University of Nottingham
UK | CHINA | MALAYSIA

Investigating lyssavirus glycoprotein interactions

Joseph Elliot Hinds,

BSc, MSc

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

November 2023

Abstract

The *Lyssavirus* genus contains the viruses responsible for the disease rabies, an acute progressive encephalomyelitis which is fatal in over 99% of human cases without treatment prior to the onset of symptoms. The genus is made up of three distinct phylogroups representing genetic distance and antigenic properties. The primary causative virus of rabies is the rabies virus which is involved in over 99% of human infections. Despite its highly pathogenic nature and the fact it causes over 59,000 deaths annually, particularly in developing nations, it is considered to be a neglected pathogen. While the human burden of lyssavirus infection is great, there are available treatments both pre and post infection. However, current treatments suffer from high costs, inconvenient and extensive treatment regimens, cold chain necessity and lack of inter-phylogroup neutralisation. These issues have generated interest in a novel vaccine.

For lyssaviruses, there is only one surface protein, the trimeric spike glycoprotein. This protein is involved in many important viral processes such as host receptor attachment and cell entry. It is because of this importance and the many remaining questions surrounding lyssaviruses that this study has focused on the glycoprotein and its interactions.

In order to study the glycoprotein in isolation a three plasmid pseudo-virus system was used. This system can be used to express a viral glycoprotein of interest on a vector such as a lentivirus in order to produce a non-replicative and safe pseudo-virus. This was used to infect immortalised cell lines such as BHK-21 to examine changes in outcome from both modifications of the glycoprotein and neutralisation using non-human primate sera or monoclonal antibodies. Additionally, this system was used to infect primary NK cells to determine their susceptibility to rabies virus infection.

Results of neutralisation assays performed by this study demonstrate that a novel medoid vaccine presented cross-phylogroup neutralising properties beyond any currently available vaccine, indicating a potential solution for divergent lyssavirus infection concerns. The vaccine was able to elicit significant neutralisation against MOKV and WCBV by week 5, and IKOV following a boost at week 50.

Glycosylation of the lyssavirus glycoprotein was also investigated and shown to be important in the infection capability and neutralisation susceptibility of the protein. The removal of fixed virus glycan Asn204 was found to significantly reduce infectivity and increase neutralisation susceptibility whereas removal of Asn37 increased infection though had no impact on neutralisation. The strongest change was in their combined removal. Asn319 removal also completely removed any infectivity. Common amino acids used in substitutions to knock out glycans aspartic acid (D), glutamine (Q) and alanine (A) were also examined for their impact

and it was found that differing amino acids produced significantly different results. Finally, NK cell infection was found to not occur, however further work needs to be done to confirm this finding.

Research into rabies is both important and inadequate with many important questions remaining unanswered. The lyssavirus glycoprotein and its interactions are essential in understanding lyssaviruses and developing novel therapies for them. This study aimed to further elucidate some of the questions surrounding the glycoprotein. It also provides a foundation for more exciting discoveries and findings.

Acknowledgements

Firstly I would like to thank and acknowledge Professor Jonathan Ball for offering me a place in his lab and supporting my work especially during a very busy and difficult pandemic. I would also like to thank my other supervisors Dr Patrick McClure and Dr Richard Urbanowicz for their help, support and humour.

Additionally, I want to hugely thank the labs post-doctoral support team Joshua Duncan, Joe Chappel and Theo Xaris for their wisdom, kindness and laughs. In particular I want to offer my deepest thanks and appreciations to Josh, as without his guidance, support, and patience answering thousands of my questions, I would have greatly struggled.

Outside of the lab I want to thank my family, Gaynor, Steve, Gary and Michael; and my friends Jack, Emma, Josh and Fin, all of whom have had my back throughout this process.

Finally and most importantly I want to thank my wife Abbie Hinds without whom this process would have been impossible. Her support has been invaluable and her patience unending. So, thank you again Abbie, I owe it all to you!

Table of contents

List of tables

List of figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 1: Introduction

1.1 Introduction

The *Lyssavirus* genus (family *Rhabdoviridae* order *Mononegavirales*) contains bullet-shaped, single stranded, negative sense-RNA viruses which are the causative pathogens for the zoonotic disease known as rabies, an acute and progressive encephalomyelitis. Lyssaviruses are widespread, being present on every continent with the exception of Antarctica and causing infections in over 150 countries [WHO]. Rabies is highly virulent, with a >99% fatality rate in humans once the onset of symptoms has occurred. Infection with any lyssavirus causes rabies, however 99% of cases are caused by the rabies virus (RABV). RABV, despite an effective vaccine, results in an estimated 59,000 human deaths annually, however this figure is likely lower than the true burden [WHO; Hampson *et al*., 2015]. These factors make lyssaviruses an important pathogen to study, yet they are considered a neglected pathogen [Fooks *et al*., 2017].

1.2 Viral classification and phylogrouping

The *Lyssavirus* genus currently contains 17 viral species that are officially recognized by the International Committee on Taxonomy of Viruses (ICTV) [ICTV, 2019] all of which are listed in table 1.1, however, this is likely to change rapidly. Over the past 20 years the genus has expanded from 8 known viruses to 17 [Banyard *et al*., 2014]. This is in part due to an increased interest in virus discovery as well as improved surveillance.

Table 1.1: List of all officially recognised lyssaviruses as stated by the ICTV. Examples sequences of each lyssavirus are provided which were sourced from the ICTV website [ICTV, 2019].

Phylogroup	Species name	Virus name	Abbreviation	Example sequence
1	*Lyssavirus aravan*	Aravan virus	ARAV	EF614259
	Lyssavirus australis	Australian bat lyssavirus	ABLV	AF081020
	Lyssavirus bokeloh	Bokeloh bat lyssavirus	BBLV	JF311903
	Lyssavirus Duvenhage	Duvenhage virus	DUVV	EU293119
	Lyssavirus Formosa	Taiwan bat lyssavirus	TWBLV	MF472710
	Lyssavirus gannoruwa	Gannoruwa bat lyssavirus	GBLV	KU244266
	Lyssavirus hamburg	European bat lyssavirus 1	EBLV-1	EF157976
	Lyssavirus helsinki	European bat lyssavirus 2	EBLV-2	EF157977
	Lyssavirus irkut	Irkut virus	IRKV	EF614260
	Lyssavirus Khujand	Khujand virus	KHUV	EF614261
	Lyssavirus rabies	Rabies virus	RABV	M13215
2	*Lyssavirus mokola*	Mokola virus	MOKV	Y09762
	Lyssavirus shimoni	Shimoni bat virus	SHIBV	GU170201
	Lyssavirus lagos	Lagos bat virus	LBV	EU293108
3	*Lyssavirus caucasicus*	West Caucasian bat virus	WCBV	EF614258
	Lyssavirus ikoma	Ikoma lyssavirus	IKOV	JX193798
	Lyssavirus Lleida	Lleida bat lyssavirus	LLEBV	KY006983

Lyssaviruses are grouped into phylogroups which are determined according to genetic distances of the nucleoprotein (N) gene, as well as whole genome phylogenetics and antigenic data [Badrane *et al*., 2001]. Through these methods there are at least two identifiable phylogroups, phylogroup I and II, with a third tentative group consisting of highly divergent lyssaviruses. The sequence identities of phylogroup II glycoproteins offer the highest similarity at 71.5%, with phylogroup I glycoproteins having a 70.3% sequence similarity. The divergent phylogroup III however, exhibits 58.2% glycoprotein sequence identity. The phylogroup III member IKOV shares only 61-62.% sequence identity with other lyssaviruses when measuring whole genome-sequences [Marston *et al*., 2012a]. This genetic distance and phylogroup distinction is important due to the difference in antigenicity. Traditionally, it has been considered that lyssavirus neutralisation with currently available vaccines is only intra, not inter-phylogroup, meaning vaccines against a phylogroup I lyssavirus will only be capable of neutralising other phylogroup I lyssaviruses [Badrane *et al*., 2001]. However, a recent study has shown limited cross-neutralisation of phylogroup II by RABV based vaccines, and that broadly neutralising antibodies which provide cross-phylogroup neutralisation are possible, offering a basis for the next-generation of vaccines for rabies [Inoue *et al*., 2024]. There has also been detection of cross-neutralisation by RABV vaccination against the phylogroup II virus MOKV, however the cause of this was unknown [Malerczyk *et al*., 2014]

While classification has most often, until recently, been performed using serological, immunological and epidemiological information, it is now more common to use sequence data to classify lyssaviruses. Currently, a sequence nucleotide identity value of 80-82% or above of the N protein is required in order to be classified as the same species [Kuzmin *et al*., 2005], this is also used by the ICTV who cite a 78-80% cut off threshold for whole N protein sequences [ICTV, 2019].

There are currently two other lyssaviruses awaiting classification, Kotalahti bat lyssavirus (KBLV) [Nokireki *et al*., 2018], and Matlo bat lyssavirus (MBLV) [Coertse et al 2020, Grobler *et al*., 2021]. The glycoprotein of KBLV has since been challenged with anti-RABV sera and antibodies which was capable of neutralizing the virus, indicating it is likely a phylogroup I lyssavirus, this is also confirmed through phylogenetics indicating a close relation to other European bat lyssaviruses such as EBLV-2 [Shipley *et al*., 2021].

1.3 Lyssavirus epidemiology

1.3.1 Distribution and primary reservoirs of lyssaviruses

The epidemiology of lyssaviruses is complex due to a number of factors including their widespread nature, often difficult to track reservoir species such as bats which have a multitude of geographic boundaries and transmission dynamics, and the currently estimated capability to infect any mammal species. Lyssaviruses can also undergo host-shift events, inter-reservoir infections and spill-over events, complicating surveillance further. Lyssaviruses can be found on every continent with the exception of Antarctica. The currently known distribution and reservoir for each lyssavirus can be found in table 1.2 [Fooks & Jackson, 2020].

Table 1.2: Distribution and reported reservoir species of all known lyssaviruses in the genus. Recreation of the table from Fooks & Jackson, 2020. *Reservoir host not confirmed. Any new information beyond the tables publication has not been included.

Species	Abbreviation	Distribution	Reservoir
Rabies lyssavirus	RABV	Worldwide with the exception of Antarctica and some islands.	*Carnivora*: many species of domestic and wild canids, foxes, mongooses, skunks and the racoon. *Chiroptera*: (Americas only) several species of insectivorous, vampire and frugivorous bats.
Aravan lyssavirus	ARAV	Only single case reported from Kyrgyzstan	Microchiroptera: Myotis blythi*
Australian bat lyssavirus	ABLV	Australia and possible areas of SE Asia	*Megachiroptera:* Pteropid species *Microchiroptera:* the insectivorous yellow-bellied sheathtail bat (*Saccolaimus flavicentris*)
Bokeloh bat lyssavirus	BBLV	Several cases recovered from Germany, France and Poland	Natterer's bat (*Myotis nattererii*)
Duvenhage virus	DUVV	African nations, including Guinea, South Africa and Zimbabwe	Microchiroptera: Single cases assigned to Miniopterus schreibersii, Nycteris gambiensis, and N. thebaica
European bat lyssavirus 1	EBLV-1	Europe, including Denmark, France, Germany, The Netherlands, Poland, Russia, Spain and Ukraine	Microchiroptera: Eptesicus serotinus
European bat lyssavirus 2	EBLV-2	Several countries in western Europe particularly The Netherlands, Switzerland, Finland and the UK	Microchiroptera: Myotis species, especially M. dasycneme and M. daubentonii

Gannoruwa bat lyssavirus	GBLV	Four cases recovered from flying foxes in Sri Lanka	Pteropus sp.
Ikoma virus	IKOV	Single case in Tanzania	Isolated from an African civet (*Civettictis civetta*) but bat reservoir is possible*
Irkut virus	IRKV	One case from Irkutsk province, Russia.	Microchiroptera: Murina leucogaster*
Khujand virus	KHUV	Single case from Tajikistan	Microchiroptera: single case from a bat (Myotis daubentonii)*
Lagos bat virus	LBV	Several African countries including Central African Republic, Ethiopia, Nigeria, Senegal, South Africa and Zimbabwe	Megachiroptera: Fruit bats, including Eidolon helvum, Rousettus aegyptiacus and Epomophorus wahlbergi
Lleida bat lyssavirus	LLEBV	Two cases, one in Spain and one in France	Isolated from Miniopterus schreibersii*
Mokola virus	MOKV	Several African countries including Cameroon, Central African Republic, Ethiopia, Nigeria, South Africa, Zimbabwe	Reservoir unknown. Single cases reported in shrews (*Crocidura sp.*) and a small rodent (*Lopyhromys sikapusi*) with most reported cases in domestic cats and dogs
Shimoni bat virus	SHIBV	Single case from a bat in Kenya	Commerson's leaf-nosed bat (*Hipposideros commersoni*)*
West Caucasian bat virus	WCBV	Single case from Krasnodar region of Russia, single case from Italy	Isolated from *Miniopteris schreibersii**and the same species is suspected in the Italian case.
Taiwan bat lyssavirus	TWBLV	Two cases in Taiwan	Isolated from Pipistrellus abramus

As can be seen from the table, phylogroup I lyssaviruses are found in several regions including Eurasia, Europe, Australia and Africa. Phylogroup II lyssaviruses have only been found to be of African origin at present. The reason for this divergent group in Africa has not yet been determined [Badrane *et al*., 2001]. In Europe, the primary reservoirs of RABV are the red fox (*Vulpes vulpes*) and raccoon dogs (*Nyctereutes procyonoides*) [WHO rabies bulletin Europe 2024]. Raccoon dogs were artificially introduced through importation to Western Russia in the 1920s where they were farmed for fur. This resulted in the raccoon dog population spreading across much of eastern Europe and northwards, reaching as far north as Finland. Their presence in Europe is important for RABVs sustained existence in the colder regions of the continent. Raccoon dogs hibernate in the winter and can incubate RABV without succumbing to infection, allowing the virus to survive harsh winters where it would otherwise be diminished significantly due to a loss of host availability following the reduction of red fox numbers [Macdonald, 1984;

Sutor *et al*., 2014]. It is important to understand these conditions in which RABV survives as lyssavirus success appears to be very dependent on ecological factors. While lyssaviruses are present in Europe, much of the continent is considered rabies-free including the UK, with the only recent cases of rabies being imported to the nation [Finnegan *et al*., 2002; Public Health England, 2019]. Rabies-free status does not mean that lyssaviruses are not present within a nation and in fact infections can occur within a rabies-free country [Leopardi *et al*., 2021]. This apparent contradiction is due to there being multiple definitions of rabies-free status. The World Organization for Animal Health (WOAH) defines a country as rabies free if it meets the criteria listed in its terrestrial animal health code chapter 8 article 14.3. The criteria includes: rabies must be a notifiable disease across the whole country, effective surveillance is in place and has been for 24 months, prevention and control measures are in place such as vaccination programmes and animal importation restrictions, no case of indigenously acquired rabies virus has occurred for the last 24 months, no imported cases of rabies infection from the Orders Carnivora or Chiroptera have been confirmed in the last 6 months. Imported human cases of rabies do not affect the rabies free status. This is important, as the WOAH only includes the classical rabies virus in its 24 month infection period criteria, therefore other lyssaviruses such as EBLV are not included. The WHO, in an effort to reach a target of zero human deaths from rabies globally, introduced a new process for verification of rabies free status for countries which are reaching zero human deaths from rabies. The WHO states that while rabies itself remains a notifiable disease in humans and all animal species, rabies free status can be achieved in relation to "dog-mediated RABV" (specifically a canine RABV variant) and "terrestrial RABV" (specifically RABV in all domestic and wild animals except bats). Under this new definition, the UK is free of dog-mediated RABV and terrestrial RABV and is therefore rabies free according to the WHO, despite the presence of rabies in bats (EBLV-1 and EBLV-2) and infections from bats reported annually, posing significant public and veterinary health risks.

Divergent non-RABV species of lyssaviruses have been found almost exclusively in bats with two main exceptions. The first is IKOV which was discovered in an Africa civet in Tanzania and the second is MOKV which has been identified in six different African countries with widespread hosts such as cats, dogs, rodents, shrews and even human infections [Marston *et al*., 2012b; Coertse *et al*., 2017]. Lyssavirus host switching is a concern and Lyssavirus strains have been observed for some time migrating into novel bat populations [Fooks *et al*., 2003], and several recent studies have identified possible emerging reservoirs for lyssavirus strains such as Australian flying fox pups [Barrett *et al*., 2020] and non-human primates in Brazil [Kotait *et al*., 2019].

1.3.2 Burden of disease

Rabies is a neglected pathogen which is highly virulent and causes suffering and death to tens of thousands of humans and an untold number of animals, as well as impacting the economies of many nations. The vast majority of human infections, totalling an estimated >59,000 a year [Hampson *et al*., 2015], involve RABV, with only 16 estimated deaths from other members of the lyssavirus genus according to the WHO [WHO, 2018] however, both these figures are likely significantly lower than the true burden of disease due to poor surveillance in the highest affected nations. Infections occur in over 150 countries [WHO, 2023] and 40% of infections affect children under the age of 15. Every year more than 29 million people receive post-exposure prophylaxis (PEP) which successfully prevents hundreds of thousands of deaths, however the cost of each course is estimated to be $108 USD which can be problematic in developing nations where incomes per person can be as little as $1-2 USD a day [WHO, 2023].

While successful vaccination and prevention programmes have been put in place since the 18th century in many nations, RABV causes more human fatalities than any other zoonotic disease [Fooks *et al*., 2014; Hampson *et al*., 2015], receiving far less coverage than conditions which cause less deaths such as Ebola (2,178 deaths) or Zika virus (13,087 deaths) [Fooks & Jackson, 2020]. Rabies is considered by many to be a disease of the poor with the highest burden of infections occurring in developing nations. It has also been labelled a neglected disease by the WHO which demonstrates the importance of continued research into the pathogen and potential therapeutics. 95% of infections occur in Africa and Asia, Asia accounting for over half of the global rabies human burden with an estimated 37,000 deaths annually occurring in the continent [Hampson *et al*., 2015]. Only 4 of the 30 countries in Asia being considered rabies free: Japan, Singapore, South Korea and Maldives. As is to be expected, the majority of infections, 99%, are RABV and canine mediated. India has the highest number of human rabies infections in the world though not the highest per capita death rate, which is estimated to be Somalia with 12.3 deaths/100,000 people. 16 of the top 20 death rates/100,000 nations are located in the African continent demonstrating that while Asia has more overall deaths, Africa bears a higher burden when factoring in population and contains no nations that are considered to be rabies-free. [Hampson *et al*., 2015].

In addition to the cost of life that lyssaviruses inflict, there are also significant economic impacts to consider. It is estimated that the global cost of canine mediated RABV infection is 8.6 billion USD in losses annually [Hampson *et al*., 2015]. Of the 8.6 billion USD, 55% of the cost is due to premature deaths from RABV, 20% is the cost of post-exposure prophylaxis, 5% due to lost time income seeking PEP, 6% in livestock losses and 1.5% through dog vaccination costs. The most effective method to reduce the cost of dog mediated rabies is the elimination of human

infections through dog vaccination programs. These programs have proven successful in Europe and the USA and have brought costs of combating rabies down despite the initial investment required for the programs themselves. It is important to note however that rabies prevention programs must be continuous and cannot be simply performed once. Without effective continuous vaccination of animal populations RABV can easily be reintroduced to vulnerable canine populations, producing human infections again [Marston *et al*., 2018c].

The true cost of RABV and lyssaviruses as a whole is hard to evaluate. The burden of disease is most heavily felt by developing nations which struggle to report accurate figures due to several factors, from economic to educational deficiencies, in addition to a lack of access to the required medical and scientific infrastructure required for consistent surveillance of disease. Studies have shown that as little as 1% of suspected human rabies infections have been investigated properly and confirmed as such using modern diagnostic methods [Cleaveland *et al*., 2002; Suraweera *et al*., 2012]. Indeed, rabies is often also misdiagnosed as other conditions such as cerebral malaria and Guillain-Barré syndrome.

A study collating clinical patient data found that out of 122 patients, 60.7% were misdiagnosed with conditions other than rabies. Most commonly mistaken for rabies were an indeterminate encephalitis (11.6%), pharyngitis (9.1%) and Guillain-Barre syndrome (5.8%). Misdiagnosis in countries without endemic canine mediated rabies was exceedingly high, with 100% and 91% of bat and dog rabies infection misdiagnosed from the sample set [Udow *et al*., 2013].

A lack of education about the dangers of rabies and the importance of completing post-exposure vaccination courses has been cited as a cause of a significant number of deaths in countries such as Brazil [Duarte, 2021], cementing the importance for increased rabies awareness in the public. This combined with a lack of political will, financial capability and knowledge of the severity of the disease produces an environment where surveillance is severely limited and human infections are allowed to continue when they are in fact wholly preventable.

Recent reporting on lyssavirus infection figures have shown a decrease in annual deaths as of 2019 with a total of 14,075 reported deaths at a 95% confidence interval [Gan *et al*., 2023]. While data such as this relies on accurate reporting which has been demonstrated to be an issue with lyssaviruses, it does present a positive picture of the success of RABV elimination programs. The study cites data from the Global Burden of Disease (GBD) which is based on reporting and suggests a downward trend of RABV infections over the last 30 years. This contrasts the estimated 59,000 deaths annually by Hampson *et al.,* 2015, which is based on multiple mathematical models such as the probability decision tree model framework, considering bite incidence and control efforts in their estimation. Though it is considerably likely that the reported fatalities are much lower than the reality, the reduction in cases still

inspires a hopeful optimism that the 2030 human rabies infection elimination goal set by the WHO and the rabies coalition may come to fruition.

Unfortunately, outside of dog mediated rabies infections it is highly unlikely that lyssaviruses could be eradicated completely. Their wide host repertoire combined with elusive and remote host ecology means that it is likely impossible to truly eradicate rabies from the world [Rupprecht *et al*., 2017]. That being said, elimination of human infection from lyssaviruses can be achieved, through improvements in reporting, introduction of vaccine programmes in susceptible areas, and creation of novel therapeutics which are cross-neutralising.

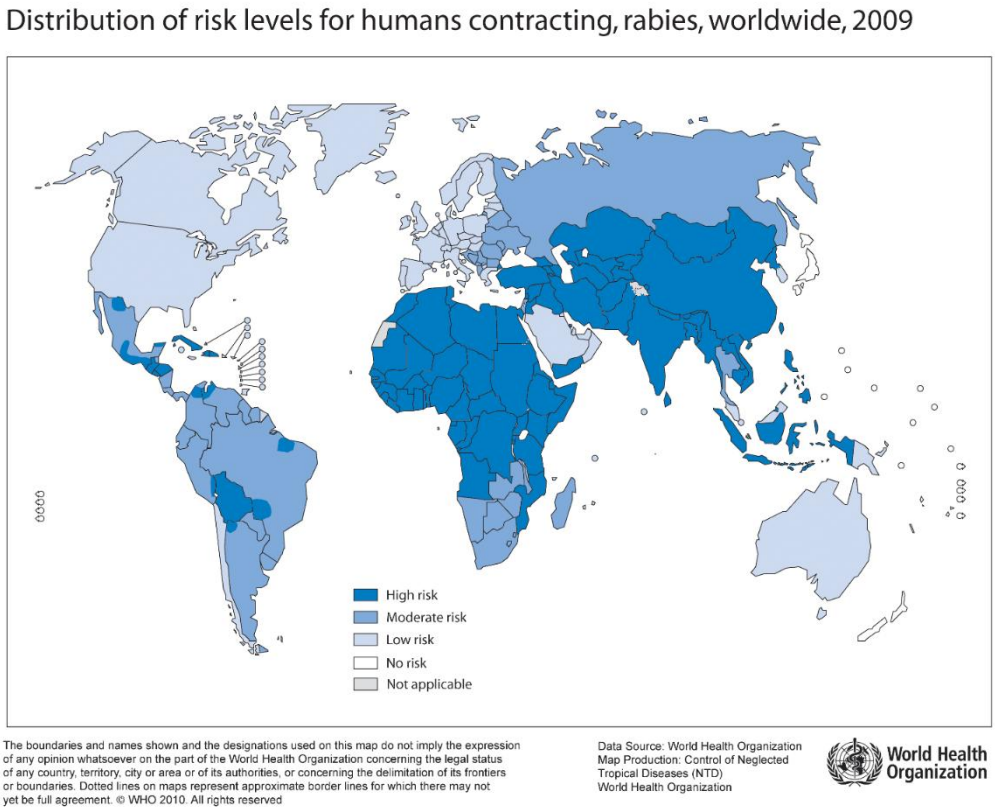


Figure 1.1: World map representing the risk of rabies infection by country [WHO, 2010].

1.4 Lyssavirus structure

1.4.1 Lyssavirus structure and organisation of the genome

Lyssavirus virions are highly structured, bullet-shaped and composed of two main structural units, the first being an internal helical nucleocapsid and the second a host-derived lipid envelope which is acquired in the budding phase of viral pathogenesis [Hummler *et al*., 1967] [figure 1.2]. Lyssaviruses consist mainly of protein, between 67%-74%, which is consistent with other rhabdoviruses. The remaining features of the virus are 20-26% lipid, 3% carbohydrate and 2-3% RNA. Lyssavirus virions are made up of 5 viral proteins, the nucleoprotein (N), phosphoprotein (P), matrix protein (M), glycoprotein (G) and RNA-dependent RNA polymerase also known as the large protein (L) [Schnell *et al*., 2010]. The virion structure surrounds an RNA genome which is tightly coiled and itself encapsulated by the phosphorylated viral N protein. Both the P protein and L protein are also contained within this tightly packed core, associated with the N protein, together forming the nucleocapsid or ribonucleoprotein (RNP). This structure is surrounded by the M protein which spans the entire RNP core. The M protein sits beneath the hosts lipid bilayer which forms the envelope of the virion, within which are embedded trimeric spike viral G proteins in a highly ordered fashion at the surface of the virus, the only surface viral protein.

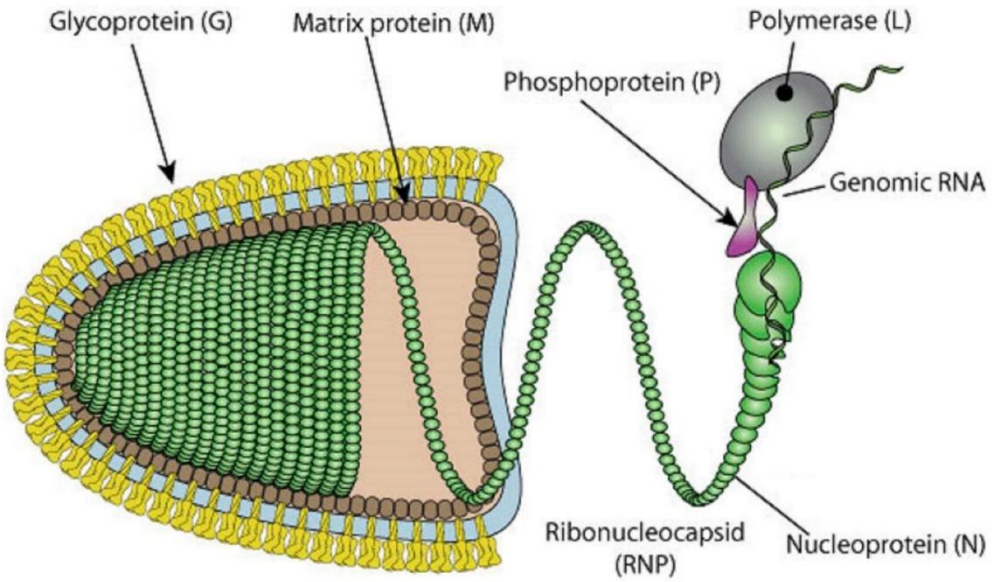


Figure 1.2: Lyssavirus virion structure with each of the viral proteins labelled [image retrieved from Dhulipala S & Uversky, 2022].

1.4.2 Viral genome structure and protein functions

The viral genome is approximately 12 kilobases in length and encodes 5 viral proteins [Schnell *et al*., 2010]. The structure of the single stranded negative sense genome consists of a 58-nucleotide leader sequence (Le) at the 3' followed by the conserved order of genes N-P-M-G-L, and a 5' terminal extragenic Tr region. The Le region of the genome is considered to be a genomic promoter which helps direct transcription of monocistronic mRNAs.

Lyssavirus viral proteins often serve multiple purposes and all play roles in generating novel virions. The N protein is approximately 450 amino acids long (RABV Pasteur virus strain (PV)) and phosphorylated, it has several functions including protecting the RNA genome from ribonuclease activity, triggering transition from RNA transcription to replication, and potentially enabling immune avoidance through the prevention of RIG-I recognition of the RNA viral genome [Kissi *et al*., 1995; Liu *et al*., 2004; Yang *et al*., 1999; Wunner *et al*., 1988; Masatani *et al*., 2010]. It is the most abundant protein in the RNP and the has the highest conservation of the 5 viral proteins.

The phosphoprotein (P) which is also part of the RNP is approximately 297 amino acids long (PV). It, like the N protein, is phosphorylated and multifunctional, interacting with N to form N-P complexes which enables it to act as a chaperone for newly synthesized viral N protein helping to prevent self-assembly (polymerisation). In addition, this N-P complex formation prevents N protein from non-specific binding to cellular RNA and helps specifically direct the N protein to encapsulate viral RNA instead [Mavrakis *et al*., 2003]. The P protein also interacts with the L protein during transcription and replication, helping to stabilize the L and position the P-L complex on the RNA template which the L protein alone has been shown to be unable to do [Chenik *et al*., 1994 & 1998; Fu *et al*., 1994]. Similar to the N protein, the P protein has host innate immune system counteraction responsibilities, specifically through interaction with the type I/III interferon system. It prevents the activation of latent interferon regulatory factors (IRF-3), which is the initial transcription factor in IFN expression [Honda & Taniguchi, 2006]. This prevention is mediated by interference by the P protein of the phosphorylation of the IRF-3 protein [Brzózka *et al*., 2005], and blocking the IFN-mediated JAK/Stat signalling in the cytoplasm [Blondel *et al*., 2015].

The M protein is approximately 202 amino acids long, making it the smallest of the 5 viral proteins. It lines the viral envelope, with one part of the matrix attaching to one end of the viral envelope and the rest sticking into the interior of the virion. This allows it to interact with the N protein and connect between the lipid bilayer and the RNP [Guichard *et al*., 2011]. It is responsible for the virions bullet shape and for the assembly and budding of the virions, interacting with the G protein to achieve this [Mebatsion *et al*., 1999]. The M protein also

performs immune system interference by restraining the NF-kB pathways of host cells and interacting with the P protein to inhibit the JAK-STAT pathway signalling pathway [Sonthonnax *et al*., 2019].

Finally, the large RNA-dependent RNA polymerase protein (L) is the largest of the viral proteins with approximately 2142 amino acids in length. It comprises 54% of the coding potential of the lyssavirus genome. As an RNA-dependent RNA polymerase, the L protein functions to catalyse the replication of RNA from the viral RNA template. It conducts the transcription of viral mRNAs through the use of its transcriptase, capping and polyadenylation processes. Viral genome replication is also performed through the L proteins replicase activity.

1.5 The lyssavirus glycoprotein

1.5.1 Glycoprotein structure

Lyssavirus glycoproteins are approximately 65 kDa in size, 524 amino acids in length and represent the only surface protein on the virion [Gaudin, 1997]. The glycoprotein formation is a trimeric spike protein which coats the surface of the virus in a highly ordered manner and is primarily used to bind to host cells amongst other functions [Buthelezi *et al*., 2016]. This trimeric formation is theorised to occur in the Golgi apparatus, however this has not been experimentally demonstrated. The pathogenicity of lyssaviruses has been attributed to the glycoprotein, with studies determining it to be the most significant factor [Luo *et al*., 2020]. The RABV glycoprotein consists of a 19 amino acid signal peptide, 21 amino acid transmembrane domain, 439 amino acid ectodomain and a 42 amino acid cytoplasmic tail [Anilionis *et al*., 1981; Wunner *et al*., 1988]. Rabies glycoproteins have been shown to aid in the budding efficiency of lyssaviruses and in mobilising RNA or RNPs [Mebatsion *et al*., 1996; Rolls *et al*., 1994]

The signal peptide is hydrophobic and located at the N-terminal of the protein. It is responsible for providing a membrane insertion signal which results in the transporting of the glycoprotein to the endoplasmic reticulum-golgi apparatus membrane where the protein undergoes the ER-Golgi secretion pathway. This results in the nascent protein transport to the cell membrane. The signal peptide is cleaved from the mature protein in the Golgi apparatus. Due to the removal of the signal peptide in mature glycoproteins, it is often not included when discussing amino acid positions of the glycoprotein, therefore the numbering of them begins at the 20^{th} residue [Fernando et al., 2016; Conzelmann et al., 1990; Marissen *et al*., 2005].

The ectodomain (ED) is positioned so that it is protruding from surface membrane of the mature virion. This domain is important for a number of functions. The ED is responsible for interaction with the host cellular receptors, inducing of the low pH fusion of the viral envelope with plasma and endosomal membranes in infected cells [Albertini *et al*., 2012], and is essential in the antigenicity of the virus. It induces the host humoral immune response and is the target for virus neutralising antibodies and helper/cytotoxic T cells [Celis *et al*., 1988, Macfarlan *et al*., 1986].

The lyssavirus glycoprotein adopts a type I single pass transmembrane topology [Yang *et al*., 2020]. This means that the protein is anchored to the host lipid membrane, which the virion incorporates during budding, and has a single transmembrane α-helix segment. It is also a class III viral fusion protein with four distinct domains in the ectodomain [Ng *et al*., 2022]. A class III fusion protein can vary in size but generally produces trimeric, elongated, rod-like structures which is the case with the RABV glycoprotein.

The glycoprotein structure is sensitive to pH changes which modify the structure of the protein. This is essential during cell entry, however unusually for viral fusion proteins the process is reversible [Gaudin *et al*., 1991]. The glycoprotein of RABV exists in at least three different structural states, the native state which is detected at pH 7 or above, the activated state which is detectable at more acidic levels, and the transient state detectable at pH 6.7 [Gaudin *et al*., 1993 & 1996]. The activated state is conformationally different to the native state with the more acidic environment resulting in glycoprotein domain re-orientation, extending from a bent hairpin conformation into a longer structure [Yang *et al*., 2020]. These conformational changes have been demonstrated in another study by Gaudin *et al*., 1993, to be a key pre-fusion transition. The importance of this transition is demonstrated through the induction of virions into a transient state, defined as half of all glycoprotein trimers being inactivated which occurs at pH 6.7. During this state, RABV is unable to commence membrane fusion, which suggests that multiple trimers are required for membrane fusion within RABV infection.

1.5.2 Antigenic regions

Four major and one minor antigenic sites have been identified within the RABV glycoprotein which have been particular targets for research [Lafon *et al*., 1983; Benmansour *et al*., 1991; Prehaud *et al*., 1988; Dietzschold *et al*., 1982; Seif *et al*., 1985]. These epitopes are continuous with the exception of antigenic site II which composed of two distinct regions (IIb, residues 34-42 and IIa, residues 198-200) [Wunner *et al*., 1988]. The remaining antigenic sites are located at: site I (263-264), site III (330-338), site IV (226-231) and site ‘a’ (342-343) [Evans *et al*., 2018]. Though essential in RABV neutralisation, it is unknown if antigenic sites present in RABV bear the same or similar function within other lyssaviruses. It has been shown that the swapping of antigenic sites between lyssaviruses of phylogroups I and II, elicits intra-, but limited inter-cross-neutralisation [Evans *et al*., 2018]. Antibodies raised against phylogroup II were able to neutralise phylogroup I pseudo viruses (PVs) which presented the phylogroup II antigenic sites. This was used to demonstrate that antigenic site II was immunodominant. It also showed that site IV was poorly neutralised after swapping. The same study further demonstrated that individual antigenic sites within lyssavirus glycoproteins can affect neutralisation profiles, indicating a need to investigate individual lyssavirus strains and their antigenic sites, as this individuality may be a contributing factor in the lack of cross-neutralisation between phylogroups.

1.5.3 Glycan presence

N-linked glycosylation (NLG) is a specific subset of glycosylation which involves the addition of a glycan moiety to an asparagine (Asn) amino acid within the protein structure, usually found in the sequon Asn-X-Ser/Thr [Shakin-Eshleman *et al*., 1992a]. This takes place during viral

replication within the ER and is N-linked due to the glycan being attached to the amide nitrogen atom of an asparagine residue. This process is beneficial in eukaryotes but is also utilised by a multitude of viruses in order to shield the surface glycoprotein from host immune responses via antigen recognition [Vallbracht *et al*., 2018]. Glycosylation can result in changes to structure, intra-cellular transportation, physiochemical properties, and function of proteins.

Due to its importance in immune response mitigation, NLG has been researched within lyssaviruses for some time. Studies such as those by Shakin-Eshleman *et al*., 1992a, and Katsuri *et al*., 1995, suggested that there are 3 sequons located in antigenic regions which are suitable for glycosylation, though only Asn^{247} and Asn^{319} are efficiently glycosylated in virions, whereas Asn^{37} is not. However, it is generally accepted now that there are two glycosylation sites in street/wild type RABV, Asn^{37} and Asn^{319}. It has also been demonstrated through sequon knockouts that NLG performs an essential role in the expression of the RABV surface glycoprotein structures with alteration of the natural viral glycosylation directly affecting surface glycoprotein levels of expression. Additionally, when knocked out, the inefficient Asn^{37} had effects on levels of expression, indicating all 3 sequons are involved in surface glycoprotein expression [Shakin-Eshleman *et al*., 1992a].

Sequon and glycosylation variation has also been demonstrated within street and fixed strains of RABV [Yamada *et al*., 2013]. Fixed strains are laboratory-based viruses which have undergone multiple passages in order to adapt them to lab conditions. As a consequence, they are often found with changes to their protein structure. Street viruses generally contain either a single sequon at Asn^{319} or two sequons at Asn^{37} and Asn^{319} whereas fixed strains often have numerous positions including Asn^{158}, Asn^{204} and Asn^{247} [Yamada *et al*., 2013]. The introduction of these additional sequons within fixed strains creates a quandary, studies have shown that these mutant sites can have a significant impact on viral production, therefore potentially affecting studies which do not take these into account when utilising fixed strains, making them less representative [Shakin-Eshleman *et al*., 1992a; Yamada *et al*., 2013].

More recent research into glycosylation has confirmed the inefficient core-glycosylation at the Asn^{37} site [Yamada *et al*., 2013], and, through mutant production, has determined that the sites function was to facilitate the production of viral progeny. The mutant panel generated also demonstrated the effects of the addition of N-glycans to sites in both fixed and street viruses. Each sequon site had a differing response to this with some outcomes being positive for the virus, some negative with additional dependencies on which cell lines infected. For example, additional N-glycans at sites Asn^{194} and Asn^{247} greatly enhanced viral production whereas at sequon Asn^{158} the effect was negligible. The study also poses important questions which represent gaps in the literature such as what the effects of further glycosylation sites are and

how the mechanisms of this glycosylation impacts progeny development and viral production levels. Indeed, lyssavirus glycan research also lacks information on the impact of glycans on infectivity and neutralisation profiles of lyssaviruses.

Viruses can be heterogeneous with the number of glycans which are present in any one glycoprotein, and indeed the types of glycans which are present also [Wojczyk *et al*., 2005]. The RABV glycoprotein is also dependent on calnexin interaction. Calnexin is a molecular chaperone within the host cell that recognises the partially trimmed monoglucosylated glycan that has been attached to the glycoprotein, binding to it and supporting protein folding [Gaudin, 1997]. The essential nature of this interaction means that the rabies glycoprotein must have glycosylation in order to produce functioning proteins, as such Asn319 is efficiently glycosylated in all members of the genus.

1.6 Lyssavirus transmission, pathogenesis and replication

1.6.1 Lyssavirus transmission and pathogenesis

Rabies is transmitted from host to host primarily through bites and scratches with the main vector being saliva. There have been multiple cases of rabies transmission through organ transplants, however this is not a primary route of transmission and is quite rare [8 Zhang *et al*., 2018]. For transmission to human hosts, 99% occur through the bites of rabid domestic dogs (*Canis lupus familiaris*) [WHO, 2013]. Upon deposition into the new hosts tissue through a bite, lyssaviruses enter the nervous system (NS) through two main entry points: a sensory nerve spindle or, more commonly, a motor neuron via the synaptic cleft of a neuromuscular junction [Burrage *et al*., 1985; Murphy *et al*., 1973]. To achieve infiltration of the nervous system, lyssaviruses make use of 4 currently known receptors: Nicotinic acetylcholine receptor (nAChR) [Lentz *et al*., 1982], neural cell adhesion molecule (NCAM/CD56) [Thoulouze *et al*., 1998], low-affinity p75 neurotrophin receptor [Tuffereau *et al*., 1998], and metabotropic glutamate receptor subtype 2 [Wang *et al*., 2018]. This use of multiple receptors by lyssaviruses adds a layer of difficulty when attempting to design therapeutics which target host receptor interactions.

The lyssavirus centripetal propagation towards the CNS occurs within motor and potentially sensory axons of peripheral nerves using retrograde fast axonal transport. This has been proven through the use of colchicine on lyssavirus infected rats, which inhibited fast axonal transport in the sciatic nerve and resulted in the halting of RABV propagation (Tsiang, 1979). This rate of transport has been determined to be between 50-100 mm/day in Human dorsal root ganglia (Tsiang *et al*., 1991). In order to enter the nervous system the virus makes use of either receptors on the surface of the host neuromuscular junction neuron, or buds across the synaptic cleft after infecting local muscle cells and replicating, as seen in figure 1.3 [Fooks *et al*., 2017]. Following the receptor mediated entry into the neuron, the virus is contained within endosomal transport vesicles [Tsiang, 1979; Lycke & Tsiang, 1987; Gillet *et al*., 1986; Kucera *et al*., 1985; Ceccaldi *et al*., 1989], where it remains until arrival at the neuronal soma via retrograde axonal transport through microtubules. Once in the neuronal soma, the viral RNP is released from the vesicles and replication occurs [Piccinotti & Whelan, 2016], resulting in novel virion assembly at the post synaptic membranes in preparation for trans-neuronal transmission via the viral glycoprotein in order for the cycle to be repeated [Charlton & Casey, 1979; Etessami *et al*., 2000].

The process continues until CNS neurons in the spinal cord or brainstem are reached, at which point the virus rapidly spreads throughout the CNS using the retrograde fast axonal transport. RABV in the CNS has been shown to have a strict preference for neurons over other potential

CNS target cells such as astrocytes and microglia [Ugolini, 2011]. Though some budding has been observed into the intracellular space, the primary movement of virions is trans neuronal and dendroaxonal [Charlton & Casey, 1979]. Rabies virus propagation does not occur at gap junctions or cell to cell, but rather at chemical synapses and neurons which have been infected with RABV remain functional and are able to express their neurotransmitters and also cotransport other tracers [Ugolini, 2011].

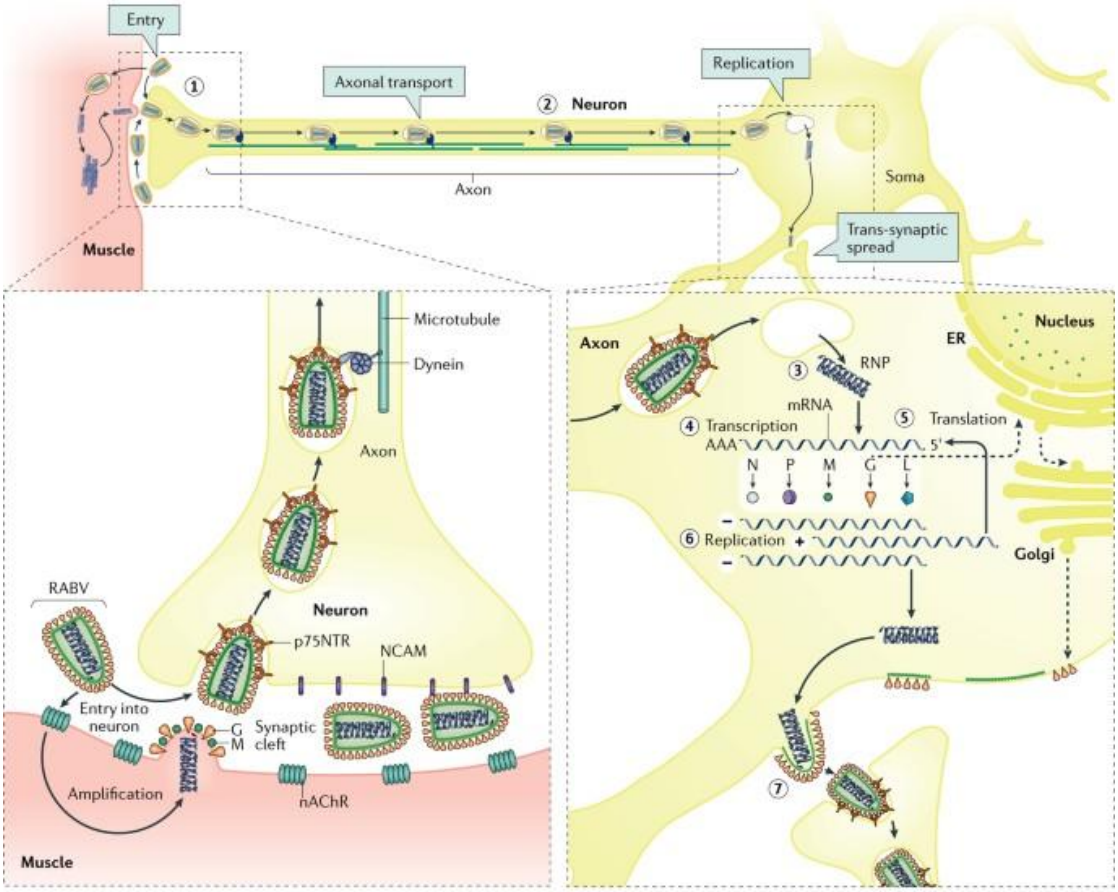


Figure 1.3: A visual representation of the early lyssavirus entry (1), transport (2), replication (3-6) and release to the following neuron (7). Image sourced from: [Fooks et al., 2017].

For lyssaviruses to be transferred to the next host there needs to be spread to the salivary glands of the host. To accomplish this, the virus spreads down facial nerve, reaching the ganglion neurons of the mandibular gland. This is followed by infection of the acinar epithelium via the salivary gland myoepithelium, resulting in non-suppurative sialadenitis [Boonsriroj *et al*., 2016] [figure 1.4]. In addition to the salivary glands, lyssaviruses spread to multiple organs from the CNS such as the skin, heart, liver, pancreas, adrenal glands and tongue [Macedo *et al*., 2006] with both neuronal and non-neural infections being demonstrated. However, it has been speculated that this spread is likely due to passive diffusion rather than active transport, though there is currently a lack of study surrounding this process [Ugolini, 2011].

Cellular entry of RABV has been difficult to characterise due to limited methods and viral models. However, recent studies such as that by Xu *et al*., 2015, and Piccinotti & Whelan, 2016, have used novel methods to resolve key details within the pathogenesis of RABV. Xu *et al*., 2015, used single virus tracking to produce real-time images of viral particles infecting cells. It was demonstrated that internalisation of RABV occurs via classical dynamin-dependent clathrin-mediated endocytosis with a requirement for intact host actin, achieved by blocking each pathway and measuring infectivity. Another important confirmation was the use of cellular filopodia by RABV to gain proximity to cellular membrane in order for attachment to occur, with it being suggested that some virus entered via attachment to the filopodia. Filopodia were also shown to increase in production in the presence of RABV, therefore assisting viral uptake. However, the mechanisms behind this and the activation of filopodia were not shown leaving this open to further investigation. In vitro, lyssaviruses have the capability to infect most cell types, however the underlying mechanisms which enable this are not yet understood [Fooks *et al*., 2017].

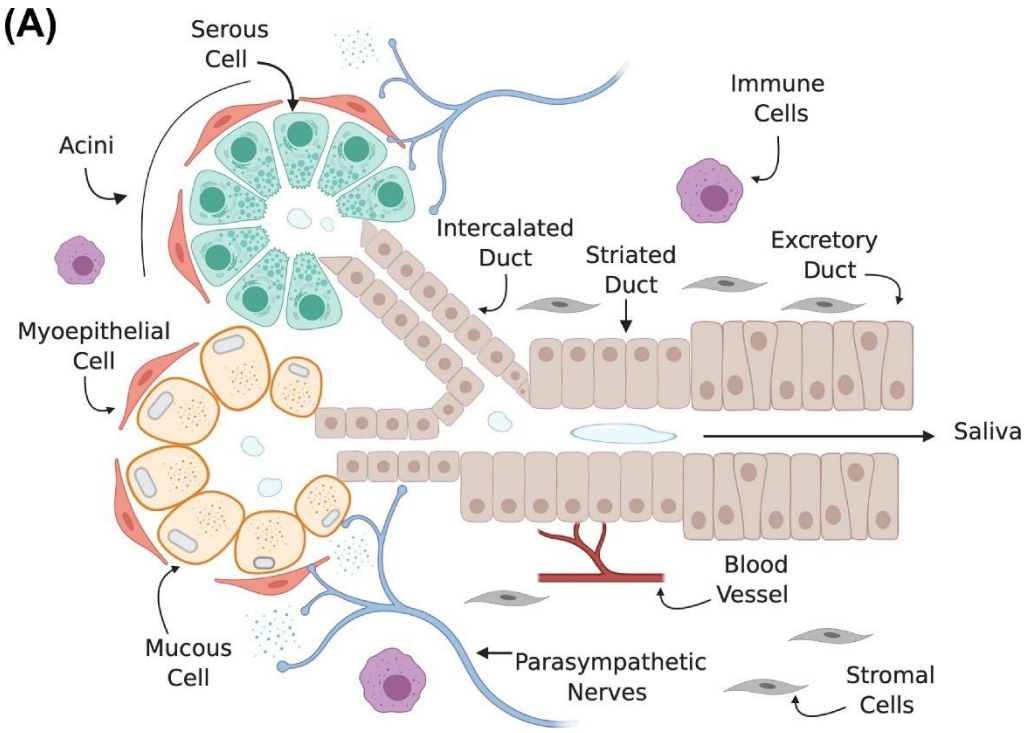


Figure 1.4: Stylised layout of the salivary region as an example of the route taken by lyssaviruses when travelling to enter the saliva. Image referenced from [Rocchi & Emmerson, 2020].

1.6.2 Lyssavirus replication

Lyssavirus replication occurs within the cytoplasm of infected neurons. Once the RNP core has been released from the endosomal vesicle and disassociated from the matrix protein during uncoating, the tightly coiled structure loosens, relaxing to form a coiled helix [Iseni *et al.,* 1998].

As RABV is a single stranded negative sense RNA genome, the initial stage of replication is transcription into mRNAs. Genes are transcribed from the RNA genome in a sequential manner in the order of the genome N, P, M, G, L [Flamand & Delagneau, 1978; Holloway & Obijeski, 1980]. Transcription of the genes occurs in a stop-start manner from the 3' end of the genome. Each gene is separated by conserved 'border' regions of nucleotides which direct the polymerase activity during transcription to stop and start. The stop/poly(A) nucleotides in the border region are characterised by, in RABV, 3'…ACUUUUUU…5'. This sequence initiates the production of a 50-150 nucleotide long poly (A) tail on the mRNA as the L protein 'stutters' on it, repeatedly adding a nucleotide without leaving the vRNA template [Barr *et al*., 1997]. Remaining still on the viral RNA, the polymerase releases the mRNA of the gene preceding the stop signal and then continues from the start signal at the end of the border region and the start of the next gene, in RABV 3'…UUGURRNGA…5'. The stop start signals are separated by intergenic spacer sequences which are not coded. As a result of this system, the earlier genes are transcribed more often as the polymerase can dissociate at each stop signal, leading to higher levels of particular proteins in a gradient fashion [Finke et al., 2000]. This has advantages such as the production of appropriate levels of required proteins in an efficient manner, reduction of cell cytotoxicity levels and reduced risk of host immune system activation [Schnell *et al*., 2010; Faber *et al*., 2002].

Host free cytoplasmic ribosomes synthesise the viral proteins. Viral glycoproteins are then glycosylated in the host endoplasmic reticulum and Golgi apparatus. As the levels of N protein increase, replication of the viral genome commences which is initiated with synthesis of full-length positive strand copies of the viral genome. This is achieved by ignoring stop codons, with the viral polymerase entering a single site at the 3' end and continuously transcribing until the 5' end. The positive strands serve as templates for the synthesis of negative strands which will be encompassed to produce an immature virion [CDC, 2020].

N, P and L proteins assemble to form a complex which encapsulates the newly synthesised negative, single strand RNA. This forms the RNP core, with the matrix protein surrounding it. This stage is followed by migration to the host plasma membrane. The M-protein then coils and binds with the glycoprotein, budding from the membrane which contains the G protein and forming a mature virion [CDC, 2020].

1.7 Lyssavirus entry receptors

1.7.1 Nicotinic acetylcholine receptor

The Nicotinic acetylcholine receptor (nAChR) receptor is a polypeptide that responds to the neurotransmitter acetylcholine and was the first receptor for lyssaviruses discovered [Lentz *et al*., 1982]. RABV antigen has been observed in neuromuscular junctions at sites coincident with nAChR. This was further confirmed by infecting myotubes that had been treated with either the irreversible binding nicotinic cholinergic antagonist α-bungarotoxin or the reversible binding to *d*-tubocurarine. This resulted in reduced infection from RABV in the treated myotubes. This data was further extrapolated on in rat myotube models using α-bungarotoxin which confirmed the reduction in infectivity following this inhibition of nAChR [Tsiang *et al*., 1986].

Once nAChR was identified as a receptor candidate, the details of binding sites were investigated. The nAChR receptors are transmembrane proteins that form pentameric structures which are assembled from multiple subunits, these produce diverse subunit combinations that can be found within different locations, for example α2-α10 are found in neuronal cells, whereas α-1 and β1 are found at neuromuscular junctions [Millar & Gotti, 2009; Millar, 2003]. Both RABV and neurotoxins bind to sites located in the nAChR which are located at amino acid residues 173-204 of the α1-subunit. The highest affinity residues, which are also virus binding determinants, are located within residues 179-192 [Lentz, 1990]. Snake venom neurotoxins have been shown to bind with high affinity to nAChR. When the amino acid sequence of the snake venom protein is compared to that of the RABV glycoprotein there is a considerable amount of sequence identity between them. Residues 151-238 of the RABV glycoprotein match closely to the entire neurotoxin protein sequences which is 71-74 residues long. This helps elucidate the potential site of nAChR recognition on the RABV glycoprotein [Lentz, 1985].

The presence of nAChR throughout the nervous system and in particular the neuromuscular junctions make it an important receptor for RABV pathogenesis as a neurotropic pathogen. Indeed, the binding of RABV to nAChR receptors has been suggested to help concentrate virus on post-synaptic cells, facilitating infiltration of the virus into cells. However, the α1-nAChR is mostly present at the post-synaptic membrane rather than the pre-synaptic [Lafon, 2005], and as RABV moves up the NS in a retrograde fashion this would mean difficulty in entering neurons using this receptor at neuromuscular junctions. The current exact mechanisms of nAChR use in RABV infection are not well understood and due to the complication of additional receptors being used, are difficult to elucidate, though it is clear that it is an important receptor in lyssavirus infection.

1.7.2 Neural cell adhesion molecule (NCAM)

Neural cell adhesion molecule (NCAM), also known as CD56, is a glycoprotein which is expressed on the surface of skeletal muscle, glia and neurons. It is a cell adhesion protein, a subset of cell surface proteins, which means it is involved in cellular binding with other cells or with the extracellular matrix. It has also been implicated in neurite outgrowth, learning and memory, and synaptic plasticity [Dallérac *et al*., 2013].

The NCAM receptor was first identified after it was found that cell lines which did not contain NCAM were resistance to RABV infection [Thoulouze *et al*., 1998]. The study also found that when cells susceptible to RABV were incubated with the virus, the surface expression of NCAM was reduced but had no impact on other essential surface proteins. This evidence points towards the internalisation of RABV through RABV-NCAM complexes that undergo adsorptive endocytosis resulting in viral entry to the cell. Blocking of the receptor with heparan sulphate also reduces the RABV infection levels as does inhibition of the receptor using monoclonal and polyclonal antibodies. [Thoulouze *et al*., 1998]. In addition to this, soluble NCAM was found to neutralise RABV infection which indicated viral particles bound to the soluble receptor were prevented from binding to target cells, rendering them neutralised. Finally, the addition of NCAM receptors through cDNA introduction introduced RABV infection susceptibility to cells which had previously been resistant. This abundance of evidence strongly supports NCAM as a RABV receptor.

However, as with nAChR, NCAM is not the sole receptor for RABV as it has been demonstrated that NCAM knockout mice can still succumb to RABV infection despite a lack of NCAM, though the infection is delayed in comparison [Thoulouze *et al*., 1998]. CD56 is also present in the same isoform as RABV on the surface of NK cells, making infection of them possible by RABV though this has not been experimentally shown.

1.7.3 Low affinity p75 neurotrophin receptor

Low affinity p75 neurotrophin receptor (P75NTR) is a neurotrophic factor receptor. These receptors under normal circumstances bind neurotrophins which can include nerve growth factor, neurotrophin-3 and Brain-derived neurotrophic factor. P75NTR functions also extend to promotion of apoptosis, neurite growth, cell cycle regulation and myelination by Schwann cells [Brady *et al*., 2012; Bamji *et al*., 1998]. The receptor is a type I transmembrane protein and does not appear to be found at neuromuscular junctions [Sheard *et al*., 2002]. It is widely expressed throughout the CNS during development; however, it is downregulated once adulthood is reached and found in only a few neuronal populations within the brain [Brady *et al*., 2012]. Interestingly, p75NTR expression is increased in both Schwann cells and motor neurons after

injury, perhaps giving RABV increased receptor availability due to the most likely introduction of RABV being through muscle injury after a bite [Pérez *et al*., 2019], however this has not been shown experimentally.

p75NTR was discovered as a RABV receptor using a random-primed cDNA library which was generated from the mRNA of neuroblastoma cells (NG108). The library was used to transfect COS7 cells which enabled the identification of a single plasmid after subcloning. This plasmid was then transfected into BSR cells which resulted in RABV glycoprotein binding [Tuffereau *et al*., 1998]. The plasmid was found to have high sequence identity with human and rat p75NTR. BSR cells cannot generally be infected by wild type (WT) RABV, as is the case with most non-neuronal cell lines. However, BSR cells which expressed p75NTR were able to bind soluble RABV glycoprotein. Though this provides promising evidence of the nature of p75NTR as a RABV receptor, it has been shown to not be essential in infection, with evidence some neurons which express p75NTR being unable to bind to RABV [Tuffereau *et al*., 2007]. Due to its lack of appearance at neuromuscular junctions and more concentrated nature at the dorsal horn and spinal cord, it has been speculated that the receptor may be involved in the trafficking of RABV via sensory pathway [Lafon, 2005]. More experimentation is needed to better define the role of p75NTR in lyssavirus infection.

1.7.4 Metabotropic glutamate receptor subtype 2

Metabotropic glutamate receptor subtype 2 (mGluR2) has been recently identified as a RABV receptor. It is abundant in the CNS, expressed on the pre-synaptic axon terminal, and is a member of the G protein-coupled receptor family [Moussawi & Kalivas, 2010]. Investigations have confirmed it to be a functional cellular entry receptor that interacts directly with the RABV glycoprotein, mediating viral entry into cells [Wang *et al*., 2018]. The study also found that, in knockout experiments removing the receptor from cells, RABV infection was significantly decreased. Additionally, antibodies which were targeted to mGluR2 were capable of preventing RABV infection in an *in vitro* experiment. Finally, mGluR2 ectodomain soluble protein was able to neutralise RABV infectivity *in vitro* and *in vivo* in a study performed on mice. Further research needs to be performed on this receptor to characterise its role in RABV infection.

1.7.5 Further receptor information

Conservation of receptors amongst species is an important factor in transmission of rabies into potential novel reservoirs. Studies have shown that NCAM is well conserved across species including mammals, birds, reptiles and other vertebrae [Moore *et al*., 1998; Albach *et al*., 2004; Hoffman *et al*., 1984]. However, none of the studies found specifically mentioned canines which, as the primary route of infection to humans, should be investigated. This was similarly

the case with the nAcHR receptor, studies such as those by Yoshioka *et al*., 1999, and Shorey-Kendrick *et al*., 2015, provide evidence of high conservation between humans, canines and non-human primates, with a 96.1% homology between human and canine nAcHR alpha-subunits. There is, however, a lack of data regarding full length sequence comparisons as well as investigation into the other subunits of nAcHR such as α2, β, γ, and δ, which leaves a more precise conclusion as to the receptor's conservation unavailable.

1.8 Clinical signs and symptoms

Rabies presents as an acute and progressive encephalomyelitis which can be characterized by the development of a range of neurological abnormalities in mammals [Jackson, 2014]. Rabies is a broad and varied disease which can cause a multitude of clinical symptoms to occur. There are two main presentations of rabies symptoms. Approximately 80% of patients will develop an encephalitic, classical form of rabies also known as furious rabies due to the behavioural changes, with the remaining 20% developing a paralytic form. At present, the cause of these divergent patient outcomes is poorly understood and not well defined [Fooks *et al*., 2017]. The furious and classical forms of rabies generate different symptoms from one another. Both forms generally take around 20-90 days for initial symptoms to manifest in humans. The duration of the asymptomatic period is variable. A study collating data on naturally acquired human cases found the median incubation time was 54 days. Incubation time differed based on which animal had been the source of infection, dog median of 64.5 and bat median of 51, likely due to the viral load delivered to the host [Udow *et al*., 2013]. Non-rabies specific prodromal symptoms can occur which take the form of pain or itching at the site of the initial infection as well as generalised symptoms such as fever, chills, malaise, fatigue, insomnia, anorexia, headaches, anxiety and irritability, all of which may last up to 10 days prior to the onset of neurological symptoms [Warrell, 1976].

Following the short prodromal phase, the patient enters the acute neurological phase. During this phase, encephalitic rabies patients commonly have episodes of hyperexcitability, confusion, agitation, hallucinations, throat spasms, hypersalivation and aggressive behaviour which typically last for short periods of time, a little as 1-5 minutes [Fooks *et al*., 2017]. 50-80% of patients with encephalitic rabies develop hydrophobia, which is not a feature of any other known diseases and is a well-known sign of rabies infection [Fooks & Jackson, 2020]. In contrast to encephalitic rabies, paralytic rabies causes muscle weakness, drowsiness and paralysis as well as a generally longer acute neurological phase [Hemachudha *et al*., 2005]. Both forms of rabies infection eventually lead to coma and the death of the patient. Interestingly, there appears to be different presentations of RABV infection in patients depending on the source if the source of the infection was bat or dog mediated [Udow *et al*., 2013]. Bat RABV infection was found to cause tremors and involuntary twitching/jerking more often, while dog strains generated hydrophobia and aerophobia. This links to the current theory that symptoms and the likelihood of developing either type of rabies is tied to a number of factors such as viral load delivered, species, strain, patient immune competence and location of the initial infection [Fooks *et al*., 2017]. However, these factors require further research as they are not well understood.

1.9 Lyssavirus treatments

1.9.1 Vaccination

Rabies vaccinations are most often given post exposure, what is known as post exposure prophylaxis (PEP). The WHO classifies rabies exposure into three major categories [WHO, 2013]. Category I is touching an animal or receiving licks, in this case treatment not necessary. Category II exposures are minor scratches or nibbles that do not result in bleeding, these require wound cleaning and administration of PEP, however if the source of category II exposure was a bat, the WHO now recommends this be treated as category III [O'Brien *et al*., 2019]. Finally, category III are all transdermal or mucosal exposures, this includes biting or licking of broken skin or mucosal membranes, this requires thorough washing and disinfection of the wound, infiltration of the wound with rabies immunoglobulin (RIG) and vaccination. Individuals that have received the complete regimen of a WHO approved rabies vaccine previously need booster immunization but not RIG. Treatment is considered a high priority and should be sought out as soon as possible. Annually, 26 million people worldwide are estimated to receive PEP [WHO, 2023]. Treatment with vaccination and RIG post exposure is nearly 100% effective, though it is not effective once symptoms set in (Sadeghi *et al*., 2015; WHO, 2013).

Vaccination for rabies has existed since the 19th century when, in 1885, Louis Pasteur injected a boy who had been bitten by a rabid dog with a vaccine he had developed. The vaccine had been produced using a crude solution of dried spinal cord tissue from rabbits infected with rabies virus stock which had originated from a rabid cow [Lafon *et al*., 1988; Pearce, 2002]. The virus was serially passaged 80 times in rabbits and 13 total injections were used on the patient who ultimately survived [Fooks & Jackson, 2020].

Following the success of this vaccine, modifications were made which inactivated the virus using chemicals. The most common inactivated vaccine was developed by Sir David Semple in India in 1911. For this vaccine, infected sheep brain was taken and the virus was inactivated with phenol which also acted as a preservative [Pearce, 2002]. So called Semple type vaccines, ones which had been derived from nerve tissues, were used for decades after. Though these types of vaccines did protect from infection, they did so with side effects such as immune responses to the nerve tissue which ultimately lead to debilitating and sometimes fatal neurological diseases in up to 1/80-200 patients. In response to this, subsequent vaccines were produced in new-born mice or rats which had naïve immune nervous systems. The WHO no longer recommends nerve tissue vaccines and nearly all countries have discontinued their use [WHO].

1.9.2 Modern vaccines

Modern vaccination currently has much more variation than the rabies vaccination of the past. There are many approved vaccines with different components, the most common being purified Vero cell rabies vaccine (PVRV), purified chicken embryo cell vaccine (PCECV) and human diploid cell vaccine (HDCV). At present, most animal rabies vaccines used are produced in BHK-21 cells as well as Vero cells. The currently produced rabies vaccines such as Rabipur, use attenuated strains which are grown in different vectors and then inactivated in order to produce the vaccine. These attenuated strains are known as 'fixed' viruses and are produced through serial passage, grown in neural tissues of rabbits, sheep, goats, mice or rats or in cell cultures. Examples of these include: Challenge virus standard (CVS), Flury low egg passage (LEP), Flury high egg passage (HEP), Kelev, Evelyn Rokitniki Abelseth (ERA), Vnukovo-32, Street Alabama Dufferin (SAD), Pasteur virus (PV), Pitmann Moore (PM) [Fooks & Jackson, 2020].

1.9.2.1 Purified chicken embryo cell vaccine

PCEVC is a primary cell vaccine sourced from chicken embryos. Avian embryos in vaccine production were first described by Goodpasture et al in 1931. They are suitable for such a task as live cells are required for virus cultivation and their closed system make them ideal for viral production in an isolated environment. Currently the main WHO approved PCEVC vaccine is Rabipur [WHO]. Rabipur is an inactivated rabies vaccine derived from the fixed-virus strain Flury LEP. It is inactivated using β-propiolactone and then purified through centrifugation. The vaccine has a shelf life of 48 months and can be stored as a lyophilised powder at 2-8°C [Bavarian Nordic A/S, 2022].

However, the use of primary cells comes with some disadvantages. The main ones being the need for continuous harvesting of new cells as they are not immortalised like cell lines. There is also a problem with inconsistent access to starting material, scalability as they rely on a product which has a finite supply, a risk of variation in permissiveness for the target virus, and a risk of contamination which may introduce adventitious agents [Barrett *et al*., 2009].

1.9.2.2 Human diploid cell vaccine

HDCV was made possible after advances in mammalian tissue culture allowed production of a human diploid cell line WI-38-derived inactivated rabies vaccine (B: Wiktor *et al*., 1978). This was licenced in Europe in 1976 and the U.S.A in 1980. Currently, the main WHO approved vaccine using the HDCV system is Imovax, manufactured by Sanofi. Imovax is a rabies vaccine which is harvested from the human diploid (HD) MRC-5 strain and uses β-propiolactone inactivation to produce inactive virus [Sanofi Pasteur SA, 2019]. The vaccine is freeze dried

upon production and must be stored at 2-8°C. HDCV vaccines have advantages over PCEVC vaccines as they undergo senescence and are non-tumorigenic. They also offer the capability of multiple expansion passages from working and stock cultures which are well characterised and form a closed system [Jordan & Sandig, 2014]. They are also free from all adventitious agents [Ma *et al*., 2015], though can become infected during in vitro propagation. However, they have some disadvantages such as difficulty upscaling, requirement for demanding growth media and difficulties in propagation without the use of serum [Barrett *et al*., 2009].

1.9.2.3 Vero cell rabies vaccine

Verorab, developed by Sanofi Pasteur, is a mammalian Vero cell-based rabies vaccine approved by the World Health Organization [WHO, 2005]. Cell line vaccines are immortalised and theoretically have an infinite life span. However, when kept at low passage numbers they are considered tumorigenic [Furesz *et al*., 1989]. They offer the advantage of a much more scale-able propagation of viruses at lower costs than alternative methods and stable well characterised systems. The Vero cell line is sourced from an African Green Monkey [Osada *et al*., 2014] and widely adopted by medical authorities for use in rabies vaccine production, chosen mainly for its high viral yields [Barrett *et al*., 2009]. Verorab is licenced in over 100 countries and provides effective vaccination to millions. However, as with the other types of vaccines Verorab requires refrigeration at 2-8°C which adds complications when factoring in developing nations which may struggle with this requirement.

Pre-exposure prophylaxis (PrEP) is rabies vaccination prior to exposure, often given to high-risk individuals such as bat handlers, or, rarely at present, distributed to communities in areas with high risk of infection [Kessels *et al*., 2017]. It is also used in preparation to travel to areas with serious rabies presence [Rupprecht *et al*., 2010; Warrell & Warrell, 2015]. The primary model of protection are virus-neutralising antibodies (VNAs). This is two parts, circulating VNAs and then increased response when booster immunization is used post exposure. PrEP simplifies treatment after exposure, circumvents the need for passive immunisation like RIG and reduces the number of vaccine doses needed for protection. In theory it also protects about unrecognised exposures and reduces risks caused by time delays between exposure and PEP treatment (Rupprecht *et al*., 2010; Warrell & Warrell, 2015; WHO, 2013). PrEP is given intramuscularly on days 0, 7, 21 and 28 although shorter schedules have been proposed [Wieten *et al*., 2013]. Intradermal route is seen as economical and acceptable alternative but needs appropriate medical training. The cost effectiveness may be limited by use of only a portion of the vaccine in the vial with the rest binned if not used within a day [Madhusudana & Mani, 2014].

1.9.2.4 Rabies immunoglobulin

PEP under WHO guidelines also involves the administration of anti-rabies immunoglobulin (RIG) which is administered to the site of initial infection where it inhibits the spread of infection [Bakker *et al*., 2008]. There are currently two types of RIG used, human RIG and equine RIG. This type of treatment is unfortunately quite costly and can often be in short supply which leaves many developing nations with only vaccination [Wilde *et al*., 2002], because of this there are currently investigations into the use of monoclonal antibodies instead of RIG which would be able to be manufactured more easily and potentially cost less [Bakker *et al*., 2008; Chao *et al*., 2017; De Benedictis *et al*., 2016].

1.9.2.5 Wildlife immunisation

Wildlife immunisation is a major contributor to reducing human infections and has been suggested as the most effective way of reaching zero human rabies infections by 2030 [Hampson *et al*., 2015]. This immunisation is performed using vaccination that are delivered through edible bait. In Europe, wildlife vaccines make use of attenuated live viruses of the SAD RABV strain [Mähl *et al*., 2014]. In North America, recombinant vaccina or adenoviral vectors are used which express the surface glycoprotein of the SAD or ERA RABV strains [Knowles *et al*., 2009; Maki *et al*., 2017]. The advantages to this method of rabies prevention are significant. The administration of the vaccine is safe with minimal animal exposure to workers, the method has been proven to work effectively and is significantly more cost effective than consistent administration of human PEP.

Though there are effective vaccines and treatments for rabies currently, there are multiple shortcomings that are inherent to them that need rectifying. A novel vaccine is needed to solve current problems with cost, a lack of cross-reactivity with other lyssaviruses and storage both when transported and at medical facilities which may lack cold storage.

1.10 Natural Killer cells

NK cells are innate lymphocytes and perform the important function of being the first line of defence against both tumour cells and viral infections. They comprise of approximately 15% of all circulating lymphocytes [Cooper *et al*., 2001]. The main mechanism by which they aid in viral infections is the cytotoxicity against host cells. NK cell functions that are used to achieve this include engagement of extracellular death receptors and exocytosis of cytolytic granules [Smyth *et al*., 2005]. NK cells also release a range of proinflammatory cytokines which have antiviral activity [Miller, 2002].

NK cells are not homogenous in population and can be split into several subtypes. Of key importance in discussing their relation to lyssaviruses is the receptor CD56. NK populations are defined by their relative expression of the markers CD16 and CD56. CD56 bright and CD56 dim relates to the levels of expression of CD56 on the surface of the NK cell. CD16 bright and CD56 dim NK cells represent at least 90% of all peripheral blood NK cells and are defined as the major circulating subset because of this [Cooper *et al*., 2001]. A maximum of 10% of the circulating NK cell population are CD56 bright. Importantly, mouse NK cells do not express CD56. Importantly, NK cell stimulation using IL-2 or IL-12 increases the cytotoxic activity of all NK cell subsets dramatically [Poli *et al*., 2009]. CD56 bright NK cells are low in cytotoxicity but have significantly higher cytokine production.

CD56 or NCAM is one of the four currently known receptors for lyssaviruses and, as discussed, is present on the surface of NK cells. NCAM can occur in multiple isoforms, the main ones being NCAM-120, NCAM-140 and NCAM-180. NK cells primarily express the NCAM-140 kDa isoform of the receptor [Gunesch *et al*., 2020]. This is important, as it has been found that RABV binds poorly NCAM-120, whereas NCAM-140 and 180 have been demonstrated to be receptors for RABV [Hotta *et al*., 2007]. This provides further evidence as to the potential capacity of RABV to infect NK cells.

NK cells have been shown to be negatively impacted by viral infection before such as in influenza, respiratory syncytial virus and HIV virus infection [Isobe *et al*., 2004; Mao *et al*., 2009; Chehimi *et al*., 1991]. Infection of NK cells can induce immune suppression through mechanisms such as downregulation of cytotoxic functions of the cell or by cell death through triggering of apoptosis which leads to depletion of the NK cell population in the host [van Erp *et al*., 2019]. Based on this, it is clear that viral infection of NK cells does occur and has been used by viruses before as a means to suppress the host immune response to prevent clearance.

The CNS is a key region of infection for lyssaviruses and is considered an immune privileged site due to the lack of dendritic cells [Shi & Ransohoff,2010], however immune cell infiltration

can and does occur during times of injury or infection, which includes NK cells. Importantly, studies on cerebrospinal fluid (CSF) immunophenotyping cells in patients with neuroimmunological diseases found that the majority of NK cells present were CD56 bright which provides more evidence that these cells have the presence and phenotypic receptor type where lyssavirus infection may occur [Han *et al*., 2014]. However, the study states that these cells are transient in nature and leave the CNS following completion of their effector functions, which does mean that while these cells may be present in infection, infection must first be detected for their recruitment from the blood to occur.

NK cells are not known to be stimulated during a lyssavirus infection [Panpanich *et al*., 1992], and presence of NK cells alone is not effective at controlling virus in the CNS [Mastraccio *et al*., 2023]. The study by Panpanich *et al*., 1992, investigated cells with natural killer activity in humans. They found that when compared to 31 normal control patients, rabies patients showed no difference in the number of killer cells. The studies data suggested NK cells are not fully stimulated which may contribute to the virulence of rabies, however the cause of the phenomenon remains unknown. As NK cells have been demonstrated to respond to other viral infections within the CNS, it is unusual that they would not be activated despite the immune privileged site of infection. This could imply that some sort of inhibition of NK cell activation, action or recruitment is occurring during lyssavirus infection, however there is currently no evidence for this and NK cell infection with lyssaviruses has so far not been demonstrated experimentally. Interestingly, NK cells have been identified as important in early lyssavirus infection control in vaccinated patients and forms an important role in successful vaccination [Horowitz *et al*., 2010].

1.11 Project aims

Lyssaviruses pose an important and significant threat to human and animal health in the world today. This project will set out to investigate three key areas which relate to the lyssavirus glycoprotein in order to contribute to the wider knowledge of lyssaviruses and potentially help in the development of future therapeutics.

It is essential that advancements in vaccine science surrounding lyssaviruses are performed. In collaboration with a group developing a novel vaccine, this study will identify the cross-neutralising capabilities of said vaccine in an effort to expand the protection that RABV vaccination offer against through neutralisation of other phylogroups. If the vaccine in question is found to be cross-neutralising, it will be the first step in total lyssavirus and therefore rabies protection.

The only surface protein present on lyssaviruses, the glycoprotein, plays an essential role in the pathogenesis of the virus. It is due to this that it is important that the mechanisms by which this protein functions and is protected from the immune system are studied. It is therefore the aim of this thesis to further elucidate the function of glycans on the surface of the glycoprotein and their impact on infection and neutralisation using a pseudo-virus system.

Though rabies has been known about for thousands of years, some basic questions are still unanswered when it comes to its pathogenesis. For example, the immune avoidance of lyssaviruses has been in contention as it is apparent that the immune privilege of the CNS is not a sufficient explanation for the lack of host response. This study will therefore investigate the potential of lyssavirus infection of NK cells in an attempt to further explain part of the lyssavirus pathogenesis.

Project summary:

- Investigate cross-neutralising capabilities of novel vaccine.
- Determine the impact of glycosylation on lyssavirus glycoproteins and if they impact infection and neutralisation profiles.
- Investigate whether the glycan profiles of divergent lyssaviruses contribute to neutralisation avoidance between phylogroups.
- Infect NK cells using a lyssavirus pseudoviral system to determine if lyssavirus infection of immune cells is possible.

Chapter 2: Establishing pseudo-virus system functionality and determining medoid vaccine cross-phylogroup neutralising potential

2.1 Introduction

2.1.1 Landscape of modern lyssavirus vaccines

Infection with a lyssavirus without any pre or post exposure treatment results in a >99% fatality rate making it the most fatal of all infectious diseases. This cements the importance of vaccination as an essential tool in preventing deaths from rabies. Though vaccines have existed for over a century, there are still tens of thousands of entirely preventable cases of rabies fatalities every year. This is due to a number of complicated factors from inadequate access to healthcare facilities, ineffective government programs, lack of compliance with the complicated schedule involved in vaccine administration, and education. However, the cost of current vaccines is also too high for many developing nations which often have a high burden of rabies infection. This has resulted in a need for a novel vaccine which has a reduced cost and simpler administration schedule.

Protection from RABV covers >99% of lyssavirus infections, however treatment for infection by other phylogroup lyssaviruses is lacking. It is generally accepted that current vaccines against RABV can confer protection against other phylogroup I lyssaviruses, however, attempts at cross-phylogroup neutralisation have shown little to no effectiveness and there are no commercially available vaccines to phylogroup II or III lyssaviruses [Klein *et al*., 2022; Badrane *et al*., 2001; Hanlon *et al*., 2001; Hanlon *et al*., 2005; Malerczyk *et al*., 2014]. In fact, even within phylogroup I the amount of protection provided can vary depending on the lyssavirus [Coertse *et al*., 2023]. With so little protection available against more divergent and less common lyssaviruses there remains a blind spot in treatment should infection with these viruses become more common through potential spill over or host shift events. Therefore, a cross-phylogroup vaccine would be an excellent tool in the treatment of lyssaviruses.

There is currently a push by the WHO in coalition with the United Nations Food and Agriculture Agency (FAO), Global Alliance for Rabies Control (GARC) and the World Organisation for Animal Health (WOAH), to eradicate dog mediated human rabies fatalities by 2030 [WHO, FAO & WOAH. 2018]. The initiative would see the most common route of transmission between animals and humans of lyssaviruses removed through vaccine programmes, wild dog population control, education and awareness, and greater disease surveillance. These efforts would greatly benefit from a cheaper and more affordable cross-neutralising vaccine.

Novel spill-over events are relatively common in RABV, particularly from bats to other non-bat mammals [Johnson *et al*., 2014]. Phylogroup II and III lyssavirus spill-over events are a cause for concern when discussing treatment of rabies [Johnson *et al*., 2010; Shinwari *et al*., 2014]. Currently, reports of infection with lyssaviruses other than RABV are rare, and when measured in humans are exceedingly rare. However, surveillance of lyssaviruses and rabies infections as a whole are poor, it has been suggested that less than 1% of human deaths by suspected rabies cases have been confirmed through modern diagnostics [Cleaveland *et al*., 2002; Suraweera *et al*., 2012] leaving open the possibility of undiagnosed cases of lyssavirus infections that are not RABV mediated, which almost certainly occur. Indeed, one of the most recent cases of EBLV-1 in France was only discovered retrospectively during a study using next generation sequencing on undiagnosed encephalitic patients [Regnault *et al*., 2022]. This report states that this is the first case of lethal encephalitis in western Europe caused by EBLV-1, however it is likely this is only the first which has been discovered, and that other cases have gone undiagnosed and undetected. Examples such as this are evidence that greater surveillance and protections are needed against lyssaviruses as a whole, not just RABV.

This lack of surveillance is further demonstrated by the rapid discovery of novel lyssavirus species in recent years. Though spill-over events are rare they still remain a cause for concern with much of the reasoning behind the lack of phylogroup II and III lyssaviruses successfully migrating to new reservoirs, as RABV has, being unknown [Klein *et al*., 2022]. This uncertainty lends further evidence to the need for a cross-neutralising vaccine.

2.1.2 A novel medoid vaccine

In order to address the need for a cheaper and more effective alternative to current treatments for RABV, a novel medoid vaccine has been developed [Napolitano *et al*., 2020]. The vaccine utilised the Chimpanzee-derived adenovirus-vectored (ChAd) platform. Viral vector vaccines can express the antigen of a virus that immunity against is required, in this case a lyssavirus glycoprotein, and delivers it to the host using a viral vector, in this case a chimpanzee-derived adenovirus. This system has multiple key advantages to current rabies vaccines. The platform has been demonstrated to avoid the issue of pre-existing anti-vector antibodies already present in patients which were developed against human adenovirus serotypes [Colloca *et al*., 2012]. Additionally, the platform has been shown to be safe and produces highly immunogenic responses [Vitelli *et al*., 2017], has a sustained presence in a transcriptionally active form enabling a persistent immune response [Tatsis *et al*., 2007], and can be used in nations which have not developed cold chain processes due to stability methods being available [Vellinga *et al*., 2014]. This vaccine also requires only a single dose to achieve immune correlate of protection levels comparable to three doses of modern commercial vaccines [Napolitano *et al*.,

2020] and is compatible with boosts of other vaccines, making it flexible. The medoid vaccine benefits solve many of the issues currently licensed vaccines incur, however cross-neutralisation would still be very desirable. To achieve this aim, the vaccine was developed using a ‘medoid’ method, where the glycoprotein cloned into the ChAd vector represented the highest average amino acid identity percentage determined from 2060 aligned G protein sequences from both phylogroup I and II lyssaviruses. This average ‘medoid’ sequence (NCBI accession number AGN94271, isolate A11_4583) shared 94% sequence identity to all lyssavirus glycoprotein sequences and thus may produce greater cross-neutralising antibody responses. This cross-neutralising potential was investigated in this chapter.

2.2 Materials and Methods

2.2.1 Cell culture

All cell lines used in experimentation (BHK-21, HEK293T, MDCK, Huh7, VeroE6 and TE671) were maintained in Dulbecco's modified Eagles medium (DMEM) which had 10% foetal bovine serum (FBS) and 1% non-essential amino acids (NEAAs) added. For all experimentation involving DMEM, this formula was used. Both the MDCK and TE671 cell lines were provided by Leah Goulding of the University of Nottingham (UoN) Vet School and Ian Mellor of the UoN School of Life Sciences respectively. All other cell lines were in-house cultures that had been previously obtained and frozen down in liquid nitrogen. Cell lines used have been previously verified as accurate to their stated genotype and all were phenotypically correct and behaved as expected. Cultures were not kept beyond passage 50. Disposal of cells was performed using Distel-Trigene. All cell cultures required incubation at 37°C and 5% CO_2.

Cell culture passage was performed twice a week on Monday and Thursday for experimental purposes. Passage was initiated with a visual inspection of cells under the microscope followed by placement into a class II biological safety cabinet with all equipment sterilised prior to use. Aseptic techniques were followed for all experiments and procedures. The culture media was poured off into Distel-Trigene and cells were washed using sterile PBS. The PBS was then also removed and replaced with Trypsin-EDTA which had been warmed in a water bath set to 37°C. Incubation time and trypsin concentration depended on cell line, for Vero E6 and MDCK cells, 0.20% trypsin was used with a 20-30 minute incubation time at 37°C and 5% CO_2. All other cell lines required 0.05% trypsin and 5 minutes of incubation at 37°C and 5% CO_2.

Following removal, the trypsin was diluted out using fresh DMEM to a volume dependent on the number of cells seeded. Higher amounts of DMEM were used for larger volumes to make counting more consistent. Cells were agitated to reduce clumping and 10µl was pipetted onto a haemocytometer for counting. The suspended cells were then centrifuged at 300xg for 7 minutes with soft acceleration and deceleration selected on a swinging bucket centrifuge. Once cell pelleting was complete, the supernatant was removed, and the cell pellet was resuspended through kinetic agitation. Media was then added to the calculated volume required of the passage and the resuspended cells were added to a new culture flask. Volumes were dependent on several factors such as number of cells needed, health of the cells and volume recovered.

Thawing procedures for cell lines involved the removal of the cell line from liquid nitrogen storage, thawing in a water bath set to 37°C for approximately 1 minute and then slow introduction of DMEM into the culture in order to dilute the freezing media which contained DMSO. The entire culture was added to a T-25 culture flask and incubated overnight at 37°C

and 5% CO_2. Following incubation, the cells were passaged and counted using standard procedure if deemed healthy.

The freezing of cell lines for long term storage involved a standard passage routine, however, after the spin stage and once cells were counted they were resuspended in freezing media (a mixture of 90% FBS and 10% DMSO) and aliquoted at approximately 6 million cells per vial in specialised cryovials. These were then frozen overnight in a -80°C freezer within a Mr Frosty (Thermo scientific) which contained 100% isopropanol, after which they were transferred to liquid nitrogen.

2.2.2 Pseudo-virus system

In order to investigate the aims of this study, a three-plasmid pseudo-virus (PV) system was utilised. The PVs used in this study were CVS-11 (EU352767), EBLV-1 (EU352768), MOKV (HM623780), WCBV (EF614258) and IKOV (JX193798) which had been cloned into the Pi.18 plasmid vector. These were kindly provided by Dr Edward Wright, University of Sussex. A positive control VSV glycoprotein was also used in the Pi.18 plasmid which was used as a non-specific virus control. The MLV gag/pol structural genes were contained within the phCMV5349 plasmid and the firefly luciferase plasmid construct pTG126 was used as a reporter.

2.2.3 Transfection of HEK293T cell line and harvesting of pseudo-particles

Transfection was performed by first seeding HEK293T cells immediately following a passage. Resuspended cells were seeded into a 10 cm^2 culture dish at a density of $1.2x10^6$. This was incubated overnight in standard cell conditions. After incubation, a plasmid mix was produced which contained 2µg of the MLV-*gag/pol* structural plasmid, luciferase reporter plasmid and glycoprotein plasmid of interest. This was made up to a volume of 300µl using Optimem. To this, a mixture of 24µl of polyethyleneimine (PEI) transfection reagent was added to 276µl of Optimem which was then added to the plasmid mix. The final 600µl mixture was incubated at room temperature for one hour. The HEK293T cells were prepared for transfection through the removal of the DMEM media and replacement with Optimem. The incubated plasmid mix was then pipetted onto the HEK293T cells and incubated for 5-6 hours at 37°C and 5% CO_2. The Optimem was then removed and replaced with fresh DMEM and the cells were incubated for 72 hours at 37°C and 5% CO_2.

Pseudo-particles were harvested after 72 hours using a 10mL syringe to remove the supernatant which was passed through a 0.45µm filter. PVs were stored at 4°C for up to 2 weeks. For longer storage, some PVs were aliquoted and frozen at -20°C and remained infectious for at least 6 months.

2.2.4 Infection assay

Cells for infection (such as BHK-21s) were seeded into a 96 well white bottom clear lid plate at a density of $2x10^4$ cells per well 24 hours prior to the infection time. Cells were incubated overnight in the standard conditions. Once ready to infect, the DMEM was removed from cells and 100µl aliquots of pseudo-virus was added in triplicate to the wells. Approximately 5 hours of incubation following infection the cells were topped up with 150µl of DMEM before a further 72-hour incubation time.

Once incubation was complete, the DMEM medium was removed, and cells were lysed with 50µl of a cell lysis solution (Promega). Cells were incubated on a rocking platform for 15 minutes at room temperature while lysing and finally agitated using a plate shaker. The infection assay was measured in luminescence using a Fluostar Omega luminometer, which added 50µl of luciferase substrate (Promega) automatically to each well. Optical gain was set to 3,600 as standard, being reduced occasionally for VSV-G positive controls due to its high infectivity.

2.2.5 Neutralisation assay

Neutralisation assays followed the same procedure as infection with the exception of an additional neutralisation stage. This was performed on the infectious PVs prior to their addition to the target cells. For single point neutralisations, 270µl of PV was incubated for 1 hour with 30µl of 1:100 NHP sera. A 100% infection control was produced using 270µl of PV and 30µl of PBS. All serum was heat inactivated at 56°C for 30 minutes. Following incubation, the PVs were added the appropriate wells in a 96 well plate as with an infection assay, 100µl per well in triplicate. The remaining stages of the infection assay were then performed.

2.2.6 Lyssavirus glycoprotein sequencing

Lyssavirus glycoproteins were sequenced prior to their use to confirm that no unexpected changes had been introduced into the plasmids during their processing. Sequences were as expected and showed no differences from their original sequence reported by Dr Edward Wright, and Dr Emma Horncastle in her thesis. Sequencing was performed by Source Bioscience.

2.2.7 Vaccine dosages

For the Macaque sera, animals were vaccinated intramuscularly with a single dose of $5x10^{10}$vp of ChAd155-RG (for group 1 and 2), or three doses (day 0, 7 and 21) of RABIPUR (group 3). For both the ChAd155-RG and RABIPUR vaccines, half the standard human dose was used with an injection volume of 0.3mL and 0.5mL respectively.

2.2.8 Animal ethics

As live animals were used to obtain the sera used in this study, ethical approval was sought by our collaborators at Aptuit srl (Verona, Italy), prior to the completion of their work. As stated in the publication, all animal protocols described by the experimental procedures were reviewed and approved by internal Aptuit Committee on Animal Research and Ethics, and under the authorisation issued by the Italian Ministry of Health (nr.984/2015-PR). Experiments were in compliance with both EU regulations (Directive 2010/63/EU) and national regulations (Italian Legislative Decree 26/2014, and French decree No. 2013-118).

2.2.9 Statistical analysis

Data was analysed using GraphPad Prism 10. Normality was assumed due to the low N value and data was statistically analysed using a one-way or two-way ANOVA followed by Tukey multiple comparisons post hoc test. A value of $P \leq 0.05$ was considered to indicate a statistically significant difference.

2.3 Results

2.3.1 Pseudo-virus system establishment

In order to establish the functionality of the pseudo-virus system, infection assays were performed in the BHK-21 cell line. Lyssavirus glycoprotein plasmids from each phylogroup were selected, CVS-11 (RABV) and EBLV-1 being phylogroup I, MOKV phylogroup II, and WCBV and IKOV being phylogroup III. These glycoprotein plasmids were then used to generate PV particles as described in the methods and used to infect BHK-21 cells as shown in figure 2.1.

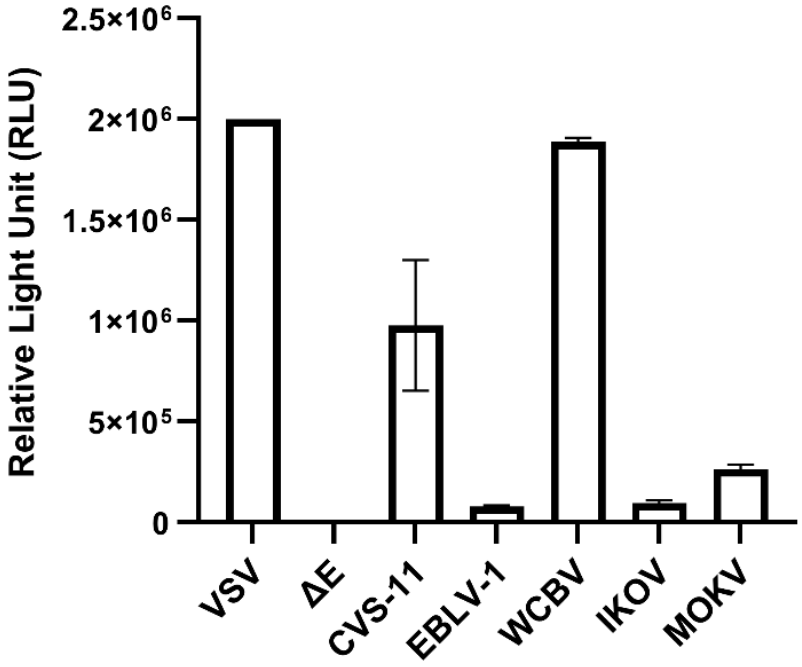


Figure 2.1: Infection assay performed in BHK-21 cells using lyssavirus pseudo-virus (PV) particles. PVs used were CVS-11 and EBLV-1 from phylogroup I, MOKV from phylogroup II and WCBV and IKOV from phylogroup III. Infectivity was measured through relative luminescence units (RLU) on a luminometer at 3600 gain. Infection assay was performed in triplicate and repeated a second time. A positive control of VSV was used in conjunction with the lyssavirus plasmids, and a negative non-enveloped PV (ΔE) was deployed as a negative control. Additionally, a cells only triplicate well was also used. Data is presented as mean±SD (assumed to be normally distributed) and statistically analysed using a one-way ANOVA followed by Tukey multiple comparisons post hoc test. All significance data is reported in Supplementary Table S.1.

The assay showed successful infection of the BHK-21 cells with each lyssavirus PV. The WCBV strain was most infective and the RABV lab strain CVS-11 was also considerably more infective than the other PVs, consistent with previous infection assays done with these plasmids. These results remained consistent in biological repeats. WCBV and CVS-11 were both re-transformed and re-transfected to rule out any contamination increasing their infectivity.

2.3.2 Infectivity comparison between BHK and Vero cells

With the successful infection of BHK-21 cells, a second cell line was chosen to examine. Vero E6 cells were infected alongside BHK-21 cells to compare their infectivity as shown in figure 2.2.

Results showed BHK-21 cells were consistently more susceptible to infection across every PV with the exception of IKOV, with all PVs producing significantly higher RLU values in comparison. A similar pattern of infectivity was seen in the BHK-21 cells compared to the initial infectivity experiments. In both cell lines WCBV infected highest followed by CVS-11. IKOV had very low infectivity in Vero E6 cells. Based on this data, BHK-21 cells were chosen to be the primary cell line in experiments going forward.

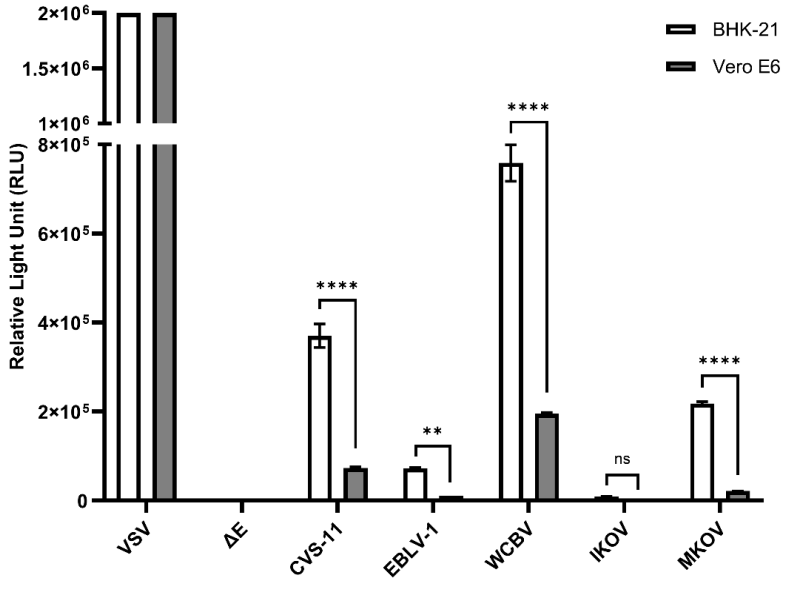


Figure 2.2: Comparative infection assay between BHK-21 and Vero E6 cell lines both infected with a panel of lyssavirus PVs CVS-11, EBLV-1, WCBV, IKOV and MOKV. Infectivity was measured in RLU. Results showed BHK-21 cells to be more susceptible to infection than Vero E6. Assay was performed in triplicate and repeated. A positive control of VSV was used in conjunction with the lyssavirus plasmids, and a negative non-enveloped PV (ΔE) was deployed as a negative control. Data is presented as mean±SD (assumed to be normally distributed) and statistically analysed using a two-way ANOVA followed by Tukey's multiple comparisons post hoc test (****P<0.0001 compared to BHK-21; ns=not significant).

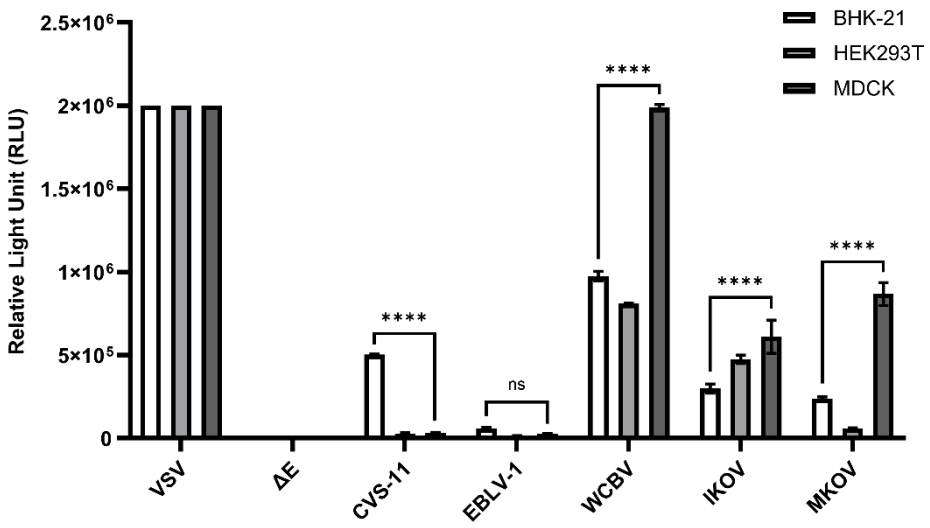
2.3.3 Lyssavirus PVs can infect numerous cell lines

Several other cell lines were examined to determine their susceptibility to infection and provide insights into the tropism of the lyssavirus glycoprotein. Infectivity assays were performed on HEK293T, MDCK, Te671 and Huh7 cell lines, however Huh7 was infected with lyssavirus PVs made using a two-plasmid system using PNL4.3. Data for all infection assays are shown in figure 2.3.

The results of figure 2.3A demonstrate that BHK-21s are the most susceptible to CVS-11 and EBLV-1, both phylogroup I lyssaviruses. WCBV, MOKV and IKOV all produce their highest RLU values in the dog kidney cell line MDCK, greater than both HEKs and BHKs, though the infectivity of both phylogroup I lyssavirus PVs is low. The results also continue displaying the high infection trend of WCBV, with it infecting the highest in all cell lines.

Figure 2.3B is an infectivity assay with the human cell line Te671 compared to BHK-21. This figure again confirms the trends found in BHK-21 experiments with WCBV infecting highest followed by CVS-11. With the exception of CVS-11, cell line comparison shows somewhat similar results. CVS-11 is significantly reduced in infectivity in Te671s when compared to the BHK-21 cell line.

The final figure, 2.3C, shows a Huh7 cell line infection with a PNL4.3 two-plasmid system-based lyssavirus PV. This infection assay displays the same pattern as BHK-21 with WCBV having the highest RLU followed by CVS-11. All other PVs do infect but produce much lower RLU values compared to CVS-11 and WCBV. There was no comparison cell line infected during experimentation with Huh7.



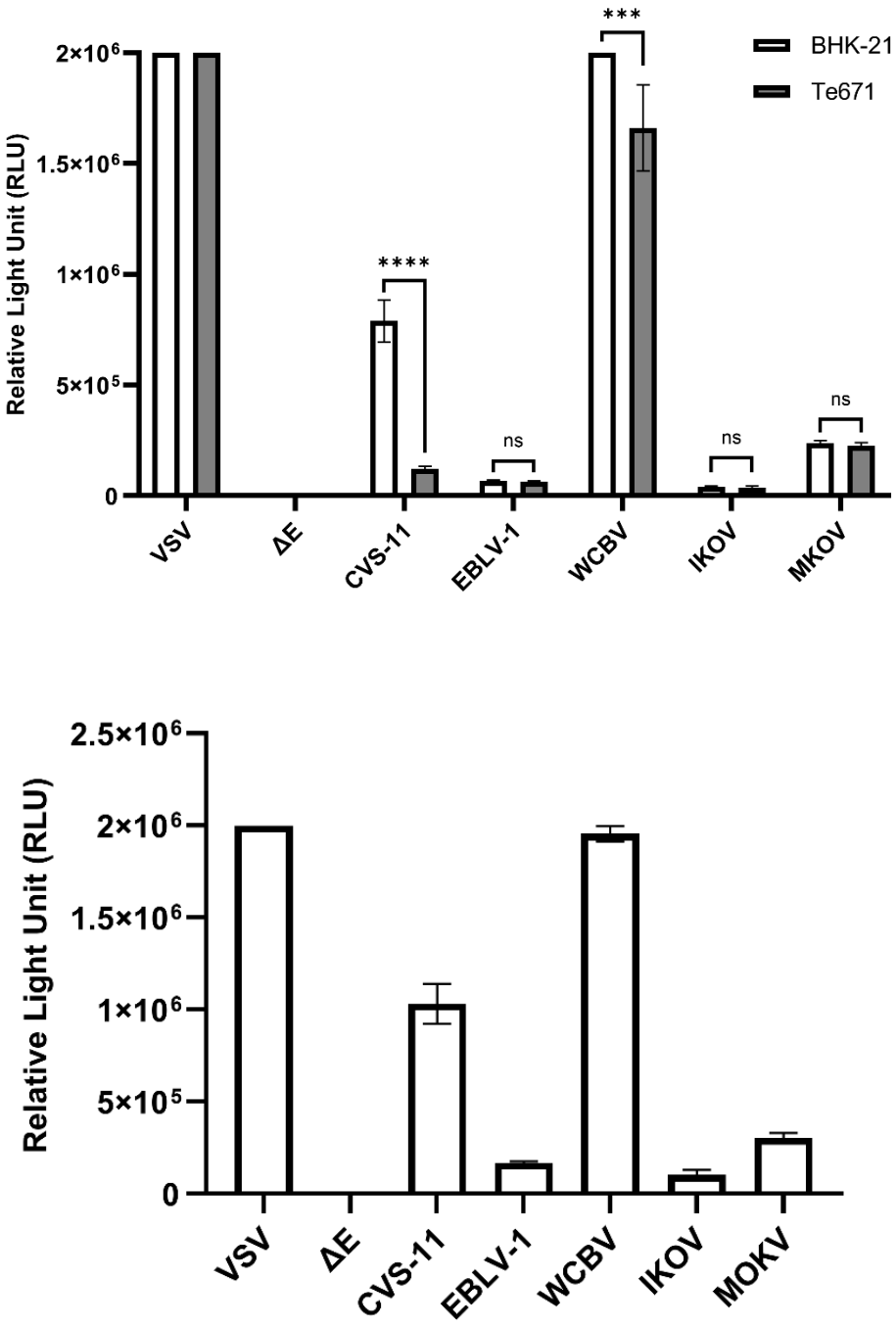


Figure 2.3: Comparison of cell lines (A): BHK-21, HEK293T and MDCK (B): BHK-21 and Te671 (C): Huh7s in an infection assay using a lyssavirus PV panel containing CVS-11, EBLV-1, MOKV, WCBV and IKOV. All experiments were performed in triplicate and repeated twice. All PVs were obtained from the same transfection cycle to provide more accurate comparison. A positive control of VSV was used in conjunction with the lyssavirus plasmids, and a negative non-enveloped PV (ΔE) was deployed as a negative control. Data is presented as mean±SD (assumed to be normally distributed) and statistically analysed using either a two-way ANOVA followed by Tukey's multiple comparisons post hoc test (A: ****P<0.0001 compared to BHK-21; ns=not significant. B: ****P<0.0001, ***P<0.001 compared to BHK-21) or a one-way ANOVA followed by Tukey multiple comparisons post hoc test (C). All significance data is reported in Supplementary Table S.2.

2.3.4 Lyssavirus panel optimisation and validation

The panel of lyssaviruses used in these experiments had been validated and optimised by a previous lab member as part of their thesis [Horncastle, 2019]. During this, the panel was developed from donated plasmids. The plasmid vector was optimised; to achieve this, the panel was cloned into a new vector, pI.18, which showed marked improvement over the old pcDNA3.1 system in infectivity assays, as can be seen in the figure H.1.

In addition to this, three of the pseudo virus panel members were tested to determine the optimal amounts of plasmids to add during transfection. Though the pseudo viruses tested were not used in this study, the results were used to inform the volume of plasmid to add. Ultimately, this study settled on 2ng of glycoprotein plasmid in order to standardise with the VSV control.

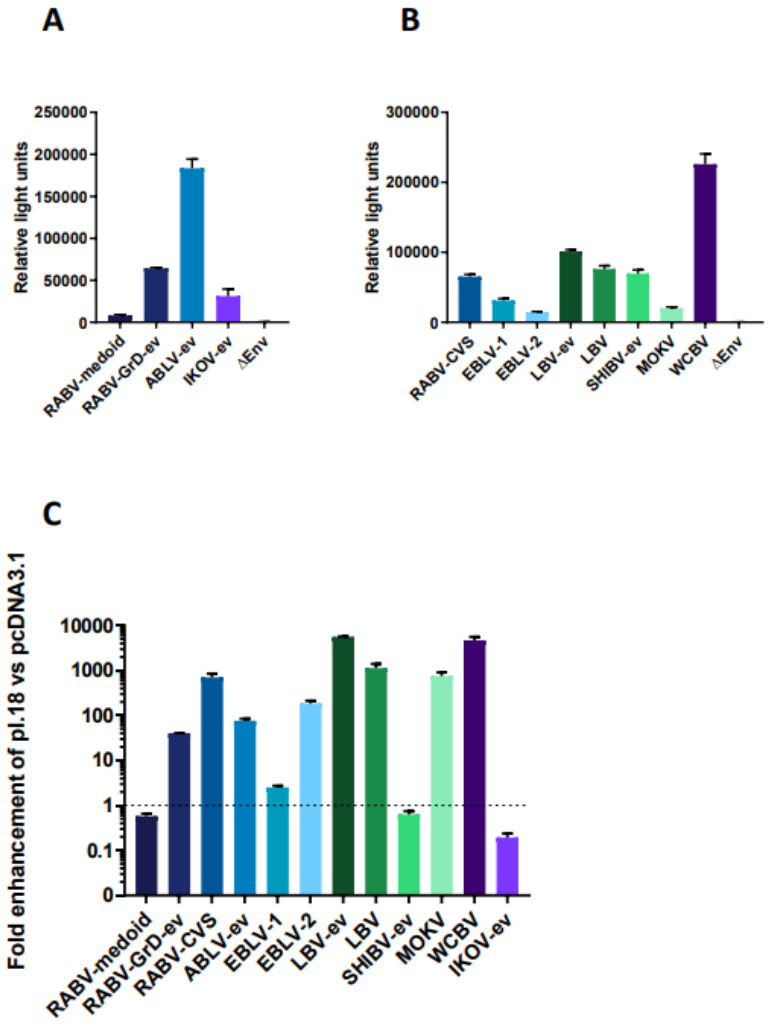


Figure H.1: Figure sourced from the thesis of Horncastle 2019. A panel of lyssavirus pseudo viruses were produced in the pI.18 plasmid system. BHK-21 cells were infected and read at A) 3,600 optical gain and B) 2,000 optical gain measuring luminescence. C) A third experiment was performed with a panel in both the pI.18 and pcDNA3.1 systems. Infection values were normalised against the corresponding pcDNA3.1 values and are expressed as fold enhancements. Error bars indicate standard deviation.

2.3.5 Cross-phylogroup reactive potential of novel medoid vaccine

Having selected BHK-21 as the primary cell line for experimentation, neutralisation assays were performed using them. Lyssavirus PVs were first neutralised using sera provided by Dr Allessandra Vitelli and Dr Alfredo Nicosia of ReiThera SRL, Italy, which was taken from non-human primates that had gone through a course of vaccination using a novel medoid vaccine ChAd155 and a standard RABIPUR course for comparison (days 0, 7 and 21).

Figure 2.4 shows four different lyssavirus PVs which were used in the neutralisation assay. The RABIPUR sera neutralised CVS-11 as expected but did not show any significant neutralising properties to the phylogroup II and III lyssaviruses except within MOKV, where a significant amount of neutralisation occurred by week 5. In comparison, by week 5 in animals vaccinated with ChAd155, MOKV, WCBV and CVS-11 all showed significant neutralisation. CVS-11 had reached complete neutralisation, with WCBV approaching 50%. Additionally, the potency of neutralisation increased further at the week 48 boost of ChAd15 for MOKV and WCBV, with the phylogroup III lyssavirus IKOV being significantly neutralised by the sera to below 50% infectivity.

EBLV-1 was tested after the conclusion of this study by a third party and the data published alongside figure 2.4. The results showed that neutralisation was achieved however through the use of neutralisation curves, showed that it was less potent than against CVS-11. Additionally, lower concentrations of ChAd155 sera were required to neutralise EBLV-1 and EBLV-2 than RABPIPUR.

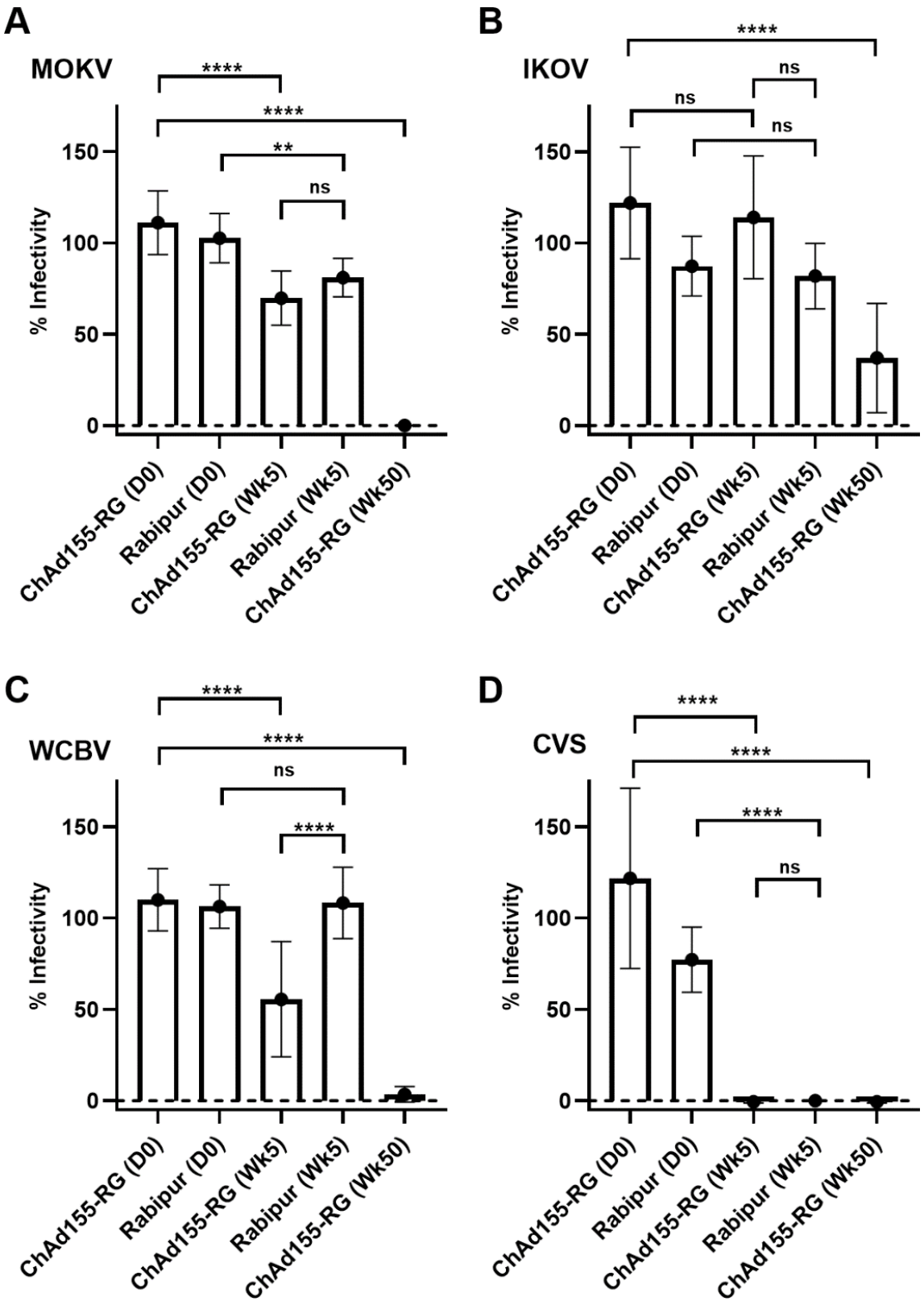


Figure 2.4: Figure of published data produced by Napolitano et al., 2020, showing point neutralisations of (A): MOKV (B): IKOV (C): WCBV and (D): CVS-11. Neutralisations were performed in triplicate with the standard VSV, ΔE and cells only controls in place. Results were read as RLU and reported as a percentage of infectivity which had been normalised against a PBS and PV standard which represented 100% infectivity.

2.3.6 Amino acid sequence identity of the chosen pseudo virus glycoproteins

Amino acid sequence identity was examined using the Geneious Prime software to produce heatmaps based on sequence alignments, shown in figure 2.5. The maps demonstrate the similarity in sequences and is comparable to the reference sequences seen in table 2.2. As can be seen from the data, The medoid strain correlates highest with the CVS-11 and EBLV-1 strains. The shared percentage identity between the divergent lyssaviruses is less prevalent, and in comparison with the other strains does not appear to be significant.

A

	CVS-11	EBLV-1	IKOV	Medoid	MOKV	WCBV
CVS-11		73.86%	52.66%	91.19%	62.22%	56.84%
EBLV-1	73.86%		54.53%	73.47%	64.66%	56.08%
IKOV	52.66%	54.53%		53.60%	55.78%	53.81%
Medoid	91.19%	73.47%	53.60%		65.91%	56.84%
MOKV	62.22%	64.66%	55.78%	65.91%		57.21%
WCBV	56.84%	56.08%	53.81%	56.84%	57.21%	

B

	CVS-11	EBLV-1	IKOV	Medoid	MOKV	WCBV
CVS-11		260	169	321	219	150
EBLV-1	260		175	371	258	148
IKOV	169	175		172	179	143
Medoid	321	371	172		263	150
MOKV	219	258	179	263		151
WCBV	150	148	143	150	151	

Figure 2.5: Heatmaps generated using the Geneious Prime software showing an alignment of sequences obtained through sangar sequencing of the lyssavirus glycoprotein used in the infection assays in figure 2.4.

2.4 Discussion

The aims of this study were to confirm the infectivity of lyssavirus plasmids in preparation for their use in assays, expand on the currently known repertoire of cell lines which the PVs have infected, and finally to confirm the potential cross-neutralising capabilities of a novel medoid vaccine. Through the use of a PV system these aims were achieved. The lyssavirus plasmid infectivity was confirmed in multiple cell lines. The plasmids behaved as expected based on previous infectivity assays performed [Horncastle, 2019], with similar patterns of infection observed, such as WCBV being the highest performing in terms of infectivity. The volumes of plasmid used were based on the same previous work which used optimisation to produce efficient levels of infectivity.

Though there are 17 officially recognised lyssaviruses, a panel of 5 was chosen for these experiments. The chosen lyssaviruses were selected to represent each phylogroup. CVS-11 is commonly used in lyssavirus research as a lab strain and represents the most common rabies virus, RABV, which made it an obvious choice for selection. EBLV-1 has the second most sequences available on Genbank and was included as a representative of phylogroup I that was not the target of commercial vaccines. MOKV was chosen due to its history of human infection and as a phylogroup II lyssavirus. IKOV and WCBV both represent two of the three phylogroup III lyssaviruses, IKOV is particularly distant in relation to the RABV genome and was therefore an interesting extreme for a cross-phylogroup neutralising study.

Rabies virus infection of cell lines such as BHK-21, Vero, TE671 and MDCK has been demonstrated previously [Nitschel *et al*., 2021; Trabelsi *et al*., 2019; Rincón *et al*., 2005], however it appears the infection of huh7 and HEK293T cells with rabies PVs is novel in the literature and expands upon the known tropism of the lyssavirus glycoprotein. Cell lines were selected for a variety of reasons. In the case BHK-21 and Veros, both cell lines are used in the production RABV vaccines. They have also been demonstrated to be infectable with the PVs used in this study [Horncastle, 2019] which made them suitable benchmarks for any anomalies. As BHK-21 and Vero E6 are both kidney cell lines, the MDCK kidney cell line was chosen in order to infect a cell type from a reservoir species, in this case canine. HEK293ts were used both in transfection of the PVs but also in an infection experiment. They were chosen for their kidney morphology and to provide an insight into infection of cells used to generate PVs. The results did not show any apparent adaptational preference in infection as the HEKs were outperformed by BHK-21s and MDCKs. TE671 cells are a cell line which has been identified as Rhabdomyosarcoma cells, a cancer of soft tissue such as muscle. This makes it an interesting cell line of choice due to the importance of the muscle cells in lyssavirus pathogenesis. Huh7

infection was performed due to the lack of literature surrounding the subject and because of their human origin.

2.4.1 Cell line infection assay analysis

Multiple cell lines including BHK-21, VeroE6, MDCK, TE671 and HEK293t's were examined for their susceptibility to infection by a panel of pseudo viruses. The results showed there to be significant differences in infection susceptibility between cell lines. This could be due to several factors, such as presence and prevalence of surface receptors, and innate cellular machinery. As cell lines are immortalised cells, they contain many differences when it comes to their surface proteins.

BHK-21 cells performed well in terms of susceptibility, and were most responsive to infection with CVS-11 out of all the cell lines. Interestingly, MDCK cells were more susceptible to infection from divergent lyssavirus glycoproteins, such as WCBV, IKOV and MOKV, when compared to BHK-21 and HEK293ts. MDCK cells originate from dog kidneys and are therefore not generally associated with any neural-specific surface proteins such as those targeted by lyssaviruses. However, their relation to canines is notable. Though the cell lines themselves may contain different levels of receptors, with the exception of TE671s [Syapin et al., 1982], there is no evidence to suggest any of the other cell lines contained unusually high levels of a lyssavirus receptors.

TE671s are the exception within these cell lines. They have been shown to express functional nicotinic cholinergic receptors amongst other neuronal like properties such as spontaneous hyperpolarisation [Syapin et al., 1982]. This would suggest that they might be more susceptible to lyssavirus infection, however it seems that BHK-21s were either equal or more susceptible. This could imply that the receptor presence had little impact, or that perhaps BHK-21s contain other receptors not yet characterised fully. These observations suggest that a deeper study into these cell lines and their receptors could be done in the context of lyssaviruses, and while these findings may be interesting, they were not the goal of this analysis and therefore was not performed.

Results showed a significant difference between cell lines, but also between which pseudo virus infected better. WCBV, for example, far outperformed all other lyssavirus glycoproteins in terms of infectivity. This could possibly be due to binding affinity, with the WCBV glycoprotein being more efficient at binding, though without further research into these differences this can only be speculation.

2.4.2 Medoid vaccine in producing cross-neutralisation

The medoid vaccine was produced using 2060 aligned glycoprotein sequences from phylogroup I and II lyssaviruses in order to produce a glycoprotein sequence which had higher amino acid similarity to all members of the genus. This was done in order to increase the potential cross-protective breadth of the vaccine. The approach produced a lyssa glycoprotein which shared 94% average sequence identity with the other glycoproteins [Napolitano *et al*., 2020]. The approach was also performed for the N protein which had a 98% similarity to 2541 N proteins. The inclusion of both of these sequences in the vaccine is interesting as a number of studies suggest that both the G and N proteins are suitable candidates for inclusion in vaccines [Drings *et al*., 1999]. The glycoprotein, being the only surface protein present in lyssavirus, is an obvious target but the N protein has been shown to stimulate T cell production [Ertl *et al*., 1991]. The N protein is also more conserved than the glycoprotein [Dietzschold *et al*., 1987], indeed, the N protein conservation between RABV and the most divergent lyssaviruses are much higher than that of the glycoprotein as can be seen in tables 2.1 and 2.2, with the RABV G protein having a 34% amino acid sequence identity with the most divergent lyssa, IKOV, and the N protein having 74%. The medoid N and G proteins share higher percentage identities when comparing RABV and IKOV, with the N protein increasing by 2% and the glycoprotein by 12%. The N protein has also been shown to stimulate the production of virus neutralising antibodies induced by classical vaccines [Dietzschold *et al*., 1987]. Unfortunately, however, the vaccine study found that inclusion of the N protein sequence resulted in decreased anti-G IgG titre and decreased VNA stimulation *in vivo* in comparison to the G protein vaccine alone. The G protein only vaccine was found to be effective and therefore despite the evidence of its potential antigenic nature, the N protein was deemed dispensable.

This leaves the glycoprotein change as the main driving factor behind the potentially cross-neutralising capability of this vaccine. As can be seen in table 2.2, the medoid vaccine sequence had significant increases to sequence identity when compared to RABV, the classical vaccine target. This is likely the source of the cross-neutralisation however the exact mechanisms surrounding the changes are not understood and require more investigation. When aligning the medoid sequence with the ICTVs classical rabies sequence the amino acid changes seen do not appear to correlate to any ‘antigenic region’ that has been described [Evans *et al*., 2018].

Table 2.1: Lyssavirus nucleoprotein sequence identity comparison using the ICTV standard reference sequences. Aligned and produced in Geneious Prime.

Phylogroup		IKOV	LLEBV	WCBV	MOKV	LBV	SHIBV	EBLV-2	IRKV	BBLV	ARAV	EBLV-1	TWBLV	DUVV	KHUV	Medoid	RABV	ABLV	GBLV
3	IKOV		87.96	78.71	76.13	78.89	76.13	74.41	73.98	73.98	75.05	75.48	75.27	77.38	73.98	75.78	73.98	74.44	73.98
	LLEBV	87.96		80	76.82	79.11	77.47	75.54	76.18	76.18	77.25	76.18	76.61	79.16	77.25	78	75.75	78.22	76.61
	WCBV	78.71	80		80.86	83.56	82.37	79.78	81.29	80.22	81.72	80.43	81.51	84.26	79.35	81.56	79.35	81.33	81.94
2	MOKV	76.13	76.82	80.86		90.89	85.74	76.94	80.67	79.41	81.65	80	80.5	84.48	79.25	82	78.69	83.11	79.79
	LBV	78.89	79.11	83.56	90.89		92	82.44	87.11	84.89	87.56	86.22	86.89	86.67	84	84.89	84	83.78	84.22
	SHIBV	76.13	77.47	82.37	85.74	92		77.36	82.56	80.88	83.33	82.95	83.86	87.14	80.92	86	82.49	84.44	80.84
1	EBLV-2	74.41	75.54	79.78	76.94	82.44	77.36		83.4	86.97	85.86	84.63	82.85	86.25	88.05	88.44	83.97	87.11	86.11
	IRKV	73.98	76.18	81.29	80.67	87.11	82.56	83.4		85.08	87.34	89.05	88.03	90.24	84.66	87.56	82.91	87.78	84.63
	BBLV	73.98	76.18	80.22	79.41	84.89	80.88	86.97	85.08		89.45	86.95	85.29	88.91	89.29	89.78	84.81	90.44	88.63
	ARAV	75.05	77.25	81.72	81.65	87.56	83.33	85.86	87.34	89.45		89.03	88.82	91.57	89.45	90.44	85.65	91.11	88.61
	EBLV-1	75.48	76.18	80.43	80	86.22	82.95	84.63	89.05	86.95	89.03		90.11	92.68	86.53	88.89	84.6	88.22	86.74
	TWBLV	75.27	76.61	81.51	80.5	86.89	83.86	82.85	88.03	85.29	88.82	90.11		94.01	86.16	88.89	84.6	89.33	85.26
	DUVV	77.38	79.16	84.26	84.48	86.67	87.14	86.25	90.24	88.91	91.57	92.68	94.01		89.14	88.89	87.58	88.67	89.14
	KHUV	73.98	77.25	79.35	79.25	84	80.92	88.05	84.66	89.29	89.45	86.53	86.16	89.14		91.11	86.5	90.44	88.63
	Medoid	75.78	78	81.56	82	84.89	86	88.44	87.56	89.78	90.44	88.89	88.89	88.89	91.11		98.44	90.67	93.11
	RABV	73.98	75.75	79.35	78.69	84	82.49	83.97	82.91	84.81	85.65	84.6	84.6	87.58	86.5	98.44		89.56	89.24
	ABLV	74.44	78.22	81.33	83.11	83.78	84.44	87.11	87.78	90.44	91.11	88.22	89.33	88.67	90.44	90.67	89.56		93.11
	GBLV	73.98	76.61	81.94	79.79	84.22	80.84	86.11	84.63	88.63	88.61	86.74	85.26	89.14	88.63	93.11	89.24	93.11	

Table 2.2: Lyssavirus glycoprotein sequence identity comparison using the ICTV standard reference sequences. Aligned and produced in Geneious Prime.

Phylogroup		IKOV	LLEBV	WCBV	MOKV	LBV	SHIBV	TWBLV	IRKV	EBLV-1	EBLV-2	ARAV	DUVV	BBLV	KHUV	Medoid	RABV	ABLV	GBLV
3	IKOV		48.33	35.19	36.33	48.76	36.7	37.68	37.2	37.29	36.74	36.6	45.34	35.81	36.38	46.49	34.76	48.11	36.51
	LLEBV	48.33		35.54	34.25	44.11	34.75	35.5	35.96	35.42	36.73	35.52	46.37	36.36	36.69	47.35	35.2	46.69	36.03
	WCBV	35.19	35.54		38.1	53.14	38.72	35.94	36.54	36.59	36.98	36.98	48.69	36.98	36.38	50.09	34.72	51.33	36.2
2	MOKV	36.33	34.25	38.1		76.25	52.92	44.21	44.24	44.51	43.48	44.85	55.53	43.97	44.49	59.16	44.17	57.14	45.96
	LBV	48.76	44.11	53.14	76.25		78.35	55.36	58.62	58.43	55.56	55.75	56.51	56.9	56.7	58.05	56.7	58.24	59
	SHIBV	36.7	34.75	38.72	52.92	78.35		42.9	43.04	44.83	44.07	44.07	57.22	44.07	44.63	58.97	42.88	60.19	45.75
1	TWBLV	37.68	35.5	35.94	44.21	55.36	42.9		53.59	57.75	55.38	54.01	66.98	53.89	54.23	66.98	49.36	66.86	52.33
	IRKV	37.2	35.96	36.54	44.24	58.62	43.04	53.59		61.88	59.63	60.43	73.92	58.87	59.15	69.08	50.85	71.62	56.48
	EBLV-1	37.29	35.42	36.59	44.51	58.43	44.83	57.75	61.88		62.84	64.91	78.05	60.6	61.17	71.56	53.7	72.19	58.89
	EBLV-2	36.74	36.73	36.98	43.48	55.56	44.07	55.38	59.63	62.84		67.05	77.86	70.33	72.62	76.15	57.41	75.62	62.32
	ARAV	36.6	35.52	36.98	44.85	55.75	44.07	54.01	60.43	64.91	67.05		77.49	67.34	65.37	75	56.31	76	62.09
	DUVV	45.34	46.37	48.69	55.53	56.51	57.22	66.98	73.92	78.05	77.86	77.49		76.74	75.98	70.61	68.11	71.05	73.36
	BBLV	35.81	36.36	36.98	43.97	56.9	44.07	53.89	58.87	60.6	70.33	67.34	76.74		70.7	77.1	58.59	75.81	64.99
	KHUV	36.38	36.69	36.38	44.49	56.7	44.63	54.23	59.15	61.17	72.62	65.37	75.98	70.7		78.24	58.89	76.76	63.95
	Medoid	46.49	47.35	50.09	59.16	58.05	58.97	66.98	69.08	71.56	76.15	75	70.61	77.1	78.24		94.66	74.81	79.96
	RABV	34.76	35.2	34.72	44.17	56.7	42.88	49.36	50.85	53.7	57.41	56.31	68.11	58.59	58.89	94.66		72.76	60.5
	ABLV	48.11	46.69	51.33	57.14	58.24	60.19	66.86	71.62	72.19	75.62	76	71.05	75.81	76.76	74.81	72.76		81.14
	GBLV	36.51	36.03	36.2	45.96	59	45.75	52.33	56.48	58.89	62.32	62.09	73.36	64.99	63.95	79.96	60.5	81.14	

A recent study performed using the purified chick embryo vaccine (PCEC) flury strain of RABV (RABIPUR) showed that it was able to induce an antibody response that fully neutralised non-RABV phylogroup I lyssaviruses EBLV-1&2, ALBV, BBLV and DUVV. Interestingly, the study also showed neutralising responses to MOKV, a phylogroup II lyssa [Malerczyk *et al*., 2014]. Unfortunately, the reasoning behind this cross-neutralising potential was not able to be determined conclusively and an attempted reproduction of the results using a different strain of MOKV was unable to replicate the findings. These results were somewhat replicated here in that MOKV was able to be partially neutralised by NHP serum taken 5 weeks after vaccination with RABIPUR. Further investigation into this phenomenon should be performed. The original assumption of the RABIPUR study was that, due to some amino acid changes in the MOKV antigenic regions II and III, the strain became susceptible to neutralisation. When the MOKV and medoid sequence used in this study are aligned there are some residues which are conserved in these regions, however this is also true of the classical RABV sequence. Unfortunately, the exact residue changes are not stated in the RABIPUR study and can therefore not be corroborated here.

The single primer dose of the medoid vaccine ChAd155-RG demonstrated an increased neutralisation to the divergent lyssaviruses WCBV and MOKV by week 5 both of which were neutralised following the week 48 boost. IKOV was neutralised to a significant degree by week 50, but showed less susceptibility which is to be expected as the most divergent of the genus, though any neutralisation at all is novel for a phylogroup I vaccine. Partial neutralisation of MOKV was also achieved by the RABIPUR vaccine however there was no week 48 boost sera available to identify if this neutralisation would follow the same pattern as the ChAd155 producing full neutralisation after the boost.

2.4.3 Alternative potential methods of cross-neutralisation

Monoclonal antibody (mAb) cocktails have been developed which show potential for RABV treatment. Antibodies are currently being explored as a potential replacement for the expensive and scarce RIG system which is recommended by the WHO to be used with all category III exposures. These cocktails of mAbs help avoid viral escape through antigenic drift. Unfortunately, they do not solve cold chain issues and so far they do not show promise in cross-neutralisation between phylogroups, only between strains [Franka *et al*., 2017].

Other novel vaccination methods are being developed to help eradicate rabies infections. Protein vaccines provide immunity through the administration of recombinant protein into the patient such as the rabies glycoprotein. To induce any neutralising antibody the glycoprotein must be in its trimeric form, correctly folded and glycosylated [Astray *et al*., 2017]. The disadvantage with protein vaccines is their low immunogenicity when compared to other systems, requiring

the use of adjuvants to produce the required immune response. Additionally, they require careful purification which does not fulfil the requirement of novel rabies vaccine to be cost effective [Fooks *et al*., 2019].

Genetic vaccines can consist of both DNA and RNA vaccines and are used to deliver a gene or transcript. They only include the required subunit of the virus and due to the fact the host cells produce the protein it is ensured that the correct folding occurs and post-translational modifications are produced [Ertl, 2019]. There has been an mRNA vaccine for rabies which has reached human testing stages, proven to be safe with only mild side-effects. Unfortunately, the immune response was also short lived, becoming significantly reduced after just 1 year, reducing its usefulness as a PrEP treatment [Alberer *et al*., 2017].

Genetically modified virus vaccines modify viral genes in order to produce attenuated rabies viruses which can be used in vaccination. As the matrix gene has been found to be essential in virion production and removal of it induces replication deficient lyssaviruses, it is an ideal candidate for this type of treatment [Cenna *et al*., 2009]. Indeed, the same study found that removal of the matrix gene also elicits a high titre of neutralising antibodies. Genetic modification can also be used to increase the number of genes, for example the glycoprotein, which produces a higher level of surface expression potentially making the virus more immunogenic [Hosokawa-Muto *et al*., 2006].

Adjuvant use is standard protocol in many vaccines and their usefulness extends to rabies vaccines. Currently there are two main adjuvants which have been identified for their use against lyssaviruses, PIKA and CV8102, both of which have shown in studies to improve antibody responses when administered alongside a vaccine [Kalimuddin *et al*., 2017; Doener *et al*., 2019].

Though these novel vaccine ideas potentially improve upon some of the issues of the current rabies treatments, none are in the running the become a replacement and at present they lack the cross-neutralising capabilities which are important moving forward in lyssavirus vaccines. It appears that the medoid method of vaccine production is the most promising in that regard with a reduction in cost, no need for cold storage and potentially cross-neutralising capabilities.

2.4.4 Risk of lyssavirus spillover or host switching

One of the major factors in the discussion of cross-neutralising vaccines is the potential for spillover events to occur, and indeed proceed into full novel reservoir generation. This has been a topic of debate surrounding pathogens in general and has been made more relevant than ever following the spillover of Sars-Cov-2 causing the Covid-19 pandemic [Sparrer *et al*., 2023]. Spillover of lyssaviruses does occur however it is quite rare. The exception is RABV which has

distributed broadly around the globe and instigated several successful spillovers to produce novel reservoirs and even adaptive variants, this is well documented [Badrane & Tordo, 2001; Rupprecht *et al*., 2011; Kuzmin *et al*., 2012; Troupin *et al*., 2016]. It has been speculated that this rapid global expansion occurred in the 15th century due to the introduction of widespread intercontinental trade and may go some way to provide and explanation for the lack of divergent lyssaviruses in the Americas as RABV is the only lyssa that has ever been isolated in the New World [Troupin *et al*., 2016]. Several canine RABV spillovers into various wildlife species from the order *Carnivora* involved no adaptive evolution, suggesting that a major factor in the risk of novel reservoirs being generated is ecological in nature and not necessarily impaired by the virus genome [Troupin *et al*., 2016]. In contrast there is evidence which shows that amino acid changes do occur when lyssaviruses host switch such as in ferret-badgers in Asia [Troupin *et al*., 2016]. This indicates that the subject of spillover possibility, especially pertaining to genetic adaption, requires more data.

Divergent lyssavirus spillovers do occur, for example, a recent study investigating an event which involved the phylogroup III lyssavirus WCBV in Italy highlights this risk [Leopardi *et al*., 2021]. In this incident, WCBV was transmitted to a pet cat from a suspected local bent-winged bat (*Miniopterus schreibersii*) population which was located in a tunnel within the city limits of Arezzo. The infection has confirmed the capability of WCBV to infect non-volant mammals and cause clinical rabies, an alarming yet expected discovery based on the history of other lyssaviruses. Cats, similar to the RABV reservoir of canines, are highly social animals and are often free roaming as pets and indeed strays, often in close proximity to humans. It is therefore conceivable that domestic cats could represent a potential novel reservoir, replicating the success of canine RABV. However, this is speculation as there is currently not enough understanding of lyssavirus spillovers to determine a species susceptibility. In any case, the study found that free-ranging cats are a potential hazard for public health from lyssavirus exposure due to this incident.

A significant outcome of this incident was the furthering of knowledge surrounding lyssavirus distribution in Europe. The WCBV virus was found over 2000km away from the original discovery of WCBV 18 years ago in the mountains of Caucasus [Kuzmin *et al*., 2008]. However, the Italian isolate shared very high sequence identity showing no signs of geographical clustering or adaptation to a novel host [Troupin *et al*., 2016]. It was also discovered in the same bat species, *Miniopterus schreibersii*. Notably, recent divergent bat lyssavirus detections in the Old World have all been associated with the *Miniopterus* bat genera [Troupin *et al*., 2016; Fooks *et al*., 2021]. This genera inhabits large regions across much of southern Europe, Africa, northern and eastern Australia, and southern and south-eastern Asia [Miller-Butterworth *et al*., 2007]. This wide range combined with the presence of divergent viruses such as WCBV

increases the threat of untreatable human rabies spill-over infections occurring. In fact, it has been suggested that because of the presence of divergent lyssaviruses in *Miniopterus* including WCBV and LLEBV and the currently unclassified Matlo bat lyssavirus, it is possible they play a major role in the evolution of divergent lyssaviruses [Coertse *et al*., 2020]. Additionally, the three bat species *Myotis myotis, Myotis blythii* and *Tadarida teniotis*, which are considered to be non-reservoir at present, have been found to be serologically reactive against EBLV-1 despite no positive confirmed cases [Leopardi *et al*., 2018]. This is important as it implies that lyssaviruses may have either higher diversity or host range than currently estimated.

Though the WCBV spillover infection represents another dead-end incursion into a novel species, it does highlight the need for development of a novel vaccine against divergent lyssaviruses in phylogroups II and III. The study concluded that it was only based on good fortune no human infections occurred, and that the economic cost of this single infection was high. This was due to the extensive animal quarantining that had to be performed as well as PEP for all potentially exposed members of the public. It also increased the potential risk to workers handling the investigation and the quarantining of the animals, all of which would have been prevented had a cross-neutralising vaccine been available.

MOKV is another candidate which has been shown multiple times to spill over into the domestic animal group. Both cats and dogs have been found to be infected with MOKV in several cases [Foggin, 1982; Foggin 1988]. This is concerning, as some of these animals were shown to be vaccinated against the rabies virus, further demonstrating the vulnerability of vaccinated individuals in relation to divergent lyssaviruses. These events have been rarely detected, as with WCBV, however once again the lack of surveillance could hide the fact that these occurrences are more common than currently reported. Additionally, the occurrence in domestic dogs should be a cause for concern, as they are the primary reservoir for currently circulating RABV.

ABLV was identified in Australia in May 1996 (Hooper *et al*., 1997) in a juvenile black flying-fox (*Pteropus alecto)* and has since been found in multiple other bat species including the little-red flying-fox (*Pteropus scapulatus*), grey-headed flying-fox (*Pteropus poliocephalus*), spectacled flying-fox (*Pteropus conspicillatus*), and the yellow-bellied sheath-tailed bat (*Saccolaimus flaviventris*). Serological examination has also found evidence ABLV has been detected as far as Thailand and the Philippine Islands [Constantine, 2009] which indicates a wide distribution of the virus. This is important as ABLV has been involved in a number of spillover events including three human cases [Allworth *et al.,* 1996; Francis *et al*., 2014; Hanna *et al*., 2000]. These cases ranged from 1996 to 2013 and involved bat transmission to humans. Dog and horse exposure has also occurred, with the dog infection being particularly important to note due to the current RABV canine reservoir [New South Wales Department of Primary

Industries, 2013; Annand & Reid, 2014]. Despite ABLV being phylogroup I and therefore less divergent than WCBV, its spillover is still a hazard, though current vaccines are deemed effective against it.

The majority of spill-over events involving RABV and other lyssaviruses are considered dead-end infections. However, lyssavirus host switching can occur and has been documented to be capable of sustained onwards transmission to a naïve host [Troupin *et al*., 2016]. RABV is the only member of the *lyssavirus* genus to repeatedly host shift and circulate successfully within non-volant mammals [Marston *et al*., 2018c]. The unique situation that RABV has with its infection of non-volant mammals is important in the effort to eradicate human rabies. While host switching has not occurred in any other lyssa, it is clearly possible as demonstrated by RABV, however the mechanisms are unknown and thus the likelihood of this occurring cannot be stated for certain. Understanding host shifting even within RABV will, in the event of canine lyssa eradication, aid in the prevention of future host-shift events which would ultimately undo the efforts of any eradication programmes [Marston *et al*., 2018c]. A cross-neutralising vaccine would also be extremely useful over simple vaccines against each phylogroup as many countries which have divergent lyssaviruses have limited access to the technologies required to identify the viral species. In these cases, a cross-neutralising vaccine would be a 'catch all'. Ultimately, despite the potential for spillover events in divergent lyssavirus species being present, the number of recognised human cases of non-RABV infections are small, as little as 15 [Fooks *et al*., 2021]. Though this is the case it cannot be determined at present if this is due to a lack of surveillance, perhaps more cases occur without being recognised, or if it is due to viral limitations both genetic and ecological. Without further studies on the issue this cannot be known.

Understanding of the risks involved with the large bat reservoirs which maintain lyssaviruses is hampered by the severe lack of data on bat ecology, immunology and pathobiology [Fooks *et al*., 2021]. However currently, in the context of human infection, lyssaviruses coming from sources other than canine represents a rare hazard. The discussion of spillover risk therefore should be on the production of novel reservoirs with a particular focus on canines, as this would hamper efforts to completely eradicate human and canine rabies. If RABV was eliminated from canine reservoirs, there could be a risk based on the current understanding of lyssaviruses, for a different member of the genus to replace it. With the fact that divergent lyssaviruses have been found in canines [Coertse *et al*., 2017], it is essential that a cross-neutralising treatment be made available in the event that these lyssaviruses replace RABV once it has been eradicated from canine populations.

Lyssavirus co-evolution with a single host reservoir likely generates more successful replication after initial deposition into the host, increasing the likelihood of reaching the CNS in addition to better adaptation for immune avoidance [Marston *et al*., 2018c]. Unfortunately, a lack of data regarding real world examples of infection means there is no real understanding of the main factors which produce onward infections in a natural environment, and this gap in lyssavirus knowledge could lead to difficulties when attempting to eradicate human infections.

In relation to the reduction in rabies within a nation, the prevention of dog mediated rabies creates the greatest impact and is clearly an essential strategy when attempting to eradicate human infection with RABV [Hampson *et al*., 2015], however a single course of preventative action is not enough, there needs to be a consistent and active effort to keep the susceptible dog population low [Marston *et al*., 2018c]. Part of this effort should be producing the means to prevent further incursions of lyssaviruses both RABV and divergent through the generation of more efficient and affordable vaccines as well as treatments to divergent lyssaviruses, cementing the need for more vaccines and studies like this. Rabies also causes high levels of suffering in the mammals it infects. A consideration should be made to the impact of these viruses on animals as well as humans and that future spillovers and infections are most likely to impact other animals, such as canines, before humans. Reducing human infections to zero also means eradicating rabies infections in many non-volant animals, the benefits of which cannot be understated.

2.4.5 Future work

Though the results of this study are exciting, there are limitations. The vaccine study suffered from an unexpected low mortality observed in the rabies infected mouse control group given the mock ChAd155 vaccination, weakening the observations here, however, as the mouse sera was not utilised it had no impact on the results of this study specifically. Additionally, the numbers of analysed sera and animals available for testing were limited which prevented the vaccine study from drawing conclusions on the efficacy of their novel vaccine in comparison to RABPIPUR. This has the knock-on effect of reducing the impact of the cross-neutralising findings. A remedy to this would be a repeat of the entire experiment with a more robust animal model and greater numbers of sera which could be used to repeat the experiments performed by this study. Fortunately, other than the limitations inherent to the use of pseudo-particles in research such as their lack of natural glycoprotein distribution on the surface of virions, this study had no obvious limitations which would call into question the results. This being said, the study would greatly benefit from expansion of lyssaviruses used, sera time points, animal sera samples, and different viral isolates to not repeat the inconsistent results of Malerczyk *et al*., 2014, where a change of strain removed the cross-neutralisation which occurred.

Chapter 3: *In silico* investigations of lyssavirus glycoprotein glycosylation

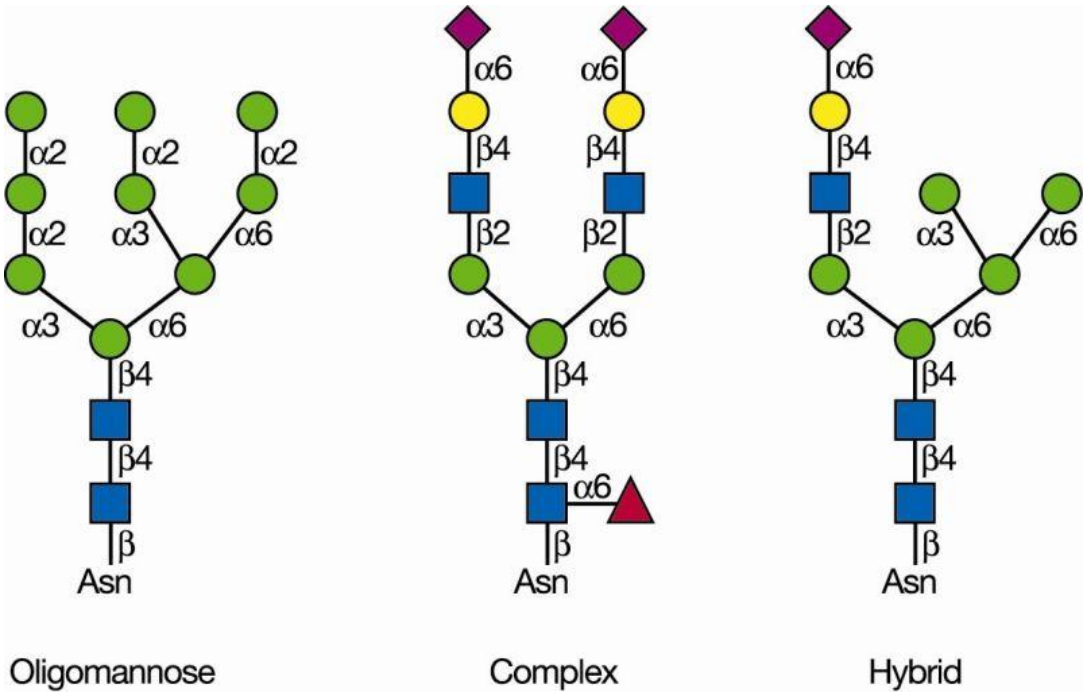
3.1 Introduction

3.1.1 Glycosylation

Glycosylation is the co/post-translational modification of proteins, attaching a 'glycan' or carbohydrate molecule to the protein. There are several types of glycosylation which can occur: *N*-linked, *O*-linked, *C*-linked, phosphoglycans and glypiation. N, O and C linked refer to the position the carbohydrate is attached, for N it is the amide nitrogen atom of asparagine (N) which the glycan is covalently bound to, and requires a lipid called dolichol phosphate [Mellquist *et al*., 1998]. O linked can be attached to the hydroxyl oxygen of several different amino acids such as serine, tyrosine and threonine; or it can also be attached to oxygens on lipids for example ceramide [Van den Steen *et al*., 1998]. C linked is far less common than N or O linked glycosylation and occurs when a carbohydrate is attached to a carbon on a tryptophan sidechain [Hofsteenge *et al*., 1999]. Phosphoglycans are linked through the use of a phosphoserine's phosphate and finally, glypiation occurs when there is an addition of a GPI anchor linking proteins to lipids through glycan linkages [Roller *et al*., 2020].

In addition to the variation in glycosylation types, there is also variation in the types of glycans which can be attached. For *N*-linked there can be high mannose, complex or hybrid glycans and O-linked can have core 1, 2, 3 or 4. All N-linked glycans share a common core sugar sequence which consists of: Manα1–6(Manα1–3)Manβ1–4GlcNAcβ1–4GlcNAcβ1-Asn-X-Ser/Thr as can be seen in figure 3.1 [Stanley *et al*., 2009]. Oligomannose or 'high mannose' is a type of N-glycan which consists of only mannose residues that are attached to the core. Mannose is a sugar monomer of the aldohexose carbohydrates group meaning that is a simple sugar with six carbon atoms total. Complex NLGs are formed when 'antennae' are attached to the core, the antennae consist of potentially many types of saccharides which can include more than the two original N-Acetylglucosamine (GlcNAc) moieties. Hybrid is, as the name suggests, a combination of the two types where mannose sugars are attached to Manα 1-6 arm and the Manα 1-3 arm has one or two complex antenna attached [Stanley *et al*., 2009]. The type of glycan that is attached to a protein is important as it can impact characteristics of the protein such as structure, which in turn impacts protein function [Higel *et al*., 2016]. This is potently demonstrated by the fact that de-glycosylated IgG was unable to be crystalised due to the impact removal of the glycans had on the flexibility of the proteins CH2 domain [Krapp *et al*., 2003].

A



B

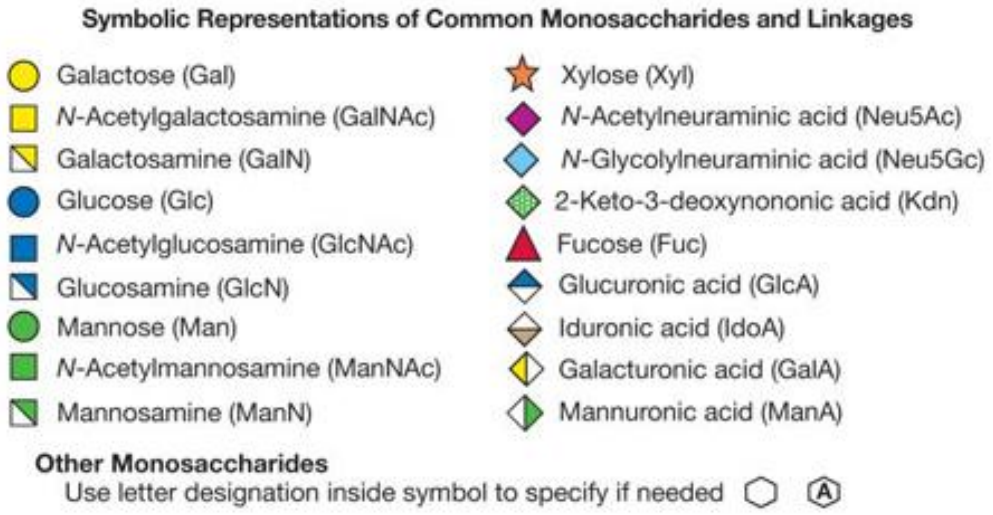


Figure 3.1: (A) The different types of N-linked glycans which can be added to a protein, these are oligomannose or high mannose, complex and hybrid. The high mannose consists of multiple additional mannose monomers added to the original glycan polysaccharide. Complex is formed from antennae which can contain multiple different sugars as demonstrated by figure B. Hybrid is a combination of the two previous types. [Stanley et al., 2009].

3.1.2 N-linked glycosylation in viruses

Glycans play an essential role in biology as a whole but are also essential in many viruses. Despite this and an increasing amount of research into glycans, some functions have not been fully characterised. N-linked glycosylation (NLG) is the most common form of protein modification and has been shown to aid or enable functions such as viral attachment, transmission, immune avoidance, infectivity, entry and assembly [Li *et al*., 2021; Haywood, 2010; Vigerust & Shepherd, 2007]. Examples of viruses and their known glycan functions can be found in table 3.1. Glycans are common in viruses and in fact, all attachment or fusion proteins of enveloped viruses are glycosylated in some way [Li *et al*., 2021].

Table 3.1: Important glycosylations in common enveloped viruses. This list is not comprehensive [Li et al., 2021].

Virus	***Viral glycoprotein***	***N-glycosylation sites***	***Viral glycoprotein functions***
HIV-1	Gp120	20-30	Attachment, transmission, glycan shield
H1N1	HA	0-11	Attachment, glycan shield
EBOV	GP	11-18	Infectivity, attachment, glycan shield
HCV	E1, E2	4-15	Infectivity, entry
RABV	GP	2	Infectivity, entry, virulence
WNV	prM, E	1-2	Entry, release
SARS-CoV-2	S	22	Assembly, attachment, entry

N-linked glycosylation (NLG) in viruses occurs when an oligosaccharide is attached via covalent bonding to the Asn residue of a sequon. The glycan is attached to the amide nitrogen atom of the asparagine. For this process to occur, a particular sequence of amino acids or a 'sequon' must be present in the protein. For NLG, this is Asn-X-thr/ser where X is any amino acid with the exception of proline [Shakin-Eshleman *et al*, 1992a, Vigerust DJ & Shepherd *et al*., 2007;]. This process begins in the endoplasmic reticulum (ER) in the early stages of viral

glycoprotein synthesis. The transfer of a pre-synthesized core oligosaccharide via oligosaccharyltransferase from a dolichol phosphate, which is anchored in the lumen of the host ER, to the viral glycoprotein occurs [Mellquist *et al*., 1998; Ortega *et al*., 2019]. The host dolichol phosphate naturally forms these glycans for use in host proteins. Later modifications occur both within the ER and the Golgi apparatus as the glycoprotein progresses. These modifications can be both the addition of carbohydrate residues or the removal or 'trimming' of carbohydrates already present. Further modification can occur in the Golgi which results in the three NLG glycan classifications, high-mannose, hybrid or complex depending on their carbohydrate consistency [Varki *et al*., 2022; Aebi, 2013]. The trimming of the carbohydrates is performed by host glycosidases, with further modifications being performed by glycotransferases which assemble complex glycans in the Golgi [Varki, 2011].

HIV glycosylation has been scrutinised for its role in viral pathogenesis and host immune response. The HIV envelope protein gp120 is essential for the virus to bind to host cells and contains between 20-30 glycosylation sites, with an average of 24, and is the most heavily glycosylated of viral glycoproteins [Stewart-Jones *et al*., 2016]. It has been demonstrated that these sites are somewhat involved in the correct protein folding of the gp120 protein, though it is not essential [Rathore *et al*., 2017; Mathys *et al*., 2014]. Particular NLGs such as Asn260 have been shown to impact viral infectivity and gp120 structure when removed [Mathys *et al*., 2014]. Additionally, the ability of the virus to bind to its CD4 host receptor is significantly reduced post mutation suggesting a role in viral entry [François & Balzarini, 2011]. Host immune avoidance has also been explored in HIV. Glycosylation plays a critical role in the avoidance of the humoral immune system and produces difficulties when attempting to target key sites with antibodies, impacting therapeutic design [Stewart-Jones *et al*., 2016; Wei *et al*., 2003]. It does this primarily through its sheer size, with the glycan 'shield' comprising of over half the glycoproteins mass, blocking key immunogenic sites [Leonard *et al*., 1990]. These findings demonstrate just how important these glycans are and their multifaceted roles in viral processes.

In the lyssavirus family *Rhabdoviridae*, glycosylation has been poorly researched with many of the virus genera within the family lacking any glycosylation characterisation [Ortega *et al*., 2019]. The *Vesiculovirus* genus is a member of the *Rhabdoviridae* family and the glycans present on its glycoprotein have been identified, and along with lyssaviruses, is the most researched genus of the *Rhabdoviridae* in relation to glycosylation [Ortega *et al*., 2019]. Similar to the lyssavirus genus, the VSV genus contains two NLGs which are present at Asn181 and Asn336, making up 10% of the glycoproteins mass [Etchison & Holland, 1974]. These have been shown to impact VSVs folding and confirmation [Ortega *et al*., 2019] and proper glycoprotein transport to the cell surface [Ortega *et al*., 2019]. However in contrast to

lyssaviruses, either glycan can be present and the proper folding and transport can be achieved [Bailey *et al*., 1989], whereas in lyssaviruses Asn319 is an obligate glycan that is present in all members of the genus [Badrane & Tordo, 2001].

3.1.3 Glycosylation in the genus Lyssavirus

Studies of lyssavirus glycosylation have been able to identify two key glycosylations which occur in most RABVs, Asn37 and Asn319. Asn319 is an obligate glycan and is present in every member of the genus, it has been shown to be essential for the production of mature virions and loss of the glycan through knockouts prevents correct glycoprotein folding, decreased viral production and inhibited membrane fusion [Yamada *et al*., 2013]. The same study reported that removal of all three lyssa glycans in their fixed RABV strain resulted in a complete loss of cell surface expression levels (CSE). Asn37 is inefficiently glycosylated meaning that not every glycoprotein which is produced during replication receives a glycan and indeed increases in glycan attachment efficiency reduces the pathogenicity of RABV, indicating that a balance of Asn37 is required for efficient pathogenicity [Ortega *et al*., 2019]. Asn37 has been shown to not be necessary for CSE [Shakin-Eshleman *et al*., 1992a]. Glycan efficiency is impacted by several factors including the X amino acid in the sequon (Asn-X-T/S) and the amino acid directly following the sequon [Shakin-Eshleman *et al*., 1992a, b and 1996]. Asn247 has been commonly identified as a lab adaptation which is present in many lab or 'fixed' strains of RABV [Wunner *et al*., 1985]. This glycan site has been speculated to be involved in the reduced pathogenicity that is commonly seen in fixed RABV strains, as well as increased efficiency for viral production in cell cultures [Yamada *et al*., 2012].

Recent studies have focused on the addition and removal of glycan sites such as the study by Yamada et al 2013 [Yamada *et al*., 2013] which investigated an extra glycan artificially introduced at Asn194 and how removal of Asn319 impacted the virus, finding it massively reduced virion production. While there has been an interest in viral production, the impact of lyssavirus glycoprotein glycosylation on infectivity and neutralisation has been limited.

A study by [Coelingh, 2021] investigating the structure of RABV using cryo-EM and mass spectrometry combined, identified glycans on the surface of the RABV glycoprotein. The study found that the Asn247 and Asn319 had a variety of glycan isoforms with a relatively high fucosylated glycan being the most common for both of the NLG sites. The study also identified a large number of biantennary glycans which is unusual, however it is explained through the previous reporting of this occurrence when using specific cell lines and medium for the transfection and expression of virus.

Studies have clearly demonstrated that lyssavirus glycosylation is undoubtedly important and are an essential factor in CSE of the RABV G. However, there are clear gaps in the literature

such as the effect of glycosylation on infection and neutralisation of lyssaviruses. In addition, the frequency of glycans, their positions and their sequon structures have not been effectively catalogued in literature. Therefore, this study aimed to collate this data through the NCBI sequence database GenBank.

Additionally, the 3D crystal structure of the RABV glycoprotein has only been produced recently [Yang *et al*., 2020] and, though the antigenic regions have been mapped they haven't been discussed in the context of glycosylation.

Aims:

- Fully characterise the sequon presence throughout the *lyssavirus* genus.
- Identify the amino acids responsible for sequon presence, measure their conservation and estimate their glycosylation efficiency based on computer models.
- Identify the antigenic regions on a 3D crystal structure of the RABV spike and image where potential glycans would be based on sequon analysis.
- Systematically review the methodology behind the site directed mutagenesis of glycosylation sites in viruses as a whole.

3.2 Materials and Methods

3.2.1 Sequence data acquisition and processing

All analyses and sequence acquisition were performed through the software Geneious Prime version 2022.2.2. Sequences were pulled from the NCBI protein database using the search term "lyssavirus glycoprotein" which returned 9,679 sequences. Full length sequences were extracted and organised into species, then trimmed to remove the signal peptide region, leaving 4785 sequences remaining. Trimmed full length sequences were then examined for PNGs manually.

Sequence data was exported from Geneious into a fasta file format, this was then uploaded to the WebLogo website version 2.8.2 [Crooks *et al*., 2004] in order to generate WebLogo plots.

3.2.2 Crystal structure acquisition and processing

Crystal structure images were obtained from [Callaway *et al*., 2022] and imported into the ChimeraX software version 1.2.0 [Meng *et al*., 2023]. Manual labelling of antigenic regions and addition of glycans to PNG was performed in ChimeraX, as was production of all figures which were produced using the software. UCSF ChimeraX was developed by the by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases [Meng *et al*., 2023].

3.2.3 Systematic review parameters and performance

A systematic review was performed investigating the amino acid substitutions used in studies to knock out N-linked glycosylation sites. This review was performed using the NCBI PubMed central database using the search string "N-linked AND virus AND substitution AND mutagenesis AND glycan". These search terms returned 1494 papers of which 186 were retained. Papers were reviewed for relevance and assessed as to if they provided data pertaining to the aims of the study. Data harvested from retained papers included their title, a link to their PubMed page, the virus being researched, the amino acid position in a glycan PNG changed, the amino acid changed and the amino acid changed to. Papers were removed from the review if they contained any of the following criteria:

- The paper was a review paper and not primary research.
- Glycans were researched in the paper but they were not N-linked.
- The paper studied glycans but did not perform a knockout study.
- Amino acids were modified to add a PNG instead of removing it.

- PNGs modified were within an organism other than a virus.
- Removal was through experimental ‘coincidence’ and not the aim of the study, for example random mutation studies.
- Removal was a choice to rescue certain escape mutants or the mutation occurred naturally through passage or other processes.

3.2.4 Web logo

Web logo analysis was performed using sequence data retrieved from GenBank. Representative sequences were selected and input into the WebLogo website creation tool version 2.8.2 [Crooks *et al*., 2004]. This was done in a fasta file format and the settings were set to default with the exception of the title and image output format which was set to PNG.

3.2.5 NetNGlyc

To generate glycosylation prediction graphs the NetNGlyc version 1.0 online tool was utilised [Gupta & Brunak, 2002]. Fasta format full length rabies virus glycoprotein sequences of interest were exported from Geneious having been obtained through the GenBank protein search function. The settings were left on the default and files were generated without the signal peptide as is standard in lyssavirus amino acid reporting.

3.3 Results

3.3.1 Identification of glycan sites in sequences available

Sequon position and frequency differ between lyssaviruses and viral isolates within a strain, therefore an investigation into sequon presence was performed on all lyssavirus sequences available on GenBank. Sequences were aligned together and their signal peptide region (19 amino acids) removed, as such, all amino acid positions are labelled without the presence of the signal peptide, this is common practice in lyssavirus research. The amino acid positions and frequencies in lyssaviruses are show in figure 3.2.

			N residue location (% of sequences containing glycosylation site)													
Phylogroup	Lyssavirus	Number of sequences	34	37	42	90	146	158	160	184	202	204	247	319	334	436
I	RABV	4499	X	81.04%	X	X	0.17%	0.35%	X	X	0.02%	1.18%	6.60%	100%	X	X
	EBLV1	100	X	X	X	X	X	X	X	X	X	X	X	100%	X	X
	EBLV2	37	X	X	X	X	X	X	X	X	X	X	X	100%	X	X
	ABLV	57	X	X	X	X	X	X	X	X	X	X	X	100%	X	X
	Bokeloh BLV	11	X	100%	X	X	X	X	X	X	X	X	X	100%	X	X
	Irkut LV	3	X	X	X	67%	X	X	X	X	X	X	X	100%	X	X
	Duvenhage LV	9	X	X	X	X	X	X	X	X	X	X	88.80%	100%	X	X
	Aravan LV	3	X	X	X	X	X	X	X	X	X	X	X	100%	X	X
	Khujand LV	3	X	X	X	X	X	X	X	X	X	X	X	100%	X	X
	Gannoruwa BLV	5	X	X	X	X	X	X	X	X	X	X	X	100%	X	X
	Medoid	1	X	100%	X	X	X	X	X	X	X	X	X	100%	X	X
II	MOKV	24	X	X	X	X	X	X	X	8%	100%	X	X	100%	X	X
	Shimoni BLV	2	X	X	X	X	X	X	X	X	100%	X	X	100%	X	X
	Lagos BLV	27	X	X	X	X	X	X	X	100%	100%	X	X	100%	29.60%	22.20%
III	IKOV	2	X	X	X	X	100%	X	100%	X	100%	X	X	100%	100%	X
	WCBV	3	X	X	X	X	X	X	X	X	100%	X	X	100%	100%	X
	Lleida BLV	3	100%	X	100%	X	100%	X	X	X	X	X	X	100%	100%	X
			Site IIb	Site IIb	Site IIb										Site III	

Figure 3.2: Sequence alignment in Geneious Prime of all lyssavirus sequences available on the NCBI GenBank database. Full length sequences only were used, no partials. Lyssaviruses are split into their phylogroups and the number of sequences for each virus was recorded. Seuqon positions are noted and represent the Asn amino acid position in the sequence. The frequency/conservation of the sequon within a species is represented by a percentage, with 100% conservation marked green. Antigenic regions are annotated to indicate sequons which coincide with their amino acid positions in the glycoprotein. The medoid sequence from chapter 2 has been included.

These results represent all known full length amino acid sequences of lyssavirus glycoproteins, partial sequences were excluded from the search results. The figure shows each officially recognised lyssavirus at the time of the study and each asparagine (N) position which is a part of a potential N-linked glycosylation site (PNG) or 'sequon'. The percentage of sequences from each virus containing the sequon is also listed at each location. Sequons which are located within a classical lyssavirus 'antigenic region' are labelled below the column.

The results demonstrate that Asn319 is conserved across all lyssavirus sequences and species without exception, also spanning every phylogroup. Phylogroup II and III lyssaviruses generally have more sequons present which are well conserved, though the lack of total sequences available makes this difficult to accurately state. In contrast, phylogroup I outside of RABV is relatively devoid of additional PNGs. In RABV there are several more sequons than

any other lyssa but most are low in prevalence compared to Asn37 and Asn319, with examples such as Asn202 only being present in 0.02% of sequences. Though Asn37 is present in RABV, it is almost entirely unique to it, only being found in a single other lyssavirus, BBLV. Asn334 appears to be an important sequon in phylogroup III lyssaviruses as it is 100% conserved in all sequences of the group. Asn202 is 100% conserved in all phylogroup II species and is present in two of the phylogroup III species.

Antigenic regions IIb and III are intersected by PNGs in phylogroup II and III lyssaviruses. It should also be noted that while certain PNGs intersected antigenic regions, Asn202 and Asn204 were located directly adjacent to antigenic region IIa (amino acids 198-200). The study also reveals that PNGs across the genus are distributed throughout the glycoprotein sequence, with no particular region being glycosylated when examining the sequence in a non-conformational format, in essence not considering the 3D nature of the protein.

Though these results reveal the sequon frequencies and positions, the amino acid structure of said PNGs had not been determined. The amino acids present in the sequon matter a great deal when it comes to core glycosylation efficiency. In order to examine the potential efficiency of each of these sequons, the lyssavirus sequences were again examined and each sequons' amino acid residues were recorded. These findings were formatted in the same fashion as figure 3.2 and can be seen in figure 3.3. The 4 amino acids follow the sequon structure Asn-X-T/S-X.

			Seuqon sequence of glycosylation sites with the following residue included														
Phylogroup	Lyssavirus	Number of sequences	34	37	42	90	146	158	160	184	202	204	247	319	334	436
I	RABV	4499	X	NLSG	X	X	NKSL	NCSG	X	X	NGSK	NK/RTC	NETK	NKTL	X	X
	EBLV1	100	X	X	X	X	X	X	X	X	X	X	X	NKTL	X	X
	EBLV2	37	X	X	X	X	X	X	X	X	X	X	X	NKTL	X	X
	ABLV	57	X	X	X	X	X	X	X	X	X	X	X	NKTL	X	X
	Bokeloh BLV	11	X	NLSG	X	X	X	X	X	X	X	X	X	NKTL	X	X
	Irkut LV	4	X	X	X	NVSA	X	X	X	X	X	X	X	NNTL	X	X
	Duvenhage LV	9	X	X	X	X	X	X	X	X	X	X	NNSE	NR/KTL	X	X
	Aravan LV	3	X	X	X	X	X	X	X	X	X	X	X	NKTL	X	X
	Khujand LV	3	X	X	X	X	X	X	X	X	X	X	X	NKTL	X	X
	Gannoruwa BLV	5	X	X	X	X	X	X	X	X	X	X	X	NKTL	X	X
	Medoid		X	NLSG	X	X	X	X	X	X	X		X	NKTL	X	X
II	MOKV	25	X	X	X	X	X	X	X	NLSL	NGSR	X	X	NGSL	X	X
	Shimoni BLV	2	X	X	X	X	X	X	X	X	NGSK	X	X	NGSL	X	X
	Lagos BLV	27	X	X	X	X	X	X	X	NLSL/M	NGSR/L	X	X	NGSL	NWSE(NWTD)	NWSF
III	IKOV	2	X	X	X	X	NKTM	X	NFSP	X	NGSK	X	X	NNTM	NWTD	X
	WCBV	3	X	X	X	X	X	X	X	X	NRTS	X	X	NGTL	NWSE	X
	Lleida BLV	3	NCTD	X	NYSE	X	NKTM	X	X	X	X	X	X	NKTL	NWSE	X
			Site IIb	Site IIb	Site IIb										Site III	

Figure 3.3: Sequence data as obtained in 3.2, however the amino acid constituents of each sequon is annotated. Frequency is represented with green being 100% and yellow under 100%. When less than 100% conservation is present, the alternative amino acid is noted. The medoid sequence from chapter 2 has been included in the analysis.

The data in figure 3.3 provides more insight into the lyssavirus sequons. A sequence conservation of 100% appears to be almost universal, with few sequons differing in amino acid sequences. However, as can be seen in Asn319, amino acid sequence between species can differ substantially, with the majority of sequons having some variation throughout the genus.

The hydroxy group in a sequon occupied by threonine and serine had differences which are present throughout the genus. Over 50% (22) of the 42 PNGs had a serine, with the remaining 20 having threonine. However, when not factoring in Asn319, threonine is only present in 7 of the remaining 26 PNGs, with the remaining 19 containing the serine amino acid.

The presence of the amino acid proline at position 4 in sequon Asn202 of IKOV indicates no glycosylation will be present here as proline blocks glycosylation. This was the only PNG which contained a proline in the genus. Unfortunately due to only having two IKOV sequences it is not currently possible to know if this is an outlier or not.

3.3.2 Amino acid alternatives when sequon is not present

For the majority of PNGs present in the genus there is 100% conservation of their presence in the sequences of a particular lyssavirus species. Asn37 in RABV however, has a degree of variation, appearing in 81% of RABV sequences. A WebLogo was therefore produced to investigate what the population of amino acids at the sequon location looked like, including when the sequon is not present. This can be seen in figure 3.4.

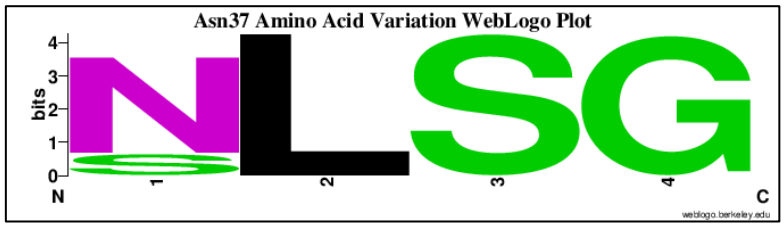


Figure 3.4: A WebLogo generated using all RABV sequences. The four amino acids are those present at position 37 of the RABV sequence.

The WebLogo demonstrates the proportionality of variation at a sequon. The second amino acid leucine is 100% conserved in the PNG. Though the serine and glycine positions both have sequences which do not contain them at Asn37, the numbers are insignificant and they therefore do not present on the WebLogo. While amino acid 37 is primarily asparagine, it is clear that there is a significant proportion of the other sequences that are a single amino acid, serine. A serine to asparagine change is potentially only a single nucleotide.

WebLogos were produced for all the PNGs within RABV sequences which had variation to elucidate the condition of amino acid variation at their locations. This includes Asn146, Asn158, Asn202, Asn204 and Asn247.

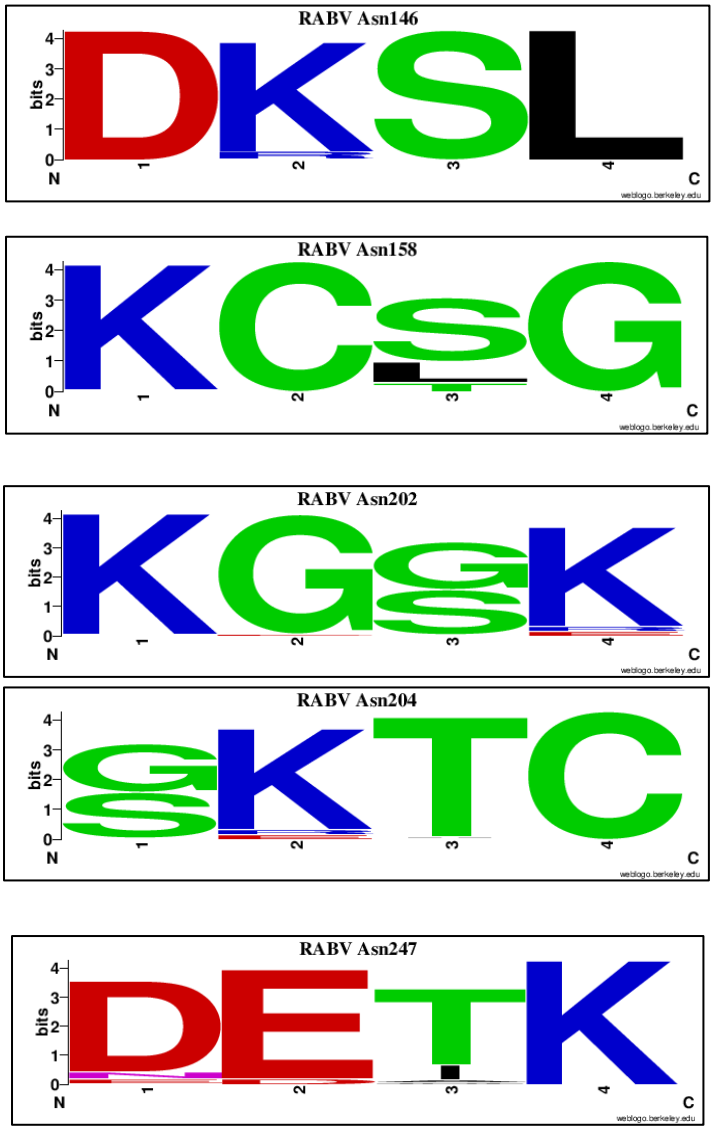


Figure 3.5: WebLogos defining the sequence variation of the amino acids at the position of detected sequons. Asn146, 158, 202, 247 and 204 from the RABV species of lyssa are shown.

Most of the positions within these sequons can be seen to be heavily conserved, however there are exceptions such as position three in Asn158 and Asn202, in addition to position one in Asn204. With the exception of Asn202, the third position amino acid is usually the serine or threonine required to produce a sequon, with the limiting factor determining the capability of the sequon to be produced being the asparagine in position one.

Though there are far fewer sequences available, Asn334 and Asn436 in LBV showed less than 100% conservation also and therefore WebLogos were also generated.

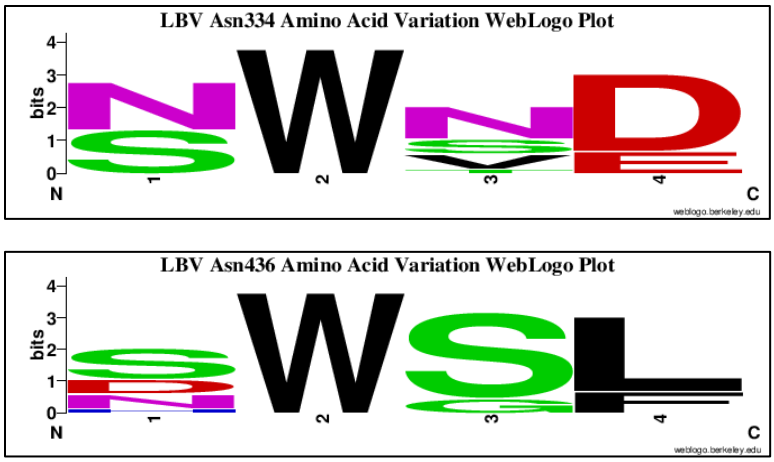
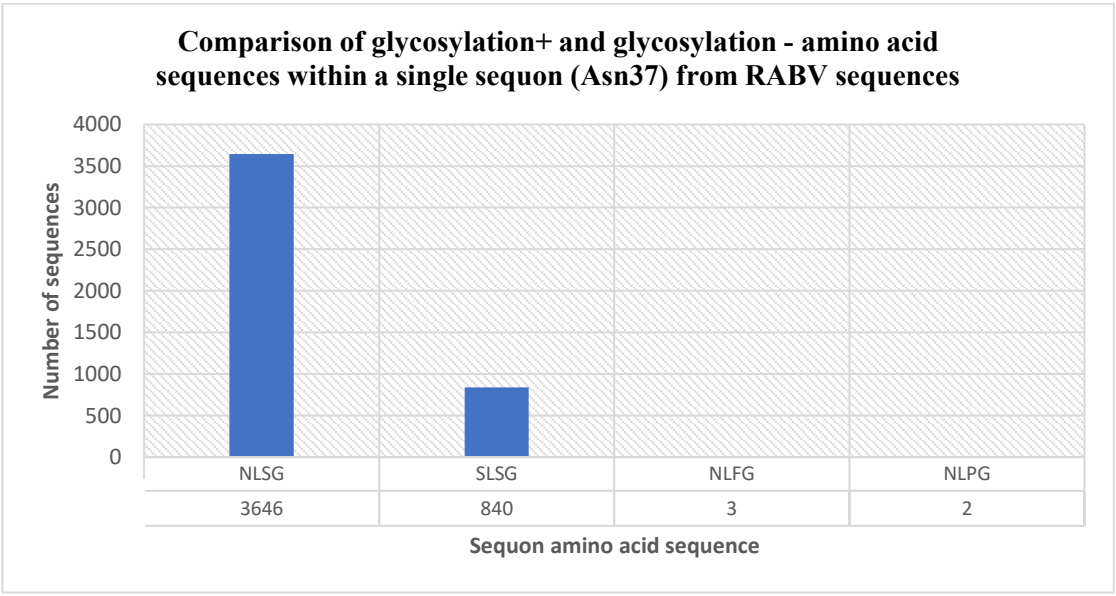
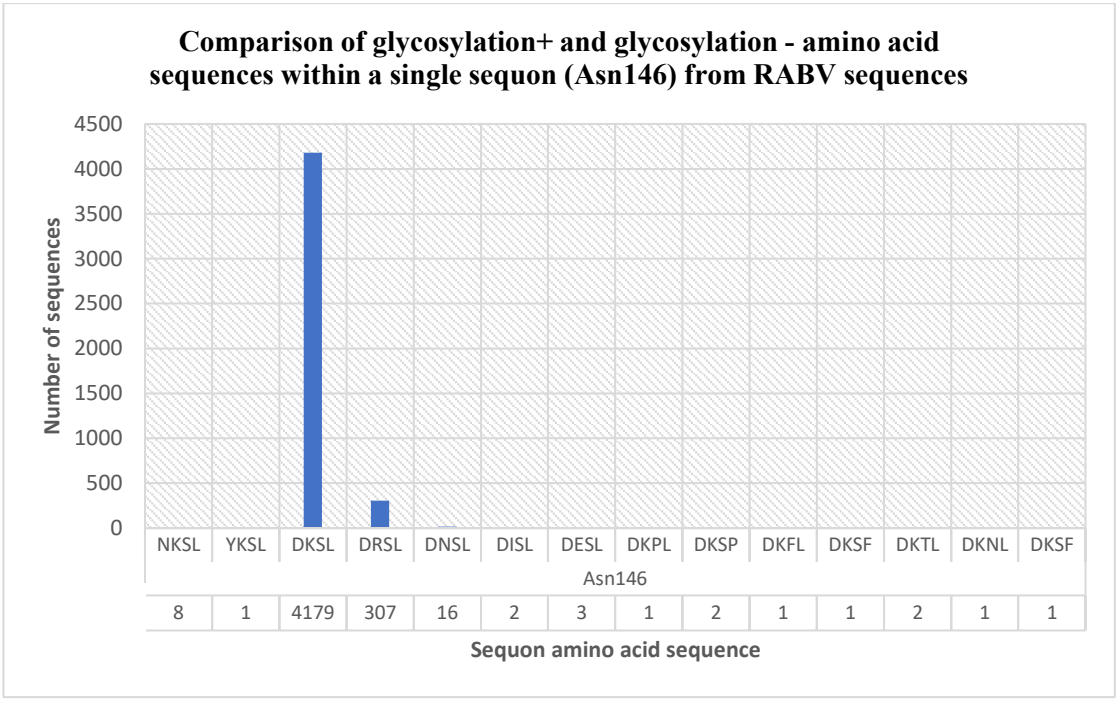


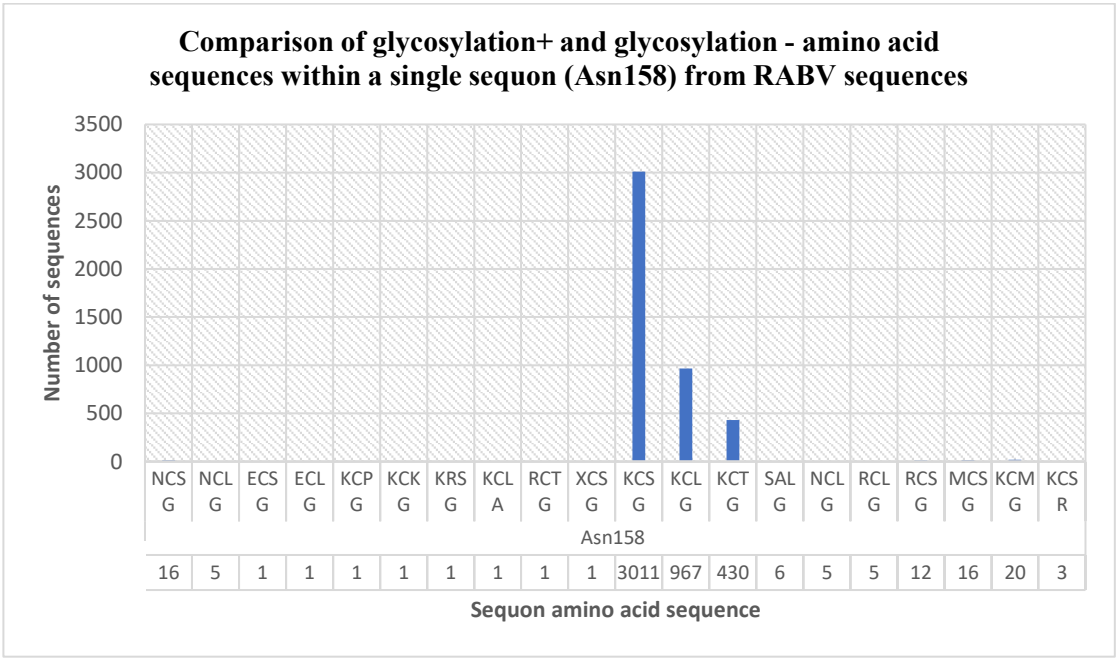
Figure 3.6: WebLogos defining the sequence variation of the amino acids at the position of detected sequons. Asn334 and Asn436 from the lyssavirus LBV were analysed.

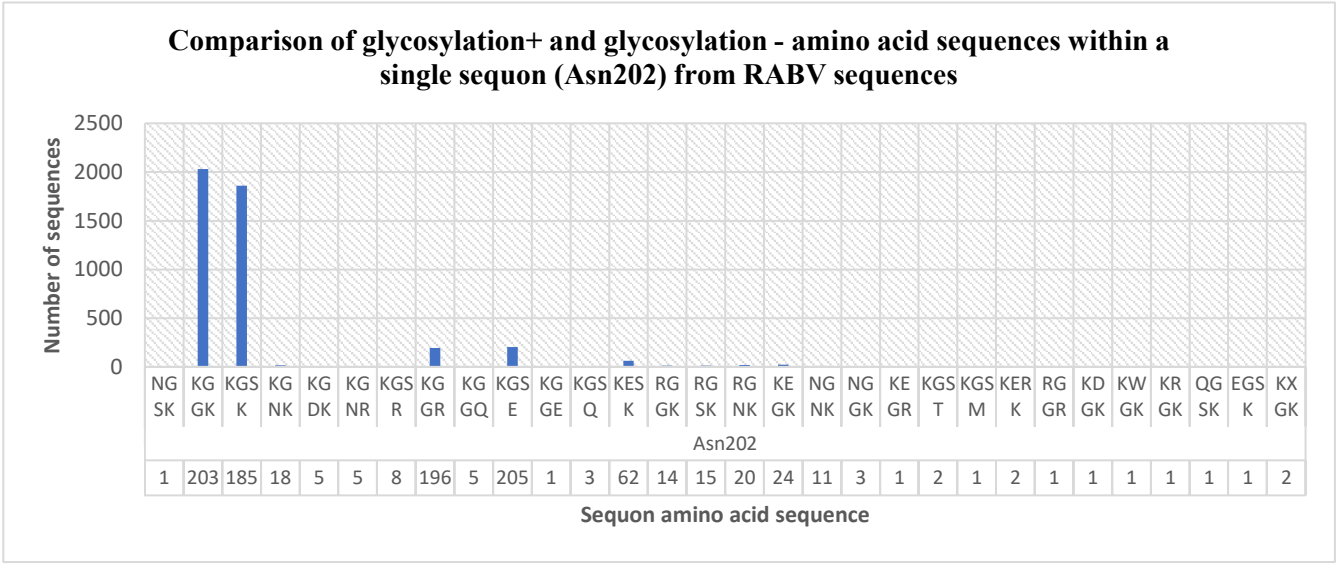
Asn334 of LBV showed that the primary substitution of asparagine was serine, with very high conservation in amino acid position two. The third amino acid position showed significant variation and though asparagine was the most common, other amino acids such as serine were commonplace. At Asn436 the position one amino acid in the PNG had high variation with the majority of sequences containing a serine in position one. The second amino acid was once again highly conserved and the third position was dominated by serine.

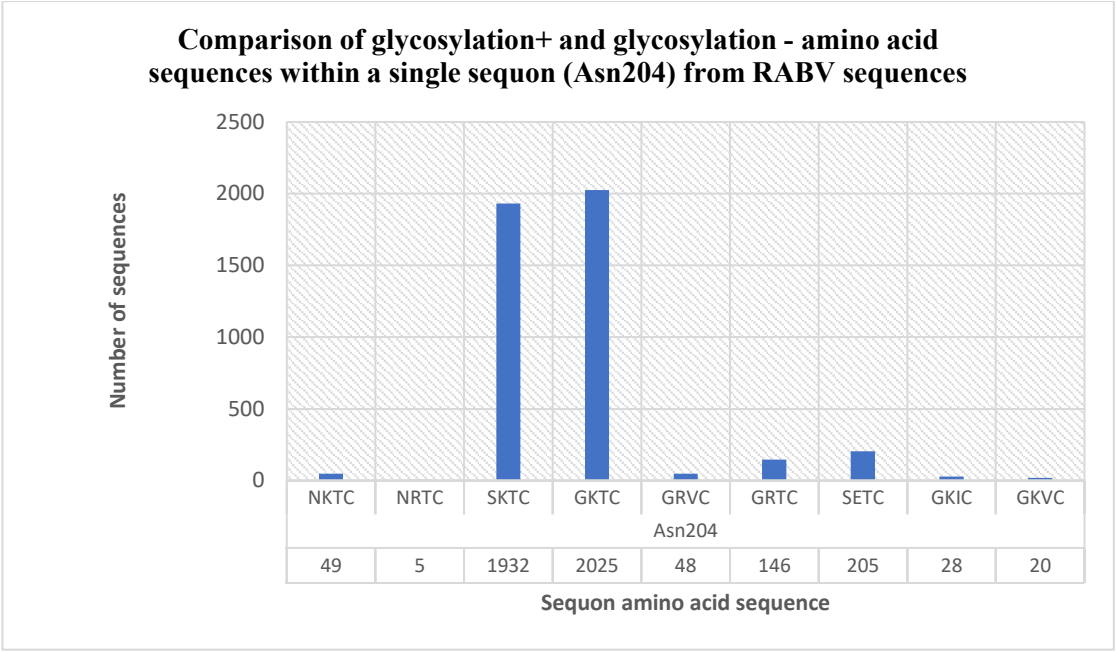
Use of WebLogos gives an overall image of the sequon as a whole, however it can be difficult to determine the exact differentiation at each position. To this end, figure 3.7 was produced which shows the number of sequences of each variant present in the sequon for all of the Asn positions discussed.




Comparison of glycosylation+ and glycosylation - amino acid sequences within a single sequon (Asn146) from RABV sequences
Number of sequences
4500
4000
3500
3000
2500
2000
1500
1000
500
0
NKSL YKSL DKSL DRSL DNSL DISL DESL DKPL DKSP DKFL DKSF DKTL DKNL DKSF
Asn146
8 1 4179 307 16 2 3 1 2 1 1 2 1 1
Sequon amino acid sequence


Comparison of glycosylation+ and glycosylation - amino acid sequences within a single sequon (Asn158) from RABV sequences
Number of sequences
3500
3000
2500
2000
1500
1000
500
0
NCSG NCLG ECSG ECLG KCPG KCKG KRSG KCLA RCTG XCSG KCSG KCLG KCTG SALG NCLG RCLG RCSG MCSG KCMG KCSR
Asn158
16 5 1 1 1 1 1 1 1 1 3011 967 430 6 5 5 12 16 20 3
Sequon amino acid sequence


Comparison of glycosylation+ and glycosylation - amino acid sequences within a single sequon (Asn202) from RABV sequences
Number of sequences
2500
2000
1500
1000
500
0
NGSK KGGK KGSK KGNK KGDK KGNR KGSR KGGR KGGQ KGSE KGGE KGSQ KESK RGGK RGSK RGNK KEGK NGNK NGGK KEGR KGST KGSM KERK RGGR KDGK KWGK KRGK QGSK EGSK KXGK
Asn202
1 203 185 18 5 5 8 196 5 205 1 3 62 14 15 20 24 11 3 1 2 1 2 1 1 1 1 1 1 2
Sequon amino acid sequence



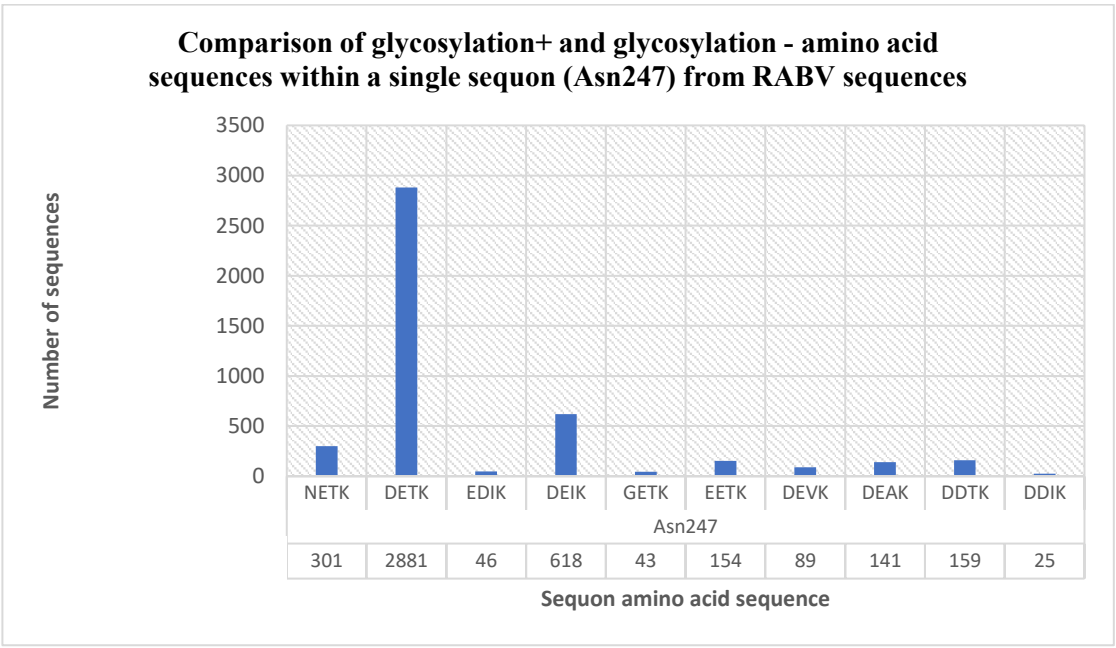


Figure 3.7: Amino acid sequences and their frequency within RABV sequons: Asn37, 146, 158, 202, 204 and 247. Sequences were retrieved from GenBank.

3.3.3 Identification of glycan sites in sequences used in chapter 2

In order for comparison to be made to the wider sequence base, the original sequences of the lyssavirus glycoproteins (used in the infection assays in chapter 2), which were obtained through sanger sequencing, were examined for potential N-linked glycan sites. Sequences were aligned against each other and then assessed for potential sites using the Asn-X-T/S sequon identifier. The amino acids were also recorded, as seen in figure 3.8.

		N residue location					
Phylogroup	Lyssavirus	37	146	160	202	204	319
X	Medoid	NLSG	X	X	X	X	NKTL
I	CVS-11 (RABV)	NLSE	X	X	X	NKTC	NKTL
	EBLV-1	X	X	X	X	X	NKTL
II	MOKV	X	X	X	NGSR	X	NGSL
III	IKOV	X	NKTM	NFSP	NGSK	X	NNTM
	WCBV	X	X	X	NRTS	X	NGTL

Figure 3.8. A sequence alignment in Geneious Prime of lyssaviruses used in infection assays in chapter 2. Lyssaviruses were split into phylogroups and their amino acids examined for potential N-linked glycosylation sites. Where sites are present, they are highlighted green, with the amino acid sequence recorded. The medoid sequence obtained from GenBank was also included, which can be found under the accession number: AGN94271.1.

3.3.4 Antigenic regions as described in the literature and their positions mapped to the 3D crystal structure of a lyssavirus

Antigenic regions have been described in the literature in lyssavirus glycoproteins. These regions are seen as important in the discussion around the antigenicity of lyssaviruses. Though these regions are important, they are not always mapped in crystal structures reported in the literature. To this end, the crystal structure from Callaway *et al*., 2022, was taken and analysed to visualise the locations of the antigenic regions and their relationship to PNGs.

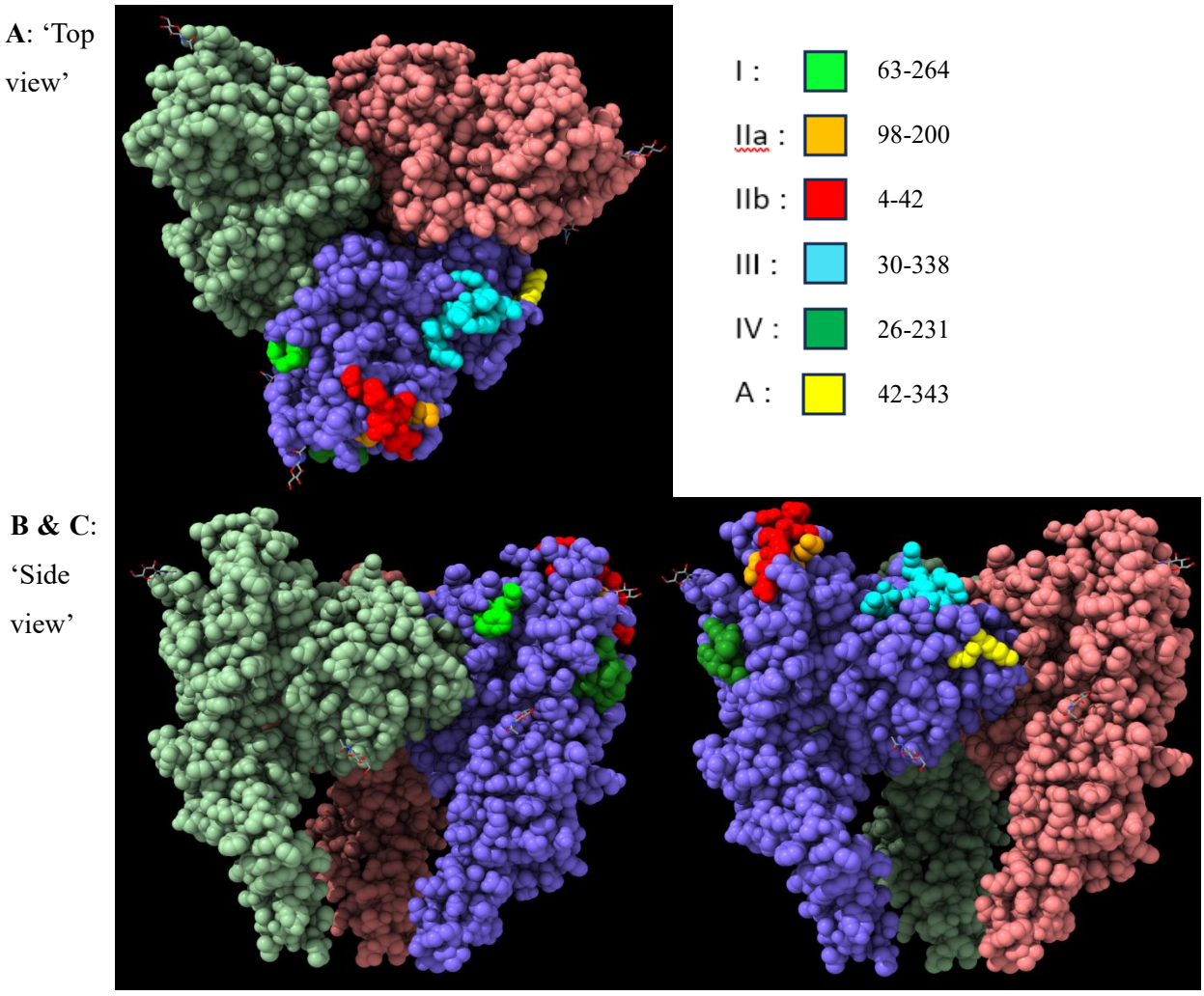


Figure 3.9: Crystal structural model 7u9g of RABV glycoprotein in the pre-fusion state, visualised using ChimeraX software. The antigenic regions of RABV G protein have been highlighted.

The results appear to show that the antigenic regions are exposed on the surface of the protein, as is to be expected as they are the binding sites for many neutralising antibodies which often prevent the change from pre-fusion to post-fusion structure. It is also apparent that the glycosylation sites present somewhat cluster around the antigenic regions with the exception of C and D.

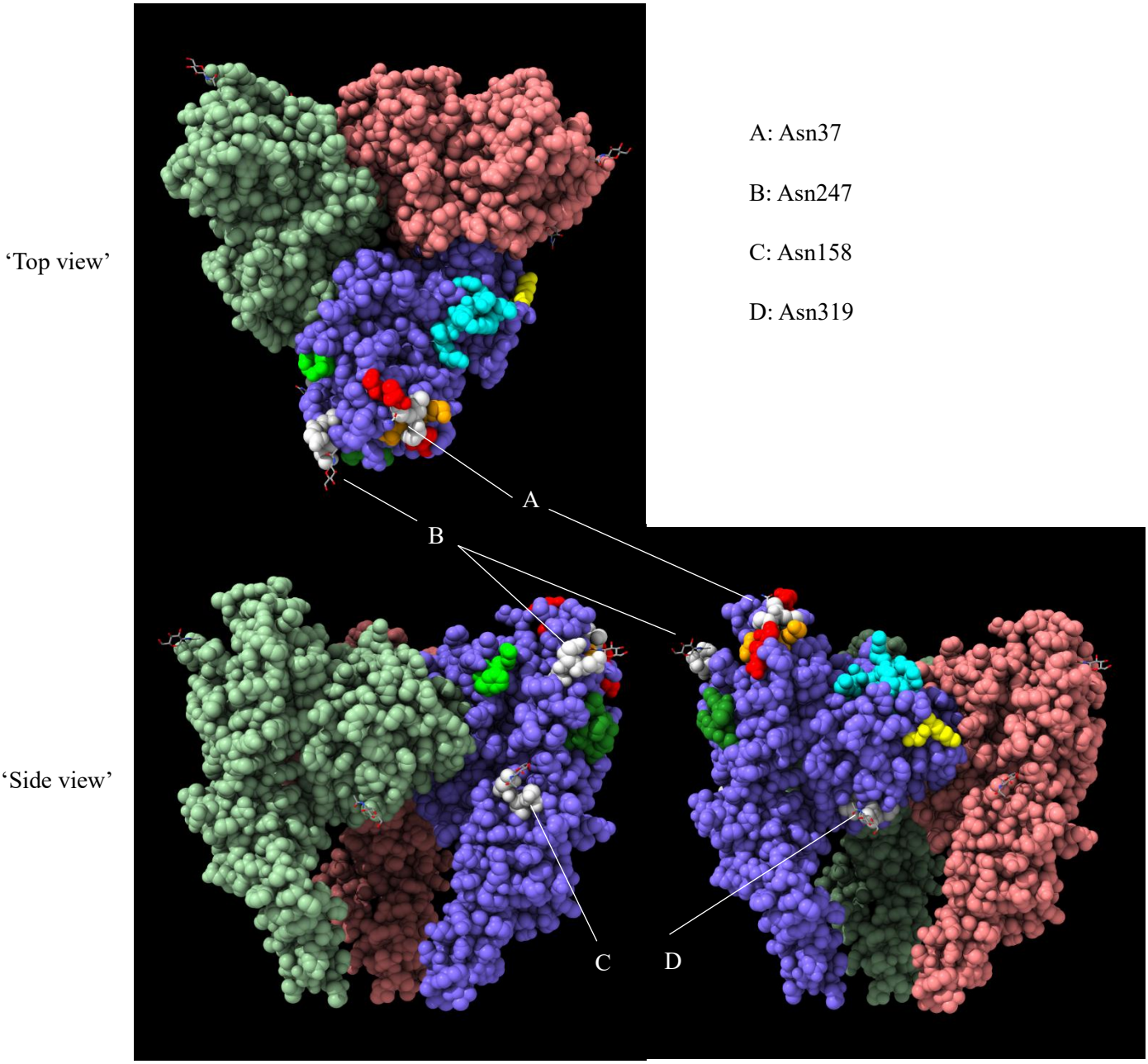


Figure 3.10: Crystal structural model 7u9g of RABV glycoprotein in the pre-fusion state, visualised using ChimeraX software. The antigenic regions of RABV G protein have been highlighted, and the PNGs present in the sequence have also been identified.

3.3.5 NetNGlyc glycan position prediction

NetNGlyc sequon glycosylation prediction software was used to produce predictions on the glycan presence at all sequons present in the *Lyssavirus* genus. The result can be seen for RABV in figure 3.11, and for all other viruses in figure 3.12.

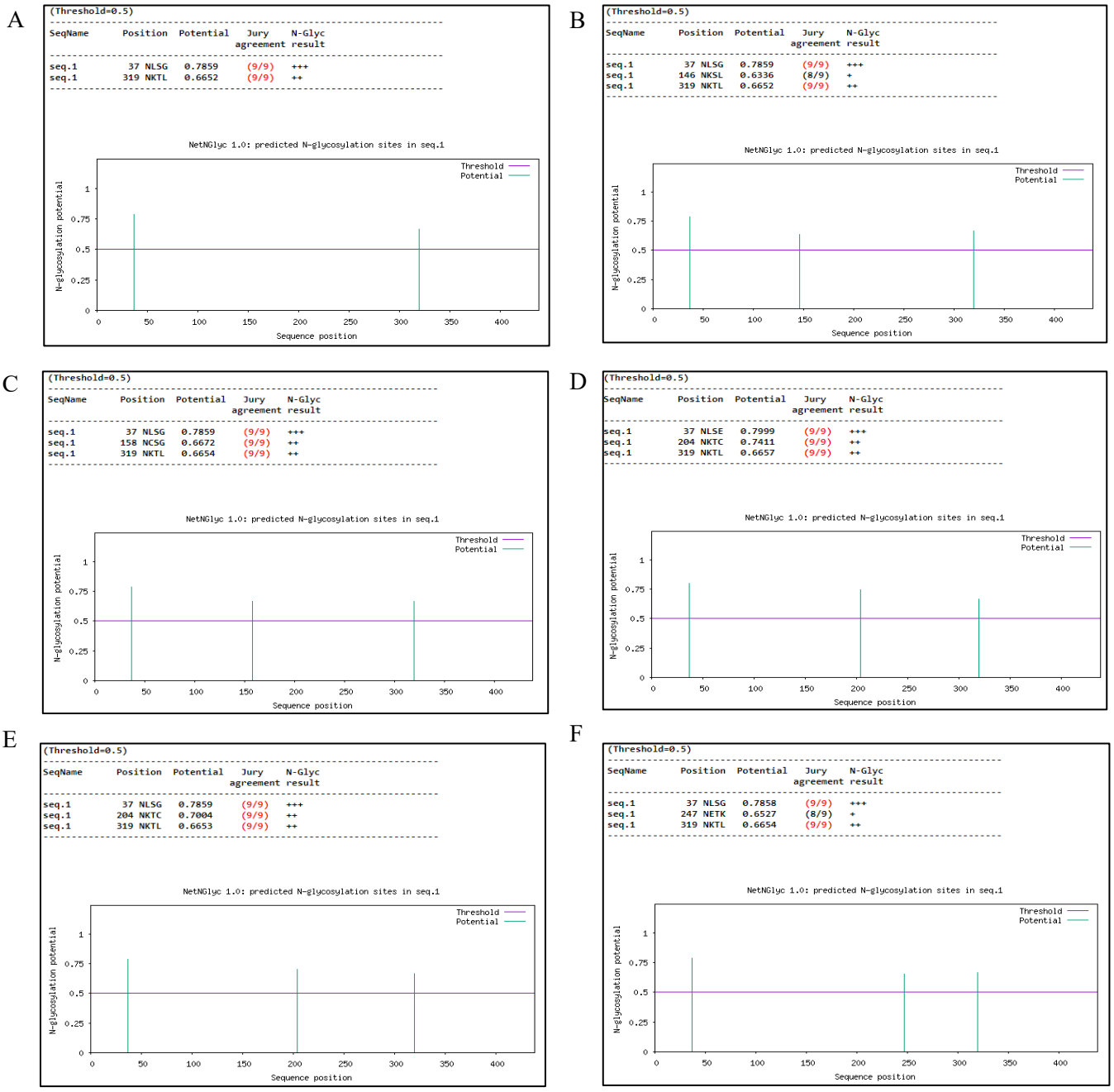
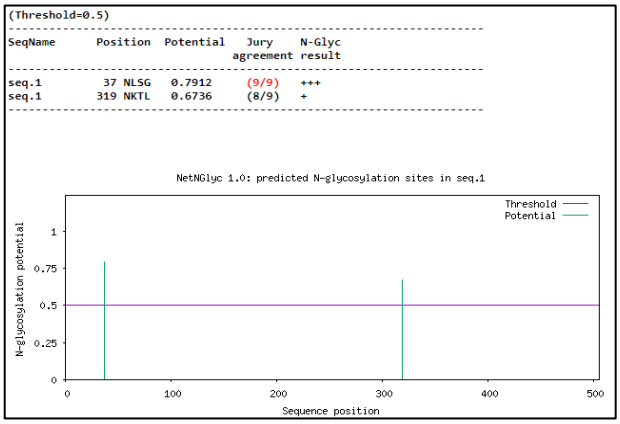


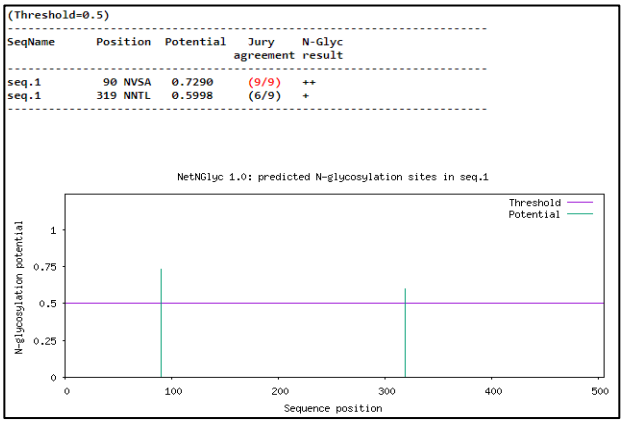
Figure 3.11: RABV NetNGlyc sequon glycosylation prediction of (A) Asn37 & Asn319 generated with consensus sequence of all sequences containing Asn37, (B) Asn146 generated with sequence AAK92080, (C) Asn158 generated with sequence ADD84790, (D) Asn202 which was not detected due to the presence of Asn204 intersecting it in sequence AAB97690 (E) Asn204 generated with sequence ACR50745, (F) Asn247 generated with consensus sequence of all sequences containing Asn247.

The RABV sequon glycosylation prediction shows that all sequons are predicted to be glycosylated when present in a RABV sequence. Confidence ratings for each sequon are high, with only a drop in significance for Asn146 and Asn247. All results fell within the range of 63-79% of potential.

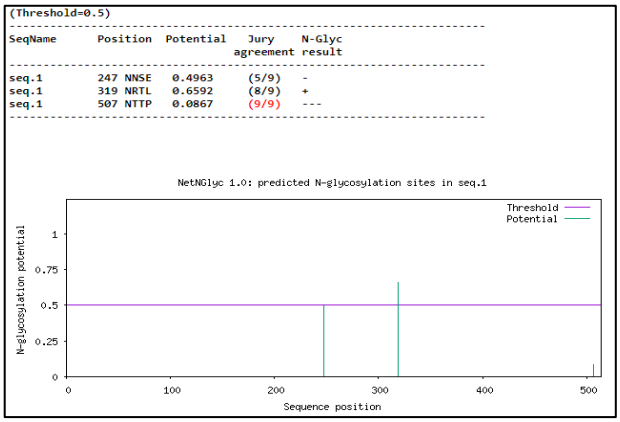
A



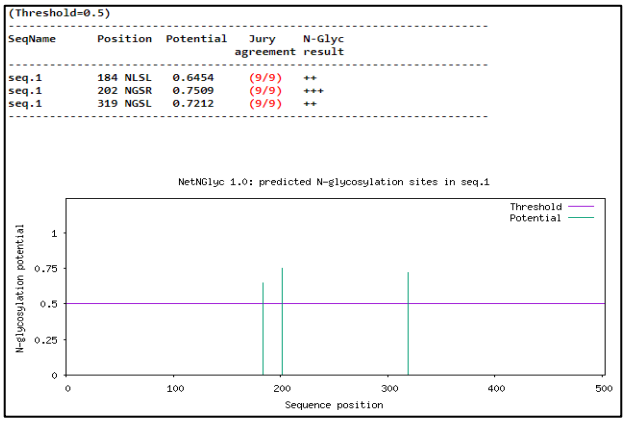
B



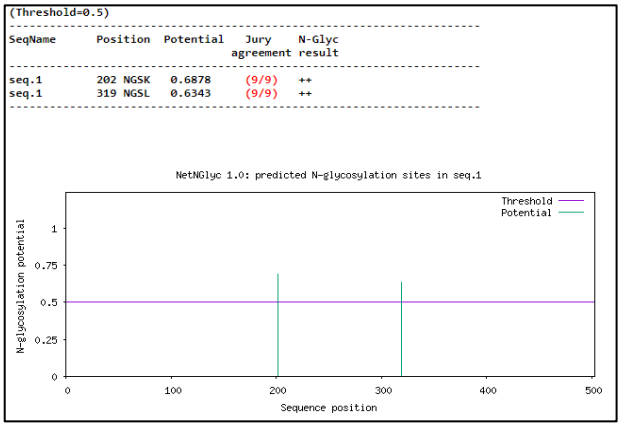
C



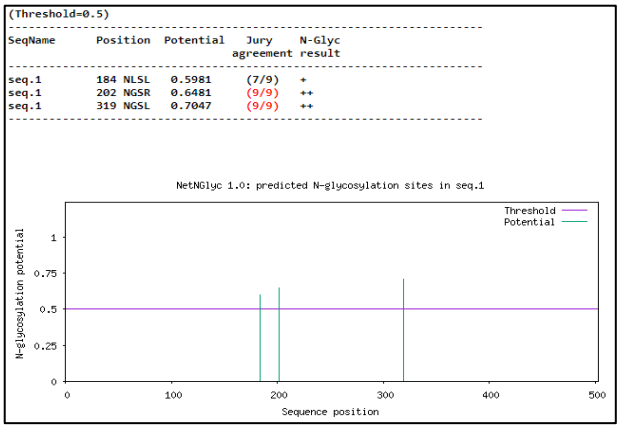
D



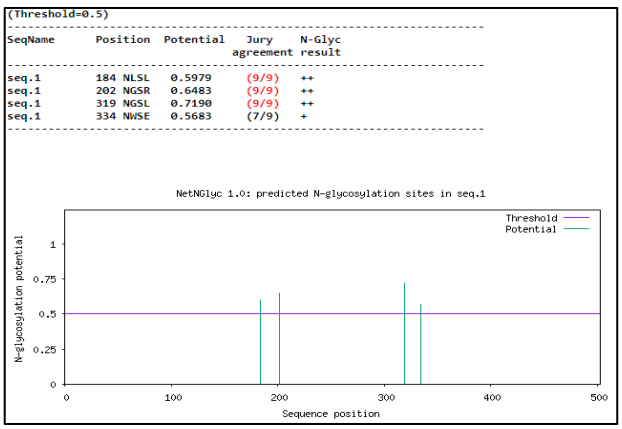
E



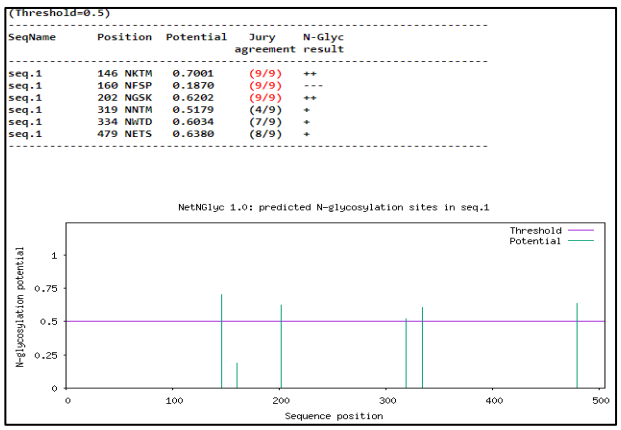
F



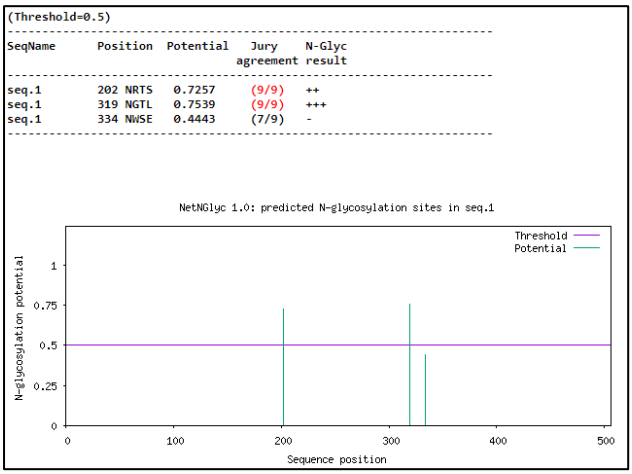
G



H



I



J

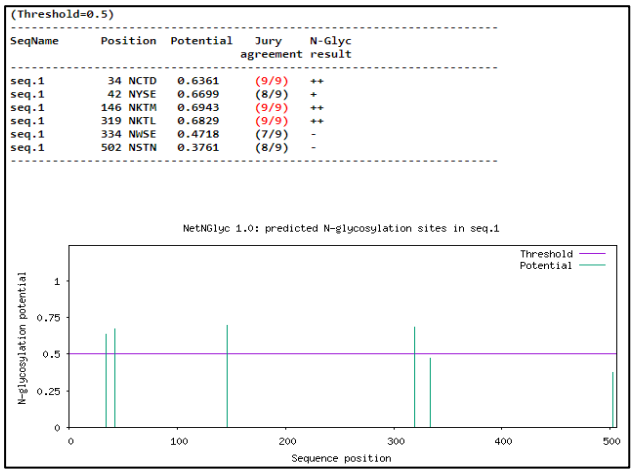


Figure 3.12: Members of the lyssavirus genus, with the exception of RABV, examined using NetNGlyc to determine the likelihood of glycosylation at each sequon through prediction based software: (A) Asn37 in BBLV generated using a consensus sequence of all BBLV, (B) Asn90 in IRKV generated using a consensus sequence of all IRKV sequences which contained Asn90, (C) Asn247 in DUVV generated using a consensus sequence of all DUVV sequences which contained Asn247, (D) Asn184, Asn202 and Asn319 in MOKV generated using the AGQ16866, (E) Asn202 in SHIBV generated using consensus sequence, (F) Asn184 and Asn202 in LBV generated using consensus sequence, (G) Asn334 in LBV generated using AAK97863, (H) Asn146, Asn160, Asn202, Asn319, Asn334, Asn479 in IKOV generated using consensus sequence, (I) Asn202, Asn319 and Asn334 in WCBV generated using consensus sequence, (J) Asn34, Asn42, Asn146 and Asn334 in LLBV generated using consensus sequence.

Glycosylation of sequons remains high in potential with the exception of a few locations. Asn247 and Asn507 in DUVV were deemed to not be glycosylated, as was the case with Asn160 in IKOV, Asn334 in WCBV, and Asn334 and Asn502 in LLBV. Sequons located above 439 amino acids would no longer be in the ectodomain of the glycoprotein and therefore not be present extracellularly.

3.3.6 Amino acid changes used in producing glycan knock outs in viruses: an in-depth analysis of past literature and their approaches.

Once glycans were determined to be of interest and their study through knockouts decided, the ideal knockout procedure needed to be determined. Upon investigation there appeared to be no standard of amino acid substitution for glycan knockouts in lyssaviruses, and in fact different amino acid substitutions were used to knock out the same sequons. A systematic review of the literature was therefore conducted to investigate how glycan knockouts were performed in virus research in general.

Figure 3.13 shows 54 viruses which were studied in 186 research papers recovered from 1494 total using the search terms provided in the methods. Results of the systematic review showed that of the 54 viruses, HIV-1 was the most studied followed by influenza with no lyssavirus

papers present in the dataset. Though a wide variety of viruses were studied, glycan knock out studies on each were low in frequency. Substitutions in the first amino acid position (N) were more common and widely used than the third position, with 211 instances of the substitution. The majority of studies selected a single substitution type for knockouts. While amino acid substitutions were still common in position 3 (T/S), totalling 57 identified, they were significantly less common than the first position. Of the studies examined, very few revealed justifications for their choices in amino acid substitutions or why position 1 or 3 was chosen.

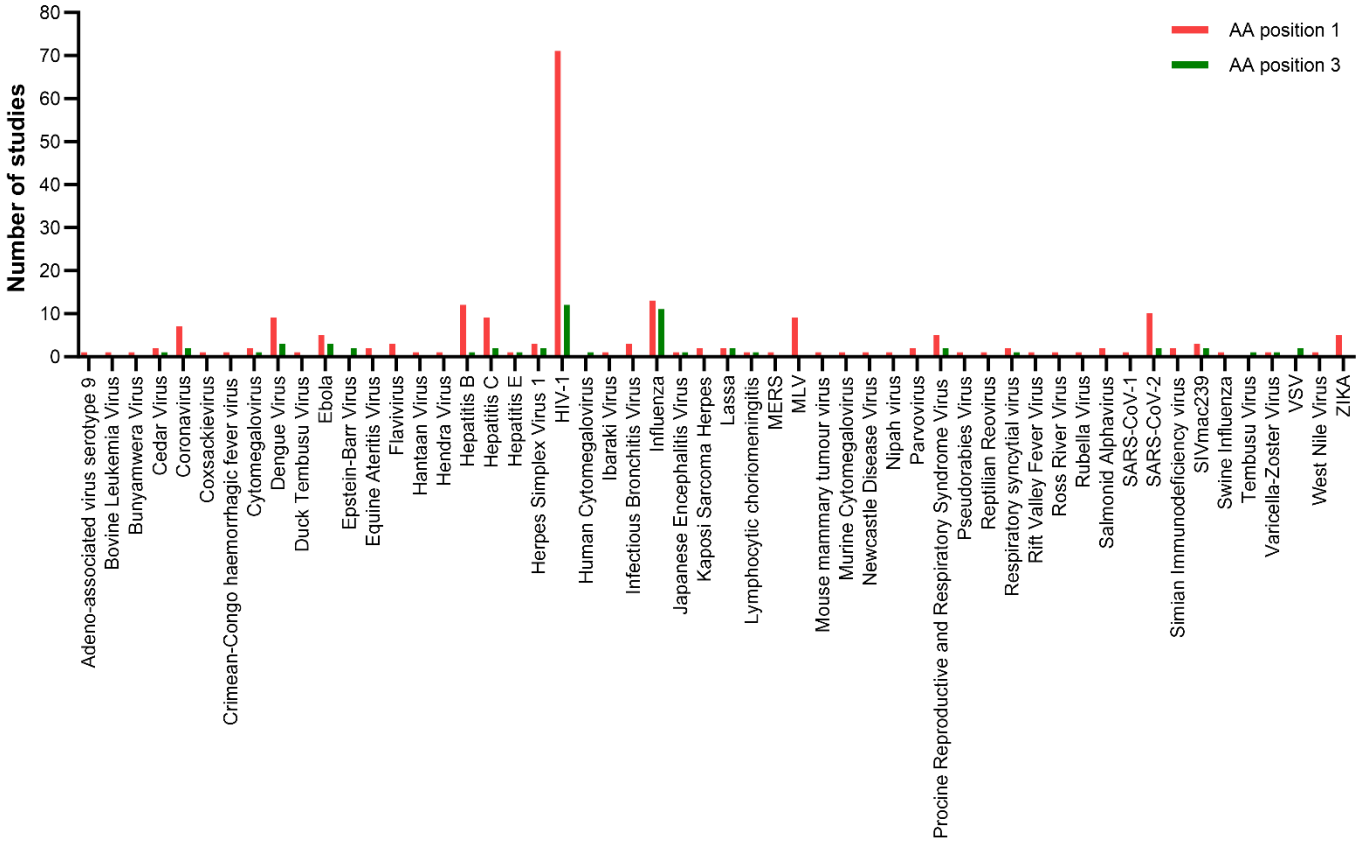


Figure 3.13: Results of a review of the literature on glycan mutation studies in viruses. The data represents the number of studies which investigated each virus listed. Studies are split into sequon substitutions in the first position of a sequon and the third position. HIV-1 was the most commonly studied, and sequon position one was the most commonly used site of substitution.

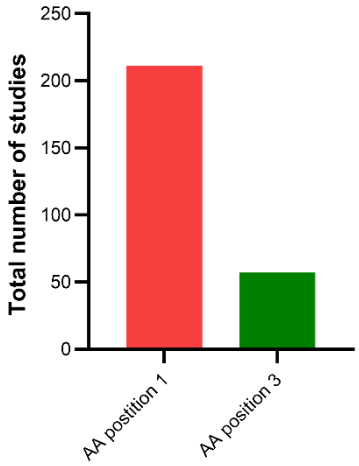


Figure 3.14: Differences in amino acid position within a sequon chosen to be substituted obtained from the literature review. 211 amino acid position one (N) substitutions were used as opposed to 57 of T/S.

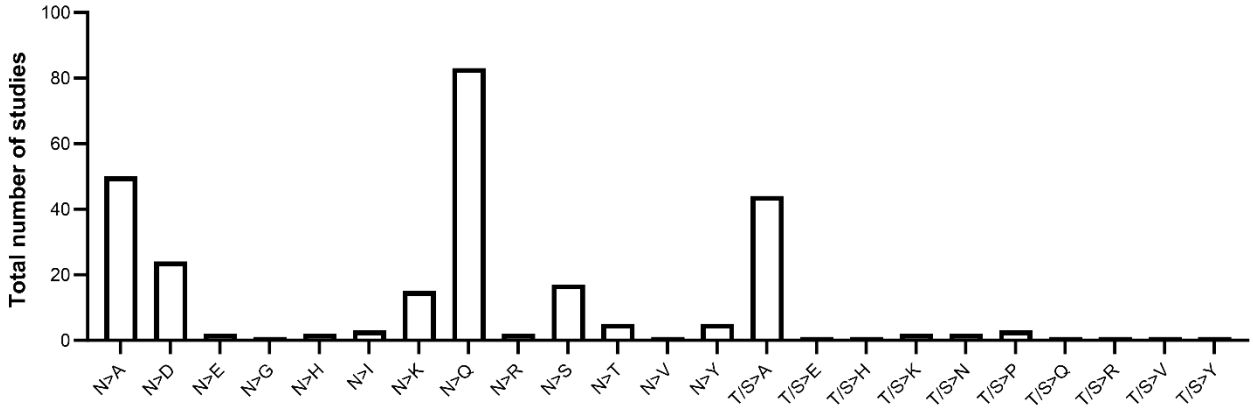


Figure 3.15: Results of a review of the literature on glycan mutation studies in viruses. This data represents the individual substitutions which had been chosen in each study recorded. Both position one (N) and position three (T/S) are presented.

The substituted amino acid was recorded for each study found which forms the data found in figure 3.15. Figure 3.15 shows a bias towards three substitutions in particular, N>Q with 83 instances, N>A with 50 instances and 44 instances of T/S>A. Other less common but still often used substitutions were N>D, N>S and N>K. While it has been shown that position 1 amino acid substitutions are the most common, there is more consensus on the procedure in position 3 with the vast majority being a single T/S>A mutant, whereas in position 1 there were high levels of variation.

Finally, the amino acid substituted, regardless of position, was measured. This produced the data in figure 3.16. The most commonly used amino acid in substitution was alanine (A) which was closely followed by glutamine (Q). These were by far the most significantly used amino acids encompassing 178 of the 267 substitutions. While aspartic acid (D), lysine (K) and serine (S) were used in several studies, they were far less significant than A and Q.

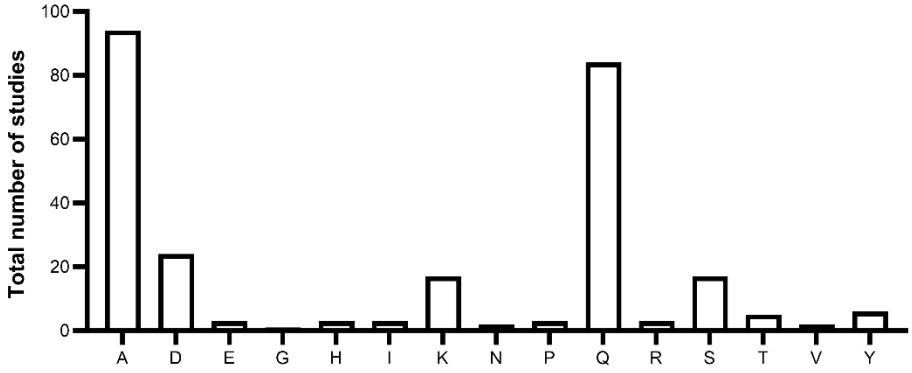


Figure 3.16: Results of a review of the literature on glycan mutation studies in viruses. The bases regardless of which position they were substituted into are presented here. Alanine becomes the most common substitution when removing the position restriction.

3.4 Discussion

Glycosylation of lyssavirus glycoproteins is an important process for effective folding and transport of the lyssavirus G protein, an essential step in virion production, as well as low pH-dependent fusion [Yamada *et al*., 2013; Li *et al*., 2021]. Despite this, research into lyssavirus glycosylation is quite limited especially compared with other viruses that have similarly important glycans. This study therefore set out to elucidate what the gamut of sequons, or potential N-linked glycosylation sites (PNGs), were in the lyssavirus genus and their amino acid formation. In addition, the study aimed to assess their context in the glycoprotein through visualisation and their relation to antigenic regions. To support these findings, it was determined how common the PNGs are in lyssaviruses and how close they are to being acquired through mutation, as well as how glycan studies perform knock outs through substitutions in preparation for a similar study in the following chapter.

3.4.1 GenBank lyssavirus genus sequon presence

Studying the sequence data obtained from GenBank has revealed that RABV contains two primary PNGs at Asn37 and Asn319 which is consistent with the findings of other studies [Yamada *et al*., 2013]. In addition to these sites there are several far less common PNGs that exist at Asn146 (0.17% of sequences), Asn158 (0.35% of sequences), Asn202 (0.02% of sequences), Asn204 (1.18% of sequences) and Asn247 (6.60% of sequences). The results also demonstrate that more divergent phylogroup II and III lyssaviruses obtain many more PNGs than phylogroup I. Glycans have been shown to have multiple functions and an increased number of them present on the G protein could improve many divergent lyssavirus functions including infectivity and immune avoidance. It has been shown that increasing the amount of glycosylation sites within RABV reduces the pathogenicity of the virus [Li *et al*., 2021]. With the genetically distant viruses of phylogroup II and III containing significantly more sites of glycosylation, perhaps this partially explains their lack of presence and reduced success in more reservoir species. The introduction of these additional sites through potential single base pair mutations may be a trade-off in order to have an increased immune avoidance but has the additional effect of reducing their ability to successfully adapt to novel reservoirs or produce more widespread infection. However, this would need to be investigated further and currently remains speculation. The reduction of virulence in response to additional glycans is also not a universal constant, as there are examples of more glycans increasing virulence such as in West Nile virus [Beasley *et al*., 2005] where glycosylation increased neurovirulence. Though it should still be noted this does not appear to have been seen in lyssaviruses.

While Asn37 is considered an important glycosylation despite the inefficiency of its addition to the glycoprotein, it is not present in over 18% of RABV sequences, further confirming the

results in the literature that it is not required for expression [Shakin-Eshleman *et al*., 1992a]. Additionally, Asn37 is only present in a single other lyssavirus, the phylogroup I BBLV. This is interesting as it appears that this glycosylation site has been lost in other lyssaviruses. When observing the amino acids present in sequences at position one of the sequon it appears that in RABV the alternative to asparagine is serine, potentially only a single nucleotide difference. An interesting follow up study could compare the presence of Asn37 in bats and dogs. Due to RABV being one of only two lyssaviruses to contain this glycan and it is uniquely adapted to many reservoirs outside of bats, perhaps this is an adaptational change in response to that and the RABV sequences which do not contain Asn37 have been obtained from bats.

3.4.1.1 Asn146

Asn146 is not considered to be a natural glycosylation site in wild type or 'street' rabies. The presence of sequences containing Asn146 may be explained by mutation introduced artificially during passages. The study by Yamada *et al*., 2014, produced mutants from RABV WT through serial passage and limited dilution cloning which introduced amino acid changes that led to the introduction of Asn146. Sequence AAK92080 sourced from GenBank has the NKSL PNG at Asn146 and was obtained from a study investigating several viral isolates. The sequence in question was taken from a bovine sample acquired in Brazil, however according to the study it was subject to 'limited passages' [Badrane & Tordo, 2001]. This combined with the evidence from Yamada *et al*., 2013, may suggest that this PNG was introduced during passages and is an adaptational change. No information is given on the cell line used to passage or the number of passages performed. However, a second study on another NKSL PNG sequence: AFN24298, also passaged in order to amplify viral samples but found no nucleotide changes when comparing sequence data between passaged and non-passaged viruses. They did not, however, perform this on all samples, only a randomly selected panel. Five of the remaining six sequences came from a single paper also taking primary viral samples from infected animals [Jamil *et al*., 2012]. Passages are not performed in this study and 5/7 sequences obtained contained the NKSL PNG, all sampled from infected bovine or goats in Bangladesh. These findings together seem to indicate the possibility of a PNG site existing in street RABV at Asn146, though its prevalence is limited. It should be noted that there is poor surveillance and monitoring in many areas with high levels of RABV and this PNG may be more common than the current sequence database suggests. Interestingly, Asn146 is present in no other lyssaviruses except 100% of IKOV and LLBV sequences, both phylogroup III lyssaviruses and therefore highly genetically divergent from RABV. This follows a trend of more PNGs present in phylogroup II and III lyssaviruses. The WebLogo figure 3.5 for Asn146 shows that the non-sequon amino acid sequence DKSL is highly conserved in this region of RABV. The 8 sequences which contain a PNG are all NKSL which potentially is only a single nucleotide change away from DKSL.

Without more phylogroup III sequences it is difficult to determine if this is a glycan site of importance for more divergent lyssaviruses.

Investigation into the PNG with NetNGlyc 1.0 provides insight into the likelihood that a glycan is attached. NetNGlyc uses artificial neural networks to measure and predict PNGs through identification of surrounding residues and sequence context. This process has been cross-validated and achieves an 86% accuracy when predicting glycosylated sequons. Overall accuracy when accounting for prediction of non-glycosylated sites, NetNGlyc achieves 76% accuracy; because of this the results shown do not achieve the standard 95% confidence interval and should not be taken as fact. However, this process does provide insight into the potential spectrum of glycosylation within lyssaviruses. According to NetNGlyc, Asn146 has a 63% potential for glycosylation with 8/9 jury's in agreement with the results. These results indicate a reasonable likelihood that Asn146 is glycosylated, and when combined with the findings of Yamada *et al*., 2014, this outcome is certain.

3.4.1.2 Asn158

Asn158 was present in 16 sequences, 0.35% of the total sequences, which is higher than Asn146 but equally not a common occurrence. The PNG consisted of NCSG amino acids in all sequences and when compared to the WebLogo figure 3.5 it appears that position one, two and four of the PNG are highly conserved. Though variation occurs in position three, both threonine and serine are present in the majority of sequences with the rest being leucine and therefore not a PNG. The lysine conserved in position one is, like in Asn146, only a single nucleotide alteration away from becoming asparagine which strengthens the possibility of this genotype occurring naturally. Investigating the studies behind the identified sequences appears to show that the Asn158 is isolated frequently in the Flury strain of RABV. Flury is a strain of RABV which is often used as a live attenuated virus strain in vaccination such as Flury high egg passage or Flury low egg passage [Tao *et al*., 2010; Bavarian Nordic A/S, 2023]. There were sequences which were not identified as the Flury strain but were passaged in cell lines and a single study with wild type isolates [Benmansour *et al*., 1992]. Glycosylation prediction using NetNGlyc found that the PNG had a 66% chance of being glycosylated, the same potential as Asn319 in this amino acid configuration which is known to be efficiently glycosylated. This evidence would suggest that this PNG is an adaptational change introduced through artificial means via serial passage which would explain its low prevalence.

3.4.1.3 Asn202

PNG Asn202 is present in 0.02% of RABV sequences, however it appears in 100% of MOKV, SHIBV and LBV of phylogroup II and IKOV and WCBV sequences from phylogroup III. The WebLogo results for Asn202 in RABV reveal the majority of sequences contain a lysine at

position one which, similar to Asn158, is potentially only a single nucleotide change from the introduction of an asparagine. However, the required thr/ser at position three of the sequon is less common with approximately 50% of sequences split between serine and glycine, this change is also potentially a single nucleotide mutation. The Asn202 present in the RABV sequence data is from a single study using CVS-24, which has been passaged for decades in suckling mice, and passaging lyssavirus in cell culture using BHK-21 cells [Morimoto *et al*., 1998]. Being present in a single sequence and that sequence being artificially produced is conclusive evidence this is not a significant RABV mutation in wild-type sequences.

In phylogroup II and III sequences which contain PNG Asn202 there is little variation between amino acid sequences in positions one, two and three with all but WCBV. The WCBV sequences contain NRT compared to all other PGII and III lyssaviruses which had NGS, a significant difference as the S mutation will greatly impact glycosylation efficiency [Kasturi *et al*, 1995]. Asn202 is unable to be predicted by NetNGlyc due to its intersection with Asn204 in RABV which the software prioritises. Whether this means that it is likely that Asn204 is glycosylated over Asn202 or not, and if the presence of both sequons intersecting impacts glycosylation remains to be seen. However, due to the lack of Asn204 in phylogroup II and III lyssaviruses the site is unobstructed and is estimated to be glycosylated at a 75% potential for MOKV, 68% SHIBV, and 64% for LBV with all 9 jury verdicts in agreement. Prior to confirmation with physical lab assessments of the glycans at these positions, it can be speculated with reasonable though not >95% certainty that these PNGs are glycosylated. This is important due to the presence of Asn202 near the antigenic region IIa in RABV, though whether this region is antigenic in phylogroup II or III lyssaviruses also requires confirmation.

3.4.1.4 Asn204

Asn204 is present in 1.18% of total RABV sequences which is significantly more than the other low frequency PNGs discussed so far, however it is still an insubstantial number of total sequences. Also novel to this PNG is variation in amino acid sequence. Of the 54 sequences which contain a PNG at Asn204, five of them were NRTC in sequence, with the other PNGs consisting of NKTC. Investigating the studies behind these sequences reveals that of the five NRTC sequences, three were reported in the same dataset. The samples were taken from infected *Cynictis penicillata* feline, and the other *Suricata suricatta* in South Africa [Van Zyl *et al*., 2010]. The samples were extracted from brain tissue directly or passaged in mouse brain tissue, no distinction is made between which samples underwent which process in the study. The final two sequences come from an unpublished manuscript and another study performing phylogenetic analysis [Fischer *et al*., 2018]. Interestingly, the sequence from Fischer *et al*., 2018, also originated from a feline sample *Feli silvestris catus* originating from South Africa.

Of the remaining NKTC sequences, 21 of them also originated from Fischer *et al*., 2018. There was also an additional sequence which contained NKTC isolated from Africa, the Botswanan AMM70623 sequence. It is conceivable that particular strains of RABV originating in the South African region have developed a novel PNG through mutation. This is supported further by the presence of a highly conserved third position threonine in the sequon, with glycine and serine both occupying the majority of position one amino acids; however, both of these are multiple nucleotide substitutions away from Asn.

The remainder of the sequences come from multiple studies, many of which use CVS strains of RABV. This is significant as Asn204 has previously been reported as a potential lab adaptation and glycosylation of Asn204 has been shown to induce transfer of RABV G proteins to the cell surface membrane in fixed virus studies [Hamamoto *et al*., 2015]. However, this study did not investigate the sites impact on infectivity and neutralisation, focusing more on the impact on glycoprotein mobility. This leaves a gap in the current understanding of the impact of this PNG on phenotypic responses such as infection that could be investigated. The implication of the results of the sequence database examination appears to be inconclusive as to the presence of Asn204 in wild type virus. Though sequences exist which have been acquired from tissue samples, in many cases it is unclear if passaging was involved and whether or not there was opportunity for strain adaptation to occur.

3.4.1.5 Asn247

Asn247 has a more significant presence in the RABV sequences than other less common sites with a 6.6% overall appearance. It is also present in 88.8% of DUVV sequences which is another phylogroup I lyssavirus. This sequon has been cited in early literature as a part of the classical RABV glycans [Shakin-Eshleman *et al*, 1992a, Wojczyk *et al*., 2005] but following continued sequencing of RABV it has become clear that this glycosylation is not essential, nor is it present in the vast majority of sequences. It has been shown that the introduction of Asn247 is important for certain cell line adaptations such as in BHK-21 and MDCK-II [Nitschel *et al*., 2021]. Asn247 also impacts the type of glycans attached to other sites such as Asn319 when present [Wojczyk *et al*., 2005]. When Asn247 is the sole glycan present on the RABV glycoprotein the type of glycan attached is 90% high mannose, in contrast when Asn319 is solely present only 35% are high mannose. When combined, the sites both contain 87% complex glycans. This interaction extends to Endo H sensitivity with different patterns depending on the presence of Asn247 with Asn319 or Asn37. However, beyond this discovery the actual mechanisms behind their interactions remain to be determined. This interaction does however need to be appreciated when performing knockout experiments.

With a significant number of sequences, 6% of RABV, it should not be ruled out that this adaptation can occur naturally and may exist commonly in some RABV strains, which is particularly important due to its potential when glycosylated to provide benefits to infection [Yamada *et al*, 2013]. An interesting follow up study could isolate these sequences and determine their source to rule out lab-based strains. Using NetNGlyc to determine the potential for glycosylation determines this site to have a 65% potential for glycosylation and it has been demonstrated before that Asn247 is efficiently glycosylated [Shakin-Eshleman *et al*., 1992a]. Asn247 which is present in DUVV showed a 49% potential which is below the 50% cut off, deeming it to be a negative result.

The inclusion of predicted glycans at PNGs on the RABV trimeric glycoprotein assists in visualising the potential immune system inhibiting functionality, showing how the glycans may cover the antigenic regions which they surround. However, it should be noted that glycans are flexible and are not rigid structures, therefore these figures should not be taken as totally representative of the real glycan structures.

3.4.2 Antigenic regions

Antigenic regions are an important part of lyssavirus neutralisation. Interestingly, as shown in the sequence data, they are present close to many PNGs. Asn34, Asn37, Asn42, and Asn334 are all located in amino acid regions commonly attributed to the antigenic regions IIb (34-42) and III (330-338). Additionally, Asn202, and Asn204 are close to antigenic site IIa (198-200). These findings along with the evidence in other viruses that glycans play a role in immune avoidance suggest that there is a possible link between these PNGs and the antigenic regions. However, this link has not yet been explored in the literature and warrants further investigation. As shown in the glycoprotein crystal structures with predicted glycans attached, the sites could easily obscure these antigenic regions from the host immune system.

Addition of an NLG at Asn336 which is in antigenic region III has been speculated to be the cause of a hindering of the interaction between the glycoprotein and neutralising antibodies [Wang *et al*., 2019]. This study found that mutations surrounding the antigenic region III (330-338) produced the strongest resistance to neutralising antibodies and sera. They express concerns that in RABV, mutations in this region could confer resistance to current vaccines. This finding perhaps links to the 100% conserved Asn334 glycans which reside in the middle of this antigenic region. It could be that the glycan in this position adds to the resistance seen by divergent lyssaviruses against current vaccines and therapies. Indeed, Arg333 has been shown to be a key determinant of virulence in most representative fixed viruses [Tao *et al*,. 2010]. This could support the idea that the glycan site at Asn334 reduces virulence in divergent

lyssaviruses, reducing their capability to produce novel reservoirs but increasing their resistance to phylogroup I neutralising factors.

3.4.3 Amino acids at position X and Y of Asn-X-T/S-Y

While investigating the presence of glycan sequons in lyssaviruses the amino acid residues which make up the sequon were also recorded. This is due to evidence that the amino acid position of X and Y in the sequon Asn-X-T/S-Y have importance and impact the efficiency of the glycosylation at that site [Shakin-Eshleman *et al*., 1996; Mellquist *et al*., 1998]. The results uncovered a proline at position 4 of Asn160 in IKOV, indicating the likelihood of this sequon being glycosylated being very low. It was also found that sequence variation at these sites was low, with only 5 out of the 42 PNGs having any differences in amino acid.

Of the third position threonine and serine amino acids, over 50% (22) of the 42 PNGs had a serine, with the remaining 20 having threonine. However, with the removal of Asn319, threonine is only present in 7 of the remaining 26 PNGs, with the remaining 19 containing the serine amino acid. This is important as the third amino acid has a large impact on the efficiency of glycosylation [Kasturi *et al*., 1995]. Threonine at the third position produces a dramatic increase in glycosylation efficiency and therefore it can be concluded that highly efficient glycosylation is not necessarily required at many of the PNGs in the genus, though Asn319 appears to bias towards the inclusion of a threonine implying it may require the opposite. Interestingly, only phylogroup II lyssaviruses have a serine in the Asn319 sequon.

3.4.4 Analysis of the literature for methodology when altering amino acids through SDM

In preparation for experimentation involving the production of lyssavirus glycan knockout mutants using site-directed mutagenesis (SDM), a systematic review of the literature was performed to determine the consensus on amino acid choices for knockouts. It is clear from the findings that glycan studies have been performed in a wide variety of viruses, however there is a focus in the literature on HIV-1, influenza and hepatitis, likely due to their importance as pathogens and highly glycosylated surface proteins. Lyssaviruses did not appear in the search results and, though there are studies investigating glycans, the review demonstrates a lack of research into the particular area of glycan knockouts.

Amino acid changes are important especially in glycans where, as has been discussed previously, each change at position X, threonine/serine, or Y in the sequence can impact efficiency and therefore the results of any knockout experiment. It is therefore important to consider knockout choice carefully. The justification for choices of particular amino acid substitutions in SDM were found to be lacking in a significant number of studies. Only a few

posed reasoning behind their decisions, often citing decisions based on: natural occurrence of the proposed mutant through cell culture passage, therefore likely to be tolerated [Prentoe *et al*., 2019], sequence data analysis identified this amino acid as natural substitute [Doran *et al*., 2014], and also alanine as a substitute due to its general low impact nature [Shakin-Eshleman *et al*., 1992a].

Though the systematic review did not include any lyssavirus papers, there have been some studies which have performed glycosylation knockouts [Yamada *et al*., 2013; Shakin-Eshleman *et al*., 1992b]. In Yamada *et al*., 2013, an N>Q substitution was used to knock out Asn37 and Asn319. The substitution was justified, unlike a significant proportion of studies, as an attempt to retain the polarity and size of the amino acid chain. As seen in figure 3.17 this substitution is appropriate and indeed mimics the size and polarity as closely as possible to the original asparagine. Though this is the case, the study by Shakin-Eshleman *et al*., 1992b, utilised a T/S>A substitution in order to investigate Asn37, Asn247 and Asn319. When these two papers are compared in the context of the systematic reviews findings, they mirror the overall state of glycan knockout research in viruses. Both studies have knocked out the same glycans and yet used differing substitutions which may have had an impact on their findings.

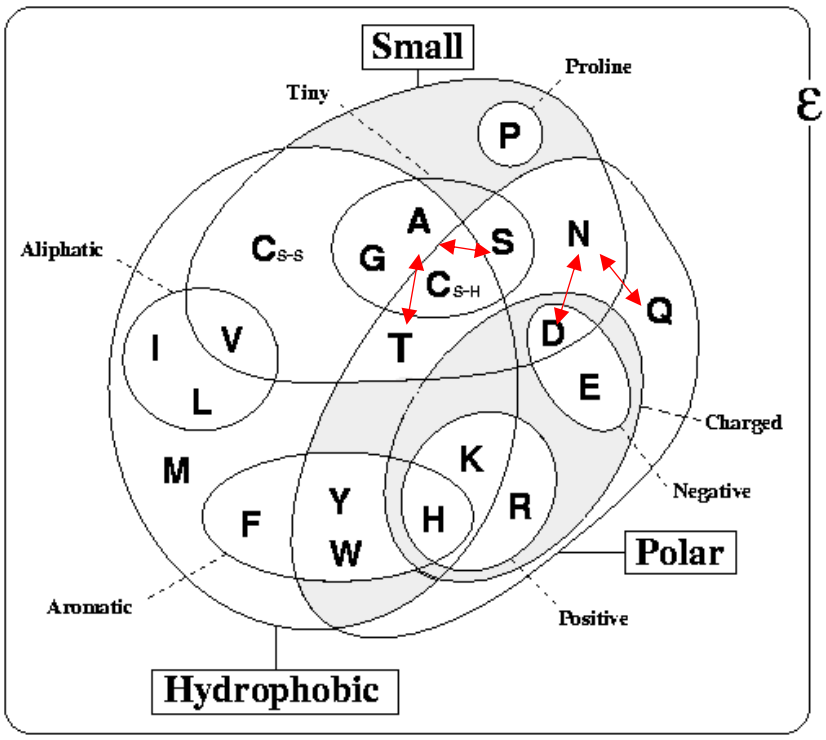


Figure 3.17: Amino acid nucleotides and properties [Livingstone & Barton, 1993].

Interestingly there were very few proline substitutions, despite this being an amino acid known to significantly reduce glycosylation in sequons, however this is likely due to the fact that proline can impact conformation and structure which would be undesirable in a knockout

experiment [Shakin-Eshleman *et al*., 1996; Bause *et al*., 1983]. Amino acid features and their similarity to asparagine appears to be the correlation with studies substitution choices. This can be seen when looking at amino acid characteristics as seen in figure 3.17, the most commonly used amino acids such as aspartic acid (D), alanine (A), and glutamine (Q) all have relatively similar characteristics to asparagine (N). However, as can be seen from the figure there are no 'perfect' substitutions, in essence one which has all of the same characteristics.

As has been shown both in *lyssavirus* research and wider virus research as a whole there are discrepancies in the choices made on how best to knock out PNGs. These discrepancies in amino acid substitution choices present a quandary, does the amino acid substitution impact the results of glycan knock out studies and if so to what proportion? There is potential for the choice of substitution to significantly impact the results of these studies and perhaps produce outcomes which otherwise may not have been the case if a different amino acid had been chosen. This was therefore investigated in the following chapter, with a focus on the most common knockout methods, N>Q, N>D and T/S>A. Additionally, glycosylation in lyssaviruses have been shown to play an important role in several viral factors. However, studies have as yet not produced infectivity and neutralisation impact data. It was therefore the aim of the following chapter to produce a full panel of CVS-11 glycosylation knockouts to investigate the changes to these phenotypic responses.

3.4.5 Future work

Though the work in this chapter manages to produce some interesting conclusions, it is not comprehensive. Future work could be performed to improve and expand on the current research here. For example, the systematic review of papers could be refined further. There are a few different parameters that would bring back interesting results or otherwise improve on the current results. For one there are a number of wider questions which could be answered through broadening the glycan knock-out to not just be related to viruses; how are glycans knocked out in bacteria and human proteins. Additionally, more search terms could be used to broaden the papers, and perhaps through the use of bioinformatics and command line programming, more papers could be sorted without the need for manual analysis. This study was limited through attempting to find a reasonable sample size which could be manually assessed and therefore undoubtably neglected to include some papers because of the narrower search terms. An example of this is the inclusion of the word substitution which, when removed, doubles the returned papers though it is unclear at present if these papers would be relevant. Additionally, the body of research into glycosylation in viruses is ever expanding, in fact during the process of this review 6 novel papers were published on the subject. This study could therefore be

performed again after some time to investigate changes in trends or confirm these results in future novel research.

This could also be performed on the lyssavirus sequence study which could be looked at regularly with new additions added to the table. This would be especially important in comparison to the systematic review as some lyssaviruses only had a few sequences available and novel sequences could dramatically change how the landscape of lyssavirus glycans looks. As lyssavirus surveillance gets better and more sequences become available, more conclusions may be drawn about the presence and importance of glycans in the lyssavirus genus.

It is clear that the bigger picture of glycosylation within lyssaviruses is far from complete, with much greater surveillance required and many more sequences acquired of less common lyssaviruses before more solid conclusions on the state of PNGs in the genus can be formed.

Chapter 4: Sequon modification in lyssaviruses impacts infectivity and neutralisation

4.1 Introduction

Glycosylation in lyssaviruses has been demonstrated to be important for the genus as a whole, with high conservation of sequons in all three phylogroups and in varied locations in the lyssa genome. The *in silico* findings of this study also suggest that there may be an important role for glycans in immune avoidance. To investigate this further, several lyssavirus PNG modifications were produced.

Lyssavirus glycan modifications have been performed in studies previously. A core text of this field is a study by Shakin-Eshleman *et al*., 1992a, which investigated lyssavirus glycosylation and the impact of its efficiency on cell surface expression. To achieve this, they produced a panel of RABV PNG knockout mutants using site-directed mutagenesis, substituting the threonine/serine for an alanine. The strain of RABV used in the study contained Asn37, Asn146 and Asn319. The findings of the study showed that all three of those sequons could be glycosylated and that Asn37 did not need to be efficiently glycosylated in order to support cell surface expression. It was concluded that sequons which did not become glycosylated or were glycosylated in an inefficient manner were still capable of supporting specific biological functions. The study also investigated the interactions between sequons when removed, attempting to determine if removal of one site impacted another; it was concluded that there was no impact and that each sequon is core glycosylated independently. Though the study contained a lot of key findings, the lack of infection and neutralisation impact left important questions unanswered. Also mentioned is the justification for the use of the alanine substitution, where it was chosen due to its perceived minimal impact on protein configuration, however the conclusion refers to a potential experiment using other amino acid substitutions to confirm their findings, until this study this was never performed.

In a series of studies, changes in the glycosylation of lyssaviruses and their impact on pathogenicity were investigated. Initially, serial passage was used to produce an attenuated RABV from a wild-type virus [Yamada *et al*., 2012]; this resulted in the addition of a novel PNG at Asn194. Following this introduction, the study investigated its impact on pathogenicity and concluded it was likely that the introduction of the novel glycan site reduced pathogenicity and was unlikely to be replicated in wild-type viruses, being negatively selected. The follow up study performed similar experiments, assessing the impact of the deletion of Asn37 and Asn319 and the addition of Asn158, Asn204 and Asn247 [Yamada *et al*., 2013]. They found that glycans were attached to Asn158, Asn196 and Asn247, however Asn204 was inefficiently glycosylated.

Asn37 was found, as had been previously shown [Shakin-Eshleman *et al*, 1992b], to be inefficiently core-glycosylated and when Asn319 was removed, had even weaker glycosylation. It was also shown that introduction of N-glycosylation enhanced virus production. This was greatly increased in Asn196 and Asn247 and slightly increased in Asn158 and Asn204. Deletion of Asn37 reduced viral production and removal of Asn319 significantly reduced viral production. Interestingly, Asn158 was shown to be potentially cell line specific, as it is produced through serial passage in chicken eggs (HEP-flury strain) however when adapted to BHK-21 cells it lost this sequon.

Due to the impact of Asn37 being removed and its importance to viral production, Yamada et al 2013 speculates that the sequon is optimised in circulating WT RABV strains, hence the conservation of amino acid residues shown in sequence data. It is further speculated that the absence of the sequon in some WT strains is also beneficial to viral circulation in bats however the mechanisms are unknown.

The final study of the series investigated the addition of Asn146 and modification of the efficiency of Asn37 through an L38R substitution. Their findings showed that the substitution provided Asn37 with efficient core-glycosylation which was related to cell adaptation as well as reduced pathogenicity when compared to the original WT. The novel PNG at Asn146 did not impact pathogenicity which is linked to this site being reported in WT samples, indicating its presence would not impact natural pathogenicity.

As demonstrated in chapter 3, antigenic sites share a potentially interesting relationship with sequon positions across the lyssavirus genus. The work by Evans *et al*., 2018, investigated how swapping the antigenic regions of lyssavirus phylogroups II and I could induce susceptibility or provide resistance to neutralisation. Their findings showed that demonstrable intra and inter-phylogroup cross-neutralisation were observed, however the inter-phylogroup cross-neutralisation was limited. PVs with phylogroup I glycoproteins had phylogroup II antigenic sites I, II, III and IV and were able to be neutralised by phylogroup II antibodies. Site II swapped PPs were neutralised to an efficient degree, however site IV was poorly neutralised. Site II appears to be immunodominant.

Studies investigating pathogenicity, cell surface expression, viral production, cell-to-cell transmission and cellular internalisation have been performed previously when investigating novel or current lyssavirus sequons [Yamada *et al*., 2012, 2013, 2014; Shakin-Eshleman *et al*., 1992a; Faber *et al*., 2005]. These studies have found there to be an important relationship between glycosylation and all these factors in lyssaviruses; however no known studies have performed a comprehensive investigation into the impact of glycosylation on infectivity and neutralisation. Furthermore, as stated in Shakin-Eshleman *et al*., 1992a, and determined in

chapter 3, it is clear that the type of amino acid substitution introduced may indeed impact the results of knock out experiments and there needs to be a comprehensive panel produced to identify if this is the case and suggest a consistent knock-out substitution selection which can be used going forward. It was therefore the aim of this chapter to investigate the impact of lyssavirus sequon knockouts on the infectivity and neutralisation outcomes, assess whether amino acid substitution selection impacted said outcomes, and finally introduce phylogroup site swaps in order to further elucidate the nature of lyssavirus glycans in immune avoidance.

4.2 Materials and Methods

4.2.1 Site-directed mutagenesis

In order to produce mutations which would lead to glycosylation knockouts and PNG introductions, site-directed mutagenesis (SDM) was performed with the New England Biolabs (NEB) Q5® Site-Directed Mutagenesis Kit on the pI.18 CVS-11 lyssavirus glycoprotein plasmid. Primers used can be found in tables 4.1 and 4.2 below. SDM Q5 PCR was performed with a total volume of 25µl. The reagents consisted of: 12.5µl Q5x2, 1.2µl forward and reverse primer (10µm), 9µl H_2O, 1µl template (10ng/µl). Annealing temperature varied and was based on the NEB predictions though 57°C was often used when products were not being produced which led to success. PCR protocols used in the SDM reaction are as follows:

- Initial denaturation: 95°C for 30 seconds
- 95°C 10 seconds
- Annealing temperature: 57°C 20 seconds
- 72°C 2 min 30 seconds

35 cycles

- Final extension 72°C for 2 min

A KLD reaction was performed following manufacturers recommendations with the exception that the reaction was performed for 1 hour instead of the recommended 5 minutes, at room temperature. After incubation was complete the resulting product was used in a transformation protocol which can be found in 4.2.2.

Primers for the NEB SDM reaction were designed using the NED primer designer software NEBaseChanger, associated with the SDM kit.

Table 4.1: Forward and reverse primers for the glycan knockout mutants produced with each substitution listed.

Mutant required	Substitution used	Forward primer (5' > 3')	Reverse primer (5' > 3')
CVS-11 - Asn37	T/S>A	TACCAACCTGgccGAGTTCTCCT	CATCCTTCATCCTCCACAACCAG
CVS-11 - Asn204	T/S>A	AGGGAACAAGgctTGCGGCTTTG	TTGGATGCTCTCTTCCCTCTG
CVS-11 - Asn319	T/S>A	ATTCAACAAAgccTTGATGGAGG	ATGGTATATGCTTTTCCAAAC
CVS-11 - Asn37	N>Q	AGGATGTACCcaaCTGTCCGAGTT	TCATCCTCCACAACCAGG

CVS-11 - Asn204	N>Q	ATCCAAAGGGcaaAAGAC TTGCG	GCTCTCTTCCCTCTGCTA
CVS-11 - Asn319	N>Q	TACCATATTCcagAAAACC TTGATGG	TATGCTTTTCCAAACCCT G
CVS-11 - Asn37	N>D	AGGATGTACCgACCTGTC CGA	TCATCCTCCACAACCAGG
CVS-11 - Asn204	N>D	ATCCAAAGGGgACAAGA CTTG	GCTCTCTTCCCTCTGCTA
CVS-11 - Asn319	N>D	TACCATATTCgACAAAAC CTTG	TATGCTTTTCCAAACCCT G

Table 4.2: Forward and reverse primers for the addition of PNGs from PG III lyssaviruses to CVS-11.

Mutant introduced to CVS-11	**Amino acid sequence**	**Forward primer (5' > 3')**	**Reverse primer (5' > 3')**
LLBV Asn34	NCTD	accgacCTGTCCGAGTTC TCCTAC	acagttTTCATCCTCCACA ACCAG
LLBV Asn42	NYSE	agcgaaCTCAAAGTGGGA TACATCTC	gtagttGAACTCGGACAG GTTGGT
LLBV + IKOV Asn146	NKTM	accatgCACTCAAGGGTC TTCCCT	tttgttATATGGGTCCAAAT CTGTC
LLBV Asn334	NWSE	CGAACTGGAGCGAGG ATCATCCC	CCTCGCTCCAGTTCGG AACTGAC
IKOV Asn160	NFSP	agcccgGTGTCCTCTACCT ACTGCTC	gaagttGCACTTTCCGCC AGGGAA
IKOV Asn202	NGSK	agcaagACTTGCGGCTTT GTGGAT	cccgttGGATGCTCTCTTC CCTCTG
IKOV + WCBV Asn334	NWTD	gagcGAGATCATCCCCTT CAAAAG	cagttCCGGAACTGACTT GTAGTG
WCBV Asn202	NRTS	accagcACTTGCGGCTTT GTGGAT	cctgttGGATGCTCTCTTC CCTCTG

All SDM reactions were sent for Sanger sequencing, following colony screen, with Source Biosciences in order to both verify the correct substitution had occurred but also to ensure no other erroneous changes to the plasmid had occurred, ensuring that the correct sequence was present with only the desired substitution.

4.2.2 Plasmid transformation

Transformation protocol was performed using the heat shock method. Chemically competent E. coli (Stellar competent cells, Takara Clontech) were removed from -80°C storage and thawed on ice. Once thawed, for SDM reactions, 5µL of the KLD reaction was added to an aliquot of 500µL of cells. The cells were carefully mixed to avoid cellular damage and then incubated on ice for 30 minutes. Heat-shock was performed by placing the Eppendorf's containing the cells in exactly 42°C water for 30s. Following this step, the cells were returned to ice for 5 min, followed by mixture with 950µL room temperature SOC media. Cells were incubated at 37°C in a shaking incubator set to 250rpm for 60 min. Cells were agitated and then streaked or pipetted onto LB-ampicillin selection media agar plates and incubated inverted overnight at 37°C. The subsequent plates could then be used for extraction or colony screen PCR.

4.2.3 Colony PCR

Colony PCR was performed on transformed bacterial colonies containing mutant plasmids. The PCR reaction mix contained 1.5µL 10x buffer (Qiagen), 0.6µL of both forward and reverse primers (10pmol/µL), 0.6µl dNTPs (2.5mM concentration), 0.075µL of HotStar *Taq* polymerase (Qiagen) and finally 11.6µL of nuclease-free water. Colonies were grown using both streaked and spread methods onto selection media plates consisting of LB-ampicillin agar. Random selection of a colony was then performed which was added to the PCR mixture. Chosen colonies were seeded into a new LB-ampicillin agar dish in a grid pattern to enable identification and selection of successful mutants for proliferation and extraction. PCR reaction protocol used was as follows:

- Initial denaturation: 95°C for 15 min
- 95°C 20s
- Annealing temperature: 60°C
- 72°C 30s/kb

35 cycles

- Final extension 72°C for 2 min

PI.18 primers are as listed below:

F: TCCATGGGTCTTTTCTGCAG

R: CAGGCGTGACACGTTTATTG

Colony seeded plates were incubated overnight at 37°C. The colony PCR output was examined using gel electrophoresis and strong bands were sequenced. Successful mutants were picked from the seeded grid plate after the overnight incubation and grown in LB-ampicillin broth in a shaking incubator at 37°C overnight in preparation for plasmid extraction. Growth media volume was dependent on the size of the extraction reaction used (50mL for midi prep, 5mL for mini prep).

4.2.4 Agarose gel electrophoresis

Agarose gels consisted of 2g of agarose powder in 100mL 1x TAE buffer for a 2% concentration gel. The mixture was boiled using a microwave and allowed to cool. 2.5µL of Ethidium bromide was added to the gel mixture after cooling but prior to setting, mixing until no longer visible. The gel was then poured and allowed to set for 45 minutes. The resulting gel was placed into a gel electrophoresis tank containing 1x TAE buffer and 5µL of ethidium bromide mixed thoroughly.

Samples were prepared for running with the addition of 2.5µl of 10x FastDigest Green Buffer (Thermo Scientific) loading dye. 5µL of the subsequent volume was then loaded onto the gel. Either side of sample sets, 3.5µL of GeneRuler DNA ladder mix (Thermo Scientific) was also loaded onto the gel.

Gels were run at 90V for 36 min. The resulting banding was visualised using UV light and imaged with a camera in the dark. For colony screening, bright clear bands of the appropriate size were noted and the corresponding sample sent for sequencing. Source Bioscience was used for all sequencing and their guidelines for sample preparation were followed.

4.2.5 Plasmid extraction using midi/mini prep kits

Plasmid extraction was performed using GenElute plasmid mini/midi-prep kits (Sigma-Aldrich) depending on the yield required. For midi-preps, bacteria were grown in 50mL LB-ampicillin broth which was a solution of pre-made LB broth and 50µL of 100µg/mL working concentration of ampicillin, mini prep volumes were divided by 10. Prior to spinning, if required, 1mL stocks of bacteria were taken from the broth and stored at -20°C with 3 drops of glycerol. Manufacturers protocol for both kits was followed precisely including optional spin steps, however optional wash stages were not performed. The final step, plasmid resuspension, was performed using 750µL of elution solution for midi-prep and 250µL elution for mini-prep. Extracted plasmid was measured using a Nanodrop spectrophotometer. All plasmids were stored at 4°C following extraction.

4.2.6 Cell lysis

Producer cells which were used to produce PVs were harvested from the culture dishes, after the supernatant had been removed, and placed in Eppendorf's. Cells were then centrifuged for 5 minutes at 300g followed by the removal of supernatant and resuspension in PBS. The cells were centrifuged a second time with the same settings, and resuspended in lysis buffer (for 250mL, 1.65g Trizma HCL, 0.24g Trizma base, 25mL 100mM EDTA p*H* 8.0, 2.1g NaCL, 2.5mL Triton x-100), which had a protease inhibitor tablet/10mL of buffer and incubated for 30 minutes. A further centrifuge step was performed at 16,000 RPM for 20 minutes. The supernatant was then aliquoted into eppendorfs and stored at -20°C.

4.2.7 Protein quantification

Supernatants extracted from PV producer cells had protein quantification performed on them for use in downstream protocols. This was done through the use of the BCA protein estimation kit (Thermo Fisher Scientific) according to the manufacturer's instructions. The assay was performed in a low-binding 96 well plate (Thermo Fisher Scientific). Cell lysate produced using the method in 4.2.6 was diluted 1:5 in lysis buffer. A standard curve was produced through serial dilution of a BSA standard (Thermo Fisher Scientific) to 2mg/mL, 1.5mg/mL, 1mg/mL, 0.75mg/mL, 0.5mg/mL, 0.25mg/mL and 0.025mg/mL using lysis buffer. 10µl of protein sample/standard was added per well followed by the addition of 200µl of BCA working reagent was added. The plate was incubated at 37°C for 30 minutes with absorbance then being measured using a FLUROstar Omega plate reader set to 595nm. From these readings, a standard curve was produced which was used to map the concentrations of unknown samples.

4.2.8 SDS-PAGE

Electrophoretic separation of protein samples was performed using pre-cast polyacrylamide gels (find the ones I used). Protein samples were mixed with 4 x Laemmli buffer (277.8mM Tris-HCL, p*H* 6.8, 44.4% (v/v) glycerol, 4.4% SDS, 10% β-ME) and heated at 95°C for 10 minutes. The resulting samples were then loaded onto the pre-cast gel which had been placed in a Bio-Rad tetra cell gel tank filled with SDS-PAGE running buffer (25 mM Tris base, 192mM glycine, 0.1% SDS). In addition to the samples, 4µl of Spectra Broad Range Molecular weight marker (Thermo Fisher Scientific) was also loaded. The gel was then run at 140V for 1 hour.

4.2.9 Western blotting

After the completion of the SDS-PAGE protocol, a 20% Mini-PROTEAN TGX Precast Protein Gel, 15-well, 15 µl, 4--20% precast polyacrylamide gel (Bio-Rad) was equilibrated in transfer buffer which had been chilled (25mM Tris, 192mM glycine, p*H* 8.3) for 10 minutes at room

temperature. A PVDF membrane (Amersham) was submerged in and activated by 100% methanol for 3 minutes. The membrane was also equilibrated om chilled transfer buffer alongside blotting paper. Following this, all three components (gel, membrane and blotting paper) were assembled onto a Trans-Blot Semi dry cell (Bio-Rad) which was run at 100 mA for 1 hour. After this transfer process was completed, the membrane was incubated in a blocking solution (PBS, 0.1% TWEEN 20, 10% skimmed milk) overnight at 4°C. The blocking buffer was then removed and 10mL of primary antibody which had been diluted to a working concentration in buffer (PBS, 0.1% Tween 20, 5% skimmed milk) was added to the membrane in a container which was incubated for 1 hour at room temperature on a rocking apparatus. This solution was removed after incubation and the membrane washed using TPBS (0.1% Tween-20) three times each for 5 minutes. A secondary antibody which was HRP-conjugated was then added to the membrane in the same solution as the first, and also incubated for 1 hour at room temperature on a rocking apparatus. This was removed following incubation and the membrane washed a further three times.

Detection of membrane bound HRP was performed using an ECL substrate (Thermo Fisher Scientific). To achieve this, 500µl of each provided ECL solution was mixed prior to the completion of the final membrane wash. The membrane was placed on a clear plastic sheet which had been cleaned and dried. The ECL substrate was then added to the membrane in droplets, ensuring that the membrane was completely covered, followed by a 1 minute incubation at room temperature. A second clear plastic sheet was then added on top of the membrane and HRP activity was visualised in a gel box.

4.2.10 Vacuum/dot blotting

Vacuum/dot blotting was attempted in order to produce results using a conformational antibody and to circumvent the difficulties which the western blots were producing. To achieve a dot blot, a PVDF membrane was activated followed by the application of 5-20µl droplets of protein which were pipetted onto the surface in a grid pattern. This membrane then followed the stages of a western blot beginning from the activation stage. Dot blotting was performed as per bio-rads recommendations when using PVDF membranes [Bio-Rad Laboratories, 2000]. A vacuum method was also attempted where the PVDF membrane was loaded onto a Bio-Rad Bio-Dot microfiltration device with a low-level vacuum applied which concentrated the droplets into more define circles. While the membrane remained intact, no results were achieved using either method.

4.2.11 Lectin capture ELISA

Lectin capture of CVS-11 G protein expressed on PVs was performed. Lectin from *Galanthus nivalis* (GNA) and wheat germ agglutinin (WGA) (Sigma-Aldritch) were diluted in 0.05 M carbonate-bicarbonate, p*H* 9.6, to achieve a final concentration of 10µg/mL. This was then transferred to a flat bottom Nunc MaxiSorp plate (Thermo Fisher Scientific) at 50µl/well volume. This was incubated overnight at 4°C followed by the removal of the coating solution. Wells were then blocked for 2 hours at room temperature using 5% skimmed milk TPBS, 200µl/well volume. Blocking solution was removed and the plate was washed with 200µl of TPBS followed by 90 mins incubation at room temperature with diluted antigen. Sample was removed and the plate was washed three times with 200µl of TPBS each wash followed by the addition of 50µl of primary antibody diluted in TPBS. This was incubated for 1 hour at room temperature. Washing was repeated and 50µl of secondary AP-conjugated antibody added which had been diluted in TPBS with a further 1 hour incubation. A final washing step was performed using TTBS. Detection of bound AP was performed through the addition of 50µl of pNPP substrate (Sigma-Aldrich) to each well. Absorbance was measured at 405 nm using a FLUROstar® Omega plate reader.

4.2.12 qPCR

qPCR was performed using cDNA produced from RNA extracted from the supernatant of PP producer HEK293T cells. Supernatant containing PVs was removed from the culture dish after the standard 72 hour incubation. This supernatant was then taken into a clean separate extraction laboratory and RNA extraction was performed with the RNEasy minikit (Qiagen) kit on ice. The process was performed as per manufacturers recommendations and additional steps. Care was taken to ensure no supernatant was lost during the process through the use of regular centrifuge spins between steps.

Following extraction, the RNA was treated with DNase (NEB). The reaction was set up on ice and consisted of 10µl of RNA, 2µl of DNase I reaction buffer (10X), 0.5µl DNase I (RNase-free) and 7.5µl of nuclease free H2O for a 20µL reaction size. This was incubated for 1 hr 10 min at 37°C in a heat block followed by a 10 minute denaturing stage at 75°C. RNA was stored at -80°C

The resulting treated RNA was then converted into cDNA using EcoDry random hexamers (Takara Clontech). Lyophilised random hexamers were resuspended in 5µL of RNase-free water, 2.5µL of which was then added to 7.5µL of RNA extract. The mixture was incubated on a Bioer XP cycler version 2008-1.3 heatblock at 42°C for 1 hour. The reaction was then terminated through a 10 min 70°C incubation. cDNA was stored at -20°C.

qPCR was performed using the resulting cDNA and a Luna universal qPCR master mix (NEB). Reaction size was 5µL, 2.5µL of Luna master mix, 0.125 of forward and reverse primers, 0.5µL of cDNA and 1.75µL of H2O. qPCR was run with the following cycling conditions:

- 95°C for 60 seconds
- 95°C for 15 seconds
- 60°C 30 seconds

40 cycles

- Melt curve 60-95°C

Melt curve analysis was performed following the qPCR run and indicated that single amplicon production was successful with no erroneous curves.

QPCR primers used:

Forward primer TTGTGCCAGAGTCCTTCGAT

Rev TAGGATCTCTGGCATGCGAG

4.2.13 Statistical analysis

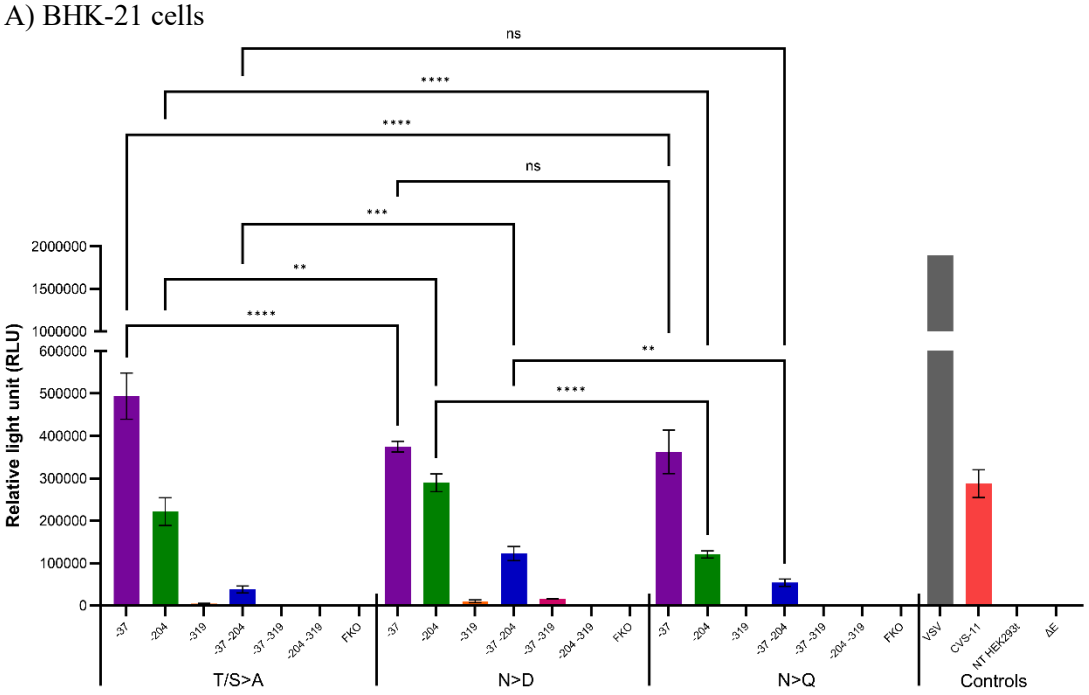
Data was analysed using GraphPad Prism 10. Normality was assumed due to the low N value and data was statistically analysed using either a one-way or two-way ANOVA followed by Tukey multiple comparisons post hoc test. A value of $P \leq 0.05$ was considered to indicate a statistically significant difference. All experiments were repeated technically and biologically.

4.3 Results

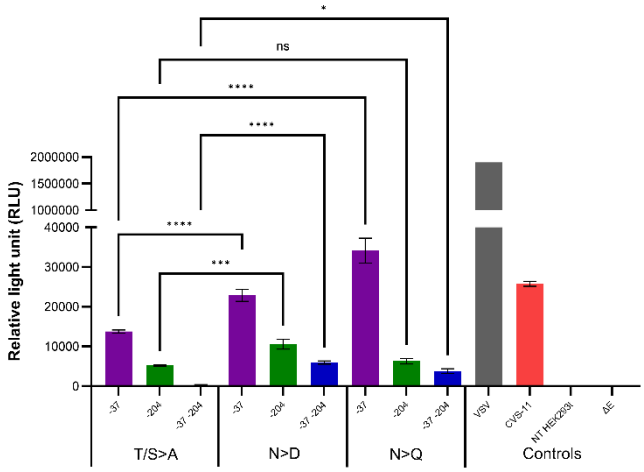
4.3.1 Infection assay with sequon knockout panel

The panel of mutant substitutions was completed and verified, each one using a different substitution to perform the knockout, T/S>A, N>D and N>Q. The panels were all possible permutations of the knockouts based on the sequons present in CVS-11. PVs were produced with the plasmids which contained the mutants and used in an infection assay in the cell lines BHK-21, HEK293t and MDCK. A further bat cell line, DesNi had an infection assay performed also; however this did not produce results. The infection assay generated the results shown in figure 4.1.

A) BHK-21 cells



B) HEK239T cells



C) MDCK cells

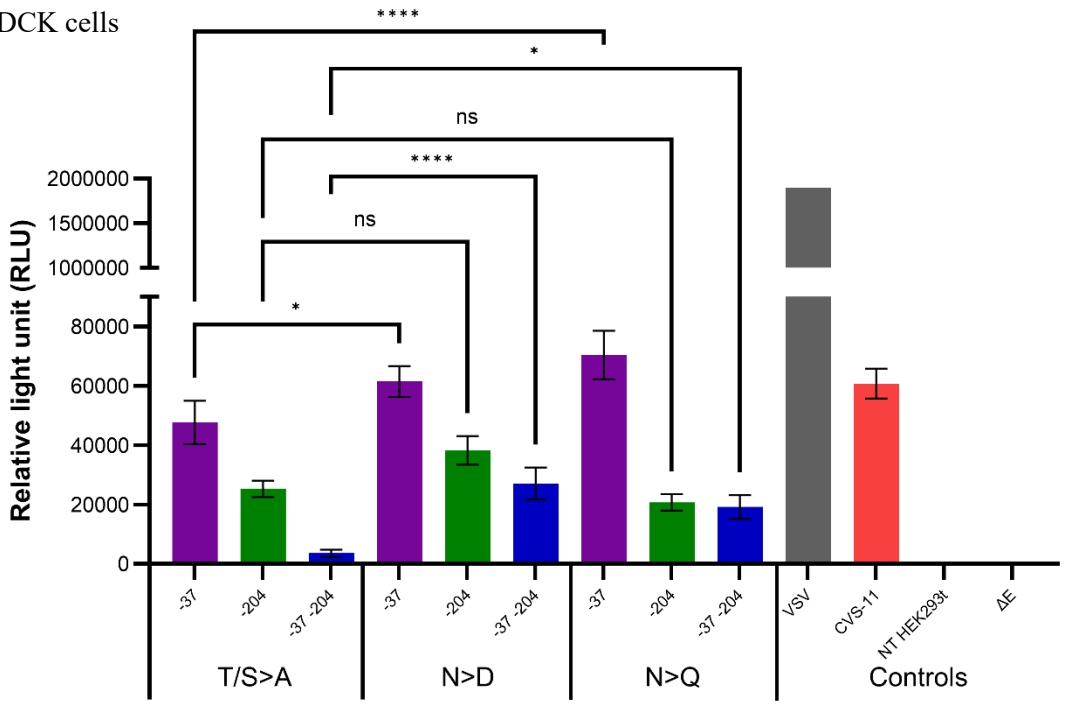


Figure 4.1: Infection assays using a full panel of CVS-11 knockout mutants produced using three different substitutions; T/S>A, N>D and N>Q. Assay was measured using relative light units and was both repeated and performed in triplicate. Three different cell lines were used (A) BHK-21, (B) HEK239T and (C) MDCK. Infections were read at 2800 gain. FKO = full knock-out. Data is presented as mean±SD (assumed to be normally distributed) and statistically analysed using a two-way ANOVA followed by Tukey multiple comparisons post hoc test (****P<0.0001; *P<0.05; ns = not significant). Full significance data can be found in supplementary tables S.3, S.4 and S.5 for A, B and C respectively. A VSV positive control was used in each experiment, alongside a delta envelope negative, and no transfection.

All infections were read on a Fluostar Omega luminometer at 2800 gain. The BHK-21 results showed a clear pattern of infection which was retained through each of the three types of substitution, with -Asn37 being the highest RLU of the mutants followed by -Asn204 and finally -Asn37 & -Asn204. The removal of Asn37 did not decrease infectivity and in fact increased it, with a more significant increase found in T/S>A substitution over the alternatives. Removal of Asn204 reduced infectivity in both T/S>A and N>Q, however the results fluctuated between them significantly and infectivity was not impacted by the N>D substitution. Any removal of Asn319 produced little to no infectivity levels. The combined removal of Asn37 and Asn204 resulted in the greatest loss of infectivity with a high reduction in all three substitution types. There was again fluctuation between the substitutions with N>D being less impacted than the alternatives. Negative controls behaved as expected and the VSV positive infected significantly higher.

4.3.2 Neutralisation curves of sequon knockouts utilising NHP sera in BHK cells

An examination of neutralisation profiles for the mutants in the panel which were infectious was also performed, shown in figure 4.2. The neutralisations were performed using NHP sera which had been produced from a RABIPUR challenge in the medoid vaccine trial. The sera was used in a three-fold dilution series beginning at 1/50. The neutralisation experiment was performed in three cell lines, BHK-21, HEK239T and MDCK. The original non-mutant CVS-11 was run alongside for comparison.

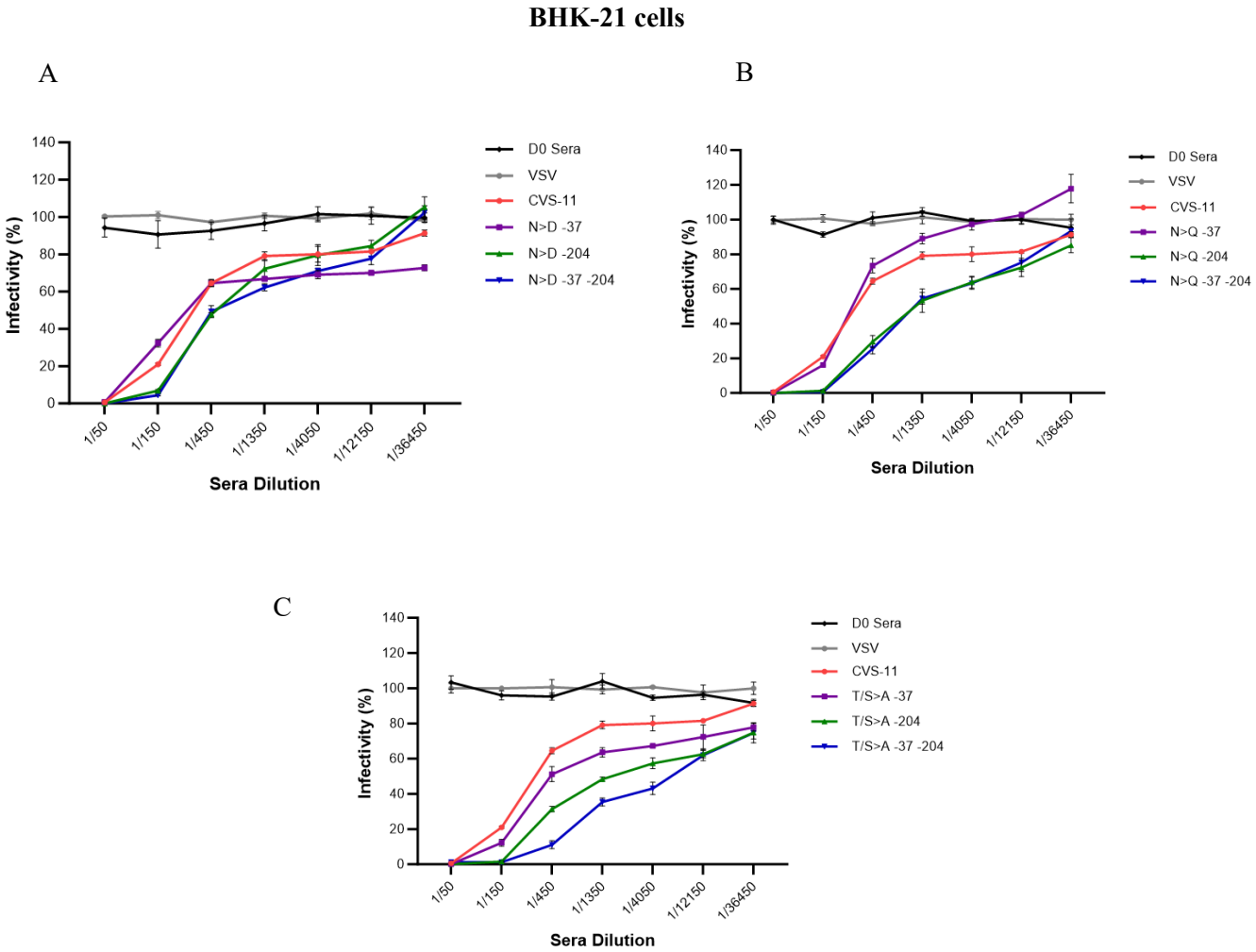
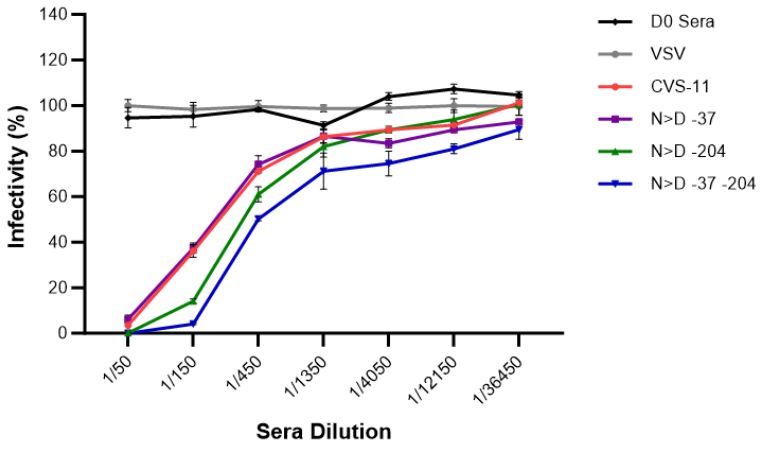


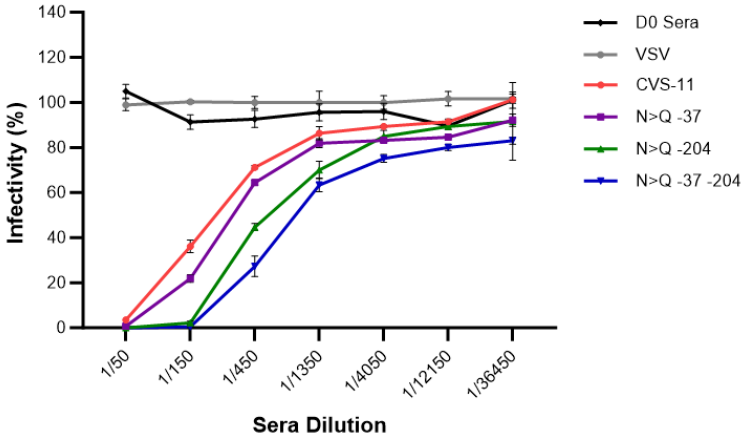
Figure 4.2: Neutralisation curves of three different sequon substitution knockouts, (A) N>D, (B) N>Q and (C) T/S>A. The assay was performed in BHK-21 cells. A three-fold dilution series was used to dilute the sera, which was NHP sera that had been generated in response to RABIPUR inoculation. The results were normalised to a PBS control to represent 100% infectivity. A negative control of day 0 sera was used, in addition to VSV pseudo-virus as a non-specific virus control. Both controls behaved as expected.

HEK239T cells

A



B



C

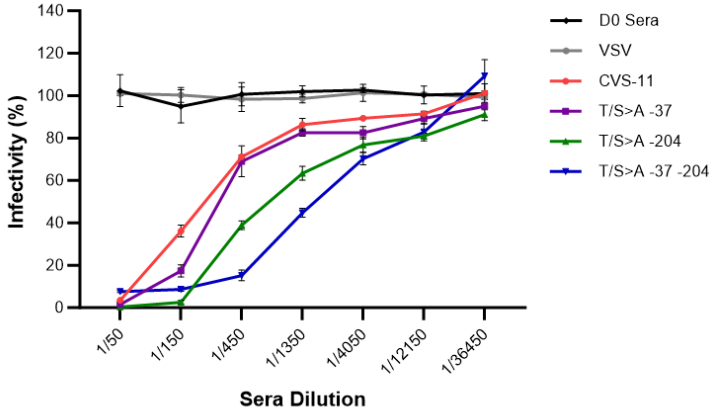
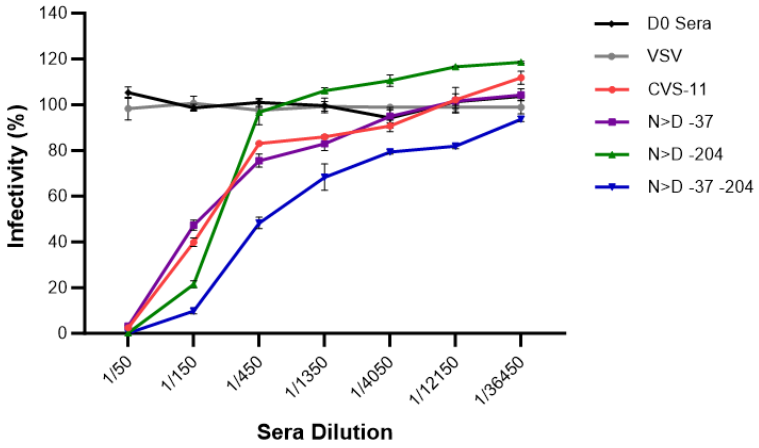


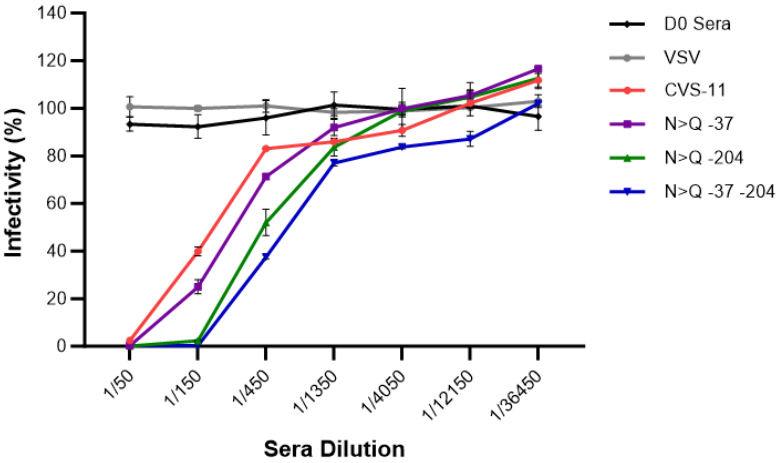
Figure 4.3: Neutralisation curves of three different sequon substitution knockouts, (A) N>D, (B) N>Q and (C) T/S>A. The assay was performed in HEK239T cells. A three-fold dilution series was used to dilute the sera, which was NHP sera that had been generated in response to RABIPUR inoculation. The results were normalised to a PBS control to represent 100% infectivity. A negative control of day 0 sera was used, in addition to VSV pseudo-virus as a non-specific virus control. Both controls behaved as expected.

MDCK cells

A



B



C

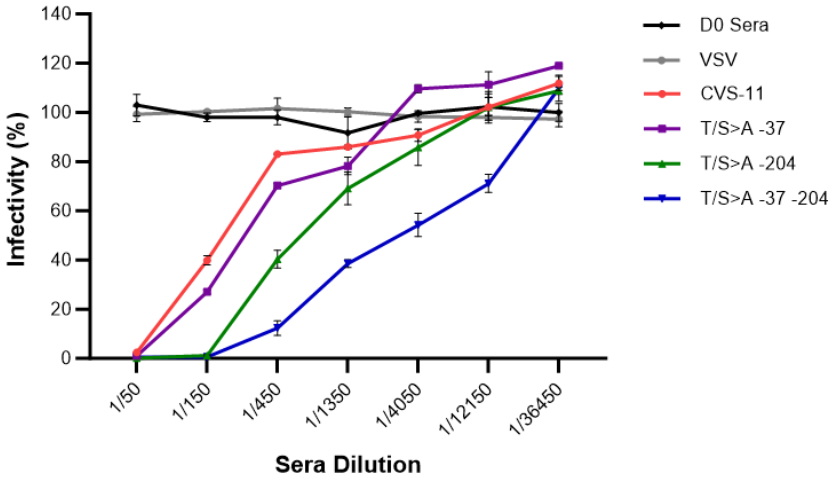


Figure 4.4: Neutralisation curves of three different sequon substitution knockouts, (A) N>D, (B) N>Q and (C) T/S>A. The assay was performed in MDCK cells. A three-fold dilution series was used to dilute the sera, which was NHP sera that had been generated in response to RABIPUR inoculation. The results were normalised to a PBS control to represent 100% infectivity. A negative control of day 0 sera was used, in addition to VSV pseudo-virus as a non-specific virus control. Both controls behaved as expected

4.3.3 IC50 values in knockouts

IC50 values represent the point in a neutralisation assay at which 50% infectivity/neutralisation occurs. The IC50 values were compared for each of the three parameters tested to examine the significance between the neutralisation assays. These parameters were: any significant differences between the mutants (-Asn37, -Asn204, -Asn37 & -Asn204), the chosen substitutions (N>D, N>Q, T/S>A) or the cell lines (BHK-21, HEK239T, MDCK).

4.3.3.1 IC50 values between mutants in BHK-21s

The mutations -Asn37, -Asn204, and -Asn37 & -Asn204 IC50 values from sera neutralisation in BHK-21s were compared for significance as can be seen in figure 4.5. Figure 4.5A demonstrates that the -Asn37 & -Asn204 knock out has no significant difference when compared to CVS-11 or -Asn37. However, -Asn204 produces significant increase in neutralisation susceptibility. The results in figure 4.5B shows a pattern of change in neutralisation profiles within the mutants, with -Asn204 and -Asn37 and -Asn204 showing increased susceptibility to neutralisation when compared to CVS-11 and -Asn37. The removal of Asn37 generated no significant impact on neutralisation using any substitution. Figure 4.5C has no significance between mutants with the exception of -Asn37 & -Asn204 which again produces a marked increase in neutralisation susceptibility, with 50% infectivity not occurring until around 1/2600 sera dilution.

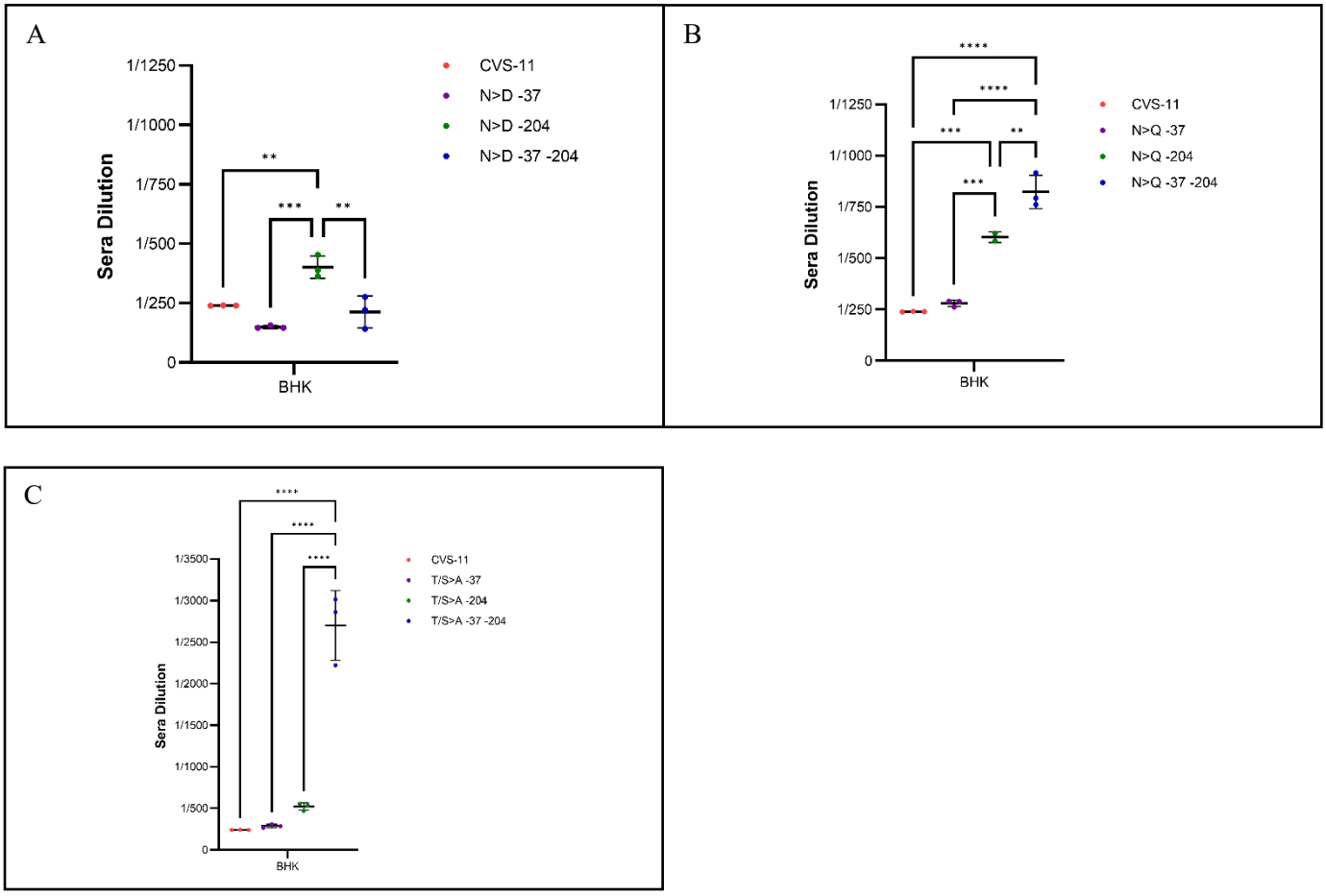


Figure 4.5: IC50 values generated from figure 4.2 neutralisation curves using NHP sera that had been produced through challenge with RABIPUR vaccination. All experiments were performed in BHK-21 cells using different substitutions (A) N>D (B) N>Q and (C) T/S>A. Data is presented as mean±SD (assumed to be normally distributed) and statistically analysed using a one-way ANOVA followed by Tukey multiple comparisons post hoc test. (****P<0.0001; ***P<0.001; **P<0.01; *P<0.05; ns = not significant). Technical repeats in triplicate were used.

4.3.3.2 IC50 values between substitutions in BHK-21s

To examine the impact amino acid substitution choice had on neutralisation, the IC50s of mutants with different substitutions were compared. The results can be seen in figure 4.6. Comparing the different substitutions reveals that, despite knocking out the same sequons, the substitution choice does often have a significant impact. In figure 4.6A the N>D substitution significantly differs in neutralisation susceptibility to both N>Q and T/S>A, which cluster closely together. This is repeated in figure 4.6B. For the -Asn37 & -Asn204 combination, the pattern changes and reveals T/S>A to have a much greater susceptibility to the sera than both other substitutions, N>D and N>Q. The results overall show that different substitutions at different sites generate varying results.

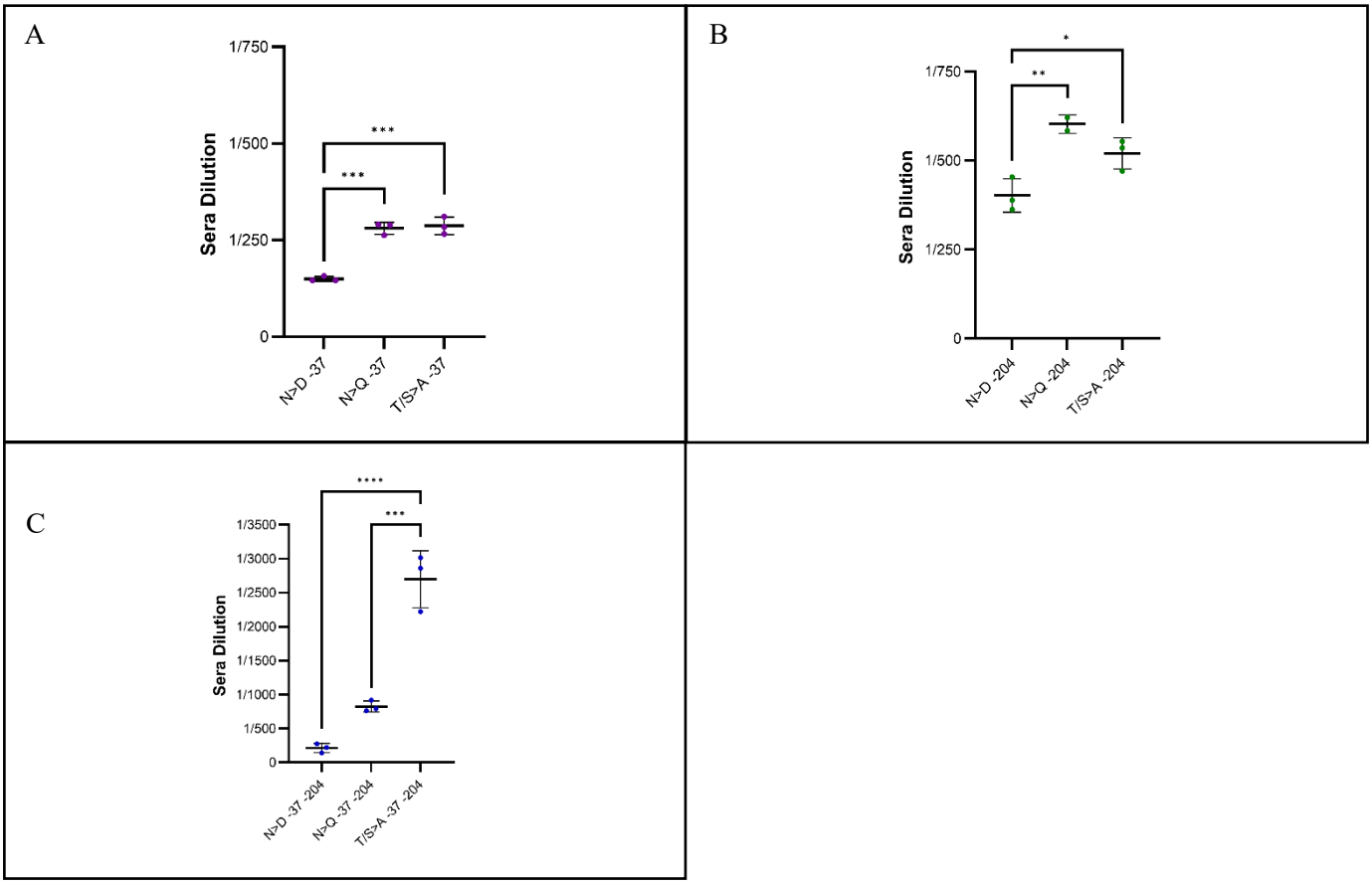


Figure 4.6: IC50 values generated from figure 4.2 neutralisation curves using NHP sera that had been produced through challenge with RABIPUR vaccination. All experiments were performed in BHK-21 cells comparing mutations between different substitutions (A) Asn-37 (B) Asn-204 and (C) Asn-37-204. Data is presented as mean±SD (assumed to be normally distributed) and statistically analysed using a one-way ANOVA followed by Tukey multiple comparisons post hoc test (*****P*<0.0001; ****P*<0.001; ***P*<0.01; **P*<0.05; ns = not significant) . Technical repeats in triplicate were used.

4.3.3.3 IC50 values between cell lines

As three different cell lines were used (BHK-21, HEK239T and MDCKs), the changes to IC50 values were also measured between them, as shown in figure 4.7. When comparing between the three cell lines it is clear that a general pattern does emerge which is consistent, with the exception of BHK-21s in N>D which is an outlier when compared to all three substitutions. The cell lines appear to be consistent in relation -Asn37, with no significance between them. However, in both -Asn204 and -Asn37 & Asn-204 there are clear and significant changes to the IC50 values found, in particular in the case of -Asn37 & -Asn204. However, the trends of which cell line produced higher or lower IC50s are not the same between cells lines when observing all substitution types.

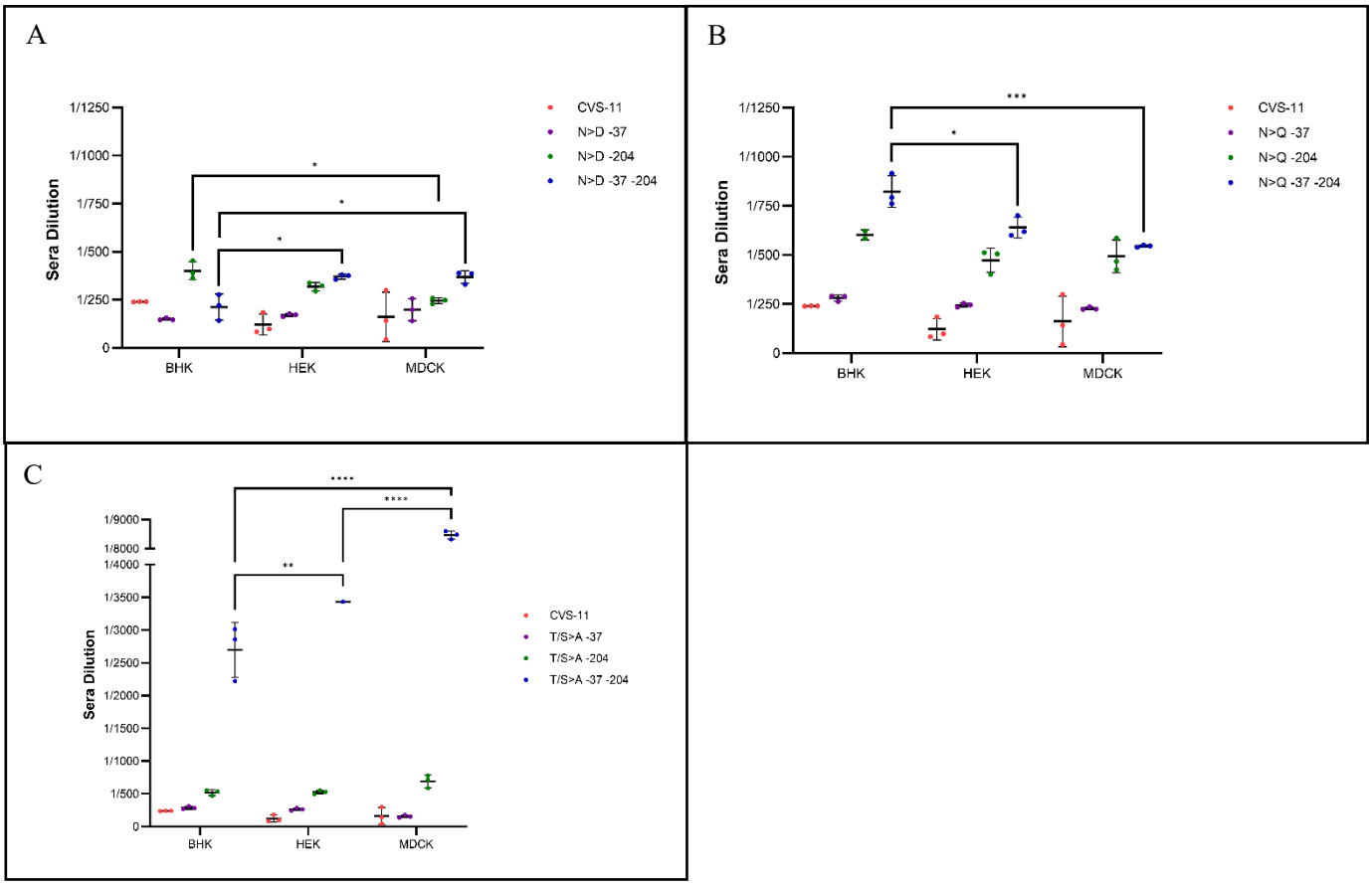
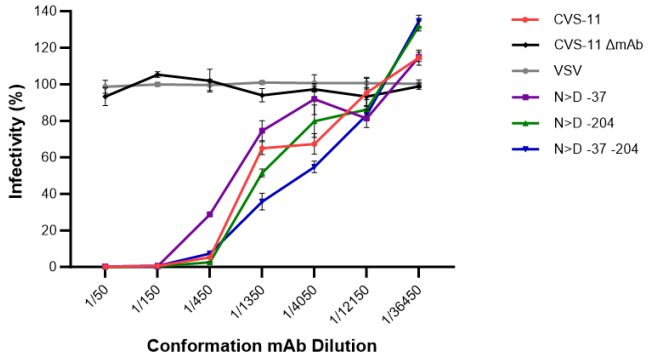


Figure 4.7: IC50 values generated from figure 4.2 neutralisation curves using NHP sera that had been produced through challenge with RABIPUR vaccination. Different substitutions were compared between different mutations (A) N>D (B) N>Q and (C) T/S>A as well as between different cell lines (BHK-21, HEK239T and MDCK cells). Data is presented as mean±SD (assumed to be normally distributed) and statistically analysed using a two-way ANOVA followed by Tukey multiple comparisons post hoc test. (****P<0.0001; ***P<0.001; **P<0.01; *P<0.05; ns = not significant). Technical repeats in triplicate were used.

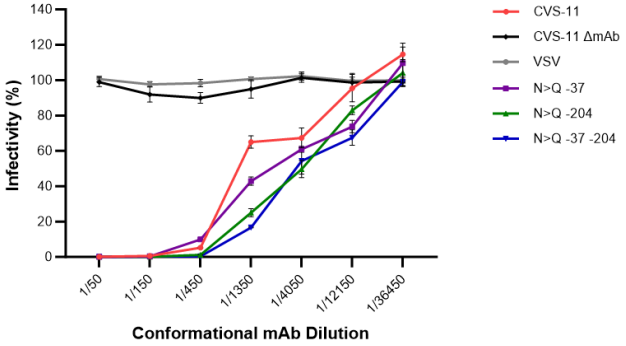
4.3.4 Neutralisation of sequon knockout panel using site specific monoclonal antibody

Following infection and neutralisation experiments, the antibody structure was tested to ensure that mutations that had been introduced had not altered the form of the glycoprotein. This was done using the conformational monoclonal antibody RVC20.

A



B



C

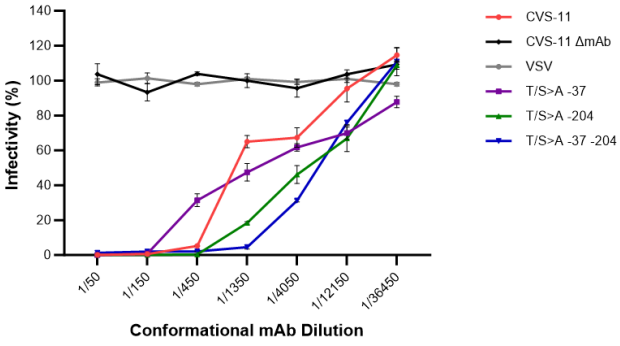


Figure 4.8: Neutralisation curves of three different sequon substitution knockouts, (A) N>D, (B) N>Q and (C) T/S>A. The assay was performed in BHK-21 cells. A three-fold dilution series was used to dilute the conformational monoclonal antibody RVC20. The results were normalised to a PBS control to represent 100% infectivity. Negative controls of VSV and a no-antibody CVS-11 infection were used and functioned as expected.

4.3.5 IC50 values of mAb neutralisation

The results of the neutralisation assays' IC50 values were examined and compared which can be seen in figure 4.9. The IC50 values based on the mAb neutralisation, similar to the sera neutralisation results, demonstrate significant differences between mutants and the original CVS-11 glycoprotein. The combined -Asn37 & -Asn204 is once again most impacted in its neutralisation profile, with all three significantly increasing in susceptibility when compared to CVS-11 and -Asn37. -Asn204 is also impacted greatly in all substitutions when compared to CVS-11 and -Asn37. Importantly, binding and neutralisation still occurred in all samples indicating that the structure remained relatively stable.

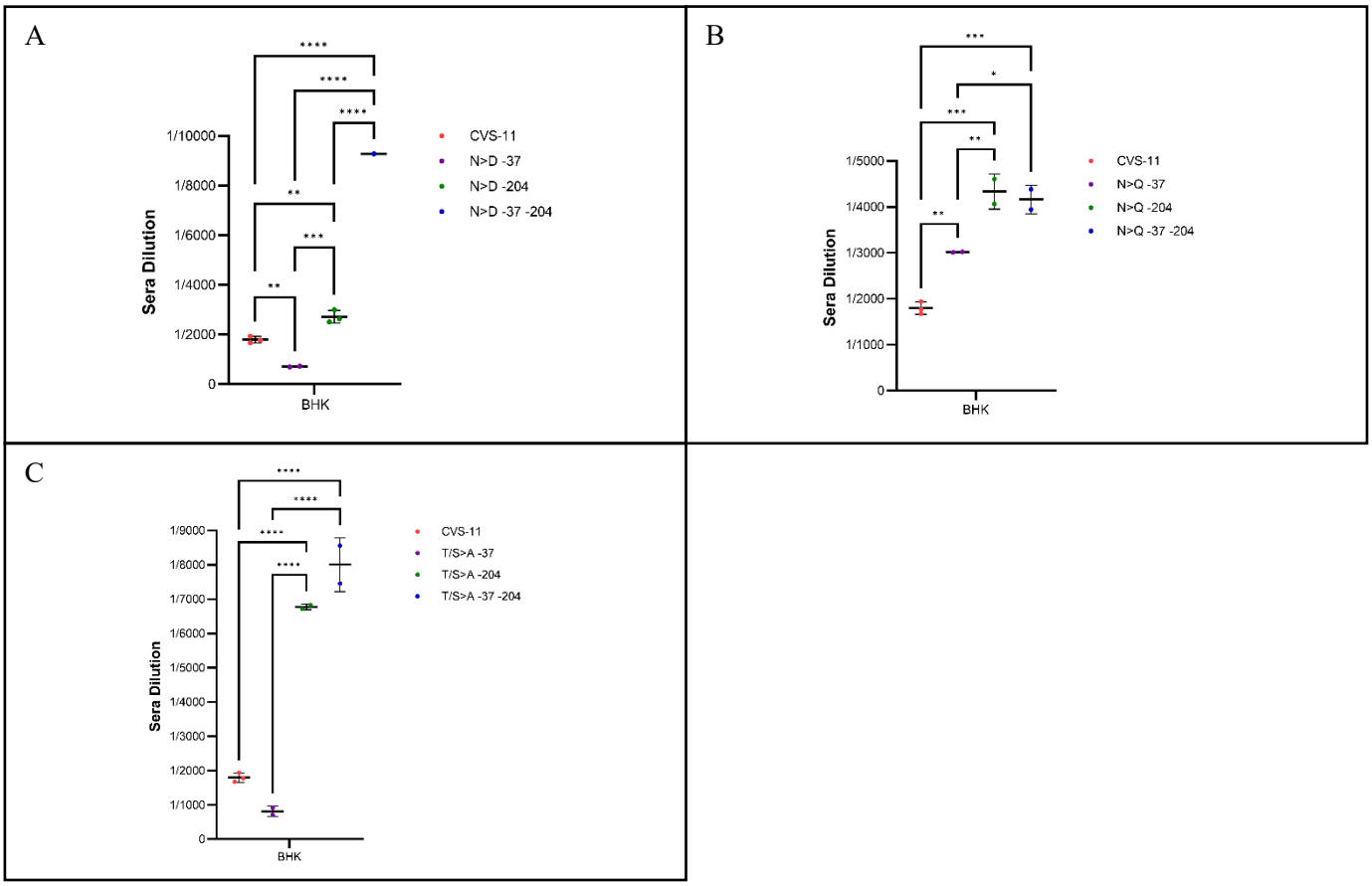


Figure 4.9: IC50 values generated from figure 4.8 neutralisation curves using the conformational monoclonal antibody, RVC20. All experiments were performed in BHK-21 cells using different substitutions (A) N>D (B) N>Q and (C) T/S>A. Data is presented as mean±SD (assumed to be normally distributed) and statistically analysed using a one-way ANOVA followed by Tukey multiple comparisons post hoc test (****P<0.0001; ***P<0.001; **P<0.01; *P<0.05; ns = not significant).

4.3.6 qPCR analysis of sequon mutant panel

qPCR was performed on cDNA of the mutant panel. This was generated using RNA retrieved from the supernatant of producer cells following the PV transfection process. The results of the qPCR are as shown below in figure 4.10. The qPCR data demonstrates a clear and consistent Cq value across the mutant panels. Both positive and negative controls were successful and all mutants were measured against both un-transfected cells as well as cells transfected with no envelope protein. VSV was also consistent when compared to the mutant panels and the positive control consisting of a luciferase plasmid produced an expectedly high Cq value.

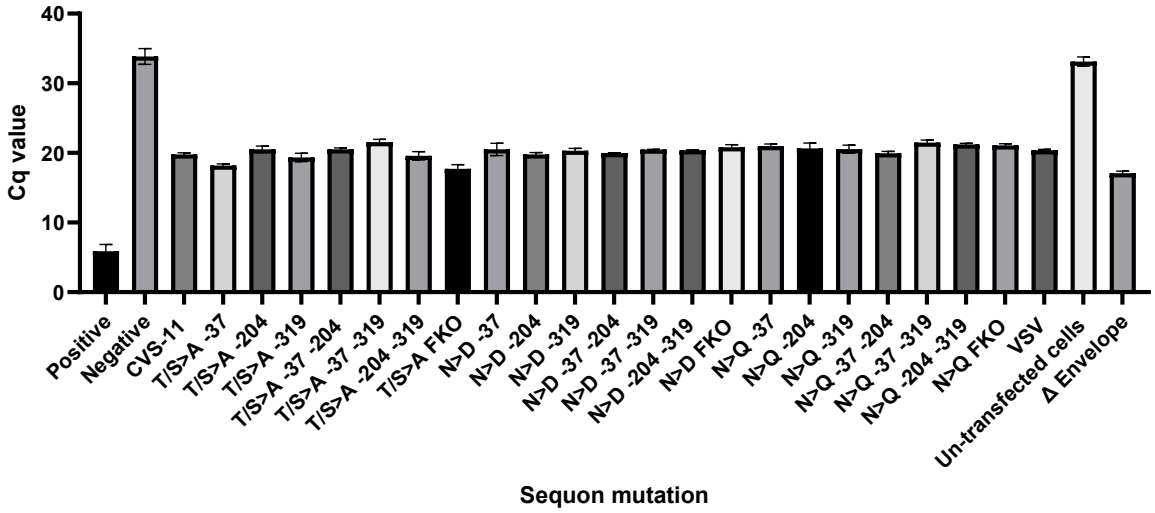


Figure 4.10: qPCR analysis on cDNA of the sequon mutant panel. The Cq values were calculated and data is presented as mean±SD (assumed to be normally distributed) and statistically analysed using a one-way ANOVA followed by Tukey multiple comparisons post hoc test. No statistically significant differences were found between the sequon mutants Cq value.

4.3.7 Infection assay with site swap mutants

In order to produce mutants which contained sequons from the divergent phylogroup III lyssaviruses WCBV, IKOV and LLBV, multiple rounds of SDM was used. This resulted in three CVS-11 plasmids which contained the phylogroup III sequons. After an infection test, these were found to be completely non-infectious. To investigate this further, the individual stages that had been used to produce the sequon site-swap plasmids were used in an infection assay to determine which impacted it. The results can be seen in figure 4.11.

The full panel of mutants showed again that CVS-11 mutants containing all of any of the phylogroup III sequons produced no infectivity whatsoever. Mutants containing LLBV Asn42, IKOV Asn334 and WCBV Asn334 produced little to no infectivity. While Asn334 mutants produced completely negative results, Asn42 did produce very low levels of infectivity. In contrast, LLBV Asn146, IKOV Asn160 and WCBV 202 produced infectious PVs. However, in the case of Asn146 and Asn160, there was a significant reduction in infectivity.

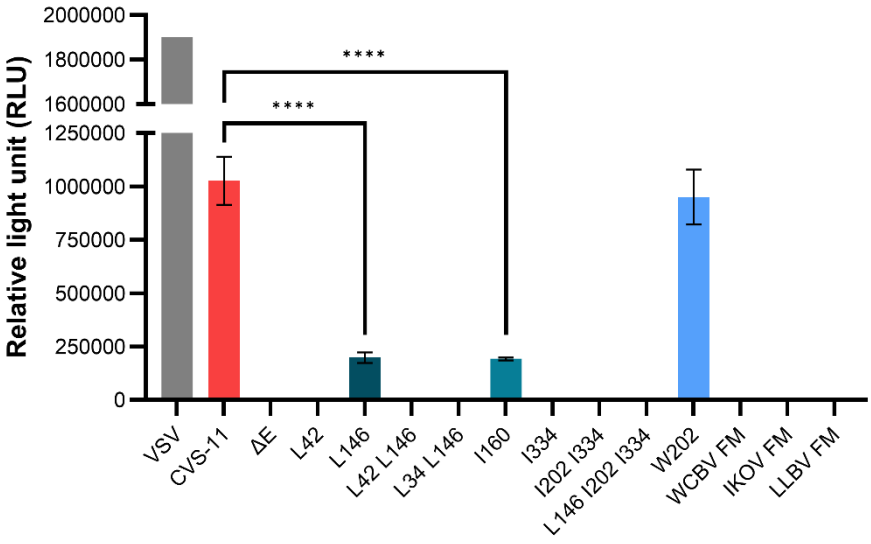


Figure 4.11: Infection assays using a full panel of site swap mutants. Assay was measured using relative light units and was both repeated and performed in triplicate. Infections were read at 2800 gain. FM = full mutant. Data is presented as mean±SD (assumed to be normally distributed) and statistically analysed using a one-way ANOVA followed by Tukey multiple comparisons post hoc test. (****P<0.0001).

4.3.8 Neutralisation curve of site swap mutants with NHP sera

A neutralisation experiment on the infectious PVs was performed using NHP sera from the medoid vaccine trial that had been produced after challenge with the RABIPUR vaccine. The results can be seen in figure 4.12.

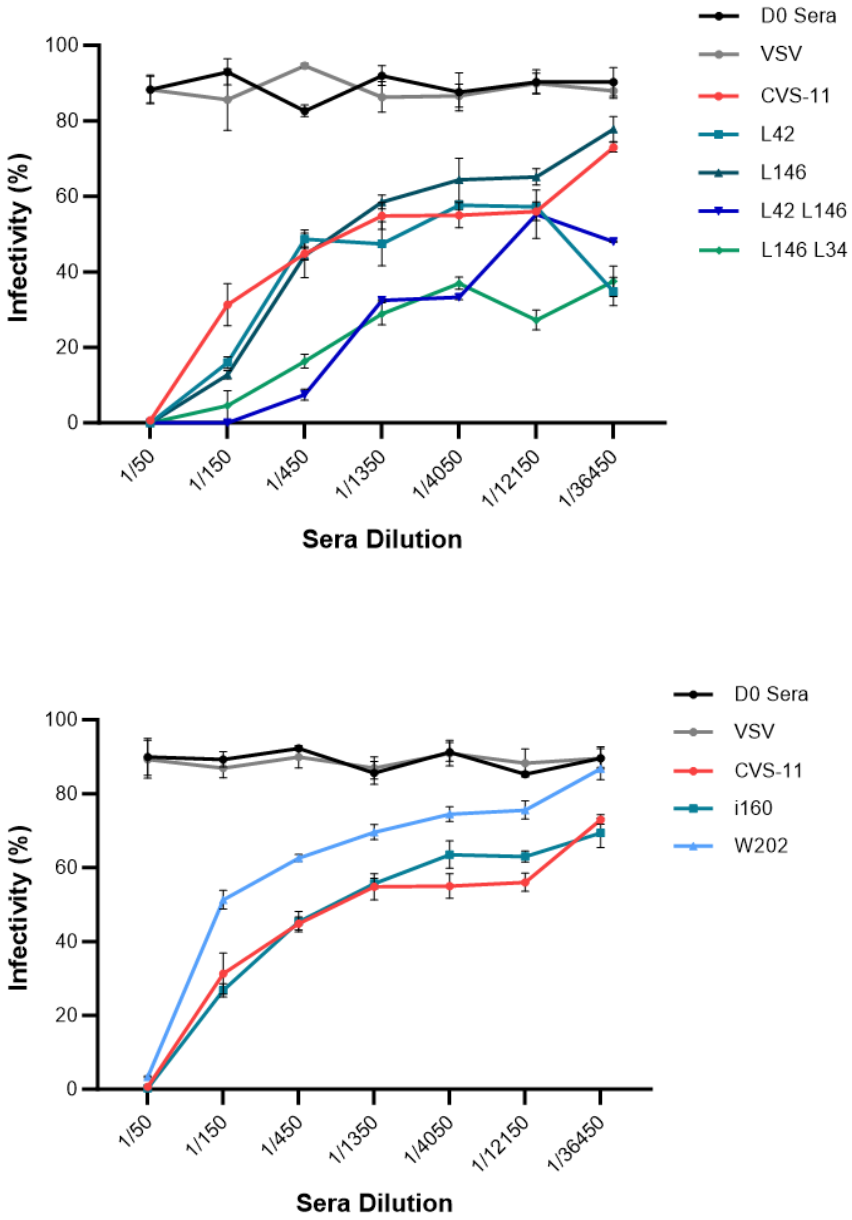


Figure 4.12: Neutralisation curves of six different site swap mutants. A three-fold dilution series was used to dilute NHP sera from the medoid vaccine trial. The results were normalised to a PBS control to represent 100% infectivity.

4.3.9 Neutralisation of site swap mutants with site specific monoclonal antibody

The conformational antibody RVC20 was used to determine if the structure had been impacted by mutational changes through a single point neutralisation assay. Results can be seen in figure 4.13.

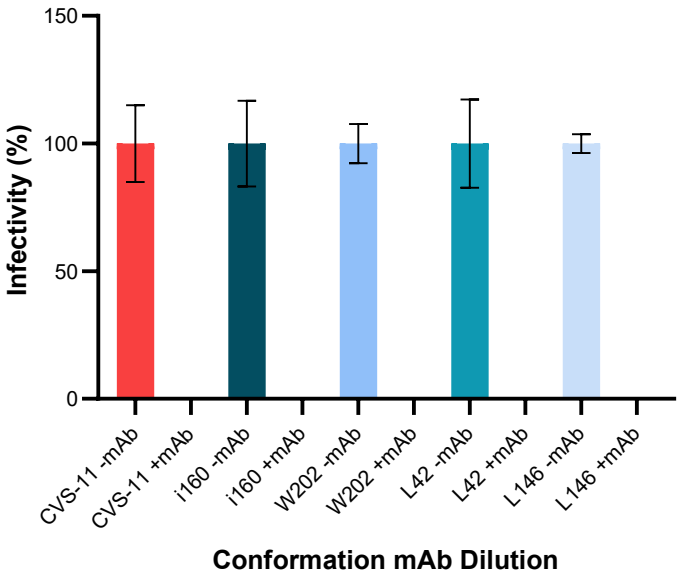


Figure 4.13: Neutralisation curves of four different site swap mutants in the presence and absence of the conformation antibody RVC20. The results were normalised to a PBS control to represent 100% infectivity.

4.3.10 IC50 values of site swap mutants with NHP sera

Comparisons of IC50 results from the neutralisation assay reveal that there was no significant difference between the neutralisation profile of CVS-11 and L42, L146, i160 and W202. However, there were significant differences between CVS-11 and the combined introduction of L42 & L146, and L34 & L146.

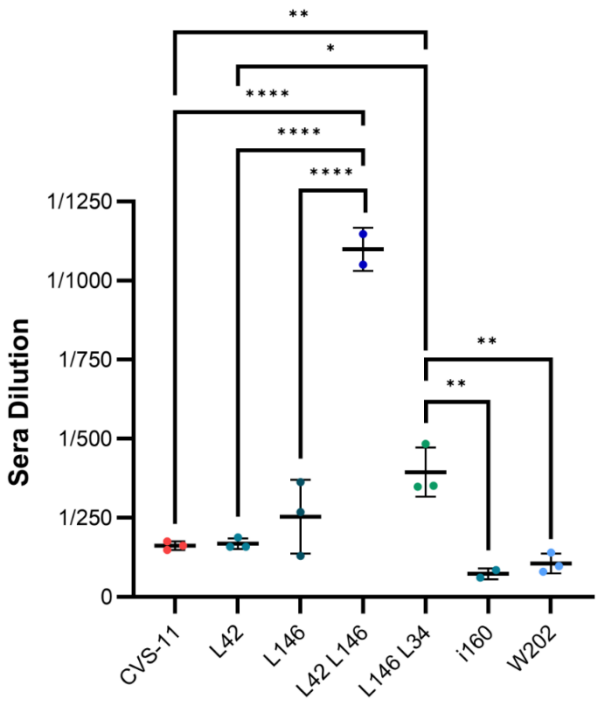


Figure 4.14: IC50 values generated from figure 4.12 neutralisation curves using NHP sera that had been produced through challenge with RABIPUR vaccination. Different site swap mutants were compared. Data is presented as mean±SD (assumed to be normally distributed) and statistically analysed using a one-way ANOVA followed by Tukey multiple comparisons post hoc test. (****P<0.0001; ***P<0.001; **P<0.01; *P<0.05; ns = not significant).

4.3.11 qPCR analysis of site swap mutant panel

The qPCR analysis of all available mutants revealed that there was no significant difference between the levels of PV in the supernatant which had been harvested from the HEK293T producer cells following transfection. All results indicated that transfection was successful and adequate levels of PVs were present in the supernatant for all plasmids used in the experiment. Positive and negative controls were implemented and behaved as expected. It is apparent that introduction of mutations did not impact production of PVs.

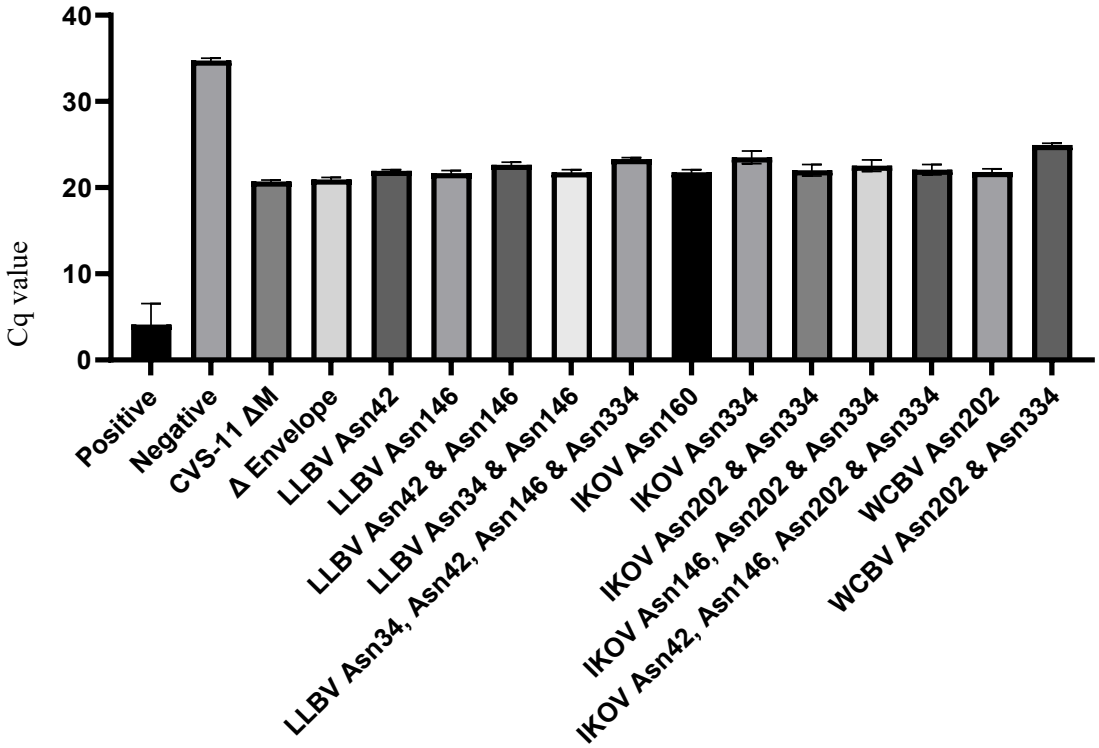


Figure 4.15: qPCR analysis on cDNA of the site swap mutant panel. The Cq values were calculated and data is presented as mean±SD (assumed to be normally distributed) and statistically analysed using a one-way ANOVA followed by Tukey multiple comparisons post hoc test. No statistically significant differences were found between the site swap mutants Cq value.

4.3.12 Visualisation of successful mutant glycoproteins

Using AI modelling provided by Robetta [Baek *et al*., 2021; Yang *et al*., 2020; Song *et al*., 2009], the sequences of mutated CVS-11 were input which produced predictive 3D models of the glycoprotein structure. Figures 4.16 to 4.22 show these outputs and their confidence ratings as reported by Robetta. As can be seen from the models, the input of the mutations appears to have made a minimal impact if any to the structure of the glycoprotein, particularly when observing the overall structure. All of the mutations shown represent an N>D mutation.

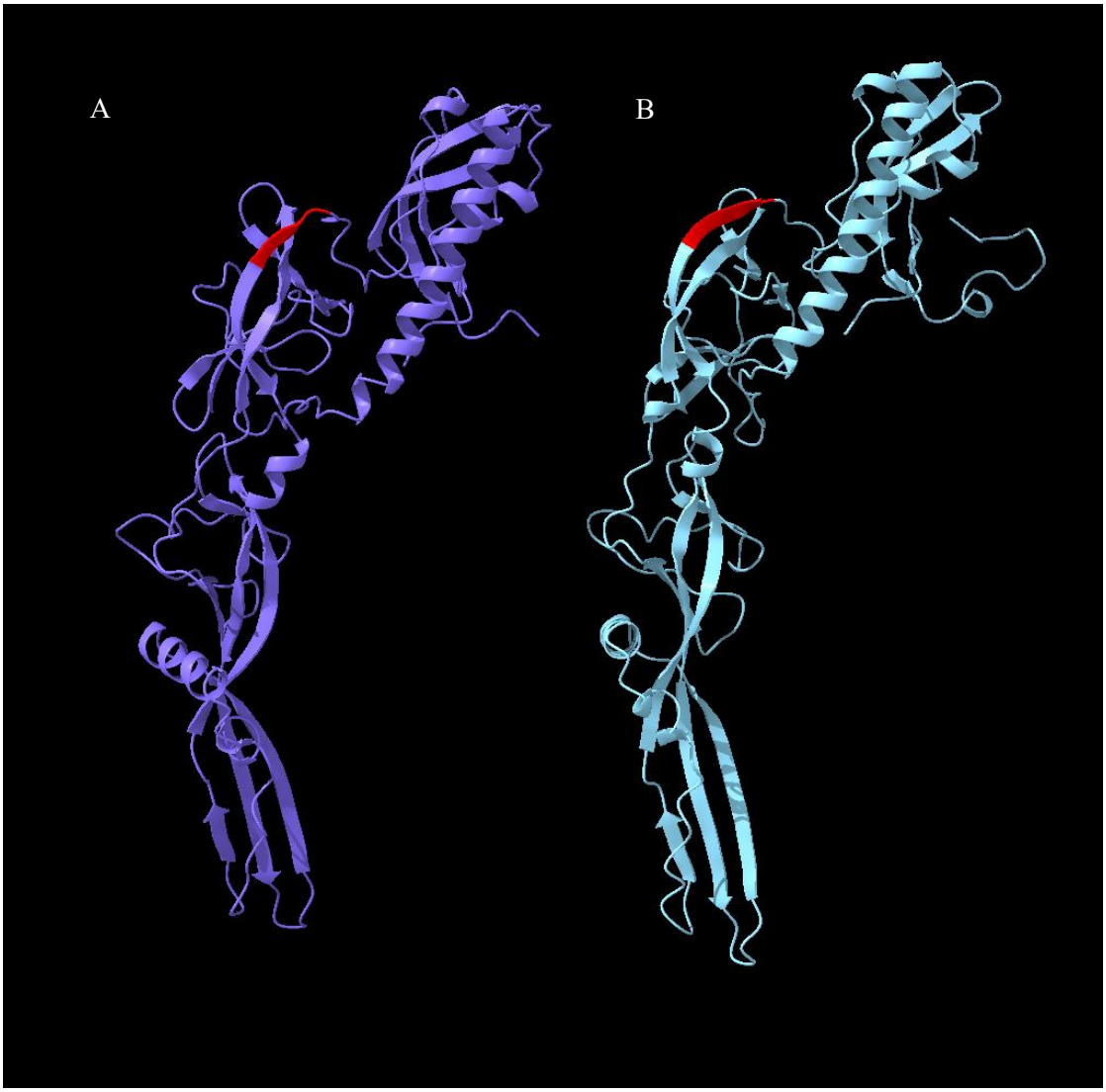


Figure 4.16 A structural comparison between (A) CVS-11 and the (B) -Asn37 mutant. The Asn37 sequon is highlighted in red. Structures were generated using the Robetta software. The confidence rating given by the software was (A) 79%, and (B) 80%.

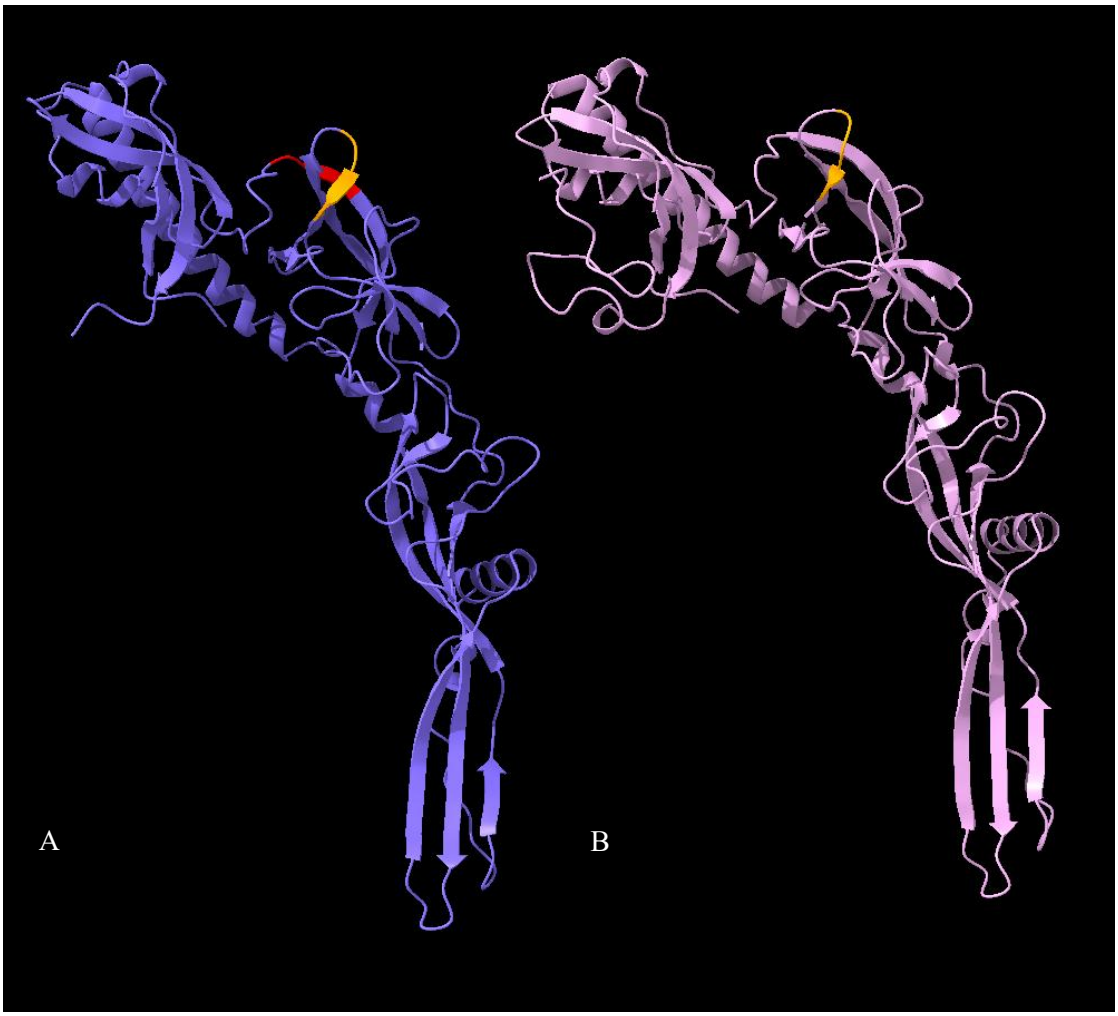


Figure 4.17 A structural comparison between (A) CVS-11 and the (B) -Asn204 mutant. The Asn204 sequon is highlighted in orange. Structures were generated using the Robetta software. The confidence rating given by the software was (A) 79%, and (B) 79%.

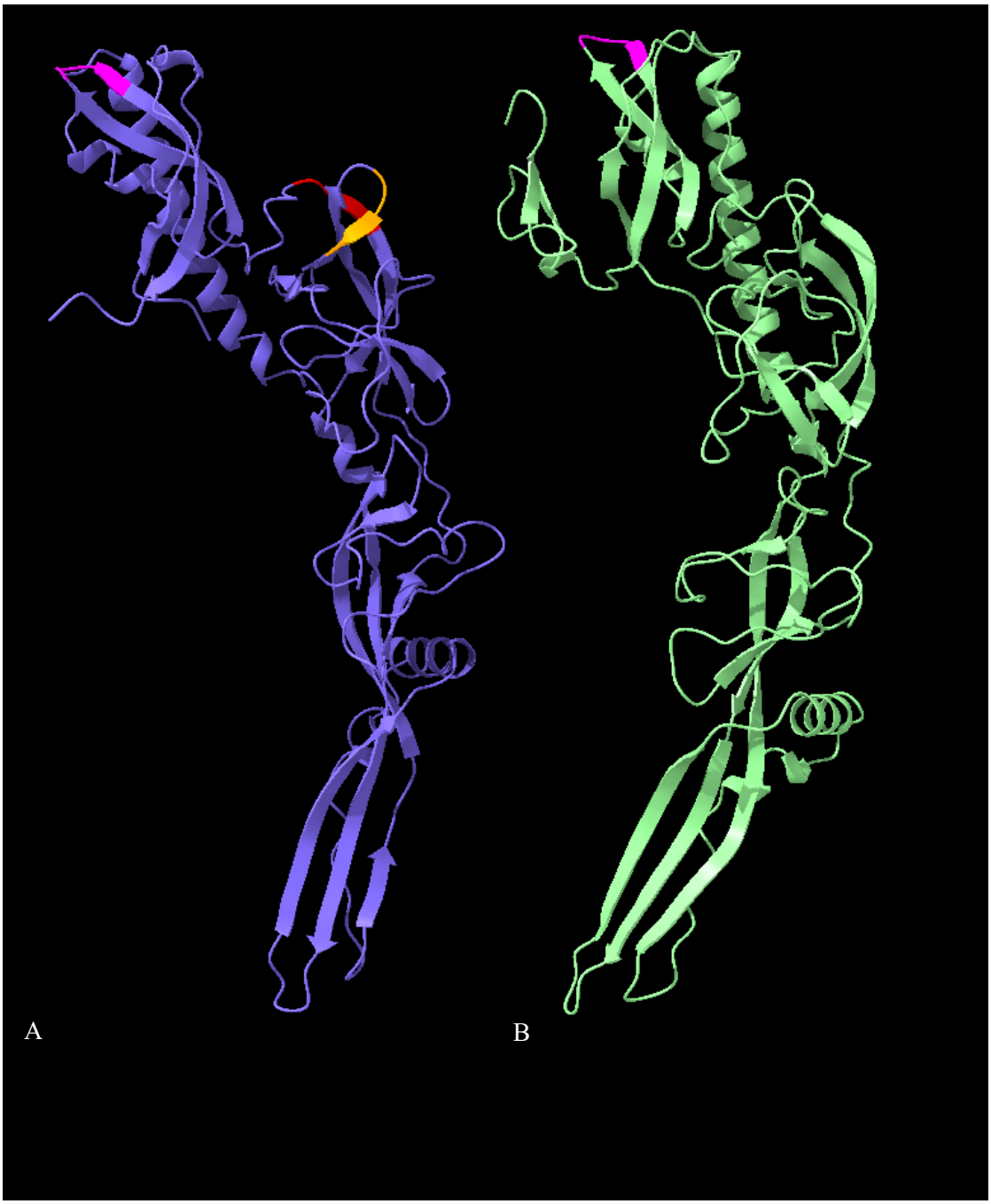


Figure 4.18 A structural comparison between (A) CVS-11 and the (B) -Asn319 mutant. The Asn319 sequon is highlighted in pink. Structures were generated using the Robetta software. The confidence rating given by the software was (A) 79%, and (B) 79%.

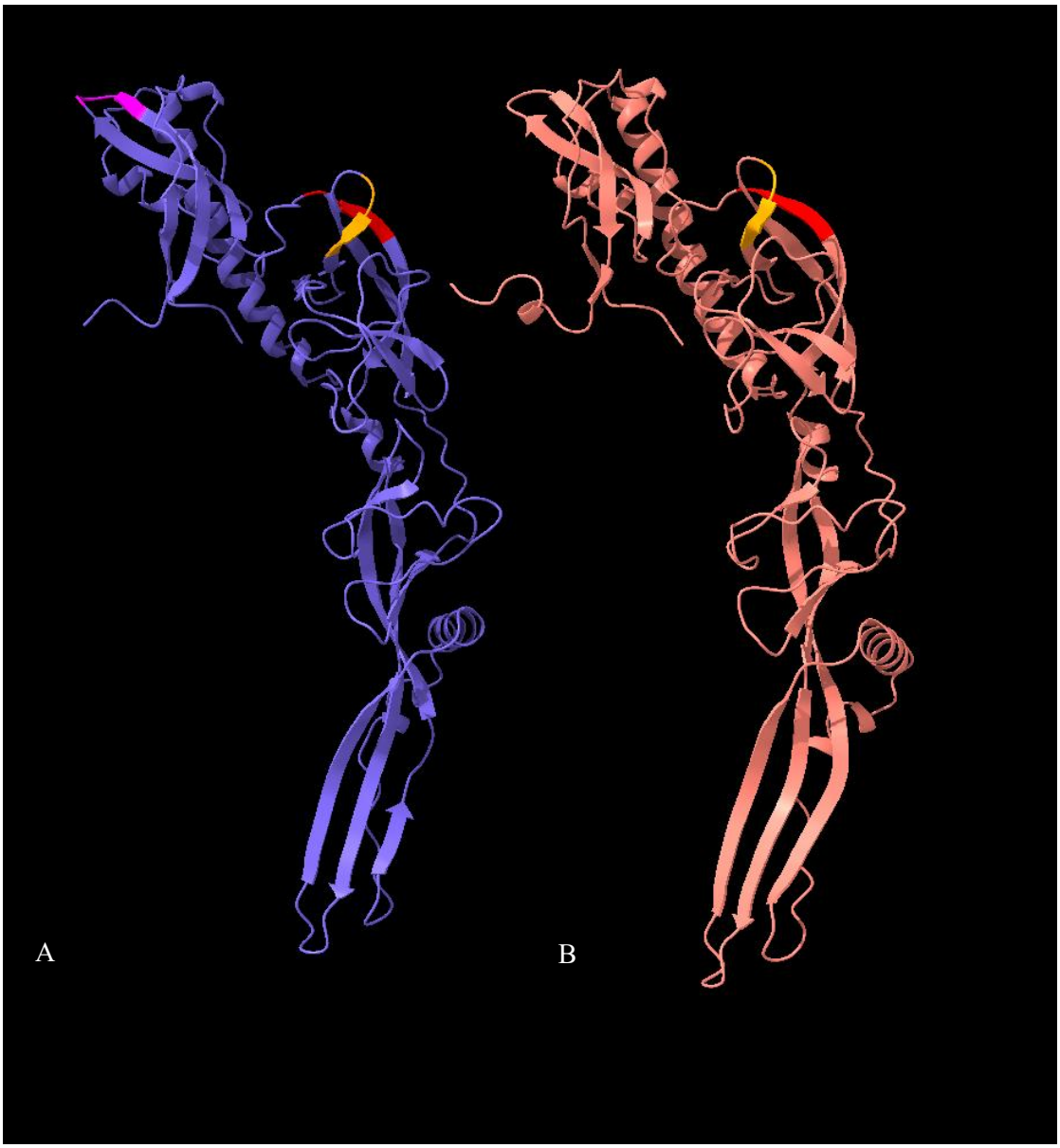


Figure 4.19 A structural comparison between (A) CVS-11 and the (B) -Asn37 & -Asn204 mutant. The Asn37 sequon is highlighted red, and the Asn204 sequon orange. Structures were generated using the Robetta software. The confidence rating given by the software was (A) 79%, and (B) 80%.

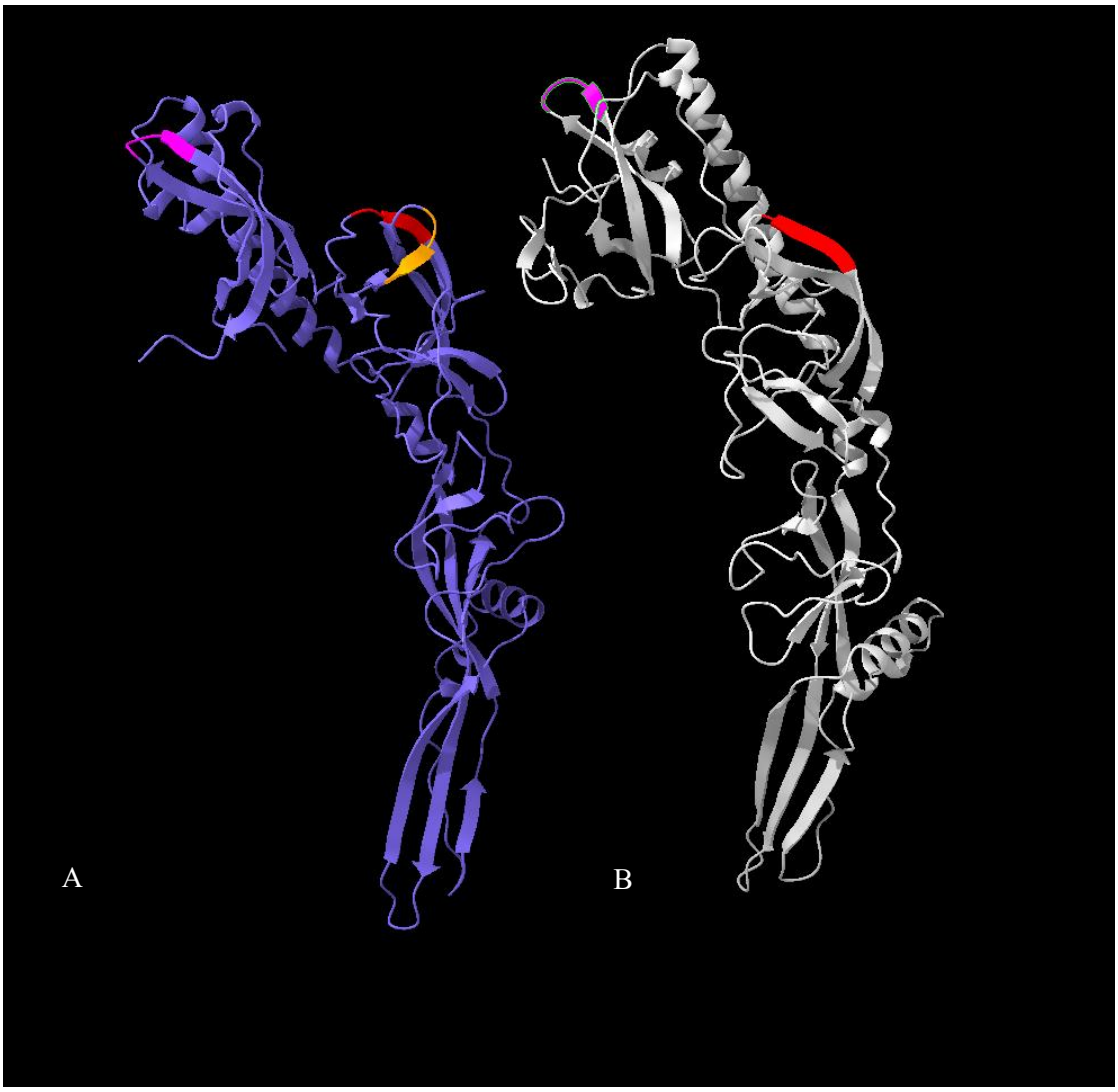


Figure 4.20 A structural comparison between (A) CVS-11 and the (B) -Asn37 & -Asn319 mutant. The Asn37 sequon is highlighted red, and the Asn319 sequon pink. Structures were generated using the Robetta software. The confidence rating given by the software was (A) 79%, and (B) 78%.

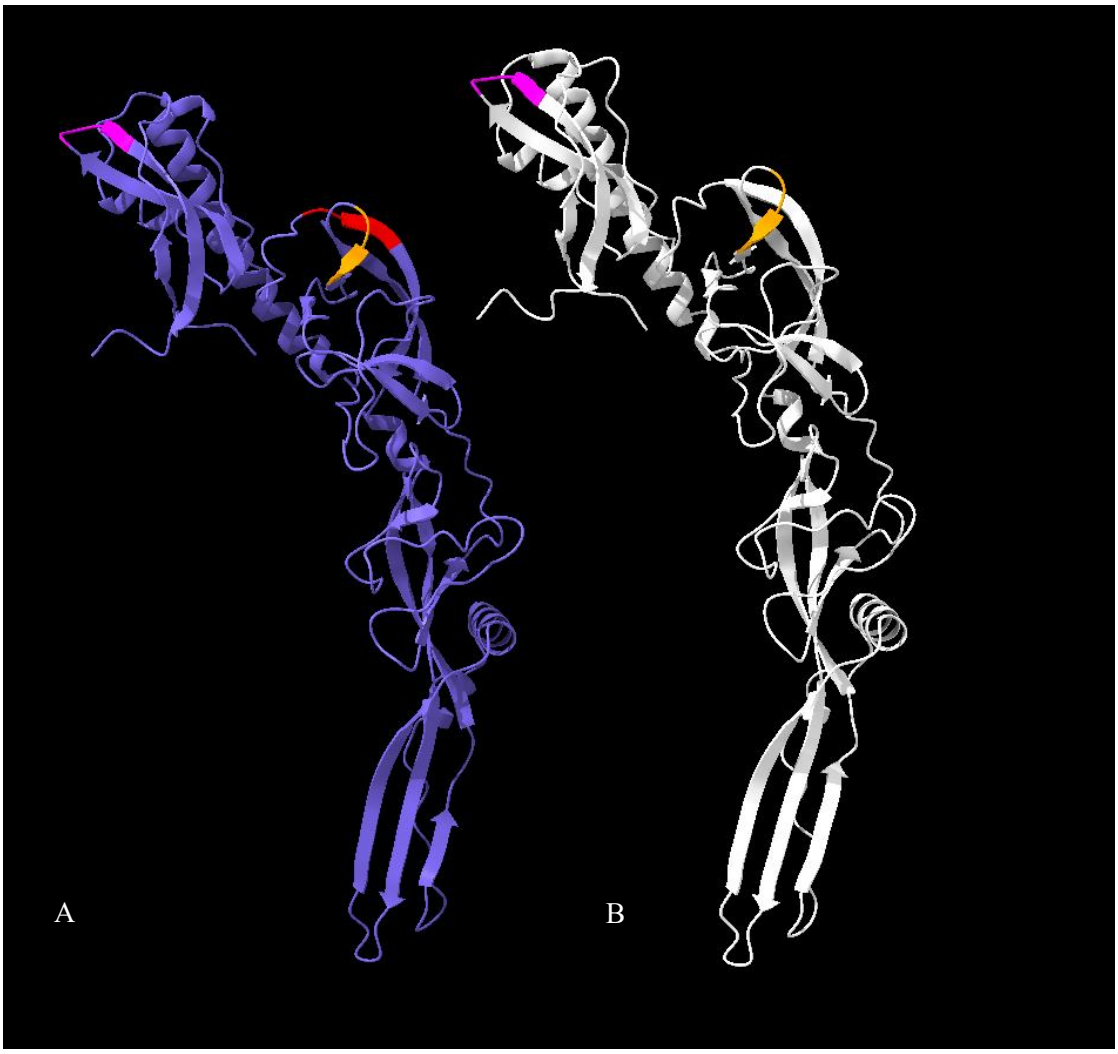


Figure 4.21 A structural comparison between (A) CVS-11 and the (B) -Asn204 & -Asn319 mutant. The Asn204 sequon is highlighted orange, and the Asn319 sequon pink. Structures were generated using the Robetta software. The confidence rating given by the software was (A) 79%, and (B) 80%.

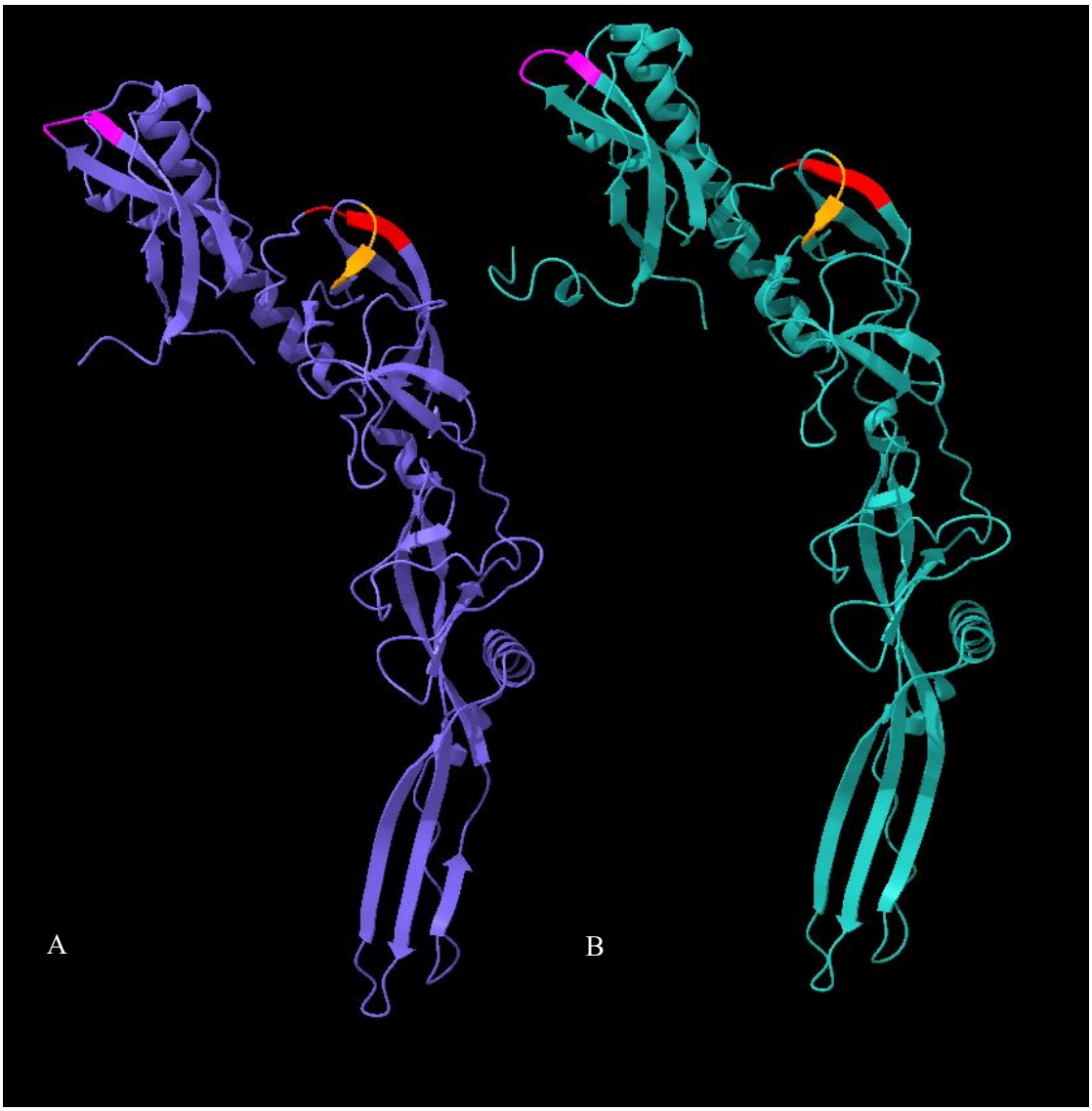


Figure 4.22 A structural comparison between (A) CVS-11 and the (B) full knockout mutant. The Asn37 is highlighted red, Asn204 sequon is highlighted orange, and the Asn319 sequon pink. Structures were generated using the Robetta software. The confidence rating given by the software was (A) 79%, and (B) 80%.

4.4 Discussion

Glycosylation has been proven to be important for the capability of viruses to infect cells, and their susceptibility to neutralisation. Research into lyssavirus glycosylation has so far focused on the impact on the production of virions and virulence. The further elucidation of the impact of these sites could lead to new targets for therapies. This study aimed to prove that, as with other viruses, lyssavirus infection and neutralisation was indeed impacted by the presence or absence of glycans.

4.4.1 Infectivity and neutralisation of RABV G protein is impacted by glycan presence

Infection assays with the glycan knockout mutants presented in PVs have shown that there are clear and significant changes in infectivity potential when certain sequons are removed. The removal of the cell culture adaptation Asn204 and, Asn204 in combination with the removal of Asn37, reduced infectivity and increased neutralisation susceptibility. This is potentially due to the proximity of the sites to antigenic regions which were perhaps obscured by glycans. These sites in RABV have been shown to be involved in increases in viral release when present [Nitschel *et al*., 2021] and that this release is the major bottleneck for adaptation to cell culture from WT strains. However, the QPCR data shows that there was not a significant difference between levels of PVs in the supernatant, indicating this is a difference in infectivity based on the change to the sequon itself. All of the substitutions which were used produced similar patterns of changes between mutants. Removal of Asn319 removed all potential of infectivity.

The removal of the sequon at Asn37 did not reduce the infectivity of the PVs but instead increased it. It also had no impact on the neutralisation susceptibility of the PVs. This is interesting and is possibly due to the inefficiency of the glycosylation at the sequon. If the glycan was beneficial in immune avoidance in a significant way it would likely be more efficiently present, however it appears that this efficiency is potentially sacrificed to improve viral propagation [Li *et* al, 2021] and that the trade-off for RABV is a modest loss of infectivity. It should also be noted that Asn37 is by no means an essential glycosylation site and is absent in nearly 20% of RABV sequences as well as not being present in any other member of the genus with the exception of BBLV.

It is well known that Asn204 is a cell culture adaptation, similar to Asn247, and that an additional glycan is not only a common occurrence but also quick to appear in cultured strains of lyssavirus [Nitschel *et al*., 2021]. Indeed, it has also been strongly argued by Nitschel *et al*., 2021, that their presence greatly impacts viral release, which is based on the number of NLG sites on the glycoprotein ectodomain, forming a sort of bottleneck. This would indicate that as

NLG sites are added, viral release is reduced and therefore infectivity. However, in a controlled environment using the PV system, infectivity went down when the additional Asn204 site was removed entirely. This indicates that the though the number of NLGs present negatively impacts viral release, the adaptations themselves increase infectivity. Asn204 presence also reduces neutralisation susceptibility, indicating another benefit for the glycoprotein from this adaptational change. However, it is likely that there are other factors which are reduced such as virulence when introducing these additional sequons due to the lack of their presence in lyssaviruses outside of lab adaptations, indicating a trade-off to their introduction in fixed strains, though this would need to be investigated further to confirm. Indeed, the presence of other lab adaptations Asn194 and Asn247 have been shown to reduce pathogenicity in peripheral infection in mice [Yamada *et al*, 2012].

Results of the infection and neutralisation of both the knockout panel and the site swap panel often showed results which increased to a greater degree when two sites were combined, regardless of the response the two sites generated before. This is an interesting phenomenon and may point to interactions between sites or, in the case of neutralisation, show that a single glycan is enough to prevent or generate neutralisation susceptibility, but when combined with the loss of another site creates a more drastic outcome. The introduction of L42 or L146 generated no significant difference individually, but their introduction together produced a strong shift in the IC50 of the mutant that meant a much greater sensitivity to neutralisation than CVS-11. This is a curious outcome and one that was preserved over biological repeats. The mechanism behind this is unknown. Interestingly, the conformational antibody continued to bind correctly indicating that the glycoprotein is not significantly malformed, though it has been demonstrated that the antibody used is somewhat resistant to small single nucleotide changes [Hellert *et al*., 2020]. In addition to this, combined introduction of L42 & L146, and L34 & L146 also reduced infectivity to a greater degree than either sequon individually. These findings parallel the knockout mutants. The combination of -Asn37 & -Asn204 almost always produced either greater susceptibility to neutralisation or a stronger reduction in infectivity when compared to their individual knockouts, in fact -Asn37 did not impact infectivity reduction at all yet in combination with -Asn204 appears to contribute to this. These findings could suggest glycan or sequon interactions which are not yet understood.

4.4.2 The substitution used to knock out a sequon impacts results to a significant degree

As discussed in the previous chapter, the amino acid substitutions used in SDM glycan knockout studies vary to a great degree. This is concerning as it introduces a potential variable in glycan knockout research that could impact the significance of certain results. The findings of this

chapter demonstrate that different substitutions used to knock out the same sequon do indeed impact the results of infection and neutralisation assays, generating results with variation. Infectivity assays using different substitutions produced similar patterns, though significant changes did occur. In addition, the IC50 values generated from neutralisation assays also differed significantly in a number of ways in every mutation. This phenomenon occurred both between positions one and three of the sequon but also within the first position with both N>D and N>Q having significant differences between each other. Though this study has demonstrated this phenomenon, it is important to continue to expand and confirm these findings using different sequons within the lyssa genus, in addition to other viruses altogether, especially as glycobiology in virology is an important and expanding field.

4.4.3 The site swap experiment and potential reasons for lack of infection

Though meaningful data was established using the site swap method, the initial aim of the experiment was not met fully. The production of full site swap mutants resulted in a complete reduction in infectivity, though the QPCR data demonstrates that there was successful production of PVs which had been released from the producer cells into the supernatant. This poses a quandary of what prevented infection. As can be seen from the infectivity data of the mutants which had been produced leading up to the full site swaps, it appears that there is a correlation between the addition of Asn334, present in all complete site swaps, which led to greatly reduced infectivity. Though there are other sites which did show no infectivity, because of the lack of a full panel that cannot be ruled on as conclusively as Asn334 which was present in all full site swaps and had individual presence also.

While individual mutants like L146 and I160 did continue to produce infective particles, their inclusion did negatively impact infectivity greatly and perhaps this is additive, with additional sequons reducing infectivity further. This would potentially make generating these kinds of CVS-11 PG II/III mutants impossible due to the number of sequons which need to be added. This is contrasted however with the WCBV mutant which, as evidenced by the addition of Asn202, retained infectivity until Asn334 was added. This points towards the modification of Asn334 as a potential limiting factor. The amino acid at position 333 has been examined extensively in lyssavirus research and appears to be an important residue with multiple functions. The recent study by [Itakura *et al*., 2022] found that the residue was a major determinant of RABV pathogenicity with the particular amino acid present changing virulence, for example an Arg333 was present in virulent strains compared to a Glu333 which was present in attenuated strains. Interestingly, the different residues also impacted the tropism of the glycoprotein, with Arg333 decreasing astrocyte tropism. The attenuated strain containing Glu333 invoked higher IFN responses in astrocytes compared to the Arg333 version. Other

studies such as that by [Tao *et al*., 2010] have identified residue 333 as important in viral-induced apoptosis in infected host cells. Further work by [Tuffereau *et al*., 1998] has shown that the residue influences the binding affinity of the glycoprotein to p75NTR indicating that changes to this residue could change infectivity or indeed tropism. While changes to this residue were not performed in this study, the introduction of a glycan at Asn334 could potentially block this important residue, or the change in amino acids surrounding the site may have caused misfolding. Though the QPCR detected PVs in the supernatant which contained the Asn334 sequon, it does not rule out a non-functional glycoprotein as evidenced by the no-envelope negative control.

The essential nature of residue 333 in RABV may extend to more divergent lyssaviruses, and its proximity to glycosylation is interesting. In combination with the findings of Li *et al*., 2021, which suggest increased numbers of glycans reduces virulence, it could be further speculated that both the increase in glycan numbers and their positions at sites such as Asn334, significantly reduces the virulence of these divergent lyssaviruses, though this would need to be examined further through more site swap and knockout experiments. Indeed, it has been reported that the R333 residue, which is essential for virulence, is replaced by D333 in phylogroup II viruses and this change is a likely cause of their reduced pathogenicity [Badrane *et al*., 2001]. These findings could indicate that divergent lyssaviruses have potentially produced more sequons through selection, perhaps selective immune related pressures, which attenuate their virulence both due to their location and because of the increase in total glycans on the glycoprotein. This could go some way to explaining their lack of host switching and spillover in comparison to RABV.

While the introduction of Asn334 does appear to have impacted infectivity, it was not the sole change that removed infectivity altogether. L42 and L34 both reduced infectivity to levels equivalent to the negative control. Both of these sites are located near the CVS-11 Asn37 which could explain some changes in infectivity, however it has been shown in the knockout study that this sequon is not required for infectivity levels to still be high. It is therefore unlikely that this is the cause. Instead, it is possible and likely that these changes caused misfolding of the protein. This could be further verified in future work using mass spectrometry or western blotting.

As previously described, the IC50 results of neutralisation using the mutants which did produce infection indicate that the combination of sites such as L42 & L146 or L34 & L146 increase susceptibility to neutralisation and reduce infectivity, which is an interesting outcome of the experiment.

Overall the conclusions of the site-swap mutation experiments are those of foundations. The potential interaction between particular sites such as 333 and glycans are compelling and with

more investigation may answer some questions as to the reason behind the 100% conserved sequons in more divergent lyssaviruses. Indeed, the site swaps also demonstrate that there is a possibility of strong inter-glycan or inter-sequon interactions which needs further investigation if the mechanisms that lyssavirus G protein glycosylations are involved in are to be fully elucidated.

4.4.4 Limitations of the studies

The study investigated glycan sites and used their removal to measure changes. One of the limitations of this type of experimentation is the difficulty in accounting for sequon interactions. While conclusions can be drawn from single site knockouts, the impact of site removal on other sites which remain can be difficult to discern, especially when dealing with more than two sequons. This appears to be demonstrated in the combined removal of Asn204 and Asn37, which decreases further below either when they are knocked out individually, in fact Asn37 removal does not impact infectivity negatively whatsoever. There has been investigation into sequon interactions before, such as in the study by Wojczyk *et al*., 2005, which found that knockouts of particular sequons in the RABV glycoprotein changed which types of glycans were attached to others and to what degree. Their findings also involved the passage generated sequon Asn247 which was found to influence Asn319 through enabling additional N-glycan processing, however ultimately concluded that the mechanisms remained to be understood and work in lyssaviruses has not continued down this avenue.

4.4.4.1 Western blotting and validation

Western blot analysis was performed using lysate from the producer cells but unfortunately yielded no results including the failure of positive controls. After repeated attempts to rectify the issue through replacement of the antibody, modification of method times such as the transfer stage, and the replacement of reagents, the experiment produced no results including that of the positive controls β-actin and anti-MLV p30 mAb which also failed. Western blot analysis would have shown the differences in protein size changing with the removal of the glycan, indicating the presence of them in CVS-11 and their successful removal in the mutants. Several other methods were used to attempt to demonstrate the presence of glycans on the PVs such as a lectin binding ELISA and vacuum/dot blotting which also unfortunately were not successful. Antibody confirmation was performed using a neutralisation assay which did confirm that the antibodies were indeed functional.

It is clear in the literature that the lack of glycosylation at Asn319 would have produced non-functional glycoprotein and when the site was knocked out this was the case. It can therefore be assumed that a glycan was present here and studies have demonstrated its efficient glycosylation [Shakin-Eshleman *et al* 1992b]. Asn247 has also been demonstrated to be efficiently

glycosylated previously [Shakin-Eshleman *et al* 1992b], therefore the experiments involving this site are also confidently stated. These sites are highly likely based on the previous findings of the literature to have glycans attached, and therefore the impact that their removal generates can be attributed to the loss of a glycan rather than the change in amino acid with reasonable confidence despite the lack of western blot analysis.

However, this lack does provide a shortcoming to the results involving Asn37 which is inefficiently glycosylated [Shakin-Eshleman *et al*., 1996]. Though it is still likely that glycans were attached and the differences seen between the mutants were due to the knocking out of the sequon. Indeed, as seen in chapter 3, glycan addition prediction indicates that the sequons examined in this study would have glycan attachment. This limitation of a lack of western blotting would be ideal future work to further confirm the outcomes of this study.

4.4.4.2 Other difficulties with validation

While western blot analysis was proving difficult to achieve results, several attempts were made to circumvent the issues using other methods of validation. A lectin binding ELISA was performed using the lectins GNA and WGA the latter of which has been shown to bind to RABV glycoproteins previously [Delagneau *et al*., 1981]. GNA was chosen due to its mannose targeting nature though it has been shown to not neutralise RABV previously [Wang *et al*., 2023]. These lectins bind to glycans and could therefore be used to demonstrate how the removal of glycans had occurred in the knockout mutants, though this also unfortunately was unsuccessful in producing results. Recent work by Wang *et al*., 2023, has demonstrated a number of lectins which have neutralising capabilities which may be interesting to use in future binding experiments on these samples.

A follow up experiment using a vacuum based dot blot assay was also performed. The benefit to this is that a conformational antibody can be used, which had been obtained and was tested but also produced negative results after multiple attempts. The antibody had been used in neutralisation assays successfully and so could be ruled as functional, and the secondary antibody had been replaced due to concerns it was non-functional, but this provided no change to the negative results. This assay had been performed as part of another thesis in the same laboratory previously, which also produced no results and was similarly abandoned.

4.4.4.3 Additional interpretations

It is important to clarify that, due to a lack of western blots or mass spectrometry confirming the presence and removal of glycans from the mutants, other explanations of these results should be acknowledged and explored. One such explanation could be the alteration of glycoprotein structure into a less efficient shape, without the successful removal of the glycan. It has been

demonstrated by Shakin-Eshleman *et al*., 1992 & 1996 and Nita-Lazar *et al*., 2005, that N-linked glycosylation is not possible without the sequon present. As the sequencing of the mutants confirmed the deletion of the sequons, the presence of glycans at those sites should not be possible. However, the lack of glycan presence to begin with should also be evaluated. While efficiency of glycosylation and the percentage likelihood has been investigated in chapter 3, it is still possible that glycosylation never occurred in both the mutants and non-mutants, particularly at variable sites such as Asn37. If this was the case, it is possible the changes seen in this chapter are due to protein structural changes, or changes in charge, both based on the artificially introduced amino acids. While experiments were performed using confirmational antibody binding, therefore demonstrating that structure was unaltered to a significant degree, this does not concretely confirm that no alteration took place. The confirmational site may have been intact, but the 3D structure of a protein is complex and the interactions between amino acids important. Therefore subtle changes such as those introduced in the mutants could have impacted performance in the downstream analysis.

There is evidence, however, that this concern may not be warranted in the removal of N-linked glycosylation sites. A study by Lee *et al*., 2015, demonstrated through PDB structure analysis, that NLG did not induce significant conformational changes in folded protein structures. Indeed, they found that in most cases the global structures of proteins were unaffected when compared to their deglycosylated counterparts, and that the structures were almost identical. However, these were based on 3D models from crystal structures, 100% sequence matched using BLAST to sequences without glycans, rather than intentionally removed using amino acid swaps such as in this study. The study also found that smaller local changes can occur, and that changes introduced did not even always appear to be close to the glycan site.

There are also other studies which posit that the removal of glycans can impact the flexibility of proteins and the amino acids surrounding the sequon. These changes in shape or flexibility may have opened up new avenues of neutralisation for the sera which would not be present naturally [Gerken *et al*., 1989]. This could explain the results even if the removal of the glycan had been successful and demonstrates that further work needs to be done to confirm this study's findings.

4.4.5 Future work

Though the panel of mutants was comprehensive for the strain of CVS-11 obtained by this study, it is not comprehensive when discussing the whole gamut of sequons in RABV and lyssaviruses as a whole. Future work in this area could include the testing of more knockout mutants of other sequons and their impact on infection and neutralisation, something which has so far not been investigated. There is also the potential for a repeat of this study to be done in more divergent

lyssaviruses, examining their phenotypic responses when sequons are knocked out. It is likely that, based on the findings of this study, the introduction of additional sites would reduce infectivity. This combined with the findings of Li *et al*., 2021, that additional glycans reduce virulence, should result in there being significant impact on infectivity and neutralisation when these glycans are removed.

While the validation of this study is considered to be satisfactory, the use of western blots to generate direct images of glycan removal would be a useful next step in the confirmation of these results. While it has been shown previously that these sites are glycosylated in multiple sources [Shakin-Eshleman *et al*., 1992 & 1996; Yamada *et al*., 2013], and predictive software NetNGlyc adds credence to this, western blot analysis is a logical next step and was only not achieved due to difficulties with the process. Western blot analysis could also lead on to experiments with Endo H which are commonly used in glycan research to assess the presence and number of glycans on a protein [Freeze & Kranz, 2010].

The site swap mutations were ultimately unsuccessful in producing the entire genotypic profile of sequons found in divergent lyssaviruses into CVS-11. The mechanisms behind this failure are unknown but there was not enough time to complete a full panel of each stage of mutation. For example, many of the mutations were only found in combination with other knockouts including L34, I202, I146, and W334. Having a full panel would help better identify some of the causes behind the failure of full sequon swaps, as well as provide interesting insight into each individual sites impact. Further to this, more combinations of mutants could also be added to a full panel. In contrast to this work, the CVS-11 sequons could be swapped into the PG II and PG III lyssaviruses in a reversal of this experiment which may perhaps be more stable with less additions. These CVS-11 mutants could then be examined with the goal of identifying if their introduction generates new susceptibilities or weaknesses to PGI targeted sera and PG II/III targeted sera.

Conclusion

In conclusion, it is clear from this study that glycans can indeed impact the infectivity and neutralisation profiles of the rabies glycoprotein similar to other viruses. The results also indicate that the type of substitution chosen can impact the results of these knockout experiments significantly, a worrying outcome when combined with the results of the systematic review which revealed there was no real standard for knockout studies or even which position, one or three, should be altered.

The unfortunate failure of the site swap study to produce CVS-11 with the sequons of PGII or III lyssaviruses did produce interesting results which form the basis for much further research to be performed. The results showed that introduction of L146 and I160 reduced infectivity in

the CVS-11 glycoprotein, and that certain introductions could remove infectivity altogether, such as asn334. The findings of this chapter contribute to the overall understanding of glycans in lyssaviruses, but much more needs to be done to understand important questions such as why the PG II and III lyssaviruses have developed more sequons and determine if this links to immune avoidance as a potential target for future therapies.

Chapter 5: Infection of natural killer cells with lyssavirus pseudo-virus particles

5.1 Introduction

With the fatality rate in humans of untreated rabies reaching almost 100%, the host immune response in absence of vaccination is clearly flawed. Numerous studies have found that within the majority of human cases there is no detectable antibody response until clinical signs and symptoms appear, even then some cases have a significantly delayed response [Johnson *et al*., 2010]. Though antibody titre does increase as disease progresses, there isn't a strong enough response for clearance in any known patient. There are multiple possible explanations which have been proposed for this lack of response. The pathogenesis of the virus results in the majority of replication occurring within the dorsal root ganglion and the CNS, areas which are generally considered 'immune privileged' or at least restricted in surveillance compared to other tissues and organs. However, this reasoning has been challenged by recent discoveries involving immune responses in the CNS [Cserr HF & Knopf, 1992; Goldman *et al*., 2006]. There are also other viral infections which affect the CNS such as Borna disease virus which is effectively controlled by the immune system despite its location [Morimoto *et al*., 1996]. Viral loads have been established as a possible reason for the lack of immune response. As the virus is introduced through bite wounds, depending on the previous host, it has been suggested that this method of delivery may assist lyssaviruses in avoiding the immune system by being introduced in low levels to local tissue, therefore avoiding much of the innate defences, though this theory has not been extensively investigated.

One of the most significant discoveries surrounding host response to infection is the immunosuppression of host defences by RABV. Though the exact mechanisms are not yet fully understood, RABV has been observed supressing interferons and depleting lymphoid cells across multiple tissues such as the thymus, spleen and lymphoid nodes [Vidy *et al*., 2005; Brzozka *et al*., 2006; Torres-Anjel *et al*., 1988; Perry *et al*., 1990]. Interferon suppression is mediated through the viral phosphoprotein and matrix, which disrupt interferon signalling by binding to STAT isoforms, preventing accumulation within cell cytoplasm [Sonthonnax *et al*., 2019]. Additionally, viral phosphoprotein has been observed inhibiting translocation of STAT dimers into the nucleus, this results in the inability to initiate transcription of interferon as well as genes which are interferon-inducible [Vidy *et al*., 2007]. Another recent study by Besson *et al*., 2017, has further demonstrated the matrix proteins involvement in host immune suppression with evidence suggesting that NF-κB signalling is disturbed in lyssavirus infection. This

interaction results in the reduction of cytokines and chemokines which NF-κB regulates, such as IFNβ, TNF and CXCL2.

It has also been suggested that both macrophages and lymphocytes can be infected by lyssaviruses and that this kind of infection has an effect on cytokine expression profiles [LIT 64, 65]. This theory is further supported by the presence of NCAM receptors on immune cells such as NK cells, γδ T cells, CD8+ T cells and dendritic cells, though this link in NK cells has not been investigated further.

Demonstration of infection of NK cells by lyssaviruses would provide an interesting insight into the potential of further immune avoidance or interference during infection. The presence of the NCAM receptors could hint at an interaction or route of infection, therefore the aim of this study was to investigate the potential for infection of primary NK cells by RABV through the use of a pseudo virus system.

5.2 Methods

5.2.1 Whole blood extraction

Whole blood was sourced from volunteers which had filled in the University of Nottingham Faculty of Medicine & Health Sciences "Healthy Volunteer's Consent Form". They were made fully aware of what the blood was to be used for, and asked if they had any rabies vaccines previously. Blood was extracted using the NHS phlebotomy clinic found at Queens Medical Centre in Nottingham with all procedures and guidelines followed. Volunteers were not used more than once and have been kept anonymous.

5.2.2 PBMC and NK cell extraction

Whole blood was collected in standard phlebotomy collection tubes. These were carefully opened in a sterile cell culture hood. Whole blood SepMate PBMC isolation tubes (STEMCELL technologies) were filled with 15mL of 1077 Histopaque® (Sigma), with the whole blood added slowly on top using a Pasteur pipette. The blood was then centrifuged for 15 minutes at 1200g. Following the centrifugation, the fluffy layer containing the PBMCs was removed using a Pasteur pipette and topped up to 50ml with PBS in a fresh falcon tube and then centrifuged again at 300g for 10 minutes with 7 acceleration and deceleration. Following centrifugation, supernatant was poured off and cells were resuspended in RPMI and a cell count performed. The centrifugation at 300g was repeated and the appropriate amount of cells was then put through the EasySep™ Human NK Cell Isolation Kit (Stemcell technologies) protocol as per the standard manufacturer's instructions.

5.2.3 Transfection differences

Transfection was performed as previously described; however a two plasmid system was used in this instance, the PNL 4.3 plasmid which contained GFP was used alongside the standard glycoprotein plasmids used in all transfection experiments. Additionally, for troubleshooting, a three plasmid system using the HIV plasmid R84.7 and PCSGW GFP, and the glycoprotein of interest, was also used. A luciferase reporter was also used instead of GFP for further troubleshooting infection assays.

5.2.4 Infection assay and cell culture differences

NK cells were cultured in RPMI with 10% FBS and 1X pen/strep (1:100). After an initial experiment with no additives, IL-2 supplementation was used at 70ng/mL of RPMI. Prior to IL-2 stimulation, due to NK cells being unable to adhere to plates as cell lines do, infection assay protocols differed somewhat. Instead of plating cells into a 96 well plate, they were distributed into 15ml falcon tubes. Whenever cell washing or infection needed to be performed, they were

spun down at 300g with low acceleration and deceleration then resuspended either in fresh RPMI or in PV containing DMEM as needed. Cells were stored in an incubator as usual but the lids left loose due to the lack of cell culture flask air holes. Following introduction of IL-2 which generates NK cell adhesion, trypsinisation was introduced to retrieve adhered cells.

5.2.5 Sample preparation for flow cytometry

Flow cytometry was performed by the University of Nottingham Flow Cytometry Facility Manager: Dr David Onion. Samples were prepared for FACS following infection using a FACS buffer (50ml sterile PBS, 3% Bovine Serum Albumin and 2mM EDTA). Preparation involved the removal of adhered cells (when IL-2 had been used, causing adherence) using 0.5% trypsin. Cells were transferred to a new falcon and RPMI added to neutralise the trypsin. Cells were then centrifuged at 300g for 5 minutes. Supernatant was removed and cells were resuspended in 5mL of FACS buffer. A further centrifuge step was performed for 5 minutes at 300g followed the removal of the supernatant and resuspension in 500µl of FACS buffer. Cells were then transferred to a FACS tube and a final centrifugation stage occurred. The supernatant was again removed and cells were resuspended in 0.5% paraformaldehyde solution with a PBS diluent in order to become fixed.

5.3 Results

5.3.1 Producer cells transfected with GFP PVs expressing fluorescence

In order to determine the successful transfection of PNL4.3 based PVs into the HEK239T producer cells, images were taken following transfection using an iRiS Digital Cell Imaging System (Logos Biosystems) scanner which produced figure 5.1. The images depict the fluorescent glow of GFP from the cells which have been overlayed using iRiS software. This demonstrates successful transfection.

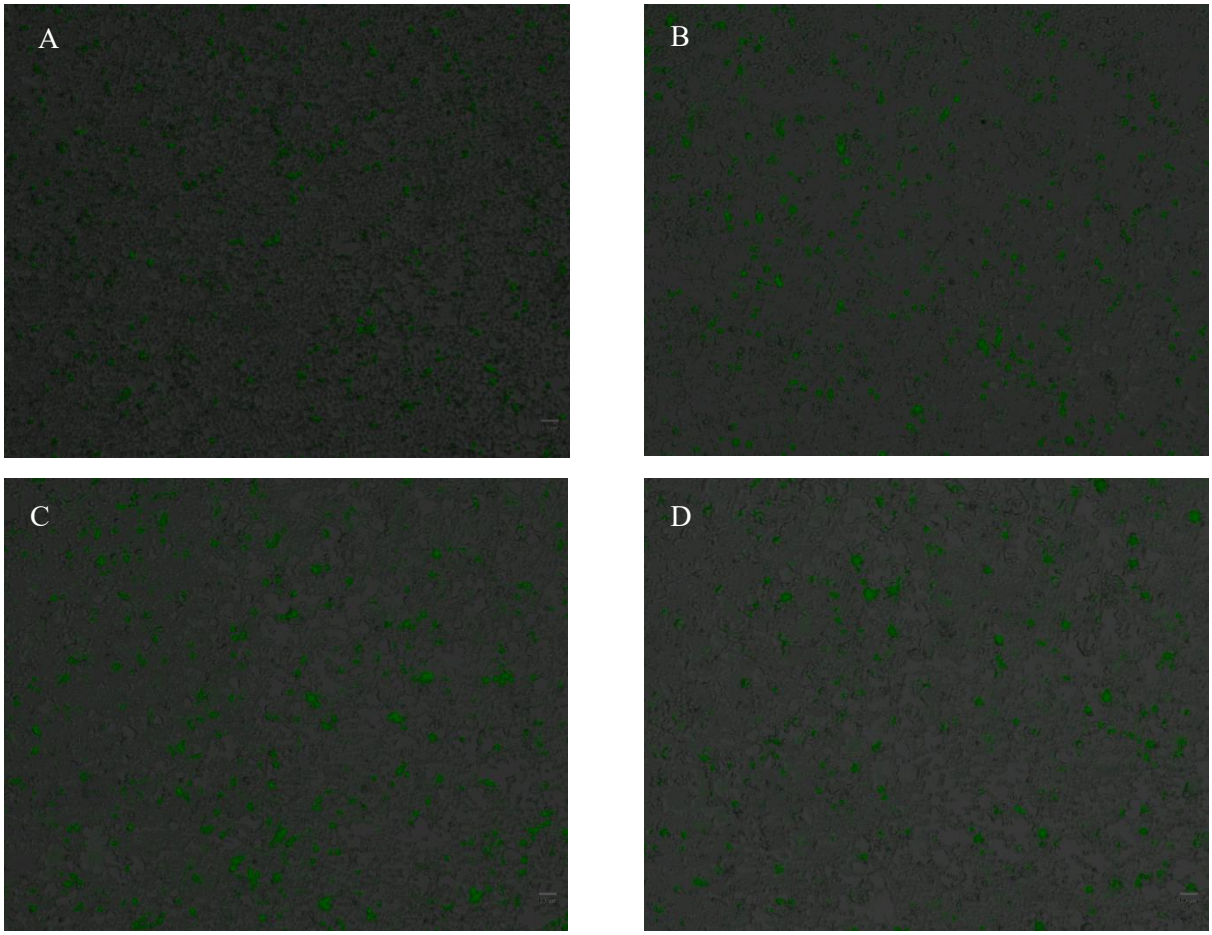


Figure 5.1: Images detected using an iRiS Digital Cell Imaging System depicting HEK239T producer cells expressing GFP indicating successful transfection. Each image represents a different PV construct (A): C15, (B): CVS-11, (C): Δ Envelope, (D): VSV.

5.3.2 Flow cytometry results of VeroE6 cell infection

VeroE6 cells were also infected in a PNL4.3 GFP assay, as they had been shown to be susceptible, and examined using the same flow cytometry protocol as the NK cells. This experiment also resulted in no levels of infection even in positive controls VSV and C15, indicating an issue with the protocol.

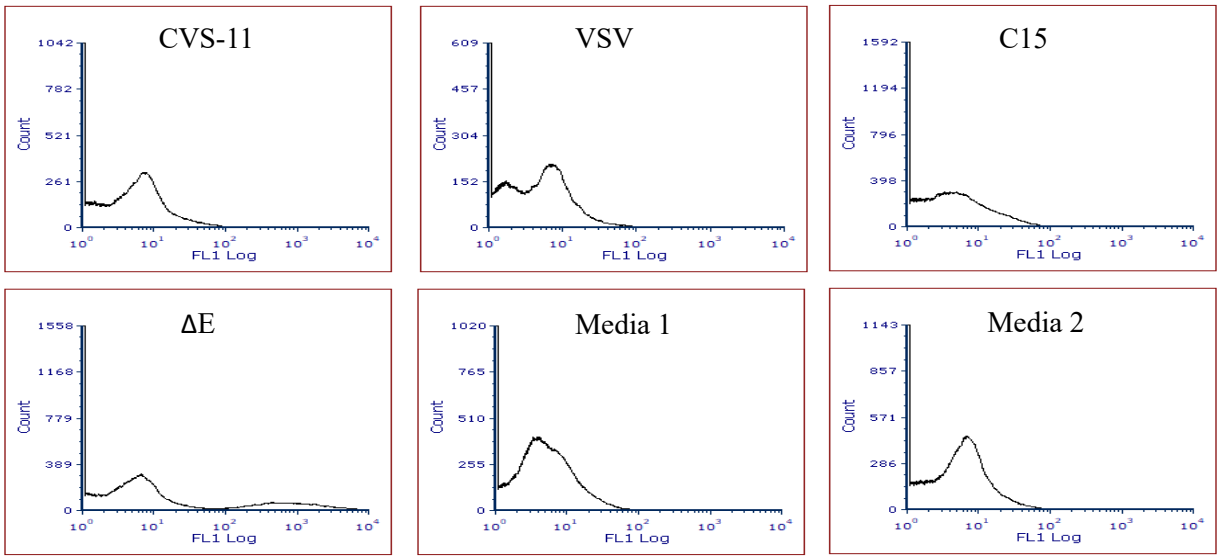
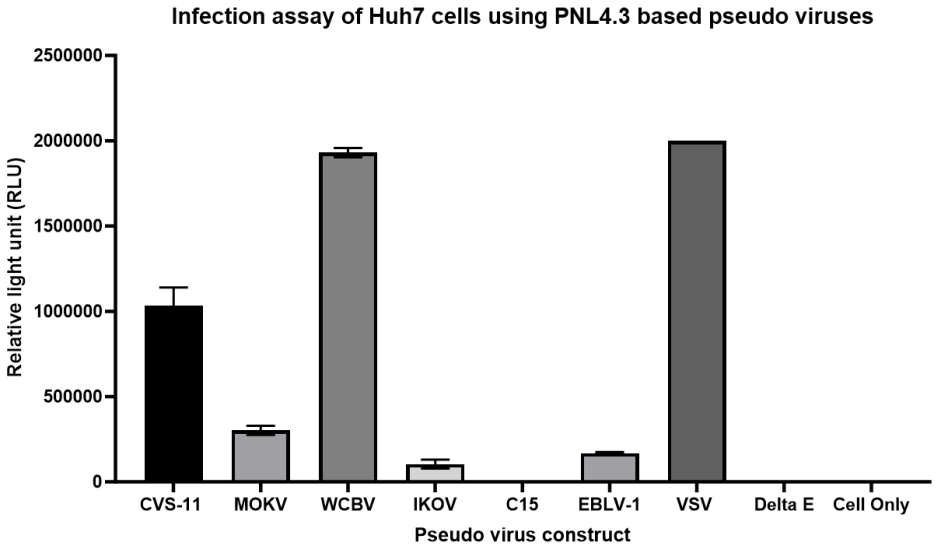


Figure 5.2: Flow cytometry histograms of a population of Vero E6 cells following an infection assay with multiple pseudo viruses including VSV, CVS and C15. Histograms 5 and 6 show media only controls. The ΔE construct was a negative control using a construct with no spike protein, thus it was unable to infect cells. The graphs display a full cell population, there was little change when populations were isolated using gates. All graphs display a negative result, with no GFP positive peak greater than the negative controls.

5.3.3 Infection assays using pseudo viruses generated for subsequent NK cell infection.

As part of the assessment of issues with the NK cell experiments, infection assays were performed using both of the pseudo virus constructs used in the NK cell infections. The results demonstrated that all pseudo viruses which had been used were indeed functional and infective, though the results of the MLV experiment suggested less than usual infectivity.

A



B

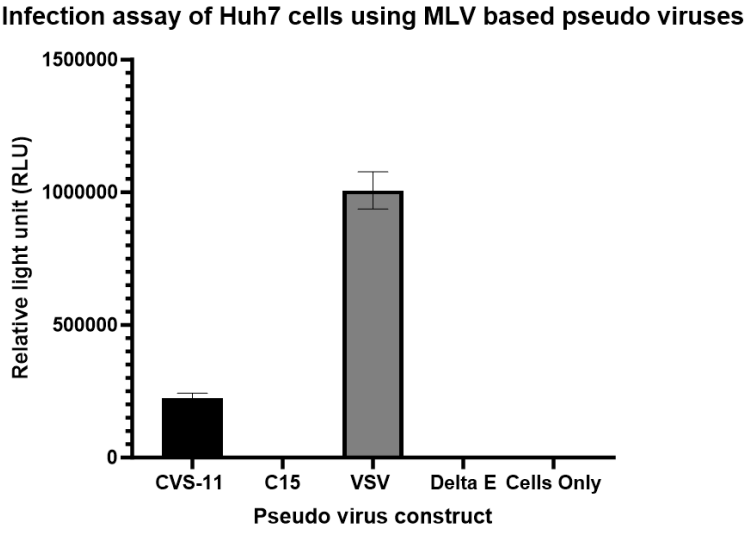


Figure 5.3 shows infection assays using (A) the HIV based PNL4.3 construct, and (B) the MLV based pi.18 construct. Both infectivity assays produced positive results, and were performed on the Huh7 cell line. A VSV positive control was also included, which is the same control used in the NK cell experiments.

5.4 Discussion

The aim of this study was to investigate the potential for NK cell infection by lyssavirus glycoprotein presenting PVs. Unfortunately, the study resulted in a lack of apparent infection. However, it is not the conclusion of this study that NK cells are not infectable by lyssaviruses, rather that the protocols performed need modification, as well as investigations using alternative methods in order to fully determine if this is the case. Indeed, failure of the positive control VSV G which has been used to perform this kind of study before successfully, Tran J & Kung, 2007, indicates this.

Initial attempts at producing functional pseudo-virus with GFP reporters in failed. In order to investigate this, a combination of different plasmid stocks was used to assess the viability of each plasmid component. After an infection experiment it appeared that the HIV vector was non-functional. This was replaced and infectious pseudo-virus was produced. Unfortunately, the subsequent GFP experiment infecting NK cells produced no significant results and it appeared that infection had not occurred. In order to troubleshoot the lack of infection, a number of methods were attempted. Producer cells were examined for GFP expression which they demonstrated, indicating successful transfection. The GFP reporter was replaced with a luciferase reporter and the NK cells were instead examined through a luciferase reaction, deemed to be more robust due to the experience of the lab in this assay. The reasoning behind this was that the GFP reporter plasmid used was anecdotally reported to have previously been inconsistent. Due to the fact that a luciferase method requires cells to adhere to the wells in plates, and NK cells do not adhere, multiple centrifugation steps were introduced when changing the media. Cells were resuspended in lysis buffer before a 15 minute incubation and then a reading on a luminescence plate reader. This once again returned negative results in all plasmids used, including the VSV G positive control.

Each glycoprotein plasmid had previously demonstrated exceptional infectivity and this was re-tested following the failure of NK infection, again proving they were functional. In an effort to increase susceptibility to infection, NK cells were incubated with IL-2 which enabled adherence and activation. Additionally, virus was concentrated using ultra-centrifugation. These infectivity assays however also did not produce results. Combined with the changes made to protocol and the ruling out of many potential causes of failure, it appears a few explanations remain. NK cells are potentially not susceptible to infection by the RABV G protein, or the experiment may need further modification to be successful and surpass the resistance NK cells have to genomic modification and infection. Evidence towards the latter of these suggestions is that an infection assay involving VeroE6 cells and the GFP system examined using flow cytometry in the same manner as the NK cells were, also ended in failure despite it being successful in previous

experimentation within the same lab. This points to issues with the protocol rather than NK cell infectivity not being possible in RABV infection.

Delivery of genes into an NK cell has been described in other studies as difficult, with the current reasons for this difficulty unknown, this has been reinforced by the findings of this study [Schmidt *et al*., 2021]. However, there have been some studies which have had moderate success [Tran & Kung, 2007; Schmidt *et al*., 2021; Mantesso *et al*., 2020]. The study by Tran & Kung, 2007 used an HIV-1 lentiviral PV system using GFP expression and achieved 40% transduction efficiency on primary NK cells, though they were sourced from mice. This NK cell transduction was performed both with and without IL-2 stimulation, indicating it is not essential to achieve these results. However, removal of IL-2 stimulation did reduce transduction efficiency. Transduction of NK cells was found not to impact NK surface phenotypes or functions such as apoptosis, cytokine production and cytotoxicity. One important finding of the study is that the type of promoter used in the PV system impacted the efficiency of transduction. The vectors used by the study were VSV-G presenting lentiviral PVs consisting of an EGFP reporter gene, either FUGW or Cppt2E, which carried the Ubi-C and RhMLV promoters respectively. While both vectors supported transduction it was found that the RhMLV promoter induced a much greater mean florescent intensity (MFI) than the Ubi-C. Indeed, a further comparison found that a phosphoglycerate kinase also outclassed Ubi-C, though the cytomegalovirus (CMV) promoter produced the lowest levels of MFI. These findings would be an excellent place to start in the continuation of this work, replacing the CMV promoter used with other promoters including the ones used by Tran & Kung, 2007. The study also demonstrated that different cytokine cocktails could be used to improve assay performance. While the experiments in this chapter used IL-2, Tran & Kung, 2007 also recommend the use of IL-15 in tandem with IL-12 or IL-12 with IL-2.

IL-2 and IL-21 stimulation has been found in another study by Sutlu *et al*., 2012 to also support efficient transduction using a VSV-G pseudo-typed lentiviral vector. However, the study went further and identified that inhibition of the innate immune receptor signalling, greatly enhanced the transduction results by 3.8 fold using the BX795 inhibitor. This inhibitor targets the TBK1/IKKε complex, acting downstream of RIH-I, MDA-5 and TLR3. These findings could be used to augment future experiments in order to demonstrate receptor binding of the lyssavirus G protein through the PV system. However, with the introduction of modifications to the NK cell antiviral defences, the representation of real-world infection may be lost. If NK cell defences are capable of preventing infection, then artificially removing them would not be particularly representative of a natural infection occurrence. The study also performed optimisation of the length of time required for the greatest transduction, which aligned with the

protocol used in this study. It was again confirmed that IL-2 stimulation was essential at reaching increased levels of transduction.

In the context of PV systems used, this study opted for the lentiviral HIV system. It has been previously shown that lentiviral transduction of NK cells can activate innate immune receptor signalling and trigger NK cell apoptosis [Mantesso *et al*., 2020]. Retroviral transduction has demonstrated improvement over the lentiviral system in NK cell lines, in particular NK-92, and even in *ex vivo* NK cells which had been expanded and activated [Imai *et al*., 2005]. However, a setback of retroviral systems is the requirement for dividing cells, something lentiviral PVs do not require.

Another potential avenue for investigation is the use of an NK cell line. These cell lines are replicative and therefore do not have the issue of non-division which excludes the use of MLV vectors. NK cell lines would also offer a more controlled and consistent testing environment compared to primary cells, with no need for volunteers or costly extraction from whole blood, and they would form an endless supply due to immortalisation. However, a drawback of such a system would be its questionable relationship with primary NK cells. As with all cell lines, the comparison between primary cells and their cell line counterparts should be scrutinised for differences. NK cells are highly regulated and present specific levels of CD16 and CD56. An NK cell line which presented artificial levels of the receptor CD56, would not be representative of the environment a lyssavirus would encounter. Indeed, only 10% of cells are CD56 bright without inflammation, which is most of the lyssavirus infection. Though this is the case, NK cell lines such as NK-92 are used in clinical trial studies, such as those on NK cells expressing chimeric antigen receptors for the treatment of cancers, which involves transduction of the cells [Mantesso *et al*., 2020]. This would support the argument that though the cell lines may not be truly representative of primary cells in population, they are appropriate alternatives. Ultimately, using an NK cell line to demonstrate the act of lyssavirus infection would be appropriate, but how common that infection is and to what degree the virus infects would not be if the population was artificial.

An additional improvement to the study would be the use of live virus. The PV system provides safety when handling such pathogenic viruses, however it does not perfectly mimic wild type virus in many ways. For one the glycoprotein structure on lyssaviruses is highly ordered, whereas in PVs they are more sporadic in nature [Evans *et al*, 2018]. Additionally, only the viral glycoprotein is used, meaning that other potential viral protein interactions cannot be studied. A live virus experiment would therefore be interesting to perform, using staining techniques to identify virus inside NK cells. Also, while the use of NK cells extracted from whole blood does provide a more useful model of study, it does introduce an issue when attempting this type of

NK cell infection. As stated, only 10% of the circulating NK cells in blood are CD56 bright, meaning they have an increased level of the lyssavirus receptor. This means that when extracting from whole blood, only a small portion of the NK cell population will retain significant CD56 levels, impacting potential infectivity. This could be circumvented with the use of NK cell lines which present as CD56 bright, improving potential infectivity.

NK cell infection could also be demonstrated through an alternative method, many of which exist. Staining techniques that could be used with PVs could include attaching a non-neutralising antibody which has been tagged with a fluorochrome, this is called an immunofluorescence assay. The assay is commonly combined with a DAPI stain of the nucleus to help visualise the cell better. This would be suitable for detecting viral proteins inside NK cells [Yi *et al*., 2019]. Electron microscopy can also be used to visualise virus within a single cell though this is more expensive and specialised.

An infectivity assay using the plaque method could produce interesting results in examining if viral infection of RABV would cause NK cell death. In a plaque assay, a monolayer of an NK cell line would be stained on an agar dish which had been mixed with diluted live virus. This is incubated and results in a monolayer with clear circular areas where cell death has occurred due to the infection, revealing if the virus does cause cell death and to what degree. This method can also be used even if the virus does not kill infected cells using focus-forming assays, where foci are the clusters of infected cells. Fluorescent molecules or substrate-cleaving enzymes, both of which tagged to antiviral antibodies, can be used. In this method, the foci would be revealed using fluorescence or colour changes in the presence of the antibodies which would be bound to the virus. This kind of non-killing method would be appropriate for PVs, though clusters would not form in the same way as PVs are non-replicative.

Additional changes to the current protocol could include other cocktails of interleukins, complementary to IL-2. IL-21 has been shown previously to synergise well with IL-2 during NK cell transduction by activating NK cells [Sutlu *et al*., 2012]. While the study by Sutlu *et al*., 2012, found adding further cytokines to this cocktail did increase transduction, it was deemed not to a degree worth the addition. IL-15 has also been used in NK cell transduction successfully, as has IL-12. It is likely that these additional changes in tandem are the appropriate route forward, as individual changes often illicit varying results in studies, with some finding much greater or lesser improvement compared to other studies utilising the same additions [Sutlu *et al*., 2012].

A further study in addition to these improvements could also be performed on population changes of NK cells within the CNS similar to those by Han *et al*., 2014, though this would require live infection with lyssavirus.

Ultimately while NK cell infection, if proven to occur in lyssavirus pathogenesis, would be interesting it would not be the 'smoking gun' of host immune avoidance and would instead be another facet of the overall immune interactions lyssaviruses have during infection. Lyssaviruses have evolved multiple strategies to become the silent invaders they are known to be. However, though this may be the case, many aspects of lyssavirus host interactions are still poorly understood and a lot more research needs to be done to decipher the full picture of this virus.

Viruses other than lyssaviruses have been shown to inhibit and escape NK cells through a variety of extracellular methods, without ever infecting the cell itself. There are several ways viruses achieve this. In HIV, the virus releases soluble NKG2D ligands via proteolytic shedding which impairs NKG2D (an essential activating receptor of NK cells) on the cells leading to reduction in NK cell cytotoxicity [Matusali *et al*., 2013]. There are also mechanisms such as the use of viral proteins, as in the Epstein-Barr Virus, which produces LMP2A that reduces expression of MICA and ULBP4 which would activate NK cells [Rancan *et al*., 2015]. This combined with evidence that the P protein of lyssaviruses is already known to downregulate immune responses and interact with the immune system, and that NK cell activation is key in vaccine response [Vidy *et al*, 2005, Vidy *et al*, 2007, Horowitz *et al*., 2010], makes for a potentially interesting avenue of investigation as to whether NK cells are downregulated or inhibited by lyssaviruses in other ways than infection. Indeed, P protein targeting of signal transducer and activator of transcription (STAT) 3 and STAT1 has been shown in lyssaviruses [Harrison *et al*., 2020]. The study demonstrated that this STAT antagonism of the P protein was conserved among diverse pathogenic lyssaviruses and correlated with pathogenesis. STAT1 is used in regulation of IFN-γ production and NK-cell cytotoxicity, while STAT3 is currently not well characterised however has been determined to be important in immunity [Harrison *et al*., 2020; Gotthardt & Sexl, 2017]. It is clear from the lack of literature on these interactions, and how lyssaviruses are able to escape significant immune response so effectively, that more needs to be done to uncover these important questions as understanding of these factors could lead to more effective therapies.

Conclusion

To conclude, it is unfortunate that, despite efforts to change the protocols of transduction, NK cells were not able to be infected by lyssavirus PVs in this study. Despite this, there are several approaches which remain that could be further explored and implemented to satisfy the conclusion of whether they can or cannot be infected by lyssaviruses and no conclusions can be reached without fully exhausting all avenues of investigation.

Chapter 6: Discussion

Lyssaviruses, which are the causative viruses of the disease rabies, represent a widespread and significant hazard to human health resulting in over 59,000 estimated deaths, and a cost of over $8 billion USD annually, yet they have been determined to be a neglected pathogen [Hampson *et al*, 2015]. Infections caused by the genus are primarily the result of a single virus, RABV, which produces over 99% of known infections, with 99% of those being induced by a bite from a rabid dog. However, any member of the genus can cause the disease rabies and multiple cases of this occurring have been documented.

Treatment of rabies is a lengthy and costly process involving the administration of vaccines and RIG over the course of multiple sessions, resulting in effective vaccine induced antibodies against the pathogen. However, the current vaccine regimes, storage and transport are costly and difficult for many developing nations, which are the most affected by RABV. Additionally, phylogroup II and III lyssaviruses currently have no treatments and infection with these viruses would be fatal regardless of intervention. It is because of these reasons that a novel vaccine needs to be produced, one which offers cross-neutralising properties in a more affordable, easier to administer, and store, form. The novel medoid vaccine examined by this study fits these criteria. Results of this study have shown that the vaccine employs potentially cross-phylogroup neutralising capabilities through its use of a medoid glycoprotein of the lyssavirus genus, a glycoprotein sequence which has higher sequence identity with all lyssavirus species. Results of neutralisation assays with vaccine induced sera found that the novel vaccine was able to partially neutralise MOKV and WCBV after a standard course, and following a boost at week 48 the MOKV and WCBV were completely neutralised and IKOV was reduced below 50% infectivity.

While the findings of cross-neutralising capabilities in the novel medoid vaccine are interesting and potentially exciting, they do come with some caveats. The study cites some shortcomings which need to be overcome in the future. The medoid vaccine study utilised animal models in the production of sera, however, the number of animals and therefore the numbers of analysed sera was not substantial enough for the study to draw conclusions as to the vaccines superiority when compared to the standard RABIPUR vaccine [Napolitano *et al*, 2020]. This subsequently impacts the conclusions of the cross-neutralisation data as the current number of samples and time points are somewhat restrictive. An increased number of sera samples and new vaccination time points would allow for greater resolution of the data.

Through the use of *in silico* data analysis, this study has demonstrated the gamut of lyssavirus PNGs present in the genus, and created predictions on the potential glycosylation efficiency of each site. The sequence variation at sequons and potential sequons were also examined using

WebLogo analysis and has created a picture of a genus with a variety of PNGs that exist in divergent lyssaviruses and are potentially only a single base mutation away from production, hinting at their genesis. Several other important conclusions were reached, such as the increase in PNGs in more divergent lyssaviruses, suggesting a possible link between them and both their resistance to phylogroup I neutralising antibodies, and their reduced virulence in the wild compared to RABV.

In addition to this, the use of a systematic literature review has uncovered a concerning trend in the generation of glycan knockout mutants through mutagenesis, demonstrating that the literature does not adequately consider the substitutions chosen, leading to potentially impacted results. Sequon substitutions were confirmed to produce different phenotypic responses in the infection and neutralisation assays performed on the mutants. These findings in tandem with the conclusions of the *in silico* investigations of chapter three, demonstrate a need for this kind of confirmatory work when examining glycosylation knockouts. This is because, while the results showed similar patterns in the cell line BHK-21, the levels of infectivity changed significantly with certain substitutions. Therefore, if work was performed with only one type of substitution it could impact the conclusions of a study and jeopardise the integrity of the work. It is currently commonplace in literature to rely on previous and often very different studies to decide on the knockout provided, with little in the way of justification or validation as demonstrated in the systematic review of chapter three. While this situations impact on literature may currently be smaller in scope, as the ever-growing field of glycobiology further matures it could worsen. The field should be based on consistent and repeatable principles backed by validation such as this. Some studies in other viruses or glycan research have performed multiple substitutions such as that by Julithe et al., 2014, in order to determine the impact on their work and to rule out amino acid choice as a cause for particular results, which is a practice which would be recommended by this study in future work, similar to that of utilising multiple cell lines in certain experiments.

In addition to the substitution results, this study was able to determine that the infectivity and neutralisation profiles of the glycoprotein are impacted when certain glycans are removed. The removal of Asn37 from the CVS-11 glycoprotein increased infectivity but did not change the neutralisation profile of the protein. In contrast, removal of Asn204 greatly reduced infectivity and increased susceptibility to neutralisation by NHP sera. When both were removed the combined knockout induced an even greater loss of infectivity and even higher susceptibility to neutralisation. This result contributes to the understanding of glycan importance in the genus.

Finally, the site swap experiments were not able to replicate the sequon genotypes of divergent lyssaviruses due to unknown causes. It appears that the important residue 333 may have been a

factor, as all three lyssaviruses had a sequon present at Asn334 which this study introduced to the CVS-11 strain. However, the results did demonstrate that the sites which were successfully implemented into the CVS-11 strain did not reduce neutralisation susceptibility but instead increased it. This evidence potentially rules these glycans out as part of the reasoning behind the phylogroup III resistance to phylogroup I sera neutralisation.

NK cell infection has been shown to be of importance in other viruses such as Influenza and HIV. This study attempted to demonstrate this same interaction may exist in lyssaviruses by infecting NK cells with lyssavirus PVs. The importance of this discovery would be in the elucidation of a currently puzzling phenomenon in lyssavirus infection which is the lack of an effective immune response. While the immune privilege of the CNS goes some way to explaining this, it is unsatisfactory due to the fact that immune cells have been shown to invade the CNS in times of infection, including NK cells [Shi & Ransohoff, 2010], and that other CNS invading viruses such as Borna Disease Virus are detected and effectively controlled [Morimoto et al, 1996]. In contrast, some viruses reach chronic conditions by becoming dormant in neurons such as herpes [Whitley, 2001], however lyssaviruses are not only generally non-dormant but rapidly replicate and spread. While some mechanisms have been explored such as the interaction of the viral P protein with the host immune system, reducing its effectiveness, the full explanations do not appear to have been uncovered. While the experiment was unsuccessful it did leave many avenues which still need to be explored before this phenomenon can be ruled out such as modifications to the experiments parameters from plasmid promoter changes and NK cell line use, to new experiments which could be performed with this framework such as immunofluorescence assays and plaque infection assays. These findings form a foundation for future work.

The pseudo-virus system is a flexible and extremely useful system when investigating highly pathogenic or deadly viruses, such as lyssaviruses. Though this system offers exceptional flexibility and safety, it does come with some drawbacks. The lyssavirus glycoprotein is highly ordered and structured on a bullet shaped virion [Evans et al, 2018]. This order is functional and is not reflected in the artificial virions of the PV system. Though this does not prevent infection or neutralisation it should be noted that the presentation of these glycoproteins is not entirely natural and has been shown to increase the sensitivity of the target to neutralisation [Evans et al., 2018; Corti et al., 2011; Wright et al., 2010]. It would therefore be beneficial that the experiments performed using neutralisation be performed with WT virus.

In the PV system only the target glycoprotein is utilised. This is a useful function for studying this protein in isolation, however, as has been demonstrated there are interactions with viral proteins and the glycoprotein in lyssaviruses which do impact virion production such as that of

the matrix [Mebatsion *et al*., 1999]. This therefore means that any interactions between viral proteins is lost within the PV system and that should be considered when assessing the data produced by said system.

The PV system has some inherent issues in reproducibility of exact RLU values. Cell lines were vulnerable to changes in phenotypic responses based on their handling, with some passages and assays producing differing RLU values. This was combated by only comparing data that was produced in full runs and treated identically. For example, if a panel of mutants was produced through transfection, this panel would only be used in experiments together, never interchanged with subsequent transfections or compared to other outcomes of experiments such as infection assays. Though this meant experiments became quite large at some stages, it removed the variation due to problems with cell line behaviour.

This study has outlined a number of areas which could use further research to both explore novel avenues but also improve and bolster the work produced here. One such instance is the cross-neutralising medoid vaccine which represents an interesting find, however the mechanisms behind the cross-reactivity beyond the fact the medoid glycoprotein was used need to be elucidated. In order to do this, structural models could be developed which identify key regions of structural change, with particular investigation into the changes involving antigenic regions. Repeating the success of the vaccine is also key, as has been demonstrated in the study by Malerczyk *et al*., 2014 involving MOKV neutralisation with RABIPUR. It is important to validate findings such as these with a completely repeated study.

The glycan work could be further bolstered with successful western blots which identified the presence of glycans on the surface of the glycoproteins used in these experiments and their removal from the glycoprotein through mutation. Though the findings are interesting on their own merit, and other studies have demonstrated the glycosylation efficiencies of these positions, it is good practice to further validate the findings. Further expansion of this type of experiment could also be performed such as examining other lyssavirus adaptations or the importance of sequons in more divergent lyssaviruses.

Additionally, the glycan work provides interesting introductions to more complex work such as identifying if different cells attach different types of glycans on the surface of the glycoprotein and how that impacts phenotypic outcomes of the virus.

Finally, the NK cell work conducted in this study paves the way for many improvements and experiments which could lead to more successful results. These include the use of new cytokine cocktails, NK cell lines such as NK-92, improvements to experimental practices, novel experiments such as cell staining methods, and use of differing promoters suggested by the

literature [Tran & Kung, 2007]. Though the results were disappointing, the potential for future work is extensive and interesting.

To conclude, lyssaviruses are an important yet neglected pathogen with numerous avenues of research not yet explored. This study contributes to the knowledge base with the discovery of cross-neutralising properties of a novel vaccine which, if perfected, could save lives and make the transition to a human lyssavirus infection free world easier. It also uncovers interesting particularities in the glycan arrangement of the genus using sequence data, pointing to a possible connection between glycosylation and neutralisation. These findings were then taken further, with a demonstration that the removal of currently known RABV PNGs impacts infectivity, and in some cases, neutralisation susceptibility. The glycan site swap represents a novel experiment, and with some further optimisation, could add to the research surrounding the question of lyssavirus phylogroup neutralisation susceptibility. Finally, the interaction between lyssaviruses and the host immune system is a topic which is far from complete with many questions still unanswered. Though this study fails to provide concrete answers, it does provide the foundation for further research to be performed in this area to confirm whether or not there is NK cell interaction during lyssavirus infection.

Professional internship reflective statement

As part of the BBSRC DTP it was required that a three-month internship placement was performed. I chose a lab at the University of Nottingham which I was aware performed significant amounts of bioinformatics work, a field I have a great deal of interest in joining. The project supervisor Dr Adam Blanchard offered a virus discovery project based on a dataset he had previously acquired. The dataset was faecal metatranscriptomic sequence data taken from several different animals including stoats and bats.

The project involved learning and using bioinformatics to interrogate the very large data set for novel viruses, first cleaning it up and then sorting and searching through it all using various command line based software. The project resulted in me finding a novel virus from one of the UK stoat samples which I was able to extract and then submit to GenBank, a very important process for any bioinformatician to have experience in. To follow this up, I was given the opportunity to first author a paper on the finding which I wrote and completed while on the project and is currently awaiting publication. This represents a fantastic success for the project and will be very useful in my career to come.

Through this project I was able to gain many useful skills. For example, the use of command line and the operating system ubuntu which was used throughout the project. These tools are essential to any budding bioinformatician and being able to use them to produce publishable results is an excellent outcome. Furthermore, the various software I was able to use such as FastP and FastQ, Kraken, MetaViralSpades, Megahit, Viral verify and Viral complete have given me an insight into how to perform work in pipelines and what each stage of a process like this looks like. In addition to the more technical skills I gained, I also got to experience an essential part of research, the writing of a paper which I had previously not been involved in. This process enabled me to better understand what was required and expected for a publication and the types of communication you need to properly produce it. I also learned how to submit a novel sequence to GenBank which is also an important skill for a virologist to have. Finally and most importantly, I learned that this area of science is one I am capable of being successful in, as prior to this I thought it too complex to achieve for me in addition to my lack of previous experience in the field.

As eluded to previously, this kind of work is exactly what I would envision myself doing following the PhD and experience in it is essential. Therefore this project and its successes will definitely positively impact my career prospects. This is in addition to the knowledge I have gained and the tools I was able to become familiar with, all things which are used regularly in the industry. Finally, in science networking is very important and the connections I made

through this PIP will also be invaluable. Overall I very much enjoyed the project and learned a lot and I am very grateful that Dr Adam Blanchard was able to give me the opportunity.

References

Aebi M. N-linked protein glycosylation in the ER. Biochim Biophys Acta. 2013;1833(11):2430-7.

Albach C, Damoe E, Denzinger T, Schachner M, Przybylski M, Schmitz B. Identification of *N*-glycosylation sites of the murine neural cell adhesion molecule NCAM by MALDI-TOF and MALDI-FTICR mass spectrometry. Analytical Bioanalytical Chem. 2004;378:1129-35.

Alberer M, Gnad-Vogt U, Hong HS, Mehr KT, Backert L, Finak G, *et al*. Safety and immunogenicity of a mRNA rabies vaccine in healthy adults: an open-label, non-randomised, prospective, first-in-human phase 1 clinical trial. Lancet. 2017;390:1511–20.

Albertini AA, Baquero E, Ferlin A, Gaudin Y. Molecular and cellular aspects of rhabdovirus entry. Viruses. 2012;4(1):117-39.

Allworth A, Murray K, Morgan J. A human case of encephalitis due to a lyssavirus recently identified in fruit bats. Comm Dis Intell. 1996;20(24):504.

Anilionis A, Wunner WH, Curtis PJ. Structure of the glycoprotein gene in rabies virus. Nature. 1981;294:275–8.

Annand EJ & Reid PA. Clinical review of two fatal equine cases of infection with the insectivorous bat strain of Australian bat lyssavirus. Aus Vet J. 2014;92(9):324-32.

Astray RM, Jorge SAC, Pereira CA. Rabies vaccine development by expression of recombinant viral glycoprotein. Arch Virol. 2017;162:323–32.

Badrane H & Tordo N. Host switching in Lyssavirus history from the Chiroptera to the Carnivora orders. J Virol. 2001;75(17):8096-104.

Badrane H, Bahloul C, Perrin P, Tordo N. Evidence of two Lyssavirus phylogroups with distinct pathogenicity and immunogenicity. J Virol. 2001;75(7):3268-76.

Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, *et al*. Accurate prediction of protein structures and interactions using a 3-track network. Science. 2021;373(6557):871-6.

Bailey MJ, McLeod DA, Kang CY, Bishop DH. Glycosylation is not required for the fusion activity of the G protein of vesicular stomatitis virus in insect cells. Virology. 1989;169(2):323-31.

Bakker ABH, Python C, Kissling CJ, Pandya P, Marissen WE, Brink MF, *et al*. First administration to humans of a monoclonal antibody cocktail against rabies virus: Safety, tolerability, and neutralizing activity. Vaccine. 2008;26:5922–7.

Bamji SX, Majdan M, Pozniak CD, Belliveau DJ, Aloyz R, Kohn J, *et al*. The p75 neurotrophin receptor mediates neuronal apoptosis and is essential for naturally occurring sympathetic neuron death. J Cell Biol. 1998;140(4):911-23.

Banyard AC, Evans JS, Luo TR, Fooks AR. Lyssaviruses and bats: emergence and zoonotic threat. Viruses. 2014;6(8):2974-90.

Barr JN, Whelan SPJ, Wertz GW. Cis-Acting signals involved in termination of vesicular stomatitis virus mRNA synthesis include the conserved AUAC and the U7 signal for polyadenylation. J Virol. 1997;71(11):8718-25.

Barrett J, Höger A, Agnihotri K, Oakey J, Skerratt LF, Field HE, *et al*. An unprecedented cluster of Australian Bat Lyssavirus in *Pteropus conspicillatus* indicates pre-flight flying fox pups are at risk of mass infection. Zoonoses Public Health. 2020;67:435-42.

Barrett PN, Mundt W, Kistner O, Howard MK. Vero cell platform in vaccine production: Moving towards cell culture-based viral vaccines. Expert Rev Vaccines. 2009;8:607-18.

Bause, E. Structural requirements of *N*-glycosylation of proteins. Biochem J. 1983;209:331-6.

Bavarian Nordic A/S. RabAvert Rabies Vaccine. 2023. [cited 2023 Nov 9]. Available from: https://www.fda.gov/media/83874/download?attachment.

Bavarian Nordic A/S. Rabipur New Zealand Data Sheet. 2022. [cited 2023 Nov 8]. Available from: https://www.medsafe.govt.nz/profs/datasheet/r/rabipurinj.pdf.

Beasley DW, Whiteman MC, Zhang S, Huang CY, Schneider BS, Smith DR, *et al*. Envelope protein glycosylation status influences mouse neuroinvasion phenotype of genetic lineage 1 West Nile virus strains. J Virol. 2005;79(13):8339-47.

Benmansour A, Brahimi M, Tuffereau C, Coulon P, Lafay F, Flamand A. Rapid sequence evolution of street rabies glycoprotein is related to the highly heterogeneous nature of the viral population. Virology. 1992;187(1):33-45.

Benmansour A, Leblois H, Coulon P, Tuffereau C, Gaudin Y, Flamand A, Lafay F. Antigenicity of rabies virus glycoprotein. J Virol. 1991;65(8):4198-203.

Besson B, Sonthonnax F, Duchateau M, Khalifa YB, Larrous F, Eun H, *et al*. Regulation of NF-κB by the p105-ABIN2-TPL2 complex and RelAp43 during rabies virus infection. PLoS Pathogens. 2017;13(10).

Bio-Rad Laboratories. Immun-Blot® PVDF membrane for protein blotting instruction manual. [Internet]. 2000. [cited 2023 Nov 9]. Available from: https://www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin_4006127A.pdf.

Blondel D, Maarifi G, Nisole S, Chelbi-Alix MK. Resistance to rhabdoviridae infection and subversion of antiviral responses. Viruses. 2015;7(7):3675-702.

Boonsriroj H, Manalo DL, Kimitsuki K, Shimatsu T, Shiwa N, Shinozaki H, *et al*. A pathological study of the salivary glands of rabid dogs in the Philippines. J Vet Med Sci. 2016;78(1):35-42.

Brady ST *et al*., editors. Basic Neurochemistry. Principles of Molecular, Cellular and Medical Neurobiology. 8th ed. London: Elsevier Inc; 2012.

Brzózka K, Finke S, Conzelmann KK. Identification of the rabies virus alpha/beta interferon antagonist: Phosphoprotein P interferes with phosphorylation of interferon regulatory factor 3. J Virol. 2005;79(12):7673-81.

Burrage TG, Tignor GH, Smith AL. Rabies virus binding at neuromuscular junctions. Virus Res. 1985;2(3):273-89.

Buthelezi SG, Dirr HW, Chakauya E, Chikwamba R, Martens L, Tsekoa TL, *et al*. The *Lyssavirus* glycoprotein: A key to cross-immunity. Virol. 2016;498:250-6.

Callaway HM, Zyla D, Larrous F, de Melo GD, Hastie KM, Avalos RD, *et al*. Structure of the rabies virus glycoprotein trimer bound to a prefusion-specific neutralizing antibody. Sci Adv. 2022;8(24).

CDC. Rabies. [Internet]. 2020. [cited 2023 Nov 8]. Available from: https://www.cdc.gov/rabies/about.html.

Ceccaldi P E, Gillet J P, Tsiang H. Inhibition of the transport of rabies virus in the central nervous system. J. Neuropathol. Exp. Neurol. 1989;48:620–30.

Celis E, Karr RW, Dietzschold B, Wunner WH, Koprowski H. Genetic restriction and fine specificity of human T cell clones reactive with rabies virus. J Immunol. 1988;141(8):2721-8.

Cenna J, Hunter M, Tan GS, Papaneri AB, Ribka EP, Schnell MJ, *et al*. Replication-deficient rabies virus–based vaccines are safe and immunogenic in mice and nonhuman primates. J Infect Dis. 2009;200:1251–60.

Chao T-YY, Ren S, Shen E, Moore S, Zhang SF, Chen L, *et al*. SYN023, a novel humanized monoclonal antibody cocktail, for post-exposure prophylaxis of rabies. PLoS Negl Trop Dis. 2017;11.

Charlton KN & Casey GA. Experimental rabies in skunks: immunofluorescence light and electron microscopic studies. Lab Invest. 1979;41(1):36-44.

Chehimi J, Bandyopadhyay S, Prakash K, Perussia B, Hassan NF, Kawashima H, *et al*. In vitro infection of natural killer cells with different human immunodeficiency virus type 1 isolates. J Virol. 1991;65(4):1812-22.

Chenik M, Chebli K, Gaudin Y, Blondel D. In-vivo interaction of rabies virus phosphoprotein (P) and nucleoprotein (N)-Existence of 2 N-binding sites on P-protein. J Gen Virol. 1994;75:2889-96.

Chenik M, Schnell M, Conzelmann KK, Blondel D. Mapping the interacting domains between the rabies virus polymerase and phosphoprotein. J Virol. 1998;72(3):1925-30.

Cleaveland, Fevre, Kaare, & Coleman 2002: Cleaveland S, Fevre EM, Kaare M, Coleman PG. Estimating human rabies mortality in the United Republic of Tanzania from dog bite injuries. Bulletin WHO. 2002;80(4):304-10.

CM, Murphy WJ, O'Brien SJ, Jacobs DS, Springer MS, Teeling EC. A family matter: conclusive resolution of the taxonomic position of the long-fingered bats, miniopterus. Mol Biol Evol. 2007;24(7):1553-61.

Coelingh B. A structural study on the rabies virus glycoprotein combining cryo-EM and mass spectrometry-based approaches. Master's Thesis. Utrecht University. 2021.

Coertse J, Grobler CS, Sabeta CT, Seamark ECJ, Kearney T, Paweska JT, Markotter W. Lyssaviruses in Insectivorous Bats, South Africa, 2003-2018. Emerg Infect Dis. 2020;26(12):3056-60.

Coertse J, Markotter W, Roux K, Stewart D, Sabeta CT, Nel LH. (2017). New isolations of the Rabies-related Mokola Virus from South Africa. BMC Vet Res. 2017;13(37).

Coertse J, Viljoen N, Weyer J, Markotter W. Comparative neutralization activity of commercial rabies immunoglobulin against diverse lyssaviruses. Vaccines (Basel). 2023;11(7):1255.

Colloca S, Barnes E, Folgori A, Ammendola V, Capone S, Cirillo A, *et al*. Vaccine vectors derived from a large collection of simian adenoviruses induce potent cellular immunity across multiple species. Sci Transl Med. 2012;4(115).

Constantine DG. Bat rabies and other lyssavirus infections. Reston, Virginia, US. Geological Survey Circular 1329. 2009.

Conzelmann KK, Cox JH, Schneider LG, Thiel HJ. Molecular cloning and complete nucleotide sequence of the attenuated rabies virus SAD B19. Virol. 1990;175(2):485-99.

Cooper MA, Fehniger TA, Caligiuri MA. The biology of human natural killer-cell subsets. Trends Immunol. 2001;22(11):633-40.

Corti D, Voss J, Gamblin SJ, Codoni G, Macagno A *et al*. A neutralizing antibody selected from plasma cells that binds to group 1 and group 2 influenza A hemagglutinins. Science. 2011;333:850-6.

Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: A sequence logo generator. Gen Res. 2004;14:1188-90.

Cserr HF & Knopf PM. Cervical lymphatics, the blood-brain barrier and the immunoreactivity of the brain: A new view. Immunol Today. 1992;13:507-12.

Dallérac G, Rampon C, Doyère V. NCAM function in the adult brain: lessons from mimetic peptides and therapeutic potential. Neurochem Res. 2013;38(6):1163-73.

De Benedictis P, Minola A, Rota Nodari E, Aiello R, Zecchin B, Salomoni A, *et al*. Development of broad-spectrum human monoclonal antibodies for rabies post-exposure prophylaxis. EMBO Mol Med. 2016;8:407–21.

Delagneau JF, Perrin P, Atanasiu P. Affinity of lectins for rabies virus: Partial structural analysis of oligosaccharides and virus purification. Annales de l'Institut Pasteur. 1981;132(4):461-71.

Dhulipala S & Uversky VN. Looking at the pathogenesis of the rabies Lyssavirus strain pasteur vaccins through a prism of the disorder-based bioinformatics. Biomol. 2022;12(10):1436.

Dietzschold B, Wang HH, Rupprecht CE, Celis E, Tollis M, Ertl H, *et al*. Induction of protective immunity against rabies by immunization with rabies virus ribonucleoprotein. Proc Natl Acad Sci USA. 1987;84(24):9165–9.

Dietzschold B, Wunner WH, Wiktor TJ, Lopes AD, Lafon M, Smith CJ, Koprowski H. Characterisation of an antigenic determinant of the glycoprotein that correlates with the pathogenicity of rabies virus. PNAS. 1982;80:70-4.

Doener F, Hong HS, Meyer I, Tadjalli-Mehr K, Daehling A, Heidenreich R, *et al*. RNA-based adjuvant CV8102 enhances the immunogenicity of a licensed rabies vaccine in a first-in-human trial. Vaccine. 2019;37:1819–26.

Doran RC, Morales JF, To B, Morin TJ, Theolis R Jr, O'Rourke SM, *et al*. Characterization of a monoclonal antibody to a novel glycan-dependent epitope in the V1/V2 domain of the HIV-1 envelope protein, gp120. Mol Immunol. 2014;62(1):219-26.

Drings A, Jallet C, Chambert B, Tordo N, Perrin P. Is there an advantage to including the nucleoprotein in a rabies glycoprotein subunit vaccine? Vaccine. 1999;17(11–12):1549–57.

Duarte NFH, Pires Neto RDJ, Viana VF, Feijão LX, Alencar CH, Heukelbach J. Clinical aspects of human rabies in the state of Ceará, Brazil: an overview of 63 cases. Rev Soc Bras Med Trop. 2021;23:54.

Ertl HC, Dietzschold B, Otvos L Jr. T helper cell epitope of rabies virus nucleoprotein defined by tri- and tetrapeptides. Eur J Immunol. 1991;21(1):1–10.

Ertl HCJ. New Rabies Vaccines for Use in Humans. Vaccines. 2019;7:54.

Etchison JR & Holland JJ. Carbohydrate composition of the membrane glycoprotein of vesicular stomatitis virus. Virology. 1974;60(1):217-29.

Etessami R, Conzelmann KK, Fadai-Ghotbi B, Natelson B, Tsiang H, Ceccaldi PE. Spread and pathogenic characteristics of a G-deficient rabies virus recombinant: an in vitro and in vivo study. J Gen Virol. 2000;81(9):2147-53.

Evans JS, Selden D, Wu G, Wright E, Horton DL, Fooks AR, Banyard AC. Antigenic site changes in the rabies virus glycoprotein dictates functionality and neutralizing capability against divergent Lyssaviruses. J Gen Virol. 2018;99:169-80.

Faber M, Faber ML, Papaneri A, Bette M, Weihe E, Dietzschold B, Schnell MJ. A single amino acid change in rabies virus glycoprotein increases virus spread and enhances virus pathogenicity. J Virol. 2005;79(22):14141-8.

Faber M, Pulmanausahakul R, Hodawadekar SS, Spitsin S, McGettigan JP, Schnell MJ, *et al*. Overexpression of the rabies virus glycoprotein results in enhancement of apoptosis and antiviral immune response. J Virol. 2002;76:3374–81.

Fernando BG, Yersin CT, José CB, Paola ZS. Predicted 3D model of the rabies virus glycoprotein trimer. Biomed Res Int. 2016;2016:1674580.

Finke S, Cox JH, Conzelmann K-K. Differential transcription attenuation of 126 rabies virus genes by intergenic regions: generation of recombinant viruses overexpressing the polymerase gene. J Virol. 2000;74:726-9.

Finnegan CJ, Brookes SM, Johnson N, Smith J, Mansfield KL, Keene VL, *et al*. Rabies in North America and Europe. J R Soc Med. 2002;95:9–13.

Fischer S, Freuling CM, Müller T, Pfaff F, Bodenhofer U, Höper D, *et al*. Defining objective clusters for rabies virus sequences using affinity propagation clustering. PLoS Negl Trop Dis. 2018;12(1).

Flamand A & Delagneau JF. Transcriptional mapping of rabies virus in vivo. J Virol. 1978;28(2):518-23.

Foggin CM. Atypical rabies virus in cats and a dog in Zimbabwe. Vet Rec. 1982;110(14):338.

Foggin CM. Rabies and rabies-related viruses in Zimbabwe: Historical, virological and ecological aspects. PhD thesis. 1988; Harare: University of Zimbabwe.

Fooks A, Cliquet F, Finke S, Freuling C, Hemachudha T, Mani RS, *et al.* Rabies. Nat Rev Dis Primers. 2017;3:17091.

Fooks AR & Jackson AC, editors. Rabies. 4th ed. London: Elsevier Inc; 2020.

Fooks AR, Banyard AC, Ertl HCJ. New human rabies vaccines in the pipeline. Vaccine. 2019;37:A140–5.

Fooks AR, Banyard AC, Horton DL, Johnson N, McElhinney LM, Jackson AC. Current status of rabies and prospects for elimination. Lancet. 2014;384(9951):1389-99.

Fooks AR, Brookes SM, Johnson N, McElhinney LM, Hutson AM. European Bat Lyssaviruses: An emerging zoonosis. Epidemiol Infect. 2003;131:1029-39.

Fooks AR, Shipley R, Markotter W, Tordo N, Freuling CM, Müller T, *et al*. Renewed public health threat from emerging lyssaviruses. Viruses. 2021;13(9):1769.

Francis JR, Nourse C, Vaska VL, Calvert S, Northill JA, McCall B, Mattke AC. Australian bat lyssavirus in a child: The first reported case. Pediatrics. 2014;133(4).

François KO & Balzarini J. The highly conserved glycan at asparagine 260 of HIV-1 gp120 is indispensable for viral entry. J Biol Chem. 2011;286(50):42900-10.

Franka R, Carson WC, Ellison JA, Taylor ST, Smith TG, Kuzmina NA, *et al*. In vivo efficacy of a cocktail of human monoclonal antibodies (CL184) against diverse North American bat rabies virus variants. Trop Med Infect Dis. 2017;2(3):48.

Freeze HH & Kranz C. Endoglycosidase and glycoamidase release of N-linked glycans. Curr Protoc Mol Biol. 2010.

Fu ZF, Zheng YM, Wunner WH, Koprowski H, Dietzschold B. Both the N-terminal and the C-terminal domains of the nominal phosphoprotein of rabies virus are involved in binding to the nucleoprotein. Virol. 1994;200(2):590-7.

Furesz J, Fanok A, Contreras G, Becker B. Tumorigenicity testing of various cell substrates for production of biologicals. Develop Biol Stand. 1989;70:233-43.

Gan H, Hou X, Wang Y, Cheng ZJ, Zhu Z, Sun B. Global burden of rabies in 204 countries and territories, from 1990 to 2019: results from the Global Burden of Disease Study 2019. Intern J Infect Dis. 2023;126:136-44.

Gaudin Y, Raux H, Flamad A, Ruigrok RWH. Identification of amino acids controlling the low-pH-induced conformational change of rabies virus glycoprotein. J Virol. 1996;70(11):7371-8.

Gaudin Y, Ruigrok RWH, Knossow M, Flamand A. Low-pH conformational changes of rabies virus glycoprotein and their role in membrane fusion. J Virol. 1993;67(3):1365-72.

Gaudin Y, Tuffereau C, Segretain D, Knossow M, Flamand A. Reversible conformational changes and fusion activity of rabies virus glycoprotein. J Virol. 1991;65:4853–9.

Gaudin Y. Folding of rabies virus glycoprotein: epitope acquisition and interaction with endoplasmic reticulum chaperones. J Virol. 1997;71:3742–50.

Gerken TA, Butenhof KJ, Shogren R. Effects of glycosylation on the conformation and dynamics of O-linked glycoproteins: carbon-13 NMR studies of ovine submaxillary mucin. Biochem. 1989;28(13):5536-43.

Gillet J P, Derer P, Tsiang H. Axonal transport of rabies virus in the central nervous system of the rat. J. Neuropathol. Exp. Neurol. 1986;45:619–34.

Goldman J, Kwidzinski E, Brandt C, Mahlo J, Richter D, Bechmann I. T-cells traffic from brain to cervical lymph nodes via the criboid plate and the nasal mucosa. J Leukocyte Biolo. 2006;80:767-801.

Goodpasture EW, Woodruff AM, Buddingh GJ. The cultivation of vaccine and other viruses in the chorioallantoic membrane of chick embryos. Science. 1931;74(1919):371-2.

Gotthardt D & Sexl V. STATs in NK-Cells: The good, the bad, and the ugly. Front Immunol. 2017;7:694.

Grobler CS, Coertse J, Markotter W. Complete Genome sequence of Matlo Bat Lyssavirus. Microbiol Resour Announc. 2021;10(20).

Guichard P, Krell T, Chevalier M, Vaysee C, Adam O, Ronzon F, Marco S. Three dimensional morphology of rabies virus studied by cryo-electron tomography. J Struct Biol. 2011;176(1):32-40.

Gunesch JT, Dixon AL, Ebrahim TA, Berrien-Elliott MM, Tatineni S, Kumar T, *et al*. CD56 regulates human NK cell cytotoxicity through Pyk2. Elife. 2020;9.

Gupta R & Brunak S. Prediction of glycosylation across the human proteome and the correlation to protein function. Pac Symp Biocomput. 2002:310-22.

Hamamoto N, Uda A, Tobiume M, Park CH, Noguchi A, Kaku Y, *et al*. Association between RABV G proteins transported from the perinuclear space to the cell surface membrane and N-Glycosylation of the sequon Asn(204). Jpn J Infect Dis. 2015;68(5):387-93.

Hampson K, Coudeville L, Lembo T, Sambo M, Kieffer A, Attlan M, *et al*. Estimating the global burden of endemic canine rabies. PLoS Negl Trop Dis. 2015;9(4).

Han S, Lin YC, Wu T, Salgado AD, Mexhitaj I, Wuest SC, *et al*. Comprehensive immunophenotyping of cerebrospinal fluid cells in patients with neuroimmunological diseases. J Immunol. 2014;192(6):2551-63.

Hanlon CA, DeMattos CA, DeMattos CC, Niezgoda M, Hooper DC, Koprowski H, *et al*. Experimental utility of rabies virus-neutralizing human monoclonal antibodies in post-exposure prophylaxis. Vaccine. 2001;19(28-29):3834-42.

Hanlon CA, Kuzmin IV, Blanton JD, Weldon WC, Manangan JS, Rupprecht CE. Efficacy of rabies biologics against new lyssaviruses from Eurasia. Virus Res. 2005;111(1):44-54.

Hanna JN, Carney IK, Smith GA, Tannenberg AE, Deverill JE, Botha JA, *et al*. Australian bat lyssavirus infection: A second human case, with a long incubation period. Med J Aus. 2000;172(12):597-9.

Harrison AR, Lieu KG, Larrous F, Ito N, Bourhy H, Moseley GW. Lyssavirus P-protein selectively targets STAT3-STAT1 heterodimers to modulate cytokine signalling. PLoS Pathog. 2020;16(9).

Haywood AM. Membrane uncoating of intact enveloped viruses. J Virol. 2010;84(21):10946-55.

Hellert J, Buchrieser J, Larrous F, Minola A, de Melo GD, Soriaga L, *et al*. Structure of the prefusion-locking broadly neutralizing antibody RVC20 bound to the rabies virus glycoprotein. Nat Commun. 2020;11(1):596.

Hemachudha T, Wacharapluesadee S, Mitrabhakdi E, Wilde H, Morimoto K, Lewis RA. Pathophysiology of human paralytic rabies. J Neurovirol. 2005;11(1):93-100.

Higel F, Seidl A, Sörgel F, Friess W. N-glycosylation heterogeneity and the influence on structure, function and pharmacokinetics of monoclonal antibodies and Fc fusion proteins. Eur J Pharm Biopharm. 2016;100:94-100.

Hoffman S, Chuong CM, Edelman, G.M. Evolutionary conservation of key structures and binding functions of neural cell adhesion molecules. PNAS. 1984;81:6881-5.

Hofsteenge J, Blommers M, Hess D, Furmanek A, Miroshnichenko O. The four terminal components of the complement system are C-mannosylated on multiple tryptophan residues. J Biol Chem. 1999;274(46):32786-94.

Holloway BP & Obijeski JF. Rabies virus-induced RNA synthesis in Bhk-21 cells. J Gen Virol. 1980;49:181-95.

Honda K & Taniguchi T. IRFs: Master regulators of signalling by Toll-like receptors and cytosolic pattern-recognition receptors. Nat Rev Immunol. 2006;6(9):644-58.

Hooper PT, Lunt RA, Gould AR, Samaratunga H, Hyatt AD, Gleeson LJ, *et al*. A new lyssavirus-the first endemic rabies-related virus recognized in Australia. Bulletin du Institut Pasteur. 1997;95:209-18.

Horncastle ER. Investigation into the infectivity and neutralisation profiles of diverse lyssaviruses using a pseudotyped virus system. Dissertation. University of Nottingham. 2019.

Horowitz A, Behrens RH, Okell L, Fooks AR, Riley EM. NK cells as effectors of acquired immune responses: effector CD4+ T cell-dependent activation of NK cells following vaccination. J Immunol. 2010;185(5):2808-18.

Hosokawa-Muto J, Ito N, Yamada K, Shimizu K, Sugiyama M, Minamoto N. Characterization of recombinant rabies virus carrying double glycoprotein genes. Microbiol Immunol. 2006;50:187–96.

Hotta K, Motoi Y, Okutani A, Kaku Y, Noguchi A, Inoue S, Yamada A. Role of GPI-anchored NCAM-120 in rabies virus infection. Microbes Infect. 2007;9(2):167-74.

Hummler K, Koprowski H, Wiktor TJ. Structure and development of rabies virus in tissue culture. J Virol. 1967;1(1):152-70.

ICTV. Genus: *Lyssavirus*. [Internet]. 2019. [cited 2023 Nov 7]. Available from: https://ictv.global/report/chapter/rhabdoviridae/rhabdoviridae/lyssavirus.

Imai C, Iwamoto S, Campana D. Genetic modification of primary natural killer cells overcomes inhibitory signals and induces specific killing of leukemic cells. Blood. 2005;106(1):376-83.

Inoue Y, Kaku Y, Harada M, Ishijima K, Kuroda Y, Tatemoto K, *et al*. Establishment of serological neutralizing tests using pseudotyped viruses for comprehensive detection of antibodies against all 18 lyssaviruses. J Vet Med Sci. 2024;86(1):128-34.

Iseni F, Barge A, Baudin F, Blondel D, Ruigrok RW. Characterization of rabies virus nucleocapsids and recombinant nucleocapsid-like structures. J Gen Virol. 1998;79(12):2909-19.

Isobe Y, Sugimoto K, Yang L, Tamayose K, Egashira M, Kaneko T, *et al*. Epstein-Barr virus infection of human natural killer cell lines and peripheral blood natural killer cells. Cancer Res. 2004;64(6):2167-74.

Itakura Y, Tabata K, Morimoto K, Ito N, Chambaro HM, Eguchi R, *et al*. Glu_{333} in rabies virus glycoprotein is involved in virus attenuation through astrocyte infection and interferon responses. iScience. 2022;25(4):104122.

Jackson AC. Rabies. Handbook Clin Neurol. 2014;123:601-18.

Jamil KM, Ahmed K, Hossain M, Matsumoto T, Ali MA, Hossain S, *et al*. Arctic-like rabies virus, Bangladesh. Emerg Infect Dis. 2012;18(12):2021-4.

Johnson N, Aréchiga-Ceballos N, Aguilar-Setien A. Vampire bat rabies: ecology, epidemiology and control. Viruses. 2014;6(5):1911-28.

Johnson N, Cunningham AF, Fooks AR. The immune response to rabies virus infection and vaccination. Vaccine. 2010;28:3896-901.

Johnson N, Vos A, Freuling C, Tordo N, Fooks AR, Müller T. Human rabies due to lyssavirus infection of bat origin. Vet Microbiol. 2010;142(3-4):151-9.

Jordan I & Sandig V. Matrix and backstage: Cellular substances for viral vaccines. Viruses. 2014;6(4):1672-1700.

Julithe R, Abou-Jaoudé G, Sureau C. Modification of the hepatitis B virus envelope protein glycosylation pattern interferes with secretion of viral particles, infectivity, and susceptibility to neutralizing antibodies. J Virol. 2014;88(16):9049-59.

Kalimuddin S, Wijaya L, Chan YFZ, Wong AWL, Oh HML, Wang LF, *et al*. A phase II randomized study to determine the safety and immunogenicity of the novel PIKA rabies vaccine containing the PIKA adjuvant using an accelerated regimen. Vaccine. 2017;35:7127–32.

Kasturi L, Eshleman JR, Wunner WH, Shakin-Eshleman SH. The hydroxy amino acid in an Asn-X-Ser/Thr sequon can influence N-linked core glycosylation efficiency and the level of expression of a cell surface glycoprotein. J Biolog Chem. 1995;270(24):14756-61.

Kessels JA, Recuenco S, Navarro-Vela AM, Deray R, Vigilato M, Ertl H, *et al*. Pre-exposure rabies prophylaxis: a systematic review. Bull World Health Organ. 2017;95(3):210-9.

Kgaladi J, Wright N, Coertse J, Markotter W, Marston D, Fooks AR, *et al*. Diversity and epidemiology of Mokola Virus. PLoS Negl Trop Dis. 2013;7(10):2511.

Kissi B, Tordo N, Bourhy H. Genetic polymorphism in the rabies virus nucleoprotein gene. Virol. 1995;209:526–37.

Klein A, Eggerbauer E, Potratz M, Zaeck LM, Calvelage S, Finke S, *et al*. Comparative pathogenesis of different phylogroup I bat lyssaviruses in a standardized mouse model. PLoS Negl Trop Dis. 2022;16(1).

Knowles MK, Nadin-Davis SA, Sheen M, Rosatte R, Mueller R, Beresford A. Safety studies on an adenovirus recombinant vaccine for rabies (AdRG1.3-ONRAB®) in target and non-target species. Vaccine. 2009;27:6619–26.

Kotait I, Oliveira RN, Carrieri ML, Castilho JG, Macedo CI, Pereira PMC, *et al*. Non-human primates as a reservoir for Rabies virus in Brazil. Zoonoses Public Health. 2019;66:47-59.

Krapp S, Mimura Y, Jefferis R, Huber R, Sondermann P. Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity. J Mol Biol. 2003;325(5):979-89.

Kucera P, Dolivo M, Coulon P, Flamand A. Pathways of the early propagation of virulent and avirulent rabies strains from the eye to the brain. J. Virol. 1985;55:158–62.

Kuzmin IV, Niezgoda M, Franka R, Agwanda B, Markotter W, Beagley JC, *et al*. Possible emergence of West Caucasian bat virus in Africa. Emerg Infect Dis. 2008;14(12):1887-9.

Kuzmin IV, Shi M, Orciari LA, Yager PA, Velasco-Villa A, Kuzmina NA, *et al*. Molecular inferences suggest multiple host shifts of rabies viruses from bats to mesocarnivores in Arizona during 2001-2009. PLoS Pathog. 2012;8(6).

Kuzmin, Hughes, Botvinkin, Orciari & Rupprecht 2005: Kuzmin IV, Hughes GJ, Botvinkin AD, Orciari LA, Rupprecht CE. Phylogenetic relationships of Irkut and West Caucasian bat viruses within the Lyssavirus genus and suggested quantitative criteria based on the N gene sequence for lyssavirus genotype definition. Virus Res. 2005;111(1):28-43.

Lafon M, Bourhy H, Sureau P. Immunity against the European bat rabies (Duvenhage) virus induced by rabies vaccines: an experimental study in 122mice. Vaccine. 1988;6:362–8.

Lafon M, Wiktor TJ, Macfarlan RI. Antigenic sites on the CVS rabies virus glycoprotein: Analysis with monoclonal antibodies. J Gen Virol. 1983;64:843-51.

Lafon M. Rabies virus receptors. J Neurovirol. 2005;11:82–7.

Lee HS, Qi Y, Im W. Effects of N-glycosylation on protein conformation and dynamics: Protein Data Bank analysis and molecular dynamics simulation study. Sci Rep. 2015;5:8926.

Lentz TL, Burrage TG, Smith AL, Crick J, Tignor GH. Is the acetylcholine receptor a rabies virus receptor? Science. 1982;215(4529):182-4.

Lentz TL. Rabies virus binding to an acetylcholine receptor α-subunit peptide. J Mol Recog. 1990;3:82-8.

Lentz TL. Rabies virus receptors. Trends Neurol Sciences. 1985;8:360-4.

Leonard CK, Spellman MW, Riddle L, Harris RJ, Thomas JN, Gregory TJ. Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells. J Biol Chem. 1990;265(18):10373-82.

Leopardi S, Barneschi E, Manna G, Zecchin B, Priori P, Drzewnioková P, *et al*. Spillover of West Caucasian Bat Lyssavirus (WCBV) in a domestic cat and westward expansion in the palearctic region. Viruses. 2021;13(10):2064.

Leopardi S, Priori P, Zecchin B, Poglayen G, Trevisiol K, Lelli D, *et al*. Active and passive surveillance for bat lyssaviruses in Italy revealed serological evidence for their circulation in three bat species. Epidemiol Infect. 2018;147.

Li Y, Liu D, Wang Y, Su W, Liu G, Dong W. The importance of glycans of viral and host proteins in enveloped virus infection. Front Immunol. 2021;12:638573.

Liu P, Yang J, Wu X, Fu ZF. Interactions amongst rabies virus nucleoprotein, phosphoprotein and genomic RNA in virus-infected and transfected cells. J Gen Virol. 2004;85:3725–34.

Livingstone CD & Barton GJ. Protein sequence alignments: a strategy for the hierarchical analysis of residue conservation. Comput Appl Biosci. 1993;9(6):745-56.

Luo J, Zhang B, Wu Y, Guo X. Amino acid mutation in position 349 of glycoprotein affect the pathogenicity of rabies virus. Front Microbiol. 2020;11(481).

Lycke E & Tsiang H. Rabies virus infection of cultured rat sensory neurons. J. Virol. 1987;61:2733–41.

Ma B, He LF, Zhang YL, Chen M, Wang LL, Yang HW, *et al*. Characteristics and viral propagation properties of a new human diploid cell line, Walvax-2, and its suitability as a candidate cell substrate for vaccine production. Human Vaccines Immunotherap. 2015;11(4):998-1009.

Macdonald D. *Encyclopedia of Mammals*. New York: Facts on File Publications, 1984.

Macedo CI, Carnieli P Jr, Brandão PE, Travassos da Rosa ES, Oliveira Rde N, Castilho JG, *et al*. Diagnosis of human rabies cases by polymerase chain reaction of neck-skin samples. Braz J Infect Dis. 2006;10(5):341-5.

Macfarlan RI, Dietzschold B, Koprowski H. Stimulation of cytotoxic T-lymphocyte responses by rabies virus glycoprotein and identification of an immunodominant domain. Molecul Immunol. 1986;23(7):733-41.

Madhusudana SN & Mani RS. Intradermal vaccination for rabies prophylaxis: Conceptualization, evolution, present status and future. Expert Rev Vaccines. 2014;13(5):641-55.

Mähl P, Cliquet F, Guiot A-L, Niin E, Fournials E, Saint-Jean N, *et al*. Twenty year experience of the oral rabies vaccine SAG2 in wildlife: A global review. Vet Res. 2014;45.

Maki J, Guiot AL, Aubert M, Brochier B, Cliquet F, Hanlon CA, et al. Oral vaccination of wildlife using a vaccinia-rabies-glycoprotein recombinant virus vaccine (RABORAL V-RG®): A global review. Vet Res. BioMed Central. 2017;48:1–26.

Malerczyk C, Freuling C, Gniel D, Giesen A, Selhorst T, Müller T. Cross-neutralization of antibodies induced by vaccination with Purified Chick Embryo Cell Vaccine (PCECV) against different Lyssavirus species. Hum Vaccin Immunother. 2014;10(10):2799-804.

Mantesso S, Geerts D, Spanholtz J, Kučerová L. Genetic engineering of natural killer cells for enhanced antitumor function. Front Immunol. 2020;11:607131.

Mao H, Tu W, Qin G, Law HK, Sia SF, Chan PL, *et al*. Influenza virus directly infects human natural killer cells and induces cell apoptosis. J Virol. 2009;83(18):9215-22.

Marissen WE, Kramer RA, Rice A, Weldon WC, Niezgoda M, Faber M, *et al*. Novel rabies Virus-neutralizing epitope recognized by human monoclonal antibody: Fine mapping and escape mutant analysis. J Virol. 2005;79:4672–8.

Marston DA, Banyard AC, McElhinney LM, Freuling CM, Finke S, de Lamballerie X, *et al*. The lyssavirus host-specificity conundrum-rabies virus-the exception not the rule. Curr Opin Virol. 2018c;28:68-73.

Marston DA, Ellis RJ, Horton DL, Kuzmin IV, Wise EL, McElhinney LM, *et al*. Complete genome sequence of Ikoma lyssavirus. J Virol. 2012a;86(18):10242-3.

Marston DA, Horton DL, Ngeleja C, Hampson K, McElhinney LM, Banyard AC, *et al*. Ikoma Lyssavirus, highly divergent novel Lyssavirus in an African Civet. Emerg Infect Dis. 2012b;18(4):664-7.

Masatani T, Ito N, Shimizu K, Ito Y, Nakagawa K, Sawaki Y, *et al*. Rabies virus nucleoprotein functions to evade activation of the RIG-I-mediated antiviral response. J Virol. 2010;84:4002–12.

Mastraccio KE, Huaman C, Coggins SA, Clouse C, Rader M, Yan L, *et al*. mAb therapy controls CNS-resident lyssavirus infection via a CD4 T cell-dependent mechanism. EMBO Mol Med. 2023;15(10).

Mathys L, François KO, Quandte M, Braakman I, Balzarini J. Deletion of the highly conserved N-glycan at Asn260 of HIV-1 gp120 affects folding and lysosomal degradation of gp120, and results in loss of viral infectivity. PLoS One. 2014;9(6).

Matusali G, Tchidjou HK, Pontrelli G, Bernardi S, D'Ettorre G, Vullo V, *et al*. Soluble ligands for the NKG2D receptor are released during HIV-1 infection and impair NKG2D expression and cytotoxicity of NK cells. FASEB J. 2013;27(6):2440-50.

Mavrakis M, Iseni F, Mazza C, Schoehn, G, Ebel C, Gentzel M, *et al*. Isolation and characterisation of the rabies virus N degrees-P complex produced in insect cells. Virol. 2003;305(2):406-14.

Mebatsion T, Konig M, Conzelmann KK. Budding of rabies virus particles in the absence of the spike glycoprotein. Cell. 1996;84(6):941-51.

Mebatsion T, Weiland F, Conzelmann KK. Matrix protein of rabies virus is responsible for the assembly and budding of bullet-shaped particles and interacts with the transmembrane spike glycoprotein G. J Virol. 1999;73(1):242-50.

Mellquist JL, Kasturi L, Spitalnik SL, Shakin-Eshleman SH. The amino acid following an asn-X-Ser/Thr sequon is an important determinant of N-linked core glycosylation efficiency. Biochemistry. 1998;37(19):6833-7.

Meng EC, Goddard TD, Pettersen EF, Couch GS, Pearson ZJ, Morris JH, Ferrin TE. UCSF ChimeraX: Tools for structure building and analysis. Protein Sci. 2023;32(11).

Millar NS & Gotti C. Diversity of vertebrate nicotinic acetylcholine receptors. Neuropharmacology. Elsevier Ltd; 2009;56:237–46.

Millar NS. Assembly and subunit diversity of nicotinic acetylcholine receptors. Biochem Soc Trans. 2003;31:869–74.

Miller JS. Biology of natural killer cells in cancer and infection. Cancer Invest. 2002;20(3):405-19.

Moore SS, Hale P, Byrne K. NCAM: A polymorphic microsatellite locus conserved across eutherian mammal species. Ani Genet. 1998;29:33-3.

Morimoto K, Hooper DC, Bornhorst A, Corisdeo A, Bette M, Fu ZF, *et al*. Intrinsic responses to Borna Disease Virus infection of the central nervous system. PNAS. 1996;93(23):13345-50.

Morimoto K, Hooper DC, Carbaugh H, Fu ZF, Koprowski H, Dietzschold B. Rabies virus quasispecies: implications for pathogenesis. Proc Natl Acad Sci USA. 1998;95(6):3152-6.

Moussawi K & Kalivas PW. Group II metabotropic glutamate receptors (mGlu2/3) in drug addiction. Eur J Pharmacol. 2010;639:115–22.

Murphy FA, Bauer SP, Harrison AK, Winn WC Jr. Comparative pathogenesis of rabies and rabies-like viruses. Viral infection and transit from inoculation site to the central nervous system. Lab Invest. 1973;28(3):361-76.

Napolitano F, Merone R, Abbate A, Ammendola V, Horncastle E, Lanzaro F, *et al*. A next generation vaccine against human rabies based on a single dose of a chimpanzee adenovirus vector serotype C. PLoS Negl Trop Dis. 2020;14(7).

New South Wales Department of Primary Industries. A dog tests antibody positive for lyssavirus. Ani Health Surv. 2013;3(3-4).

Ng WM, Fedosyuk S, English S, Augusto G, Berg A, Thorley L, *et al*. Structure of trimeric pre-fusion rabies virus glycoprotein in complex with two protective antibodies. Cell Host Microbe. 2022;30(9):1219-30.

Nita-Lazar M, Wacker M, Schegg B, Amber S, Aebi M. The N-X-S/T consensus sequence is required but not sufficient for bacterial N-linked protein glycosylation. Glycobiology. 2005;15(4):361-7.

Nitschel S, Zaeck LM, Potratz M, Nolden T, Te Kamp V, Franzke K, *et al*. Point mutations in the glycoprotein ectodomain of field rabies viruses mediate cell culture adaptation through improved virus release in a host cell dependent and independent manner. Viruses. 2021;13(10):1989.

Nokireki T, Tammiranta N, Kokkonen UM, Kantala T, Gadd T. Tentative novel lyssavirus in a bat in Finland. Transbound Emerg Dis. 2018;65(3):593-6.

O'Brien KL, Nolan T; SAGE WG on Rabies. The WHO position on rabies immunization - 2018 updates. Vaccine. 2019;37(Suppl 1):A85-A87.

Ortega V, Stone JA, Contreras EM, Iorio RM, Aguilar HC. Addicted to sugar: roles of glycans in the order Mononegavirales. Glycobiology. 2019;29(1):2-21.

Osada N, Kohara A, Yamaji T, Hirayama N, Kasai F, Sekizuka T, *et al*. The genome landscape of the African green monkey kidney-derived vero cell line. DNA Res. 2014;21:673-83.

Panpanich T, Hemachudha T, Piyasirisilp S, Manatsathit S, Wilde H, Phanuphak P. Cells with natural killer activity in human rabies. Clin Exp Immunol. 1992;89(3):414-8.

Pearce JMS. Louis Pasteur and rabies: a brief note. J Neurol Neurosurg Psychiatry. 2002;73.

Pérez V, Bermedo-Garcia F, Zelada D, Court FA, Pérez MÁ, Fuenzalida M, *et al*. The p75NTR neurotrophin receptor is required to organize the mature neuromuscular synapse by regulating synaptic vesicle availability. Acta Neuropathol Commun. 2019;7(1):147.

Perry LL, Hotchkiss JD, Lodmell DL. Murine susceptibility to street rabies virus is unrelated to induction of host lymphoid depletion. J Immunol. 1990;144:3553-7.

Piccinotti S & Whelan SPJ. Rabies internalizes into primary peripheral neurons via clathrin coated pits and requires fusion at the cell body. PLoS Pathogens. 2016;12(7).

Poli A, Michel T, Thérésine M, Andrès E, Hentges F, Zimmer J. CD56bright natural killer (NK) cells: an important NK cell subset. Immunol. 2009;126(4):458-65.

Prehaud C, Coulon P, Lafay F, Thiers C, Flamand A. Antigenic site II of the rabies virus glycoprotein: Structure and role in viral virulence. J Virol. 1988;62(1):1-7.

Prentoe J, Velázquez-Moctezuma R, Augestad EH, Galli A, Wang R, Law M, *et al*. Hypervariable region 1 and N-linked glycans of hepatitis C regulate virion neutralization by modulating envelope conformations. Proc Natl Acad Sci USA. 2019;116(20):10039-47.

Public Health England. Rabies: epidemiology, transmission and prevention. [Internet]. 2019. [cited 2023 Nov 7]. Available from: https://www.gov.uk/guidance/rabies-epidemiology-transmission-and-prevention.

Rancan C, Schirrmann L, Hüls C, Zeidler R, Moosmann A. Latent membrane protein LMP2A impairs recognition of EBV-infected cells by CD8+ T Cells. PLoS Pathog. 2015;11(6).

Rathore U, Saha P, Kesavardhana S, Kumar AA, Datta R, Devanarayanan S, *et al*. Glycosylation of the core of the HIV-1 envelope subunit protein gp120 is not required for native trimer formation or viral infectivity. J Biol Chem. 2017;292(24):10197-219.

Rincón V, Corredor A, Martínez-Gutíerrez M, Castellanos JE. Fluorometric cell-ELISA for quantifying rabies infection and heparin inhibition. J Virol Methods. 2005;127(1):33-9.

Rocchi C & Emmerson E. Mouth-watering results: Clinical need, current approaches, and future directions for salivary gland regeneration. Trends Mol Med. 2020;26(7):649-69.

Roller RF, Malik A, Carillo MA, Garg M, Rella A, Raulf MK, *et al*. Semisynthesis of functional glycosylphosphatidylinositol-anchored proteins. Angew Chem Int Ed Engl. 2020;59(29):12035-40.

Rolls MM, Webster P, Balba NH, Rose JK. Novel infectious particles generated by expression of the vesicular stomatitis virus glycoprotein from a self-replicating RNA. Cell. 1994;79(3):497-506

Rupprecht C, Kuzmin I, Meslin F. Lyssaviruses and rabies: current conundrums, concerns, contradictions and controversies. F1000 Res. 2017;6:184.

Rupprecht CE, Briggs D, Brown CM, Franka R, Katz SL, Kerr HD, *et al*. Use of a reduced (4-dose) vaccine schedule for postexposure prophylaxis to prevent human rabies: Recommendations of the advisory committee on immunization practices. MMWR Recomm Reports. 2010;59:1-9.

Rupprecht CE, Turmelle A, Kuzmin IV. A perspective on lyssavirus emergence and perpetuation. Curr Opin Virol. 2011;1(6):662-70.

Sadeghi M, Moallem SA, Yousefi-Abdolmaleki E, Montazeri M. The rabies early death phenomenon: A report of ineffective administration of rabies vaccine during symptomatic disease. Indian J Crit Care Med. 2015;19(7):422-4.

Sanofi Pasteur SA. Imovax Rabies. 2019. [cited 2023 Nov 8]. Available from: https://www.fda.gov/media/75709/download.

Schmidt P, Raftery MJ, Pecher G. Engineering NK cells for CAR therapy - recent advances in gene transfer methodology. Front Immunol. 2021;11:611163.

Schnell MJ, McGettigan JP, Wirblich C, Papaneri A. The cell biology of rabies virus: using stealth to reach the brain. Nat Rev Microbiol. 2010;8(1):51-61.

Seif I, Coulon P, Rollin PE, Flamand A. Rabies virulence: Effect on pathogenicity and sequence characterization of rabies virus mutations affecting antigenic site III of the glycoprotein. J Virol. 1985;53(3):926-34.

Shakin-Eshleman SH, Remaley AT, Eshleman JR, Wunner WH, Spitalnik SL. N-linked glycosylation of rabies virus glycoprotein. J Biolog Chem. 1992a;267(15):10690-8.

Shakin-Eshleman SH, Remaley AT, Eshleman JR, Wunner WH, Spitalnik SL. N-linked glycosylation of rabies virus glycoprotein. Individual sequons differ in their glycosylation efficiencies and influence on cell surface expression. J Biol Chem. 1992b;267(15):10690-8.

Shakin-Eshleman SH, Spitalnik SL, Kasturi L. The amino acid at the X position of an Asn-X-Ser sequon is an important determinant of N-linked core-glycosylation efficiency. J Biol Chem. 1996;271(11):6363-6.

Sheard PW, Musaad K, Duxson MJ. Distribution of neurotrophin receptors in the mouse neuromuscular system. Int J Dev Biol. 2002;46:569–75.

Shi FD & Ransohoff RM. Nature killer cells in the central nervous system. Natural Killer Cells. 2010:373–83.

Shinwari MW, Annand EJ, Driver L, Warrilow D, Harrower B, Allcock RJ, *et al*. Australian bat lyssavirus infection in two horses. Vet Microbiol. 2014;173(3-4):224-31.

Shipley R, Wright E, Lean FZX, Selden D, Horton DL, Fooks AR, Banyard AC. Assessing Rabies vaccine protection against a novel Lyssavirus, Kotalahti Bat Lyssavirus. Viruses. 2021;13(5):947.

Shorey-Kendrick LE, Ford MM, Allen DC, Kuryatov A, Lindstrom, J, Wilhelm L, *et al*. Nicotinic receptors in non-human primates: Analysis of genetic and functional conservation with humans. Neuropharmacology. 2015;96:263-73.

Smyth MJ, Cretney E, Kelly JM, Westwood JA, Street SE, Yagita H, *et al*. Activation of NK cell cytotoxicity. Mol Immunol. 2005;42(4):501-10.

Song Y, DiMaio F, Wang RYR, Kim D, Miles C, Brunette TJ, *et al*. High resolution comparative modeling with RosettaCM. Structure. 2013;21(10):1735-42.

Sonthonnax F, Besson B, Bonnaud E, Jouvion G, Merino D, Larrous F, Bourhy H. Lyssavirus matrix protein cooperates with phosphoprotein to modulate the Jak-Stat pathway. Sci Rep. 2019;9(1):12171.

Sparrer MN, Hodges NF, Sherman T, VandeWoude S, Bosco-Lauth AM, Mayo CE. Role of spillover and spillback in SARS-CoV-2 transmission and the importance of one health in understanding the dynamics of the COVID-19 pandemic. J Clin Microbiol. 2023;61(7).

Stanley P, Schachter H, Taniguchi N. Essentials of Glycobiology. 2nd ed. Chapter 8: N-Glycans. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2009.

Stewart-Jones GB, Soto C, Lemmin T, Chuang GY, Druz A, Kong R, *et al*. Trimeric HIV-1-Env structures define glycan shields from clades A, B, and G. Cell. 2016;165(4):813-26.

Suraweera et al 2012: Suraweera W, Morris SK, Kumar R, Warrell DA, Warrell MJ, Jha P. Deaths from symptomatically identifiable furious rabies in India: A nationally representative mortality survey. PLoS Neglec Trop Dis. 2012;6(10).

Sutlu T, Nyström S, Gilljam M, Stellan B, Applequist SE, Alici E. Inhibition of intracellular antiviral defense mechanisms augments lentiviral transduction of human natural killer cells: implications for gene therapy. Hum Gene Ther. 2012;23(10):1090-100.

Sutor A, Schwarz S, Conraths FJ. The biological potential of the raccoon dog (*Nyctereutes procyonoides*, Gray 1834) as an invasive species in Europe—new risks for disease spread? Acta Theriol. 2014;59:49–59.

Syapin PJ, Salvaterra PM, Engelhardt JK. Neuronal-like features of TE671 cells: presence of a functional nicotinic cholinergic receptor. Brain Res. 1982;231(2):365-77.

Tao L, Ge J, Wang X, Zhai H, Hua T, Zhao B, *et al*. Molecular basis of neurovirulence of flury rabies virus vaccine strains: importance of the polymerase and the glycoprotein R333Q mutation. J Virol. 2010;84(17):8926-36.

Tatsis N, Fitzgerald JC, Reyes-Sandoval A, Harris-McCoy KC, Hensley SE, Zhou D, *et al*. Adenoviral vectors persist in vivo and maintain activated CD8+ T cells: implications for their use as vaccines. Blood. 2007;110(6):1916-23.

Thoulouze MI, Lafage M, Schachner M, Hartmann U, Cremer H, Lafon M. The neural cell adhesion molecule is a receptor for rabies virus. J Virol. 1998;72(9):7181-90.

Torres-Anjel MJ, Volz D, Torres MJ, Turk M, Tshikuka, J.G. Failure to thrive, wasting syndrome, and immunodeficiency in rabies: A hypophyseal/thymic axis effect of rabies virus. Rev Infect Dis. 1988;10:710-25.

Trabelsi K, Ben Zakour M, Kallel H. Purification of rabies virus produced in Vero cells grown in serum free medium. Vaccine. 2019;37(47):7052-7060.

Tran J & Kung SK. Lentiviral vectors mediate stable and efficient gene delivery into primary murine natural killer cells. Mol Ther. 2007;15(7):1331-9.

Troupin C, Dacheux L, Tanguy M, Sabeta C, Blanc H, Bouchier C, *et al*. Large-scale phylogenomic analysis reveals the complex evolutionary history of rabies virus in multiple carnivore hosts. PLoS Pathog. 2016;12(12).

Tsiang H, Ceccaldi PE, Lycke E. Rabies virus infection and transport in human sensory dorsal root ganglia neurons. J Gen Virol. 1991;72(5):1191-4.

Tsiang H, Delaporte S, Ambroise DJ, Derer M, Koenig J. Infection of cultured rat myotubes and neurons from the spinal-cord by rabies virus. J Neuropathol and Exp Neurol. 1986;45(1):28-42.

Tsiang H. Evidence for an intraaxonal transport of fixed and street rabies virus. J Neuropathol Exper Neurol. 1979;38:286-96.

Tuffereau C, Benejean J, Blondel D, Kieffer B, Flamand A. Low-affinity nerve-growth factor receptor (P75NTR) can serve as a receptor for rabies virus. Eur Mol Biol Organ J. 1998;17(24):7250-9.

Tuffereau C, Schmidt K, Langevin C, Lafay F, Dechant G, Koltzenburg M. The rabies virus glycoprotein receptor p75NTR is not essential for rabies virus infection. J Virol. 2007;81:13622–30.

Udow SJ, Marrie RA, Jackson AC. Clinical features of dog- and bat-acquired rabies in humans. Clin Infect Dis. 2013;57(5):689-96

Ugolini G. Rabies virus as a transneuronal tracer of neuronal connections. Adv Virus Res. 2011;79:165-202.

Vallbracht M, Rehwaldt S, Klupp BG, Mettenleiter TC, Fuchs W. Functional role of N-linked glycosylation in pseudorabies virus glycoprotein gH. J Virol. 2018;92(9).

Van den Steen P, Rudd PM, Dwek RA, Opdenakker G. Concepts and principles of O-linked glycosylation. Crit Rev Biochem Mol Biol. 1998;33(3):151-208.

van Erp EA, van Kampen MR, van Kasteren PB, de Wit J. Viral infection of human natural killer cells. Viruses. 2019;11(3):243.

Van Zyl N, Markotter W, Nel LH. Evolutionary history of African mongoose rabies. Virus Res. 2010;150(1-2):93-102.

Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, *et al*, editors. Essentials of Glycobiology. 4th ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2022.

Varki A. Evolutionary forces shaping the Golgi glycosylation machinery: why cell surface glycans are universal to living cells. Cold Spring Harb Perspect Biol. 2011;3(6).

Vellinga J, Smith JP, Lipiec A, Majhen D, Lemckert A, van Ooij M, *et al*. Challenges in manufacturing adenoviral vectors for global vaccine product deployment. Hum Gene Ther. 2014;25(4):318-27.

Vidy A, Bougrini JE, Chelbi-Alix M, Blondel D. The nucleocytoplasmic rabies virus P protein counteracts interferon signalling by inhibiting both nuclear accumulation and DNA binding to STAT1. J Virol. 2007;81:4255-63.

Vidy,A, Chelbi-Alix M, Blondel D. Rabies virus P protein interacts with STAT1 and inhibits interferon signal transduction pathways. J Virol. 2005;79:14411-20.

Vigerust DJ & Shepherd VL. Virus glycosylation: role in virulence and immune interactions. Trends Microbiol. 2007;15(5):211-8.

Vitelli A, Folgori A, Scarselli E, Colloca S, Capone S, Nicosia A. Chimpanzee adenoviral vectors as vaccines - challenges to move the technology into the fast lane. Expert Rev Vaccines. 2017;16(12):1241-52.

Wang J, Wang Z, Liu R, Shuai L, Wang X, Luo J, *et al*. Metabotropic glutamate receptor subtype 2 is a cellular receptor for rabies virus. PLoS Pathogens. 2018;14(7).

Wang W, Ma J, Nie J, Li J, Cao S, Wang L, *et al*. Antigenic variations of recent street rabies virus. Emerg Microbes Infect. 2019;8(1):1584-92.

Wang X, Terrie L, Wu G, Van Damme EJM, Thorrez L, Fooks AR, *et al*. *Urtica dioica* agglutinin prevents rabies virus infection in a muscle explant model. Pharmaceutics. 2023;15(5):1353.

Warrell DA. The clinical picture of rabies in man. Transactions Royal Society Trop Med Hyg.1976;70:188-95.

Warrell MJ & Warrell DA. Rabies: The clinical features, management and prevention of the classic zoonosis. Clin Med. 2015;15(1):78-81.

Wei X, Decker JM, Wang S, Hui H, Kappes JC, Wu X, *et al*. Antibody neutralization and escape by HIV-1. Nature. 2003;422(6929):307-12.

Whitley, RJ. 2001. Fields virology, 4th ed. Lippincott Williams & Wilkins, Philadelphia, Pa; 2001: Herpes simplex viruses, p. 2461-2509.

WHO, FAO & OIE. Zero by 30. The global strategic plan to end human deaths from dog-mediated rabies by 2030. [Internet]. 2018. [cited 2023 Nov 8]. Available from: https://www.woah.org/fileadmin/Home/eng/Media_Center/docs/Zero_by_30_FINAL_online_version.pdf.

WHO. Distribution of risk levels for human contracting, rabies, worldwide, 2009. [Internet]. 2010. [cited 2023 Nov 7]. Available from: https://www.who.int/data/gho/map-gallery-search-results?&maptopics=910b5dfc-ce2e-4440-8b43-8d83f4a85485&term=Rabies.

WHO. Epidemiology of Rabies. [Internet]. [cited 2024 Nov 21]. Available from: https://www.who-rabies-bulletin.org/site-page/epidemiology-rabies.

WHO. Rabies. [Internet]. [cited 2023 Nov 7]. Available from: https://www.who.int/health-topics/rabies#tab=tab_1.

WHO. Rabies. [Internet]. 2023. [cited 2023 Nov 7]. Available from: https://www.who.int/news-room/fact-sheets/detail/rabies.

WHO. Verorab. [Internet]. 2005. [cited 2023 Nov 8]. Available from: https://web.archive.org/web/20221124032320/https://extranet.who.int/pqweb/content/verorab.

WHO. WHO expert consultation on Rabies. Third report WHO technical report series. 2018. No. 1012. Vol. 2018.Geneva: WHO. p183.

WHO. WHO Expert Consultation on Rabies: Second Report. World Health Organization Technical Report Series. 2013.

Wieten RW, Leenstra T, van Thiel PP, van Vugt M, Stijnis C, Goorhuis A, Grobusch MP. Rabies vaccination: Are abbreviated intradermal schedules the future? Clin Infect Dis. 2013;56(3):414-9.

Wiktor TJ, Plotkin SA, Koprowski H. Development and clinical trials of the new human rabies vaccine of tissue culture (human diploid cell) origin. Develop Biol Stand. 1978;40:3-9.

Wilde H, Khawplod P, Hemachudha T, Sitprija V. Postexposure treatment of rabies infection: Can it be done without immunoglobulin? Clin Infect Dis. 2002;34:477–80.

Wojczyk BS, Takahashi N, Levy MT, Andrews DW, Abrams WR, Wunner WH, Spitalnik SL. N-glycosylation at one rabies virus glycoprotein sequon influences N-glycan processing at a distant sequon on the same molecule. Glycobiol. 2005;15(6):655-66.

Wright E, Hayman DT, Vaughan A, Temperton NJ, Wood JL *et al*. Virus neutralising activity of African fruit bat (Eidolon helvum) sera against emerging lyssaviruses. Virol. 2010;408:183-9.

Wunner WH, Dietzschold B, Smith CL, Lafon M, Golub E. Antigenic variants of CVS rabies virus with altered glycosylation sites. Virology. 1985;140(1):1-12.

Wunner WH, Larson JK, Dietzschold B, Smith CL. The molecular biology of rabies viruses. Rev Infect Dis. 1988;10:S771–84.

Xu H, Hao X, Wang S, Wang Z, Cai M, Jiang J, *et al*. Real-time imaging of rabies virus entry into living vero cells. Sci Rep. 2015;5:11753.

Yamada K, Noguchi K, Nishizono A. Efficient N-glycosylation at position 37, but not at position 146, in the street rabies virus glycoprotein reduces pathogenicity. Virus Res. 2014;179:169-76.

Yamada K, Noguchi K, Nonaka D, Morita M, Yasuda A, Kawazato H, Nishizono A. Addition of a single N-glycan to street rabies virus glycoprotein enhances virus production. J Gen Virol. 2013;94:270-5.

Yamada K, Park CH, Noguchi K, Kojima D, Kubo T, Komiya N, *et al*. Serial passage of a street rabies virus in mouse neuroblastoma cells resulted in attenuation: potential role of the additional N-glycosylation of a viral glycoprotein in the reduced pathogenicity of street rabies virus. Virus Res. 2012;165(1):34-45.

Yang F, Lin S, Ye F, Yang J, Qi J, Chen Z, *et al*. Structural analysis of rabies virus glycoprotein reveals pH-dependent conformational changes and interactions with a neutralizing antibody. Cell Host Microbe. 2020;27(3):441-53.

Yang J, Anishchenko I, Park H, Peng Z, Ovchinnikov S, Baker D. Improved protein structure prediction using predicted interresidue orientations. PNAS. 2020;117(3):1496-1503.

Yang J, Koprowski H, Dietzschold B, Fu ZF. Phosphorylation of rabies virus nucleoprotein regulates viral RNA transcription and replication by modulating leader RNA encapsidation. J Virol. 1999;73:1661–4.

Yi X, Verbeke EJ, Chang Y, Dickinson DJ, Taylor DW. Electron microscopy snapshots of single particles from single cells. J Biol Chem. 2019;294(5):1602-8.

Yoshioka T, Uzuka Y, Tanabe S, Sarashina T, Ishiguro N. Molecular cloning of the canine nicotinic acetylcholine receptor α-subunit gene and development of the ELISA method to diagnose myasthenia gravis. Vet Immunol Immunopathol. 1999;72:315-24.

Zhang J, Lin J, Tian Y, Ma L, Sun W, Zhang L, *et al*. Transmission of rabies through solid organ transplantation: a notable problem in China. BMC Infect Dis. 2018;18(1):273.

Supplementary data

Table S.1: A table of significance for figure 2.1

Tukey's Multiple Comparisons	Significance
VSV vs. ΔE	****
VSV vs. CVS-11	****
VSV vs. EBLV-1	****
VSV vs. WCBV	ns
VSV vs. IKOV	****
VSV vs. MOKV	****
ΔE vs. CVS-11	****
ΔE vs. EBLV-1	ns
ΔE vs. WCBV	****
ΔE vs. IKOV	ns
ΔE vs. MOKV	ns
CVS-11 vs. EBLV-1	****
CVS-11 vs. WCBV	****
CVS-11 vs. IKOV	****
CVS-11 vs. MOKV	****
EBLV-1 vs. WCBV	****
EBLV-1 vs. IKOV	ns
EBLV-1 vs. MOKV	ns
WCBV vs. IKOV	****
WCBV vs. MOKV	****
IKOV vs. MOKV	ns

Table S.2: A table of significance for figure 2.3 C

Tukey's Multiple Comparisons	Significance
VSV vs. ΔE	****
VSV vs. CVS-11	****
VSV vs. EBLV-1	****
VSV vs. WCBV	ns
VSV vs. IKOV	****
VSV vs. MOKV	****
ΔE vs. CVS-11	****
ΔE vs. EBLV-1	**
ΔE vs. WCBV	****
ΔE vs. IKOV	ns
ΔE vs. MOKV	****
CVS-11 vs. EBLV-1	****
CVS-11 vs. WCBV	****
CVS-11 vs. IKOV	****

CVS-11 vs. MOKV	****
EBLV-1 vs. WCBV	****
EBLV-1 vs. IKOV	ns
EBLV-1 vs. MOKV	*
WCBV vs. IKOV	****
WCBV vs. MOKV	****
IKOV vs. MOKV	**

Table S.3: A table of significance for figure 4.1 A

Tukey's multiple comparisons test	**Significance**
T/S>A -37 vs. T/S>A -204	****
T/S>A -37 vs. T/S>A -319	****
T/S>A -37 vs. T/S>A -37 -204	****
T/S>A -37 vs. T/S>A -37 -319	****
T/S>A -37 vs. T/S>A -204 -319	****
T/S>A -37 vs. T/S>A FKO	****
T/S>A -37 vs. N>D -37	****
T/S>A -37 vs. N>D -204	****
T/S>A -37 vs. N>D -319	****
T/S>A -37 vs. N>D -37 -204	****
T/S>A -37 vs. N>D -37 -319	****
T/S>A -37 vs. N>D -204 -319	****
T/S>A -37 vs. N>D FKO	****
T/S>A -37 vs. N>Q -37	****
T/S>A -37 vs. N>Q -204	****
T/S>A -37 vs. N>Q -319	****
T/S>A -37 vs. N>Q -37 -204	****
T/S>A -37 vs. N>Q -37 -319	****
T/S>A -37 vs. N>Q -204 -319	****
T/S>A -37 vs. N>Q FKO	****
T/S>A -37 vs. CVS-11	****
T/S>A -37 vs. NT HEK293t	****
T/S>A -37 vs. ΔE	****
T/S>A -37 vs. VSV	****
T/S>A -204 vs. T/S>A -319	****
T/S>A -204 vs. T/S>A -37 -204	****
T/S>A -204 vs. T/S>A -37 -319	****
T/S>A -204 vs. T/S>A -204 -319	****
T/S>A -204 vs. T/S>A FKO	****
T/S>A -204 vs. N>D -37	****
T/S>A -204 vs. N>D -204	**

T/S>A -204 vs. N>D -319	****
T/S>A -204 vs. N>D -37 -204	****
T/S>A -204 vs. N>D -37 -319	****
T/S>A -204 vs. N>D -204 -319	****
T/S>A -204 vs. N>D FKO	****
T/S>A -204 vs. N>Q -37	****
T/S>A -204 vs. N>Q -204	****
T/S>A -204 vs. N>Q -319	****
T/S>A -204 vs. N>Q -37 -204	****
T/S>A -204 vs. N>Q -37 -319	****
T/S>A -204 vs. N>Q -204 -319	****
T/S>A -204 vs. N>Q FKO	****
T/S>A -204 vs. CVS-11	*
T/S>A -204 vs. NT HEK293t	****
T/S>A -204 vs. ΔE	****
T/S>A -204 vs. VSV	****
T/S>A -319 vs. T/S>A -37 -204	ns
T/S>A -319 vs. T/S>A -37 -319	ns
T/S>A -319 vs. T/S>A -204 -319	ns
T/S>A -319 vs. T/S>A FKO	ns
T/S>A -319 vs. N>D -37	****
T/S>A -319 vs. N>D -204	****
T/S>A -319 vs. N>D -319	ns
T/S>A -319 vs. N>D -37 -204	****
T/S>A -319 vs. N>D -37 -319	ns
T/S>A -319 vs. N>D -204 -319	ns
T/S>A -319 vs. N>D FKO	ns
T/S>A -319 vs. N>Q -37	****
T/S>A -319 vs. N>Q -204	****
T/S>A -319 vs. N>Q -319	ns
T/S>A -319 vs. N>Q -37 -204	ns
T/S>A -319 vs. N>Q -37 -319	ns
T/S>A -319 vs. N>Q -204 -319	ns
T/S>A -319 vs. N>Q FKO	ns
T/S>A -319 vs. CVS-11	****
T/S>A -319 vs. NT HEK293t	ns
T/S>A -319 vs. ΔE	ns
T/S>A -319 vs. VSV	****
T/S>A -37 -204 vs. T/S>A -37 -319	ns
T/S>A -37 -204 vs. T/S>A -204 -319	ns
T/S>A -37 -204 vs. T/S>A FKO	ns
T/S>A -37 -204 vs. N>D -37	****
T/S>A -37 -204 vs. N>D -204	****
T/S>A -37 -204 vs. N>D -319	ns
T/S>A -37 -204 vs. N>D -37 -204	***

T/S>A -37 -204 vs. N>D -37 -319	ns
T/S>A -37 -204 vs. N>D -204 -319	ns
T/S>A -37 -204 vs. N>D FKO	ns
T/S>A -37 -204 vs. N>Q -37	****
T/S>A -37 -204 vs. N>Q -204	***
T/S>A -37 -204 vs. N>Q -319	ns
T/S>A -37 -204 vs. N>Q -37 -204	ns
T/S>A -37 -204 vs. N>Q -37 -319	ns
T/S>A -37 -204 vs. N>Q -204 -319	ns
T/S>A -37 -204 vs. N>Q FKO	ns
T/S>A -37 -204 vs. CVS-11	****
T/S>A -37 -204 vs. NT HEK293t	ns
T/S>A -37 -204 vs. ΔE	ns
T/S>A -37 -204 vs. VSV	****
T/S>A -37 -319 vs. T/S>A -204 -319	ns
T/S>A -37 -319 vs. T/S>A FKO	ns
T/S>A -37 -319 vs. N>D -37	****
T/S>A -37 -319 vs. N>D -204	****
T/S>A -37 -319 vs. N>D -319	ns
T/S>A -37 -319 vs. N>D -37 -204	****
T/S>A -37 -319 vs. N>D -37 -319	ns
T/S>A -37 -319 vs. N>D -204 -319	ns
T/S>A -37 -319 vs. N>D FKO	ns
T/S>A -37 -319 vs. N>Q -37	****
T/S>A -37 -319 vs. N>Q -204	****
T/S>A -37 -319 vs. N>Q -319	ns
T/S>A -37 -319 vs. N>Q -37 -204	ns
T/S>A -37 -319 vs. N>Q -37 -319	ns
T/S>A -37 -319 vs. N>Q -204 -319	ns
T/S>A -37 -319 vs. N>Q FKO	ns
T/S>A -37 -319 vs. CVS-11	****
T/S>A -37 -319 vs. NT HEK293t	ns
T/S>A -37 -319 vs. ΔE	ns
T/S>A -37 -319 vs. VSV	****
T/S>A -204 -319 vs. T/S>A FKO	ns
T/S>A -204 -319 vs. N>D -37	****
T/S>A -204 -319 vs. N>D -204	****
T/S>A -204 -319 vs. N>D -319	ns
T/S>A -204 -319 vs. N>D -37 -204	****
T/S>A -204 -319 vs. N>D -37 -319	ns
T/S>A -204 -319 vs. N>D -204 -319	ns
T/S>A -204 -319 vs. N>D FKO	ns
T/S>A -204 -319 vs. N>Q -37	****
T/S>A -204 -319 vs. N>Q -204	****
T/S>A -204 -319 vs. N>Q -319	ns

T/S>A -204 -319 vs. N>Q -37 -204	ns
T/S>A -204 -319 vs. N>Q -37 -319	ns
T/S>A -204 -319 vs. N>Q -204 -319	ns
T/S>A -204 -319 vs. N>Q FKO	ns
T/S>A -204 -319 vs. CVS-11	****
T/S>A -204 -319 vs. NT HEK293t	ns
T/S>A -204 -319 vs. ΔE	ns
T/S>A -204 -319 vs. VSV	****
T/S>A FKO vs. N>D -37	****
T/S>A FKO vs. N>D -204	****
T/S>A FKO vs. N>D -319	ns
T/S>A FKO vs. N>D -37 -204	****
T/S>A FKO vs. N>D -37 -319	ns
T/S>A FKO vs. N>D -204 -319	ns
T/S>A FKO vs. N>D FKO	ns
T/S>A FKO vs. N>Q -37	****
T/S>A FKO vs. N>Q -204	****
T/S>A FKO vs. N>Q -319	ns
T/S>A FKO vs. N>Q -37 -204	ns
T/S>A FKO vs. N>Q -37 -319	ns
T/S>A FKO vs. N>Q -204 -319	ns
T/S>A FKO vs. N>Q FKO	ns
T/S>A FKO vs. CVS-11	****
T/S>A FKO vs. NT HEK293t	ns
T/S>A FKO vs. ΔE	ns
T/S>A FKO vs. VSV	****
N>D -37 vs. N>D -204	***
N>D -37 vs. N>D -319	****
N>D -37 vs. N>D -37 -204	****
N>D -37 vs. N>D -37 -319	****
N>D -37 vs. N>D -204 -319	****
N>D -37 vs. N>D FKO	****
N>D -37 vs. N>Q -37	ns
N>D -37 vs. N>Q -204	****
N>D -37 vs. N>Q -319	****
N>D -37 vs. N>Q -37 -204	****
N>D -37 vs. N>Q -37 -319	****
N>D -37 vs. N>Q -204 -319	****
N>D -37 vs. N>Q FKO	****
N>D -37 vs. CVS-11	***
N>D -37 vs. NT HEK293t	****
N>D -37 vs. ΔE	****
N>D -37 vs. VSV	****
N>D -204 vs. N>D -319	****
N>D -204 vs. N>D -37 -204	****

N>D -204 vs. N>D -37 -319	****
N>D -204 vs. N>D -204 -319	****
N>D -204 vs. N>D FKO	****
N>D -204 vs. N>Q -37	**
N>D -204 vs. N>Q -204	****
N>D -204 vs. N>Q -319	****
N>D -204 vs. N>Q -37 -204	****
N>D -204 vs. N>Q -37 -319	****
N>D -204 vs. N>Q -204 -319	****
N>D -204 vs. N>Q FKO	****
N>D -204 vs. CVS-11	ns
N>D -204 vs. NT HEK293t	****
N>D -204 vs. ΔE	****
N>D -204 vs. VSV	****
N>D -319 vs. N>D -37 -204	****
N>D -319 vs. N>D -37 -319	ns
N>D -319 vs. N>D -204 -319	ns
N>D -319 vs. N>D FKO	ns
N>D -319 vs. N>Q -37	****
N>D -319 vs. N>Q -204	****
N>D -319 vs. N>Q -319	ns
N>D -319 vs. N>Q -37 -204	ns
N>D -319 vs. N>Q -37 -319	ns
N>D -319 vs. N>Q -204 -319	ns
N>D -319 vs. N>Q FKO	ns
N>D -319 vs. CVS-11	****
N>D -319 vs. NT HEK293t	ns
N>D -319 vs. ΔE	ns
N>D -319 vs. VSV	****
N>D -37 -204 vs. N>D -37 -319	****
N>D -37 -204 vs. N>D -204 -319	****
N>D -37 -204 vs. N>D FKO	****
N>D -37 -204 vs. N>Q -37	****
N>D -37 -204 vs. N>Q -204	ns
N>D -37 -204 vs. N>Q -319	****
N>D -37 -204 vs. N>Q -37 -204	**
N>D -37 -204 vs. N>Q -37 -319	****
N>D -37 -204 vs. N>Q -204 -319	****
N>D -37 -204 vs. N>Q FKO	****
N>D -37 -204 vs. CVS-11	****
N>D -37 -204 vs. NT HEK293t	****
N>D -37 -204 vs. ΔE	****
N>D -37 -204 vs. VSV	****
N>D -37 -319 vs. N>D -204 -319	ns
N>D -37 -319 vs. N>D FKO	ns

N>D -37 -319 vs. N>Q -37	****
N>D -37 -319 vs. N>Q -204	****
N>D -37 -319 vs. N>Q -319	ns
N>D -37 -319 vs. N>Q -37 -204	ns
N>D -37 -319 vs. N>Q -37 -319	ns
N>D -37 -319 vs. N>Q -204 -319	ns
N>D -37 -319 vs. N>Q FKO	ns
N>D -37 -319 vs. CVS-11	****
N>D -37 -319 vs. NT HEK293t	ns
N>D -37 -319 vs. ΔE	ns
N>D -37 -319 vs. VSV	****
N>D -204 -319 vs. N>D FKO	ns
N>D -204 -319 vs. N>Q -37	****
N>D -204 -319 vs. N>Q -204	****
N>D -204 -319 vs. N>Q -319	ns
N>D -204 -319 vs. N>Q -37 -204	ns
N>D -204 -319 vs. N>Q -37 -319	ns
N>D -204 -319 vs. N>Q -204 -319	ns
N>D -204 -319 vs. N>Q FKO	ns
N>D -204 -319 vs. CVS-11	****
N>D -204 -319 vs. NT HEK293t	ns
N>D -204 -319 vs. ΔE	ns
N>D -204 -319 vs. VSV	****
N>D FKO vs. N>Q -37	****
N>D FKO vs. N>Q -204	****
N>D FKO vs. N>Q -319	ns
N>D FKO vs. N>Q -37 -204	ns
N>D FKO vs. N>Q -37 -319	ns
N>D FKO vs. N>Q -204 -319	ns
N>D FKO vs. N>Q FKO	ns
N>D FKO vs. CVS-11	****
N>D FKO vs. NT HEK293t	ns
N>D FKO vs. ΔE	ns
N>D FKO vs. VSV	****
N>Q -37 vs. N>Q -204	****
N>Q -37 vs. N>Q -319	****
N>Q -37 vs. N>Q -37 -204	****
N>Q -37 vs. N>Q -37 -319	****
N>Q -37 vs. N>Q -204 -319	****
N>Q -37 vs. N>Q FKO	****
N>Q -37 vs. CVS-11	**
N>Q -37 vs. NT HEK293t	****
N>Q -37 vs. ΔE	****
N>Q -37 vs. VSV	****
N>Q -204 vs. N>Q -319	****

N>Q -204 vs. N>Q -37 -204	*
N>Q -204 vs. N>Q -37 -319	****
N>Q -204 vs. N>Q -204 -319	****
N>Q -204 vs. N>Q FKO	****
N>Q -204 vs. CVS-11	****
N>Q -204 vs. NT HEK293t	****
N>Q -204 vs. ΔE	****
N>Q -204 vs. VSV	****
N>Q -319 vs. N>Q -37 -204	ns
N>Q -319 vs. N>Q -37 -319	ns
N>Q -319 vs. N>Q -204 -319	ns
N>Q -319 vs. N>Q FKO	ns
N>Q -319 vs. CVS-11	****
N>Q -319 vs. NT HEK293t	ns
N>Q -319 vs. ΔE	ns
N>Q -319 vs. VSV	****
N>Q -37 -204 vs. N>Q -37 -319	ns
N>Q -37 -204 vs. N>Q -204 -319	ns
N>Q -37 -204 vs. N>Q FKO	ns
N>Q -37 -204 vs. CVS-11	****
N>Q -37 -204 vs. NT HEK293t	ns
N>Q -37 -204 vs. ΔE	ns
N>Q -37 -204 vs. VSV	****
N>Q -37 -319 vs. N>Q -204 -319	ns
N>Q -37 -319 vs. N>Q FKO	ns
N>Q -37 -319 vs. CVS-11	****
N>Q -37 -319 vs. NT HEK293t	ns
N>Q -37 -319 vs. ΔE	ns
N>Q -37 -319 vs. VSV	****
N>Q -204 -319 vs. N>Q FKO	ns
N>Q -204 -319 vs. CVS-11	****
N>Q -204 -319 vs. NT HEK293t	ns
N>Q -204 -319 vs. ΔE	ns
N>Q -204 -319 vs. VSV	****
N>Q FKO vs. CVS-11	****
N>Q FKO vs. NT HEK293t	ns
N>Q FKO vs. ΔE	ns
N>Q FKO vs. VSV	****
CVS-11 vs. NT HEK293t	****
CVS-11 vs. ΔE	****
CVS-11 vs. VSV	****
NT HEK293t vs. ΔE	ns
NT HEK293t vs. VSV	****
ΔE vs. VSV	****

Table S.4: A table of significance for figure 4.1 B

Tukey's multiple comparisons test	Significance
T/S>A -37 vs. T/S>A -204	****
T/S>A -37 vs. T/S>A -37 -204	****
T/S>A -37 vs. N>D -37	****
T/S>A -37 vs. N>D -204	ns
T/S>A -37 vs. N>D -37 -204	****
T/S>A -37 vs. N>Q -37	****
T/S>A -37 vs. N>Q -204	****
T/S>A -37 vs. N>Q -37 -204	****
T/S>A -37 vs. VSV	****
T/S>A -37 vs. CVS-11	****
T/S>A -37 vs. NT HEK293t	****
T/S>A -37 vs. ΔE	****
T/S>A -204 vs. T/S>A -37 -204	***
T/S>A -204 vs. N>D -37	****
T/S>A -204 vs. N>D -204	***
T/S>A -204 vs. N>D -37 -204	ns
T/S>A -204 vs. N>Q -37	****
T/S>A -204 vs. N>Q -204	ns
T/S>A -204 vs. N>Q -37 -204	ns
T/S>A -204 vs. VSV	****
T/S>A -204 vs. CVS-11	****
T/S>A -204 vs. NT HEK293t	***
T/S>A -204 vs. ΔE	***
T/S>A -37 -204 vs. N>D -37	****
T/S>A -37 -204 vs. N>D -204	****
T/S>A -37 -204 vs. N>D -37 -204	****
T/S>A -37 -204 vs. N>Q -37	****
T/S>A -37 -204 vs. N>Q -204	****
T/S>A -37 -204 vs. N>Q -37 -204	*
T/S>A -37 -204 vs. VSV	****
T/S>A -37 -204 vs. CVS-11	****
T/S>A -37 -204 vs. NT HEK293t	ns
T/S>A -37 -204 vs. ΔE	ns
N>D -37 vs. N>D -204	****
N>D -37 vs. N>D -37 -204	****
N>D -37 vs. N>Q -37	****
N>D -37 vs. N>Q -204	****
N>D -37 vs. N>Q -37 -204	****
N>D -37 vs. VSV	****
N>D -37 vs. CVS-11	ns

N>D -37 vs. NT HEK293t	****
N>D -37 vs. ΔE	****
N>D -204 vs. N>D -37 -204	**
N>D -204 vs. N>Q -37	****
N>D -204 vs. N>Q -204	**
N>D -204 vs. N>Q -37 -204	****
N>D -204 vs. VSV	****
N>D -204 vs. CVS-11	****
N>D -204 vs. NT HEK293t	****
N>D -204 vs. ΔE	****
N>D -37 -204 vs. N>Q -37	****
N>D -37 -204 vs. N>Q -204	ns
N>D -37 -204 vs. N>Q -37 -204	ns
N>D -37 -204 vs. VSV	****
N>D -37 -204 vs. CVS-11	****
N>D -37 -204 vs. NT HEK293t	****
N>D -37 -204 vs. ΔE	****
N>Q -37 vs. N>Q -204	****
N>Q -37 vs. N>Q -37 -204	****
N>Q -37 vs. VSV	****
N>Q -37 vs. CVS-11	****
N>Q -37 vs. NT HEK293t	****
N>Q -37 vs. ΔE	****
N>Q -204 vs. N>Q -37 -204	ns
N>Q -204 vs. VSV	****
N>Q -204 vs. CVS-11	****
N>Q -204 vs. NT HEK293t	****
N>Q -204 vs. ΔE	****
N>Q -37 -204 vs. VSV	****
N>Q -37 -204 vs. CVS-11	****
N>Q -37 -204 vs. NT HEK293t	*
N>Q -37 -204 vs. ΔE	*
VSV vs. CVS-11	****
VSV vs. NT HEK293t	****
VSV vs. ΔE	****
CVS-11 vs. NT HEK293t	****
CVS-11 vs. ΔE	****
NT HEK293t vs. ΔE	ns

Table S.5: A table of significance for figure 4.1 C

Tukey's multiple comparisons test	Significance
T/S>A -37 vs. T/S>A -204	***
T/S>A -37 vs. T/S>A -37 -204	****
T/S>A -37 vs. N>D -37	*
T/S>A -37 vs. N>D -204	ns
T/S>A -37 vs. N>D -37 -204	***
T/S>A -37 vs. N>Q -37	****
T/S>A -37 vs. N>Q -204	****
T/S>A -37 vs. N>Q -37 -204	****
T/S>A -37 vs. VSV	****
T/S>A -37 vs. CVS-11	ns
T/S>A -37 vs. NT HEK293t	****
T/S>A -37 vs. ΔE	****
T/S>A -204 vs. T/S>A -37 -204	***
T/S>A -204 vs. N>D -37	****
T/S>A -204 vs. N>D -204	ns
T/S>A -204 vs. N>D -37 -204	ns
T/S>A -204 vs. N>Q -37	****
T/S>A -204 vs. N>Q -204	ns
T/S>A -204 vs. N>Q -37 -204	ns
T/S>A -204 vs. VSV	****
T/S>A -204 vs. CVS-11	****
T/S>A -204 vs. NT HEK293t	****
T/S>A -204 vs. ΔE	****
T/S>A -37 -204 vs. N>D -37	****
T/S>A -37 -204 vs. N>D -204	****
T/S>A -37 -204 vs. N>D -37 -204	****
T/S>A -37 -204 vs. N>Q -37	****
T/S>A -37 -204 vs. N>Q -204	**
T/S>A -37 -204 vs. N>Q -37 -204	*
T/S>A -37 -204 vs. VSV	****
T/S>A -37 -204 vs. CVS-11	****
T/S>A -37 -204 vs. NT HEK293t	ns
T/S>A -37 -204 vs. ΔE	ns
N>D -37 vs. N>D -204	****
N>D -37 vs. N>D -37 -204	****
N>D -37 vs. N>Q -37	ns
N>D -37 vs. N>Q -204	****
N>D -37 vs. N>Q -37 -204	****
N>D -37 vs. VSV	****
N>D -37 vs. CVS-11	ns

N>D -37 vs. NT HEK293t	****
N>D -37 vs. ΔE	****
N>D -204 vs. N>D -37 -204	ns
N>D -204 vs. N>Q -37	****
N>D -204 vs. N>Q -204	**
N>D -204 vs. N>Q -37 -204	***
N>D -204 vs. VSV	****
N>D -204 vs. CVS-11	****
N>D -204 vs. NT HEK293t	****
N>D -204 vs. ΔE	****
N>D -37 -204 vs. N>Q -37	****
N>D -37 -204 vs. N>Q -204	ns
N>D -37 -204 vs. N>Q -37 -204	ns
N>D -37 -204 vs. VSV	****
N>D -37 -204 vs. CVS-11	****
N>D -37 -204 vs. NT HEK293t	****
N>D -37 -204 vs. ΔE	****
N>Q -37 vs. N>Q -204	****
N>Q -37 vs. N>Q -37 -204	****
N>Q -37 vs. VSV	****
N>Q -37 vs. CVS-11	ns
N>Q -37 vs. NT HEK293t	****
N>Q -37 vs. ΔE	****
N>Q -204 vs. N>Q -37 -204	ns
N>Q -204 vs. VSV	****
N>Q -204 vs. CVS-11	****
N>Q -204 vs. NT HEK293t	***
N>Q -204 vs. ΔE	***
N>Q -37 -204 vs. VSV	****
N>Q -37 -204 vs. CVS-11	****
N>Q -37 -204 vs. NT HEK293t	***
N>Q -37 -204 vs. ΔE	***
VSV vs. CVS-11	****
VSV vs. NT HEK293t	****
VSV vs. ΔE	****
CVS-11 vs. NT HEK293t	****
CVS-11 vs. ΔE	****
NT HEK293t vs. ΔE	ns