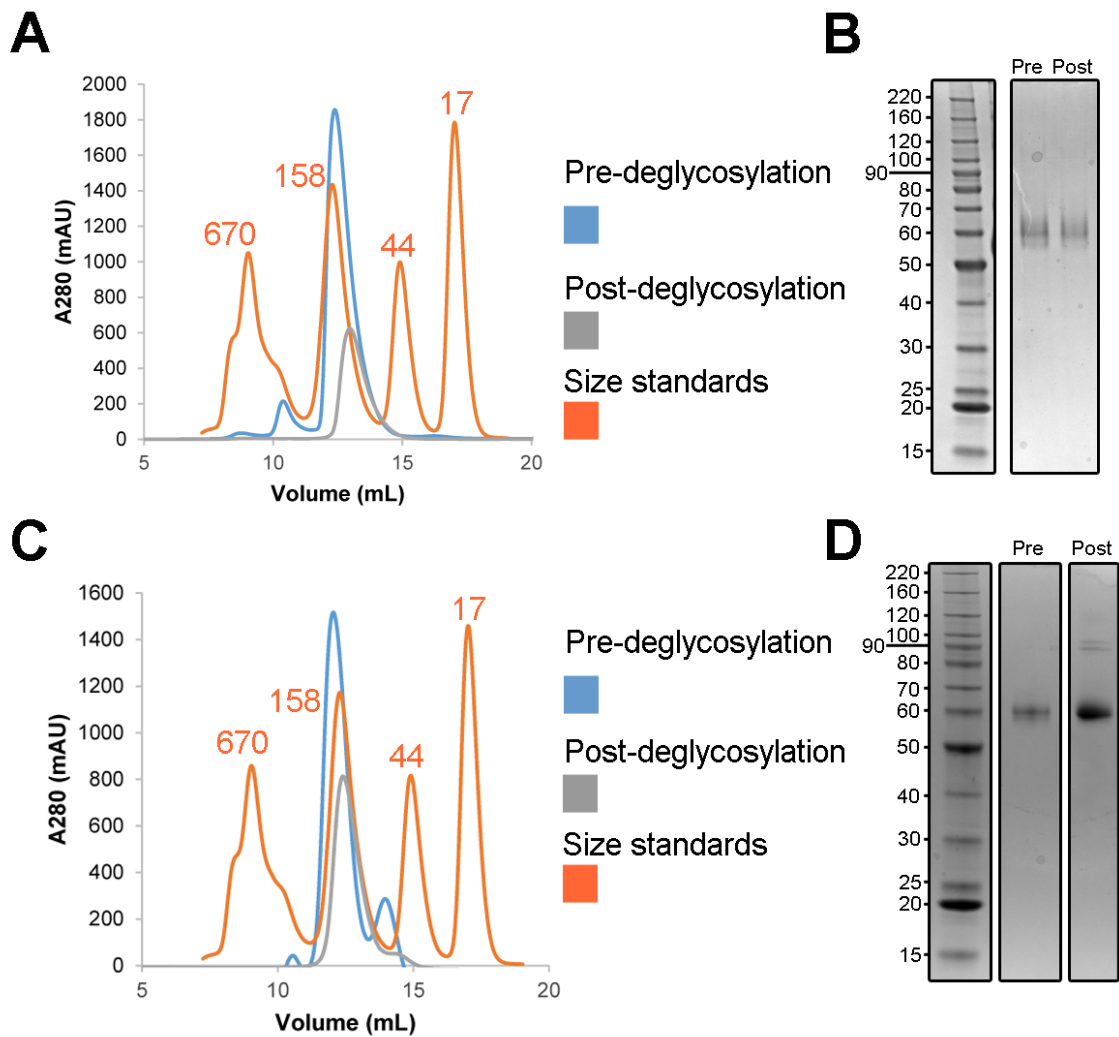


Figure 5.3.2i: Purification of JPV-RBP and BeiV(HMC)-RBP. (a) SEC of JPV-RBP(Glu149–Asn709) expressed transiently in the presence of kifunensine, pre- (blue) and post-deglycosylation (grey) with EndoF₁, aligned against gel filtration standards (BioRad, orange). (b) The SDS-PAGE analysis of JPV-RBP purification pre- and post-deglycosylation does not indicate a shift following treatment with EndoF₁. (c) SEC of BeiV(HMC)-RBP(Asn145–Glu734) expressed stably in the presence of kifunensine, pre- (blue) and post-deglycosylation (grey) with EndoF₁, aligned against gel filtration standards (BioRad, orange). (d) SDS-PAGE analysis of BeiV-RBP pre- and post-deglycosylation reveals no shift following treatment with EndoF₁. In both panel B and D, samples were run under reducing conditions, alongside Benchmark™ Protein Ladder (Thermo, left).

(Asn145–Glu734), nickel-purified protein was further purified using size exclusion chromatography (SEC) with a Superdex™ 200 Increase 10/300 GL column (GE Healthcare) (**Fig 5.3.2i**), prior to partial deglycosylation with the enzyme endoglycosidase F₁ (EndoF₁), yielding protein bearing single N-acetylglucosamine (GlcNAc) residues. EndoF₁ was removed from JPV-RBP(Glu149–Asn709) and BeiV(HMC)-RBP(Asn145-Glu-734) samples through a second gel filtration step.

In both cases, SEC and SDS-PAGE analysis showed that the samples were largely monodisperse following the initial nickel-affinity chromatography and gel filtration steps (**Fig 5.3.2i**). The bands observed upon SDS-PAGE analysis correspond to the approximate size of monomeric the JPV-RBP and BeiV(HMC)-RBP constructs, which are predicted to be 63 kilodalton (kDa) and 67 kDa, respectively. In addition, the observed peaks on the SEC traces are consistent with a dimeric organisation of both JPV-RBP and BeiV(HMC)-RBP. The expressed JPV-RBP and BeiV(HMC)-RBP constructs are predicted to encode three N-linked glycosylation sites (**Fig 5.3.1iA and C**). However, following treatment of JPV-RBP and BeiV(HMC)-RBP with EndoF₁, SDS-PAGE analysis revealed no reduction in size was observed (**Fig 5.3.2iB and D**), suggestive that either deglycosylation was not successful or that the protein was not glycosylated during post-translational modification in the HEK293T cells (Jones, Krag, and Betenbaugh 2005).

The BeiV(rat)-RBP(Asn145–Pro1046) containing supernatant underwent purification through nickel-affinity chromatography as described above, however, as protein yields were low, a SEC was not performed preceding deglycosylation with EndoF₁. In addition, due to the large size of the protein, gel filtration following overnight deglycosylation was performed on a Superose 6 Increase 10/300 column (GE healthcare), which has a broad fractionation range of 5 kDa to 5,000 kDa. The size exclusion trace revealed the elution of a large heterogenous sample between ~17 kDa and the upper limit of the column, indicating the presence of multiple species, aggregating protein and potential expression system-derived protein contamination (Molodenskiy et al. 2017; Johnson et al. 2018) (**Fig 5.3.2iiA**). These findings were similarly reflected upon SDS-PAGE and western blot analysis, which indicated that BeiV(wt)-RBP formed multiple oligomeric species or N-linked

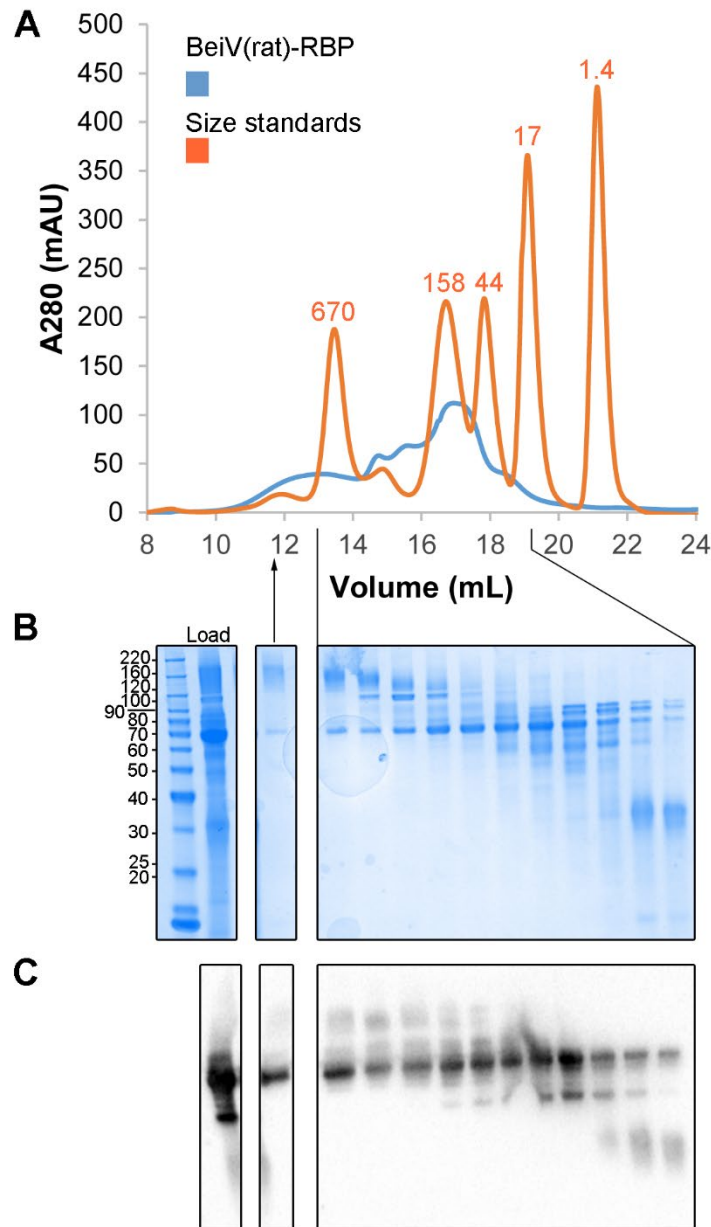


Figure 5.3.2ii: Purification of BeiV(rat)-RBP. (a) Size exclusion chromatogram of BeiV(rat)-RBP(Asn145–Pro1046) expressed transiently in the presence of kifunensine (blue), aligned against gel filtration standards (BioRad, orange). (b) The reducing SDS-PAGE analysis of BeiV(rat)-RBP purification reveals the purification of a heterogeneous sample, with unclear separation of species. (c) Anti-histidine tag western blotting detects the sample throughout multiple heterogeneous peaks, indicating that BeiV(rat)-RBP is likely presented as multiple different oligomeric species. In both panel B and C, samples were run under reducing conditions, alongside Benchmark™ Protein Ladder (Thermo, left).

and O-linked glycoforms (**Fig 5.3.2iiB** and **C**). Optimisation of the protein expression and purification protocol for this construct or screening of different constructs would likely be extensive, and given the probable issues structurally characterising the P/T/S rich BeiV(rat)-RBP (Chang et al. 2007), this work was deprioritised. Therefore, in the proceeding sections work involving BeiV-RBP will largely concern the BeiV-RBP isolated from the HMC line, unless otherwise stated.

5.3.3 Crystallisation and X-ray diffraction data collection

Utilising a pre-crystallisation test (PCT™, Hampton), the appropriate concentrations for crystallisation of JPV-RBP and BeiV-RBP were determined to be 3.5 mg/mL and 3.2 mg/mL, respectively. JPV-RBP crystals grew with a large needle morphology after 1 d in the following precipitant mix: 0.1 M carboxylic acid, 0.1 M tris/bicine pH 8.5, 6% sucrose, 0.2 M ammonium sulphate, 37.5% Morpheus® (Molecular dimensions) precipitant mix 4, consisting of 25% v/v 2-methyl-2,4-pentanediol (MPD), 25% w/v PEG 1000 (P1k) and 25% w/v polyethylene glycol 3350 (PEG3350). The crystal was immersed in a solution consisting of reservoir solution with an additional 10% MPD, prior to cryo-cooling. X-ray data collection was performed at a wavelength of 0.9282 Å on beamline I04-1, Diamond Light Source (DLS), Didcot UK. Reflections were indexed, integrated and scaled using the xia2 package (Winter 2010), with the data processing statistics presented in **Table 5.3.4i**.

BeiV-RBP crystals formed after 2 d in the following precipitant mix: 0.2 M potassium thiocyanate, 0.1 M sodium cacodylate pH 6.5, 8% w/v poly-γ-glutamic acid polymer (PGA), 6% 1,5-diaminopentane di-HCl. Crystals were harvested and cryo-cooled in a 20% glycerol and precipitant mix, before data were collected at a wavelength of 0.9686 Å on beamline I24, DLS. Subsequently, reflections from three different crystals were indexed, integrated and scaled using the xia2 package (Winter 2010), with data processing statistics present in **Table 5.3.4ii**.

Table 5.3.4i: Data collection and refinement statistics for JPV-RBP

JPV-RBP	
Data collection statistics	
Beamline	DLS I04-1
Spacegroup	$P 6_1$
Cell dimensions	
a, b, c (Å)	164.2, 164.2, 112.5
α, β, γ (°)	90.0, 90.0, 120.0
Resolution range (Å)	60.11–2.20 (2.24–2.20) ^a
Wavelength (Å)	0.9282
R_{merge}	0.18 (2.34)
R_{meas}	0.19 (2.46)
R_{pim}	0.06 (0.76)
I/σ I	9.6 (1.3)
CC_{1/2}	0.99 (0.48)
Completeness (%)	100 (100)
Multiplicity	10.2 (10.4)
Refinement statistics	
Resolution (Å)	46.40–2.20 (2.28–2.20)
No. reflections	86,384
R_{work} / R_{free}^b (%)	16.3/19.7
No. atoms	
Protein	6,915
Ligand/ion	0
Water	293
B-factors	
Protein	55.1
Ligand/ion	N/A
Water	53.4
R.m.s deviations^c	
Bond lengths (Å)	0.014
Bond angles (°)	1.28
Ramachandran analysis^d	
Favoured region (%)	95.3
Allowed region (%)	4.3

^aNumbers in parentheses refer to the outer resolution shell.

^bR_{free} is calculated as for R_{work}, but using only 5% of the data which were separated prior to refinement.

^cr.m.s deviations: root mean square deviation from ideal geometry.

^dDetermined using the Molprobit server (Chen et al. 2010).

Table 5.3.4ii: Data collection and refinement statistics for BeiV-RBP

BeiV-RBP	
Data collection statistics	
Beamline	DLS I24
Spacegroup	$C 2 2 2_1$
Cell dimensions	
a, b, c (Å)	88.0, 162.4, 226.1
α, β, γ (°)	90.0, 90.0, 90.0
Resolution range (Å)	73.18–3.50 (3.56–3.50) ^a
Wavelength (Å)	0.9686
R_{merge}	0.14 (0.95)
R_{meas}	0.15 (1.00)
R_{pim}	0.04 (0.29)
I/σ I	13.8 (2.5)
CC_{1/2}	1.00 (0.87)
Completeness (%)	99.8 (99.2)
Multiplicity	12.9 (11.5)
Refinement statistics	
Resolution (Å)	53.98–3.50 (3.63–3.50)
No. reflections	20,764
R_{work} / R_{free}^b (%)	24.7/27.5
No. atoms	
Protein	8,507
Ligand/ion	0
Water	1
B-factors	
Protein	118.4
Ligand/ion	N/A
Water	82.9
R.m.s deviations^c	
Bond lengths (Å)	0.003
Bond angles (°)	0.67
Ramachandran analysis^d	
Favoured region (%)	96.0
Allowed region (%)	3.7

^aNumbers in parentheses refer to the outer resolution shell.

^bR_{free} is calculated as for R_{work}, but using only 5% of the data which were separated prior to refinement.

^cr.m.s deviations: root mean square deviation from ideal geometry.

^dDetermined using the Molprobit server (Chen et al. 2010).

5.3.4 Structure determination

The JPV-RBP structure was solved by molecular replacement using the structure of human parainfluenza virus 3 RBP (PDB: 1V2I) (Lawrence et al. 2004) as a search model. The structure was subsequently built and refined using COOT (Emsley and Cowtan 2004) and PHENIX (Adams et al. 2010) software packages. The refinement statistics are presented in **Table 5.4.3i**.

The BeiV-RBP structure was initially solved in PHASER (McCoy et al. 2007) by molecular replacement with JPV-RBP. However, following building and refinement in COOT (Emsley and Cowtan 2004) and PHENIX (Adams et al. 2010), the C-terminus could not be resolved. Following discussions with Dr Kamel El-Omari, density modification was implemented to improve the initial phase estimates. Utilising the core β -propeller of BeiV-RBP as a search model so as to not introduce bias from the partially built C-terminal extension, molecular replacement was performed once again. This enabled application of the output density into the density modification programme, Parrot (Cowtan 2010) and 'Phase and Build' (Terwilliger et al. 2009; Adams et al. 2010), which also included automatic building of the structure. Care was taken throughout to maintain a consistent set of Rfree flags. Through iterative building and refinement in COOT (Emsley and Cowtan 2004) and PHENIX (Adams et al. 2010), the model of BeiV-RBP was built, including a considerable portion of the extended C-terminus (residue Lys575–Asn687). Data refinement statistics are presented in **Table 5.3.4ii**.

5.4 STRUCTURE ANALYSIS

5.4.1 Structure overview of the BeiV and JPV RBPs

Structure determination of JPV and BeiV RBPs confirmed that both proteins exhibited the typical six-bladed β -propeller fold characteristic of paramyxoviral RBP proteins (**Fig 5.4.1**), with each blade (β 1 to β 6) formed from four antiparallel β -strands. In both cases, two near-identical β -propeller head domains were organised as a putative dimer in the asymmetric unit. The interface between JPV-RBP monomers (root-mean-square deviation [RMSD] of 0.21 Å over 435 C α atoms) is mediated through

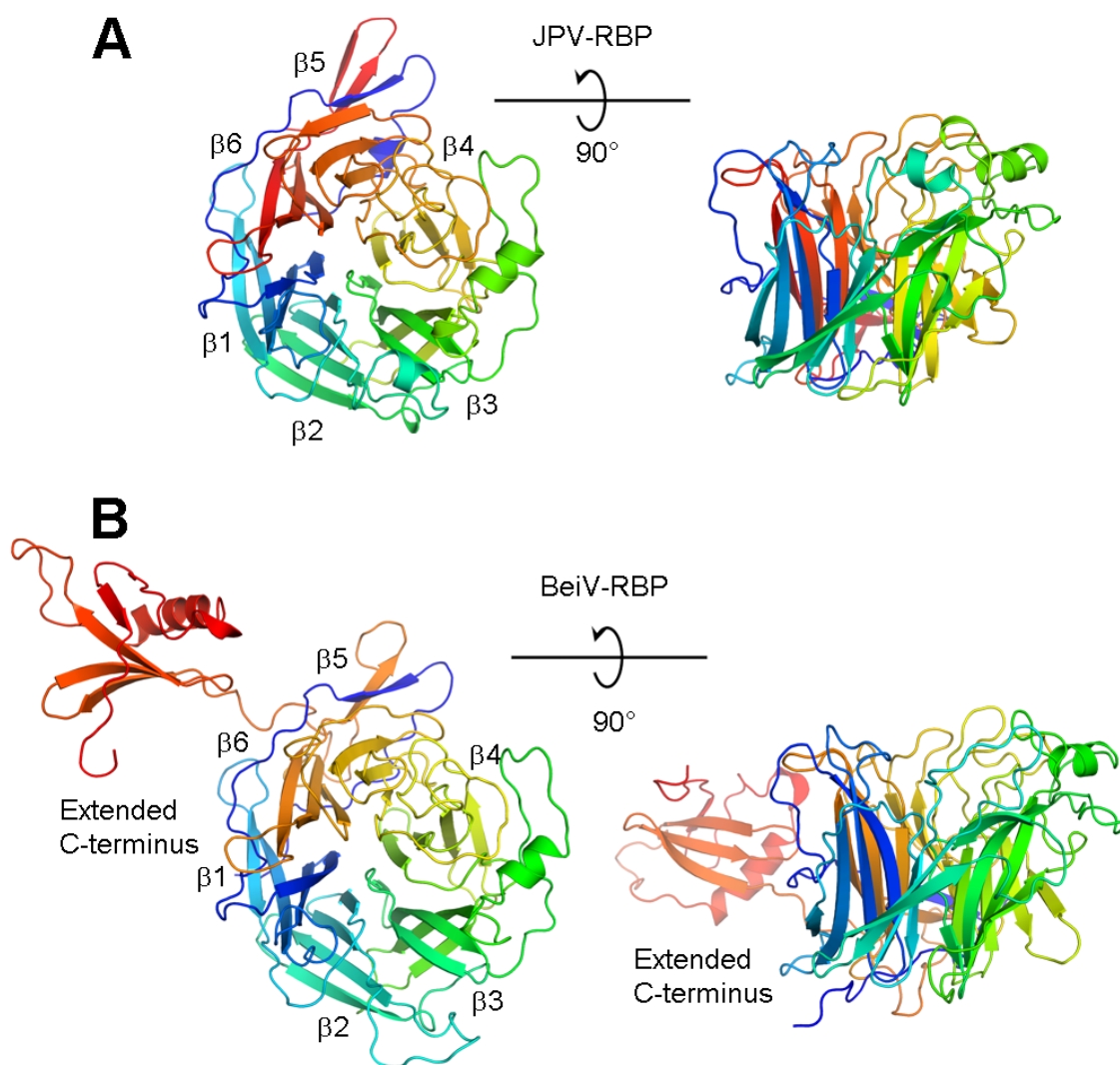


Figure 5.4.1: JPV-RBP and BeiV-RBP β -propeller monomers. (a) The six-bladed β -propeller monomer of JPV-RBP is shown in cartoon representation with blades labelled (coloured blue to red, from N- to C-terminus), viewing from both top down and the side. (b) The six-bladed β -propeller monomer of BeiV-RBP, with the C-terminal extension, is viewed from both top down and the side, and shown in cartoon representation with blades labelled (coloured blue to red, from N- to C-terminus).

an interaction between the first ($\beta 1$) and sixth ($\beta 6$) blades of the propeller. The JPV-RBP structure contained a short portion of the stalk region (Glu149–Met160), the β -propeller (Cys161–Cys573), and only a short section of the C-terminal extension (Gln574–Gly587), with the remaining terminus (Gly588–Asn709) disordered and directed towards the solvent channels in the crystal. The BeiV-RBP structure consisted of a short portion of the N-terminal stalk region (Pro146–Gln162), the β -propeller (Cys163–Cys575), and a large section of the C-terminal extension (Lys576–Asn687) which extends across the dimerically organised β -propeller domains, with the remaining C-terminus (Ile688–Glu734) disordered and directed towards the solvent channels of the crystal. Similarly, the dimeric interface between BeiV-RBP monomers (RMSD of 0.78 Å over 542 C α atoms), is mediated through an interaction driven by blades $\beta 1$ and $\beta 6$. However, the BeiV-RBP C-terminal extension also significantly contributes to the buried surface area, as discussed further in **section 5.4.2.1**.

Concerning the β -propeller only, JPV-RBP and BeiV-RBP display a pattern of disulphide bonding consistent with the closest structurally characterised relative, the PIV3-RBP structure (1V2I) (Lawrence et al. 2004), a hemagglutinin-neuraminidase (HN)-type paramyxoviral RBP. However, unlike other paramyxoviral RBPs, which typically demonstrate extensive N-linked glycosylation across the β -propeller (Bowden, Crispin, et al. 2008; Sawatsky and von Messling 2010; Hu et al. 1994; Segawa et al. 2003; Ng, Hiebert, and Lamb 1990; McGinnes and Morrison 1995; Panda et al. 2004; Chu et al. 2013), both JPV-RBP and BeiV-RBP display few predicted N-linked glycosylation sequons (NXS/T, where X $\neq$ P) across the propeller (**Fig 5.3.1**). Within the construct boundaries, JPV-RBP and BeiV-RBP are predicted to have three N-linked glycosylation sites, all contained within the C-terminal extensions (Asn619^{JPV-RBP}, Asn636^{JPV-RBP}, Asn677^{JPV-RBP}, Asn579^{BeiV-RBP}, Asn691^{BeiV-RBP}) with the exception of Asn515^{BeiV-RBP}, which is within the β -propeller. No electron density corresponding to N-linked glycosylation at sites Asn619^{JPV-RBP}, Asn636^{JPV-RBP}, Asn677^{JPV-RBP}, Asn691^{BeiV-RBP}, across the C-terminal extension was observed. Additionally, density consistent with N-linked glycosylation at the surface presented, Asn579^{BeiV-RBP}

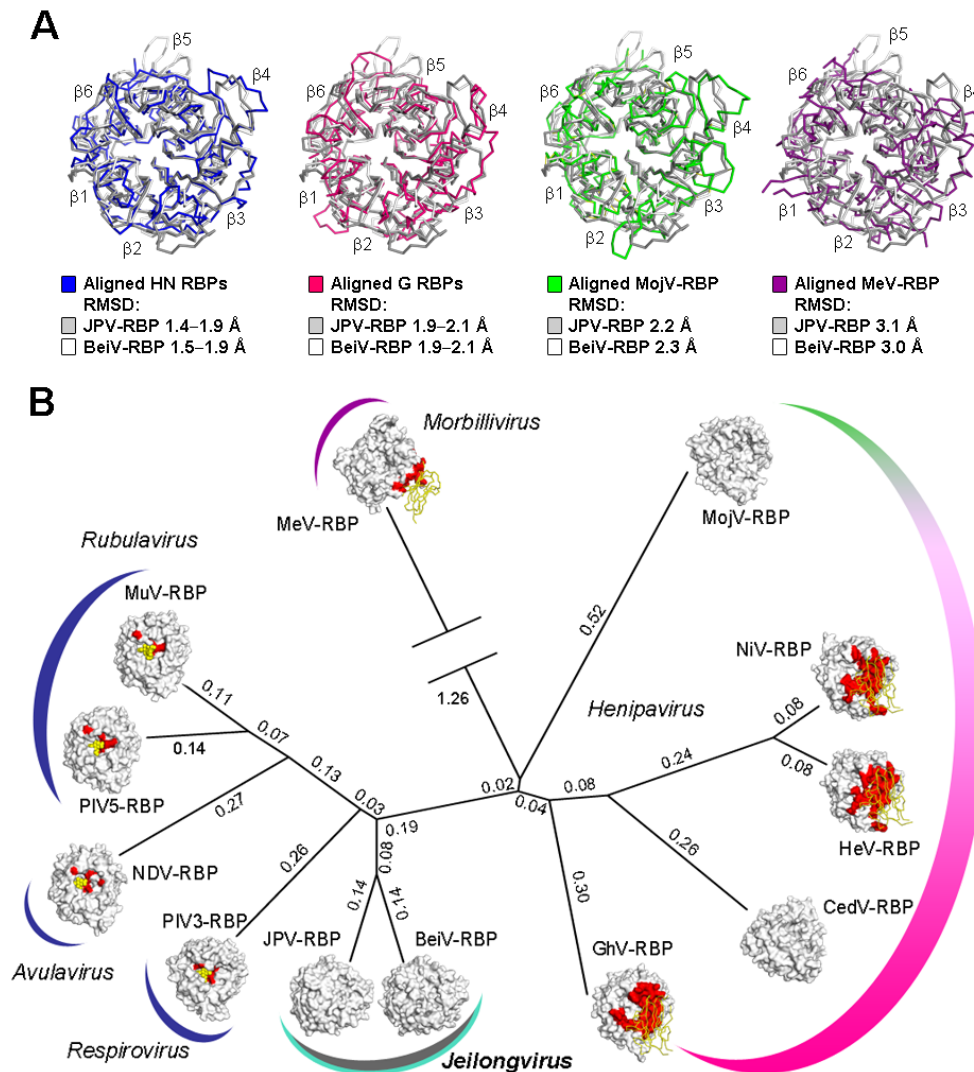


Figure 5.4.2.1i: The structural relationship of JPV and BeiV RBP β -propellers with other paramyxovirus RBPs. (a) Structure overlays of JPV-RBP (grey) and BeiV-RBP (white) with other paramyxoviral RBP structures: from left to right, NDV (blue, 1E8V) (Crennell et al. 2000), NiV, Nipah virus (pink, 2VWD) (Bowden, Crispin, et al. 2008), MojV, Mojiang virus (green, 5NOP) (Rissanen et al. 2017), MeV, measles virus (purple, 2ZB5) (Hashiguchi et al. 2007). α trace rendered and root-mean-square deviation (RMSD) annotated. (b) Structure-based phylogenetic analysis of paramyxoviral RBP monomers: JPV; BeiV; PIV3, parainfluenzavirus 3 (1V2I) (Lawrence et al. 2004); NDV (1E8V); PIV5, parainfluenza virus 5 (4JF7) (Welch et al. 2013); MuV (5B2C) (Kubota et al. 2016); MojV (5NOP); GhV, Ghana virus (4UF7) (Lee et al. 2015); NiV (2VWD); HeV, Hendra virus (2X9M) (Bowden et al. 2010); CedV, Cedar virus (PDB pending) (Pryce et al. 2019); MeV (2ZB5). The structural Homology Program (SHP) (Stuart et al. 1979) was utilised to calculate evolutionary distance matrices by pairwise superposition of RBP structures. The resultant matrices were used to plot an unrooted tree in PHYLIP (Felsenstein 1989). RBP surfaces are represented with cognate receptor binding sites coloured (red), and receptors presented either as yellow ribbon (protein) or spheres (carbohydrate). Calculated evolutionary distances are indicated beside the branches.

and Asn515^{BeiV-RBP} sites was not observed (**Fig 5.4.1**). The lack of observable occupancy at N-linked glycosylation sequons could be due to intrinsic flexibility of the protein, especially across C-terminal extension. However, consistent with SDS-PAGE analysis (**Fig 5.3.2iB and D**), the lack of observed N-linked glycosylation could be due to low or absent occupancy at these predicted sequons (Jones, Krag, and Betenbaugh 2005).

5.4.2 Comparison of BeiV-RBP and JPV-RBP to different paramyxoviral RBPs

5.4.2.1 Macro-comparisons

When comparing the six-bladed β -propeller head domains of the various paramyxoviral RBPs, pairwise superposition reveals that JPV-RBP and BeiV-RBP align very closely, with a calculated RMSD of 1.04 Å over 410 C α atoms. Beyond the genus *Jeilongvirus*, JPV-RBP and BeiV-RBP most closely align with the HN-type RBPs (1.39–1.90 Å and 1.51–1.91 Å RMSD, respectively) (**Fig 5.4.2.1iA**). This is consistent with the observation that the JPV-RBP and BeiV-RBP β -propellers display a pattern of disulphide bonding comparable to the HN-type RBPs and is supported by structure-based phylogenetic analysis (Stuart et al. 1979), which reveals the structural relatedness of the jeilongviral β -propellers (**Fig 5.4.2.1iB**). JPV-RBP and BeiV-RBP display a more distant relationship to the henipaviral RBPs (1.92–2.16 Å and 1.94–2.28 Å RMSD, respectively) and the measles virus (MeV) RBP (3.14 Å and 2.97 Å RMSD, respectively) (**Fig 5.4.2.1iA**).

The angle of association between the six-bladed β -propellers of the dimeric JPV-RBP and BeiV-RBP, is calculated to be 64.7° and 72.0°, respectively (**Fig 5.4.2.1ii**). This is reminiscent of the oligomeric arrangement of the HN-type RBPs, where the angle of association ranges from 52.6–63.5° (Crennell et al. 2000; Lawrence et al. 2004; Kubota et al. 2016; Yuan et al. 2008; Yuan et al. 2005). However, whilst the oligomeric arrangement of the JPV-RBP and BeiV-RBPs relies on the association of the β 1 and β 6 blades, the interface area is much smaller, at 1,225 Å² and 1,311 Å² (calculated by Proteins, Interfaces, Surfaces and Assemblies (PISA) server), respectively, as opposed to the average interface area exhibited by HN-type RBP dimers of 1,790 Å². This may be partly due to the contribution from the extended C-terminus. Indeed, in the BeiV-RBP structure, which includes

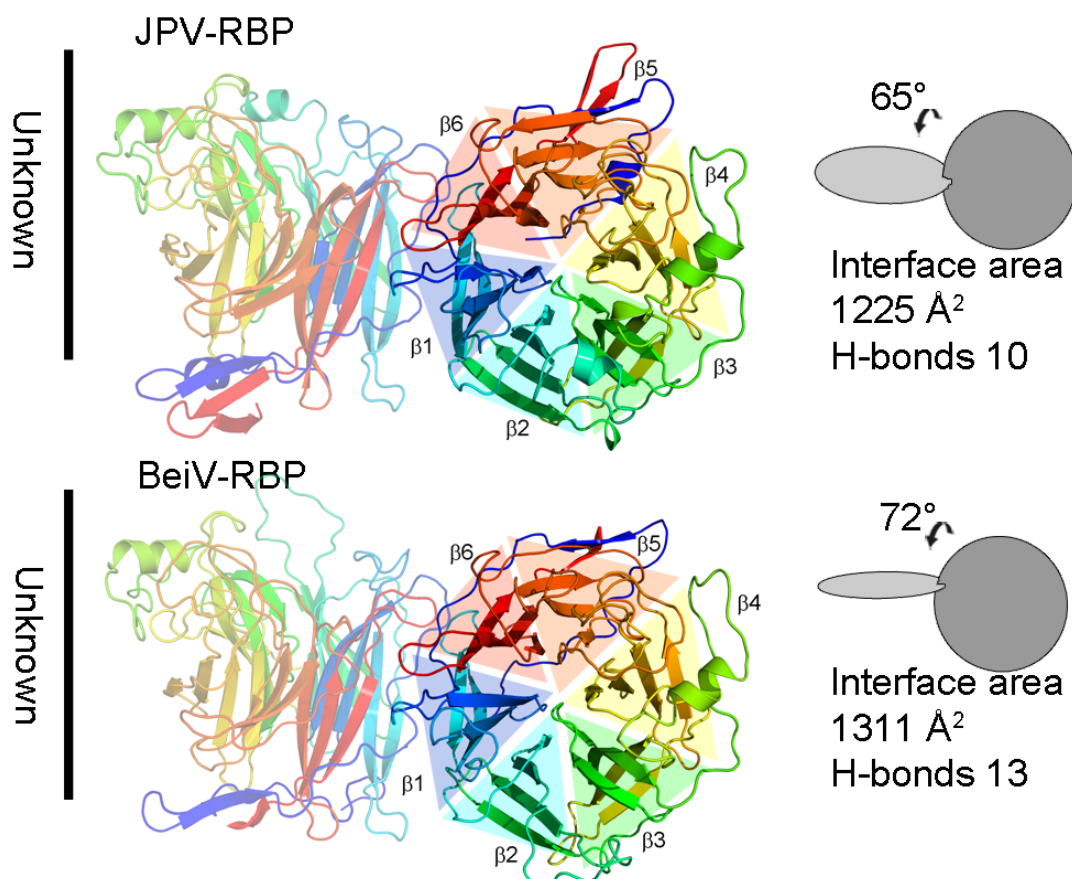


Figure 5.4.2.1ii: Dimeric assembly of MosV-RBP and NarV-RBP. The crystallographic MosV-RBP and NarV-RBP dimer shown in cartoon representation with monomers coloured rainbow from the N-terminus to C-terminus (blue to red, respectively). Calculated angle of association, interface area and number of hydrogen bonds is displayed to the right-hand side of the relevant dimer.

the short portion of N-terminal stalk (Pro146–Gln162), the six-bladed β -propeller (Cys163–Cys575), and the extended C-terminus (Lys576–Asn687), the total overall dimer interface area is 4,248 Å². This extensive homodimerisation interface is, in part, due to the extension of the structurally observed BeiV-RBP C-terminus from the underside to the top of the β -propeller to form a ‘hat-like’ domain on the adjacent protomer of the putative dimer, which is stabilised by 10 hydrogen bonds, and 4 salt bridges. The C-terminal extension is reminiscent of the dimeric organisation of the NDV-RBP from the Ulster strain (Yuan et al. 2012), where the C-terminus is stabilised by an interchain disulphide (**Fig 5.2iC**). Both the NDV Ulster strain RBP and BeiV-RBP

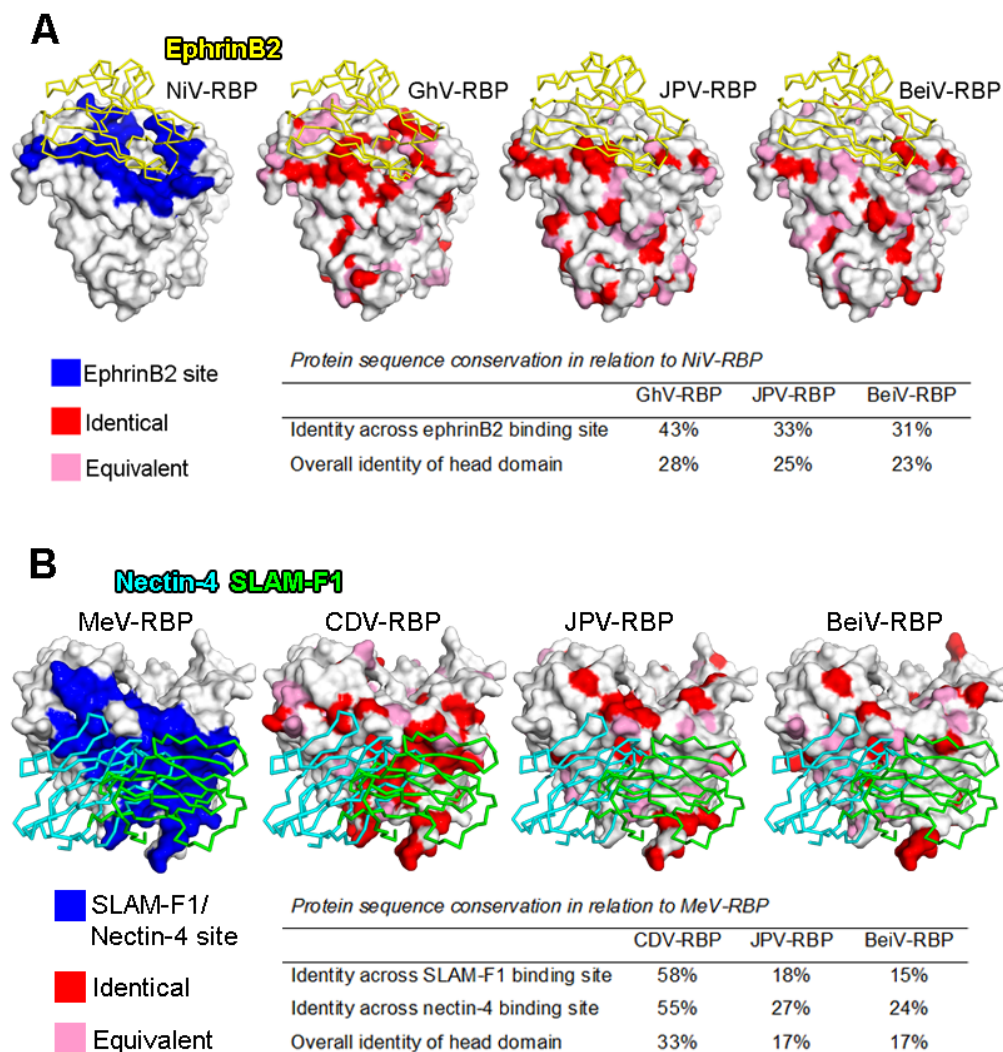


Figure 5.4.2.2i: Mapping of JPV-RBP and BeiV-RBP sequence conservation onto henipaviral and morbilliviral RBPs. (a) NiV-RBP (PDB ID 2VSM) (Bowden, Aricescu, et al. 2008; Zaitsev et al. 2004) is shown as a surface with ephrinB2 shown as a yellow ribbon and the cognate receptor-binding footprint coloured blue (far left). Sequence conservation of GhV-RBP (second left), JPV-RBP (second right), and BeiV-RBP (far right), with NiV-RBP mapped onto the surface of NiV-RBP. Consistent with the known shared usage of ephrinB2, GhV-RBP and NiV-RBP share an elevated level of sequence conservation at the receptor binding site. This feature of elevated sequence conservation was not observed using the same approach with NiV-RBP and the narmoviral RBPs, suggestive that JPV-RBP and BeiV-RBP do not utilise ephrin receptors. Residues coloured red are identical, those that are pink are equivalent residues. (b) MeV-RBP (PDB ID 3ALZ) (Hashiguchi et al. 2011; Woo et al. 2016b) is shown in surface representation with SLAM-F1 (green) and nectin-4 (cyan) shown in ribbon representation and the cognate joint receptor-binding footprint coloured blue (far left). Sequence conservation of CDV-RBP (second left), JPV-RBP (second right), and BeiV-RBP (far right), with MeV-RBP is mapped onto the surface of MeV-RBP. Residues coloured red are identical, those coloured pink are equivalent. Consistent with the shared usage of SLAMF1/nectin-4, MeV-RBP and CDV-RBP share an elevated level of sequence conservation at the receptor binding sites. This feature of elevated sequence conservation was not observed using the same approach with MeV-RBP and the jeilongviral RBPs, suggestive that JPV-RBP and BeiV-RBP do not utilise SLAMF1/nectin-4 receptors.

serve as examples of the different means by which an extended C-terminus may contribute to an unusually large and possibly more stable RBP homodimer interface.

5.4.2.1 *Comparisons to characterised receptor binding sites*

Mapping of jeilongviral RBP sequence conservation with NiV-RBP (2VSM) (Bowden, Aricescu, et al. 2008) and MeV-RBP (3ALZ) (Hashiguchi et al. 2011) onto their respective structures, reveals that JPV-RBP and BeiV-RBP do not exhibit a particularly elevated level of identity across the regions associated with henipaviral (Bowden, Aricescu, et al. 2008; Xu et al. 2008) and morbilliviral receptor binding (Zhang et al. 2013; Hashiguchi et al. 2011) (**Fig 5.4.2.2i**). Indeed, JPV-RBP and BeiV-RBP only demonstrated an overall identity with NiV-RBP of 25% and 23%, respectively, a value raised to only 33% and 31% across the potential receptor binding region (**Fig 5.4.2.2iA**). Furthermore, evaluation of the region corresponding to the henipaviral receptor binding site through superposition of ephrinB2 (2VSM) (Bowden, Aricescu, et al. 2008) and ephrinB3 (3D12) (Xu et al. 2008), reveals that the jeilongviral RBPs lack a pocket required to accommodate F120^{ephrinB2}/Y120^{ephrinB3}. The putative SLAM-F1 and nectin-4 binding sites are similarly not conserved, with JPV-RBP only demonstrating 18% and 27% identity across the receptor binding sites, respectively, and BeiV-RBP only demonstrating 15% and 24% identity, respectively (**Fig Fig 5.4.2.2iB**). Consistent with this, with the jeilongviral RBPs lack the β 4- β 5 groove, required by morbilliviruses to facilitate receptor interactions (Zhang et al. 2013; Hashiguchi et al. 2011; Santiago et al. 2010).

Importantly, the jeilongviral RBPs display many of the features required to interact with sialic acid, including the six of the seven sialidase residues, Arg₁, Glu₄, Arg₄, Arg₅, Tyr₆ and Glu₆ (**Fig 5.4.2.2iiA**). However, both JPV-RBP and BeiV-RBP lack Asp₁, with this residue being substituted with Ser and Glu, respectively. In addition, JPV-RBP and BeiV-RBP exhibit a somewhat different hexapeptide motif, presenting an Arg at the 3rd position as opposed to a Lys (**Fig 5.4.2.2iiA**). However, with the exception of the Asp to Ser substitution in JPV-RBP, it seems unlikely that these relatively conserved substitutions will impact the overall structure and chemical composition of the putative sialic acid binding pocket. Similarly, both JPV-RBP and BeiV-RBP present a putative

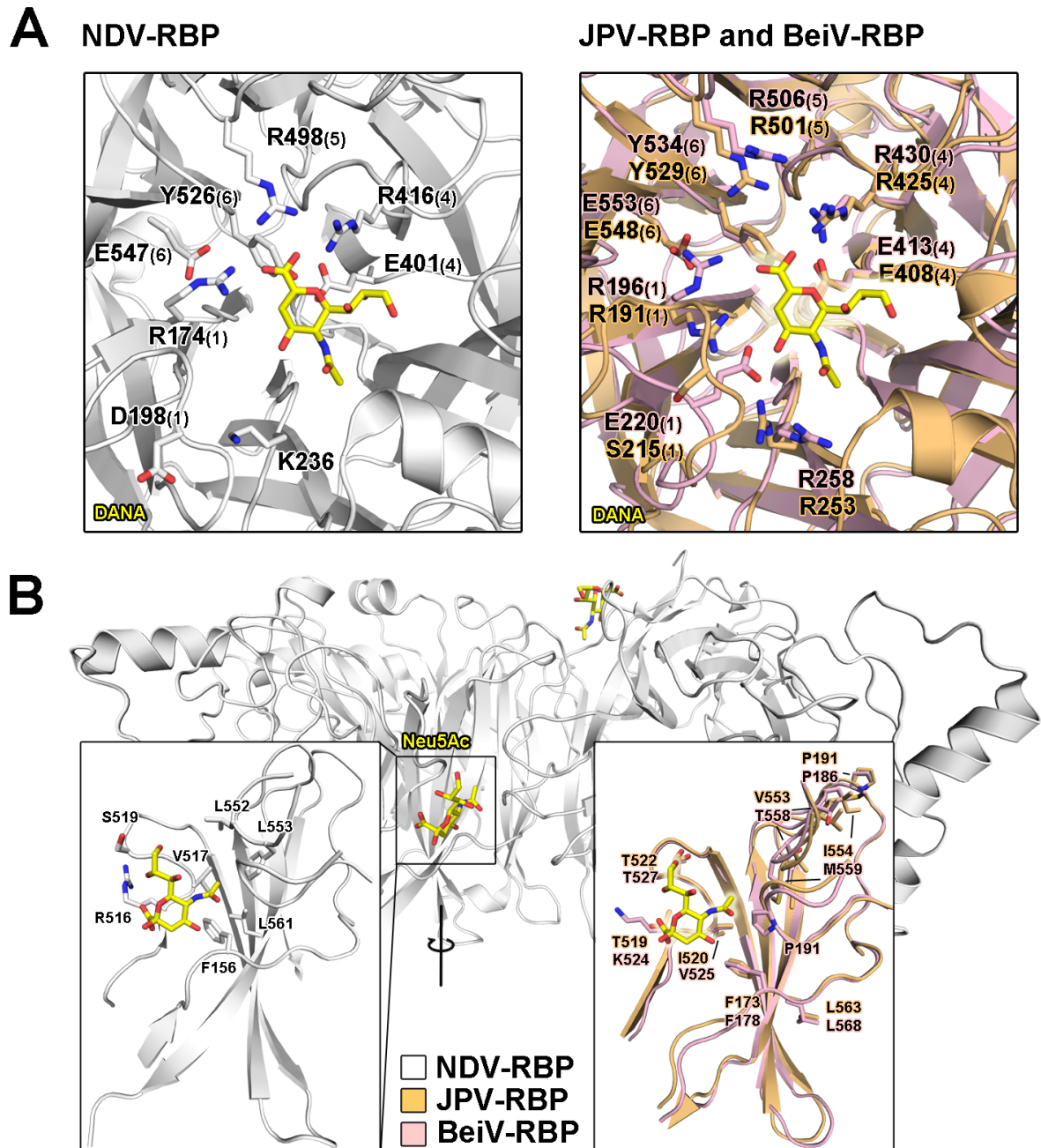


Figure 5.4.2.2ii: JPV-RBP and BeiV-RBP present putative primary and secondary binding sites that are conserved with known modes of sialic acid recognition (a) A close-up of the NDV-RBP β -propeller (Crennell et al. 2000) (left) with the primary sialic acid binding site with N-acetyl-2,3-dehydro-2-deoxyneuraminic acid (DANA) shown as yellow sticks (1E8V). The equivalent position of JPV-RBP (orange) and BeiV-RBP (pink) is shown on the right hand side with the hypothetical position of DANA. β -propellers are rendered in cartoon representation, and important residues are annotated and shown using stick representation. **(b)** The second sialic acid binding site of the NDV-RBP (left) with N-acetylneuraminic acid (Neu5Ac) (Zaitsev et al. 2004). The analogous site on JPV-RBP and BeiV-RBP is shown on the right hand side with the hypothetical position of DANA. The blades are shown using cartoon representation, Neu5Ac as yellow sticks, and important residues annotated and shown as sticks.

second sialic acid binding pocket lined with residues either identical or equivalent those identified in NDV-RBP structures (Zaitsev et al. 2004; Yuan et al. 2012; Yuan et al. 2011) (**Fig 5.4.2.2iiB**). Such findings are suggestive that both JPV-RBP and BeiV-RBP may also have the capability to bind sialic acid via this secondary site. Given this, sialic acid interactions would be expected, however previous studies indicate that JPV-RBP lacks hemagglutinin and hemadsorption abilities (Jun, Karabatsos, and Johnson 1977). By analogy to the extended C-terminus of the RBP from the NDV Ulster strain (Yuan et al. 2012), it may be that the jeilongviral RBP may encode C-terminal segment with auto-inhibitory potential.

5.4.3 The extended C-terminus

Unlike the C-terminal extension of the NDV Ulster strain RBP, which wraps round the side of each monomer (Yuan et al. 2012), the structurally observed BeiV-RBP C-terminus extends from the underside of the six-bladed β -propeller head domain to form a ‘hat-like’ domain on the adjacent protomer of the putative dimer (**Fig 5.4.3A and B**). Firstly, residues Leu573–Gly577 of C-terminus form a β -ribbon with residues Met168–Leu172 of the N-terminus (**Fig 5.4.3C**), stabilised by a disulphide (Cys163–Cys575), before a long flexible linker composed of residues Lys578–Glu604 traverses across the dimer interface (**Fig 5.4.3D**). This flexible linker is stabilised by extensive polar interactions with the residues at the pinnacles of loops β 6L12 and β 1L34 and the N-terminus, which is likewise directed beneath the β -propeller, in a mode similar to that observed in tetrameric HN-type RBP structures (Yuan et al. 2005; Yuan et al. 2008; Welch et al. 2013; Yuan et al. 2011). The long linker is directed across the dimer, where a disulphide conserved throughout the majority of the jeilongviral RBPs, Cys605–Cys648, stabilises the ‘hat-like’ domain on the proximal monomer (**Fig 5.4.3B**). The ‘hat-like’ domain is composed of an α -helix, a β -sheet, consisting of four antiparallel β -strands, and a long anchoring arm which buries into the putative sialic acid binding cavity (**Fig 5.4.3E**), before the chain becomes disordered. Using the DALI server (Holm 2019), a number of proteins found in prokaryotes and eukaryotes were found to have a similar fold (Z-score >2) (Holm et al. 2008) to the BeiV-RBP ‘hat-like’ domain, including chaperones, RNA binding proteins,

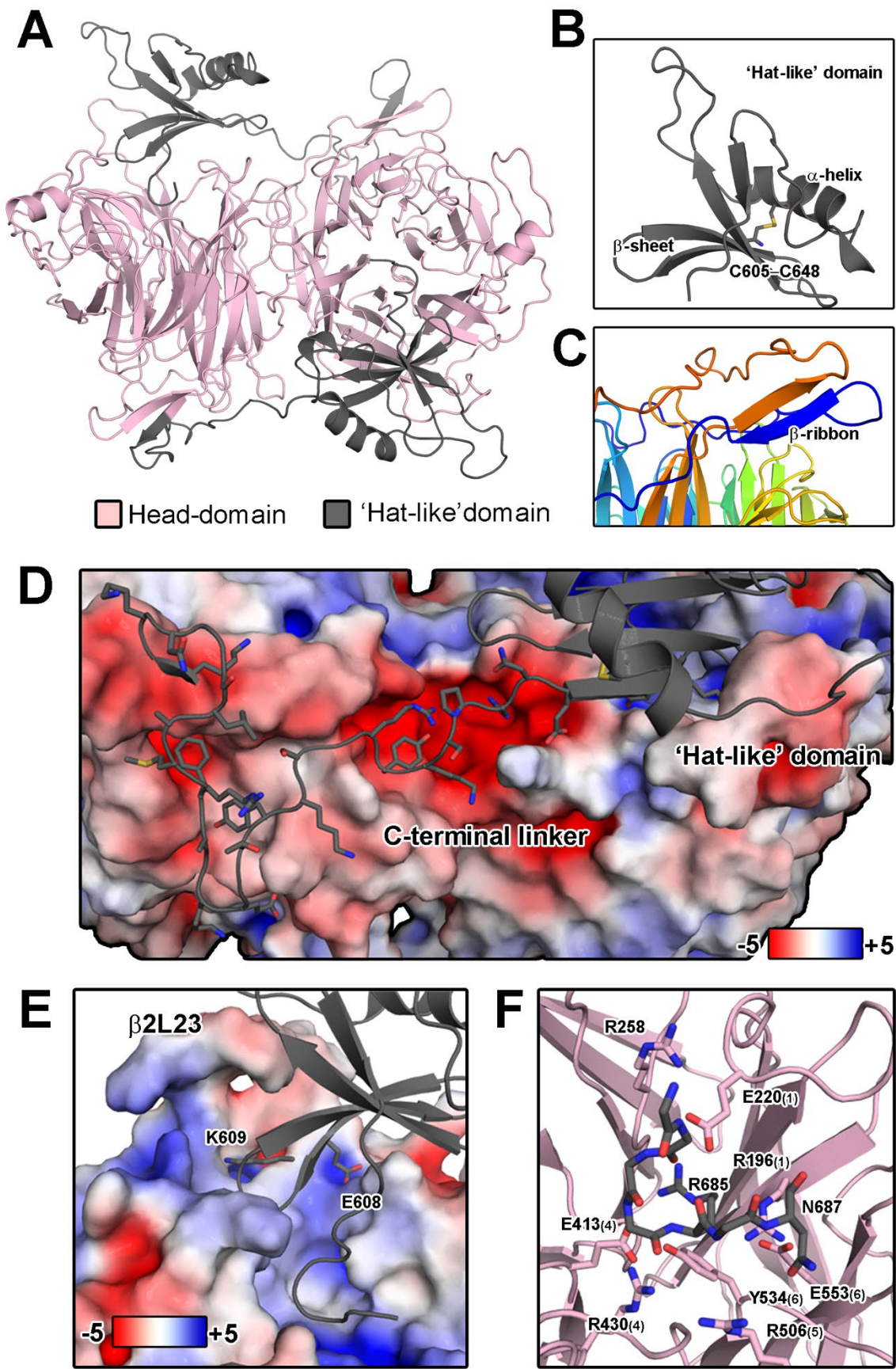


Figure 5.4.3: The structural organisation of the BeiV-RBP C-terminal extension. (a) The BeiV-RBP dimer shown in cartoon representation with the head-domain coloured pink and the ‘hat-like’ domain coloured grey. (b) The ‘hat-like’ domain in isolation, rendered as a cartoon, with the β -sheet, α -helix and disulphide bond annotated. (c) The β -ribbon that forms between the N-terminus of BeiV-RBP and the beginning of the C-terminal extension, shown in cartoon representation and coloured rainbow (blue to red, N- to C- terminus, respectively). (d) The dimeric head-domain of BeiV-RBP rendered as a surface, with the electrostatic charges calculated using the Adaptive Poisson-Boltzmann Solver (Jurrus et al. 2018), represented using surface by colouring residues from red to blue (-5 kT e-1 to +5 kT e-1). The C-terminal extension across the dimeric interface is represented using cartoon, with residue side chains shown, demonstrating the polar nature of the interaction of the BeiV-RBP head and ‘hat’ domains. (e) The BeiV-RBP head-domain rendered as a surface and coloured red to blue according to electrostatic charge, with the ‘hat-like’ domain rendered as a cartoon. The ‘hat-like’ domain displaces the loop β 2L23 through the anchoring of residues Glu608 and Lys609, enabling the protrusion of the C-terminal ‘hook-like’ end into the putative sialic acid binding site. (f) Interaction of the C-terminal ‘hook-like’ extension (grey), rendered as a C α main chain, with the putative sialic acid binding site (pink), rendered in cartoon representation. The conserved sialic acid binding residues interact extensively with the main chain of the projection and residues Arg685 and Asn687.

serine/threonine protein kinases, calcium/calmodulin-dependent protein kinases (Weber et al. 2018; Zhao et al. 2008; Ye et al. 2018; Tanwar et al. 2013; Chao et al. 2011). However, the ‘hat-like’ domain only matches small regions of larger proteins and complexes, and beyond the context of a greater assembly is unlikely to perform the functions known for proteins that present a similar structure. Nonetheless, it might be interesting to explore the potential of the BeiV-RBP ‘hat-like’ domain utilising these functionalities.

The ‘hat-like’ domain of BeiV-RBP, displaces loop β 2L23 of the propeller, which in turn envelopes the domain, likely through polar interactions between Lys609 of the ‘hat-like’ domain and the side chains and main chain carbonyls of residues Ser286 and Thr289. The pushing of the ‘hat-like’ domain against the strands and loops of blade β 2 is further cemented by the formation of a positively charged pocket by the side chain and main chain amide of residue Asn256, of the variant hexapeptide motif, and the main chain amide of Thr277, enabling the anchoring of Glu608 (**Fig 5.4.3E**). This displacement of loop β 2L23 widens the putative sialic acid interacting pocket on the surface of BeiV-RBP, enabling projection of the C-terminal hook of the ‘hat-like’ domain into this cavity. Interactions between the main chain carbonyl groups and the positively charged surfaces

facilitate this insertion, assisting Arg685 to interact with residues, Glu220₍₁₎, Glu413₍₄₎, and Tyr534₍₆₎, a polar interaction, which may disrupt potential sialic acid associated activity. Furthermore, the Arg residues of the triarginyl motif, Arg196₍₁₎, Arg430₍₄₎, and Arg506₍₅₎, interact with the various mainchain carbonyls of the ‘hook-like’ terminus (**Fig 5.4.3F**). Additionally, in contrast to the C-terminal extension observed in the structure of the RBP from NDV Ulster strain, the BeiV-RBP C-terminal extension does not appear to further occlude the second sialic acid binding site.

Further density corresponding to the remaining C-terminus of BeiV-RBP (Ile688–Glu734) is observable sandwiched between the ‘hat-like’ domain and the crystallographic symmetry mate, and at the pinnacle of loop β 5L34. However, due to disorder and the distance from the continuous chain, residues could not be unambiguously built into this density. Nonetheless, this finding indicates that the remaining C-terminus of BeiV-RBP was not cleaved. Similarly, in the case of JPV-RBP, despite the C-terminal extension likely being disordered, a short portion of density consistent with a C α backbone, is bound within the putative sialic acid site (**Fig 5.4.3.1A**). Binding appears to engage the residues putatively involved in mediating sialic acid interactions, Arg191₍₁₎, Glu408₍₄₎, Arg425₍₄₎, Arg501₍₅₎, Tyr529₍₆₎, Glu548₍₆₎, and the further positively charged residues Arg557, and Arg253 associated with the variant hexapeptide motif. However, loop β 2L23 of JPV-RBP is not displaced as it is in the BeiV-RBP structure, instead it forms a well-ordered α -helix proximal to the putative sialic acid binding site, consistent with the typical orientation observed for paramyxoviral RBP structures (Bowden, Aricescu, et al. 2008; Crennell et al. 2000; Lawrence et al. 2004; Kubota et al. 2016; Yuan et al. 2005; Pryce et al. 2019; Rissanen et al. 2017; Lee et al. 2015; Hashiguchi et al. 2007). Superposition of the ‘hat-like’ domain from BeiV-RBP onto JPV-RBP, reveals that whilst loop β 2L23 would cause a steric obstruction to the formation of the ordered β -sheet, it would not obstruct entry of the ‘hook-like’ terminus.

The observed differences between the JPV-RBP and BeiV-RBP structures, could be the product of many reasons. For example, it is possible that the C-terminal structure observed for BeiV-

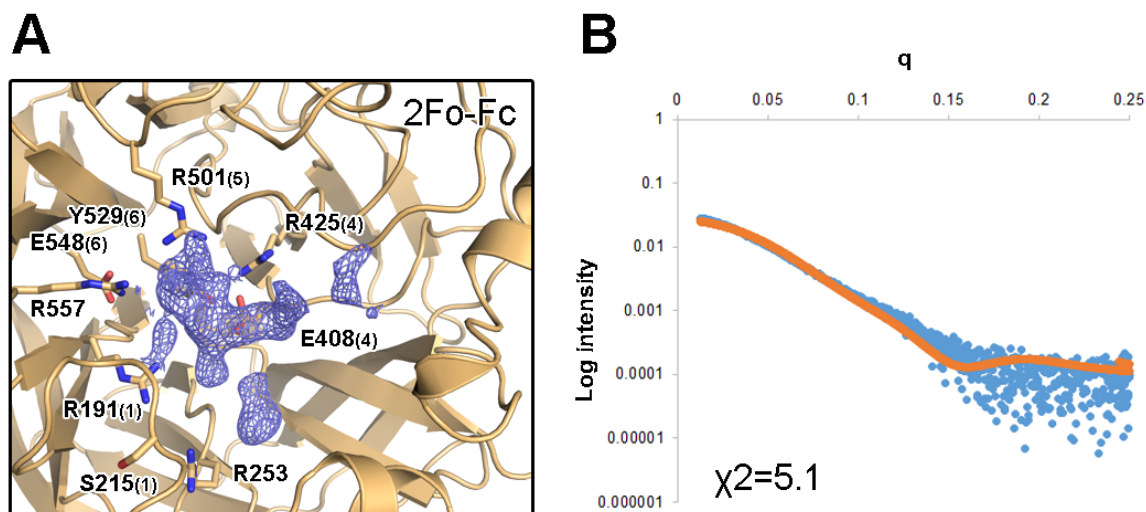


Figure 5.4.3.1: Density within the putative sialic acid binding site and SAXS analysis support the retention of the C-terminal segment by JPV-RBP. (a) Unidentified electron density (2Fo–Fc) within the putative JPV-RBP sialic acid binding site rendered at 2.0 sigma. JPV-RBP is rendered in cartoon representation with side chains of important residues shown as sticks and annotated. (b) Experimental JPV-RBP SAXS data (blue) compared to theoretical data calculated from the JPV-RBP structure (orange), plotted log intensity ($I(q)$) against the Porod invariant (q) (Schneidman-Duhovny, Hammel, and Sali 2010). The χ^2 value indicates the goodness of fit between experimental and theoretical data. The poor χ^2 value is suggestive that the JPV-RBP model is not an appropriate representation of the protein in solution, supporting the possibility that the C-terminal extension is retained by JPV-RBP.

RBP, and the displacement of loop β 2L23, are artificially introduced during crystallisation, or that the C-terminal domain of JPV-RBP simply forms a different structure, which does not displace β 2L23. Alternatively, it may be that JPV-RBP does not retain the C-terminal extension, and the density observed within the putative sialic acid binding pocket is an unknown ligand, potentially a carbohydrate product from HEK293T cell culture.

5.4.3.1 Small-angle X-ray scattering analysis of JPV-RBP

To further characterise the oligomeric organisation of JPV-RBP bearing the C-terminal extension (residues Glu149–Asn709), the solution state characteristics of JPV-RBP were assessed by tandem high performance liquid chromatography (HPLC) and small-angle X-ray scattering (SAXS) analysis, at beamline B21, DLS, Didcot, UK. Using the software package ‘ScÅtter’, the radius of gyration (R_g) using Guinier analysis, the volume-of-correlation (V_c) and the distance pair ($P(r)$) distribution

were determined, enabling calculation of molecular mass (Rambo and Tainer 2013b, 2013a; Putnam 2016). The approximate mass calculated for a deglycosylated homodimer of JPV-RBP bearing the C-terminal extension was 120 kDa, which with consideration of a 10% error (Rambo and Tainer 2013a), was consistent with the 126 kDa predicted. Comparison of the experimental scattering data to a theoretical scattering profile calculated from the dimeric JPV-RBP structure, using the 'Fast X-ray Scattering' (FoXS) server (Schneidman-Duhovny, Hammel, and Sali 2010), generated a poor χ^2 value of 5.1 (**Fig 5.4.3.1B**). This discrepancy between experimental and theoretical solution scattering data could be impacted by the lack of an observable C-terminal segment in the crystal structure. Overall, these data suggest that similar to BeiV-RBP, purified JPV-RBP also retains the C-terminal segment following purification. Furthermore, whilst it is still possible for cleavage to have occurred during crystallography, the presence of density within the putative sialic acid binding site analogous to that observed in the BeiV-RBP structure, is suggestive that the JPV-RBP C-terminal extension also binds at this point. These data provide a platform for further investigation into the structural and functional role of the C-terminal segment.

5.5 ASSESSMENT OF JPV-RBP AND BEIV(HMC AND RAT)-RBP SIALIC

ACID ASSOCIATED ACTIVITY

Previous assessments confirmed that the *Jeilongvirus*, JPV, does not appear to interact with sialic acid (Jun, Karabatsos, and Johnson 1977), however, these data contrast our structural findings, which indicate that the putative sialic acid binding pocket presented by JPV-RBP and BeiV-RBP resembles those borne by the HN-type RBPs (Crennell et al. 2000; Kubota et al. 2016; Lawrence et al. 2004; Yuan et al. 2005). Whilst the C-terminal extension of JPV-RBP has not been structurally characterised, electron density indicating the presence of a putative C α chain in the sialic acid interaction site, in combination with the elucidation of the BeiV-RBP C-terminal segment, support a model where these viruses may utilise auto-inhibitory mechanisms to potentially mediate infection.

In support of this, I also observed that HEK293T cells displaying either full-length (IV-TM-Stalk-head-hat) JPV-RBP, BeiV(HMC)-RBP, and BeiV(rat)-RBP, exhibited a lack of hemadsorptive

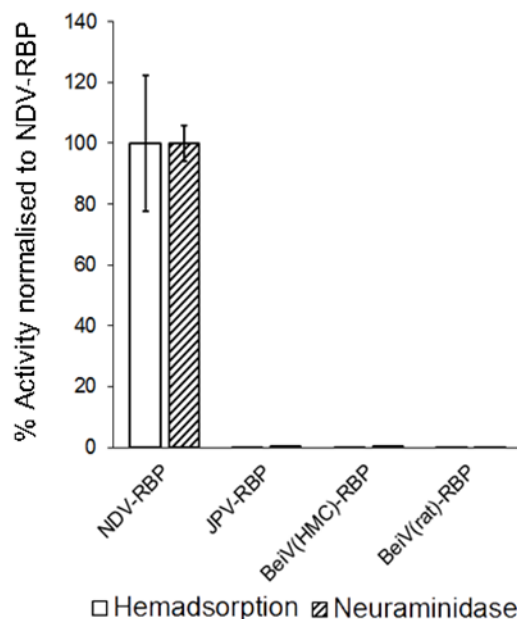


Figure 5.5: JPV-RBP, BeiV(HMC)-RBP and BeiV(rat)-RBP display no hemadsorption and neuraminidase activity. JPV-RBP, BeiV(HMC)-RBP and BeiV(rat)-RBP neuraminidase (Potier et al. 1979) and hemadsorption (Morrison and McGinnes 1989) activity normalized to cell surface expression and a NDV-RBP control. For both hemadsorption (n = 12) and neuraminidase (n = 6) analysis error bars represent the standard deviation across all replicates.

ability and neuraminidase activity, when compared to an NDV-RBP control (**Fig 5.5**) (Potier et al. 1979; Morrison and McGinnes 1989). However, these data could also reflect that JPV-RBP and BeiV-RBP exhibit novel receptor tropism characteristics. Therefore, to assess whether an auto-inhibitory mechanism was utilised by JPV-RBP and BeiV-RBP, truncated versions of protein were required.

5.5.1 Design and small-scale screening of truncated JPV and BeiV RBP constructs

In order to try to determine the role of the C-terminal segment, in particular its impact on sialic acid associated interactions, truncations lacking the extension were required. The appropriate point to truncate the JPV-RBP and BeiV-RBP was determined utilising structural knowledge of the region where the β -propeller ends and where the C-terminal segment starts (**Fig 5.5.1A**). Truncated JPV-RBP and BeiV(rat)-RBP constructs were generated with primers specific to the N-terminal start site,

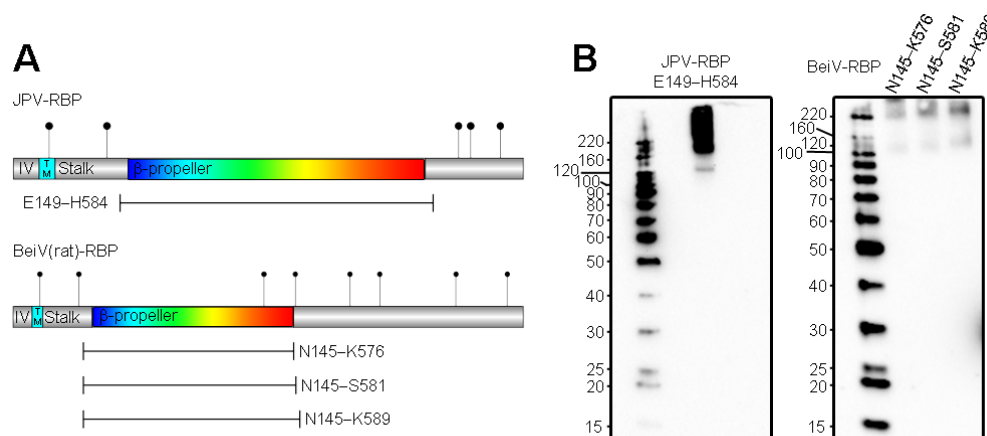


Figure 5.5.1: Construct design and small-scale expression test of truncated JPV-RBP and BeiV(rat)-RBP. (a) Gene diagram displaying features of the JPV-RBP (top) and BeiV(rat)-RBP (bottom), including the IV region, TM domain, stalk, six-bladed β -propeller receptor binding head (rainbow), and C-terminal extension. Predicted N-linked glycosylation sites (NXT/S, where X $\neq$ P) are marked with pins. The designed truncated constructs are illustrated below the schematics. (b) Western blot of small scale expression test of truncated JPV-RBP (left) and BeiV(rat)-RBP (right) constructs, where an SDS-PAGE gel of HEK293T cell supernatants was performed under reducing conditions. Western blot analysis was performed alongside the Benchmark™ Protein Ladder (Thermo, left).

Glu149 and Asn145, respectively, and selected C-terminal end sites, prior to ligation into a pHLsec vector encoding an N-terminal octa-histidine tag and SUMO tag. For JPV-RBP, construct Glu149–His584, and for BeiV(rat)-RBP, construct Asn145–Lys576, Asn145–Ser581 and Asn145–Lys589 were generated, with all constructs exhibiting soluble expression upon small-scale expression screening in human embryonic kidney (HEK)293T cells (**Fig 5.5.1B**). However, the timing of this work meant that further screening of the neuraminidase activity of these engineered proteins, using the Warren method (Warren 1959), was not possible.

5.6 DISCUSSION

The genus *Jeilongvirus* represents one of the newly classified groups of paramyxoviruses for which there is a lack of both functional and structural data (Amarasinghe et al. 2019). A number of jeilongviruses have been isolated from rodents, however, jeilongvirus sequences have also been identified in bat populations. Both rodents and bats are important reservoirs of disease, due to their extensive intra- and interspecies behaviours (Drexler et al. 2012). Continued viral surveillance strategies have begun to expand our understanding of the diversity of jeilongviral genome structures, (Jack et al. 2005; Woo et al. 2016b; Woo, Lau, Wong, Wu, et al. 2012; Li et al. 2006b; Vanmechelen, Bletsa, Laenen, Lopes, Vergote, Beller, Deboutte, Korva, Avšič Županc, et al. 2018; Woo et al. 2011b). The unique jeilongviral genomes includes features such as the additional proteins SH and TM, and the ORF termed 'X', in addition to an RBP much larger than analogous members found throughout other genera of the *Paramyxoviridae* family. Here, sequence analysis revealed that jeilongviral RBPs are likely closely related to HN-type RBPs across the IV, TM, stalk and β -propeller regions, but also encode a C-terminal extension of varying length and diverse sequence. Given the importance of the paramyxoviral RBP in dictating cellular and species tropism, an understanding of the molecular determinants of host-cell attachment are central to defining the propensity of a virus to undergo interspecies transmission (Thibault et al. 2017). To this end, I proceeded to structurally characterise JPV-RBP and BeiV-RBP, with a focus on discerning the organisation and potential role of the C-terminal extension.

Alignment of jeilongviral RBP C-terminus primary sequences reveals a relatively high level of conservation within the first portion of the extension (575^{BeiV-RBP}–700^{BeiV-RBP}) (**Fig 5.2iiB**), followed by a sequence enriched with P/T/S residues, a feature associated with mucin-like domains (Carraway and Hull 1991). The C-terminal extensions of JPV-RBP and BeiV(HMC)-RBP are shorter than the RBPs of other jeilongviruses, as they lack this second segment enriched with P/T/S, and therefore the associated extensive O-linked glycosylation (**Fig 5.2iiA and C**). This may rationalise the issues associated with expression and purification of the longer BeiV(rat)-RBP, where different glycoforms may result in many peaks eluting from a size exclusion column. Additionally, given the

likely propensity of BeiV(rat)-RBP and other jeilongviral RBP C-termini for O-linked glycosylation and P/T/S associated disorder, structural characterisation of these regions utilising X-ray crystallography and electron microscopy is likely not easily possible.

In line with sequence conversation, structure based phylogenetic analysis and pairwise structure overlay alignments of the six-bladed β -propellers of the paramyxoviral RBPs place both JPV and BeiV RBPs as close structural relatives of the HN-type RBPs borne by the viruses within the *Respiro*-, *Orthoavula*-, and *Rubulavirus* genera (**Fig 5.4.2.1i**) (Crennell et al. 2000; Kubota et al. 2016; Lawrence et al. 2004; Yuan et al. 2005). In addition, JPV-RBP and BeiV-RBP display a dimer organisation similar to that of the RBPs from viruses within the *Respiro*-, *Orthoavula*-, and *Rubulavirus* genera (**Fig 5.4.2.1ii**) (Crennell et al. 2000; Kubota et al. 2016; Lawrence et al. 2004; Yuan et al. 2005; Yuan et al. 2008). Consistent with this close structural relationship, JPV-RBP and BeiV-RBP present the sequence and structural features associated with the primary and secondary sialic acid sites displayed by the respiroviruses and orthoavulaviruses (**Fig 5.2i** and **5.4.2.2ii**) (Crennell et al. 2000; Langedijk, Daus, and VanOirschot 1997; Iorio et al. 2001). Additionally, the JPV and BeiV RBPs lack sequence and structural features required to interact with characterised morbilliviral and henipaviral receptors (**Fig 5.4.2.2i**) (Santiago et al. 2010; Zhang et al. 2013; Hashiguchi et al. 2011; Bowden, Aricescu, et al. 2008; Xu et al. 2008; Lee et al. 2015; Pryce et al. 2019). However, the apparent relatedness of the JPV-RBP and BeiV-RBP structures to the RBPs presented by sialic acid interacting viruses is contrasted by our data, which confirms a lack of hemadsorption and neuraminidase activity (**Fig 5.5**) (Jun, Karabatsos, and Johnson 1977). However, the observed BeiV-RBP C-terminal segment potentially provides a structure-based rationale for the observed lack of sialic acid associated activity, via the means of auto-inhibition effected by the binding of the ‘hat-like’ domain to the putative sialic acid binding pocket (**Fig 5.4.3A**). The BeiV-RBP C-terminal segment extends from the underneath of the six-bladed β -propeller to traverse the dimeric interface, stabilised by extensive polar interactions (**Fig 5.4.3D**), to form a domain that sits on the site analogous to the sialic acid binding site (**Fig 5.4.3B**). The domain, composed of a β -sheet consisting of four anti-parallel β -strands and single α -helix (**Fig 5.4.3B**), anchors into and pushes

against loop β 2L23 of the propeller, enabling a ‘hook-like’ projection to access the putative binding site (**Fig 5.4.3E**). Here, residues Arg685 and Asn687 interact with the numerous conserved sialic acid binding residues (**Fig 5.4.3F**). The JPV-RBP structure in contrast, lacks an observable C-terminal extension, however presence of density consistent with a C α chain indicates that a similar mechanism is employed (**Fig5.4.3.1A**). This is supported by SAXS data, which calculates a size for soluble JPV-RBP consistent with a dimeric, full-length construct (**Fig 5.4.1B**). However, this data does not preclude the possibility of the JPV-RBP C-terminal segment being cleaved during crystallisation, and therefore the density being caused by the binding of an unidentified ligand during crystallisation.

Whilst structural data indicates that BeiV-RBP and JPV-RBP, and likely other jeilongviral RBPs, utilise an auto-inhibitory mechanism to mediate infectivity, this apparent lack of sialic acid associated activity cannot be definitively rationalised without functionally characterising the truncated RBPs. The production of truncated JPV-RBP and BeiV-RBP (**Fig 5.5.1**), will enable the assessment of the impact of C-terminal extension on neuraminidase functionality. If such work finds no association between the C-terminus and sialic acid associated activity, it might be interesting to explore the potential of the BeiV-RBP ‘hat-like’ domain to perform functions associated with similar folds, as identified with the DALI server (Holm 2019).

This work also doesn’t provide information about the P/T/S rich segment encoded by the jeilongviruses MMLPV-1, MMLPV-2, PMPV-1, TlmV and BeiV(rat) (**Fig 5.2iiA**), which is located beyond the structurally characterised ‘hat-like’ domain. Analysis of the sequence suggests this segment represents a mucin-like domain with extensive O-linked glycosylation, similar to those presented by the pneumoviruses (Leyrat et al. 2014). Whilst high resolution structural characterisation is unlikely to be possible, further work to describe the function of this region is required to understand how it impacts the activity of the RBP and the virus. Additionally, until the sialic acid associated activity of the jeilongviral RBPs can be fully understood, due to the large size of the jeilongviral RBPs and the overall structural plasticity of the paramyxoviral β -propellers

(Bowden, Crispin, Jones, et al. 2010), the possibility of novel receptor interactions cannot be excluded.

5.7 CONCLUSION

The structural characterisation of a completely novel paramyxoviral RBP domain unique to the jeilongviral RBPs, illustrates the diversity of receptor binding architectures employed by these viruses. Whilst the specific role(s) of this domain are yet to be understood, our structural data provides insight into a potential convergent auto-inhibitory mechanism analogous to that utilised by the NDV Ulster strain for mediation of viral infectivity (Yuan et al. 2012). To grasp the potential risk of jeilongviral interspecies transmission and the disease they may induce in humans or domestic animal species, further work to better understand sialic acid associated activity of these viruses is imperative, in addition to characterisation of the P/T/S rich domain.

CHAPTER 6

FUTURE PERSPECTIVES AND CONCLUDING REMARKS

6 FUTURE PERSPECTIVES AND CONCLUDING REMARKS

6.1 FUTURE PERSPECTIVES

6.1.1 Novel receptor identification strategies

In this thesis, I have determined the three-dimensional structures of the receptor binding proteins (RBPs) from recently classified *Pararubulavirus* (SosV), *Narmovirus* (MosV and NarV), and *Jeilongvirus* (JPV and BeiV) genera of the *Paramyxoviridae* family. Despite the numerous insights into the structural relationships of these RBPs with those of well-characterised H, HN, and G functionality, the receptor tropism characteristics of these groups of viruses remains a mystery. Indeed, my structural studies of the RBPs of narmoviruses and Sosuga virus (SosV), in particular, reveal that these viruses likely utilise a novel receptor-mediated host-cell entry pathway, and that further study is required to determine whether the jeilongviruses use sialic acid independent mechanisms. However, a true understanding of the interspecies transmission potential of these viruses may only be achieved upon identification of the receptors utilised by these viruses. Alongside structural characterisation, I have also attempted to identify the receptors utilised by these viruses. Given the diversity of both carbohydrate and proteinaceous receptors recognised by paramyxoviruses, a number of receptor hunting techniques were explored, namely glycan array screening, co-immunoprecipitation and haploid genetic screening.

6.1.1.1 Glycan array screening

The surface of cells is coated with a glycocalyx, a complex mixture of glycolipids and glycoproteins, each presenting diverse glycan structures and linkages (Rambourg and Leblond 1967; Smith and Cummings 2014). A number of viruses present lectin-type molecules on their surface capable of utilising this complex mixture to bind and infect cells, with types of glycan and their linkages often dictating tissue and species tropisms (Smith and Cummings 2014; Van Breedam et al. 2014; Wasik, Barnard, and Parrish 2016; Neu, Bauer, and Stehle 2011).

Sialic acids are typically presented as the terminal monosaccharides of glycoconjugates and are bound by a variety of viruses either as a primary receptor, or as a secondary receptor, which aids infection (Wasik, Barnard, and Parrish 2016; Li et al. 2017). Viruses that interact with sialic acid span the Baltimore classification, with members of the families *Orthomyxoviridae*, *Paramyxoviridae*, *Picornaviridae*, *Coronaviridae*, *Reoviridae*, *Polyomaviridae*, *Adenoviridae*, and *Parvoviridae*, encoding proteins with sialic acid specific lectin capabilities (Li et al. 2017; Wasik, Barnard, and Parrish 2016; Ströh and Stehle 2014). A number of types of sialic acid exist, and due to the varying expression of sialic acid modification enzymes in different species and tissues, their presentation varies greatly (Varki A 2009). In the context of humans, the 5-N-acetylneuraminic acid (Neu5Ac) form of sialic acid is most common, and ubiquitously expressed (Ströh and Stehle 2014; Neu, Bauer, and Stehle 2011). In contrast to other mammals, humans lack the sialic acid form, 5-N-glycolylneuraminic acid (Neu5Gc), due to the loss of the enzyme cytidine monophosphate-N-acetylneuraminic acid hydroxylase (CMAH) (Wasik, Barnard, and Parrish 2016), impacting pathogen specificity between different species (Byres et al. 2008; Martin et al. 2005). For example, the polyomaviruses JC, BK and Merkel Cell polyomavirus all are tropic towards Neu5Ac, whereas the primate polyomavirus, simian virus 40, has a more flexible specificity and recognises both Neu5Ac and Neu5Gc (O'Hara, Stehle, and Garcea 2014). Additionally, sialic acid linkages vary across human and animal tissues (Varki A 2009), impacting linkage dependent virus interactions (Wasik, Barnard, and Parrish 2016). This specificity dictates both tissue and species tropism, and therefore many of the pathobiological features of disease. A classic example is afforded by the presentation of α 2,3-linked Neu5Ac in lower human lung tissue, precipitating severe disease during outbreaks of the α 2,3-linked Neu5Ac tropic avian influenza viruses (de Wit et al. 2008; Kuiken et al. 2012). Furthermore, whilst sialic acid is not the primary receptor used by middle eastern respiratory syndrome coronaviruses (MERS-CoV), low affinity interacts augment infection *in vitro*, potentially contributing to the definition of MERS-CoV host range (Li et al. 2017). These examples demonstrate the diverse sialic acid structures bound by viruses and their different roles in dictating the host range and pathobiological features of disease.

Besides sialic acid based interactions, other glycans may be bound by viruses, such as the histo-blood group antigens (HBGAs) found on red blood cells and the epithelial cells of the gastrointestinal and respiratory tracts (Van Breedam et al. 2014; Ströh and Stehle 2014). A variety of HBGAs are bound by members of the *Caliciviridae* and *Rotaviridae* (Tan and Jiang 2010; Hu et al. 2012). In the case of the noroviruses, by targeting the array of different ABO, secretor and Lewis HBGAs types, these antigens act as determinants of norovirus tissue specificity (Tan and Jiang 2010). Similarly, the sialidase-insensitive P [14] genotype of the *Rotavirus* genus, *Reoviridae* family, also utilises A-type HBGA, interacting via the site homologous to sialic acid binding (Hu et al. 2012). In addition to the HBGAs and sialic acids, many viruses bind glycosaminoglycans (GAGs), a diverse group of polyanionic linear glycans bound a large number of viruses, including members of the *Flaviviridae*, *Togaviridae*, *Retroviridae*, *Papillomaviridae*, *Hepresviridae*, *Paramyxoviridae*, and *Pneumoviridae* (Ströh and Stehle 2014). Our understanding of how viruses and GAGs interact is limited by the lack of high-resolution structure and their intrinsic chemical heterogeneity which complicates analytical techniques (Ströh and Stehle 2014; Van Breedam et al. 2014).

The development of carbohydrate microarrays, where synthetic and natural glycans are immobilised on a surface, have enable the high-throughput analysis of the lectin properties of various viral proteins, including influenza A viruses (Gulati et al. 2014), and rotaviruses (Hu et al. 2012; Smith and Cummings 2014). Given the propensity for the promiscuous paramyxoviral RBP to bind diverse receptors (Bowden, Crispin, Jones, et al. 2010), it is possible that the uncharacterised paramyxoviruses potentially interact with a variety of glycans beyond the sialic acids (Ströh and Stehle 2014; Van Breedam et al. 2014). With this in mind, we began to consider use of a commercially available glycan array screen from Raybiotech, Peachtree Corners, GA, USA. The microarray developed features 100 synthetic glycans immobilised on a glass microarray slide, and utilises biotin-based detection (Hwang et al. 2018). Purified recombinant MosV (Glu202–Thr632), NarV (Glu184–Pro657) and JPV (Glu149–Asn709) RBPs were non-specifically biotinylated by targeting primary amine groups and applied to the glass slide coated with an array of glycans, alongside the biotinylated positive control, a lectin from soybeans (glycine max), expected to bind

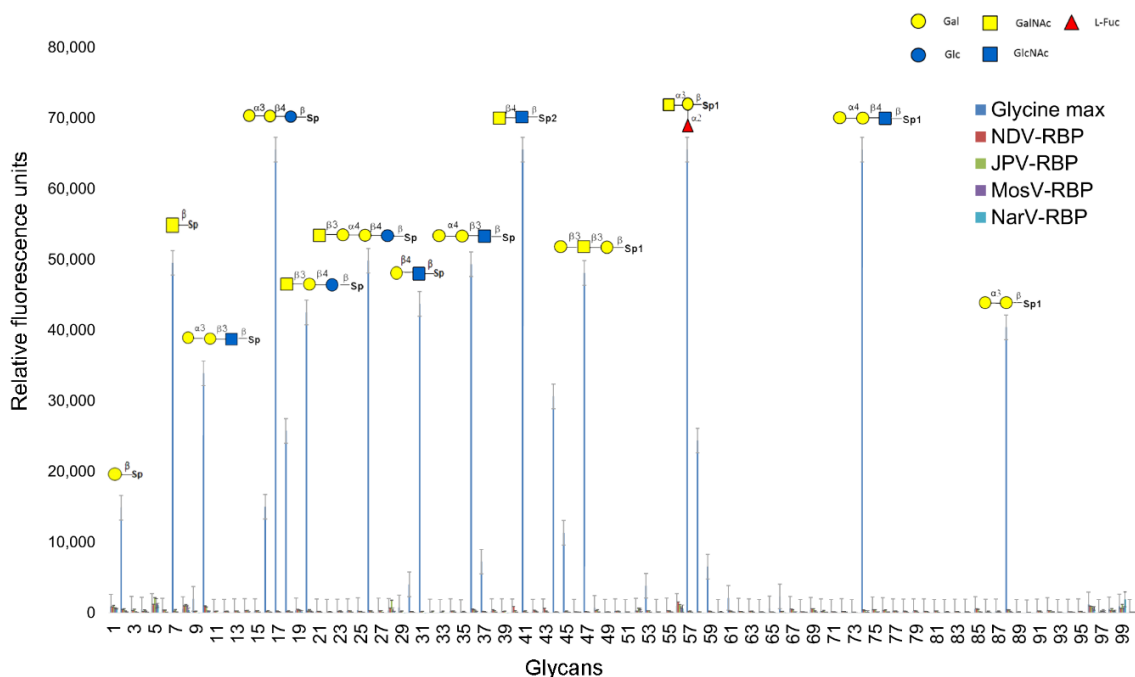


Figure 6.1.1.1: Glycan microarray analysis does not detect binding of NDV-RBP to its known receptor, sialic acid. The fluorescent signal, plotted in a bar format, measured following incubation of a positive control lectin from glycine max (soybean) (ID: L1395, Sigma) (blue), Newcastle disease virus (NDV)-RBP (red), J paramyxovirus (JPV)-RBP (Glu149–Asn709) (green), Mossman virus (MosV)-RBP (Glu202–Thr632) (purple) and Nariva virus (NarV)-RBP (Glu184–Pro657) (cyan) on a microarray with 100 immobilised glycans (<https://www.raybiotech.com/glycan-array-100/>). Relative fluorescence units were calculated using an array analysis program (RayBioTech). A selection of glycans bound by lectin from glycine max have their structures represented, demonstrating binding of lectin from glycine max to residues presenting a terminal galactose (gal), or N-acetylgalactosamine (galNAc) residue.

glycoconjugates terminating with a galactose (gal), or N-acetylgalactosamine (galNAc) residue (Gade et al. 1981), and NDV-RBP, expected to bind terminal Neu5Ac residues (Corfield et al. 1983). The panel of glycans utilised by the RayBiotech microarray represented the following types of glycan; monosaccharides, disaccharides, HBGAs, fucosylated oligosaccharides, sialylated oligosaccharides, natural oligosaccharides, milk oligosaccharides, globo series, GAGs, gangliosides, *O*-glycans, *N*-glycans, α -galactose, aminoglycosides (<https://www.raybiotech.com/glycan-array-100/>).

Following washing, bound protein was detected using a Cy3 equivalent dye conjugated to streptavidin. Interestingly, whilst binding of a positive control lectin from glycine max was detected

(ID: L1395, Sigma), my glycan array analysis did not detect binding of NDV-RBP (Ser47–Gly577) and the purified RBPs from MosV (Glu202–Thr632), NarV (Glu184–Pro657), JPV (Glu149–N709) was not detected (**Fig 6.1.1.1**). Lack of observable NDV-RBP binding to the glycan array was unexpected, as it is known that NDV-RBP is tropic towards glycoconjugates presenting a terminal Neu5Ac (Corfield et al. 1983), a carbohydrate represented within the panel of glycans. I hypothesise that this lack of observable binding could be due to several reasons. Firstly, the NDV-RBP construct was predicted to be a tetramer, a form of the molecule that lacks the avidity associated with the interaction between host-cells and the native virion, which presents multiple tetramERICALLY associated RBPs. Secondly, a rationale for the lack of observed NDV-RBP binding, is potentially the non-specific biotinylation process that could introduce a biotin molecule to the primary amine group of a lysine side chain integral to carbohydrate interactions (Mattson et al. 1993), such as the important Lys236 of the hexapeptide motif (**Fig 1.5.1**). Thirdly, the dual hemagglutinin and neuraminidase functionality of NDV-RBP may result in cleavage of terminal sialic acid (Crennell et al. 2000). As the experiment could not be carried out at temperatures of 4°C, and delivery of the microarray to the USA for fluorescence intensity scanning took place under variable temperature conditions, the enzymatic activity was likely not inhibited, potentially resulting in a loss of NDV-RBP from the array. However, a previous study exploring the sialic acid specificity of human parainfluenza viruses 1 and 3 (PIV1 and PIV3) using glycan array screening implemented by the Consortium for Functional Glycomics (CFG), indicated that the neuraminidase activity of the RBP was inactive at a pH greater than 7.0 (Amonsens et al. 2007). Our experiment was performed using phosphate buffered saline (PBS) pH 8.0, consistent with a NDV-RBP neuraminidase inactive state.

If MosV, NarV and JPV were capable of interacting with sialic acid via the modes understood for other paramyxoviruses, the associated enzymatic capability could produce a false negative result. Additionally, the impact of non-specific biotinylation of Lys residues cannot be known. Therefore, as we cannot confirm the interactions of the related NDV-RBP positive control with the panel of glycans, we cannot definitively conclude that MosV-RBP, NarV-RBP, and JPV-RBP do not bind carbohydrates, at least within the context of this assay.

Whilst glycan array analysis is a potentially effective and high-throughput method of determining virus–glycan interactions, the mode of investigation using a commercial glycan array, where the composition of kit components is unknown, alongside non-specific biotinylation of tetrameric and dimeric RBPs, is not ideal. Future work involving glycan array analysis would benefit from specific biotinylation, usage of samples with higher avidity such as pseudotyped virus or virus-like-particles (VLPs) (Gulati et al. 2014; Amonsén et al. 2007), and experimentation and analysis at 4°C.

6.1.1.2 Co-immunoprecipitation experiments

A common method for identification of novel host-cell receptors bound by viruses is co-immunoprecipitation in combination with mass spectrometry (Pillay and Carette 2015). Immunoprecipitation requires incubation of an Fc conjugated viral protein of interest, with a heterogeneous sample of solubilised membrane protein. The viral protein and any potential interacting partners are then pulled from solution by way of immunoprecipitation, taking advantage of the high affinity interactions between the conjugated Fc tag and protein A or G from *Staphylococcus sp* (Sjöbring, Björck, and Kastern 1991; Sasso, Silverman, and Mannik 1991). The use of this ‘baiting’ method has precedence in the literature, being successful in the identification of ephrinB2 as the receptor for Nipah virus (NiV) (Negrete et al. 2005), transferrin (TfR1) as the receptor for Machupo virus (MacV) and other New World arenaviruses (Radoshitzky et al. 2007), angiotensin-converting enzyme 2 (ACE2) as the primary receptor of SARS-CoV (Li et al. 2003) and dipeptidyl peptidase 4 (DPP4) as the primary receptor of MERS-CoV (Raj et al. 2013). Given the success of this technique in identifying a number of receptors utilised by viruses, including paramyxoviruses, this method was explored for the identification of MosV-RBP binding partners.

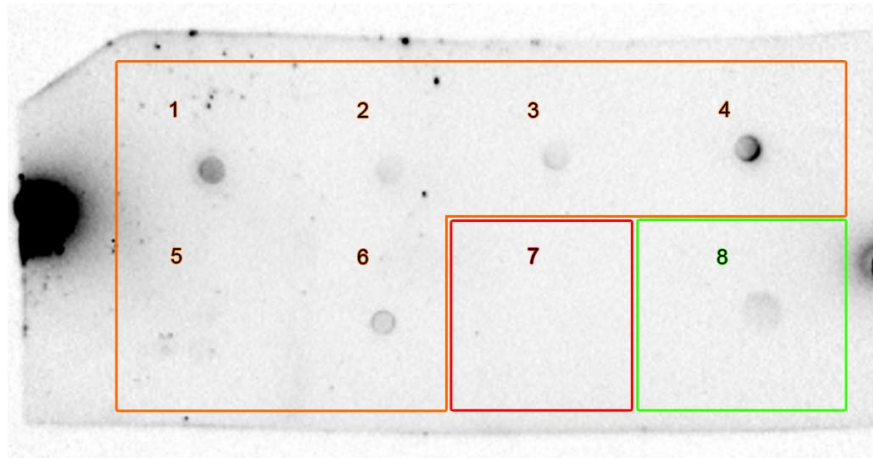
To generate a solution of solubilised membrane protein for the co-immunoprecipitation, baby hamster kidney (BHK)21 cells were grown to confluency and harvested. As initial virus isolation utilised the BHK21 cell line (Miller et al. 2003), it was selected for subsequent pulldown analysis. Cells were lysed using the detergent 1% n-dodecyl- β -D-maltopyranoside (DDM) sol-grade (ID: D301S, Anatrace) and ultrasound based sonication. The solubilised and biotinylated membrane

protein was then exchanged into a biotinylated A8-35 amphipol (ID: BAM01, Anatrace) solution and detergent removed using biobeads (ID: 152-3920, Biorad) (Byrne et al. 2016). The amphipols substitute the detergents, stabilising the membrane protein solution, whilst in their biotinylated form providing a means to tag membrane protein without the disruption to protein-protein interactions potentially caused by non-specific biotinylation (Mattson et al. 1993). The solution was then pre-cleared using unbound protein beads, and beads bound with IgG, to remove membrane proteins from the solution that interact non-specifically with the sample.

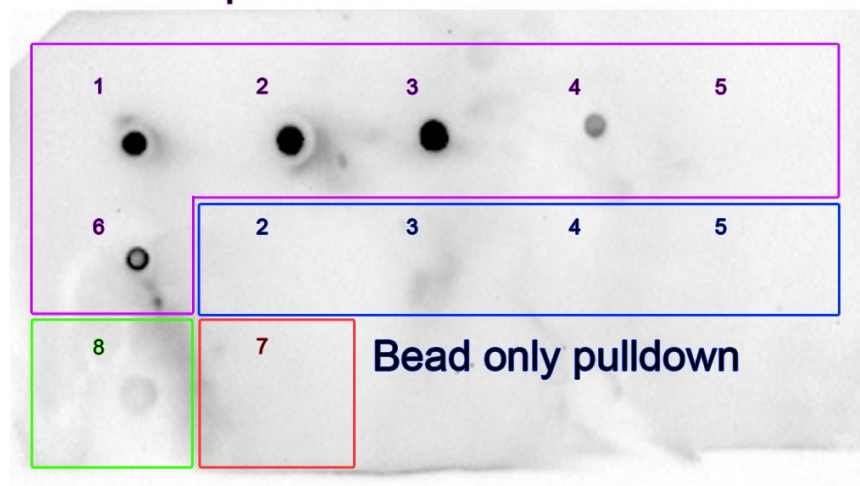
Protein G beads were prepared and bound with Fc-tagged MosV-RBP and NiV-RBP, which had been previously produced in human embryonic kidney (HEK) 293T cells and purified using protein G affinity chromatography. The pre-cleared membrane protein mixture was added to parallel solutions of bound and unbound protein G beads. Following incubation, the supernatant-containing unbound membrane protein was collected and beads washed several times to remove any non-specifically bound membrane protein, and analysed using dot-blotting (**Fig 6.1.1.2**). The RBPs along with any binding partners were subsequently eluted from the beads using an acid wash, before immediate neutralisation. Additionally, to control for the potential background signal caused by contaminating protein G, which has the capability to bind the Fc region of antibodies used during dot-blot analysis, an experiment with unbound protein G beads was performed in parallel and similarly analysed. A biotinylated molecular weight marker was used as a positive control, and PBS as a negative to confirm successful staining.

Binding of novel proteins was determined utilising dot-blot analysis, with an anti-biotin antibody (ID: ab19221, Abcam). The supernatant, wash and elution samples obtained from unbound protein G beads revealed that these steps did not result in a protein G contamination that led to false positives. To confirm the binding interaction between the positive control NiV-RBP and its cognate receptor, ephrinB2, a dot-blot was performed with an antibody specific to ephrinB2 (ID: ab150411,

NiV-RBP pulldown



MosV-RBP pulldown



1 BHK21 cell lysate
2 Unbound protein
supernatant

3 Wash 1
4 Wash 2
5 Wash 3

6 Bead eluate
7 Negative
8 Positive

Fig 6.1.1.2: Dot-blotting indicates MosV-RBP has bound a protein following incubation with baby hamster kidney (BHK)21 cell membrane protein lysate. Nipah virus (NiV)-RBP bound beads (orange), MosV-RBP bound beads (purple) and beads alone (blue) were incubated with a soluble BHK21 membrane protein mixture, prior to washing and elution in acid, with samples collected for analysis throughout. The dot-blot is labelled with numbers indicating the sample applied. Samples applied includes: BHK21 lysate (1), supernatant containing protein which did not bind coated or uncoated beads (2), washes 1, 2 and 3 (3, 4, 5, respectively), the bead eluate potentially containing pulled down protein (6), a PBS only negative control (7), and a biotinylated molecular mass marker as a positive control (8). Both the NiV-RBP and MosV-RBP eluates indicate the presence of a biotinylated protein, potentially a partner pulled from the BHK21 cell lysate.

Abcam). Analysis revealed that the NiV-RBP sample did not appear to pulldown ephrinB2, however as no suitable purified ephrinB2 control was available at the time, potential issues arising from the staining procedure could not be discounted. This inability to detect ephrinB2 binding was contrasted by the signal observed upon anti-biotin staining, where a biotinylated sample was detected in the NiV-RBP bead eluate. Signal was observed upon dot-blot analysis of the MosV-RBP sample, indicating the detection of a putative binding partner. Utilising LysC and trypsin, the whole eluted samples were digested and analysed using tandem liquid chromatography with mass spectrometry (LC-MS-MS) (Grebe and Singh 2011), performed by the Advanced Proteomics Facility, University of Oxford, UK. Unfortunately, whilst mass spectrometry revealed likely non-specific interactions with a list of common background contaminants (Mellacheruvu et al. 2013), analysis failed to identify hits for MosV-RBP and importantly identified no presence of ephrinB2 bound by the NiV-RBP positive control.

The lack of bonafide binding partners precipitated during this experiment could be due to an array of reasons. Despite BHK21 cells expressing the NiV-RBP receptor, ephrinB2 (Negrete et al. 2005; Schountz et al. 2019), no evidence of the receptor was observed during following analysis of the NiV-RBP-Fc co-immunoprecipitation sample. Potentially, the mode of membrane protein isolation was not ideal for pulldown of ephrinB2, and likewise other proteins including any possible MosV-RBP receptors. Amphipols are likely not a ‘one size fits all’ option for production of a stable membrane protein solution, and optimisation of a heterogeneous sample is complex. The potential issues arising from non-specific cell-surface biotinylation of the membrane protein pool was avoided through use of the biotinylated amphipols, however. Nonetheless, the use of biotinylated amphipol necessitated analysis of samples using dot-blotting, in place of the more informative technique, western-blotting.

Subsequent to this work, collaborators in the lab of Prof. Benhur Lee (Icahn School of Medicine at Mount Sinai), analysed binding of MosV-RBP against a panel of cell lines. Their results contrasted previous work (Miller et al. 2003), and indicated that MosV-RBP did not bind BHK21 cells stably expressing T7 RNA polymerase (BSR-T7). This work found that MosV-RBP instead

bound to Vero cells, therefore future pulldown experimentation (with both non-specific biotinylation and biotinylated amphipols) should be performed in this cell line.

6.1.1.3 Haploid screening

Knockout genetic screening approaches can be used to assess proteins and pathways vital to a virus lifecycle, including those involved in attachment and entry. However, the efficient generation of true knockouts is only practical in a haploid system (Pillay and Carette 2015). The haploid genetic screening technique was initially developed in yeast models with the capacity to exist in a haploid form. Studies performed in these yeast models have been useful in understanding questions relating to eukaryotic biology and identifying genes integral to host–pathogen interactions, such as those that support the lifecycle of the alpha-virus-like plant virus, brome mosaic virus (Janda and Ahlquist 1993). However, the majority of viruses that infect mammals are unable to replicate in a yeast model (Pillay and Carette 2015).

Development of the HAP1 cell line expanded capability of haploid genetic screens to explore factors integral to mammalian viruses. The near haploid, HAP1 cell line was adapted from the KBM-7 cell line that was isolated from the bone marrow of a patient suffering with chronic myeloid leukaemia (Andersson et al. 1995; Kotecki, Reddy, and Cochran 1999). The genome-scale cell libraries which contain knockouts in all non-essential genes can be generated through transduction of a retroviral gene-trap vector into the haploid cells, in a process called random insertional mutagenesis (Carette et al. 2009). Phenotypic selection of HAP1 cells with relevant genetic knockouts can be achieved using reporter-based strategies, where resistant cells no longer expressing factors key to virus binding or entry are selected (Pillay and Carette 2015). For example, a pseudotyped vesicular stomatitis virus (VSV) bearing the Ebola virus (EboV) glycoprotein (GP) and engineered green fluorescent protein (eGFP) was utilised in a haploid genetic screen which led to the identification of Niemann-Pick C1 (NPC1) as the receptor for EboV (Carette et al. 2011). Here, HAP1 cells bearing 39 independent mutations in NPC1 were found to not interact with the fluorescently tagged pseudovirus, and therefore were resistant to attachment by EboV-GP. This haploid genetic screening approach was also useful in the identification of the Lassa virus (LASV)

lysosomal receptor, the lysosomal-associated membrane protein-1 (LAMP-1) (Jae et al. 2014). This finding explained the incapability of LASV to infect chicken cells, despite the presentation of the cell surface α -dystroglycan (Lukashevich, Maryankova, and Fidarov 1983; Pillay and Carette 2015).

Given the success of this recently discovered technique, we proceeded to characterise binding of SosV-RBP to HAP1 cells using flowcytometry. For use in flow cytometry and potential haploid screening, the 8xHis-SUMO-SosV-RBP (Asn125–Lys582) was sub-cloned into a pHLsec vector with a C-terminal AviTag™. SosV-RBP was subsequently co-transfected into HEK293T cells, with a pDisplay™ (Thermo) vector encoding the *E. coli* biotin ligase, BirA. The cell culture media was supplemented with biotin, enabling site specific biotinylation of the Avitag™ conjugated to SosV-RBP. SosV-RBP was purified using nickel-affinity and size exclusion chromatography (SEC), and the success of biotinylation determined with western-blot analysis (**Fig 6.1.1.3A**).

Prior to staining with biotinylated SosV-RBP and control antibodies, HAP1 cells were treated with an Fc blocking solution consisting of Human TruStain FcX™ (ID: 422301, Biolegend), mouse IgG (ID:12-371, MerckMillipore), and human IgG (ID: I4506, Sigma), formulated to block the Fc receptor and therefore prevent non-specific binding of staining antibodies. Following washing, cells were treated with different staining mixtures outlined in **Fig 6.1.1.3B**. The protein CD59 was selected as a cell marker for HAP1 cells to ensure correct compensation, to account for spectral overlap of the two fluorochromes, phycoerythrin (PE) and Alexafluor®-647.

Effective staining was observed upon treatment of HAP1 cells with anti-CD59 antibodies conjugated to Alexafluor®-647 (ID: ab187769, Abcam) and PE (ID: 304707, Biolegend), contrasting the lack of signal detected when HAP1 cells were stained with varying dilutions of SosV-RBP (0.5–5 μ g) and streptavidin conjugated to PE (ID: 1475, Biolegend) (**Fig 6.1.1.3C**). To account for potential issues arising from the streptavidin-PE detection, a subset of HAP1 cells were treated initially with a biotinylated anti-CD59 antibody (ID: 555762, BD Biosciences), before the method of staining with a 1:3,000 dilution of streptavidin-PE was confirmed to be appropriate. In parallel, a subset of HAP1 cells were stained with streptavidin-PE only, which demonstrated some non-specific

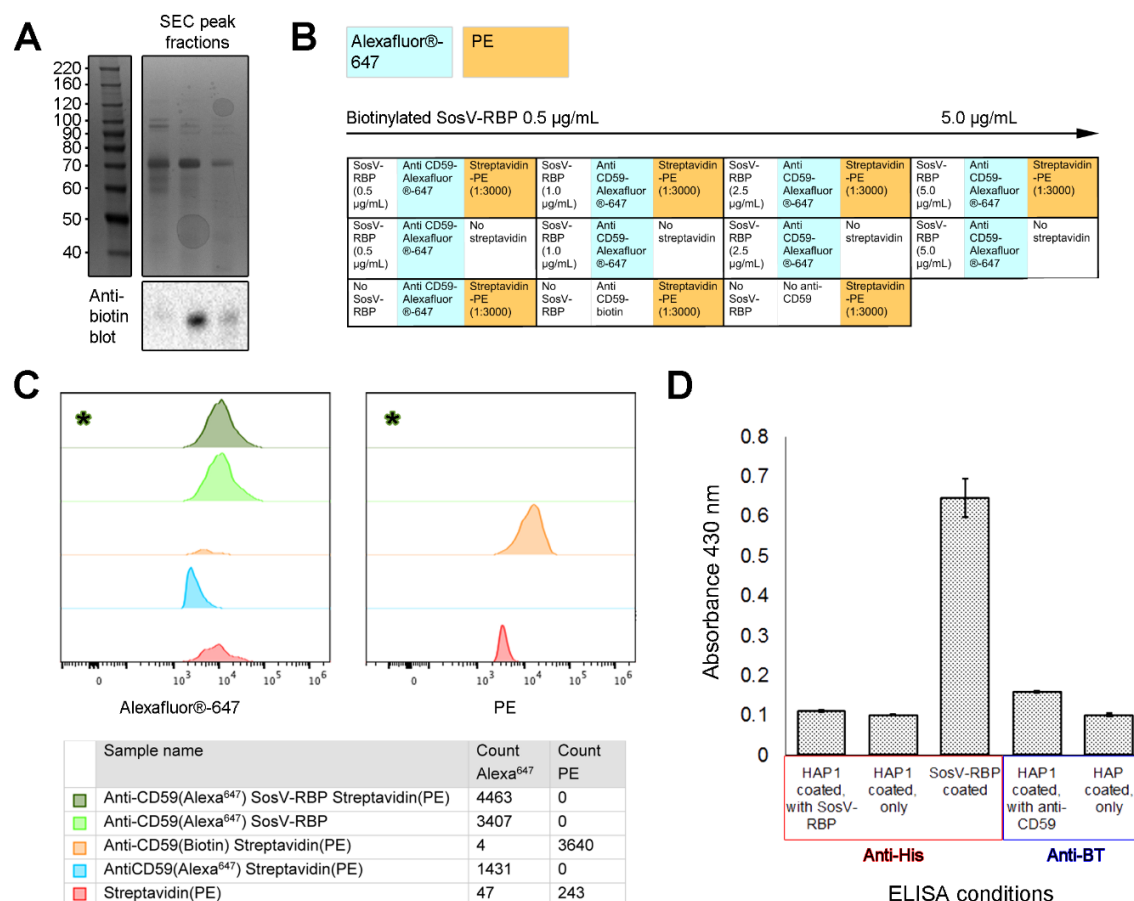


Figure 6.1.1.3: Flow cytometry and enzyme-linked immunosorbent assay (ELISA) data indicate that Sosuga virus (SosV)-RBP is unable to bind HAP1 cells. (a) Following size exclusion chromatography (SEC) of *in vivo* biotinylated SosV-RBP, fractions were analysed using SDS-PAGE and western blotting. The anti-biotin western blot indicated that the sample had been biotinylated, confirming the sample was suitable for subsequent flow cytometry analysis. (b) The mixtures incubated with HAP1 cells, and subsequently analysed using flow cytometry. Mixtures variously contain biotinylated SosV-RBP, antibodies specific to CD59, which are either conjugated to phycoerythrin (PE) (Biolegend), Alexafluor®-647 (Abcam) or biotin (BD Biosciences). If required for detection of biotinylated SosV-RBP or antibody, a secondary incubation was performed with streptavidin-PE. Cyan denotes Alexafluor®-647 staining, and orange staining using PE. (c) HAP1 cells were incubated with specifically biotinylated SosV-RBP detected using streptavidin-PE (*), and the anti-CD59-Alexafluor®-647 (*) antibody. Subsequent, flow cytometry analysis indicated that SosV-RBP does not bind HAP1 cells, following gating no PE signal was detected in the case of SosV-RBP bound cells. This is contrasted against the biotinylated anti-CD59 antibody, which following incubation with streptavidin-PE, demonstrates binding. (d) An ELISA was performed to confirm flow cytometry data. Wells were coated with either HAP1 cells, or directly with SosV-RBP and were variously incubated with SosV-RBP and biotinylated anti-CD59 antibody. Detection was performed with anti-histidine antibody (Abcam) and anti-rabbit antibody conjugated to horseradish peroxidase (HRP) (Vector laboratories), or using an anti-biotin antibody conjugated to HRP. Error bars represent the standard deviation (n=3).

signal when analysed using flow cytometry at wavelength consistent with both PE and Alexafluor®-647.

A basic enzyme-linked immunosorbent assay (ELISA) was setup to confirm the results obtained during flowcytometry. The assay was performed with two different detection methods in parallel, anti-histidine (ID: ab9108; Abcam) and anti-biotin, as no biotinylated SosV-RBP was available, and additionally we did not have a histidine tagged positive control capable of binding HAP1 cells. Following incubation of SosV-RBP on immobilised HAP1 cells, minimal signal was observed, when compared to a HAP1 only negative control and HAP1 cells coated with the anti-CD59 positive control (**Fig 6.1.1.3E**). This signal observed indicates that there may be a limited amount of SosV-RBP binding to the cells, however, the interaction may be non-specific. Alternatively, low signal could be due to poor expression of the receptor or low affinity binding between SosV-RBP and its cognate receptor. Whether or not the signal is genuine, it is nonetheless not detectable in the context of a flow cytometry experiment, which is key to haploid screening, which utilises fluorescence activated cell sorting (FACS). To improve detection of low affinity and/or quantity interactions, flow cytometry may benefit from the use of a VLP or pseudotyped virus, where both the avidity of the RBP-receptor interactions and RBP-detection interactions are increased. Production of SosV pseudotypes has precedence in the literature (Welch et al. 2018), and is a project currently being pursued by collaborators, in the group of Prof. Benhur Lee.

6.1.2 Characterisation of the determinants of receptor interactions

Paramyxovirus RBPs interact with a number of diverse carbohydrate and proteinaceous host-cell presented receptors (Bowden, Aricescu, et al. 2008; Xu et al. 2008; Hashiguchi et al. 2011; Santiago et al. 2010; Zhang et al. 2013; Yuan et al. 2005; Lawrence et al. 2004; Crennell et al. 2000; Kubota et al. 2016; Bowden, Crispin, Jones, et al. 2010). Viruses of *Orthoavula*-, *Metavula*-, *Paraavula*-, *Respiro*-, and *Rubulaviruses* genera all utilise a mechanism highly conserved across six-bladed β -propeller structures to bind and hydrolyse sialic acid (Crennell et al. 2000; Lawrence et al. 2004; Yuan et al. 2005; Kubota et al. 2016; Langedijk, Daus, and VanOirschot 1997). The viruses of the genus, *Morbillivirus*, interact with the immunoglobulin-like molecules signalling lymphocyte

activation molecule F1 (SLAMF1) and nectin-4. The viruses of the genus *Henipavirus* utilise the highly conserved B-class ephrins, ephrinB1, ephrinB2 and ephrinB3 as host-cell receptors (Negrete et al. 2005; Pryce et al. 2019; Negrete et al. 2006b). The determinants of these receptor interactions have been studied extensively utilising both structural and functional techniques. This information has been used to predict the specificity of novel viruses to characterised receptors (Zeltina, Bowden, and Lee 2016), or alternatively a lack of specificity, such as was the case for the paramyxoviruses, Mojiang and Sosuga (MojV and SosV, respectively) (Rissanen et al. 2017; Stelfox and Bowden 2019). Additionally, despite the receptor tropism of the narmoviral, jeilongviral and pararubulaviral RBPs remaining unknown, knowledge of potentially important motifs and structural features provides a means to explore the functional impact of alterations, enabling us to discern the key determinants of specificity.

6.1.2.1 Characterisation of Jeilongvirus sialic acid interactions

The JPV and beilong virus (BeiV) RBP structures maintain the dual functionality sialic acid binding and neuraminidase site at the top of the β -propeller (Kubota et al. 2016; Crennell et al. 2000; Yuan et al. 2005; Lawrence et al. 2004), in addition to a pocket consistent with the second sialic acid binding site (Porotto et al. 2012a; Zaitsev et al. 2004). With this understanding of the motifs central to dictating the hemagglutinin-neuraminidase (HN) activity of paramyxoviral RBPs, we would predict that both JPV and BeiV would demonstrate this archetypal paramyxoviral receptor interaction. However, our work in combination with initial studies of JPV from the 1970s (Jun, Karabatsos, and Johnson 1977) support a putative sialic acid independent mode of entry for the JPV, BeiV, and likely other jeilongviruses. Analogous to the C-terminal extension of NDV-RBP from the Ulster strain which has auto-inhibitory capabilities (Yuan et al. 2012), the C-terminal segment of BeiV-RBP forms a domain that occludes the putative sialic acid binding site. Given this finding, we sought to test whether truncation of this C-terminal region would result in a convalescence of sialic acid associated activity. However, despite successful soluble production of truncated JPV-RBP and BeiV-RBP (Fig 5.5.1), the timing of this thesis has not allowed for subsequent neuraminidase activity analysis using the Warren method (Warren 1959). If enzymatic activity is observed, such

engineered constructs may be readily transplanted to pseudotype of VLP systems, which would enable screening against cells with and without neuraminidase treatment, to assess infection capability beyond attachment. Alternatively, if no such sialic acid interactions are observed, truncated mutants may provide a means to test whether the C-terminal domain is integral to interacting with an array of cell lines.

6.1.2.2 Validation of the determinants of narmoviral receptor specificity

Comparisons to other RBPs, reveals that the narmoviral RBPs lack all known paramyxoviral receptor binding motifs (Bowden, Aricescu, et al. 2008; Xu et al. 2008; Hashiguchi et al. 2011; Santiago et al. 2010; Zhang et al. 2013; Yuan et al. 2005; Lawrence et al. 2004; Crennell et al. 2000; Kubota et al. 2016). Yet, comparative methods exploiting the observed trend for the maintenance of receptor binding sites within individual paramyxoviral genera despite high sequence variation, may also be useful for the identification of novel receptor binding sites on the RBPs of these more divergent viruses (Zeltina, Bowden, and Lee 2016). Pairwise sequence conservation mapping between MosV-RBP and NarV-RBP has enabled the identification of a region of elevated identity, which may constitute a receptor binding site. Flow cytometry analysis performed by collaborators in the lab of Prof Benhur Lee (Icahn School of Medicine at Mount Sinai) confirmed the capability of MosV-RBP to interact with mouse lung adenoma LA-4 and African green monkey kidney epithelial Vero cell lines. In combination with generation of mutant MosV-RBP constructs presenting different surface properties across the potential receptor binding site (**Fig 4.6**), this information provides a means to aid identification of important sites on the MosV-RBP surface despite receptor tropism remaining unknown. These works are planned to take place in the laboratory of Prof Benhur Lee. Additionally, future mutants generated may benefit from consideration of the related bank vole virus (BaVV) (**Fig 4.6.2**), a virus from Russia only recently characterised at the molecular level.

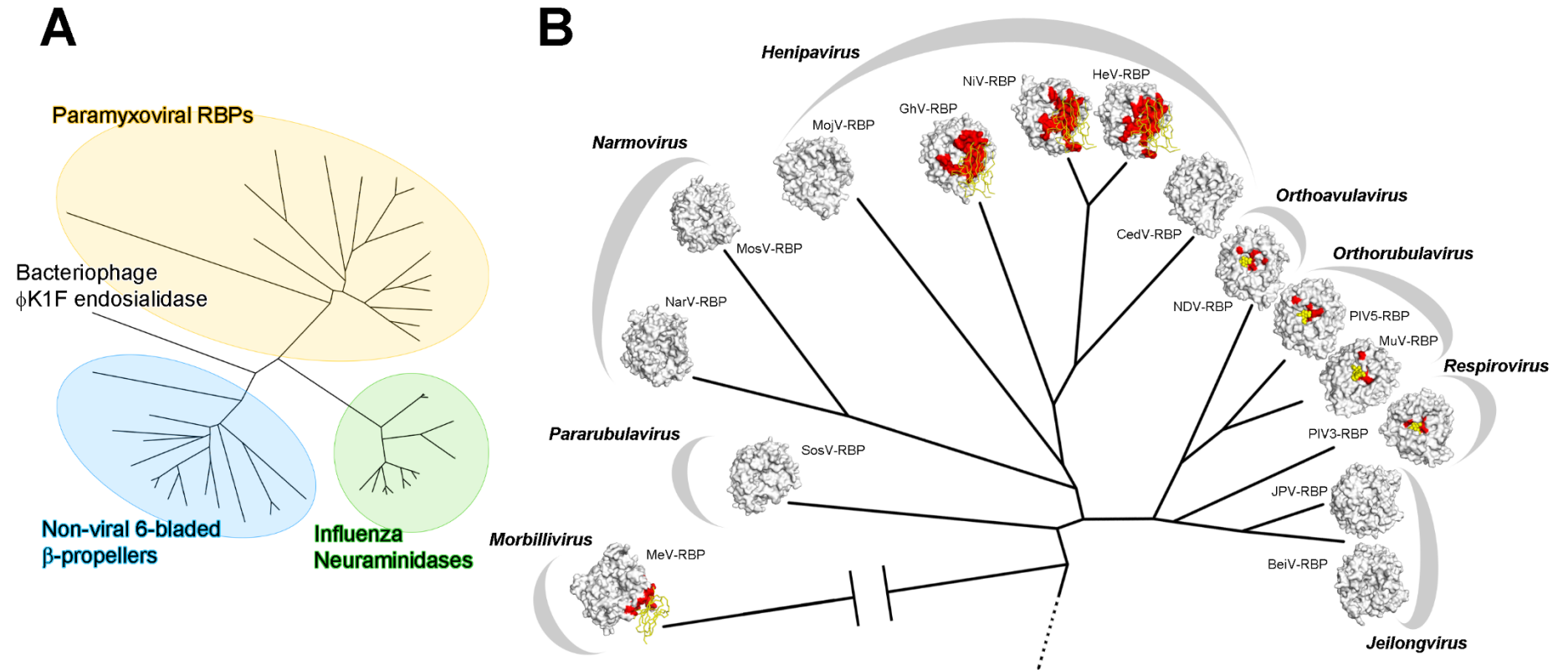


Figure 6.2: Structure based phylogenetic analysis of the six-bladed β -propeller structures of the sialidase superfamily (Wilson et al. 2009). (a) Homologous six-bladed β -propeller structures were identified by application of the NDV-RBP (PDB no. 1E8T) into the DALI server. Structure-based phylogenetic analysis was subsequently performed with the following structures: MeV-RBP, measles virus RBP (2RKC); SosV-RBP (6SG8); NarV-RBP; MosV-RBP; MojV-RBP, Mòjiāng virus RBP (5NOP); GhV-RBP, Ghana virus RBP (4UF7); NiV-RBP (2VSM); HeV-RBP, Hendra virus RBP (2X9M); NDV-RBP; PIV5-RBP, parainfluenza virus 5 RBP (1Z4Y); MuV-RBP, mumps virus RBP (5B2C); PIV3-HN, parainfluenza virus 3 HN (1V2I); JPV-RBP; BeiV-RBP; influenza B virus NA molecules B/Perth (3K36), B/Lyon (4CPO), B/Brisbane (4CPL); influenza A virus NA molecules A/Vietnam N1 (2HTY), A/Tokyo N2 (1IVG), A/Missouri N3 (4HZV), A/Sweden N4 (2HTV), A/Alberta N5 (3SAL), A/England N6 (1V0Z), A/ALB N7 (4QN3), A/Ukraine N8 (2HT5), A/Australia N9 (7NN9), A/Guatemala NL10 (4FVK), A/Peru NL11 (4K3Y); *S. typhimurium* sialidase (3SIL); *T. rangeli* sialidase (1N1S); *T. cruzi* trans-sialidase (1MR5); *V. cholerae* sialidase (1KIT); *S. pneumoniae* NanA (2YA4); *C. perfringens* NanI (2VK5); *S. pneumoniae* NanC (4YZL); *S. pneumoniae* NanB (2VW0); *R. gnavus* intramolecular trans-sialidase (4X47); *M. decora* intramolecular trans-sialidase (1SLL); *H. sapien* Neu2 (1SNT); *M. virifaciens* sialidase (1EUT); *A. fumigatus* sialidase (2XCY); *B. thetaiotaomicron* sialidase (4BBW); *P. aeruginosa* pseudaminidase (2W38); and bacteriophage K1F endosialidase (1V0F). Using pairwise structural superimposition of the six-bladed β -propeller structures an evolutionary distance matrix was calculated by SHP (Stuart et al. 1979), and used to generate an unrooted tree with PHYLIP (Felsenstein 1989). Groups of sialidase and sialidase-like six-bladed β -propellers are designated as follows: paramyxoviral RBPs (orange), influenza neuraminidases (green), non-viral 6-bladed β -propellers (blue), with the bacteriophage ϕ K1F distinct from these three groups. (b) A detailed view of the paramyxovirus RBP arm of the sialidase superfamily structure-based phylogeny. The RBP structures are rendered in surface representation (white) with known receptor footprints colored red. Where relevant receptors are coloured yellow, with carbohydrate receptors rendered as spheres and proteinaceous receptors shown as a ribbon. The calculated evolutionary distances are indicated beside the branches.

6.2 THE PARAMYXOVIRAL RECEPTOR BINDING PROTEINS IN CONTEXT WITH THE SIALIDASE SUPERFAMILY

The six-bladed β -propeller fold is purposed for a diversity of functions, and represents the core fold utilised by a number of protein superfamilies according to the ‘Structural Classification of Proteins’ (SCOP) database (Wilson et al. 2009). Importantly, the fold includes the sialidase superfamily of proteins, an enzyme type encoded by both viruses and organisms across all the domains of life (Roggentin et al. 1989; Quistgaard and Thirup 2009). Humans encode the lysosomal NEU1, cytosolic NEU2 and membrane-bound NEU3 and NEU4 sialidases, which through regulation of sialic acid presentation, aid a diversity of essential pathways, such as cell stress response, protein folding, trafficking, signalling regulation, cell differentiation, cell-cell interactions, and others (Cirillo et al. 2016; Maurice et al. 2016; Chavas et al. 2005). The essential nature of these proteins is demonstrated upon their dysregulation or deficiency, which precipitates cancer or lysosomal storage diseases (De Potter and Schelfhout 1995; Yogalingam et al. 2008). In the case of bacteria such as *Vibrio cholerae*, *Salmonella typhimurium*, and *Clostridium perfringens*, sialidases perform parasitic nutritional functions which contribute to virulence (Corfield 1992). The trypanosomal parasites that cause sleeping sickness and Chagas produce trans-sialidases which are key to invasive phenotypes which lead to pathogenesis in mammalian hosts (Pereira et al. 1996), and are required for survival in the insect host (Nagamune et al. 2004). Additionally, as discussed throughout this thesis, sialidases are utilised by the members of the *Paramyxoviridae* and *Orthomyxoviridae* as a means of host-cell attachment and escape (Kubota et al. 2016; Crennell et al. 2000; Yuan et al. 2005; Lawrence et al. 2004).

Despite the varying specificity and different functions displayed by the sialidases, the residues associated with hydrolysing sialic acid are stringently conserved throughout the sialidase superfamily. These residues include a nucleophilic tyrosine, a triarginyl motif tasked with binding the carboxy group, and three aspartate and/or glutamate residues, which are often numbered according to their blade location within the propeller (Newstead et al. 2008; Kao, Lerner, and Warner 1997; Manzoni et al. 2007; Kobayashi et al. 2000; Kubota et al. 2016; Crennell et al. 2000; Yuan et

al. 2005; Lawrence et al. 2004; Vavricka et al. 2013a; Langedijk, Daus, and VanOirschot 1997; Tailford et al. 2015; Roggentin et al. 1989). A number of sialidase-like molecules utilise the stable yet pliable six-bladed β -propeller fold for adaptation to novel functionality (Stummeyer et al. 2005; Bowden, Crispin, Jones, et al. 2010; Wu, Wu, et al. 2014). An example of smaller variation is afforded the endosialidase from bacteriophage ϕ K1F, where residues critical to the canonical exosialidase activity, such as D₁ and Y₆, are not presented (Stummeyer et al. 2005). To a greater degree, the protein binding RBPs of the *Paramyxoviridae* have diversified from sialidase functionality, and present an array of structures capable of interacting with numerous receptors. This pliability has enabled the morbilliviruses to interact with nectin-4 and signalling lymphocyte activation molecule (SLAM)-F1 (Zhang et al. 2013; Hashiguchi et al. 2011), and the henipaviruses to interact with the B-class ephrin receptors (Bowden, Aricescu, et al. 2008; Zhang et al. 2013; Pryce et al. 2019). Likewise, many RBPs from paramyxoviruses recently characterised at a molecular level are predicted to lack the canonical sialidase functionality or ability to enter cells via the established protein-dependent mechanisms. This adaptation to diverse functionality is reflected by structure-based phylogenetic analysis, which reveals that despite the relatedness of the paramyxoviruses, their sialidase-like structures have a much greater structural diversity than the structurally characterised non-viral sialidases, encoded by both prokaryotes and eukaryotes (**Fig 6.2**). The isolation of the viruses MosV, NarV, JPV, BeiV and SosV and the subsequent structural characterisation of their virion exposed RBP six-bladed β -propellers, has expanded our understanding of the diversity of paramyxoviral RBP architectures.

6.3 CONCLUDING REMARKS

The role of paramyxoviruses as major aetiological agents and perpetrators of significant human and animal outbreaks (Fields, Knipe, and Howley 2013; Griffin, Lin, and Pan 2012; Branche and Falsey 2016; Eaton et al. 2006), raises questions about the potential impact of the sudden expansion of our understanding of the paramyxovirus family (Amarasinghe et al. 2019). An understanding of interactions between a virus and its host(s) can provide vital insights into the pathobiological

characteristics of paramyxoviruses (Rosenberg 2015; Cunningham, Daszak, and Wood 2017). As a protein that mediates the initial stages of infection, the RBP represents a major determinant of paramyxoviral tissue and host specificity, and therefore is a component that defines the potential for interspecies transmission (Thibault et al. 2017). Using X-ray crystallography, I structurally defined RBPs from *Narmovirus*, *Jeilongvirus* and *Pararubulavirus* genera, and sought to use this information to explore the proteins functionally. In parallel, whilst still in a nascent stage, receptor identification strategies have been pursued, laying the foundations for future study. However, despite the receptors utilised by the narmoviruses, jeilongviruses and pararubulaviruses remaining unknown, we have begun to reveal the distinct features unique to these RBPs. Through application of structural methods, I have rationalised why SosV-RBP demonstrates a lack of sialic acid dependent interactions, found that the narmoviral RBPs present a unique surface incompatible to established modes of receptor engagement, and defined the structure of the novel jeilongviral RBP domain, providing a means of testing its functional role in the future. Altogether, these data demonstrate the diversity of architectures adopted by the pliable 6-bladed β -propeller fold, and serve to illuminate the potential for these viruses to continually adapt to new host species. The high propensity of paramyxoviruses to undergo interspecies transmission (Kitchen, Shackelton, and Holmes 2011; Holmes 2010), coupled with an unprecedented diversity of structures inherently linked to virus tropism (Bowden, Crispin, Jones, et al. 2010; Zeltina, Bowden, and Lee 2016; Thibault et al. 2017), necessitates the continued study of this family, including characterisation of members yet to be associated with human disease.

Word count: ~41,500 excluding abstract, table of contents, figures, references and appendix.

APPENDIX A: LIST OF PUBLICATIONS

Publications arising from work undertaken during the course of this DPhil:

Beaty, S., Nachbagauer, R., Hirsh, A., Vigant, F., Duehr, J., Azarm, K. D., **Stelfox, A. J.**, Bowden, T. A., Duprex, W. P., Krammer, F., & Lee, B. 2017. Cross-Reactive and Cross-Neutralizing Activity of Human Mumps Antibodies Against a Novel Mumps Virus From Bats. *J Infect Dis*, 215(2): 209–213. DOI: 10.1093/infdis/jiw534.

Stelfox, A. J. & Bowden, T. A. 2019. A structure-based rationale for sialic acid independent host-cell entry of Sosuga virus. *PNAS*, 201906717. DOI: 10.1073/pnas.1906717116.

Ng, W. M., **Stelfox, A. J.**, Bowden, T. A. 2019. Unraveling viral relationships with structure-based phylogenetic classification. *To be submitted to Virus Evol.*

APPENDIX B: BIBLIOGRAPHY

- Abdullah, Nurshariza, James T. Kelly, Stephen C. Graham, Jamie Birch, Daniel Gonçalves-Carneiro, Tim Mitchell, Robin N. Thompson, Katrina A. Lythgoe, Nicola Logan, Margaret J. Hosie, Vassiliy N. Bavro, Brian J. Willett, Michael P. Heaton, and Dalan Bailey. 2018. 'Structure-Guided Identification of a Nonhuman Morbillivirus with Zoonotic Potential', *Journal of Virology*, 92: e01248-18.
- Abraham, M., N. M. Arroyo-Diaz, Z. Li, J. Zengel, K. Sakamoto, and B. He. 2018. 'Role of Small Hydrophobic Protein of J Paramyxovirus in Virulence', *J Virol*, 92.
- AbuBakar, Sazaly, Li-Yen Chang, A. R. Mohd Ali, S. H. Sharifah, Khatijah Yusoff, and Zulkeflie Zamrod. 2004. 'Isolation and molecular identification of Nipah virus from pigs', *Emerging Infectious Diseases*, 10: 2228-30.
- Adams, P. D., P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, and P. H. Zwart. 2010. 'PHENIX: a comprehensive Python-based system for macromolecular structure solution', *Acta Crystallographica Section D-Biological Crystallography*, 66: 213-21.
- Adu-Gyamfi, Emmanuel, Lori S. Kim, Theodore S. Jardetzky, and Robert A. Lamb. 2016. 'Flexibility of the Head-Stalk Linker Domain of Paramyxovirus HN Glycoprotein Is Essential for Triggering Virus Fusion', *Journal of Virology*, 90: 9172-81.
- Afonine, P. V., R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart, and P. D. Adams. 2012. 'Towards automated crystallographic structure refinement with phenix.refine', *Acta Crystallographica Section D-Structural Biology*, 68: 352-67.
- Afonso, Claudio L., Gaya K. Amarasinghe, Krisztián Bányai, Yīmíng Bào, Christopher F. Basler, Sina Bavari, Nicolás Bejerman, Kim R. Blasdel, François-Xavier Briand, Thomas Bries, Alexander Bukreyev, Charles H. Calisher, Kartik Chandran, Jiāsēn Chéng, Anna N. Clawson, Peter L. Collins, Ralf G. Dietzgen, Olga Dolnik, Leslie L. Domier, Ralf Dürwald, John M. Dye, Andrew J. Easton, Hideki Ebihara, Szilvia L. Farkas, Juliana Freitas-Astúa, Pierre Formenty, Ron A. M. Fouchier, Yànpíng Fù, Elodie Ghedin, Michael M. Goodin, Roger Hewson, Masayuki Horie, Timothy H. Hyndman, Dàohóng Jiāng, Elliot W. Kitajima, Gary P. Kobinger, Hideki Kondo, Gael Kurath, Robert A. Lamb, Sergio Lenardon, Eric M. Leroy, Ci-Xiu Li, Xian-Dan Lin, Lijiāng Liú, Ben Longdon, Szilvia Marton, Andrea Maisner, Elke Mühlberger, Sergey V. Netesov, Norbert Nowotny, Jean L. Patterson, Susan L. Payne, Janusz T. Paweska, Rick E. Randall, Bertus K. Rima, Paul Rota, Dennis Rubbenstroth, Martin Schwemmle, Mang Shi, Sophie J. Smither, Mark D. Stenglein, David M. Stone, Ayato Takada, Calogero Terregino, Robert B. Tesh, Jun-Hua Tian, Keizo Tomonaga, Noël Tordo, Jonathan S. Towner, Nikos Vasilakis, Martin Verbeek, Viktor E. Volchkov, Victoria Wahl-Jensen, John A. Walsh, Peter J. Walker, David Wang, Lin-Fa Wang, Thierry Wetzel, Anna E. Whitfield, Ji Tāo Xiè, Kwok-Yung Yuen, Yong-Zhen Zhang, and Jens H. Kuhn. 2016. 'Taxonomy of the order Mononegavirales: update 2016', *Archives of Virology*, 161: 2351-60.
- Aguilar, Hector C., Bryce A. Henderson, J. Lizbeth Zamora, and Gunner P. Johnston. 2016. 'Paramyxovirus Glycoproteins and the Membrane Fusion Process', *Current clinical microbiology reports*, 3: 142-54.
- Akerstrom, B., and L. Bjorck. 1986. 'A PHYSICOCHEMICAL STUDY OF PROTEIN-G, A MOLECULE WITH UNIQUE IMMUNOGLOBULIN-G-BINDING PROPERTIES', *Journal of Biological Chemistry*, 261: 240-47.
- Alayyoubi, Maher, George P. Leser, Christopher A. Kors, and Robert A. Lamb. 2015. 'Structure of the paramyxovirus parainfluenza virus 5 nucleoprotein-RNA complex', *Proceedings of the National Academy of Sciences of the United States of America*, 112: E1792-E99.

- Albarino, C. G., M. Foltzer, J. S. Towner, L. A. Rowe, S. Campbell, C. M. Jaramillo, B. H. Bird, D. M. Reeder, M. E. Vodzak, P. Rota, M. G. Metcalfe, C. F. Spiropoulou, B. Knust, J. P. Vincent, M. A. Frace, S. T. Nichol, P. E. Rollin, and U. Stroher. 2014. 'Novel Paramyxovirus Associated with Severe Acute Febrile Disease, South Sudan and Uganda, 2012', *Emerging Infectious Diseases*, 20: 211-16.
- Alkhovsky, S., A. Butenko, A. Ereman, and A. Shchetinin. 2018. 'Genetic characterization of bank vole virus (BaVV), a new paramyxovirus isolated from kidneys of bank voles in Russia', *Archives of Virology*, 163: 755-59.
- Alymova, I. V., G. Taylor, V. P. Mishin, M. Watanabe, K. G. Murti, K. Boyd, P. Chand, Y. S. Babu, and A. Portner. 2008. 'Loss of the N-linked glycan at residue 173 of human parainfluenza virus type 1 hemagglutinin-neuraminidase exposes a second receptor-binding site', *J Virol*, 82: 8400-10.
- Amarasinghe, Gaya K., María A. Ayllón, Yīmíng Bào, Christopher F. Basler, Sina Bavari, Kim R. Blasdel, Thomas Briese, Paul A. Brown, Alexander Bukreyev, Anne Balkema-Buschmann, Ursula J. Buchholz, Camila Chabi-Jesus, Kartik Chandran, Chiara Chiapponi, Ian Crozier, Rik L. de Swart, Ralf G. Dietzgen, Olga Dolnik, Jan F. Drexler, Ralf Dürwald, William G. Dundon, W. Paul Duprex, John M. Dye, Andrew J. Easton, Anthony R. Fooks, Pierre B. H. Formenty, Ron A. M. Fouchier, Juliana Freitas-Astúa, Anthony Griffiths, Roger Hewson, Masayuki Horie, Timothy H. Hyndman, Dàohóng Jiāng, Elliott W. Kitajima, Gary P. Kobinger, Hideki Kondō, Gael Kurath, Ivan V. Kuzmin, Robert A. Lamb, Antonio Lavazza, Benhur Lee, Davide Lelli, Eric M. Leroy, Jiànróng Lǐ, Piet Maes, Shin-Yi L. Marzano, Ana Moreno, Elke Mühlberger, Sergey V. Netesov, Norbert Nowotny, Are Nylund, Arnfinn L. Økland, Gustavo Palacios, Bernadett Pályi, Janusz T. Pawęska, Susan L. Payne, Alice Prosperi, Pedro Luis Ramos-González, Bertus K. Rima, Paul Rota, Dennis Rubbenstroth, Mǎng Shī, Peter Simmonds, Sophie J. Smither, Enrica Sozzi, Kirsten Spann, Mark D. Stenglein, David M. Stone, Ayato Takada, Robert B. Tesh, Keizō Tomonaga, Noël Tordo, Jonathan S. Towner, Bernadette van den Hoogen, Nikos Vasilakis, Victoria Wahl, Peter J. Walker, Lin-Fa Wang, Anna E. Whitfield, John V. Williams, F. Murilo Zerbini, Tǎo Zhāng, Yong-Zhen Zhang, and Jens H. Kuhn. 2019. 'Taxonomy of the order Mononegavirales: update 2019', *Archives of Virology*.
- Amman, B. R., C. G. Albarino, B. H. Bird, L. Nyakarahuka, T. K. Sealy, S. Balinandi, A. J. Schuh, S. M. Campbell, U. Stroher, M. E. B. Jones, M. E. Vodzak, D. M. Reeder, W. Kaboyo, S. T. Nichol, and J. S. Towner. 2015. 'A Recently Discovered Pathogenic Paramyxovirus, Sosuga Virus, is Present in Rousettus aegyptiacus Fruit Bats at Multiple Locations in Uganda', *Journal of Wildlife Diseases*, 51: 774-79.
- Amonsén, Mary, David F. Smith, Richard D. Cummings, and Gillian M. Air. 2007. 'Human parainfluenza viruses hPIV1 and hPIV3 bind oligosaccharides with alpha2-3-linked sialic acids that are distinct from those bound by H5 avian influenza virus hemagglutinin', *Journal of Virology*, 81: 8341-45.
- Anderson, D. E., E. J. Dubovi, M. Yu, L. F. Wang, and R. W. Renshaw. 2012. 'Genome characterization of Salem virus reveals its evolutionary intermediate status in the subfamily Paramyxovirinae', *Arch Virol*, 157: 1989-93.
- Andersson, B. S., V. P. Collins, R. Kurzrock, D. W. Larkin, C. Childs, A. Ost, A. Cork, J. M. Trujillo, E. J. Freireich, M. J. Siciliano, and et al. 1995. 'KBM-7, a human myeloid leukemia cell line with double Philadelphia chromosomes lacking normal c-ABL and BCR transcripts', *Leukemia*, 9: 2100-8.
- Aricescu, A. R., W. X. Lu, and E. Y. Jones. 2006. 'A time- and cost-efficient system for high-level protein production in mammalian cells', *Acta Crystallographica Section D-Structural Biology*, 62: 1243-50.
- Baker, Kate S., Shawn Todd, Glenn A. Marsh, Gary Crameri, Jennifer Barr, Alexandra O. Kamins, Alison J. Peel, Meng Yu, David T. S. Hayman, Behzad Nadjm, George Mtove, Benjamin Amos, Hugh Reyburn, Edward Nyarko, Richard Suu-Ire, Pablo R. Murcia, Andrew A. Cunningham, James L. N. Wood, and Lin-Fa Wang. 2013. 'Novel, potentially zoonotic paramyxoviruses from the African straw-colored fruit bat Eidolon helvum', *Journal of Virology*, 87: 1348-58.

- Bandaranayake, Ashok D., Colin Correnti, Byoung Y. Ryu, Michelle Brault, Roland K. Strong, and David J. Rawlings. 2011. 'Daedalus: a robust, turnkey platform for rapid production of decigram quantities of active recombinant proteins in human cell lines using novel lentiviral vectors', *Nucleic Acids Research*, 39: e143-e43.
- Baron, Michael D., Bouna Diop, Felix Njeumi, Brian J. Willett, and Dalan Bailey. 2017. 'Future research to underpin successful peste des petits ruminants virus (PPRV) eradication', *J Gen Virol*, 98: 2635-44.
- Barr, J., S. Todd, G. Crameri, A. Foord, G. Marsh, L. Frazer, J. Payne, J. Harper, K. S. Baker, A. A. Cunningham, J. L. N. Wood, D. Middleton, and L. F. Wang. 2018. 'Animal infection studies of two recently discovered African bat paramyxoviruses, Achimota 1 and Achimota 2', *Scientific Reports*, 8.
- Barr, Jennifer A., Craig Smith, Glenn A. Marsh, Hume Field, and Lin-Fa Wang. 2012. 'Evidence of bat origin for Menangle virus, a zoonotic paramyxovirus first isolated from diseased pigs', *Journal of General Virology*, 93: 2590-94.
- Basler, Christopher F., Adolfo García-Sastre, and Peter Palese. 2005. 'A novel paramyxovirus?', *Emerging Infectious Diseases*, 11: 108-12.
- Battisti, Anthony J., Geng Meng, Dennis C. Winkler, Lori W. McGinnes, Pavel Plevka, Alasdair C. Steven, Trudy G. Morrison, and Michael G. Rossmann. 2012. 'Structure and assembly of a paramyxovirus matrix protein', *Proceedings of the National Academy of Sciences*, 109: 13996-4000.
- Beineke, Andreas, Wolfgang Baumgärtner, and Peter Wohlsein. 2015. 'Cross-species transmission of canine distemper virus-an update', *One health (Amsterdam, Netherlands)*, 1: 49-59.
- Bergfors, Terese. 2007. "Succeeding with seeding: some practical advice." In *Evolving Methods for Macromolecular Crystallography*, edited by Randy J. Read and Joel L. Sussman, 1-10. Dordrecht: Springer Netherlands.
- Bester, J. C. 2016. 'Measles and Measles Vaccination: A Review', *JAMA Pediatr*, 170: 1209-15.
- Biering, Scott B., Andrew Huang, Andy T. Vu, Lindsey R. Robinson, Birgit Bradel-Tretheway, Eric Choi, Benhur Lee, and Hector C. Aguilar. 2012. 'N-Glycans on the Nipah Virus Attachment Glycoprotein Modulate Fusion and Viral Entry as They Protect against Antibody Neutralization', *Journal of Virology*, 86: 11991-2002.
- Bieringer, Maria, Jung Woo Han, Sabine Kendl, Mojtaba Khosravi, Philippe Plattet, and Jürgen Schneider-Schaulies. 2013. 'Experimental Adaptation of Wild-Type Canine Distemper Virus (CDV) to the Human Entry Receptor CD150', *PLOS ONE*, 8: e57488.
- Birch, Jamie, Nicholas Juleff, Michael P. Heaton, Ted Kalbfleisch, James Kijas, and Dalan Bailey. 2013. 'Characterization of ovine Nectin-4, a novel peste des petits ruminants virus receptor', *Journal of Virology*, 87: 4756-61.
- Bonaparte, Matthew I., Antony S. Dimitrov, Katharine N. Bossart, Gary Crameri, Bruce A. Mungall, Kimberly A. Bishop, Vidita Choudhry, Dimiter S. Dimitrov, Lin-Fa Wang, Bryan T. Eaton, and Christopher C. Broder. 2005. 'Ephrin-B2 ligand is a functional receptor for Hendra virus and Nipah virus', *Proceedings of the National Academy of Sciences of the United States of America*, 102: 10652-57.
- Bose, S., A. S. Song, T. S. Jardetzky, and R. A. Lamb. 2014. 'Fusion activation through attachment protein stalk domains indicates a conserved core mechanism of paramyxovirus entry into cells', *J Virol*, 88: 3925-41.
- Bose, S., A. Zokarkar, B. D. Welch, G. P. Leser, T. S. Jardetzky, and R. A. Lamb. 2012. 'Fusion activation by a headless parainfluenza virus 5 hemagglutinin-neuraminidase stalk suggests a modular mechanism for triggering', *Proc Natl Acad Sci U S A*, 109: E2625-34.
- Bousse, T. L., G. Taylor, S. Krishnamurthy, A. Portner, S. K. Samal, and T. Takimoto. 2004. 'Biological significance of the second receptor binding site of Newcastle disease virus hemagglutinin-neuraminidase protein', *J Virol*, 78: 13351-5.
- Bousse, T., and T. Takimoto. 2006. 'Mutation at residue 523 creates a second receptor binding site on human parainfluenza virus type 1 hemagglutinin-neuraminidase protein', *J Virol*, 80: 9009-16.
- Bousse, T., T. Takimoto, W. L. Gorman, T. Takahashi, and A. Portner. 1994. 'REGIONS ON THE HEMAGGLUTININ-NEURAMINIDASE PROTEINS OF HUMAN PARAINFLUENZA

- VIRUS TYPE-1 AND SENDAI VIRUS IMPORTANT FOR MEMBRANE-FUSION', *Virology*, 204: 506-14.
- Boussif, O, F Lezoualc'h, M A Zanta, M D Mergny, D Scherman, B Demeneix, and J P Behr. 1995. 'A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine', *Proceedings of the National Academy of Sciences*, 92: 7297-301.
- Bowden, T. A., A. R. Aricescu, R. J. Gilbert, J. M. Grimes, E. Y. Jones, and D. I. Stuart. 2008. 'Structural basis of Nipah and Hendra virus attachment to their cell-surface receptor ephrin-B2', *Nat Struct Mol Biol*, 15: 567-72.
- Bowden, T. A., M. Crispin, D. J. Harvey, A. R. Aricescu, J. M. Grimes, E. Y. Jones, and D. I. Stuart. 2008. 'Crystal Structure and Carbohydrate Analysis of Nipah Virus Attachment Glycoprotein: a Template for Antiviral and Vaccine Design', *Journal of Virology*, 82: 11628-36.
- Bowden, T. A., M. Crispin, D. J. Harvey, E. Y. Jones, and D. I. Stuart. 2010. 'Dimeric architecture of the Hendra virus attachment glycoprotein: evidence for a conserved mode of assembly', *J. Virol.*, 84: 6208-17.
- Bowden, T. A., M. Crispin, E. Y. Jones, and D. I. Stuart. 2010. 'Shared paramyxoviral glycoprotein architecture is adapted for diverse attachment strategies', *Biochem Soc Trans*, 38: 1349-55.
- Bowden, Timothy R., Marcel Westenberg, Lin-Fa Wang, Bryan T. Eaton, and David B. Boyle. 2001. 'Molecular Characterization of Menangle Virus, a Novel Paramyxovirus which Infects Pigs, Fruit Bats, and Humans', *Virology*, 283: 358-73.
- Bradel-Tretheway, B. G., Q. Liu, J. A. Stone, S. McNally, and H. C. Aguilar. 2015. 'Novel Functions of Hendra Virus G N-Glycans and Comparisons to Nipah Virus', *J Virol*, 89: 7235-47.
- Bragg William, Henry, and Lawrence Bragg William. 1913. 'The reflection of X-rays by crystals', *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 88: 428-38.
- Branche, A. R., and A. R. Falsey. 2016. 'Parainfluenza Virus Infection', *Semin Respir Crit Care Med*, 37: 538-54.
- Brindley, M. A., R. Suter, I. Schestak, G. Kiss, E. R. Wright, and R. K. Plemper. 2013. 'A stabilized headless measles virus attachment protein stalk efficiently triggers membrane fusion', *J Virol*, 87: 11693-703.
- Brown, Vienna R., and Sarah N. Bevins. 2017. 'A review of virulent Newcastle disease viruses in the United States and the role of wild birds in viral persistence and spread', *Veterinary research*, 48: 68-68.
- Burmeister, W. P., R. W. Ruigrok, and S. Cusack. 1992. 'The 2.2 Å resolution crystal structure of influenza B neuraminidase and its complex with sialic acid', *Embo j*, 11: 49-56.
- Burroughs, Amy L., Mary Tachedjian, Gary Crameri, Peter A. Durr, Glenn A. Marsh, and Lin-Fa Wang. 2015. 'Complete genome sequence of teviot paramyxovirus, a novel rubulavirus isolated from fruit bats in australia', *Genome announcements*, 3: e00177-15.
- Büssow, Konrad. 2015. 'Stable mammalian producer cell lines for structural biology', *Current Opinion in Structural Biology*, 32: 81-90.
- Byres, E., A. W. Paton, J. C. Paton, J. C. Lofling, D. F. Smith, M. C. Wilce, U. M. Talbot, D. C. Chong, H. Yu, S. Huang, X. Chen, N. M. Varki, A. Varki, J. Rossjohn, and T. Beddoe. 2008. 'Incorporation of a non-human glycan mediates human susceptibility to a bacterial toxin', *Nature*, 456: 648-52.
- Byrne, Eamon F. X., Ria Sircar, Paul S. Miller, George Hedger, Giovanni Luchetti, Sigrid Nachtergaele, Mark D. Tully, Laurel Mydock-McGrane, Douglas F. Covey, Robert P. Rambo, Mark S. P. Sansom, Simon Newstead, Rajat Rohatgi, and Christian Siebold. 2016. 'Structural basis of Smoothed regulation by its extracellular domains', *Nature*, 535: 517.
- Carette, J. E., C. P. Guimaraes, M. Varadarajan, A. S. Park, I. Wuethrich, A. Godarova, M. Kotecki, B. H. Cochran, E. Spooner, H. L. Ploegh, and T. R. Brummelkamp. 2009. 'Haploid genetic screens in human cells identify host factors used by pathogens', *Science*, 326: 1231-5.
- Carette, Jan E., Matthijs Raaben, Anthony C. Wong, Andrew S. Herbert, Gregor Obernosterer, Nirupama Mulherkar, Ana I. Kuehne, Philip J. Kranzusch, April M. Griffin, Gordon Ruthel, Paola Dal Cin, John M. Dye, Sean P. Whelan, Kartik Chandran, and Thijn R. Brummelkamp.

2011. 'Ebola virus entry requires the cholesterol transporter Niemann-Pick C1', *Nature*, 477: 340-43.
- Carraway, K. L., and S. R. Hull. 1991. 'Cell surface mucin-type glycoproteins and mucin-like domains', *Glycobiology*, 1: 131-8.
- Chang, Andres, and Rebecca E. Dutch. 2012. 'Paramyxovirus fusion and entry: multiple paths to a common end', *Viruses*, 4: 613-36.
- Chang, Tao-Hsin, Fu-Lien Hsieh, Matthias Zebisch, Karl Harlos, Jonathan Elegheert, and E. Yvonne Jones. 2015. 'Structure and functional properties of Norrin mimic Wnt for signalling with Frizzled4, Lrp5/6, and proteoglycan', *eLife*, 4: e06554.
- Chang, Veronica T., Max Crispin, A. Radu Aricescu, David J. Harvey, Joanne E. Nettleship, Janet A. Fennelly, Chao Yu, Kent S. Boles, Edward J. Evans, David I. Stuart, Raymond A. Dwek, E. Yvonne Jones, Raymond J. Owens, and Simon J. Davis. 2007. 'Glycoprotein structural genomics: solving the glycosylation problem', *Structure (London, England : 1993)*, 15: 267-73.
- Chant, K., R. Chan, M. Smith, D. E. Dwyer, and P. Kirkland. 1998. 'Probable human infection with a newly described virus in the family Paramyxoviridae. The NSW Expert Group', *Emerging Infectious Diseases*, 4: 273-75.
- Chao, L. H., M. M. Stratton, I. H. Lee, O. S. Rosenberg, J. Levitz, D. J. Mandell, T. Kortemme, J. T. Groves, H. Schulman, and J. Kuriyan. 2011. 'A mechanism for tunable autoinhibition in the structure of a human Ca²⁺/calmodulin- dependent kinase II holoenzyme', *Cell*, 146: 732-45.
- Chaudhary, Sarika, John E. Pak, Franz Gruswitz, Vinay Sharma, and Robert M. Stroud. 2012. 'Overexpressing human membrane proteins in stably transfected and clonal human embryonic kidney 293S cells', *Nature Protocols*, 7: 453.
- Chavas, L. M., C. Tringali, P. Fusi, B. Venerando, G. Tettamanti, R. Kato, E. Monti, and S. Wakatsuki. 2005. 'Crystal structure of the human cytosolic sialidase Neu2. Evidence for the dynamic nature of substrate recognition', *J Biol Chem*, 280: 469-75.
- Chen, Vincent B., W. Bryan Arendall, 3rd, Jeffrey J. Headd, Daniel A. Keedy, Robert M. Immormino, Gary J. Kapral, Laura W. Murray, Jane S. Richardson, and David C. Richardson. 2010. 'MolProbity: all-atom structure validation for macromolecular crystallography', *Acta crystallographica. Section D, Biological crystallography*, 66: 12-21.
- Chu, F. L., H. L. Wen, G. H. Hou, B. Lin, W. Q. Zhang, Y. Y. Song, G. J. Ren, C. X. Sun, Z. M. Li, and Z. Wang. 2013. 'Role of N-linked glycosylation of the human parainfluenza virus type 3 hemagglutinin-neuraminidase protein', *Virus Res*, 174: 137-47.
- Chua, K. B., W. J. Bellini, P. A. Rota, B. H. Harcourt, A. Tamin, S. K. Lam, T. G. Ksiazek, P. E. Rollin, S. R. Zaki, W. Shieh, C. S. Goldsmith, D. J. Gubler, J. T. Roehrig, B. Eaton, A. R. Gould, J. Olson, H. Field, P. Daniels, A. E. Ling, C. J. Peters, L. J. Anderson, and B. W. Mahy. 2000. 'Nipah virus: a recently emergent deadly paramyxovirus', *Science*, 288: 1432-35.
- Chua, K. B., L.-F. Wang, S. K. Lam, and B. T. Eaton. 2002. 'Full length genome sequence of Tioman virus, a novel paramyxovirus in the genus Rubulavirus isolated from fruit bats in Malaysia', *Archives of Virology*, 147: 1323-48.
- Chua, Kaw Bing, Chong Lek Koh, Poh Sim Hooi, Kong Fatt Wee, Jenn Hui Khong, Beng Hooi Chua, Yee Peng Chan, Mou Eng Lim, and Sai Kit Lam. 2002. 'Isolation of Nipah virus from Malaysian Island flying-foxes', *Microbes and Infection*, 4: 145-51.
- Cirillo, Federica, Andrea Ghiroldi, Chiara Fania, Marco Piccoli, Enrica Torretta, Guido Tettamanti, Cecilia Gelfi, and Luigi Anastasia. 2016. 'NEU3 Sialidase Protein Interactors in the Plasma Membrane and in the Endosomes', *Journal of Biological Chemistry*, 291: 10615-24.
- Cointe, Didier, Yves Leroy, and Frédéric Chirat. 1998. 'Determination of the sialylation level and of the ratio α -(2 $\rightarrow$ 3)/ α -(2 $\rightarrow$ 6) sialyl linkages of N-glycans by methylation and GC/MS analysis', *Carbohydrate Research*, 311: 51-59.
- Connaris, H., T. Takimoto, R. Russell, S. Crennell, I. Moustafa, A. Portner, and G. Taylor. 2002. 'Probing the sialic acid binding site of the hemagglutinin-neuraminidase of Newcastle disease virus: identification of key amino acids involved in cell binding, catalysis, and fusion', *J Virol*, 76: 1816-24.

- Corfield, Anthony P., Herman Higa, James C. Paulson, and Roland Schauer. 1983. 'The specificity of viral and bacterial sialidases for $\alpha(2-3)$ - and $\alpha(2-6)$ -linked sialic acids in glycoproteins', *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, 744: 121-26.
- Corfield, T. 1992. 'Bacterial sialidases--roles in pathogenicity and nutrition', *Glycobiology*, 2: 509-21.
- Cowtan, K. 2010. 'Recent developments in classical density modification', *Acta Crystallogr D Biol Crystallogr*, 66: 470-8.
- Crennell, S. J., E. F. Garman, W. G. Laver, E. R. Vimr, and G. L. Taylor. 1993. 'Crystal structure of a bacterial sialidase (from *Salmonella typhimurium* LT2) shows the same fold as an influenza virus neuraminidase', *Proc Natl Acad Sci U S A*, 90: 9852-6.
- Crennell, S., T. Takimoto, A. Portner, and G. Taylor. 2000. 'Crystal structure of the multifunctional paramyxovirus hemagglutinin-neuraminidase', *Nature Structural Biology*, 7: 1068-74.
- Cunningham, Andrew A., Peter Daszak, and James L. N. Wood. 2017. 'One Health, emerging infectious diseases and wildlife: two decades of progress?', *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 372: 20160167.
- Dalby, Brian, Sharon Cates, Adam Harris, Elise C. Ohki, Mary L. Tilkins, Paul J. Price, and Valentina C. Ciccarone. 2004. 'Advanced transfection with Lipofectamine 2000 reagent: primary neurons, siRNA, and high-throughput applications', *Methods*, 33: 95-103.
- Daszak, Peter, Andrew A. Cunningham, and Alex D. Hyatt. 2000. 'Emerging Infectious Diseases of Wildlife-- Threats to Biodiversity and Human Health', *Science*, 287: 443-49.
- De Potter, C. R., and A. M. Schelfhout. 1995. 'The neu-protein and breast cancer', *Virchows Arch*, 426: 107-15.
- de Wit, Emmie, Yoshihiro Kawaoka, Menno D. de Jong, and Ron A. M. Fouchier. 2008. 'Pathogenicity of highly pathogenic avian influenza virus in mammals', *Vaccine*, 26 Suppl 4: D54-D58.
- Dhiman, N., R. M. Jacobson, and G. A. Poland. 2004. 'Measles virus receptors: SLAM and CD46', *Rev Med Virol*, 14: 217-29.
- Dorig, R. E., A. Marcil, A. Chopra, and C. D. Richardson. 1993. 'The human CD46 molecule is a receptor for measles virus (Edmonston strain)', *Cell*, 75: 295-305.
- Dosztányi, Zsuzsanna. 2018. 'Prediction of protein disorder based on IUPred', *Protein science : a publication of the Protein Society*, 27: 331-40.
- Drexler, J. F., V. M. Corman, M. A. Muller, G. D. Maganga, P. Vallo, T. Binger, F. Gloza-Rausch, A. Rasche, S. Yordanov, A. Seebens, S. Oppong, Y. A. Sarkodie, C. Pongombo, A. N. Lukashev, J. Schmidt-Chanasit, A. Stocker, A. J. B. Carneiro, S. Erbar, A. Maisner, F. Fronhoffs, R. Buettner, E. K. V. Kalko, T. Kruppa, C. R. Franke, R. Kallies, E. R. N. Yandoko, G. Herrler, C. Reusken, A. Hassanin, D. H. Kruger, S. Matthee, R. G. Ulrich, E. M. Leroy, and C. Drosten. 2012. 'Bats host major mammalian paramyxoviruses', *Nature Communications*, 3.
- Duignan, P. J., M. F. Van Bresse, J. D. Baker, M. Barbieri, K. M. Colegrove, S. De Guise, R. L. de Swart, G. Di Guardo, A. Dobson, W. P. Duprex, G. Early, D. Fauquier, T. Goldstein, S. J. Goodman, B. Grenfell, K. R. Groch, F. Gulland, A. Hall, B. A. Jensen, K. Lamy, K. Matassa, S. Mazzariol, S. E. Morris, O. Nielsen, D. Rotstein, T. K. Rowles, J. T. Saliki, U. Siebert, T. Waltzek, and J. F. Wellehan. 2014. 'Phocine distemper virus: current knowledge and future directions', *Viruses*, 6: 5093-134.
- Eaton, B. T., C. C. Broder, D. Middleton, and L. F. Wang. 2006. 'Hendra and Nipah viruses: different and dangerous', *Nat. Rev. Microbiol.*, 4: 23-35.
- Elbein, A. D., J. E. Tropea, Marie Mitchell, and Gur Kaushal. 1990. *Kifunensine, a potent inhibitor of the glycoprotein processing mannosidase I*.
- Elegheert, Jonathan, Ester Behiels, Benjamin Bishop, Suzanne Scott, Rachel E. Woolley, Samuel C. Griffiths, Eamon F. X. Byrne, Veronica T. Chang, David I. Stuart, E. Yvonne Jones, Christian Siebold, and A. Radu Aricescu. 2018. 'Lentiviral transduction of mammalian cells for fast, scalable and high-level production of soluble and membrane proteins', *Nature Protocols*, 13: 2991-3017.
- Ellis, J. A. 2010. 'Bovine parainfluenza-3 virus', *Vet Clin North Am Food Anim Pract*, 26: 575-93.

- Emsley, P., and K. Cowtan. 2004. 'Coot: model-building tools for molecular graphics', *Acta Crystallographica Section D-Biological Crystallography*, 60: 2126-32.
- Emsley, P., B. Lohkamp, W. G. Scott, and K. Cowtan. 2010. 'Features and development of Coot', *Acta Crystallographica Section D-Biological Crystallography*, 66: 486-501.
- Enders, G. 1996. 'Paramyxoviruses.' in th and S. Baron (eds.), *Medical Microbiology* (University of Texas Medical Branch at Galveston
- The University of Texas Medical Branch at Galveston.: Galveston (TX)).
- Fairhead, Michael, and Mark Howarth. 2015. 'Site-specific biotinylation of purified proteins using BirA', *Methods in molecular biology (Clifton, N.J.)*, 1266: 171-84.
- Felsenstein, J. 1989. 'PHYLIP - Phylogeny Inference Package (Version 3.2)', *Cladistics*, 5: 164-66.
- Fields, Bernard N., David M. Knipe, and Peter M. Howley. 2013. *Fields virology* (Wolters Kluwer Health/Lippincott Williams & Wilkins: Philadelphia).
- Francica, Joseph R., Angel Varela-Rohena, Andrew Medvec, Gabriela Plesa, James L. Riley, and Paul Bates. 2010. 'Steric shielding of surface epitopes and impaired immune recognition induced by the ebola virus glycoprotein', *Plos Pathogens*, 6: e1001098-e98.
- Fujii, Yutaka, Takemasa Sakaguchi, Katsuhiko Kiyotani, and Tetsuya Yoshida. 1999. 'Comparison of Substrate Specificities against the Fusion Glycoprotein of Virulent Newcastle Disease Virus between a Chick Embryo Fibroblast Processing Protease and Mammalian Subtilisin-Like Proteases', *Microbiology and Immunology*, 43: 133-40.
- Gade, W., M. A. Jack, J. B. Dahl, E. L. Schmidt, and F. Wold. 1981. 'The isolation and characterization of a root lectin from soybean (*Glycine max* (L), cultivar Chippewa)', *J Biol Chem*, 256: 12905-10.
- Ghai, Ria R., Mary I. Thurber, Azza El Bakry, Colin A. Chapman, and Tony L. Goldberg. 2016. 'Multi-method assessment of patients with febrile illness reveals over-diagnosis of malaria in rural Uganda', *Malaria Journal*, 15: 460.
- Glennon, Emma E., Olivier Restif, Silke Riesle Sbarbaro, Romain Garnier, Andrew A. Cunningham, Richard D. Suu-Ire, Richard Osei-Amponsah, James L. N. Wood, and Alison J. Peel. 2018. 'Domesticated animals as hosts of henipaviruses and filoviruses: A systematic review', *The Veterinary Journal*, 233: 25-34.
- Grebe, Stefan Kg, and Ravinder J. Singh. 2011. 'LC-MS/MS in the Clinical Laboratory - Where to From Here?', *The Clinical biochemist. Reviews*, 32: 5-31.
- Griffin, Diane E., Wen-Hsuan Lin, and Chien-Hsiung Pan. 2012. 'Measles virus, immune control, and persistence', *FEMS Microbiology Reviews*, 36: 649-62.
- Grimsley, G. R., and C. N. Pace. 2004. 'Spectrophotometric determination of protein concentration', *Curr Protoc Protein Sci*, Chapter 3: Unit 3.1.
- Grobe, K., B. Sartori, C. Traving, R. Schauer, and P. Roggentin. 1998. 'Enzymatic and molecular properties of the *Clostridium tertium* sialidase', *Journal of Biochemistry*, 124: 1101-10.
- Grueninger-Leitch, F., A. D'Arcy, B. D'Arcy, and C. Chène. 1996. 'Deglycosylation of proteins for crystallization using recombinant fusion protein glycosidases', *Protein science : a publication of the Protein Society*, 5: 2617-22.
- Gulati, Shelly, Yi Lasanajak, David F. Smith, Richard D. Cummings, and Gillian M. Air. 2014. 'Glycan array analysis of influenza H1N1 binding and release', *Cancer biomarkers : section A of Disease markers*, 14: 43-53.
- Gut, H., S. J. King, and M. A. Walsh. 2008. 'Structural and functional studies of *Streptococcus pneumoniae* neuraminidase B: An intramolecular trans-sialidase', *Febs Letters*, 582: 3348-52.
- Guttman, Miklos, and Kelly K. Lee. 2016. 'Site-Specific Mapping of Sialic Acid Linkage Isomers by Ion Mobility Spectrometry', *Analytical chemistry*, 88: 5212-17.
- Hadac, Elizabeth M., Kah-Whye Peng, Takafumi Nakamura, and Stephen J. Russell. 2004. 'Reengineering paramyxovirus tropism', *Virology*, 329: 217-25.
- Harrison, Megan S., Takemasa Sakaguchi, and Anthony P. Schmitt. 2010. 'Paramyxovirus assembly and budding: building particles that transmit infections', *The international journal of biochemistry & cell biology*, 42: 1416-29.

- Hashiguchi, T., M. Kajikawa, N. Maita, M. Takeda, K. Kuroki, K. Sasaki, D. Kohda, Y. Yanagi, and K. Maenaka. 2007. 'Crystal structure of measles virus hemagglutinin provides insight into effective vaccines', *Proc Natl Acad Sci U S A*, 104: 19535-40.
- Hashiguchi, T., T. Ose, M. Kubota, N. Maita, J. Kamishikiryo, K. Maenaka, and Y. Yanagi. 2011. 'Structure of the measles virus hemagglutinin bound to its cellular receptor SLAMF', *Nat Struct Mol Biol*, 18: 135-41.
- Hashiguchi, Takao, Yoshinari Fukuda, Rei Matsuoka, Daisuke Kuroda, Marie Kubota, Yuta Shirogane, Shumpei Watanabe, Kouhei Tsumoto, Daisuke Kohda, Richard Karl Plemper, and Yusuke Yanagi. 2018. 'Structures of the prefusion form of measles virus fusion protein in complex with inhibitors', *Proceedings of the National Academy of Sciences*, 115: 2496-501.
- He, Biao, Reay G. Paterson, Nicola Stock, Joan E. Durbin, Russell K. Durbin, Stephen Goodbourn, Richard E. Randall, and Robert A. Lamb. 2002. 'Recovery of Paramyxovirus Simian Virus 5 with a V Protein Lacking the Conserved Cysteine-rich Domain: The Multifunctional V Protein Blocks both Interferon- β Induction and Interferon Signaling', *Virology*, 303: 15-32.
- Heckman, Karin L., and Larry R. Pease. 2007. 'Gene splicing and mutagenesis by PCR-driven overlap extension', *Nature Protocols*, 2: 924.
- Herren, Michael, Neeta Shrestha, Marianne Wyss, Andreas Zurbriggen, and Philippe Plattet. 2018. 'Regulatory Role of the Morbillivirus Attachment Protein Head-to-Stalk Linker Module in Membrane Fusion Triggering', *Journal of Virology*, 92: e00679-18.
- Holm, L., S. Kääriäinen, P. Rosenström, and A. Schenkel. 2008. 'Searching protein structure databases with DaliLite v.3', *Bioinformatics (Oxford, England)*, 24: 2780-81.
- Holm, Liisa. 2019. 'Benchmarking fold detection by DaliLite v.5', *Bioinformatics*.
- Holmes, Edward C. 2010. 'The comparative genomics of viral emergence', *Proceedings of the National Academy of Sciences*, 107: 1742-46.
- Hu, A., R. Cattaneo, S. Schwartz, and E. Norrby. 1994. 'Role of N-linked oligosaccharide chains in the processing and antigenicity of measles virus haemagglutinin protein', *J Gen Virol*, 75 (Pt 5): 1043-52.
- Hu, L., S. E. Crawford, R. Czako, N. W. Cortes-Penfield, D. F. Smith, J. Le Pendu, M. K. Estes, and B. V. Prasad. 2012. 'Cell attachment protein VP8* of a human rotavirus specifically interacts with A-type histo-blood group antigen', *Nature*, 485: 256-9.
- Hwang, Hyun-Ju, Jin-Woo Han, Gwang Hoon Kim, and Jong Won Han. 2018. 'Functional Expression and Characterization of the Recombinant N-Acetyl-Glucosamine/N-Acetyl-Galactosamine-Specific Marine Algal Lectin BPL3', *Marine drugs*, 16: 13.
- Iorio, R. M., G. M. Field, J. M. Sauvron, A. M. Mirza, R. T. Deng, P. J. Mahon, and J. P. Langedijk. 2001. 'Structural and functional relationship between the receptor recognition and neuraminidase activities of the Newcastle disease virus hemagglutinin-neuraminidase protein: Receptor recognition is dependent on neuraminidase activity', *Journal of Virology*, 75: 1918-27.
- Jack, P. J. M., D. B. Boyle, B. T. Eaton, and L. F. Wang. 2005. 'The complete genome sequence of J virus reveals a unique genome structure in the family Paramyxoviridae', *Journal of Virology*, 79: 10690-700.
- Jae, Lucas T., Matthijs Raaben, Andrew S. Herbert, Ana I. Kuehne, Ariel S. Wirchnianski, Timothy K. Soh, Sarah H. Stubbs, Hans Janssen, Markus Damme, Paul Saftig, Sean P. Whelan, John M. Dye, and Thijn R. Brummelkamp. 2014. 'Virus entry. Lassa virus entry requires a trigger-induced receptor switch', *Science (New York, N.Y.)*, 344: 1506-10.
- Janda, M., and P. Ahlquist. 1993. 'RNA-dependent replication, transcription, and persistence of brome mosaic virus RNA replicons in *S. cerevisiae*', *Cell*, 72: 961-70.
- Jardetzky, Theodore S., and Robert A. Lamb. 2014. 'Activation of paramyxovirus membrane fusion and virus entry', *Current opinion in virology*, 5: 24-33.
- Jentoft, Neil. 1990. 'Why are proteins O-glycosylated?', *Trends in Biochemical Sciences*, 15: 291-94.
- Johnson, Rebecca I., Mary Tachedjian, Bronwyn A. Clayton, Rachel Layton, Jemma Bergfeld, Lin-Fa Wang, and Glenn A. Marsh. 2019. 'Characterization of Teviot virus, an Australian bat-borne paramyxovirus', *Journal of General Virology*, 100: 403-13.

- Johnson, Sabine, Jun X. Wheeler, Robin Thorpe, Mary Collins, Yasuhiro Takeuchi, and Yuan Zhao. 2018. 'Mass spectrometry analysis reveals differences in the host cell protein species found in pseudotyped lentiviral vectors', *Biologicals*, 52: 59-66.
- Jones, J., S. S. Krag, and M. J. Betenbaugh. 2005. 'Controlling N-linked glycan site occupancy', *Biochim Biophys Acta*, 1726: 121-37.
- Jones, K. E., N. G. Patel, M. A. Levy, A. Storeygard, D. Balk, J. L. Gittleman, and P. Daszak. 2008. 'Global trends in emerging infectious diseases', *Nature*, 451: 990-3.
- Jun, M. H., N. Karabatsos, and R. H. Johnson. 1977. 'A new mouse paramyxovirus (J virus)', *Aust J Exp Biol Med Sci*, 55: 645-7.
- Jurrus, E., D. Engel, K. Star, K. Monson, J. Brandi, L. E. Felberg, D. H. Brookes, L. Wilson, J. H. Chen, K. Liles, M. J. Chun, P. Li, D. W. Gohara, T. Dolinsky, R. Konecny, D. R. Koes, J. E. Nielsen, T. Head-Gordon, W. H. Geng, R. Krasny, G. W. Wei, M. J. Holst, J. A. McCammon, and N. A. Baker. 2018. 'Improvements to the APBS biomolecular solvation software suite', *Protein Science*, 27: 112-28.
- Kao, Y. H., L. Lerner, and T. G. Warner. 1997. 'Stereoselectivity of the Chinese hamster ovary cell sialidase: sialoside hydrolysis with overall retention of configuration', *Glycobiology*, 7: 559-63.
- Karakus, Umut, Thiprampai Thamamongood, Kevin Ciminski, Wei Ran, Sira C. Günther, Marie O. Pohl, Davide Eletto, Csaba Jeney, Donata Hoffmann, Sven Reiche, Jan Schinköthe, Reiner Ulrich, Julius Wiener, Michael G. B. Hayes, Max W. Chang, Annika Hunziker, Emilio Yángüez, Teresa Aydillo, Florian Krammer, Josua Oderbolz, Matthias Meier, Annette Oxenius, Anne Halenius, Gert Zimmer, Christopher Benner, Benjamin G. Hale, Adolfo García-Sastre, Martin Beer, Martin Schwemmle, and Silke Stertz. 2019. 'MHC class II proteins mediate cross-species entry of bat influenza viruses', *Nature*, 567: 109-12.
- Karplus, P. Andrew, and Kay Diederichs. 2012. 'Linking crystallographic model and data quality', *Science (New York, N.Y.)*, 336: 1030-33.
- Kaufmann, S. H. E. 2019. 'Highly affordable vaccines are critical for our continued efforts to reduce global childhood mortality', *Hum Vaccin Immunother*: 1-6.
- Keravala, Annahita, and Michele P. Calos. 2008. 'Site-Specific Chromosomal Integration Mediated by ϕ C31 Integrase.' in Gregory D. Davis and Kevin J. Kayser (eds.), *Chromosomal Mutagenesis* (Humana Press: Totowa, NJ).
- Kitchen, Andrew, Laura A. Shackelton, and Edward C. Holmes. 2011. 'Family level phylogenies reveal modes of macroevolution in RNA viruses', *Proceedings of the National Academy of Sciences of the United States of America*, 108: 238-43.
- Kobayashi, T., M. Ito, K. Ikeda, K. Tanaka, and M. Saito. 2000. 'Purification and characterization of a membrane-bound sialidase from pig liver', *J Biochem*, 127: 569-75.
- Kolakofsky, Daniel. 2016. 'Paramyxovirus RNA synthesis, mRNA editing, and genome hexamer phase: A review', *Virology*, 498: 94-98.
- Kotecki, M., P. S. Reddy, and B. H. Cochran. 1999. 'Isolation and characterization of a near-haploid human cell line', *Exp Cell Res*, 252: 273-80.
- Krissinel, E., and K. Henrick. 2007. 'Inference of macromolecular assemblies from crystalline state', *Journal of Molecular Biology*, 372: 774-97.
- Kubota, M., K. Takeuchi, S. Watanabe, S. Ohno, R. Matsuoka, D. Kohda, S. Nakakita, H. Hiramatsu, Y. Suzuki, T. Nakayama, T. Terada, K. Shimizu, N. Shimizu, M. Shiroishi, Y. Yanagi, and T. Hashiguchi. 2016. 'Trisaccharide containing alpha 2,3-linked sialic acid is a receptor for mumps virus', *Proceedings of the National Academy of Sciences of the United States of America*, 113: 11579-84.
- Kuiken, T., B. Riteau, R. A. M. Fouchier, and G. F. Rimmelzwaan. 2012. 'Pathogenesis of influenza virus infections: the good, the bad and the ugly', *Current opinion in virology*, 2: 276-86.
- Kumar, Naveen, Sunil Maherchandani, Sudhir Kumar Kashyap, Shoor Vir Singh, Shalini Sharma, Kundan Kumar Chaubey, and Hinh Ly. 2014. 'Peste des petits ruminants virus infection of small ruminants: a comprehensive review', *Viruses*, 6: 2287-327.
- Kurth, Andreas, Claudia Kohl, Annika Brinkmann, Arnt Ebinger, Jennifer A. Harper, Lin-Fa Wang, Kristin Mühlendorfer, and Gudrun Wibbelt. 2012. 'Novel Paramyxoviruses in Free-Ranging European Bats', *PLOS ONE*, 7: e38688.

- Laksono, Brigitta M., Rory D. de Vries, Stephen McQuaid, W. Paul Duprex, and Rik L. de Swart. 2016. 'Measles Virus Host Invasion and Pathogenesis', *Viruses*, 8: 210.
- Lambeth, L. S., M. Yu, D. E. Anderson, G. Cramer, B. T. Eaton, and L. F. Wang. 2009. 'Complete genome sequence of Nariva virus, a rodent paramyxovirus', *Archives of Virology*, 154: 199-207.
- Langedijk, J. P. M., F. J. Daus, and J. T. VanOirschot. 1997. 'Sequence and structure alignment of Paramyxoviridae attachment proteins and discovery of enzymatic activity for a morbillivirus hemagglutinin', *Journal of Virology*, 71: 6155-67.
- Langedijk, Johannes P. M., Jozef Janda, Francesco C. Origgi, Claes Örvell, Marc Vandeveld, Andreas Zurbriggen, and Philippe Plattet. 2011. 'Canine distemper virus infects canine keratinocytes and immune cells by using overlapping and distinct regions located on one side of the attachment protein', *Journal of Virology*, 85: 11242-54.
- Lau, Susanna K. P., Patrick C. Y. Woo, Beatrice H. L. Wong, Annette Y. P. Wong, Hoi-Wah Tsoi, Ming Wang, Paul Lee, Huifang Xu, Rosana W. S. Poon, Rongtong Guo, Kenneth S. M. Li, Kwok-Hung Chan, Bo-Jian Zheng, and Kwok-Yung Yuen. 2010. 'Identification and complete genome analysis of three novel paramyxoviruses, Tuhoko virus 1, 2 and 3, in fruit bats from China', *Virology*, 404: 106-16.
- Lawrence, M. C., N. A. Borg, V. A. Streltsov, P. A. Pilling, V. C. Epa, J. N. Varghese, J. L. McKimm-Breschkin, and P. M. Colman. 2004. 'Structure of the haemagglutinin-neuraminidase from human parainfluenza virus type III', *J. Mol. Biol.*, 335: 1343-57.
- Lee, B., O. Pernet, A. A. Ahmed, A. Zeltina, S. M. Beaty, and T. A. Bowden. 2015. 'Molecular recognition of human ephrinB2 cell surface receptor by an emergent African henipavirus', *Proc Natl Acad Sci U S A*, 112: E2156-65.
- Leyrat, Cedric, Guido C. Paesen, James Charleston, Max Renner, and Jonathan M. Grimes. 2014. 'Structural Insights into the Human Metapneumovirus Glycoprotein Ectodomain', *Journal of Virology*, 88: 11611-16.
- Li, Q., X. M. Sun, Z. X. Li, Y. Liu, C. J. Vavricka, J. X. Qi, and G. F. Gao. 2012. 'Structural and functional characterization of neuraminidase-like molecule N10 derived from bat influenza A virus', *Proceedings of the National Academy of Sciences of the United States of America*, 109: 18897-902.
- Li, W., M. J. Moore, N. Vasilieva, J. Sui, S. K. Wong, M. A. Berne, M. Somasundaran, J. L. Sullivan, K. Luzuriaga, T. C. Greenough, H. Choe, and M. Farzan. 2003. 'Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus', *Nature*, 426: 450-4.
- Li, Wentao, Ruben J. G. Hulswit, Ivy Widjaja, V. Stalin Raj, Ryan McBride, Wenjie Peng, W. Widagdo, M. Alejandra Tortorici, Brenda van Dieren, Yifei Lang, Jan W. M. van Lent, James C. Paulson, Cornelis A. M. de Haan, Raoul J. de Groot, Frank J. M. van Kuppeveld, Bart L. Haagmans, and Berend-Jan Bosch. 2017. 'Identification of sialic acid-binding function for the Middle East respiratory syndrome coronavirus spike glycoprotein', *Proceedings of the National Academy of Sciences*, 114: E8508-E17.
- Li, Z., M. Yu, H. Zhang, D. E. Magoffin, P. J. M. Jack, A. Hyatt, H. Y. Wang, and L. F. Wang. 2006a. 'Beilong virus, a novel paramyxovirus with the largest genome of non-segmented negative-stranded RNA viruses', *Virology*, 346: 219-28.
- Li, Zhuo, Cher Hung, Reay G. Paterson, Frank Michel, Sandra Fuentes, Ryan Place, Yuan Lin, Robert J. Hogan, Robert A. Lamb, and Biao He. 2015. 'Type II integral membrane protein, TM of J paramyxovirus promotes cell-to-cell fusion', *Proceedings of the National Academy of Sciences of the United States of America*, 112: 12504-09.
- Li, Zhuo, Meng Yu, Hong Zhang, Danielle E. Magoffin, Philippa J. M. Jack, Alex Hyatt, Hai-Yan Wang, and Lin-Fa Wang. 2006b. 'Beilong virus, a novel paramyxovirus with the largest genome of non-segmented negative-stranded RNA viruses', *Virology*, 346: 219-28.
- Liang, Xiubin, Hong Zhang, Anyu Zhou, and Haiyan Wang. 2003. 'AngRem104, an Angiotensin II-Induced Novel Upregulated Gene in Human Mesangial Cells, Is Potentially Involved in the Regulation of Fibronectin Expression', *Journal of the American Society of Nephrology*, 14: 1443-51.

- Liang, Zhongxiang, Ruyi Yuan, Lei Chen, Xueliang Zhu, and Yongxi Dou. 2016. 'Molecular Evolution and Characterization of Hemagglutinin (H) in Peste des Petits Ruminants Virus', *PLOS ONE*, 11: e0152587-e87.
- Lin, Liang-Tzung, and Christopher D. Richardson. 2016. 'The Host Cell Receptors for Measles Virus and Their Interaction with the Viral Hemagglutinin (H) Protein', *Viruses*, 8: 250.
- Lipkin, W. Ian. 2013. 'The changing face of pathogen discovery and surveillance', *Nature reviews. Microbiology*, 11: 133-41.
- Lisabeth, Erika M., Giulia Falivelli, and Elena B. Pasquale. 2013. 'Eph receptor signaling and ephrins', *Cold Spring Harbor perspectives in biology*, 5: 10.1101/cshperspect.a009159 a59.
- Liu, Q., J. A. Stone, B. Bradel-Trethaway, J. Dabundo, J. A. Benavides Montano, J. Santos-Montanez, S. B. Biering, A. V. Nicola, R. M. Iorio, X. Lu, and H. C. Aguilar. 2013. 'Unraveling a three-step spatiotemporal mechanism of triggering of receptor-induced Nipah virus fusion and cell entry', *PLoS Pathog*, 9: e1003770.
- Liu, Y., N. Li, S. Zhang, F. Zhang, H. Lian, and R. Hu. 2015. 'Parainfluenza Virus 5 as Possible Cause of Severe Respiratory Disease in Calves, China', *Emerg Infect Dis*, 21: 2242-4.
- Luft, J. R., and G. T. DeTitta. 1999. 'A method to produce microseed stock for use in the crystallization of biological macromolecules', *Acta Crystallogr D Biol Crystallogr*, 55: 988-93.
- Lukashevich, I. S., R. F. Maryankova, and F. M. Fidarov. 1983. 'Reproduction of Lassa virus in different cell cultures', *Acta Virol*, 27: 282-5.
- Luo, Y., S. C. Li, Y. T. Li, and M. Luo. 1999. 'The 1.8 angstrom structures of leech intramolecular trans-sialidase complexes: Evidence of its enzymatic mechanism', *Journal of Molecular Biology*, 285: 323-32.
- Maganga, G. D., M. Bourgarel, J. Obame Nkoghe, N. N'Dilimabaka, C. Drosten, C. Paupy, S. Morand, J. F. Drexler, and E. M. Leroy. 2014. 'Identification of an unclassified paramyxovirus in *Coleura afra*: a potential case of host specificity', *PLOS ONE*, 9: e115588.
- Maley, Frank, Robert B. Trimble, Anthony L. Tarentino, and Thomas H. Plummer. 1989. 'Characterization of glycoproteins and their associated oligosaccharides through the use of endoglycosidases', *Analytical Biochemistry*, 180: 195-204.
- Manzoni, M., P. Colombi, N. Papini, L. Rubaga, N. Tiso, A. Preti, B. Venerando, G. Tettamanti, R. Bresciani, F. Argenton, G. Borsani, and E. Monti. 2007. 'Molecular cloning and biochemical characterization of sialidases from zebrafish (*Danio rerio*)', *Biochem J*, 408: 395-406.
- Marsh, G. A., C. de Jong, J. A. Barr, M. Tachedjian, C. Smith, D. Middleton, M. Yu, S. Todd, A. J. Foord, V. Haring, J. Payne, R. Robinson, I. Broz, G. Crameri, H. E. Field, and L. F. Wang. 2012. 'Cedar Virus: A Novel Henipavirus Isolated from Australian Bats', *Plos Pathogens*, 8.
- Marsh, G. A., and L. F. Wang. 2012. 'Hendra and Nipah viruses: why are they so deadly?', *Current opinion in virology*, 2: 242-47.
- Martin, M. J., J. C. Rayner, P. Gagneux, J. W. Barnwell, and A. Varki. 2005. 'Evolution of human-chimpanzee differences in malaria susceptibility: relationship to human genetic loss of N-glycolylneuraminic acid', *Proc Natl Acad Sci U S A*, 102: 12819-24.
- Mattson, G., E. Conklin, S. Desai, G. Nielander, M. D. Savage, and S. Morgensen. 1993. 'A practical approach to crosslinking', *Molecular Biology Reports*, 17: 167-83.
- Maurice, Pascal, Stéphanie Baud, Olga V. Bocharova, Eduard V. Bocharov, Andrey S. Kuznetsov, Charlotte Kawecki, Olivier Bocquet, Beatrice Romier, Laetitia Gorisse, Maxime Ghirardi, Laurent Duca, Sébastien Blaise, Laurent Martiny, Manuel Dauchez, Roman G. Efremov, and Laurent Debelle. 2016. 'New Insights into Molecular Organization of Human Neuraminidase-1: Transmembrane Topology and Dimerization Ability', *Scientific Reports*, 6: 38363.
- McCoy, Airlie J., Ralf W. Grosse-Kunstleve, Paul D. Adams, Martyn D. Winn, Laurent C. Storoni, and Randy J. Read. 2007. 'Phaser crystallographic software', *Journal of Applied Crystallography*, 40: 658-74.
- McGinnes, L. W., and T. G. Morrison. 1995. 'The role of individual oligosaccharide chains in the activities of the HN glycoprotein of Newcastle disease virus', *Virology*, 212: 398-410.
- McPherson, Alexander, and Jose A. Gavira. 2013. 'Introduction to protein crystallization', *Acta crystallographica. Section F, Structural biology communications*, 70: 2-20.

- Melero, José A., Vicente Mas, and Jason S. McLellan. 2017. 'Structural, antigenic and immunogenic features of respiratory syncytial virus glycoproteins relevant for vaccine development', *Vaccine*, 35: 461-68.
- Mellacheruvu, Dattatreya, Zachary Wright, Amber L. Couzens, Jean-Philippe Lambert, Nicole A. St-Denis, Tuo Li, Yana V. Miteva, Simon Hauri, Mihaela E. Sardu, Teck Yew Low, Vincentius A. Halim, Richard D. Bagshaw, Nina C. Hubner, Abdallah Al-Hakim, Annie Bouchard, Denis Faubert, Damian Fermin, Wade H. Dunham, Marilyn Goudreault, Zhen-Yuan Lin, Beatriz Gonzalez Badillo, Tony Pawson, Daniel Durocher, Benoît Coulombe, Ruedi Aebersold, Giulio Superti-Furga, Jacques Colinge, Albert J. R. Heck, Hyungwon Choi, Matthias Gstaiger, Shabaz Mohammed, Ileana M. Cristea, Keiryn L. Bennett, Mike P. Washburn, Brian Raught, Rob M. Ewing, Anne-Claude Gingras, and Alexey I. Nesvizhskii. 2013. 'The CRAPome: a contaminant repository for affinity purification-mass spectrometry data', *Nature methods*, 10: 730-36.
- Mesina, J. E., R. S. Campbell, J. S. Glazebrook, D. B. Copeman, and R. H. Johnson. 1974. 'The pathology of feral rodents in North Queensland', *Tropenmed Parasitol*, 25: 116-27.
- Miller, P. J., D. B. Boyle, B. T. Eaton, and L. F. Wang. 2003. 'Full-length genome sequence of Mossman virus, a novel paramyxovirus isolated from rodents in Australia', *Virology*, 317: 330-44.
- Mirza, A. M., R. T. Deng, and R. M. Iorio. 1994. 'SITE-DIRECTED MUTAGENESIS OF A CONSERVED HEXAPEPTIDE IN THE PARAMYXOVIRUS HEMAGGLUTININ-NEURAMINIDASE GLYCOPROTEIN - EFFECTS ON ANTIGENIC STRUCTURE AND FUNCTION', *Journal of Virology*, 68: 5093-99.
- Mishin, V. P., M. Watanabe, G. Taylor, J. Devincenzo, M. Bose, A. Portner, and I. V. Alymova. 2010a. 'N-linked glycan at residue 523 of human parainfluenza virus type 3 hemagglutinin-neuraminidase masks a second receptor-binding site', *J Virol*, 84: 3094-100.
- Mishin, Vasilii P., Makiko Watanabe, Garry Taylor, John Devincenzo, Michael Bose, Allen Portner, and Irina V. Alymova. 2010b. 'N-linked glycan at residue 523 of human parainfluenza virus type 3 hemagglutinin-neuraminidase masks a second receptor-binding site', *Journal of Virology*, 84: 3094-100.
- Mohd Nor, M. N., C. H. Gan, and B. L. Ong. 2000. 'Nipah virus infection of pigs in peninsular Malaysia', *Rev Sci Tech*, 19: 160-5.
- Molday, L. L., and R. S. Molday. 2014. '1D4: a versatile epitope tag for the purification and characterization of expressed membrane and soluble proteins', *Methods Mol Biol*, 1177: 1-15.
- Molodenskiy, Dmitry, Evgeny Shirshin, Tatiana Tikhonova, Andrey Gruzinov, Georgy Peters, and Francesco Spinozzi. 2017. 'Thermally induced conformational changes and protein-protein interactions of bovine serum albumin in aqueous solution under different pH and ionic strengths as revealed by SAXS measurements', *Physical Chemistry Chemical Physics*, 19: 17143-55.
- Morrison, T. G., and L. W. McGinnes. 1989. 'AVIAN CELLS EXPRESSING THE NEWCASTLE-DISEASE VIRUS HEMAGGLUTININ-NEURAMINIDASE PROTEIN ARE RESISTANT TO NEWCASTLE-DISEASE VIRUS-INFECTION', *Virology*, 171: 10-17.
- Morse, Stephen S., Jonna A. K. Mazet, Mark Woolhouse, Colin R. Parrish, Dennis Carroll, William B. Karesh, Carlos Zambrana-Torrel, W. Ian Lipkin, and Peter Daszak. 2012. 'Prediction and prevention of the next pandemic zoonosis', *The Lancet*, 380: 1956-65.
- Mortlock, M., M. Dietrich, J. Weyer, J. T. Paweska, and W. Markotter. 2019. 'Co-Circulation and Excretion Dynamics of Diverse Rubula- and Related Viruses in Egyptian Rousette Bats from South Africa', *Viruses*, 11.
- Mühlenhoff, Martina, Matthias Eckhardt, and Rita Gerardy-Schahn. 1998. 'Polysialic acid: three-dimensional structure, biosynthesis and function', *Current Opinion in Structural Biology*, 8: 558-64.
- Murray, K., P. Selleck, P. Hooper, A. Hyatt, A. Gould, L. Gleeson, H. Westbury, L. Hiley, L. Selvey, B. Rodwell, and et al. 1995. 'A morbillivirus that caused fatal disease in horses and humans', *Science*, 268: 94-7.

- Nagai, Yoshiyuki, and Hans-Dieter Klenk. 1977. 'Activation of precursors to both glycoproteins of Newcastle disease virus by proteolytic cleavage', *Virology*, 77: 125-34.
- Nagamune, K., A. Acosta-Serrano, H. Uemura, R. Brun, C. Kunz-Renggli, Y. Maeda, M. A. Ferguson, and T. Kinoshita. 2004. 'Surface sialic acids taken from the host allow trypanosome survival in tsetse fly vectors', *J Exp Med*, 199: 1445-50.
- Negrete, Oscar A., Ernest L. Levroney, Hector C. Aguilar, Andrea Bertolotti-Ciarlet, Ronen Nazarian, Sara Tajyar, and Benhur Lee. 2005. 'EphrinB2 is the entry receptor for Nipah virus, an emergent deadly paramyxovirus', *Nature*, 436: 401-05.
- Negrete, Oscar A., Mike C. Wolf, Hector C. Aguilar, Sven Enterlein, Wei Wang, Elke Mühlberger, Stephen V. Su, Andrea Bertolotti-Ciarlet, Ramon Flick, and Benhur Lee. 2006a. 'Two Key Residues in EphrinB3 Are Critical for Its Use as an Alternative Receptor for Nipah Virus', *Plos Pathogens*, 2: e7.
- . 2006b. 'Two key residues in ephrinB3 are critical for its use as an alternative receptor for Nipah virus', *Plos Pathogens*, 2: e7-e7.
- Neu, Ursula, Johannes Bauer, and Thilo Stehle. 2011. 'Viruses and sialic acids: rules of engagement', *Current Opinion in Structural Biology*, 21: 610-18.
- Newstead, Simon L., Jane A. Potter, Jennifer C. Wilson, Guogang Xu, Chin-Hsiang Chien, Andrew G. Watts, Stephen G. Withers, and Garry L. Taylor. 2008. 'The structure of Clostridium perfringens NanI sialidase and its catalytic intermediates', *J Biol Chem*, 283: 9080-88.
- Ng, D T, S W Hiebert, and R A Lamb. 1990. 'Different roles of individual N-linked oligosaccharide chains in folding, assembly, and transport of the simian virus 5 hemagglutinin-neuraminidase', *Molecular and Cellular Biology*, 10: 1989-2001.
- Noh, Ji Yeong, Dae Gwin Jeong, Sun-Woo Yoon, Ji Hyung Kim, Yong Gun Choi, Shien-Young Kang, and Hye Kwon Kim. 2018. 'Isolation and characterization of novel bat paramyxovirus B16-40 potentially belonging to the proposed genus Shaanvirus', *Scientific Reports*, 8: 12533.
- Noton, Sarah L., and Rachel Fearn. 2015. 'Initiation and regulation of paramyxovirus transcription and replication', *Virology*, 479-480: 545-54.
- Noyce, R. S., D. G. Bondre, M. N. Ha, L. T. Lin, G. Sisson, M. S. Tsao, and C. D. Richardson. 2011. 'Tumor Cell Marker PVRL4 (Nectin 4) Is an Epithelial Cell Receptor for Measles Virus', *Plos Pathogens*, 7.
- Noyce, Ryan S., Sebastien Delpeut, and Christopher D. Richardson. 2013. 'Dog nectin-4 is an epithelial cell receptor for canine distemper virus that facilitates virus entry and syncytia formation', *Virology*, 436: 210-20.
- O'Hara, S. D., T. Stehle, and R. Garcea. 2014. 'Glycan receptors of the Polyomaviridae: structure, function, and pathogenesis', *Current opinion in virology*, 7: 73-8.
- Ohishi, Kazue, Rintaro Suzuki, Taro Maeda, Miwako Tsuda, Erika Abe, Takao Yoshida, Yasuyuki Endo, Maki Okamura, Takashi Nagamine, Hanae Yamamoto, Miya Ueda, and Tadashi Maruyama. 2014. 'RECENT HOST RANGE EXPANSION OF CANINE DISTEMPER VIRUS AND VARIATION IN ITS RECEPTOR, THE SIGNALING LYMPHOCYTE ACTIVATION MOLECULE, IN CARNIVORES', *Journal of Wildlife Diseases*, 50: 596-606, 11.
- OIE, FAO. 2015. 'Global control and eradication of peste des petits ruminants', *Investing in veterinary systems, food security and poverty alleviation. PPRV summary document. World Organization for Animal Health (OIE)*.
- Pager, C. T., W. W. Craft, Jr., J. Patch, and R. E. Dutch. 2006. 'A mature and fusogenic form of the Nipah virus fusion protein requires proteolytic processing by cathepsin L', *Virology*, 346: 251-7.
- Panavas, Tadas, Carsten Sanders, and Tauseef R. Butt. 2009. 'SUMO Fusion Technology for Enhanced Protein Production in Prokaryotic and Eukaryotic Expression Systems.' in Helle D. Ulrich (ed.), *SUMO Protocols* (Humana Press: Totowa, NJ).
- Panda, Aruna, Subbiah Elankumaran, Sateesh Krishnamurthy, Zhuhui Huang, and Siba K. Samal. 2004. 'Loss of N-Linked Glycosylation from the Hemagglutinin- Neuraminidase Protein Alters Virulence of Newcastle Disease Virus', *Journal of Virology*, 78: 4965-75.

- Paraskevakou, G., C. Allen, T. Nakamura, P. Zollman, C. D. James, K. W. Peng, M. Schroeder, S. J. Russell, and E. Galanis. 2007. 'Epidermal growth factor receptor (EGFR)-retargeted measles virus strains effectively target EGFR- or EGFRvIII expressing gliomas', *Mol Ther*, 15: 677-86.
- Park, Jaehong, Andrea L. Throop, and Joshua LaBaer. 2015. 'Site-specific recombinational cloning using gateway and in-fusion cloning schemes', *Current protocols in molecular biology*, 110: 3.20.1-3.20.23.
- Pearce-Duvet, Jessica M. C. 2006. 'The origin of human pathogens: evaluating the role of agriculture and domestic animals in the evolution of human disease', *Biological Reviews*, 81: 369-82.
- Pereira, M. E., K. Zhang, Y. Gong, E. M. Herrera, and M. Ming. 1996. 'Invasive phenotype of *Trypanosoma cruzi* restricted to a population expressing trans-sialidase', *Infect Immun*, 64: 3884-92.
- Pernet, Olivier, Yao E. Wang, and Benhur Lee. 2012. 'Henipavirus Receptor Usage and Tropism.' in Benhur Lee and Paul A. Rota (eds.), *Henipavirus: Ecology, Molecular Virology, and Pathogenesis* (Springer Berlin Heidelberg: Berlin, Heidelberg).
- Peroutka, Raymond J., Nabil Elshourbagy, Tara Piech, and Tauseef R. Butt. 2008. 'Enhanced protein expression in mammalian cells using engineered SUMO fusions: secreted phospholipase A2', *Protein science : a publication of the Protein Society*, 17: 1586-95.
- Pettersen, E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin. 2004. 'UCSF Chimera--a visualization system for exploratory research and analysis', *J Comput Chem*, 25: 1605-12.
- Philbey, A. W., P. D. Kirkland, A. D. Ross, R. J. Davis, A. B. Gleeson, R. J. Love, P. W. Daniels, A. R. Gould, and A. D. Hyatt. 1998. 'An apparently new virus (family Paramyxoviridae) infectious for pigs, humans, and fruit bats', *Emerging Infectious Diseases*, 4: 269-71.
- Pillay, Sirika, and Jan E. Carette. 2015. 'Hunting Viral Receptors Using Haploid Cells', *Annual review of virology*, 2: 219-39.
- Poole, E., B. He, R. A. Lamb, R. E. Randall, and S. Goodbourn. 2002. 'The V proteins of simian virus 5 and other paramyxoviruses inhibit induction of interferon-beta', *Virology*, 303: 33-46.
- Popot, J.-L., E. A. Berry, D. Charvolin, C. Creuzenet, C. Ebel, D. M. Engelman, M. Flötenmeyer, F. Giusti, Y. Gohon, P. Hervé, Q. Hong, J. H. Lakey, K. Leonard, H. A. Shuman, P. Timmins, D. E. Warschawski, F. Zito, M. Zoonens, B. Pucci, and C. Tribet. 2003. 'Amphipols: polymeric surfactants for membrane biology research', *Cellular and Molecular Life Sciences CMLS*, 60: 1559-74.
- Porotto, M., M. Fornabaio, G. E. Kellogg, and A. Moscona. 2007. 'A second receptor binding site on human parainfluenza virus type 3 hemagglutinin-neuraminidase contributes to activation of the fusion mechanism', *J Virol*, 81: 3216-28.
- Porotto, M., Z. Salah, I. DeVito, A. Talekar, S. G. Palmer, R. Xu, I. A. Wilson, and A. Moscona. 2012a. 'The second receptor binding site of the globular head of the Newcastle disease virus hemagglutinin-neuraminidase activates the stalk of multiple paramyxovirus receptor binding proteins to trigger fusion', *J Virol*, 86: 5730-41.
- Porotto, Matteo, Zuhair Salah, Ilaria DeVito, Aparna Talekar, Samantha G. Palmer, Rui Xu, Ian A. Wilson, and Anne Moscona. 2012b. 'The second receptor binding site of the globular head of the Newcastle disease virus hemagglutinin-neuraminidase activates the stalk of multiple paramyxovirus receptor binding proteins to trigger fusion', *Journal of Virology*, 86: 5730-41.
- Potier, M., L. Mameli, M. Belisle, L. Dallaire, and S. B. Melancon. 1979. 'FLUOROMETRIC ASSAY OF NEURAMINIDASE WITH A SODIUM (4-METHYLBELLIFERYL-ALPHA-D-N-ACETYLNEURAMINATE) SUBSTRATE', *Analytical Biochemistry*, 94: 287-96.
- Pryce, Rhys, Kristopher Azarm, Ilona Rissanen, Karl Harlos, Thomas A. Bowden, and Benhur Lee. 2019. 'A key region of molecular specificity orchestrates unique ephrin-B1 utilization by Cedar virus', *bioRxiv*: 724138.
- Putnam, Christopher D. 2016. 'Guinier peak analysis for visual and automated inspection of small-angle X-ray scattering data', *Journal of Applied Crystallography*, 49: 1412-19.

- Quistgaard, Esben M., and Søren S. Thirup. 2009. 'Sequence and structural analysis of the Asp-box motif and Asp-box beta-propellers; a widespread propeller-type characteristic of the Vps10 domain family and several glycoside hydrolase families', *BMC Structural Biology*, 9: 46.
- Radoshitzky, S. R., J. Abraham, C. F. Spiropoulou, J. H. Kuhn, D. Nguyen, W. Li, J. Nagel, P. J. Schmidt, J. H. Nunberg, N. C. Andrews, M. Farzan, and H. Choe. 2007. 'Transferrin receptor 1 is a cellular receptor for New World haemorrhagic fever arenaviruses', *Nature*, 446: 92-6.
- Raj, V. S., H. Mou, S. L. Smits, D. H. Dekkers, M. A. Muller, R. Dijkman, D. Muth, J. A. Demmers, A. Zaki, R. A. Fouchier, V. Thiel, C. Drosten, P. J. Rottier, A. D. Osterhaus, B. J. Bosch, and B. L. Haagmans. 2013. 'Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC', *Nature*, 495: 251-4.
- Raman, Rahul, Kannan Tharakaraman, V. Sasisekharan, and Ram Sasisekharan. 2016. 'Glycan-protein interactions in viral pathogenesis', *Current Opinion in Structural Biology*, 40: 153-62.
- Rambo, Robert P. 2017. 'Considerations for Sample Preparation Using Size-Exclusion Chromatography for Home and Synchrotron Sources', *Advances in experimental medicine and biology*, 1009: 31-45.
- Rambo, Robert P., and John A. Tainer. 2013a. 'Accurate assessment of mass, models and resolution by small-angle scattering', *Nature*, 496: 477-81.
- . 2013b. 'Super-Resolution in Solution X-Ray Scattering and Its Applications to Structural Systems Biology', *Annual Review of Biophysics*, 42: 415-41.
- Rambourg, A., and C. P. Leblond. 1967. 'ELECTRON MICROSCOPE OBSERVATIONS ON THE CARBOHYDRATE-RICH CELL COAT PRESENT AT THE SURFACE OF CELLS IN THE RAT', *The Journal of Cell Biology*, 32: 27-53.
- Rendon-Marin, Santiago, Renata da Fontoura Budaszewski, Cláudio Wageck Canal, and Julian Ruiz-Saenz. 2019. 'Tropism and molecular pathogenesis of canine distemper virus', *Virology Journal*, 16: 30.
- Reynard, Olivier, Malgorzata Borowiak, Valentina A. Volchkova, Sebastien Delpout, Mathieu Mateo, and Viktor E. Volchkov. 2009. 'Ebola virus Glycoprotein GP Masks both Its Own Epitopes and the Presence of Cellular Surface Proteins', *Journal of Virology*, 83: 9596-601.
- Rima, B., P. Collins, A. Easton, R. Fouchier, G. Kurath, R. A. Lamb, B. Lee, A. Maisner, P. Rota, and L. F. Wang. 2018. 'Problems of classification in the family Paramyxoviridae', *Archives of Virology*, 163: 1395-404.
- Rissanen, I., A. A. Ahmed, K. Azarm, S. Beaty, P. Hong, S. Nambulli, W. P. Duprex, B. Lee, and T. A. Bowden. 2017. 'Idiosyncratic Mojang virus attachment glycoprotein directs a host-cell entry pathway distinct from genetically related henipaviruses', *Nature Communications*, 8.
- Rockx, Barry. 2014. 'Recent developments in experimental animal models of Henipavirus infection', *Pathogens and Disease*, 71: 199-206.
- Roeder, Peter L. 2011. 'Rinderpest: The end of cattle plague', *Preventive Veterinary Medicine*, 102: 98-106.
- Roggentin, P., B. Rothe, J. B. Kaper, J. Galen, L. Lawrisuk, E. R. Vimr, and R. Schauer. 1989. 'CONSERVED SEQUENCES IN BACTERIAL AND VIRAL SIALIDASES', *Glycoconjugate Journal*, 6: 349-53.
- Rosenberg, Ronald. 2015. 'Detecting the emergence of novel, zoonotic viruses pathogenic to humans', *Cellular and molecular life sciences : CMLS*, 72: 1115-25.
- Rubin, Steven, Michael Eckhaus, Linda J. Rennick, Connor G. G. Bamford, and W. Paul Duprex. 2015. 'Molecular biology, pathogenesis and pathology of mumps virus', *The Journal of pathology*, 235: 242-52.
- Santiago, C., M. L. Celma, T. Stehle, and J. M. Casasnovas. 2010. 'Structure of the measles virus hemagglutinin bound to the CD46 receptor', *Nature Structural & Molecular Biology*, 17: 124-U52.
- Santos-Lopez, G., T. Scior, T. Borraz-Arguello Mdel, V. Vallejo-Ruiz, I. Herrera-Camacho, J. Tapia-Ramirez, and J. Reyes-Leyva. 2009. 'Structure-function analysis of two variants of mumps virus hemagglutinin-neuraminidase protein', *Braz J Infect Dis*, 13: 24-34.

- Sasaki, M., W. Muleya, A. Ishii, Y. Orba, B. M. Hang'ombe, A. S. Mweene, L. Moonga, Y. Thomas, T. Kimura, and H. Sawa. 2014. 'Molecular epidemiology of paramyxoviruses in Zambian wild rodents and shrews', *J Gen Virol*, 95: 325-30.
- Sasso, E. H., G. J. Silverman, and M. Mannik. 1991. 'Human IgA and IgG F(ab')₂ that bind to staphylococcal protein A belong to the VHIII subgroup', *J Immunol*, 147: 1877-83.
- Sato, Hiroki, Misako Yoneda, Tomoyuki Honda, and Chieko Kai. 2012. 'Morbillivirus receptors and tropism: multiple pathways for infection', *Frontiers in microbiology*, 3: 75-75.
- Satterfield, Benjamin A., Robert W. Cross, Karla A. Fenton, Krystle N. Agans, Christopher F. Basler, Thomas W. Geisbert, and Chad E. Mire. 2015. 'The immunomodulating V and W proteins of Nipah virus determine disease course', *Nature Communications*, 6: 7483-83.
- Savitsky, Pavel, James Bray, Christopher D. O. Cooper, Brian D. Marsden, Pravin Mahajan, Nicola A. Burgess-Brown, and Opher Gileadi. 2010. 'High-throughput production of human proteins for crystallization: The SGC experience', *Journal of Structural Biology*, 172: 3-13.
- Sawatsky, Bevan, and Veronika von Messling. 2010. 'Canine Distemper Viruses Expressing a Hemagglutinin without N-Glycans Lose Virulence but Retain Immunosuppression', *Journal of Virology*, 84: 2753-61.
- Schneidman-Duhovny, D., M. Hammel, and A. Sali. 2010. 'FoXS: a web server for rapid computation and fitting of SAXS profiles', *Nucleic Acids Res*, 38: W540-4.
- Schomacker, Henrick, Peter L. Collins, and Alexander C. Schmidt. 2004. 'In silico identification of a putative new paramyxovirus related to the Henipavirus genus', *Virology*, 330: 178-85.
- Schomacker, Henrick, Anne Schaap-Nutt, Peter L. Collins, and Alexander C. Schmidt. 2012. 'Pathogenesis of acute respiratory illness caused by human parainfluenza viruses', *Current opinion in virology*, 2: 294-99.
- Schountz, Tony, Corey Campbell, Kaitlyn Wagner, Joel Rovnak, Cynthia Martellaro, Blair L. DeBuysscher, Heinz Feldmann, and Joseph Prescott. 2019. 'Differential Innate Immune Responses Elicited by Nipah Virus and Cedar Virus Correlate with Disparate In Vivo Pathogenesis in Hamsters', *Viruses*, 11: 291.
- Segawa, H., A. Inakawa, T. Yamashita, and H. Taira. 2003. 'Functional analysis of individual oligosaccharide chains of Sendai virus hemagglutinin-neuraminidase protein', *Biosci Biotechnol Biochem*, 67: 592-8.
- Seiradake, E., Y. Zhao, W. Lu, A. R. Aricescu, and E. Y. Jones. 2015. 'Production of cell surface and secreted glycoproteins in mammalian cells', *Methods Mol Biol*, 1261: 115-27.
- Seiradake, E.; Zhao, Y.; Lu, W.; Aricescu, A. R.; Jones, E. Y. 2015. 'Production of Cell Surface and Secreted Glycoproteins in Mammalian Cells', *Methods Mol Biol*, 1261: 115-27.
- Shaw, M. L., W. B. Cardenas, D. Zamarin, P. Palese, and C. F. Basler. 2005. 'Nuclear localization of the Nipah virus W protein allows for inhibition of both virus- and toll-like receptor 3-triggered signaling pathways', *J Virol*, 79: 6078-88.
- Sjobring, U., L. Bjorck, and W. Kastern. 1991. 'Streptococcal protein G. Gene structure and protein binding properties', *J Biol Chem*, 266: 399-405.
- Smith, David F., and Richard D. Cummings. 2014. 'Investigating virus-glycan interactions using glycan microarrays', *Current opinion in virology*, 7: 79-87.
- Smith, K. F., M. Goldberg, S. Rosenthal, L. Carlson, J. Chen, C. Chen, and S. Ramachandran. 2014. 'Global rise in human infectious disease outbreaks', *J R Soc Interface*, 11: 20140950.
- Spiro, R. G. 1960. 'Studies on fetuin, a glycoprotein of fetal serum. I. Isolation, chemical composition, and physiochemical properties', *J Biol Chem*, 235: 2860-9.
- Stelfox, Alice J., and Thomas A. Bowden. 2019. 'A structure-based rationale for sialic acid independent host-cell entry of Sosuga virus', *Proceedings of the National Academy of Sciences*: 201906717.
- Stoler, Justin, and Gordon A. Awandare. 2016. 'Febrile illness diagnostics and the malaria-industrial complex: a socio-environmental perspective', *BMC infectious diseases*, 16: 683-83.
- Ströh, Luisa J., and Thilo Stehle. 2014. 'Glycan Engagement by Viruses: Receptor Switches and Specificity', *Annual review of virology*, 1: 285-306.
- Stuart, D. I., M. Levine, H. Muirhead, and D. K. Stammers. 1979. 'Crystal-Structure of Cat Muscle Pyruvate-Kinase at a Resolution of 2.6 Å', *Journal of Molecular Biology*, 134: 109-42.

- Stummeyer, K., A. Dickmanns, M. Muhlenhoff, R. Gerardy-Schahn, and R. Ficner. 2005. 'Crystal structure of the polysialic acid-degrading endosialidase of bacteriophage K1F', *Nat Struct Mol Biol*, 12: 90-6.
- Sutummaporn, Kripitch, Kazuhiko Suzuki, Noboru Machida, Tetsuya Mizutani, Eun-sil Park, Shigeru Morikawa, and Tetsuya Furuya. 2019. 'Association of feline morbillivirus infection with defined pathological changes in cat kidney tissues', *Veterinary Microbiology*, 228: 12-19.
- Tailford, L. E., C. D. Owen, J. Walshaw, E. H. Crost, J. Hardy-Goddard, G. Le Gall, W. M. de Vos, G. L. Taylor, and N. Juge. 2015. 'Discovery of intramolecular trans-sialidases in human gut microbiota suggests novel mechanisms of mucosal adaptation', *Nature Communications*, 6.
- Tan, M., and X. Jiang. 2010. 'Norovirus gastroenteritis, carbohydrate receptors, and animal models', *PLoS Pathog*, 6: e1000983.
- Tanwar, A. S., V. D. Goyal, D. Choudhary, S. Panjekar, and R. Anand. 2013. 'Importance of hydrophobic cavities in allosteric regulation of formylglycinamide synthetase: insight from xenon trapping and statistical coupling analysis', *PLOS ONE*, 8: e77781.
- Tatsuo, H., N. Ono, K. Tanaka, and Y. Yanagi. 2000. 'SLAM (CDw150) is a cellular receptor for measles virus', *Nature*, 406: 893-97.
- Tatsuo, H., N. Ono, and Y. Yanagi. 2001. 'Morbilliviruses use signaling lymphocyte activation molecules (CD150) as cellular receptors', *J Virol*, 75: 5842-50.
- Taylor, Garry L. 2010. 'Introduction to phasing', *Acta crystallographica. Section D, Biological crystallography*, 66: 325-38.
- Terwilliger, T. C., P. D. Adams, R. J. Read, A. J. McCoy, N. W. Moriarty, R. W. Grosse-Kunstleve, P. V. Afonine, P. H. Zwart, and L. W. Hung. 2009. 'Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard', *Acta Crystallogr D Biol Crystallogr*, 65: 582-601.
- Thibault, P. A., R. E. Watkinson, A. Moreira-Soto, J. F. Drexler, and B. Lee. 2017. 'Zoonotic Potential of Emerging Paramyxoviruses: Knowns and Unknowns.' in M. Kielian, T. C. Mettenleiter and M. J. Roossinck (eds.), *Advances in Virus Research, Vol 98*.
- Tidona, C. A., H. W. Kurz, H. R. Gelderblom, and G. Darai. 1999. 'Isolation and molecular characterization of a novel cytopathogenic paramyxovirus from tree shrews', *Virology*, 258: 425-34.
- Tong, Suxiang, Xueyong Zhu, Yan Li, Mang Shi, Jing Zhang, Melissa Bourgeois, Hua Yang, Xianfeng Chen, Sergio Recuenco, Jorge Gomez, Li-Mei Chen, Adam Johnson, Ying Tao, Cyrille Dreyfus, Wenli Yu, Ryan McBride, Paul J. Carney, Amy T. Gilbert, Jessie Chang, Zhu Guo, Charles T. Davis, James C. Paulson, James Stevens, Charles E. Rupprecht, Edward C. Holmes, Ian A. Wilson, and Ruben O. Donis. 2013. 'New world bats harbor diverse influenza A viruses', *Plos Pathogens*, 9: e1003657-e57.
- Uchida, Shotaro, Ryo Horie, Hiroki Sato, Chieko Kai, and Misako Yoneda. 2018. 'Possible role of the Nipah virus V protein in the regulation of the interferon beta induction by interacting with UBX domain-containing protein1', *Scientific Reports*, 8: 7682.
- Van Breedam, W., S. Pohlmann, H. W. Favoreel, R. J. de Groot, and H. J. Nauwynck. 2014. 'Bitter-sweet symphony: glycan-lectin interactions in virus biology', *FEMS Microbiol Rev*, 38: 598-632.
- Van Bresse, Marie-Françoise, Pádraig J. Duignan, Ashley Banyard, Michelle Barbieri, Kathleen M. Colegrove, Sylvain De Guise, Giovanni Di Guardo, Andrew Dobson, Mariano Domingo, Deborah Fauquier, Antonio Fernandez, Tracey Goldstein, Bryan Grenfell, Kátia R. Groch, Frances Gulland, Brenda A. Jensen, Paul D. Jepson, Ailsa Hall, Thijs Kuiken, Sandro Mazzariol, Sinead E. Morris, Ole Nielsen, Juan A. Raga, Teresa K. Rowles, Jeremy Saliki, Eva Sierra, Nahiid Stephens, Brett Stone, Ikuko Tomo, Jianning Wang, Thomas Waltzek, and James F. X. Wellehan. 2014. 'Cetacean morbillivirus: current knowledge and future directions', *Viruses*, 6: 5145-81.
- Vanmechelen, B., M. Bletsa, L. Laenen, A. R. Lopes, V. Vergote, L. Beller, W. Deboutte, M. Korva, T. A. Zupanc, J. G. de Bellocq, S. Gryseels, H. Leirs, P. Lemey, B. Vrancken, and P. Maes. 2018. 'Discovery and genome characterization of three new Jeilongviruses, a lineage of paramyxoviruses characterized by their unique membrane proteins', *Bmc Genomics*, 19.

- Vanmechelen, Bert, Magda Bletsa, Lies Laenen, Ana Rita Lopes, Valentijn Vergote, Leen Beller, Ward Deboutte, Miša Korva, Tatjana Avšič Županc, Joëlle Goüy de Bellocq, Sophie Gryseels, Herwig Leirs, Philippe Lemey, Bram Vrancken, and Piet Maes. 2018. 'Discovery and genome characterization of three new Jeilongviruses, a lineage of paramyxoviruses characterized by their unique membrane proteins', *Bmc Genomics*, 19: 617-17.
- Varki A, Schauer R. Sialic Acids. In: Varki A, Cummings RD, Esko JD, et al., editors. 2009. *Essentials of Glycobiology. Chapter 14*. (Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY)).
- Vavricka, C. J., Y. Liu, H. Kiyota, N. Sriwilaijaroen, J. Qi, K. Tanaka, Y. Wu, Q. Li, Y. Li, J. Yan, Y. Suzuki, and G. F. Gao. 2013a. 'Influenza neuraminidase operates via a nucleophilic mechanism and can be targeted by covalent inhibitors', *Nat Commun*, 4: 1491.
- Vavricka, Christopher J., Yue Liu, Hiromasa Kiyota, Nongluk Sriwilaijaroen, Jianxun Qi, Kosuke Tanaka, Yan Wu, Qing Li, Yan Li, Jinghua Yan, Yasuo Suzuki, and George F. Gao. 2013b. 'Influenza neuraminidase operates via a nucleophilic mechanism and can be targeted by covalent inhibitors', *Nature Communications*, 4: 1491.
- Vizoso, A. D., and W. E. Thomas. 1981. 'Paramyxoviruses of the morbilli group in the wild hedgehog *Erinaceus europeus*', *British journal of experimental pathology*, 62: 79-86.
- Walter, T. S., J. Diprose, J. Brown, M. Pickford, R. J. Owens, D. I. Stuart, and K. Harlos. 2003. 'A procedure for setting up high-throughput nanolitre crystallization experiments. I. Protocol design and validation', *Journal of Applied Crystallography*, 36: 308-14.
- Walter, T. S., J. M. Diprose, C. J. Mayo, C. Siebold, M. G. Pickford, L. Carter, G. C. Sutton, N. S. Berrow, J. Brown, I. M. Berry, G. B. Stewart-Jones, J. M. Grimes, D. K. Stammers, R. M. Esnouf, E. Y. Jones, R. J. Owens, D. I. Stuart, and K. Harlos. 2005. 'A procedure for setting up high-throughput nanolitre crystallization experiments. Crystallization workflow for initial screening, automated storage, imaging and optimization', *Acta Crystallogr D Biol Crystallogr*, 61: 651-7.
- Walter, Thomas S., Erika J. Mancini, Jan Kadlec, Stephen C. Graham, René Assenberg, Jingshan Ren, Sarah Sainsbury, Raymond J. Owens, David I. Stuart, Jonathan M. Grimes, and Karl Harlos. 2008. 'Semi-automated microseeding of nanolitre crystallization experiments', *Acta crystallographica. Section F, Structural biology and crystallization communications*, 64: 14-18.
- Warren, L. 1959. 'The thiobarbituric acid assay of sialic acids', *J Biol Chem*, 234: 1971-5.
- Wasik, Brian R., Karen N. Barnard, and Colin R. Parrish. 2016. 'Effects of Sialic Acid Modifications on Virus Binding and Infection', *Trends in Microbiology*, 24: 991-1001.
- Waterson, A. P., Rudolf Rott, and Gisela Ruckle-Enders. 1963. "The Components of Measles Virus and their Relation to Rinderpest and Distemper." In *Zeitschrift für Naturforschung B*, 377.
- Waugh, David S. 2011. 'An overview of enzymatic reagents for the removal of affinity tags', *Protein expression and purification*, 80: 283-93.
- Weber, G., G. T. DeKoster, N. Holton, K. B. Hall, and M. C. Wahl. 2018. 'Molecular principles underlying dual RNA specificity in the Drosophila SNF protein', *Nat Commun*, 9: 2220.
- Weis, M., L. Behner, M. Hoffmann, N. Kruger, G. Herrler, C. Drosten, J. F. Drexler, E. Dietzel, and A. Maisner. 2014. 'Characterization of African bat henipavirus GH-M74a glycoproteins', *J Gen Virol*, 95: 539-48.
- Welch, B. D., Y. Liu, C. A. Kors, G. P. Leser, T. S. Jardetzky, and R. A. Lamb. 2012. 'Structure of the cleavage-activated prefusion form of the parainfluenza virus 5 fusion protein', *Proc Natl Acad Sci U S A*, 109: 16672-7.
- Welch, B. D., P. Yuan, S. Bose, C. A. Kors, R. A. Lamb, and T. S. Jardetzky. 2013. 'Structure of the parainfluenza virus 5 (PIV5) hemagglutinin-neuraminidase (HN) ectodomain', *PLoS Pathog*, 9: e1003534.
- Welch, Stephen R., Ayan K. Chakrabarti, Lisa Wiggleton Guerrero, Harley M. Jenks, Michael K. Lo, Stuart T. Nichol, Christina F. Spiropoulou, and César G. Albariño. 2018. 'Development of a reverse genetics system for Sosuga virus allows rapid screening of antiviral compounds', *PLOS Neglected Tropical Diseases*, 12: e0006326.
- Wertz, G. W., P. L. Collins, Y. Huang, C. Gruber, S. Levine, and L. A. Ball. 1985. 'Nucleotide sequence of the G protein gene of human respiratory syncytial virus reveals an unusual type

- of viral membrane protein', *Proceedings of the National Academy of Sciences of the United States of America*, 82: 4075-79.
- Wignall-Fleming, Elizabeth B., David J. Hughes, Sreenu Vattipally, Sejal Modha, Steve Goodbourn, Andrew J. Davison, and Richard E. Randall. 2019. 'Analysis of paramyxovirus transcription and replication by high-throughput sequencing', *Journal of Virology*: JVI.00571-19.
- Wilkins, M. R., E. Gasteiger, A. Bairoch, J. C. Sanchez, K. L. Williams, R. D. Appel, and D. F. Hochstrasser. 1999. 'Protein identification and analysis tools in the ExPASy server', *Methods Mol Biol*, 112: 531-52.
- Wilson, D., R. Pethica, Y. Zhou, C. Talbot, C. Vogel, M. Madera, C. Chothia, and J. Gough. 2009. 'SUPERFAMILY--sophisticated comparative genomics, data mining, visualization and phylogeny', *Nucleic Acids Res*, 37: D380-6.
- Wilson, R. L., S. M. Fuentes, P. Wang, E. C. Taddeo, A. Klatt, A. J. Henderson, and B. He. 2006. 'Function of small hydrophobic proteins of paramyxovirus', *J Virol*, 80: 1700-9.
- Winter, G. 2010. 'xia2: an expert system for macromolecular crystallography data reduction', *Journal of Applied Crystallography*, 43: 186-90.
- Wong, K. T., and K. C. Ong. 2011. 'Pathology of acute henipavirus infection in humans and animals', *Patholog Res Int*, 2011: 567248.
- Woo, P. C. Y., S. K. P. Lau, B. H. L. Wong, A. Y. P. Wong, R. W. S. Poon, and K. Y. Yuen. 2011a. 'Complete Genome Sequence of a Novel Paramyxovirus, Tailam Virus, Discovered in Sikkim Rats', *Journal of Virology*, 85: 13473-74.
- Woo, P. C. Y., A. Y. P. Wong, B. H. L. Wong, C. S. F. Lam, R. Y. Y. Fan, S. K. P. Lau, and K. Y. Yuen. 2016a. 'Comparative genome and evolutionary analysis of naturally occurring Beilong virus in brown and black rats', *Infection Genetics and Evolution*, 45: 311-19.
- Woo, Patrick C. Y., Susanna K. P. Lau, Beatrice H. L. Wong, Rachel Y. Y. Fan, Annette Y. P. Wong, Anna J. X. Zhang, Ying Wu, Garnet K. Y. Choi, Kenneth S. M. Li, Janet Hui, Ming Wang, Bo-Jian Zheng, K. H. Chan, and Kwok-Yung Yuen. 2012. 'Feline morbillivirus, a previously undescribed paramyxovirus associated with tubulointerstitial nephritis in domestic cats', *Proceedings of the National Academy of Sciences*, 109: 5435-40.
- Woo, Patrick C. Y., Susanna K. P. Lau, Beatrice H. L. Wong, Annette Y. P. Wong, Rosana W. S. Poon, and Kwok-Yung Yuen. 2011b. 'Complete genome sequence of a novel paramyxovirus, Tailam virus, discovered in Sikkim rats', *Journal of Virology*, 85: 13473-74.
- Woo, Patrick C. Y., Susanna K. P. Lau, Beatrice H. L. Wong, Ying Wu, Carol S. F. Lam, and Kwok-Yung Yuen. 2012. 'Novel variant of Beilong Paramyxovirus in rats, China', *Emerging Infectious Diseases*, 18: 1022-24.
- Woo, Patrick C. Y., Annette Y. P. Wong, Beatrice H. L. Wong, Carol S. F. Lam, Rachel Y. Y. Fan, Susanna K. P. Lau, and Kwok-Yung Yuen. 2016b. 'Comparative genome and evolutionary analysis of naturally occurring Beilong virus in brown and black rats', *Infection, Genetics and Evolution*, 45: 311-19.
- Wu, Ying, Yan Wu, Boris Tefsen, Yi Shi, and George F. Gao. 2014. 'Bat-derived influenza-like viruses H17N10 and H18N11', *Trends in Microbiology*, 22: 183-91.
- Wu, Z. L., X. Huang, A. J. Burton, and K. A. Swift. 2016. 'Probing sialoglycans on fetal bovine fetuin with azido-sugars using glycosyltransferases', *Glycobiology*, 26: 329-34.
- Wu, Z. Q., L. Yang, F. Yang, X. W. Ren, J. Y. Jiang, J. Dong, L. L. Sun, Y. F. Zhu, H. N. Zhou, and Q. Jin. 2014. 'Novel Henipa-like Virus, Mojiang Paramyxovirus, in Rats, China, 2012', *Emerging Infectious Diseases*, 20: 1064-66.
- Xu, K., K. R. Rajashankar, Y. P. Chan, J. P. Himanen, C. C. Broder, and D. B. Nikolov. 2008. 'Host cell recognition by the henipaviruses: Crystal structures of the Nipah G attachment glycoprotein and its complex with ephrin-B3', *Proceedings of the National Academy of Sciences of the United States of America*, 105: 9953-58.
- Xu, Kai, Christopher C. Broder, and Dimitar B. Nikolov. 2012. 'Ephrin-B2 and ephrin-B3 as functional henipavirus receptors', *Seminars in cell & developmental biology*, 23: 116-23.
- Xu, Kai, Yee-Peng Chan, Birgit Bradel-Tretheway, Zeynep Akyol-Ataman, Yongqun Zhu, Somnath Dutta, Lianying Yan, YanRu Feng, Lin-Fa Wang, Georgios Skinotis, Benhur Lee, Z. Hong Zhou, Christopher C. Broder, Hector C. Aguilar, and Dimitar B. Nikolov. 2015. 'Crystal

- Structure of the Pre-fusion Nipah Virus Fusion Glycoprotein Reveals a Novel Hexamer-of-Trimers Assembly', *Plos Pathogens*, 11: e1005322.
- Yamaguchi, Mayu, Yoshinori Kitagawa, Min Zhou, Masae Itoh, and Bin Gotoh. 2014. 'An anti-interferon activity shared by paramyxovirus C proteins: Inhibition of Toll-like receptor 7/9-dependent alpha interferon induction', *Febs Letters*, 588: 28-34.
- Ye, Fei, Fanli Yang, Ruijie Yu, Xi Lin, Jianxun Qi, Zhujun Chen, Yu Cao, Yuquan Wei, George F. Gao, and Guangwen Lu. 2018. 'Molecular basis of binding between the global post-transcriptional regulator CsrA and the T3SS chaperone CesT', *Nature Communications*, 9: 1196-96.
- Yin, H. S., R. G. Paterson, X. Wen, R. A. Lamb, and T. S. Jardetzky. 2005. 'Structure of the uncleaved ectodomain of the paramyxovirus (hPIV3) fusion protein', *Proc Natl Acad Sci U S A*, 102: 9288-93.
- Yin, H. S., X. Wen, R. G. Paterson, R. A. Lamb, and T. S. Jardetzky. 2006. 'Structure of the parainfluenza virus 5 F protein in its metastable, prefusion conformation', *Nature*, 439: 38-44.
- Yogalingam, Gouri, Erik J. Bonten, Diantha van de Vlekkert, Huimin Hu, Simon Moshiah, Samuel A. Connell, and Alessandra d'Azzo. 2008. 'Neuraminidase 1 is a negative regulator of lysosomal exocytosis', *Developmental cell*, 15: 74-86.
- Yoneda, Misako, Vanessa Guillaume, Hiroki Sato, Kentaro Fujita, Marie-Claude Georges-Courbot, Fusako Ikeda, Mio Omi, Yuri Muto-Terao, T. Fabian Wild, and Chieko Kai. 2010. 'The Nonstructural Proteins of Nipah Virus Play a Key Role in Pathogenicity in Experimentally Infected Animals', *PLOS ONE*, 5: e12709.
- Yuan, P., G. P. Leser, B. Demeler, R. A. Lamb, and T. S. Jardetzky. 2008. 'Domain architecture and oligomerization properties of the paramyxovirus PIV 5 hemagglutinin-neuraminidase (HN) protein', *Virology*, 378: 282-91.
- Yuan, P., R. G. Paterson, G. P. Leser, R. A. Lamb, and T. S. Jardetzky. 2012. 'Structure of the Ulster Strain Newcastle Disease Virus Hemagglutinin-Neuraminidase Reveals Auto-Inhibitory Interactions Associated with Low Virulence', *Plos Pathogens*, 8.
- Yuan, P., K. A. Swanson, G. P. Leser, R. G. Paterson, R. A. Lamb, and T. S. Jardetzky. 2011. 'Structure of the Newcastle disease virus hemagglutinin-neuraminidase (HN) ectodomain reveals a four-helix bundle stalk', *Proc Natl Acad Sci U S A*, 108: 14920-5.
- Yuan, P., T. B. Thompson, B. A. Wurzburg, R. G. Paterson, R. A. Lamb, and T. S. Jardetzky. 2005. 'Structural studies of the parainfluenza virus 5 hemagglutinin-neuraminidase tetramer in complex with its receptor, sialyllactose', *Structure*, 13: 803-15.
- Zaitsev, Viatcheslav, Mark von Itzstein, Darrin Groves, Milton Kiefel, Toru Takimoto, Allen Portner, and Garry Taylor. 2004. 'Second Sialic Acid Binding Site in Newcastle Disease Virus Hemagglutinin-Neuraminidase: Implications for Fusion', *Journal of Virology*, 78: 3733-41.
- Zanin, M., P. Baviskar, R. Webster, and R. Webby. 2016. 'The Interaction between Respiratory Pathogens and Mucus', *Cell Host Microbe*, 19: 159-68.
- Zeltina, A., T. A. Bowden, and B. Lee. 2016. 'Emerging Paramyxoviruses: Receptor Tropism and Zoonotic Potential', *PLoS Pathog*, 12: e1005390.
- Zhang, X., G. Lu, J. Qi, Y. Li, Y. He, X. Xu, J. Shi, C. W. Zhang, J. Yan, and G. F. Gao. 2013. 'Structure of measles virus hemagglutinin bound to its epithelial receptor nectin-4', *Nat Struct Mol Biol*, 20: 67-72.
- Zhao, Baoguang, Angela Smallwood, Jingsong Yang, Kristin Koretke, Kelvin Nurse, Amy Calamari, Robert B. Kirkpatrick, and Zhihong Lai. 2008. 'Modulation of kinase-inhibitor interactions by auxiliary protein binding: crystallography studies on Aurora A interactions with VX-680 and with TPX2', *Protein science : a publication of the Protein Society*, 17: 1791-97.
- Zhu, Xueyong, Hua Yang, Zhu Guo, Wenli Yu, Paul J. Carney, Yan Li, Li-Mei Chen, James C. Paulson, Ruben O. Donis, Suxiang Tong, James Stevens, and Ian A. Wilson. 2012. 'Crystal structures of two subtype N10 neuraminidase-like proteins from bat influenza A viruses reveal a diverged putative active site', *Proceedings of the National Academy of Sciences of the United States of America*, 109: 18903-08.

APPENDIX C: LIST OF ABBREVIATIONS

°C	Degrees Celcius	DMSO	Dimethyl sulphoxide
Å	Ångströms = 10 ⁻¹⁰ metres	DNA	Deoxyribonucleic Acid
ACE2	Angiotensin-converting enzyme 2	DPP4	Didpeptidyl peptidase 4
AchPV	Achimota virus	DRC	Democratic Republic of the Congo
APBS	Adaptive Poisson-Boltzmann server	E. coli	Escherichia coli
APMV	Avian paramyxovirus	EboV	Ebola virus
APV	Antarctic penguin virus	ECL	Electrochemiluminescence
AsaPV	Atlantic salmon paramyxovirus	EDTA	Ethylenediaminetetraacetic acid
AU	Absorbance units	EFN	Ephrin
BavV	Bank vole virus	EGFR	Epidermal growth factor receptor
BeiV	Beilong paramyxovirus	ELISA	Enzyme-linked immunosorbent-assay
BHK21	Baby hamster kidney	em	Emission
BirA	E. coli biotin ligase	EndoF₁	Endoglycosidase-F ₁
BMV	Bat mumps virus	ex	Excitation
BSA	Bovine serum albumin	F	Fusion protein (Paramyxovirus)
BSR-T7	BHK21 cells stably expressing T7 RNA polymerase	F	Fusion protein
CaCl₂	Calcium Chloride	FACS	Fluorescence activated cell sorting
CDC	Center for disease control	Fc	Fragment crystallisable
CDV	Canine distemper virus	FCS	Fetal Calf Serum
CDXX	Cluster of differentiation XX	FDLV	Fer-de-lance virus
CedV	Cedar paramyxovirus	FeMV	Feline morbillivirus
CeMV	Cetcean morbillivirus	FoXS	Fast X-ray scattering
CFG	Consortium for Functional Glycomics	FP	Fusion peptide
CMAH	Cytidine monophosphate-N-acetylneuraminic acid hydroxylase	FSA	Free sialic acid
CMV	Cytomeglovirus	g	Gram
CO₂	Carbon dioxide	g	Gravity
Cryo-EM	Cryogenic electron-microscopy	GAG	Glycosaminoglycans
d	day	GhV	Ghanaian henipavirus (aka Kumasi virus)
Da	Daltons	GlcNAc	N-acetyl glucosamine
DANA	N-Acetyl-2,3-dehydro-2-deoxyneuraminic acid	GMEM	Glasgow's minimum essential medium
DDM	n-dodecyl-β-D-maltopyranoside	GOF	Gain-of-function
DLS	Diamond light source	GOI	Gene of interest

DMEM	Dulbecco's Modified Eagle's Medium	GP	Glycoprotein
h	Hour	m	Meter
H	Hemagglutinin receptor binding protein	MALS	Multi-angle light scattering
HAP1	Near haploid cell line	Man	Mannose
HBGA	Histo-blood group antigen	MapV	Mapuera virus
HCl	Hydrochloric acid	MCS	Multiple cloning site
HEK293T	Human Embryonic Kidney 293T cells	MenV	Menagle virus
HeV	Hendra virus	MERS-CoV	Middle East respiratory syndrome virus
His	Histidine	MeV	Measles virus
HIV	Human immunodeficiency virus	MgCl₂	Magnesium chloride
hMPV	human metapneumovirus	Min	Minutes
HN	Hemagglutinin-neuraminidase receptor binding protein	MME	Methyl ether
HNV	Henipavirus	MMLV	Mount Mabu lophuromys virus
HPLC	High preformance liquid chromatography	MojV	Mójiāng virus
HR	Heptad repeat	MosV	Mossman paramyxovirus
HRP	Horse radish peroxidase	MU	Methylumbelliferyl
HRV	Human rhinovirus	MuV	Mumps virus
ICTV	International Committee on Taxonomy of Viruses	MWCO	Molecular Weight Cut-Off
IMDM	Iscove's modified dulbecco's medium	N	Nucleoprotein
IV	Intraviral	n	Nano (10 ⁻⁹)
JPV	J paramyxovirus	n.c.s	Non-crystallographic symmetry
k	Kilo (10 ³)	NaCl	Sodium Chloride
K	Kelvin	NarV	Nariva paramyxovirus
K₂PtCl₄	Potassium tetrachloroplatinate	NDV	Newcastle disease virus
Kb	Kilobase	NDV	Newcastle disease virus
KBM-7	Human chronic myeloid leukemia cell line	NEAA	Non-Essential Amino Acids
L	Litre	Neu5Ac	5-N-acetylneuraminic acid
L	RNA-dependent RNA-polymerase	Neu5Gc	5-N-glycolylneuraminic acid
l	Wavelength	NiV	Nipah virus
LAMP-1	lysosomal-associated membrane protein-1	NPC1	Neimann-Pick C1
LASV	Lassa virus	ORF	Open reading frame
LB	Luria Broth	P	Phosphoprotein
LC/MS/MS	Liquid chromatography with tandem mass spectrometry	P/T/S	Proline, threonine, serine
LLG	Log likelihood gain	PBS	Phosphate buffered saline
LPMV	La Piedad Michoacán Mexico virus	PCR	Polymerase Chain Reaction

LTR	Long terminal repeat	SosV	Sosuga paramyxovirus
m	Milli (10^{-3})	SP	Signal peptide
M	Molar concentration	SUMO	Small ubiquitin-like modifier
M	Matrix protein	PCT	Pre-crystallisation test
PEI	Polyethylenimine	PDB	Protein Data Bank
PGA	Poly-g-glutamic acid	SV41	Simian virus 41
PISA	Proteins, Interfaces, Structures and Assemblies	TAE	Tris Acetic acid EDTA
PIV	Parainfluenzavirus	PDV	Phocine distemper virus
PIV	Parainfluenzavirus	TBS	10 mM Tris pH 8, 150 mM NaCl
PMPV	Pohorje Myodes paramyxovirus	TevPV	Teviot paramyxovirus
POI	Protein of interest	ThkPV	Tuhoko paramyxovirus
PPRV	Peste-des-petits-ruminants virus	TioV	Tioman virus
PPT	Polypurine tract	PEG	Polyethylene glycol
RBCs	Red blood cells	TlmV	Tailam virus
RBP	Receptor binding protein	TLS	Translation/Libration/Screw
RMSD	Root-mean-square deviation	TM	Transmembrane
RNA	Ribonucleic acid	Tm	Melting temperature (of DNA duplexes)
RNP	Ribonucleoprotein	TNF	Tumour necrosis factor
Rpm	Revolutions per minute	Tris	Tris(hydroxymethyl)aminomethane
RRE	Rev response element	TupV	Tupaia paramyxovirus
RSV	Respiratory syncytial virus	V	Volts
RV	Rinderpest virus	VLP	Virus-like-particle
SAD	Single-wavelength anomalous diffraction	VSV	Vesicular stomatitis virus
SalV	Salem virus	WHO	World health organisation
SARS-CoV	Severe acute respiratory syndrome-related coronavirus	WHPV	Wēnlíng hoplichthys paramyxovirus
SAXS	Small-angle X-ray scattering	WPSSP V	Wēnzhōu pacific spadenose shark paramyxovirus
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis	WTLPV	Wēnlíng triplecross lizardfish paramyxovirus
Sec	Seconds	WTSPV	Wēnlíng tonguesole paramyxovirus
SEC	Size Exclusion Chromatography	y	psi
SeV	Sendai virus	μ	Micro (10^{-6})
SH	Small hydrophobic	χ²	Chi-squared
SHP	Structural homology program		
SLAM	signalling lymphocyte activation molecule		
SOC	Super Optimal Broth		
SOE	Splice by overlap		

APPENDIX D: LIST OF GENBANK ACCESSION NUMBERS

Virus	Gene(s) used	Accession code	Chapter(s)	Section(s)
AchPV1	L	YP_009094458.1	1	1.3
AchPV1	RBP	YP_009094457.1	1, 3	1.3, 3.2
AchPV2	L	YP_009094466.1	1	1.3
AchPV2	RBP	YP_009094465.1	1	1.3
APMV2	L	BAW94627.1	1	1.3
APMV2	RBP	AAM61887.1	1	1.3
APMV2	Genome	LC187305.1	1	1.4
APMV2	RBP	AQQ11616.1	1, 3, 4, 5	1.5, 3.2. 4.2, 5.2
APMV3	L	YP_009094217.1	1	1.3
APMV3	RBP	YP_009094216.1	1	1.3
APMV3	Genome	EU782025.1	1	1.4
APMV3	RBP	AWU68199.1	1, 3, 4, 5	1.5, 3.2. 4.2, 5.2
APMV4	L	AQZ26706.1	1	1.3
APMV4	RBP	AQZ26705.1	1	1.3
APMV5	L	BAV25267.1	1	1.3
APMV5	RBP	BAV25266.1	1	1.3
APMV6	L	BAM64955.1	1	1.3
APMV6	RBP	BAM64954.1	1	1.3
APMV7	L	YP_009094060.1	1	1.3
APMV7	RBP	YP_009094059.1	1	1.3
APMV8	L	AFX97643.1	1	1.3
APMV8	RBP	AFX97642.1	1	1.3
APMV9	L	YP_009094364.1	1	1.3
APMV9	RBP	YP_009094363.1	1	1.3
APMV10	L	AQQ11616.1	1	1.3
APMV10	RBP	AQQ11615.1	1	1.3
APMV11	L	YP_009094483.1	1	1.3
APMV11	RBP	YP_009094482.1	1	1.3
APMV12	L	YP_009094173.1	1	1.3
APMV12	RBP	YP_009094172.1	1	1.3
APMV13	L	AMR70815.1	1	1.3
APMV13	RBP	AMR70814.1	1	1.3
APMV14	L	YP_009508498.1	1	1.3
APMV14	RBP	YP_009508497.1	1	1.3
APMV15	L	YP_009380509.1	1	1.3
APMV15	RBP	YP_009380508.1	1	1.3
APMV16	L	AWU68199.1	1	1.3
APMV16	RBP	AWU68198.1	1	1.3
APMV20	L	YP_009553495.1	1	1.3

APMV20	RBP	YP_009553494.1	1	1.3
APV-A	L	YP_009508510.1	1	1.3
APV-A	RBP	ARU83010.1	1	1.3
APV-B	L	ARU83017.1	1	1.3
APV-B	RBP	ARU83016.1	1	1.3
AsaPV	L	YP_009094152.1	1	1.3
AsaPV	RBP	YP_009094151.1	1	1.3
BaVV	L	ATW63190.1	1	1.3
BaVV	RBP	ATW63189.1	1, 4	1.3. 4.2
BeiV	L	YP_512254.1	1	1.3
BeiV(HMC)	RBP	YP_512253.1	1	1.3
BeiV(HMC)	Genome	DQ100461.1	1	1.4
BeiV(HMC)	RBP	NC_007454.1	5	5.2, 5.3
BeiV(rat)	RBP	AOV81769.1	1	1.3
BeiV(rat)	Genome	KX940961.1	1	1.4
BeiV(rat)	RBP	NC_007803.1	5	5.2, 5.3
BPIV3	L	NP_037646.1	1	1.3
BPIV3	RBP	NP_037645.1	1	1.3
CDV	L	AAK54670.1	1	1.3
CDV	RBP	AAK54669.1	1	1.3
CedV	L	YP_009094087.1	1	1.3
CedV	RBP	YP_009094086.1	1	1.3
CeMV	L	AYR16899.1	1	1.3
CeMV	RBP	AYR16898.1	1	1.3
CPIV3	L	QBF53203.1	1	1.3
CPIV3	RBP	QBF53202.1	1	1.3
FDLV	L	NP_899661.1	1	1.3
FDLV	RBP	NP_899660.1	1	1.3
FeMV	L	YP_009512964.1	1	1.3
FeMV	RBP	YP_009512963.1	1	1.3
GhV	L	YP_009091839.1	1	1.3
GhV	RBP	YP_009091838.1	1	1.3
HeV	L	NP_047113.3	1	1.3
HeV	RBP	NP_047112.2	1	1.3
HeV	Genome	AF017149.3	1	1.4
HPIV1	L	AAL89409.1	1	1.3
HPIV1	RBP	AAL89408.1	1	1.3
HPIV2	L	AZS52544.1	1	1.3
HPIV2	RBP	AZS52543.1	1	1.3
HPIV3	L	AZS52615.1	1	1.3
HPIV3	RBP	AZS52614.1	1	1.3
HPIV4a	L	YP_008378665.1	1	1.3
HPIV4a	RBP	YP_008378664.1	1	1.3
HPIV4b	L	AFB82778.1	1	1.3
HPIV4b	RBP	AFB82777.1	1	1.3

JPV	L	YP_338085.1	1	1.3
JPV	RBP	YP_338084.1	1	1.3
JPV	RBP	AOV81769.1	5	5.2, 5.3
LPMV	L	YP_001331035.1	1	1.3
LPMV	RBP	YP_001331034.1	1	1.3
MapV	L	YP_001249278.1	1	1.3
MapV	RBP	YP_001249277.1	1	1.3
MenV	L	YP_009512971.1	1	1.3
MenV	RBP	AAK62284.1	1	1.3
MenV	Genome	AF326114.2	1	1.4
MenV	RBP	AAK62284.1	3	3.2
MeV	L	NP_056924.1	1	1.3
MeV	RBP	NP_056923.1	1	1.3
MeV	Genome	AY486083.1	1	1.4
MMLV1	L	AVM86019.1	1	1.3
MMLV1	RBP	AVM86018.1	1	1.3
MMLV1	RBP	YP_009508497.1	5	5.2
MMLV2	L	AVM86027.1	1	1.3
MMLV2	RBP	AVM86026.1	1	1.3
MMLV2	RBP	AWU68198.1	5	5.2
MojV	L	YP_009094096.1	1	1.3
MojV	RBP	YP_009094095.1	1	1.3
MosV	L	NP_958055.1	1	1.3
MosV	RBP	NP_958054.1	1, 4	1.3, 4.2
MosV	RBP	AY286409.1	4	4.3
MuV	L	AAL83746.1	1	1.3
MuV	RBP	AAL83745.1	1	1.3
MuV	Genome/RBP	MG460606.1	1, 3, 4, 5	1.4, 1.5, 3.2, 4.2, 5.2
NarV	L	YP_006347589.1	1	1.3
NarV	RBP	YP_006347588.1	1, 4	1.3, 4.2
NarV	Genome/RBP	FJ362497.2	1, 4	1.4, 4.3
NDV	L	AFV96177.1	1	1.3
NDV	RBP	AFV96176.1	1	1.3
NDV	Genome	AB524405.1	1	1.4
NDV	RBP	YP_009512963.1	1, 3, 4, 5	1.5, 3.2, 4.3, 5.2
NiV	L	NP_112028.1	1	1.3
NiV	RBP	NP_112027.1	1	1.3
PDV	L	YP_009177604.1	1	1.3
PDV	RBP	BAA01207.1	1	1.3
PIV5	L	AWA45301.1	1	1.3
PIV5	RBP	AGJ98318.1	1	1.3
PMPV1	L	AVM86048.1	1	1.3
PMPV1	RBP	AVM86047.1	1	1.3
PPIV1	L	AVO62728.1	1	1.3
PPIV1	RBP	AVO62727.1	1	1.3

PPRV	L	CAJ01701.1	1	1.3
PPRV	RBP	CAJ01700.1	1	1.3
RPV	L	YP_087126.2	1	1.3
RPV	RBP	YP_087125.2	1	1.3
SalV	L	YP_009094339.1	1	1.3
SalV	RBP	YP_009094338.1	1	1.3
SeV	L	AAB06283.1	1	1.3
SeV	RBP	AAB06282.1	1	1.3
SeV	Genome/RBP	NC_001552.1	1, 3, 4, 5	1.4, 1.5, 3.2, 4.2, 5.2
SosV	L	YP_009094034.1	1	1.3
SosV	RBP	YP_009094033.1	1, 3	1.3, 3.2
SosV	RBP	NC_025343	3	3.3
SV41	L	YP_138510.1	1	1.3
SV41	RBP	YP_138509.1	1	1.3
TevPV	L	AZK16201.1	1	1.3
TevPV	RBP	AZK16200.1	1	1.3
ThkPV1	L	ADI80715.1	1	1.3
ThkPV1	RBP	ADI80714.1	1	1.3
ThkPV2	L	ADI80722.1	1	1.3
ThkPV2	RBP	YP_009094066.1	1	1.3
ThkPV3	L	YP_009094080.	1	1.3
ThkPV3	RBP	ADI80728.1	1	1.3
ThkPV3	RBP	YP_009094079.1	3	3.2
TioV	L	NP_665871.1	1	1.3
TioV	RBP	NP_665870.1	1	1.3
TlmV	L	YP_009094120.1	1	1.3
TlmV	RBP	YP_009094119.1	1	1.3
TlmV	RBP	AVM87356.1	5	5.2
TupV	L	NP_054697.1	1	1.3
TupV	RBP	NP_054696.1	1	1.3
WHPV	L	AVM87400.1	1	1.3
WPSSPV	L	AVM87362.1	1	1.3
WPSSPV	RBP	AVM87356.1	1	1.3
WTLPV	L	AVM87391.1	1	1.3
WTSPV	RBP	AVM87377.1	1	1.3
WTSPV	L	AVM87378.1	1	1.3