Antibody Engineering and Antigen Discovery against Avian Paramyxovirus:



Matthew A R Llewellyn

Student Number: 1147165

Sommerville College

University of Oxford

A Thesis submitted for the degree of

Doctor of Philosophy

Trinity 2023

Declaration:

I, Matthew Adam Rhys Llewellyn confirm that the work presented in this thesis is my own except where mentioned otherwise.

October 2023

Acknowledgements:

This piece of work was entirely made possible by the support of my supervisors, colleagues, friends and family. The first thank you, of, course must go to Professor Jonathan Grimes, in whom I could not ask for a more supportive supervisor, who when I had an issue, all I needed do was ask for help.

Secondly, for all their support in the cryo-EM work, I have to thank Drs Loic Carrique and Jeremy Keown, who taught me, from scratch and were always available to help. Similarly, thanks must go to Drs Mattia Bertinelli, Max Renner and Guido Paesan who showed me the ropes when I first joined Strubi and were always happy to help. For her help reading up on class I fusion proteins and supplying several positive controls I would also thank Dr Vicky Avanzato. Margaret Jones and Weixien Lu deserve plaudits for their tireless work and helping our research flow. My time at the Pirbright institute would not have been possible without the tireless work, and sanity checking, of Dr Yaoyao Zhang, who helped me muddle my way through a very trying time.

A very substantial thanks also needs go to Lieutenant Colonel Stewart QARANC, Major Curtis-Rouse RCS, Major Copeland REME, Captain Major OBE RAMC, Lieutenant Thody RLC, Staff Corporal Thompson, Lance Corporal Firth, Trooper Gore, Gunner Havins and Guardsman Calderwood. Through their leadership, friendship and support I was able to earn the Queen's Commission in parallel to my research.



An unreserved thanks needs to go to my family who have put up with me being constantly busy, yet supportive none-the less. A final thanks must go to my Fiancée, who has been my rock throughout the last five years.

Index:

Abstract: 6
Call Nan Cearc 7
Chapter I Introduction: 8
Emergence of Newcastle Disease Virus 9
Impact 12
Recent Outbreaks 14
The Virus 15
The Fusion Protein 18
Overview of avian Immunology 20
Structural studies to-date 22
Summary 25
Hypothesis: 26
Impact of the Pandemic 26
Chapter II Materials and Methods: 27
Design of the construct: 27
Protein expression: 28
Protein purification: 30
Crystal tray preparation and crystallography: 31
Electron Microscopy: 32
Cross-linking, Chameleon and Pegylation: 33
Preparing F gene construct for insertion into HVT: 35
Immunofluorescence: 36
Inserting F construct into HVT 36
ELISA: 39
Nanobody expression and Immunoprecipitation: 40
Cloning nanobodies and large-scale expression: 41
Chapter III Introduction: 43
Results 47
Expressing and Modelling of the F protein 48
Purification of the F protein 58
Thermofluor experiments 60
Setting up Crystal Trays 62
X-Ray Crystallography 66
SAXs Experiments 68
Discussion 77
Chapter IV Introduction: 81
Negative Stain and initial Cryo EM 84
Cross Linking 91
Optimisation of Cryo EM 93
Chameleon 96
UNcle Assays 100
EM with Pegylated F 106
Discussion 125
Chapter V Introduction: 126
Results 130
Cloning the F protein into the donor vector 130

Inserting the construct into HVT 135
Animal Experiment 138
Small-scale nanobody expression and binding 139
Large-scale nanobody expression and binding 143
Discussion 149
Chapter VI Overall conclusions: 153
Future work 157
Appendix A (Bibliography) 160
Appendix B (List of genes) 176
Appendix C (Permissions for reproducing figures) 182

Abstract:

Twenty-five years since the first structure of Avian Orthoavulavirus (AOav) F protein was solved in the post-fusion conformation, the structure of the pre-fusion conformation remains elusive. AOaV is a major pathogen of poultry, especially chickens, and presents major challenges to small scale farmers and larger industry alike. Vaccination is one of the most effective ways of protecting against the pathogen. The primary antigen used, the F protein, is known to transition between 2 conformations, the meta-stable pre-fusion conformation and the stable post-fusion conformation. The structure of the post-fusion conformation is known, but the structure of the pre-fusion conformation is unknown.

The first aim of this project was to solve the pre-fusion structure of NDV F. This structure could then be used to develop a stabilised F antigen, which would allow for the creation of a more effective vectored-vaccine by preventing the conformational change. Additionally, epitope mapping could have been performed to allow for identification of regions on the antigen that were common to multiple genotypes of the virus. All this put together would have created a candidate antigen which could be inserted into a Turkey Herpesvirus (HVT) vectored vaccine. This prototype vaccine could then be compared with current vaccines in a series of immunisation assays and eventually, challenge assays.

The project made an initial success, with protein crystallising, and then the crystals examined with X-ray crystallography. However, the structure could not be determined from this dataset. Following this, further attempts to grow crystals were made, all without success. Thus, a pivot was made to cryo-EM, again with some initial promise. The task proved problematic due to difficulties in sample preparation and the preparation of grids for cryo-EM. Meanwhile a SAXs dataset was collected and compared with some models of the protein, generated with Alpha-Fold, and other class I fusion proteins and our own results. The cryo-EM work was interrupted by the pandemic, and due to other considerations, the attempt to insert the stabilised F protein into an HVT vector was brought forwards. This attempt failed. Furthermore, the nanobodies generated in the start of the project from expressed F protein were also cloned into suitable vectors and expressed. However, it appeared that when incubated with the F

protein, the nanobodies also appeared not to bind. This impacted efforts to determine the structure of the F protein as they could have been used to stabilise the protein for crystallography or cryo-EM work. Other attempts to stabilise the F protein with additives and pegylation were also initially promising, but ultimately unsuccessful. Taken as a whole, the project provides a framework for future structural studies into a protein that remains a high priority target.

Call nan, Cearc

Is ann an so a tha'n seanchas;

Here is the story;

Is e air feadh gach baile

It is all over town

Eadar bodaich is paisdean

Between old men and children

'S gu seachd araid aig mnathan,

And to seven strange women,

Is aid nionagan oga

They are young girls

Is iad an comhnaigh ag gearain,

They are always complaining,

Ged nach canadh iad riums'e,

Although they would not say to me,

Is mor an iondrainn a th'aca

They miss them dearly

Air call nan cearc

On the loss of the chickens

John Campbell

Chapter I

Introduction

Newcastle Disease Virus, also known as Avian Orthoavulavirus (AOaV), continues to present a considerable threat to food security despite attempts at eradication and extensive vaccination (1). It is a superlatively heterogeneous virus with over 20 different genotypes and limited cross-protection between strains (2). When it infects poultry, with sufficient virulence, it is termed Newcastle Disease (ND). Most commonly it will variably manifest with respiratory distress and rales, enteric lesions, green diarrhoea, swollen eye sockets, egg-drop or deformation and reduced weight-gain. Often, the most virulent cases will exhibit 'torticollis' or 'wry-neck' -the contorting of the neck- and other neurological symptoms (3, 4). It is highly infectious and capable of infecting over 200 species of bird, including poultry, waterfowl and exotic birds. In chickens it can have up to 100% lethality, while turkeys and other birds have a level of immunity (4).

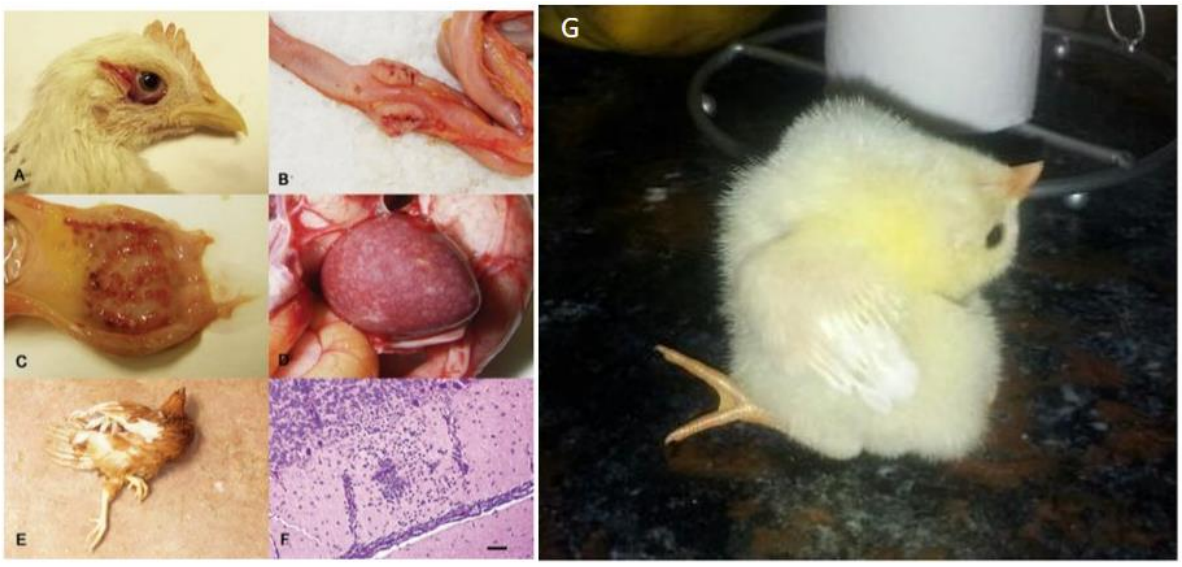


Figure 1: a breakdown of the common symptoms of NDV. 1A: Haemorrhage of eye lymph nodes. 1B: Caecal tonsil Haemorrhage. 1C. Necrosis of lymphoid tissue around the proventriculus. 1D: Spleen

lesions. 1E: Weakness on one side of the body. 1F cerebral lesions and extensive gliosis (3). 1G: The torticollis typical of velogenic NDV.

Emergence of the virus:

The origins of NDV are not known and as one of the myriad viruses that cause respiratory and enteric infections in chickens, its emergence is hard to date, but it is estimated to have emerged approximately 250 years ago (6). Historical reports of fatal illnesses sweeping through flocks of poultry are common and even feature in a Gallic poem by John Campbell (7). Indeed, analysis of this poem (8) led to MacPherson's 1956 hypothesis that an 1898 outbreak of disease in Skye was NDV. This was based on the similarity of the symptoms described, also that relief organisations had gifted healthy birds and eggs to the islanders to replace losses (6). Additionally, at the time of MacPhearson's study, there was an on-going outbreak of ND on the islands, matching much of the characteristics described by Campbell's poem (see figure 2). MacPhearson confirmed the presence of ND in 1958, but was not able to prove that the 1898 outbreak was ND (8). More recent Bayesian Coalescent inference traced the most common ancestor of class II NDV to approximately 1883 (9).

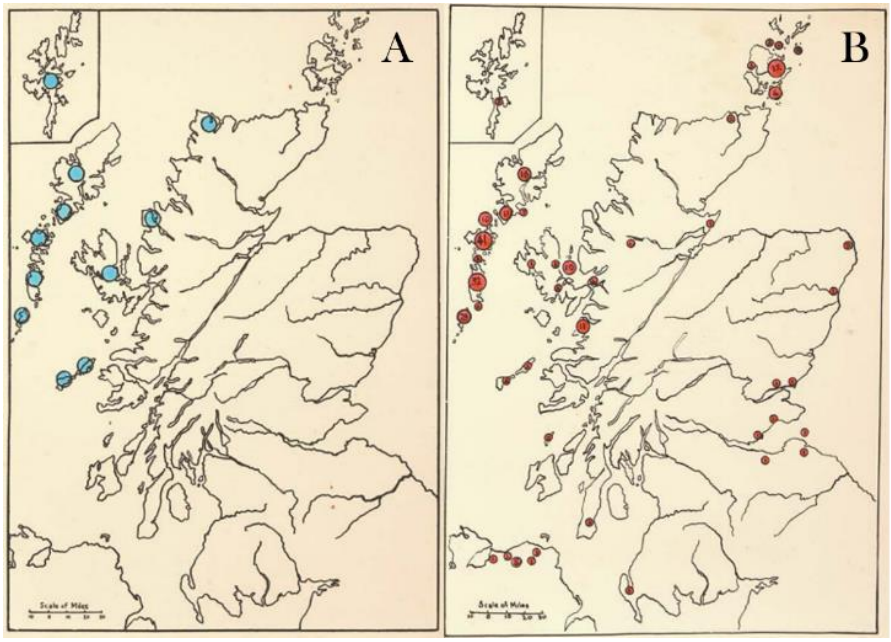


Figure 2: Maps from MacPhearson's report. Pannel A show locations where the "Congested Districts Board" (An organisation for the relief of the poor and charitable works) supplied replacement chickens

to in 1898. Pannel B shows confirmed outbreaks of NDV during the 1949-51 epizootic in Scotland. Reproduced from MacPhearson's 1958 report (8).

NDV was initially considered to be a variant of Fowl Pox, termed 'Roupe,' due to the similarities in the clinical presentation (4). In the late 1920s the Dutch East Indies (present day Indonesia), the Philippines, Korea and the UK experienced epizootics with similar symptoms, but these were more lethal (10-12). Doyle and Minnet recognised it to be a distinct organism, which could pass through a bacteria-excluding filter and had hemagglutinating properties. After being characterised, it was named 'Newcastle Disease Virus (NDV), after the UK area most afflicted. Debate continued to rage as to whether-or-not it was an avian influenza, and whether subsequent outbreaks were Influenza or NDV. Often the presentation of neurological symptoms served as the key differentiating factor (10-12). It was also noted not to have an adverse effect on ducks (13).

Epizootics in the mid-1930s, 1943 and in 1949 spurred the introduction of the first vaccines and attempts to eradicate it, in the manner of rinderpest, though complete eradication was impossible due to the heterogeneity of the virus and a large wild-animal reservoir (14,15). Despite this, a combination of culling, quarantines, surveillance and vaccination allowed many nations to eradicate the virus (16). Most European nations mandate the vaccination of all poultry and maintain a level of veterinary surveillance. Unfortunately, it remained endemic in many less-affluent nations. This, coupled with the persistence in migratory wild flocks ensure that sporadic outbreaks continue to occur to this day (17).

By the 1960s it was reclassified as part of a family of avian paramyxoviruses. The onset of DNA sequencing of viruses led to the re-evaluation of viral phylogeny (18). On account of the diversity, the family was renamed 'Avulavirus' in 2002 and NDV was renamed Avian Avulavirus 1 in 2016. A review in 2018 led to the stratification into Ortho, Meta and Para-avulavirus, with NDV becoming Avian Orthoavulavirus 1. (See Phylogeny in Figure 3) (2, 19).

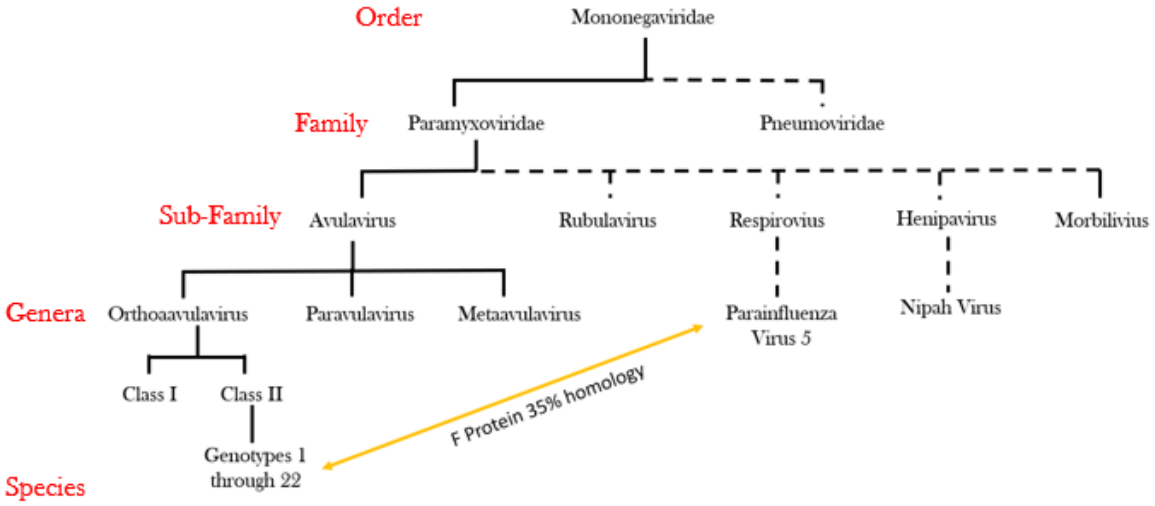


Figure 3: A phylogenetic tree showing the relationship between some key members of the Mononegaviridae family. Indicated is the homology between the F protein of NDV and that of PIV5, the most closely related member of the family that has been solved, using DNA sequence alignment. For simplicity, several sub-families are omitted.

NDV infections are classified as being 'Lentogenic,' 'Mesogenic' or 'Velogenic,' depending on the Intra-Cerebral Pathogenicity Index score (ICPI), which is calculated by inoculating 1-day old Specific Pathogen Free (SPF) chicks with the virus (2, 20). See table 1. Only mesogenic or velogenic infections are classed as ND and subject to environmental control measures (21,22).

Virulence:	Severity:	ICPI:	Tropism:
Lentogenic	Mild	ICPI < 0.7	Respiratory, Enteric
Mesogenic	Moderate	0.7 ≤ ICPI < 1.5	Respiratory, Enteric
Velogenic	Severe	1.5 ≤ ICPI	Respiratory, Enteric, Neurological

Table 1: A table surmising the severity, Inter-Cerebral-Pathogenicity-Index (ICPI) and tropism of each virulence category of the virus. ICPI is measured by averaging the mortality and morbidity of Specific Pathogen Free embryonated eggs.

Mesogenic and Velogenic strains are rarely detected in wealthy, developed nations, but will cause up to 100% mortality within a flock and necessitate the culling of birds on affected farms, as well as those in close proximity (22). Additionally, poultry farmers not directly affected by infections or culling, will be constrained by no-movement orders, embargos and food safety concerns. The mere threat of a

notifiable disease outbreak can lead to all 'Free Range' birds being confined to houses, so the products can no-longer be marketed as free-range, as seen in recent Avian Influenza outbreaks (23).

Further Phylogeny:

There are two classes of NDV, Class I are almost exclusively lentogenic or non-pathogenic, in chickens, while Class II contains viruses of varying virulence. As of January 2021, there are 22 genotypes, with genotype II, known as 'La Sota,' serving as the principal vaccine strain (2, 24).

In recent years several other strains have risen to prominence, including genotype V, which caused a large epizootic in California in 2019 and Genotype VII (GTVII) which emerged from Asia and the middle east (25, 26). In 2018 there was a GTVII.2 outbreak in Belgium, the Netherlands and Luxembourg, showing that Miller et al's 2015 warning that GTVII.2 could pose a panzootic threat was accurate (27, 28).

Impact:

Worldwide the poultry industry accounted for $310.7 billion in 2020, made up nearly 40% of meat produced by tonne and is considered one of the fastest growing food industries (29, 30). In 2019, over 118 million tonnes of chicken and nearly 1.6 trillion eggs were produced worldwide (31).

Chicken is rapidly becoming the most eaten meat, partially due to not being subject to some religious dietary laws and is seen as being 'less polluting than beef' by an increasingly environmentally conscious society (32, 33). Meat is only a fraction of the picture; there is a constantly growing demand for eggs, with the UK alone producing 10.4 billion in 2022 and is firmly a staple of many diets, both on its own and as an ingredient (34).

Infection with NDV need not be fatal to have an impact: infected birds do not grow and develop at the same rate and the lesions caused to the liver and other internal organs is sufficient for the broiler to be rejected at the abattoir as unfit for human consumption (35). Furthermore, in laying flocks, the number of eggs produced reduces (known as egg drop) and those that are produced are smaller, and have weaker, or absent shells (36). Typical measures instituted for infection control require the destruction of the entire flock, and any nearby flocks and domestic birds within a certain radius (37).

The death of birds, either caused by the infection itself, or on the orders of local authorities is particularly upsetting for the keepers of domesticated birds, such as parrots, or hobby flocks of chickens. Under UK law, a backyard/hobby flock is up to 50 birds and does not need registration with the Department for the Environment, Food and Rural Affairs (DEFRA) (38). It further poses a threat to endangered birds, such as the penguins of the Falkland Islands, or those in zoos (39).

Whilst it is certainly of high economic consequence in the west, in less affluent countries, the loss of chickens can cause severe hardship for a family. Often, poultry are the first animals acquired by subsistence farmers, being relatively inexpensive to keep (40). Their loss can cause immense hardship for already impoverished families as poultry, or their eggs, are bartered for goods and services: the eggs can supplement the diet and the bird will be a source of meat.

Since the introduction of the first NDV vaccines in 1949, there have been continual efforts for improvement, to give more complete protection, minimise risk and to protect against greater spread of variants (41). Until the advent of gene-editing technology, all vaccines used were either live attenuated or inactivated viruses. Despite the limitations of these systems (see chapter V for further details), use of these vaccines and strict biosecurity measures were sufficient to eradicate NDV in most of Western Europe and North America (42, 43).

However, someone must pay to vaccinate these flocks, commonly the farmers themselves. One of the most prevalent and effective vaccines is the live-attenuated 'La Sota' Vaccine (24, 44). This provides cellular and humoral-based immunity to genotype II virus. While the necessity is clear, $0.16 a dose quickly adds up when you consider large flocks, requiring multiple doses, and that the farmers themselves have a low income (45). Furthermore, many of these vaccines are thermo-sensitive and require cold-chain storage, something not always available (or prohibitively expensive) in low-income regions where the virus is endemic (46). The La Sota vaccine itself is noted as being prone to 'reverting to virulence,' where un-vaccinated birds are in proximity to newly vaccinated birds which are shedding live virus (47). In such cases the unvaccinated birds are liable to develop clinical disease. As such, only inactivated or recombinant vaccines (such as the bi/trivalent Innovax) are used in the UK (48).

To frustrate matters further, one of the more common strains circulating in North Africa, the Middle East and Southern Asia is a genotype VII (GTVII) even causing several outbreaks in Europe (49).

GTVII is antigenically distinct from genotype II: the La Sota vaccine does provide some protection against GTVII, though this is only partial (50). As such there is much room for vaccine development.

Beyond the cost of vaccines, the mere risk posed by NDV itself imposes additional costs. Organisations like the British Animal and Plant Health Agency (APHA) are required to maintain vigilance: dedicating staff time and resources to clinical surveillance (51). Some of the measures imposed can lead to acrimony between trading partners (i.e import restrictions on poultry) and disrupt food supply chains. Meanwhile in more affluent countries, it also presents a major cost for commercial farmers, who own exceptionally large flocks. It is likely that such costs, along with those of the overseeing veterinarian, will be passed on to the consumer. The canonical lesions associated with enteric NDV, even of lentogenic strains, will be very apparent during inspection of the liver and intestines at an abattoir. When such lesions are found the whole consignment is condemned and destroyed (35). This presents another economic loss to the poultry industry.

Besides the food and economic dependence on poultry products, there is also the biosecurity aspect to consider. Many live-attenuated vaccines, for human use, are produced in chicken eggs, due to their availability and relative cost-effectiveness (52). This includes vaccines recommended for childhood immunisations and the seasonal flu vaccine (53). In cases of epidemic or pandemic flu, it is not unreasonable to see eggs used to mass-produce vaccine. While cell-based methods are becoming far more developed, egg-based methods remain viable and useful, despite limitations and vulnerabilities. However, movement restrictions, trading restrictions and ring-fence culls, due to a NDV outbreak, would turn any mass-production effort into a logistical morass.

<u>Recent outbreaks:</u>

The recent California outbreak of NDV sparked particular controversy as domestic birds were included in the 1.2 million culled to prevent the spread of the virus (54). Many of the culled animals included chickens kept as pets, as well as other domesticated birds, resulting in a social media furore and a campaign with an unfortunate abbreviation: ‘Save Our Birds’.

In 2018 there was also an outbreak in the Belgium countries: with 20 cases being reported, most being in hobby flocks, less two professional flocks (55). Since the late 1990’s, vaccination against NDV has

been compulsory in Belgium and a robust level of surveillance maintained. A rapid implementation of no move orders, quarantine and ring-fence vaccination was able to minimise the spread of the virus. Retrospective epidemiology indicated that it was a genotype VII.2 strain of the virus, with the cases originating within the Grand Duchy of Luxembourg traceable back to Belgium (56). There was some suggestion that traders selling to hobby flocks in live bird markets could have been to blame, though evidence was inconclusive.

Migratory waterfowl, as with avian influenza, are often a route of introduction to commercial flocks, and act as a natural reservoir of the virus where it is spread through the faecal-oral route (57). Both pigeons and waterfowl can be infected, though with minimal pathology. In the UK, a legal loop-hole exists for so-called 'Pidgeon Paramyxovirus:' the detection of NDV in pigeons, without the necessity of closing-down the poultry market overnight (58). While this might be convenient for exporting poultry, it does present a risk to those who keep birds where they might encounter pigeons.

The virus:

NDV is an enveloped, negative strand RNA virus with a genome approximately 14kb in length, which codes for six major viral proteins (59). The viral RNA (vRNA) genome is encapsidated by multiple copies of the Nucleocapsid protein (N) each of which binds 6 bases and oligomerises to form helical nucleocapsids (60). The nucleocapsid is surrounded by a lipid bilayer envelope, maintained by the Matrix protein (M) (61). Bound to the nucleocapsid are multiple copies of the L-P complex composed of the Phosphatase (P) and the RNA dependent RNA Polymerase (L) that both replicates and transcribes the vRNA genome (62). Inserted in the viral envelope are the Haemaglutanin-Neuraminidase (HN) and the Fusion (F) proteins which mediate attachment and entry to the cell (63, 64). Interestingly, while most paramyxovirus membrane fuse directly with the cell membrane, NDV is first endocytosed by the cell, then uses pH dependant fusion to escape the endosome (65).

Two minor proteins, V and W, are generated through RNA Editing of the P gene, by the addition of one or two guanine bases respectively (66). V and W are responsible for host cell take-over and modulation of the avian immune response: the V protein interacts with the N protein to control vRNA

production, while also inhibiting INFα and INFβ through targeting STAT1 for degradation and interacting with MDA5, respectively (67)

The P, L, V and W proteins mediate replication and facilitate host-cell take over. Replication of the vRNA takes place in the cytoplasm (68). Phylogenic classification of the virus into genotypes is carried out via the vRNA sequence of the F protein (2).

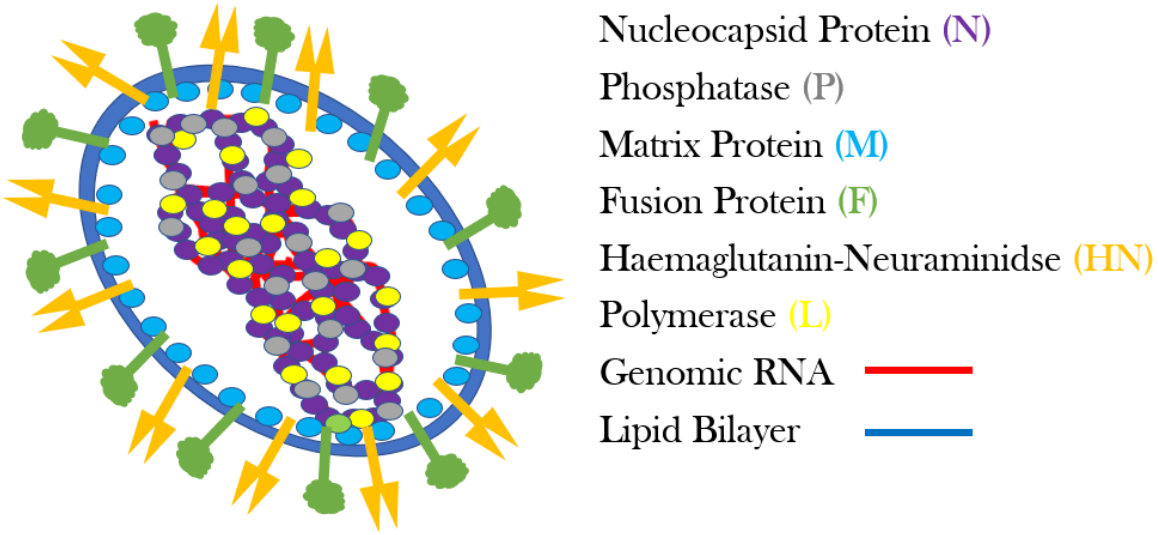


Figure 4: A cartoon showing the structure of NDV virions with representations of the proteins (Less V and W proteins).

A the 3' of each viral gene is a 'Gene Start' sequence. This re-initiates replication as it reaches each gene, with decreasing efficacy (69). Thus, the virus ends up with more messenger vRNA copies of the genes at the 3' end of the genome, namely the N, P and M genes, see figure 5 (70).

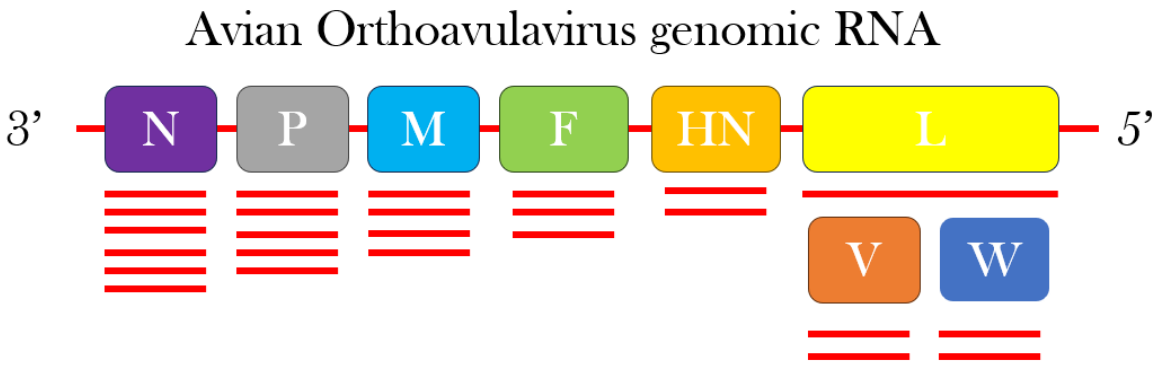


Figure 5: The organisation of the NDV genome and relative number of copies of transcription into mRNA. Modelled on the diagram in Shin-Hee Kim and Siba K. Samal's 2016 review (70).

The HN Protein is a 75KDa homo-tetramer, which mediates attachment to sialic acid residues, and also has neuraminidase activity (71). The interactions between itself and the F protein are imperfectly understood, but are thought to mediate the activity of the F protein (see figure 6) (72, 73).

An overview of Viral Fusion proteins:

Viral fusion proteins are type I trans-membrane proteins which mediate the entry of the virus into the new host cell and can be split into three distinct classes (74). Often they operate in conjunction with an attachment protein (sometimes integral) and exhibit an energetically favourable conformation change which serves to merge the viral envelope and the membrane of the cell to be infected (74).

Class I:

Class I fusion proteins, such as the NDV F, Nipah F and Influenza HA are metastable proteins that project perpendicular to the membrane (75-77). They are trimeric structures composed principally of α-helices with hydrophobic fusion peptides nestled at the interface between sub-units (78). These peptides are what interface with the membrane of the cell to be infected and are usually proximal to the N-terminus of each sub-unit and proteolytic processing is required to 'unsheathe' the peptide (78). When the virus particle is in proximity to the host membrane, the pre-fusion globular conformation shifts to a rod-like post-fusion conformation, thereby fusing the viral envelope with the cell membrane (78). In some cases, this is pH dependant, in others it is receptor dependant (65, 78). Further information is presented later in this chapter and in chapter III.

Class II:

Class II fusion proteins, by contrast, are often found in flaviviruses, can lie parallel to the membrane and are largely composed of β-sheets. They associate with a 'protector' protein in the Endoplasmic Reticulum, which is Furin-cleaved in the trans-Golgi network (79, 80). Their highly conserved fusion peptides sit at the end of the β-sheets extending from the protein and will insert into the membrane of a host cell as part of the entry process. Like some Class I fusion proteins, low pH plays a part in activation, but unlike class I fusion proteins, they are monomeric in the pre-fusion conformation (79). The fusion peptides contain hydrophobic residues which anchor the protein into the bilayer of the cell

membrane. Examples include the E proteins of Tick Bourne Encephalitis, Dengue and Yellow Fever viruses (79-81).

Class III:

Class III Fusion proteins are the most recently discovered class of Fusion protein and are commonly found on large double-stranded DNA viruses such as HSV and EBV, but also on negative sense RNA viruses like VSV. As such they are very variable and ectodomains can be as long as 770 residues or as short as 440 (82-84). Like class I fusion proteins they form trimers, which often are perpendicular to the envelope, but feature both α-helices and β-sheets prominently. Additionally, they do not require proteolytic processing to become competent. Some have been suggested to be pH dependant for fusion, using histadine residues as pH sensors. Others are receptor dependant (85).

The NDV F protein:

The NDV F protein is a Class I fusion protein and is the major surface antigen of the virus (86). It is a 150kDa homotrimer which exists in a meta-stable pre-fusion or a stable post-fusion conformation (87). This transition from pre- to post-fusion conformation facilitates cell entry and *in-vivo* is thought to involve association with the HN protein, once the HN has attached the virus particle to a cell (63). There has been debate as to whether triggering the conformational change in F is by association with HN, or dissociation from HN. The association model is considered to be more likely for NDV, as reported by Melanson and Iorio: who showed that a mutation in HN that lowers its affinity of association with F also reduces fusogenic activity (see figure 6) (88). Dissociation has been suggested as being more likely for Nipah and Measles viruses as, if the affinity between their F and HN analogues is too high, there is less viral fusion (89, 90).

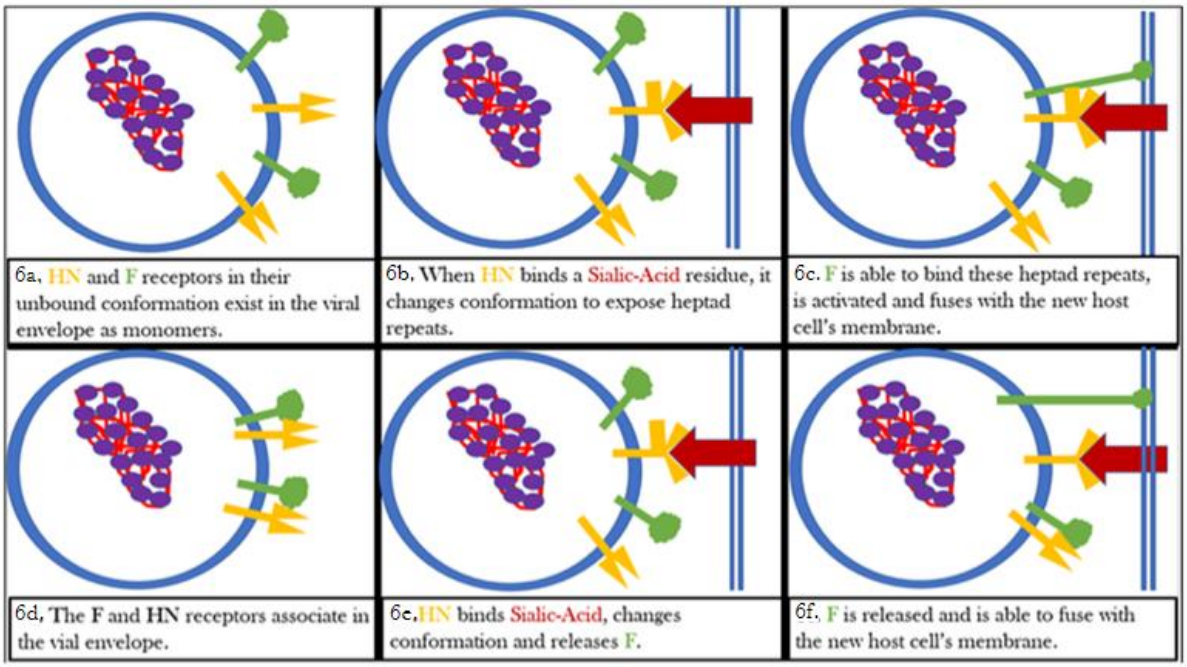


Figure 6: A cartoon showing the two possible methods by which NDV attaches to and enters a host cell by pH dependant fusion. Boxes 6a, b and c show the attach-associate-enter model while 6d, e and f show the attach-dissociate-enter model (88).

In-vitro the transition in conformation from pre-fusion to post-fusion has been observed to occur spontaneously, frustrating efforts to study the protein and produce recombinant vaccines (91). There is some evidence of the transition occurring *in-vivo* among other *Mononegaviridae*, possibly forming part of the viral strategy to evade the immune response (92). Antibodies raised against pre-fusion RSV F tend to be more neutralising and give better protection against live virus than post-fusion F antibodies (93).

Each of the three constituent monomers of F are synthesised as inactive precursors, F0. Each F0 is then cleaved by Furin-like-proteases, into F1 and F2, priming the trimer for conformational change (87). There has been some debate as to whether cleavage is necessary for conformational change, indeed, if it is the cause. Recent work on RSV using Furin inhibitors has shown the inhibition of fusogenic activity when the F protein is not cleaved (94). The purpose of this cleavage is to un-sheath a hydrophobic domain which, during fusion, integrates with the membrane of the cell that the viral membrane will fuse

with during viral infection (95). Just to the C terminus of the fusion peptide, and to the N terminus of the trans-membrane-domain are the heptad repeats (95).

The composition of this cleavage-site is one of the factors that determine the virulence: lentogenic and vaccine strains of the virus have up to three basic residues, while mesogenic and velogenic strains tend to be polybasic (95). The effect of this is that the Furin-like proteases can cleave velogenic cleavage sites more efficiently, consequently, they can infect cells more rapidly. The cleavage site sits approximately 110 AA residues from the N-terminus of the protein (96). One of the noted differences between genotype II (La Sota) and genotype VII (GTVII) is the presence of the polybasic cleavage site: ^{112}RRRKGF^{117}, compared with the tri-basic cleavage site of La Sota: ^{112}GRQGR^{116} (97). Overall, there is approximately 17% divergence in nucleotide sequence, and a 9% divergence the amino acid sequence between the sequences of La Sota F and GTVII F.

There are other factors that affect the virulence of the virus, though many of these are mediated by the other genes, especially the HN protein which has up to a 20% divergence in nucleotide sequence between the sequences of La Sota HN and GTVII HN (50).

Impact of the viral polymerase:

The evolution rate of NDV is high, with 1.059×10^{-3} substitutions/site/year, SARS-CoV-2 has a mutation rate of 1.12×10^{-3} substitutions/site/year, while influenza A was reported at 1.5×10^{-3} substitutions/site/year (98-100). Both these human pathogens are infamous for their speed of mutation, with NDV's is comparable. This allows NDV to diversify so that escape mutations to immune responses can occur rapidly.

Putting the 'B' in 'B-Cell': Avian adaptive immunity.

As the virus is principally dangerous to poultry, it is necessary to examine the avian immune system, and how it differs to the mammalian immune system that many prospective readers are more familiar with. When mammals and birds diverged from reptiles, their immune systems also diverged, which explains a number of distinct differences, most notably birds possess an organ known as the 'Bursar of Fabricius,'

which give the B-Cell its name (101, 102). In the 1950s it was observed as the site where avian immunoglobulin was synthesised, later being identified as a component of gut-associated lymphoid tissue and the central lymphoid organ of birds, where B-cells differentiate and mature (101).

The avian functional analogue to IgG is known as 'IgY' (103). The 'Y' standing for 'Yolk,' named thus because after secretion into the blood, the IgY would accumulate in the yolk of the egg, serving a function not unlike the transfer of maternal antibodies from mother to child. Avian immunoglobulins also differ in structure from the mammalian immunoglobulins (see figure 21). Avian IgY lacks the long hinge region seen on mammalian IgG, making it less flexible, while having an additional constant domain, see figure 7 (104). IgY also lacks the ability to activate complement, nor does it interact with Rheumatoid factor, protein A or protein C (104). Avian B-Cells also have access to fewer different variable genes to select from than mammalian B-Cells (104).

IgY has been shown to have a much lower degree of cross-reactivity than IgG, though tends to have a higher degree of binding avidity, is produced in greater quantities and more rapidly (106). Indeed, upon exposure to a known antigen, an avian memory B-Cells will proliferate and begin expressing antibodies much more quickly than a mammalian B-Cell (107).

Another difference is the increased primacy of T-cell mediated immunity, mediated by the increased variety of their cytotoxic T-Cells: avian T-Cells have access to more options in their variable domain than human T-Cells (108). Birds have one class of heterophils in lieu of neutrophils, basophils and eosinophils. Furthermore, nucleated erythrocytes also play a role in innate immunity (109).

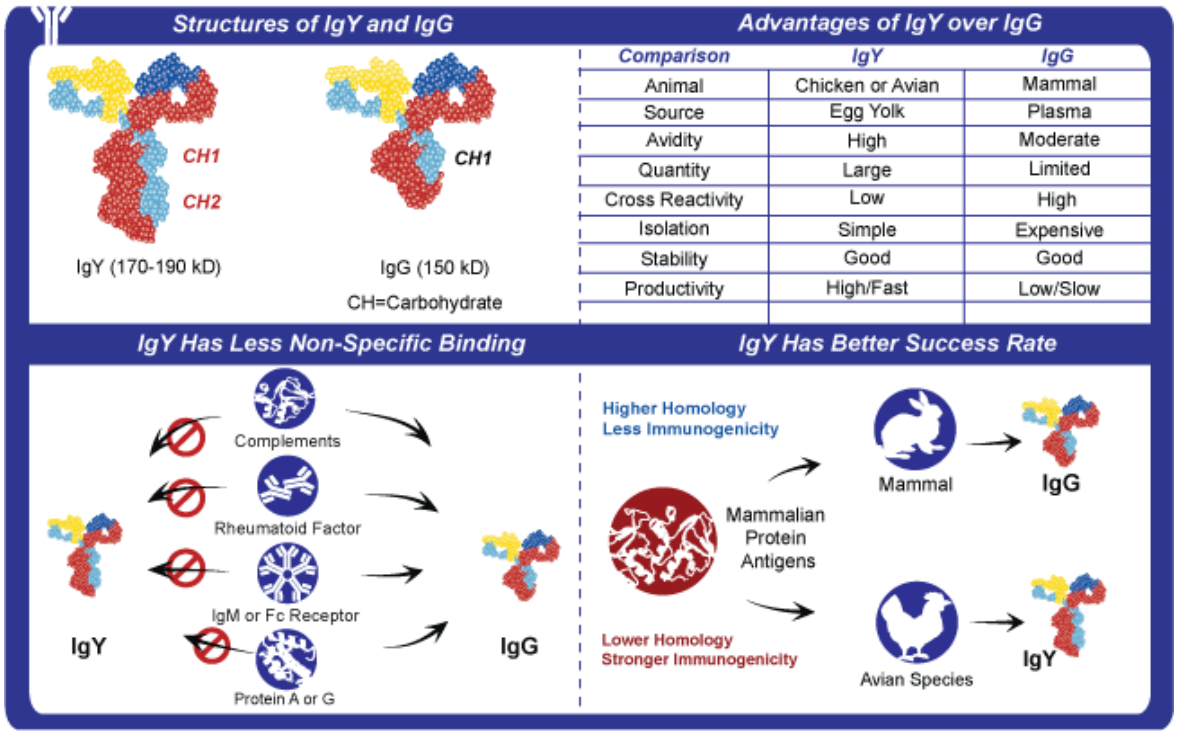


Figure 7: Some of the differences between IgG and IgY. (110).

Selection of the immunogen:

It recently emerged that the 2013 article by Kim et al (Newcastle Disease Virus Fusion Protein Is the Major Contributor to Protective Immunity of Genotype-Matched Vaccine) was retracted (110). This does raise several questions as to the efficacy of any F protein alone used as an antigen. However, it is one of the antigens used in the Innovax trivalent Vaccine (48). Additionally, challenge studies using recombinant GTVII F in several vectors showed efficacy when the birds vaccinated with these prototype recombinant vaccines expressing NDV F protein were challenged with live virus (112, 113).

Structural studies to date:

Structural studies of viral antigens are nothing new, with the influenza HA and the E protein of flavivirus being among the first (114, 115). To date only the structure of the post-fusion conformation of NDV has been determined: Chen et al were the first to crystallise and determine a partial structure of the head region, by X-ray crystallography (116). They missed a portion of the tail and transmembrane regions but were able to dissect the composition of the head region, see figure 8.

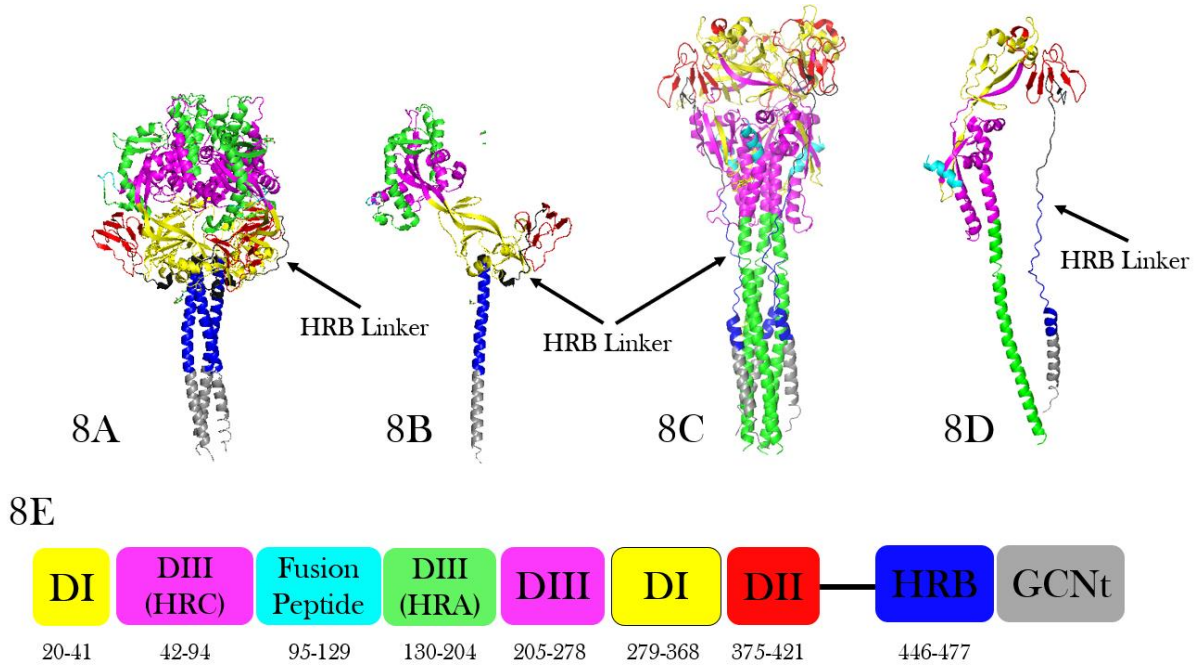


Figure 8A and 8B: The pre-fusion structures of PIV5 in trimeric and monomeric forms with each of the domains (DI, DII and DIII marked, along with the HRB domain (118). Figure 8C the post fusion structures of NDV generated by Swanson et al in 2010 (117) 8D: a monomer of Swanson et al's post fusion structure. 8E a diagram showing how each region of the peptide corresponds to the linear, unfolded peptide. The colours of each region in figure 8A-D correspond with those in 8E.

Swanson et al were able to build on this, accounting for the remainder of the protein and its trimeric nature (117). Their structure illustrates the trimeric nature of the protein and its characteristic 'Golf tee' shape. No pre-fusion structure of NDV F currently exists and it would be beneficial to fill this gap in the literature.

Each peptide that makes up the F protein has 3 principal domains, DI, DII and DIII (117), see figure 8. Additionally there are three heptad repeat (HR) regions, HRA HRB and HRC. A linker domain connects HRB to DII (124). In the post fusion structure of NDV the HRB domain of each monomer is a helix and forms the C-terminal end of the protein and anchors the protein into the viral membrane. Parallel to this is the much longer helix of the HRA domain which runs after the cleavage site, through the DIII domain. DI and DII form the head of the structure, with a number of β-sheets and connecting regions (123). A long HRB-linker connects the HRB at the C-terminus to the DII domain, marked on

figure 8. The HRC containing portion of DIII and a part of DI that is to the N-terminal of the cleavage site also forms part of the head domain. Three of these bundled together form the post-fusion F protein (116, 117).

Pre Fusion F structures:

Numerous other viruses with Class I Fusion Proteins have been solved in the pre-fusion conformation, including Respiratory-Syncitial-virus (RSV), Nipah virus, Hendra virus and the virus that shares the most homology to NDV: Parainfluenza V (PIV5) (118-21). The structures of these homologues are a useful guide to the solving of NDV, however PIV5 Fusion protein only shares 35% homology (by nucleotide) with that of Newcastle Disease. Despite this, it is predicted that they share much functional homology. Influenza HA and SARS-CoV Spike are class I fusion proteins, but distinct in look and function (114, 122).

In PIV5 (and most class I fusion proteins) each monomer, within the trimer, in the pre-fusion conformation is integrated into the viral membrane by a trans-membrane domain, at the C-terminus of the protein. This extends up into the helical HRB domain see figure 8A and 8B (116). The HRB-linker protrudes off at approximately 90° and connects to the DII domain, which consists of a number of β-sheets. This is connected to part of DI, which sits atop the HRB domain and also consists of β-sheets and some less ordered loops. DIII is connected to the N-terminal side of DI (the part already discussed) and folds back on itself. A second part of DI is at the N-terminus of DIII, anchoring the monomer together. DIII principally consists of α-helical bundles, connected with loops, which extend to form the HRA domain during fusion (118). This could be compared to the erection of a tent-pole. The three domains together form a roughly crescent shape, from which the HRB domain protrudes. Three of these monomers assemble into a globular trimer, with the concave part of each crescent nestling into the convex aspect of its neighbour.

A key part of DIII is the cleavage site, which forms a loop to the exterior of the trimer, accessible to cellular proteases. Once cleaved, the hydrophobic domains are unsheathed. The mechanism of re-folding to fuse the two membranes was also suggested by Hsien-Sheng et al, Harrison et al and others (123, 124). This involves the α-helical bundles of Heptad repeat A in DIII extending out and form an

intermediate rod, affixed to the viral envelope at one end and the cell membrane at the other. DII functions as a hinge domain, through which, mechanical force is applied to the cellular membrane and allows the viral envelope to fuse with (in the case of NDV) the membrane of the endosome. This releases the viral genome and its accessory proteins into the cellular cytoplasm (125).

Another class I fusion protein of note is the Spike protein (S) of SARS-CoV-2, a protein that would have a profound impact on the project in multiple aspects (121). While there are many differences between SARS-CoV-2 S and NDV F (S is larger and interacts with a receptor directly through the S subunit S1), many of the same difficulties presented by S are also presented by F. It is only recently that structures of S were produced, many of these leveraging the affinity of S for the (human) ACE2 receptor to create complexes which are easier to work with in Cryo-EM (126-8). Additionally, there have been numerous antibodies reported that bind strongly that could be leveraged as FAB for structural work. Independently, Walls et al and Wrapp et al were able to solve the structure of SARS-CoV-2 S using cryo-EM to 3.5Å, with Lan et al managing to crystallise the protein bound to ACE2, and solving the structure to 2.4 Å (126).

Summary:

Through the solving of these structures, a better understanding of how they function is possible, allowing for the development of more stable antigens for use in vaccines. With both PIV5 and RSV, various stabilising mutations were tried to create a protein less likely to shift into the post-fusion conformation quite so readily (See chapter III for more information) (116, 117). Such a protein would be much more thermostable and may improve the efficacy of vaccines, as outlined in the work of Crank et al (129). Additionally, the current vaccines for NDV available commercially do not offer complete protection against some of the more divergent genotypes. Little is known about which epitopes are of import, and which- if any- are common between multiple genotypes. A basic structural understanding of the NDV F protein could pave the way for work on epitope mapping and thus the development of a far more effective vaccine.

Hypothesis:

Through determining the structure and mapping epitopes of antibody binding and neutralisation, it will be possible to design a suitable antigen for use in a recombinant, vectored vaccine against the GTVII strain. Knowing where the most neutralising antibodies bind will allow for the vaccine to be made more effective, through structure-based vaccine design. It is predicted that antibodies targeting the apex of the pre-fusion conformation will be more effective, as in the case of RSV.

To ensure that the maximum number of F proteins are pre-fusion, they will need to be engineered to become unable to refold to the post-fusion, increasing the number of anti-pre-fusion antibodies. Stabilisation could also make the vaccine more resistant to the rigours of field use, and the higher core temperature of chicken's bodies. After this proof of principle, other emergent strains of NDV could also have specific vaccines designed in this manner. It would also pave the way for understanding the structural basis for the lack of cross-protection between genotypes.

A note on the impact of the pandemic:

In many ways, the pandemic has severely impacted the completion of my project. At the onset of the first lockdown I was excluded from labs for five months and able to achieve very little. When I could resume practical work as part of the DPhil programme I then had move to the Pirbright Institute, where the necessities of social distancing meant that I worked with restricted hours with minimal direct supervision. Between this and the pressure to 'make up for lost time' I very rapidly burnt myself out, for relatively little gain.

Chapter II

Materials & Methods

A note on who did what during the project:

During the course of the project, I received some invaluable assistance:

Dr Guido Paesan designed and ordered the constructs that were used during the project. Dr Karl Harlos harvested and froze the protein crystals, while Dr Mattia Bertinelli operated the data collection at Beamline I03 remotely for the x-ray crystallography. Dr Nathan Cowieson operated Beamline 21 for the SAXs data collection. The negative stain EM data-set was collected by Dr Melissa Webby, while the data-sets for cryo EM were collected by Dr Jeremy Keown and Dr Loic Carrique. The Chameleon was operated, and the dataset collected by Dr Dave Owens. The immunisation of the chickens, collection of blood samples and ELISA mentioned in Chapter V was carried out by Dr Yaoyao Zhang. A very sincere thanks to those that assisted in this project.

Construct design:

The two principle NDV F constructs used in this work are an attenuated genotype VII construct (GTVII) and a La Sota construct (LaS), GenBank accession numbers: KY073880.1 and AJ629062.1 respectively. The gene for the F protein from each strain is preceded by a GP64 export motif, allowing the protein to be exported into the media, as was used by Wong et al in their work on the Hendra Fusion glycoprotein (1). The trans-membrane-domain has been replaced by a HIV GCNT domain to lock the structure in the pre-fusion conformation, as done in the work on PIV5 (2, 3). A 6-His tag is on the C-terminus to aid purification. The construct is in a pOPIN-G plasmid, which confers ampicillin resistance (4). The GTVII construct was attenuated by mutating the polybasic cleavage site to a tri-basic cleavage site because there is a possibility that these constructs will be introduced into live chickens, later-on. The risk of a genetic recombination event was deemed too dangerous to use the virulent

cleavage site. These constructs were created and ordered from GeneArt by Dr Guido Paesan, prior to the start of the project, and can be found in Appendix B.

Modelling of the F protein:

The amino acid structure of the F protein was inserted into Swiss-model and later Alpha-Fold (5, 6). In both cases the amino acid sequences were copied and pasted into the server and the model was generated. In the case of the former, the PIV5 PDB sequence was used as a template (2), while in the latter the structure was created de-novo. Furthermore Swiss-model was able to build the model as a trimer while Alpha-Fold built as a monomer. The PIV5 structure was then used as a template to assemble the trimer: a local alignment was performed with each monomer of NDV F to a PIV5 F monomer, the latter was then subtracted.

Confirmation of protein presence:

A 6-well plate of HEK-293T cells, seeded at $1.3x10^6$ cells/ml was transfected using Lipofectamine (Product number 18324012) with 1ug of the GTVII and LaS protein constructs and incubated for 48 hours at 37°c. The cells were harvested with 0.05% Gibco trypsin EDTA (Product no 11580626), hereafter referred to as 'trypsin.' The cells were lysed with lysis buffer. Presence of His-tagged protein was confirmed by Western Blot, using a NuPage Bis-Tris 12% gel, (catalog no: NP0341). Following transfer to a membrane, and blocking in 12% milk powder in PBS, the membrane was incubated with mouse anti-6-His, incubated for 1 hour at room temperature, and the secondary antibodies were HRP-tagged goat anti-mouse, incubated overnight at 4°C.

Protein production:

Competent DH5α cells were transformed with the LaS or GTVII pOPIN TTG plasmids by heats-shock and grown for 24 hours in a starter culture of 10ml LB with 100 μg/ml ampicillin at 37°C shaken at 180rpm. This was then used to inoculate 10x 100ml flasks of LB with 100 μg/ml ampicillin and incubated for 24 hours at 37°C, shaken at 220rpm. The culture was spun at 4000x g and media discarded. The plasmid was extracted from the pellet using a Qiagen Giga kit (Cat No./ID: 12391).

Meanwhile HEK-293T or 293S cells were seeded at 2.5x10^6 cells/ml in T75 flasks, one per roller bottle desired. In practice, this meant 5. The media used was 10% FBS DMEM with 1%NEAA and 1% L-Glut. These were then incubated at 37°C for 48 hours before the cells were detached using trypsin

These were re-seeded in a 2300ml ribbed roller bottle in 200ml 2% DMEM 1% L-Glut, 1% NEAA. These were incubated in a roller incubator at 37°C for 24 hours.

0.5mg of plasmid DNA was used to transiently transfect each roller-bottle of HEK cells, using 6mg of PEI, then 2.25mg of kifunensine (a mannosidase I inhibitor) was added (7). 293S cells were sometimes substituted for 293T cells as they lack α-mannosidase (8). When 293S were used, no kifunensine was added. As part of one condition during the deglycossylation experiments, the kifunensine was omitted from a 293T culture. Both result in protein expressed with shorter, simpler glycans. See the introduction to chapter III for further details.

Expression of F protein in SF9 Insect cells:

Expression was carried out in SF9 insect cells infected with baculovirus. SF9 cells in 1.5% Serum Gibco SF900 media (product no 10902088) were transfected with the pOPIN ttg plasmid containing the La Sota F construct, using Celfectine II (Thermo: Prod no.13062-100). This was carried out in a 6-well plate with 8x10^5 cells/well. After 5 hours incubating at 27°C, the media was replaced with 10%FBS SF900 media and the cells and incubated at the same temperature for 72 hours. After this, 30µl of Baculovirus (grown in house) was added to each well.

24 hours-post-infection the media from these infected wells was added to a 20ml culture of SF9 cells in 10% FBS SF900 media at a density of 5x10^6 cells/ml and incubated at 27°C shaking at 180 rpm for a further 24 hours. After this, the cells were passed into a larger flask with 50ml of media for a further 24-48 hours. Once confluent, the media was collected and spun down at 1000xg for 10 minutes. The media, containing baculovirus virions was collected and filtered by a 0.2 micron filter for storage at 4°C.

When expression was required, a 35ml culture of SF9 cells at 5x10^6 cells/ml was infected with 1ml of the stored virion containing supernatant. After 24 hours this starter culture was divided between 5 larger flasks containing 100ml of SF900 media and incubated at 27°C shaking at 180 rpm for 72 hours.

Protein Purification:

After 4 days incubation (Mammalian expression) or 3 days (SF9), the media was collected, spun down at 4000x g and filtered to remove cell debris using a 0.2 micron vacuum filter. A Watson-Marlow Quickstand was used to buffer exchange the DMEM or SF900 for pH8 0.15M Tris 0.15M NaCl solution, then to concentrate the solution to 200ml. The column was a GE Healthcare Hollow Fibre Cartridge (model no. UFP-10-C-4X2MA) with a 10,000 mw cut-off. For every 100ml of media harvested, 200ml of Tris-NaCl buffer was added to the Quickstand.

Following buffer-exchange, the protein was then loaded onto a GE 5ml His-Trap (product no. GE17-5255-01) using an Äkta side pump. A wash step of 50ml of pH8 0.15M Tris 0.15M NaCl was carried out. The protein was then eluted on an Äkta using a 50% split of 1M Imidazole 0.15M Tris 0.15M NaCl pH8, and 0.15M Tris 0.15M NaCl pH8, into a 96 well plate fractionator in 0.5ml fractions.

The protein fractions were then assessed for purity on a protein gel, before being concentrated to 0.5ml aliquots at no greater than 8 mg/ml and injected into an Äkta Pure for gel filtration with pH8 0.15M Tris 0.15M NaCl pH8 solution, using a GE healthcare Superose 6 10/300 (product no. 29-0915-96). The loop size was 500ul. Everything that passed through the column was collected in a 96-well plate.

The protein was then concentrated using a 100k Amicon Ultra-4 centrifugal filter (product no. UFC810024) in swing out buckets, spinning in increments of 5 minutes at no more than 2000 x g.

Later a Cytiva Superose S200 10/300 (product no17517501) was used as it was able to distinguish between monomeric and trimeric forms of the protein. Additionally, size exclusion was subsequently carried out on an Äkta Pure in a 4°C cabinet. When this column was not available, a S200 16/ 600 (product no 28-9893-35) was used in lieu.

Enzyme digest:

Aliquots of protein were digested with de-glycosylating enzymes, NEB Endo H (product no. P0702S) and NEB PNGase F (product no. P0704s), at 1µl of enzyme per 20µg of protein for 12 hours at 4°C during the small scale deglycosylation. This was increased to 2µl of enzyme per 20µg of protein

incubated overnight at 4°C during the larger scale deglycosylation, for setting up crystallisation trays. Size Exclusion Chromatography was used to remove the enzyme, before being plated on crystal trays, so as to not interfere with the homogeneity of the sample.

SAXS:

GTVII and LaS F samples were prepared at 5mg/ml in 0.1M Tris 0.15M NaCl solution and also in pH5 0.1M citrate buffer and taken to Diamond Light Facility (9), where 25µl was loaded on to the superpose-6, as above. Size-exclusion-chromatography-coupled-Saxs was performed with matched buffers and the data extracted. Analysis was performed using Chromixs (10). During the homology modelling stage, it was compared with the known post-fusion structure of NDV (PDB: 3MAW). Furthermore, comparisons were made with models of NDV F protein in the pre-fusion confirmation, derived from Alpha-Fold and also with pre-fusion PIV5 (PDB:4WSG)(2).

Raising Nanobodies:

8 aliquots of LaS F at 1mg/ml in PBS were sent to Chinese collaborators and used to immunise an Alpaca. PBLs were collected and the nanobody RNA extracted and were made into cDNA. See Appendix B. More details are not available, at present.

Crystal-tray preparation

Greener Bio-One Masterblock 96-well crystal-trays (product no. 780285) were set up, the reservoir being loaded by a Robbins Scientific Hydra 96 machine. A 100nl drop of protein was dropped on to the centre well, and 100nl of the crystallisation solution from the reservoir being spotted on-top of this by a Cartesian technologies Cartesian machine. Trays were then placed into a Formulatrix Rock Imager and imaged periodically. Different conditions were set-up, with protein at different concentrations and at both 4°C and 25°C. See Chapter III for a list of all the conditions explored. When an additive was used, See Table 2 in Chapter III) this was added to the protein aliquot immediately prior to being loaded on the Cartesian.

X-ray crystallography:

The protein crystals were harvested into cryo-mounts and flash frozen in liquid nitrogen. These were loaded onto pucks and transferred to the Diamond Light Source, Beamline IO3, where they were imaged remotely from within STRUBI. All the crystals were imaged, but only one diffracted to 4.5 Å. The beam width was 20x20µm with a wavelength of 0.98 Å was used to collect 3600 images with an exposure of 0.05s. The data was processed with XIA2Dials and PHENIX (11, 12). See Chapter III.

Hampton Additive and Thermostability Screen:

A Hampton Research Additive Screen (Product no HR2-138, see website for full list of additives) was carried out using 2µg of protein, 3µl of dye (product no ThermoFisher Scientific 4461146), and 10µl of additive to a total well volume of 50µl. The plate was then put into a qPCR machine. The temperature was then increased in a stepwise, serrated manner, by 1°C each time, returning to 25°C to measure absorbance after each increase in temperature, as indicated by the product information. Absorbance was measured at 550nm.

Negative stain EM

Grids were prepared with varying concentration of protein, but 0.05mg/ml provided a good level of coverage. Grids used were Agar Scientific Formovar/carbon 300 mesh copper Product no (S162-3H). Grids were glow discharged, using a Harrick Plasma Plasma cleaner (M-PDC-32G) under high power for 45 seconds. 4ul of protein was added to the grid and left to absorb for 30 seconds. The protein solution was then blotted by hand, touching a corner of blotting paper against the edge of the grid. The grid was then touched (face down) against a 50ul droplet of deionised water, and blotted as described above, twice, using a fresh droplet each time. The grid was then touched against a droplet of Uranyl Formate for 20 seconds and blotted, twice, in the manner described for the water. The grid was then left to air-dry. The imaging of the negative stain grids was performed by Dr Melissa Webby.

Cryo-EM:

Protein was diluted to concentrations between 1 mg/ml and 0.2 mg/ml and loaded on to, 200 mesh, Quantifoil gold grids (product no. UF22200Au50). Later, other grids used included: gold UltrAuFoil R1.2/1.3 300 mesh gold grids, QUANTIFOIL R1.2/1.3 copper 300 mesh grids. The former were glow discharged for 1 minute on high power, while the latter were glow discharged for 45 seconds. The prepared grids were placed into a Vitribot at room temperature with the humidity setting set to 100%. 3.5ul of protein was pipetted onto the grid and blotted for 3.5 seconds at –25 blot strength, on both sides of the grid. In an attempt to solve the problem of the protein detrimerising during grid-preparation, other methods of grid preparation were used: a Leica Microsystems GP2 was used to prepare grids due to its ability to back-blot only. Blot times used were 8 and 10 seconds. The same humidity and volume of protein were used.

These were plunged into liquid ethane and 'clipped,' before being stored in liquid nitrogen. They were screened a 200kv Glacios Cryo-electron-microscope. The 0.6 mg/ml grid had a better coverage, and was in better condition, so was screened: see chapter IV for the screening parameters. The data was processed using cryoSPARK (14). Subsequent screenings were carried out on both the Glacios and Kryos, using various combinations of additives. See Chapter IV.

Uncle thermostability assays:

A series of protein stability assays were carried out using the Unchained UNcle (13). Aliquots in triplicate of each condition to be tested were prepared at 0.3mg/ml, 0.6mg/ml, 1.2mg/ml and 2.4mg/ml. 9ul of each condition was pipetted into the 'Uni's wells. The Uni was then encased in the cassette and loaded into the UNcle. A Dynamic Light Scattering (DLS) at 660nm reading was taken, then a temperature-ramp then began at 25°C, culminating at 95°C. At each stage of the ramp, a fluorescence and Static Light Scattering (SLS) reading was taken at both 266nm and 456nm. A DLS reading was taken at the beginning and end of the run. Post-use the Unis were cleaned in a 1% Hellemanx soulution, in a heated sonicating water-bath at 55°C for one hour. The additives, NDSB 195, NDSB201, NDSB256, L-arginine, L-glutamate and trehalose were at final concentrations of 50mM.

Chameleon:

La Sota F protein at 4mg/ml, grown in SF9 cells, was transported on dry ice to Diamond Light Source and handed over to the Chameleon beamline scientists (15). This was then used to prepare grids at 4mg/ml, 2mg/ml and 1mg/ml. These were screened, and data collections set up, as detailed in chapter IV.

Cross-linking

HEK293T expressed F protein was purified by size exclusion on a Superdex S200 10/300, while buffer exchanged into PBS at pH8. Fractions collecting the trimeric protein were collected and diluted to 0.1mg/ml. A gradient of gluteraldehyde dilutions (product no 354400) was created and incubated at room temperature for 1 hour. Subsequently a series of time-course experiments were run with 0.12% glutaraldehyde and dimethyl pimelimidate (DMP) (Thermo Fisher Product Number: 21666). All were quenched by adding 1M Tris to a final concentration of 100mM. Following cross-linking, a further round of size exclusion was carried out using the Superdex S200 10/300 equilibrated with 50mM Tris 150mM NaCl pH8 and the fractions collected.

Pegylation

SF9 expressed F protein was purified by size exclusion on a Superdex S200 16/600, while buffer exchanged into PBS at pH8. Fractions corresponding to trimeric F protein were collected and concentrated to 1mg/ml.

3ul of methyl-PEG4-NHS per 250ul of F protein was added to the solution and incubated for 60 minutes on ice. During the incubation, the Akta and S200 were re-equilibrated with 50mM Tris 150mM NaCl pH8. The reaction was quenched with 25µl of 1M Tris per 250ul of protein. The protein was loaded on to the Akta and purified as previously described.

Preparation of the gene for insert into HVT

The plasmids containing both LaS and GTVII with GCN(t) motifs were transported to the Pirbright Institute on ice. Genes lacking the GCN(t) domain were also synthesised by Eurogene (see Appendix

B). Primers were designed to add a NotI restriction enzyme site to the 3' and 5' ends of the genes (see Appendix B). All 4 constructs of NDV F were cloned in parallel.

A PCR was run using NEB Q5 High fidelity Polymerase (Product number M0491S), with relevant dNTPs, denaturing for 15 seconds at 94°C, annealing for 30 seconds at 58°C and elongating for 1 minute at 72°C, for a total of 35 cycles. Each of the four conditions were run in duplicate 50ul reactions. A 1% agarose gel was cast, the PCR products mixed with 10ul of Thermofisher 6x loading dye (product no R0611) and run at 100eV for 40 minutes. The gel was imaged on a Biorad gel imager, before the bands were excised by scalpel on a UV light box.

A Quiagen Gel extraction kit (Product no 28706X4) was used to extract the DNA from the agarose gel. NEB NotI (R0189S) was used to digest the gene fragments, and also the pCDNA3.1 vector for 1 hour at room temperature. The vector was then prepared for the gene to be ligated with the addition of CIP (M0290) to digest for a further hour.

On completion of the enzymatic digest, a further 1% agarose gel was cast and run, as described above, the bands being excised and the DNA extracted in the same manner. T4 DNA Ligase (M0202S) was used to ligate the F protein genes into the pCDNA3.1 vector, in a 20ul reaction, for 10 minutes at room temperature. The ligated construct was then transformed into cloning efficiency DH5α cells by heat-shock, incubated in SOC medium for one hour in a shaking incubator at 37°C, before being plated onto ampicillin agar plates, at 37°C overnight.

Where colonies formed, the colonies were picked and amplified in a 5ml culture of LB, with 5ul of Ampicillin at 1mg/ml in a 37°C shaking incubator overnight. 5 hours post-inoculation, 5ul of LB was siphoned from each culture, and used to set up a 20ul PCR reaction using DreamTaq Green PCR mix (product no K1081) in order to screen for successful insertion. A 1% agarose gel was run, as previously described, using the products of this diagnostic PCR.

Cultures that were positive for the inserted genes by PCR were collected, centrifuged at 4000xg and miniprepped, as per the manufacturer's instructions. The gene fragment in the pCDA3.1 vector was then subjected to digest with SfiI (R0123S), which would extract the F protein gene from the vector,

flanked by sequences necessary for the subsequent insertion into HVT. In parallel with the SfiI digestion of the gene, the new vector: PGEM-teasy-Vp2-GFP was also digested with SfiI. The digested products were run on a 1% agarose gel, excised and extracted as previously described. The gene and new vector were ligated, transformed, incubated in SOC and plated out on ampicillin-agar, as-above. Colonies were picked into 5ml LB cultures and screened for presence of the inserted gene, with the positive cultures being miniprepped, as previously described.

Freshly isolated primary Chicken Embryonic Fibroblast (CEF) cells were plated into 24 well plates at $1.5x10^5$ Cells/ml, and grown with CEF growth medium. When confluent they were transfected with 1ug of the F in PGEM-teasy-Vp2-GFP using Lipofectamine, as per the manufacturer's instructions, diluted in OptiMEM (product no 31985062).

Immunofluorescent staining:

After 24 hours, the cell media was aspirated and the cells were washed with PBS, before being fixed in 4% paraformaldehyde (PFA) for 20 minutes at room temperature. Mouse Anti-His IgG was diluted in PBS and incubated with the cells for 1 hour. After washing with PBS, RFP-tagged goat, anti-mouse IgG was incubated for a further hour and imaged under a fluorescent microscope. Constructs, when transfected into CEF, positive for both GFP and RFP were selected for insertion into HVT.

Insertion into HVT

A 12-well plate was seeded with fresh CEF cells at $4.1x10^5$ Cells/ml. Once 80% confluent they were transfected with 1ug of the new F Protein constructs, with 1ug of a plasmid containing the guide RNA sequences and 1ug of the CRISPR-CAS9 construct (see Appendix B), using TransIT X2 (Product no MIR6003).

Later, electroporation was used in-lieu of X2 to insert the plasmid into the cells. CEF from a confluent T75 flask were harvested with trypsin, counted and resuspended at $1x10^7$ Cells/ml in electroporation buffer. 1ml of cells were pulse electroporated at 100v for 5 seconds with 1ug of each of the F construct in the PGEM-Teasy-VP2-GFP plasmid, the guide RNA and the CRISPR-CAS9 construct. After

electroporation they were mixed with 2ml of CEF Growth medium, plated in 6-well plated and incubated for 24 hours at 37°C.

6 hours post-transfection (X2) or 24 hours (electroporation), when the GFP was being expressed in the cells, they were infected with 50ul HVT /well (Passage no 2), equating to 0.01 Plaque-forming-units/cell, and incubated for 24 hours. Meanwhile one 6-well plates per construct was seeded with CEF as previously described.

Following the 24-hour period, the media in the infected plate was aspirated and the cells detached with 400ul trypsin. The detached cells were placed in an Eppendorf and the trypsin quenched by adding 50ul serum. This was spun down at 200xg and the supernatant aspirated. The pellet was gently resuspended in 400ul of CEF Growth media. 50ul of the resuspended cells was added to each well of the CEF in 6 well plates and incubated for 48-hours at 37°C.

36 hours into the incubation, the plates were inspected for the formation of GFP expressing viral plaques under the fluorescent microscope. A well of a 24-well plate was seeded, plus a few for fudge-factor, was seeded with CEF, as above.

At the 48-hour mark, a 20ul pipette was used to detach each fluorescent viral plaque individually into an Eppendorf containing 40ul of trypsin. This was neutralised with the addition of 100ul of CEF growth medium, and pipetted up and down to break up the plaque. The contents of the Eppendorf were then added to a well of the 24 well plate seeded at the 36-hour mark.

The 24-well plates were returned to the incubator for 24 hours. Wells with GFP positive plaques were then passed to 6-well plates freshly seeded with CEF, as previously described. Once confluent, approximately 36 hours later, the media in the 6 well plates was aspirated, washed with PBS and 1ml Trypsin used to detach the cells. The cells from each well were transferred to an Eppendorf, the trypsin neutralised with 50ul of FBS. This was spun at 200xg for 5 minutes on a bench-top centrifuge, the supernatant discarded, and the pellet resuspended in 1ml PBS with 1%FBS. A haemocytometer was used to count the cells in each Eppendorf. More 1% FBS PBS was added-where necessary to dilute the cells to $1x10^6$ cells/ml. The cells were then pipetted into a polystyrene sorting tube through its integral

strainer-cap. The tubes of cells, each with two 96-well plates seeded at 4.5 x10^4 cells/ml was sent for cell sorting, with GFP-positive cells being sorted into the 96-well plate. Upon return of these plates, they were placed into the 37°C incubator for 3 days.

Each GFP-positive well of the 96-well plate had the plaque removed using 50ul of trypsin, neutralised with 50ul of CEF growth medium. This was then added to a well of a freshly seeded 6-well plate and incubated for 3 days. At this point, the wells were then harvested, 100ul set aside, and added to 1ml of freezing medium. The main aliquot was then frozen at –80°C, while the cells set aside were spun for 5 minutes at 200xg, the supernatant discarded, and the pellet frozen at -20°C.

The pellet was defrosted and resuspended in 50ul of squishing buffer (see appendix 2), heated at 65°c for 30 minutes, then 95°C for 2 minutes. A 20ul PCR reaction was prepared using 1ul of the DNA from the squished cells, 10ul of GoTaq 2x PCR mix, 0.5 of the junctional primers (see appendix B), and 8ul of nuclease free water. It was amplified at 95°C for 2 minutes- then a further 30 seconds, 55°C for 30 seconds, 72°C for 40 seconds- for 35 cycles, followed by 72°C for 7 minutes. The product was then loaded on a 1% agarose gel and run at 100mv for 40 minutes.

Based on the results of the PCR, constructs with the F gene properly inserted into the virus had a well of a 6-well plate seeded with CEF. Once confluent, X2 was used to transfect the cells with 2ug of plasmid containing Cre-recombinase. 12 hours post-transfection, 50ul of defrosted HVT with inserted F was used to infect each well. This was then incubated for 3 days 37°C, before being prepared for FACS into 96-well plates, as previously described, however the non-fluorescing cells were collected this time, on account of the GFP gene being removed by the Cre-recombinase. 3 days later, non-fluorescent viral plaques were picked from the 96-well plates into 6-well plates (as previously described) and incubated for 3 days. The contents were harvested, using trypsin and prepared for freezing at –80°C, less a 50ul aliquot used to infect a well of a freshly seeded 24-well plate, in triplicate. Meanwhile uninfected wells were left thus or transfected with the PCDA3.1 plasmid containing the inserted F gene, for use as a control, with lipofectamine. After incubating for 48 hours, the cells were fixed with 4%PFA, and immuno-stained as previously described.

Animal Experiment and ELISA:

F protein aliquots, expressed in HEK293 T cells, flash frozen in liquid nitrogen, were transported to the Pirbright institute, thawed and used to immunise 8 SPF Day old chicks on day 0. Each dose consisted of 50ug of F protein and 70ul of inactivated oil adjuvant. A Negative control group of 8 SPF chicks received saline and adjuvant only. A positive control received the La Sota B1 vaccine.

Booster doses were given at days 14 and 28. At each timepoint wing pricks were used to draw blood. A standard lympho-prep was carried out and an ELISA was performed using an ID Screen Newcastle Disease Indirect ELISA Kit, Product no: NDVS-10P. This experiment was carried out by Dr Yaoyao Zhang of the Pirbright institute.

Expression of nanobodies at small scale:

Sequences of three non-neutralising nanobodies were received from our collaborator and synthesised with a C-terminal V5 tag, in a PGEM vector (16). See Appendix B for sequences.

Freshly isolated CEF cells were plated into 6 well plates at $1.3x10^6$ cells/ml. When confluent they were transfected with 2ug of nanobody plasmid, using Invitrogen Lipofectamine (product no: 18324012), as per the manufacturer's instructions.

After 48 hours incubation, the cells were fixed, as described above and stained with a mouse- antiV5 – antibody for 1 hour at room temperature. After washing it was incubated with a RFP tagged goat-anti-mouse antibody at 4°C overnight.

A second 6-well plate of CEF transfected with the nanobody constructs (VHH) was prepared as above. After 48 hours incubation, the cells were harvested and lysed using lysis buffer, for 30 minutes on ice. The lysate was then denatured at 92°C, in LDS (reducing) loading buffer and loaded on a NuPage Bis-Tris 12% gel for 1 hour at 100V. Following transfer to a membrane and blocking in 10% milk-PBS, the membrane was incubated with the anti-V5 antibody diluted 1:250 in milk-PBS for an hour at room temperature, followed by HRP-tagged goat-anti-mouse secondary antibody at 4°C overnight.

A second round of expression was carried out, with the 3 VHH constructs being expressed in the wells of 1x 6-well plate, and LaS in the PCDA3.1 plasmid in another (as previously described). After 48 hours incubation at 37°C, the media was aspirated and retained. The cells were washed 3 times with PBS and fixed with 4% PFA as previously described. The media from the wells transfected with the VHH constructs was added to the wells transfected with LaS F, while the media from the wells transfected with LaS was incubated with the wells transfected with the VHH constructs. Both incubations were 1 hour at room temperature.

The media was then aspirated and the cells washed with PBS. The wells transfected with F constructs were stained with a primary antibody of mouse anti V5, while the constructs transfected with the VHH constructs were stained with the mouse-anti6xHis for one hour at room temperature. Both used a RFP tagged goat-anti-mouse secondary, overnight at 4°C.

Immuno-precipitation of nanobodies:

6-well plates of CEF were transfected with each of the nanobody constructs, as previously described, and incubated for 3 days at 37°C. At this point the cells were harvested and lysed using lysis buffer for 30 minutes on ice. The lysed cells were spun for 10 minutes at 13,000Xg in a bench top centrifuge, cooled to 4°C. The supernatant was then collected, and mixed with 100ug of purified F protein, (frozen after purification by size exclusion). This was incubated overnight at 4°C.

50ul aliquots of Magnetic Protein G Dynabeads (Thermo Fisher Scientific product no 10003D) were prepared, as per the manufacturer's instructions and incubated with 2ul of the mouse-anti-His antibody, diluted in 200ul of PBS for 10 minutes at room temperature. The beads were pelleted by magnet, supernatant aspirated and washed in PBS-0.02%tween 20, re-pelleted and the wash aspirated thrice. 200ul of F protein-VHH mixture was then added to the Dynabeads. The two were incubated for 10 minutes at room temperature on a rotating stand. Following this, the beads were washed 3 times, as previously described. Elution was carried out using 30ul of 50mM glycine pH2.8, rotating for 2 minutes at room temperature. The beads were pelleted by magnet and the supernatant added to 2ul of 1M Tris. 10ul of LDS reducing buffer was added and warmed to 92°C for 5 minutes. The samples were then

loaded onto a 12%Bis-Tris Gel, and a western blot carried out as previously described, in this case, with the anti-V5 antibody serving as the primary antibody.

The immunoprecipitation was also carried out the other way around with anti-V5 antibodies bound to the Dynabeads and the anti-His as the primary antibody.

Cloning nanobodies for large scale expression

A pMESy4 vector was used to express the nanobodies. Primers (see appendix B) were used to amplify the nanobody gene and add a PstI site at the 5' end and a 6-His-tag at the 3' end. The PCR program was denaturing for 15 seconds at 94°C, annealing for 30 seconds at 65°c and elongating for 1 minute at 72°C, for a total of 30 cycles. The product was run on a 1% agarose gel, the band excised and extracted from the gel as previously described. A second PCR was run, using a different reverse primer (See Appendix B) to add the remainder of the His tag and a BstEII site. The product was run on an agarose gel and extracted as above.

The PCR products and the pMESy4 vector were digested for one hour with PstI-HF and BstEII-HF (product nos R3140S and R3162S respectively), with the vector having rSAP (product no M0371S) added for a further hour of digestion.

Following the digest, the DNA was run on an agarose gel, the bands extracted, and the two ligated at a ration of 1ng of insert to 3ng of vector, using T4 DNA ligase, as previously described. The product of the ligation was then transformed into cloning efficiency DH5α cells by heat-shock, incubated in SOC media for 1 hour at 37°C before being plated onto 1% Ampicillin agar plates. After 24 hours of incubation at 37°C, the plates were inspected for colonies. Colonies were picked into small scale LB cultures as above. The following day the cultures had the DNA extracted by mini-prep.

A PCR was carried out using the primers used to amplify the VHH construct in order to check insertion of the sequence using the parameters stated above. The constructs that did have the sequence properly inserted were then sent for sequencing by the Europhins Light-run service.

Large scale expression of nanobodies

BL21 (Sigma-Aldrich: CMC0014) and Rossetta (Novogen-Sigma-Aldrich: 70953-3) expression cells were used for large scale expression of the nanobodies. The cells were transformed as previously described, incubated for an hour with SOC medium at 37°C, before being plated onto LB-ampicillin plates fortified by glucose at 100ug/ml and returned to the incubator overnight. The next day colonies were picked into 10ml solutions of LB with 100ug/ml of ampicillin supplemented with 1ml of 20% glucose and 5ul of 2M $MgCl_2$, incubated overnight at 37°C, shaken at 200RPM.

5ml of the starter culture was inoculated into 500ul of TB augmented with ampicillin at 100ug/ml, 0.1% glucose, 2 mM MgCl2, 0.4% glycerol in 2L conical flasks, shaken at 180rpm at 37°C. When the cells grew to an OD-600 of 0.7, they were moved to 4°c for 30 minutes to cool down, before being induced with 500ul of 1M IPTG. The flasks were returned to the incubator, once it had been cooled to 27°C, at 180rpm, overnight.

The next day the cells were spun down at 4000x g for 10 minutes and the supernatant discarded, with each pellet being resuspended in 15ml of cooled TES buffer, and left at 4°C, stirring, for 1 hour. To administer the osmotic shock, 30ml of diluted TES buffer (1 in 4) was added to the resuspended pellet and left stirring for a further 45 minutes at the same temperature.

The mix was then centrifuged at 8000 x g for 30 minutes at 4°C, the pellet discarded, and the supernatant retained. A His trap attached to an Akta was used, as detailed above, to extract the nanobodies for the supernatant. In a similar manner, a Superdex S75 10/300 (product no 29148721) was used to perform size exclusion.

Where nanobodies were mixed with F protein, it was always at a molecular ratio of 3:1, performed on ice for 1 hour.

Chapter III:

The hunting of the Snark

Introduction:

X-Ray Crystallography (XRC) has long been the standard method used to determine the structure of proteins, able to discern structures to very high resolutions. Not long after the discovery of NDV, the first diffractions of X-rays by protein crystals were being reported: Insulin in 1932 and myoglobin in 1934 (1). 1949 saw the use of X-ray crystallography to investigate the structure of penicillin, but it wasn't until the mid-1980s that the notion of 'rational drug design' began to appear (2, 3). The idea of designing better ligands for receptors within the pharmaceutical industry understandably spilled over into the development of vaccines, most famously, the search for a universal flu vaccine (4).

XRC also has the advantages of being less expensive than cryo-EM (at least recently) and is more capable of determining the structure of proteins smaller than 200kDa and can even solve structures smaller than the 50kDa limit of cryo-EM (5, 6). The development of synchrotrons, freezing the crystals (to reduce the impact of the radiation) and more efficient data collection have made it a more accessible and cost-effective (7). XRC has been used to solve many class 1 fusion proteins: Chen et al, then Swanson et al solved the post-fusion conformation of NDV F with XRC, though both were unable to form crystals of pre-fusion F (8, 9). In determining the structure of PIV3 it was noted that the protein being anchored in the membrane was necessary for the maintenance of the pre-fusion structure and that the HN protein mediated activation of F (10). However, Kim and Harbury noted that a GCNt domain (and its GCN4 derivatives), if placed at the C-terminal end of the HRB domain, could provide the same effect as the trans-membrane domain embedded in membrane (11). Furthermore, this domain is soluble, rendering expression and purification simpler.

This Leucine zipper was then applied to other class I fusion proteins, such as the work by Yin et al, then Welch et al on Parainfluenza 5 (PIV5) F (4WSG), McLellan's work on Respiratory Syncytial virus

(RSV) F (5UDE) and Xu et al's structure of Nipah F (5EVM) was all carried out using X-ray crystallography (See Chapter I) (12-15). Using some of the lessons from the solving of these structures, it might be possible to finally solve the structure of NDV F in the pre-fusion conformation.

However, it does have limitations: first and foremost of which is the necessity of crystalising protein. Obtaining crystals of the protein remains the principal hurdle to solving structures using this technique. Specifically, it has been reported that obtaining crystals of F-protein in the pre-fusion conformation is even more difficult than those of the protein in the post fusion conformation (8-10).

Use of the GCN(t) domain:

Many of the F proteins solved in the pre-fusion conformation made use of a GCN(t) domain in order to stabilise the protein in said conformation. This serves two purposes: Firstly, the approach was successful for generating the structure of the other class 1 fusion proteins. It makes the protein less likely to flip conformation so the sample is more homogenous and, thus more likely to crystalise (10-11). Additionally, should the project ever get to the stage of having a stabilised F protein being inserted into a vector, a protein stabilised in the pre-fusion conformation could be more immunogenic, as stated in the hypothesis.

In 2002 Yang et al reported using a trimerisation domain from T4 bacteriophage Fibrin to stabilise HIV GP41 glycoproteins, continued by Du et al, later being used in the solving of the PIV5 and RSV F proteins (12, 14, 16, 17). As such, it was chosen for use in this project. Specifically: Win et al reported that it forms a bundle of coiled-coils when attached the HRB domain of the protein, also serving to mimic the trans-membrane domain of the F protein (16). Removing the need for the trans-membrane domain also allows the protein to be expressed in solution. As-such, it was chosen to stabilise the F-protein with the GCNt domain, see chapter II.

Modelling

In recent years, with more and more structures of proteins being solved, it is increasingly feasible to use areas of homology between protein as templates for modelling other proteins of that class (18). Such was the approach of services like Swiss-model, which uses an extensive template library for homology

modelling. This approach, while it can be useful, could suffer from 'induced fit,' forcing the model to take a certain form, just because the template has those features (19). For a quick and rough model generation, it has some use.

Recently, AlphaFold2 has risen to the fore and has been lauded as one of the most accurate modelling software packages currently available for use (20, 21). It uses a DeepMind AI system to generate an ab-initio model of a protein, based on its amino-acid sequence. It is currently considered the gold-standard of protein modelling: they claim a global distance test score of 92.4, while their closest competitor (RosettaFold) scored 72.8 (22). This score is a measure of the similarity between two proteins with aligned primary structure but different tertiary structure, made using the α-Carbon, with the closer the value is to 100, the better the score (23). They also have a repository of nearly 200 million protein structures generated using this software (20). However, AlphaFold 2 doesn't predict post-translational modifications, antibody binding, it struggles with proteins that have multiple conformations and the effects of point mutations (22). As such, structural determination remains necessary.

<u>F protein glycobiology:</u>

Each monomer of the F protein has 4 asparagine (N) residues, which may, theoretically, be glycosylated (24). N-linked glycosylation sites, or 'sequons,' consist of an asparagine, followed by any residue (except proline), then either a serine or threonine are glycosylated in the endoplasmic reticulum (25, 26). A covalent bond is formed between the asparagine's amide and the carbohydrate units of the glycans. As the protein transits though the Golgi apparatus, the high-mannose glycans increasingly become less-high-mannose, having them cleaved and replaced with Glc-NAcs, Galactose and sialic acid (27).

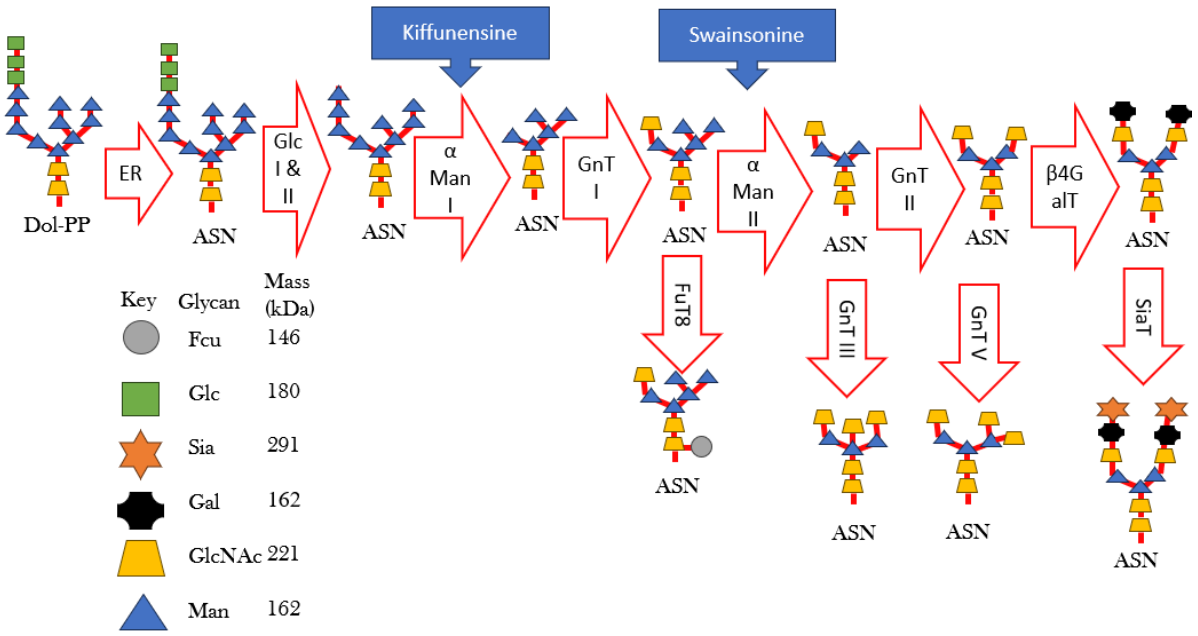


Figure 9: a summary of the pathway by which proteins are glycosylated, their approximate mass and the means by which they can be inhibited (27).

With 4 glycosylation sites, per monomer (see appendix B) glycans could well make up 20% of the protein's mass. *In-vivo*, more than half of Eukaryotic proteins are glycosylated to some degree, these serving to mediate protein trafficking, masking antigenic sites, preventing proteolysis and stabilising the protein (28). Increasingly, they are thought to play a role in mediating protein folding and functioning as a form of 'Molecular glue (29).' Pathogens, including viruses, will often have surface antigens with multiple residues that can be glycosylated to varying degrees. These perform multiple functions: they help direct transportation within the cell, prevent interference from intrinsic proteolytic factors and to evade the attentions of the immune system (30). Usually, the host's own glycosylation machinery will be diverted for the manufacture of viral glycoproteins. Inducing changes in the glycosylation sites is achievable for a virus with a change of a single nucleotide, and for a relatively small change in the viral genome, the entire antigenic profile of the virus can shift (31).

Glycosylation is one of the many factors that can impact the collection of useful crystallographic data: Indeed, the consensus is that, between the flexibility of the glycans and the subtle differences between them, glycans can negatively affect the formation of ordered crystals (32, 33). As-such the lab standard

protocol for the expression of extracellular proteins using mammalian culture includes adding kifunensine to the culture (see chapter II). This is a mannosidase inhibitor, which prevents the cellular machinery from producing proteins with complex glycans (34). HEK293-S cells are N-acetylglucosaminyltransferase I deficient, preventing formation of complex N-kinked Glycans and allowing for cleavage by endoglycosidases (35).

SAXs:

Small Angle X-ray Scattering (SAXs) is a technique that involves passing X-rays through a chamber holding protein in aqueous solution and measuring the scatter pattern (36). As the sample passing through the chamber must be mono-disperse, this is often performed coupled with size-exclusion chromatography (SEC) (37). Analysis of the scatter pattern can provide insight into the properties of the sample, namely approximate molecular weight (MW), relative order, and the conformation-type of the protein. For example, it is able to distinguish between linear and globular proteins (amongst others) which is important for the biophysical characterisation of the F protein prior to structural characterisation (38). In some cases, it can even propose low-resolution models of the structure. Within the context of this project, it will be useful for validating the construct and any models generated, to ensure it is sensible and as an element of quality control.

Results:

A note on the chronology of this work: If written chronologically, this aspect of the project would appear to have been conducted backwards. Following small scale expression tests, a trial-run of the large-scale expression was carried out, using the La Sota construct. The product of this was then purified in the manner described below and in chapter II. This initial batch of F protein was then used to set up three crystal trays. Five days later, the tray with the Morpheus screen contained two wells that had grown crystals. These crystals were harvested and imaged at Diamond Light Source (39), but diffracted poorly and made processing difficult. It was decided to try and optimise the conditions in which the crystal trays were set up, accounting for the myriad factors detailed below. However they failed to produce more crystals. This process of optimisation was included within the segment of protein purification and

production of the crystal trays to keep similar experiments together and make for more coherent reading. Similarly, the crystal trays set up using protein produced in insect cells (see chapter IV), were also included in table 2 for completeness.

Protein Expression:

It was necessary to confirm expression of the LaD and GTVII F protein in small scale cell-culture. Media from the small scale expression was used to perform a Western Blot using anti-6-His primary antibodies to bind to the 6-His tag on the F protein See figure 10. This was then scaled up to large-scale cultures of HEK-293T and later the N-acetyl-glucosaminyl transferase I deficient HEK293-S cells.

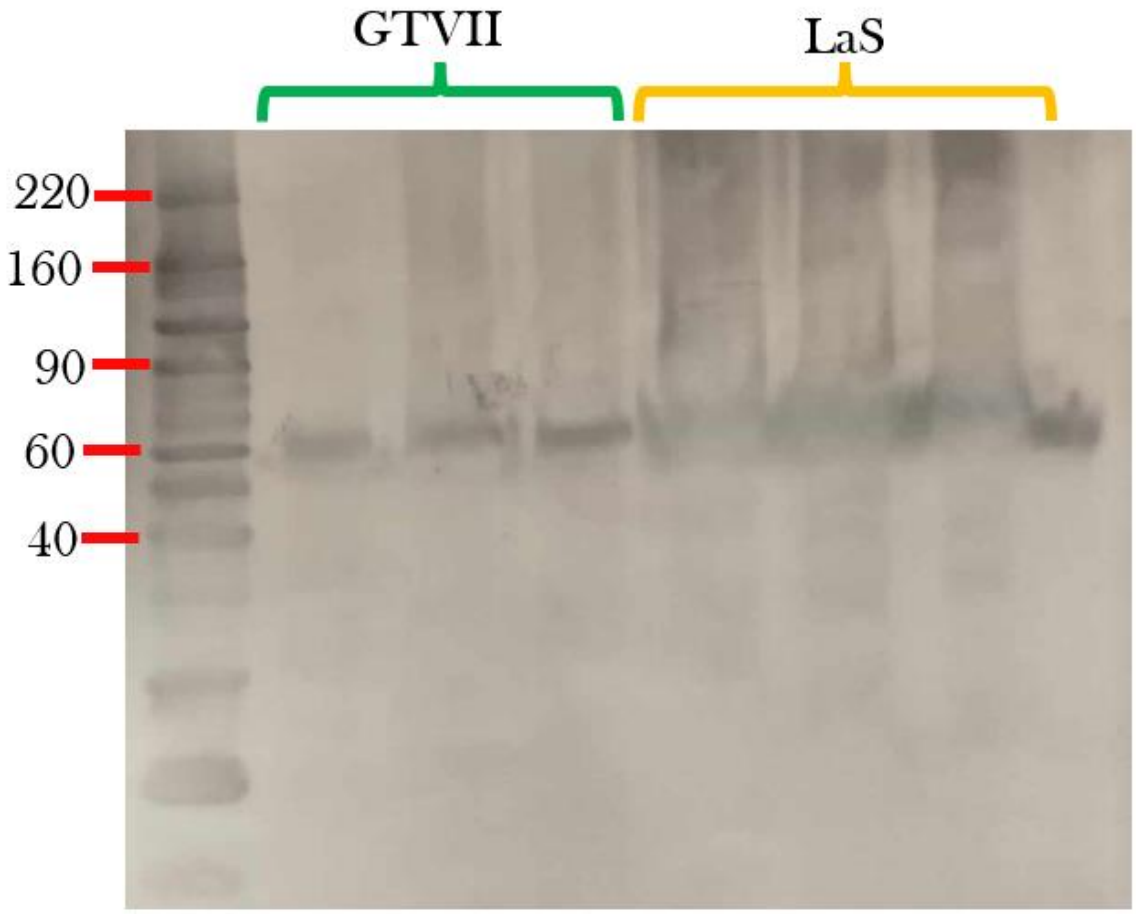


Figure 10: A western blot showing the presence of both His-tagged F protein constructs in the media of a transiently transfected HEK293T culture.

F Protein modelling:

Protein modelling was carried out to generate an image of what the NDV F protein might look like in the pre-Fusion conformation. Modelled structures could also be used to assist with the structural determination through homology reconstruction or to map stabilising mutations from known pre-fusion structures to NDV. The models generated by Swiss-model, based on the structure of PIV5 (4WSG) looks as one might expect (see figures 11 and 12). It has the globular form with a stalk usually associated with Class I fusion proteins. From a top-down view it also shares the approximately triangular shape of its counterparts. In the model, the red zones represent the areas of which the program is less confident. The seemingly disordered loops at the top of the structure where the helices of DIII fold back on themselves are among the areas of lower confidence, as are the β-sheets of DI, the loop of DIII around the cleavage site and the most N-terminal region of the helix that forms the stalk domain.

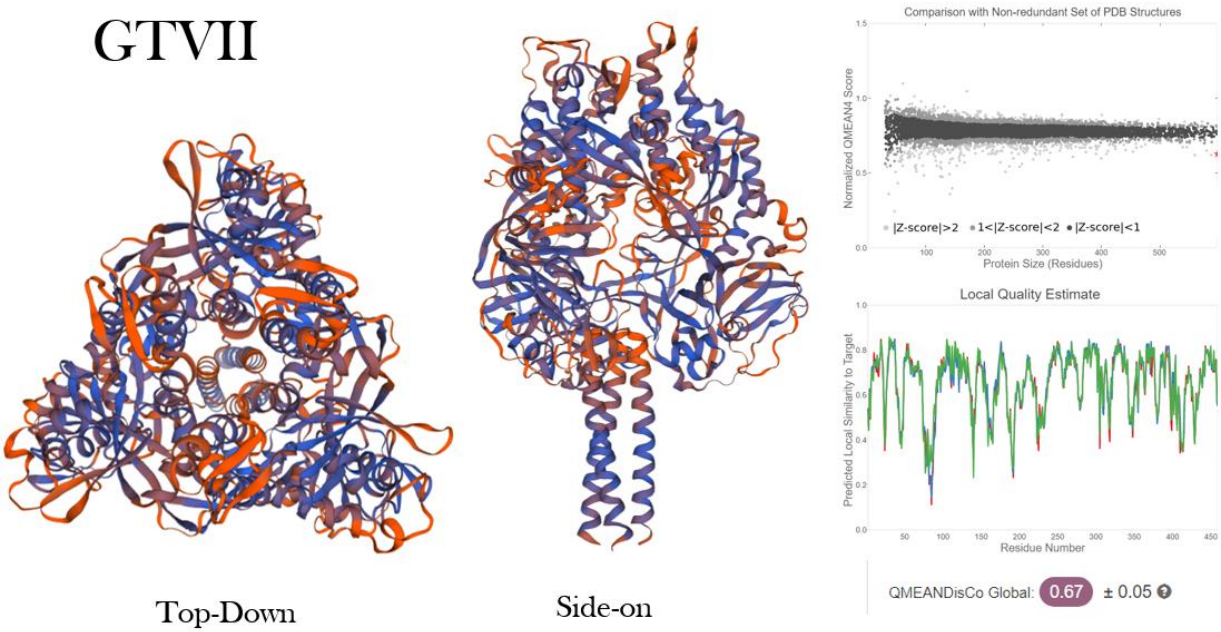


Figure 11: The output from SwissModel showing a potential trimeric structure for GTVII NDV. Areas of higher confidence are marked in blue, areas of lower confidence in red. On the lower right is a quality estimate, which shows the confidence that the software has for each residue it placed. Above this is a comparison estimate, which gives a judgement on the quality of the model, through comparison with high-resolution XRC data from the PDB.

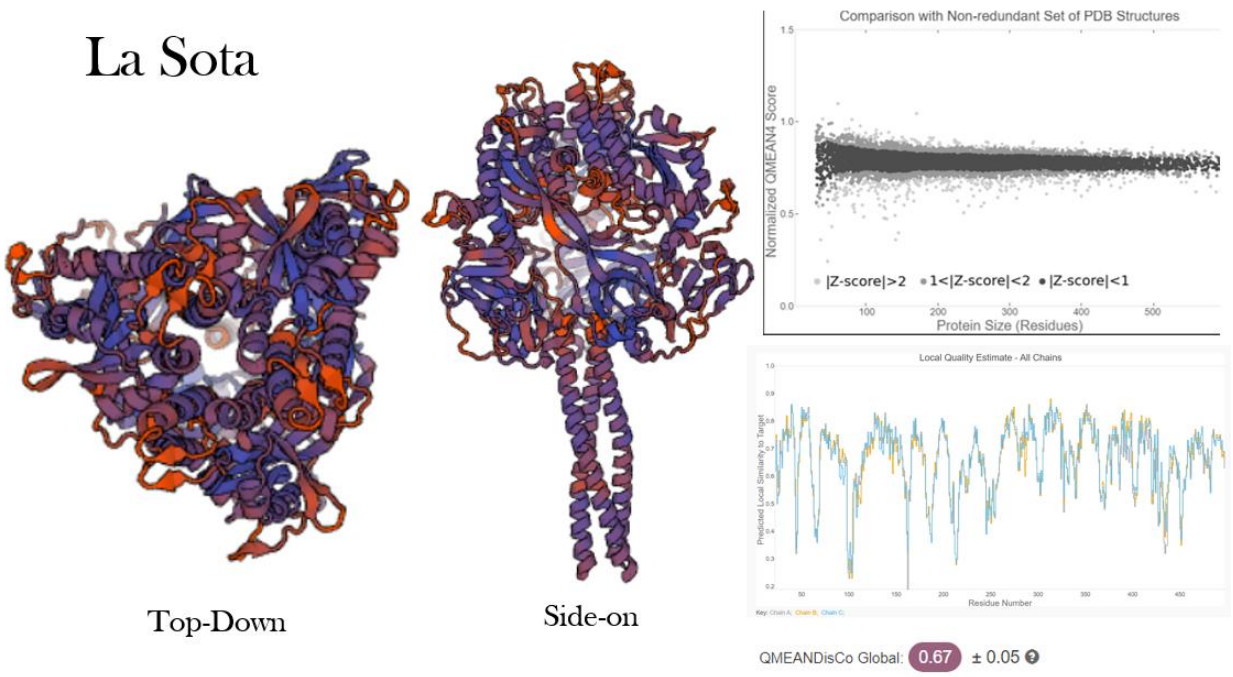


Figure 12: The output from SwissModel showing a potential trimeric structure for La Sota NDV. Areas of higher confidence are marked in blue, areas of lower confidence in red. On the lower right is a quality estimate, which shows the confidence that the software has for each residue it placed. Above this is a comparison estimate, which gives a judgement on the quality of the model, through comparison with high-resolution XRC data from the PDB.

The La Sota F protein model is, unsurprisingly, not that different to the model of GTVII. The areas of lower confidence in the model are, by and large, shared.

Alpha-Fold

Turning to Alpha-Fold, it was possible to generate a model de-novo, where-as the models generated by Swiss-Model are generated using known structures. However, Alpha-Fold struggled to process the full trimeric structure of the F protein, so a single monomer was submitted to the server for both genotypes of the virus. The program would crash while running the trimeric sequence, due to computing limitations. For ease of reference, Figure 8b, showing the domains of the F protein is included below. Each model the program produces also contains a measure of confidence for each residue in the form

of predicted local distance difference test (plDDT), with 0 being the lowest confidence and 100 being the highest confidence.

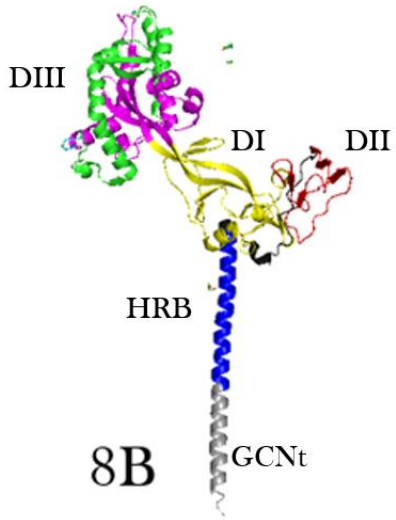


Figure 8: A reminder of the structure of the PIV5 F protein monomer, in the pre-fusion conformation colour coded to show DI, DII and DIII (13).

Looking at the model of La Sota produced (See figure 14), much of the parts which constitute the head domain of the F protein have scored a Very High plDDT (coloured blue), with only a few areas of Very Low or Low confidence in the plot. Among this latter group are the loop that extends in the middle of DIII, the α-helix extending from DII and the HRB-linker-domain, which links the HRB domain to DI. The α-helix extending from DII may be discounted as this is a signal peptide cleaved during export from the cell.

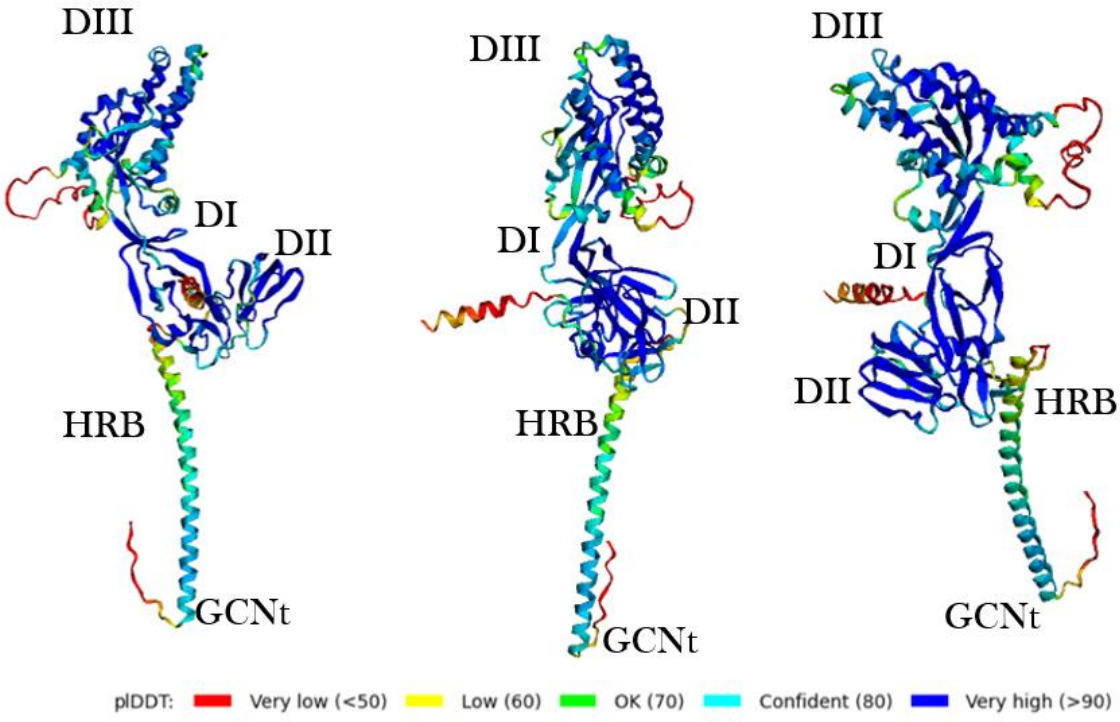


Figure 14: Multiple orientations of the Alpha-Fold generated model of La Sota, colour-mapped to show areas of lower and higher confidence.

Similarly Alpha-Fold reports high confidence in DI, DII and DIII, with lower confidence in the loop in DIII, and also the HRB-linker domain. However, it reports less confidence in the middle of the HRB domain. In generating the model, the signal peptide attached to the N-terminus of the protein was removed from the amino-acid sequence input into the program, hence its absence in the model.

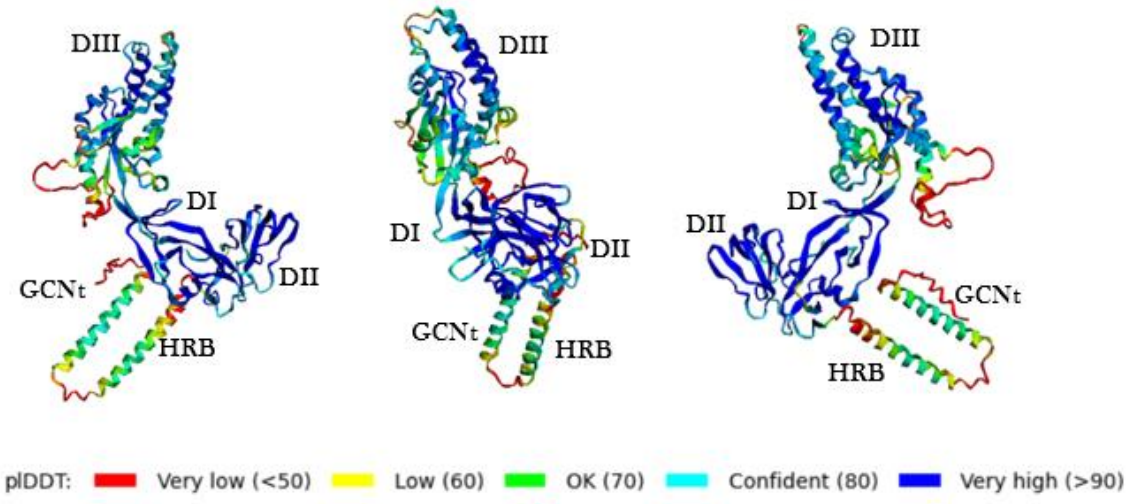


Figure 15: Multiple orientations of the Alpha-Fold generated model of GTVII, colour-mapped to show areas of lower and higher confidence.

Comparing the models to monomers of PIV5, La Sota and PIV5 monomer appear distinctly alike (see figures 16 and 17). The helices that make up the DIII region of La Sota are not exactly aligned, but are close to those equivalent helices, of PIV5. The same can be said for the β-Sheets that make up much of DI, though there is a little more divergence between the two viruses in the β-Sheets of DII. However, both proteins have very similar secondary structures in the same sort of area, reflecting their functional homology. The other major difference is in the position of the HRB and stalk domain. This, however, could well be explained by the low confidence in the relatively disordered region between the HRB domain and DI domain of the model of the La Sota monomer. The latter shares the sweeping curve of the PIV5 monomer, suggesting it will fit into the trimeric structure.

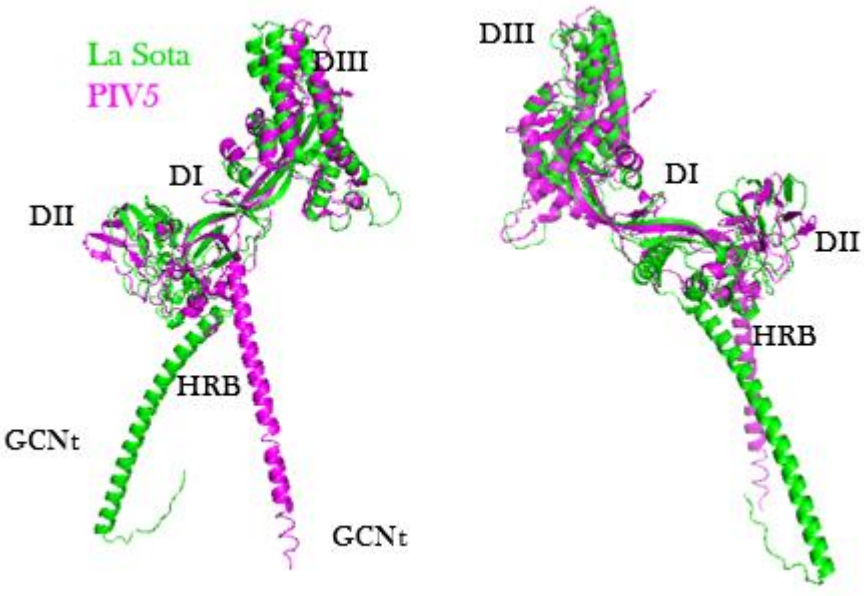


Figure 16: an overlay of the La Sota AlphaFold2 monomer on the PIV5 monomer. La Sota is in green, PIV5 in purple. Performed using the align function of Pymol.

The GTVII and PIV5 monomers, when fitted together share much of the similarities outlined above, perhaps with a better fit in the D1 β-Sheets of the two viral glycoproteins. As with the model of La Sota, the model of GTVII has a hinge region to the HRB domain which AlphaFold2 has very little confidence in, causing the stalk to protrude at an angle inconsistent with the monomer of PIV 5 (see figure 17). Furthermore the stalk domain has folded back in on itself. This is a region that the AlphaFold2 AI had lower confidence in.

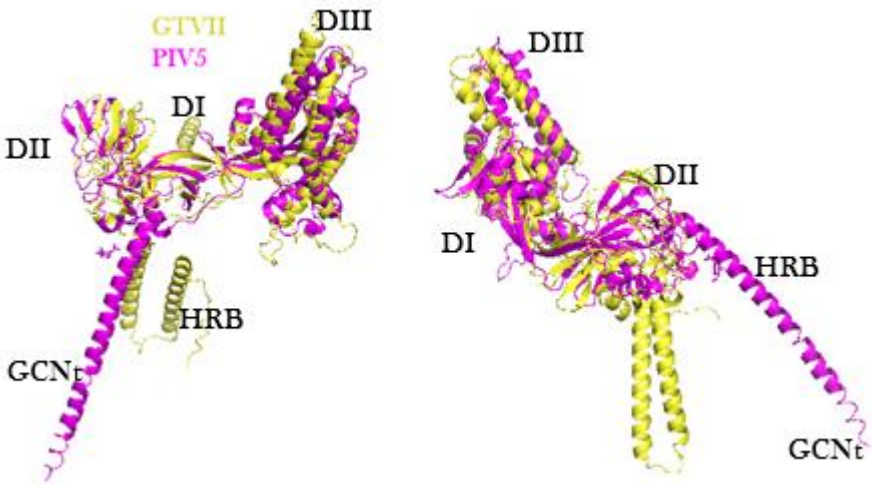


Figure 17: an overlay of the GTVII AlphaFold monomer on the PIV5 monomer. GTVII is in yellow, PIV5 in purple. the Alignment was carried out using Pymol's Align function.

Using the PIV5 trimer to assemble a NDV trimer: The La Sota monomers assemble into a trimer that does not look vastly different from the PIV5 trimer, less the splayed legs of the monomeric model. The β-Sheets that make up DII protrude more from the side of the model than PIV5 and the α-helical bundles that make up DI do not exactly match up, but neither are they especially divergent.

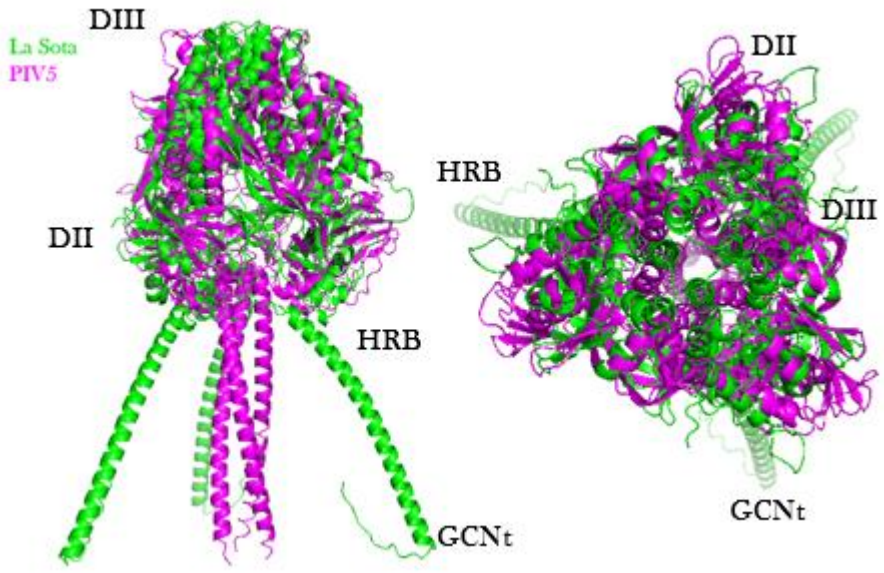


Figure 18: an overlay of the La Sota AlphaFold trimer on the PIV5 trimer. La Sota is in green, PIV5 in purple.

GTVII, in the trimeric form exhibits many of the same similarities and differences to the La Sota F trimer, when compared with the PIV5 model. The splayed stalk domains appear to be the most egregious disparity.

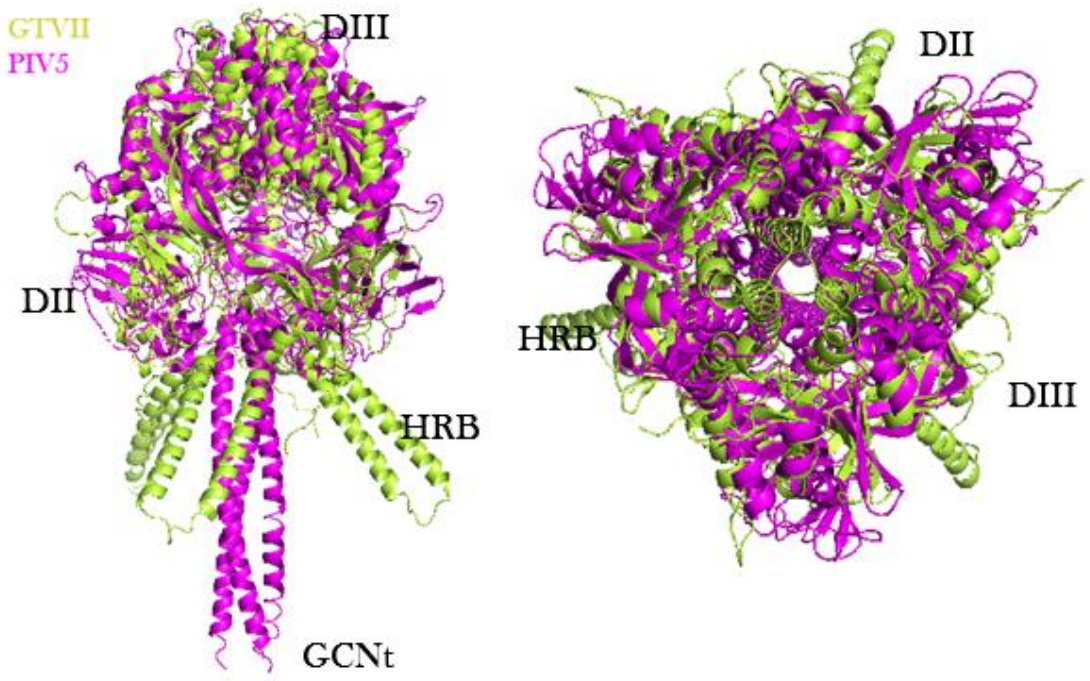


Figure 19: an overlay of the GTVII AlphaFold trimer on the PIV5 trimer. GTVII is in green, PIV5 in purple.

Subtracting the PIV5 structure to look at La Sota on its own (figure 20), it appears that the three monomers fit together without the α-helices of DIII from one monomer overlapping with those of its neighbour. The trimer of crescent shaped monomers also possesses the central cavity seen in PIV5 and other class I fusion proteins.

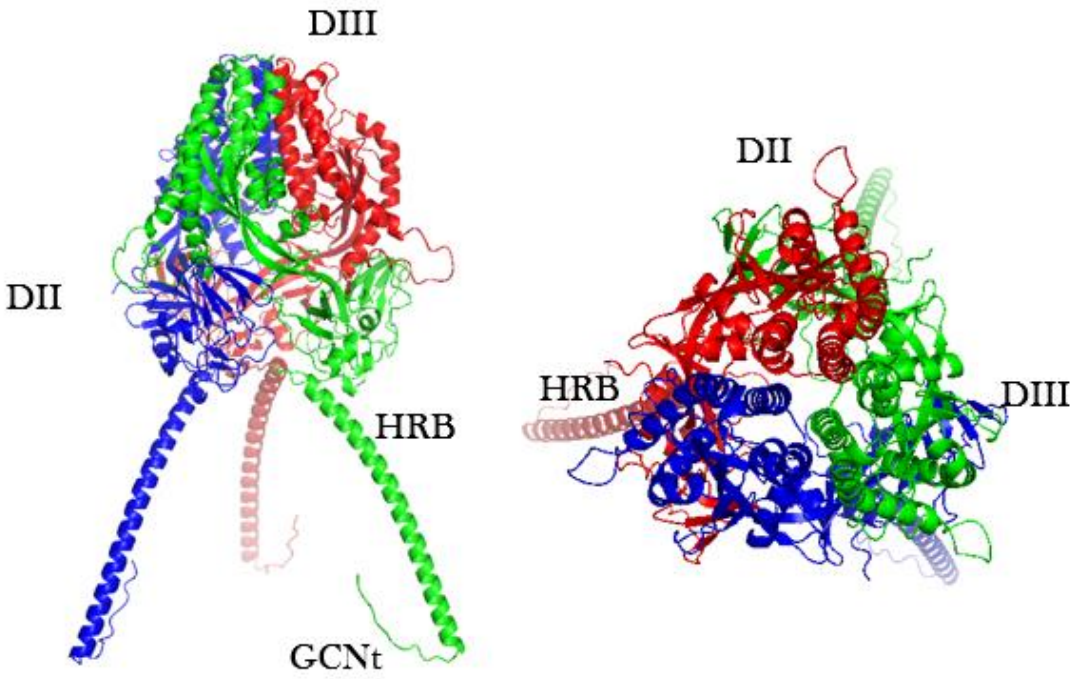


Figure 20: The trimer assembled from the La Sota AlphaFold monomers, Eash individual monomer coloured differently to show their overlap.

The trimer formed from the monomers of GTVII F protein (figure 21), without the PIV5 overlay, resembles that of La Sota, except for the doubled-over HRB domain. In both models it is clear how the sweep of the underside of DIII fits into the alcove formed by DII and DIII around DI.

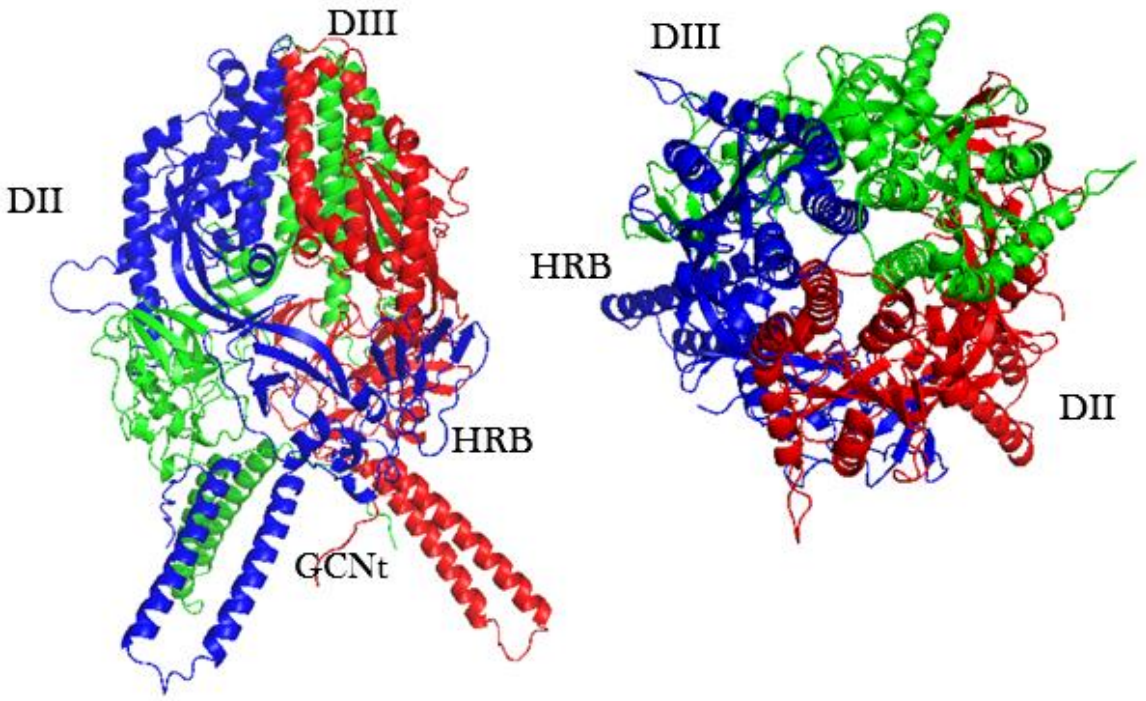


Figure 21: The trimer assembled from the GTVII AlphaFold monomers, Eash individual monomer coloured differently to show their overlap.

Purification:

Transient expression of F protein (both genotype VII and La Sota), was carried out by transiently transfecting plasmids containing the gene construct into HEK293T and 293S cells (see chapter II). The plasmids contained a signal peptide at the N terminus of the F protein, that would see the protein exported out of the cell. At the C-terminus of the F protein was a GCN(t) domain, followed by a 6-His tag.

Following buffer exchange, a two-step purification was performed to achieve a homogenous product. A His-trap was used to perform affinity chromatography. The elute of this was then subject to size-exclusion chromatography, initially using a superose-6 column. See figure 22 for the AKTA traces for the elution from the His-trap and figure 22C for the image from the SDS-PAGE gels of the product of the purification.

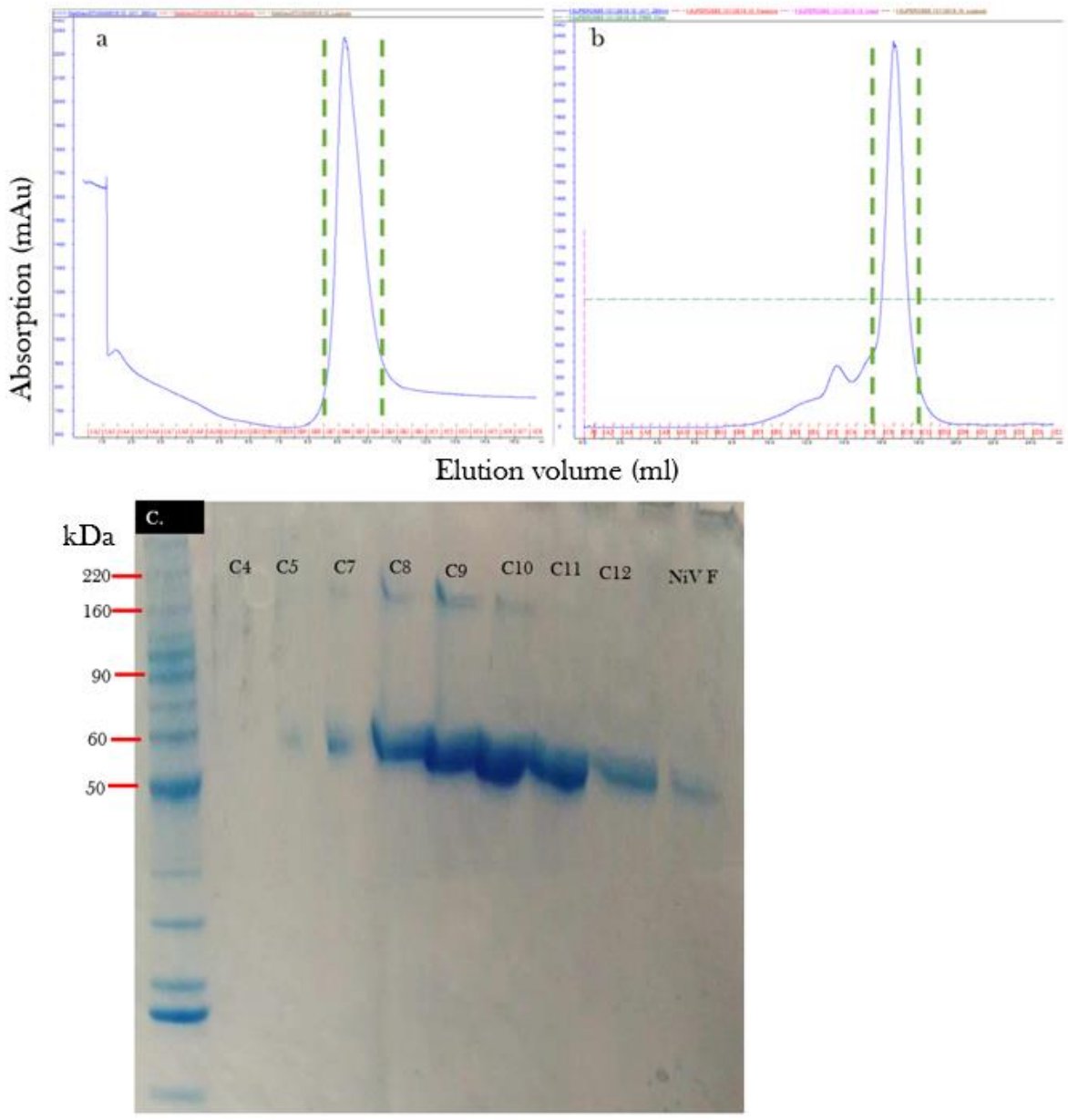


Figure 22: Typical traces generated on AKTAs from A: Elution from the His-Trap. B: Size exclusion using a Superose 6 10/300. C: The SDS-PAGE gel run using the fractions C4 to C12 from the size exclusion, using Nipah (NiV) F protein supplied by Dr Vicky Avanzato as a positive control.

The single peak during SEC and the lone band during SDS-PAGE suggested that the protein had been purified to a degree where crystal-trays could be set up. It was later realised that a Superose 6 could not distinguish between protein of approximately 55kDa and protein of 150kDa, i.e a monomer and a trimer, based on the specification of the column and the expertise of members of other postdoctoral researchers in the group. Running a series of molecular weight markers would have established this at

an earlier date and given an indication of what volume the F protein would elute at. Additionally, a wash step with imidazole should have been undertaken, but was not due to user error.

It was decided to use an S200 for further SEC. It was discovered that what appeared to be a single, relatively clean, peak on the Superose 6 was in-fact multiple peaks, each representing a protein of a different size. While this discovery was made during the work being undertaken in chapter VI, it is included here for completeness of describing purification. Additionally, some of the product of these purifications was put into crystal trays, in the hope that it would crystallise.

Thermofluor experiments: pH Screen

Once the protein had been purified experimentation was needed to determine at which pH it was most stable, and assist in finding the optimum pH for forming crystals. This was determined by performing a thermofluor (40) experiment. Concentrated La Sota F protein was diluted into Tris/NaCl solutions of different pHs (see chapter II) and mixed with Sypro Orange dye to a final concentration of 0.5mg/ml. This was then subjected to a temperature gradient inside a thermocycler, with the pH at which the protein is more stable, being more resistant to the effects of temperature. See figure 23.

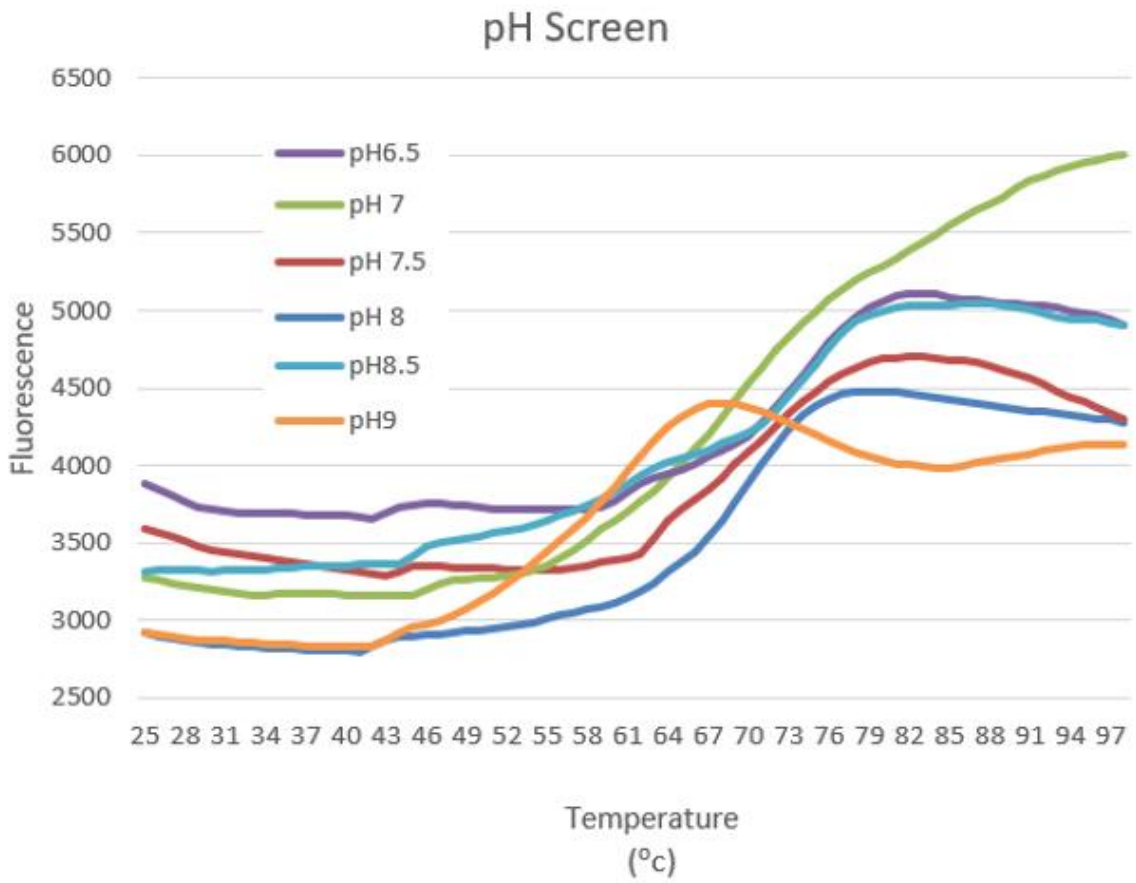


Figure 23: The effects of pH on the stability of La Sota F protein over a change in temperature. Each reaction was carried out in triplicate, and the curves shown represent the mean values. The T(m) was calculated by looking at the inflection point of the first exponential phase and reading off the temperature.

Looking at the results of the pH screen, pH8 seems to have the highest T(m), being as much as 10°c difference caused by going from pH7 to pH8. While the method of calculating the T(m) gave an indication, it is not reliable. In hindsight, plotting the inverse derivative of the graphs would give a clearer, less ambiguous value for the T(M).

Hampton Additive Screen:

A screen was prepared, with La Sota F protein, to explore the use of additives known to increase or decrease the stability of proteins in solution, as described above (40). Any additives that increased the

stability of the protein might help induce the formation of crystals. The additives were at 0.2um in solution. For brevity only the additives that modified the T(M) of the protein are shown. See figure 24.

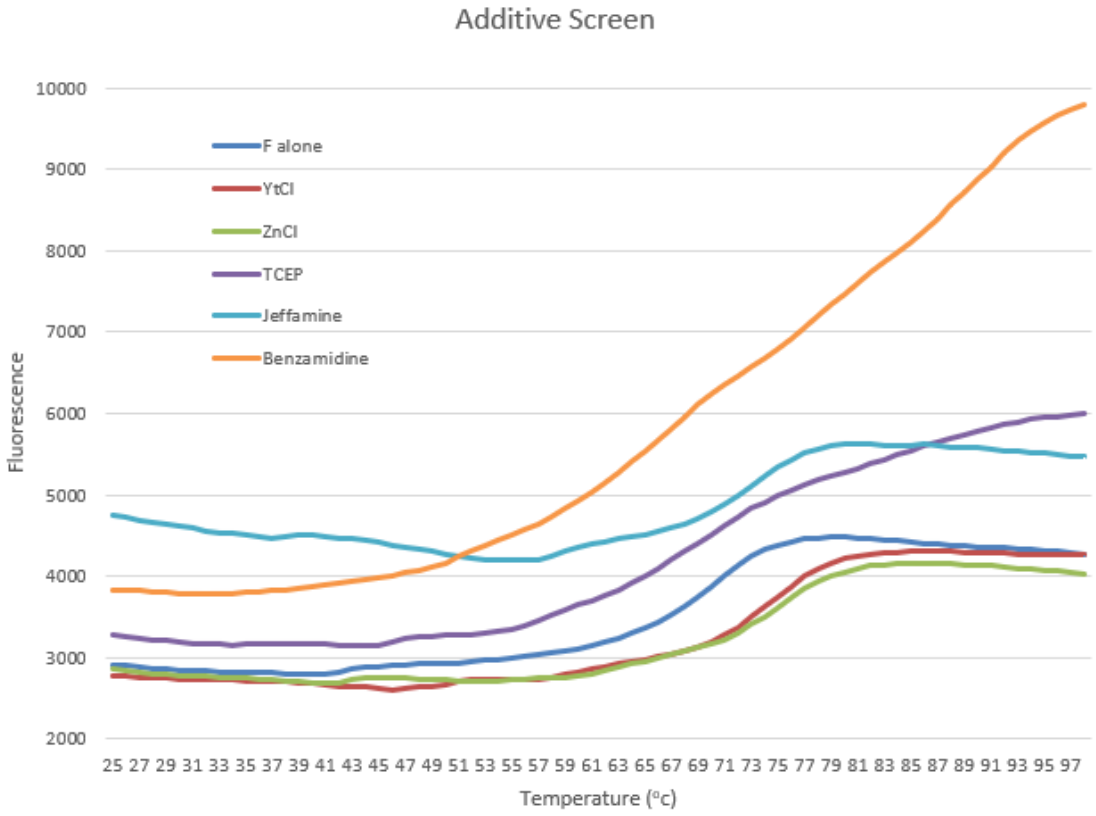


Figure 24: The effects of various additives on the stability of La Sota F protein over a change in temperature.

Yttrium chloride and zinc chloride were shown to increase the T(m) of the protein from 67°c to 75°c, suggesting that they would make the protein more stable in solution. S such crystal trays were set up with these additives. Jeffamine, benzamidine and tris(2-carboxyethyl)phosphine hydrochloride (TCEP), however, reduced the T(m) by 4 to 10°c.

Crystal Trays:

To grow the crystals, putatively purified F protein (based on the SEC trace and SDS-PAGE in figure 22) was plated into 96-well crystal trays with crystallisation solutions (Supplied by Molecular Dimensions) and placed in 'hotels' where they were regularly screened for the growth of crystals. Nearly 200 crystal trays have been set up, see table 2 for a breakdown of construct and conditions put into trays and Chapter II for the methodology. Crystals of the LaS construct grew in only 1 of those: screen

'Morpheus' (Product number MD1-76) wells A12 and C10 with divalent ligands at 0.06M tris (base); bicine buffer at 0.1M pH8.5 MPD_P1K_P3350 37.5%, and C10 1-34 nitrate phosphate sulphate ligands at 0.09M, tris (base); bicine buffer at 0.1M pH8.5 by day 5. (See figure 23).

Construct	Conc. of protein (mg/ml)	Screens	Buffer conditions	Temperature	Crystals?	Remarks
LaS	3.8	Morpheus, Index Screen 1	0.1M Tris 0.15M NaCl	RT	Yes	Crystals in Morpheus tray
LaS	5	All	Ditto	RT	No	Endo F Deglycosylated
GTVII	6	All	Ditto	RT	No	Native GTVII
LaS	10	All	Ditto	RT	No	Without Kifunensine
GTVII	5	All	Ditto	RT	No	Endo F Deglycosylated
GTVII, LaS	9.5	All	Ditto	4 °C	No	Native
GTVII, LaS	20	Morpheus, Index, Screen 1	Ditto	RT	No	PNGase digested
GTVII, LaS	12.5	All	Ditto	RT	No	Added 0.01 % trypsin
GTVII, LaS	10	All	Ditto	RT	No	Produced in 293S
GTVII, LaS	14	2 Morpheus, 2 Index	0.1M Tris 0.5M NaCl	2 at 4 °C, 2 at RT	No	Increased salt concentration

GTVII, LaS	20	1, 3, Morpheus, Index	0.1M Tris 0.5M NaCl	RT	No	0.01M ZnCl
GTVII, LaS	20	1, 3, Morpheus, Index	0.1M Tris 0.5M NaCl	4 °C	No	0.01M ZnCl
GTVII, LaS	44	Morpheus, Index	0.1M Tris 0.5M NaCl	RT	No	Increased protein concentration
GTVII' LaS	14	Morpheus, Index	0.1M Tris 0.5M NaCl	2 at 4 °C, 2 at RT	No	PNGase digested then gel filtrated.
GTVII, LaS	20	Morpheus, Index, 1, 3, SRX	0.1M Tris 0.5M NaCl	4 °C	No	0.01M $NiSO_4$ additive
LaS	20	All	0.1M Tris 0.5M NaCl	4 °C	No	Expressed using SF9 cells.
LaS	10	All	0.1M Tris 0.5M NaCl	4 °C	No	3-fold molar excess of VHH 2 or VHH 4. (See chapter V)

Table 2: A breakdown of the conditions in which crystal trays were set up, and the result. Where 'All' is reported, a crystal tray with each of the following Molecular Dimensions screens was set up: 1, 2, 3, 4N, 5, Morpheus, Morpheus II, Pro-Plex, SRX, Nuc-Pro and PRX.

Crystals from the Morpheus screen were directly flash frozen (without cryoprotectant) from the drop and stored in pucks in preparation for screening at Diamond Light Source.

The remaining trays, set up subsequently, have not grown any crystals. The conditions that led to crystals initially were replicated as far as was practical. Attempts to find other conditions that induce crystal growth have not been successful. The construct appeared to be very soluble and can tolerate concentration to 44 mg/ml. Once concentrated above 4 mg/ml, the protein in solution takes on a

distinct red hue. This could have been due to bovine serum albumin in the media of the HEK293T cells binding to the His-trap and not being removed by a wash step. A Western blot should have been undertaken on the protein fractions from SEC to ensure that this was not the case.

The addition of divalent cations, changes in salt content of the buffer, enzymatic digestion and temperature changes have not met with success.

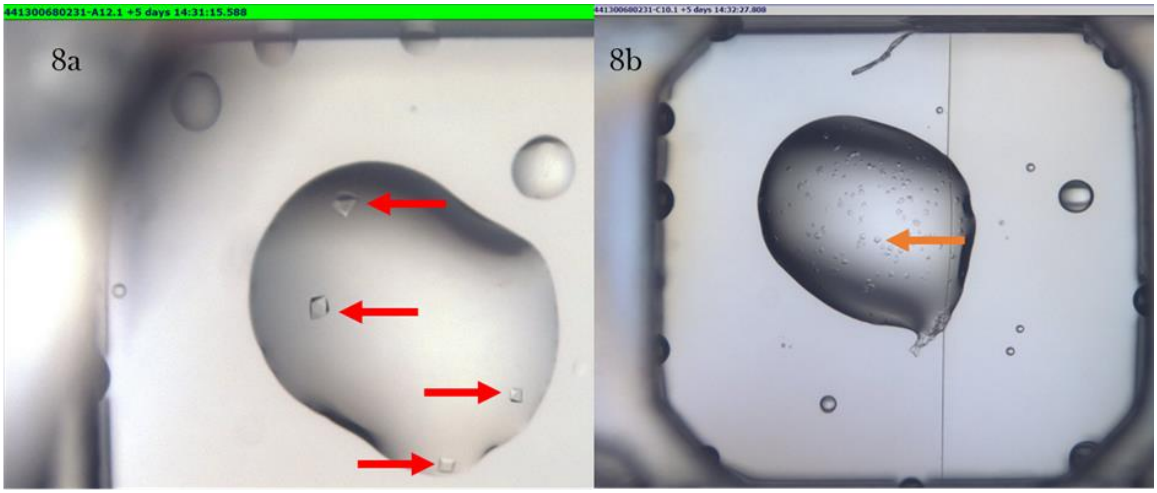


Figure 23a: The crystals formed in well A12 of the Morpheus Screen during the initial crystallisation effort.

Figure 23b: The crystals formed in Well C10 of the same screen during the initial crystallisation effort.

Enzymatic digestion:

Heterogenous glycosylation of protein can hamper the formation of protein crystals, with each monomer of NDV F having 4 potential N-linked glycosylation loci (33). Enzymatic deglycosylation was attempted to improve homogeneity and improve the chances of crystals forming. An initial digestion was carried out using 10ug of the F protein for 12 hours at 4°C. with ENDO-F or PNGase, not under denaturing conditions, then run on a gel (see figure 24). It became apparent that ENDO-F was effective at removing some glycans, but PNGase was much more radical. It also appeared that not all the protein was deglycosylated during this process.

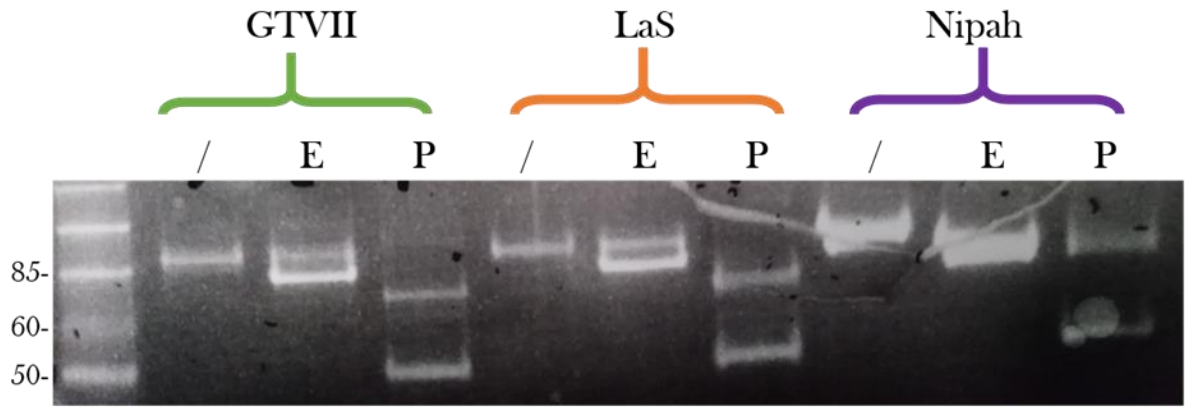


Figure 24: A SDS-PAGE gel showing the effects of enzymatic deglycosylation on both constructs of NDV F protein and NiV F. E = Endo-F, P = PNGase, / = undigested F protein. NiV F (supplied by Victoria Avanzato) was used as a positive control due to its availability in the department and similarity to NDV F (42).

This digest was then scaled up, doubling the ratio of enzyme to protein (Both had been as-per manufacturer's guidelines) and digested for 24 hours at 4°C. This temperature was chosen to minimise degradation of the protein due to temperature. After enzymatic digestion, the protein was run through another round of SEC to remove the enzyme before being used to set up crystal trays. While the digests were effective, the attempts to grow crystals with enzymatically deglycosylated protein were not.

X-Ray Crystallography:

Table 3: A summary of the data collection during X-ray crystallography using crystals of LaS F protein at Diamond Light Source

Beamline	I03
Beam_size (µ)	20x20
Space Group	P 43 3 2
Unit Cell (Å/ °)	397.107, 397.107, 397.107, 90, 90, 90,
Wavelength (Å)	0.97624
Resolution Range (Å)	110.14 - 4.5 (29.7 – 7.06)

Total Reflections	1309191 (86276)
Unique Reflections	16819 (1196)
Completeness (%)	100.00 (97.7)
Multiplicity	77.8 (72.1)
<I>/<σ(I)>	29.1 (7.3)
Rmerge (%)	00.213 (0.913)
CC half	1.0 (1.0)

Crystals were exposed to X-rays at Diamond Light Source (39), see Material and Methods and table 3 for the parameters and initial data. The crystals diffracted weakly to 4.5 Å with an estimated 4 trimers in the crystal with the Matthew's coefficient. Data was processed in IA2 on DIALS via the ISPYB interface (41, 42): X-triage did not report any issues. When molecular replacement was attempted in PHENIX (43), it was unable to find solutions for the number of trimers returning a negative log likelihood gain value. PIV5, Nipah and RSV were being used as potential models, while P43 3 2 and P4 3 2 space groups were tried unsuccessfully. The post-fusion NDV structure was also tried to ensure the analysis was failing due to the sample being in the post-fusion conformation.

The issue was attributed to the low resolution, so it was decided to try and grow more crystals, that would diffract to a higher resolution. Unfortunately after much effort and the testing of myriad conditions on both genotypes, the crystallisation of the protein was unrepeatable. Deglycosylation of the protein, using additives, altering the protein concentration, temperature, buffer conditions had no effect. Further analysis of the data, using MOLREP (44), to calculate the Self Rotation Function revealed a strong peak on the Chi=120° section, suggesting that the trimer of NDV F would lie along the crystallographic 3-fold axis (see figure 26). The Chi=120 plot shows four peaks to correspond with the number of oligomeric proteins asymmetric unit. This could suggest that the crystal is formed from contaminants.

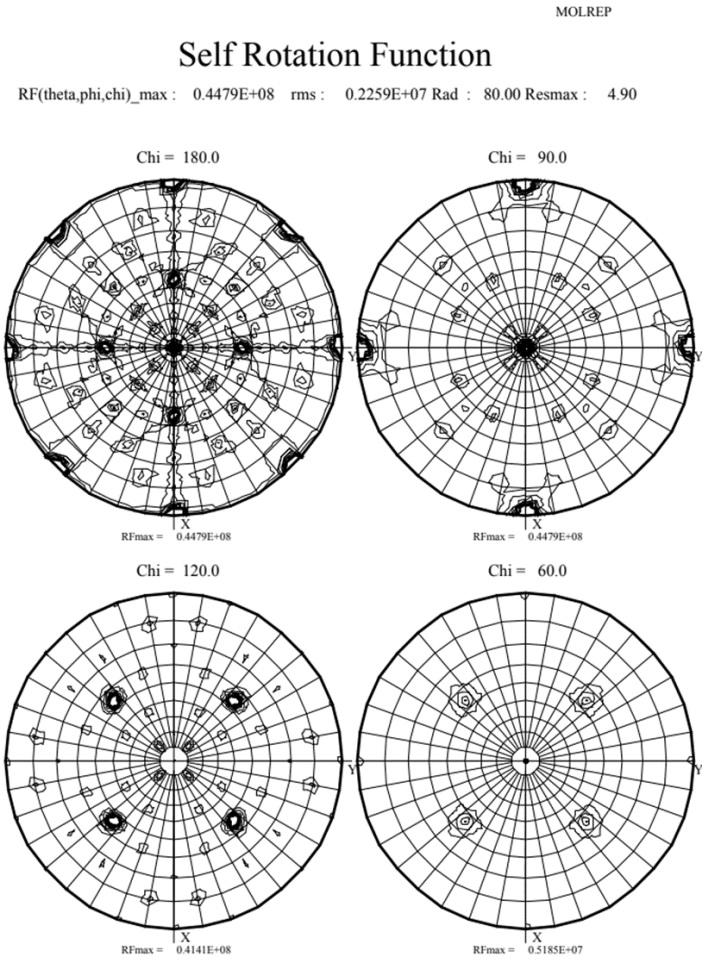


Figure 26: The self-rotation function charts produced by MOLREP during the processing of the crystallography data, Showing the symmetry axis, Chi = 180 showing 2-fold axis, Chi=90 showing a 4-fold axis.

SAXs:

SAXs offered an opportunity to learn more about the structure of the sample while it was still within an aqueous form. The key things to determine were whether the sample resembled other fusion proteins was in the pre-fusion or post fusion conformation. Additionally it was beneficial to see if pH had an impact on protein conformation, as the switch from prefusion to post-fusion F conformation was pH dependant. NDV F protein of both genotypes was expressed in HEK293T cells in the presence of kifunensine, with half of each being subjected to a second round of size exclusion to transfer them from a pH 8 buffer to a pH5 buffer. The samples were then kept on ice until the following day.

SAXs was able to provide some useful data when analysed using CHROMIXS (45): Firstly, the predicted molecular weight of the La Sota sample matched that predicted based on the protein sequence and shown by SDS-PAGE. (See figure 27).

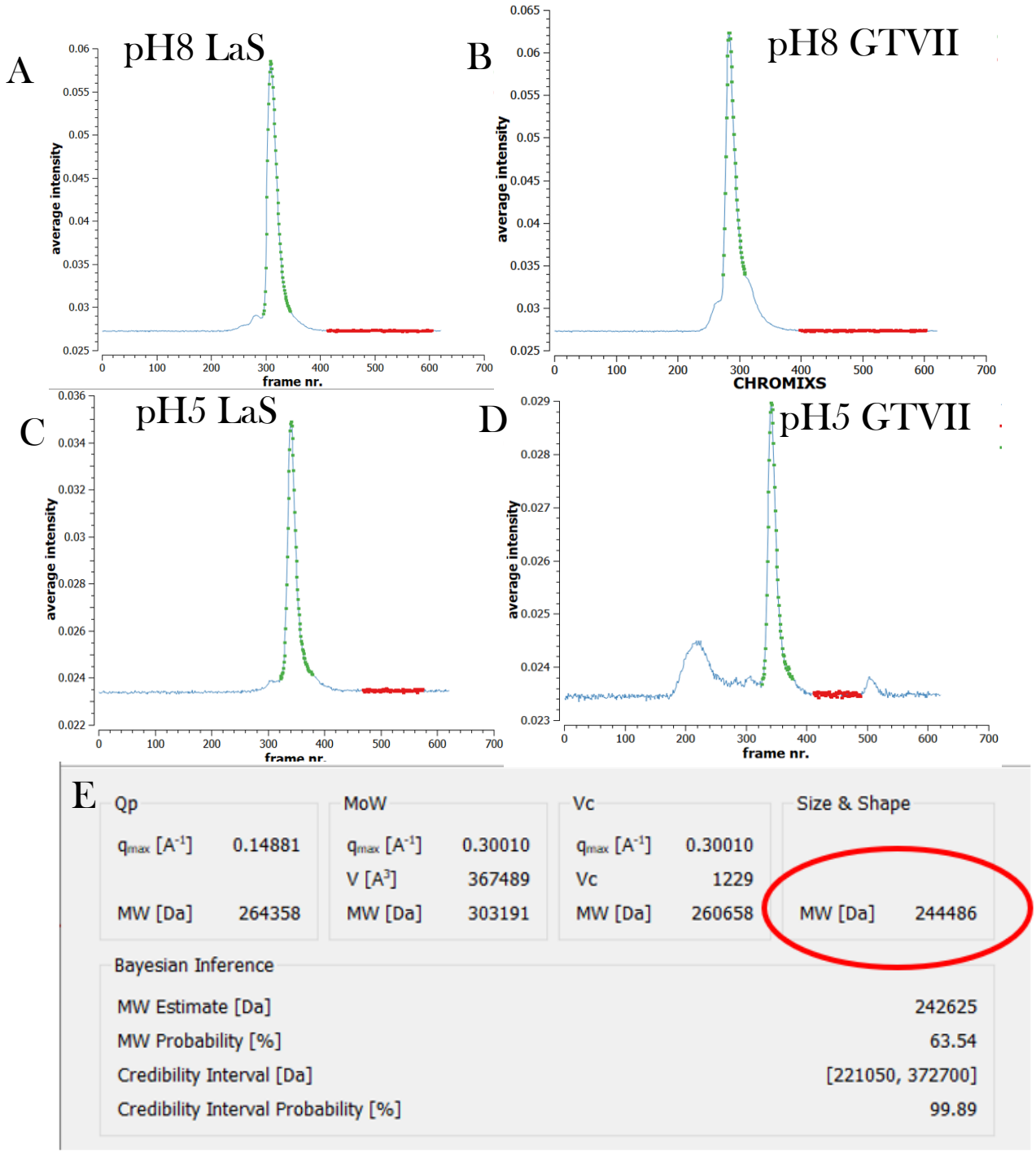


Figure 27A-D: The SEC traces for the LaS and GTVII constructs at pH8 and pH5, measuring the intensity of scattering induced by synchrotron radiation as the protein left a Superose 6 10/300 column. The green squares indicate the data-points used in future processing, while the red squares indicate the

data-points used to represent the buffer and used to exclude the background. Figure 27 E: The output of the CHROMIXS MW calculation function showing the predicted MW for LaS at pH8.

Kratky plots serve to show to what extent proteins are folded or unfolded (46). The form of the Kratky plots for the samples at pH8 resemble those of folded, globular proteins, suggesting that the F protein is in its globular pre-fusion conformation rather than the linear form. The plots for the protein at pH5, meanwhile shows partial unfolding that could be consistent with moving from a globular to a linear confirmation.

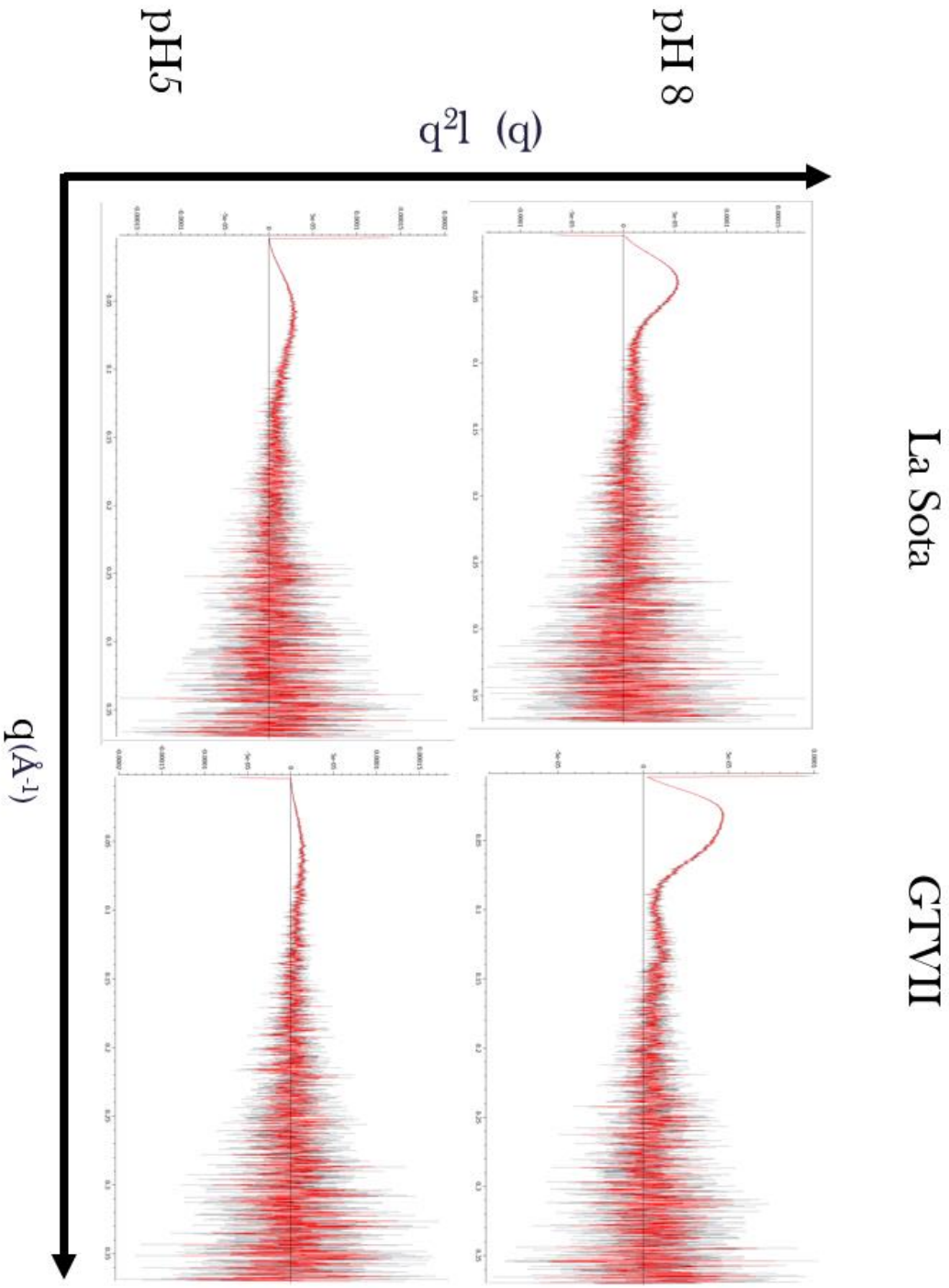
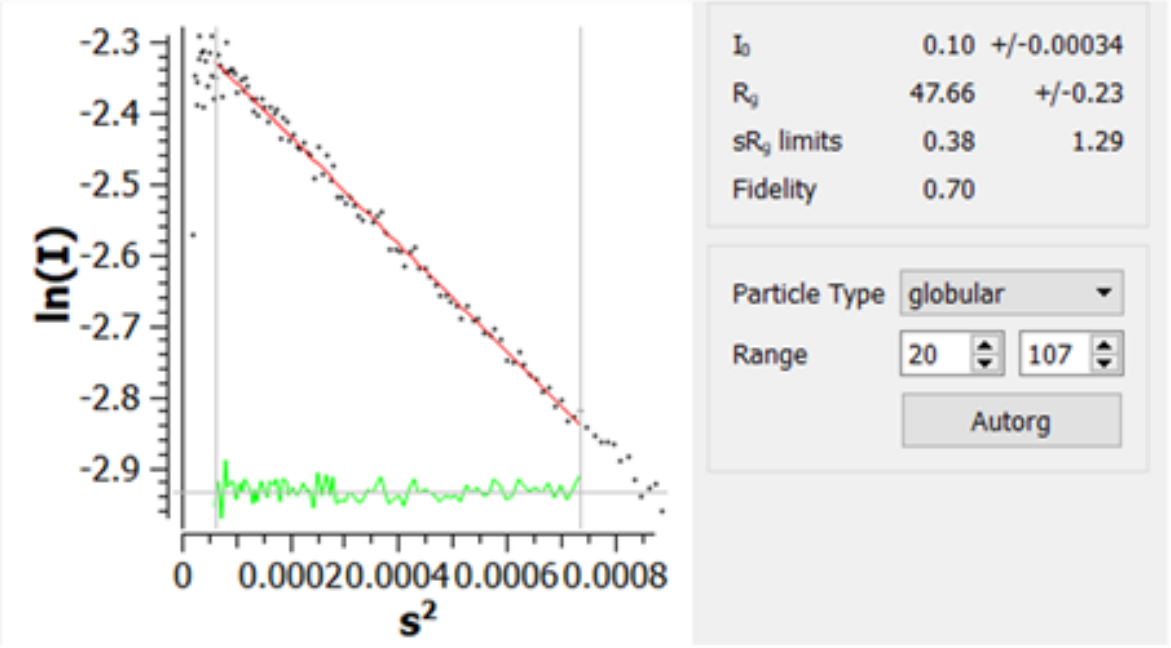


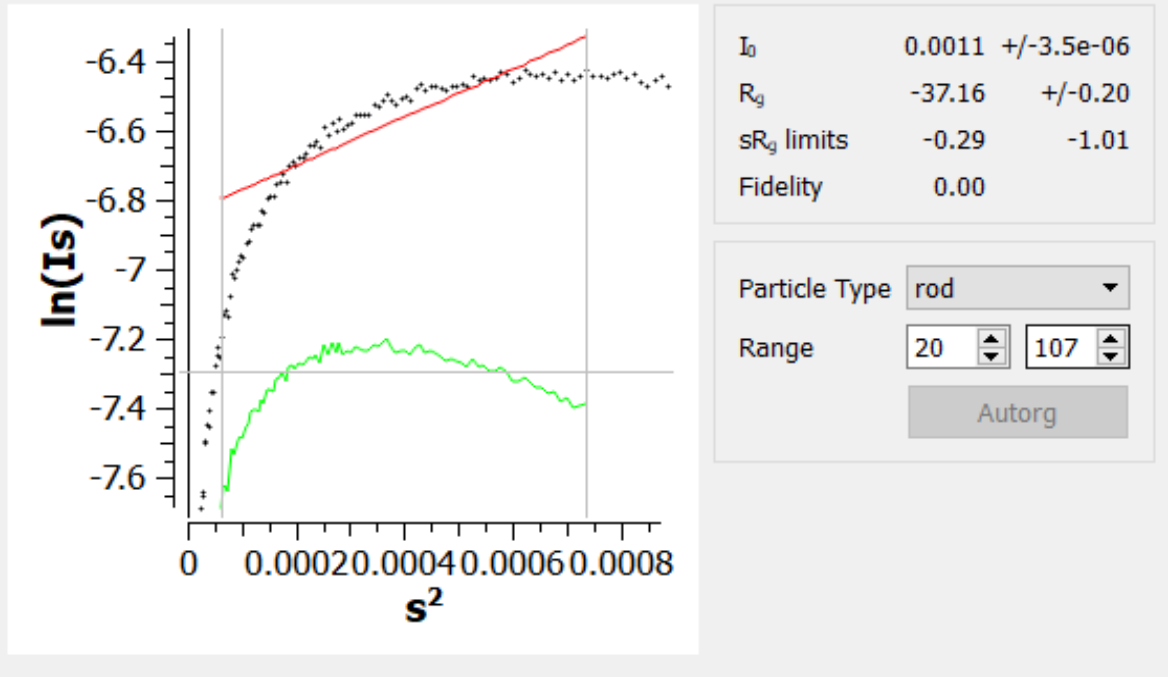
Figure 28A: (Previous page) The Kratky plots generated by CHROMIXS from the SAXs data seen in figure 27, comparing the effects of pH on LaS and GTVII. The Kratky plots suggest that at pH8, both constructs are folded, but at pH5, there is partial unfolding (43).

This is further complemented by the trend lines of the Guinier Plots of the two constructs at pH8 most matching the globular line of best fit (see figure 29). There is a small cluster of data on the top left of both globular graphs, suggesting a very limited amount of aggregation.

With the aggregation excluded, the q R_g max offers a quantification of globularity (reported in the box as sR_g. For both, GTVII and La Sota (in the globular conformation) the score is under 1.3, thus they appear credibly globular. However, the measure of fidelity for the GTVII is only 0.25 in the globular chart. For comparison, the La Sota has a Fidelity score of 0.7 in the globular chart, but 0 for the Rod conformation chart.

La Sota


-2.3
-2.4
-2.5
-2.6
-2.7
-2.8
-2.9
ln(I)
0
0.0002
0.0004
0.0006
0.0008
s²
I₀ 0.10 +/-0.00034
Rg 47.66 +/-0.23
sRg limits 0.38 1.29
Fidelity 0.70
Particle Type globular
Range 20 107
Autorg


-6.4
-6.6
-6.8
-7
-7.2
-7.4
-7.6
ln(Is)
0
0.0002
0.0004
0.0006
0.0008
s²
I₀ 0.0011 +/-3.5e-06
Rg -37.16 +/-0.20
sRg limits -0.29 -1.01
Fidelity 0.00
Particle Type rod
Range 20 107
Autorg

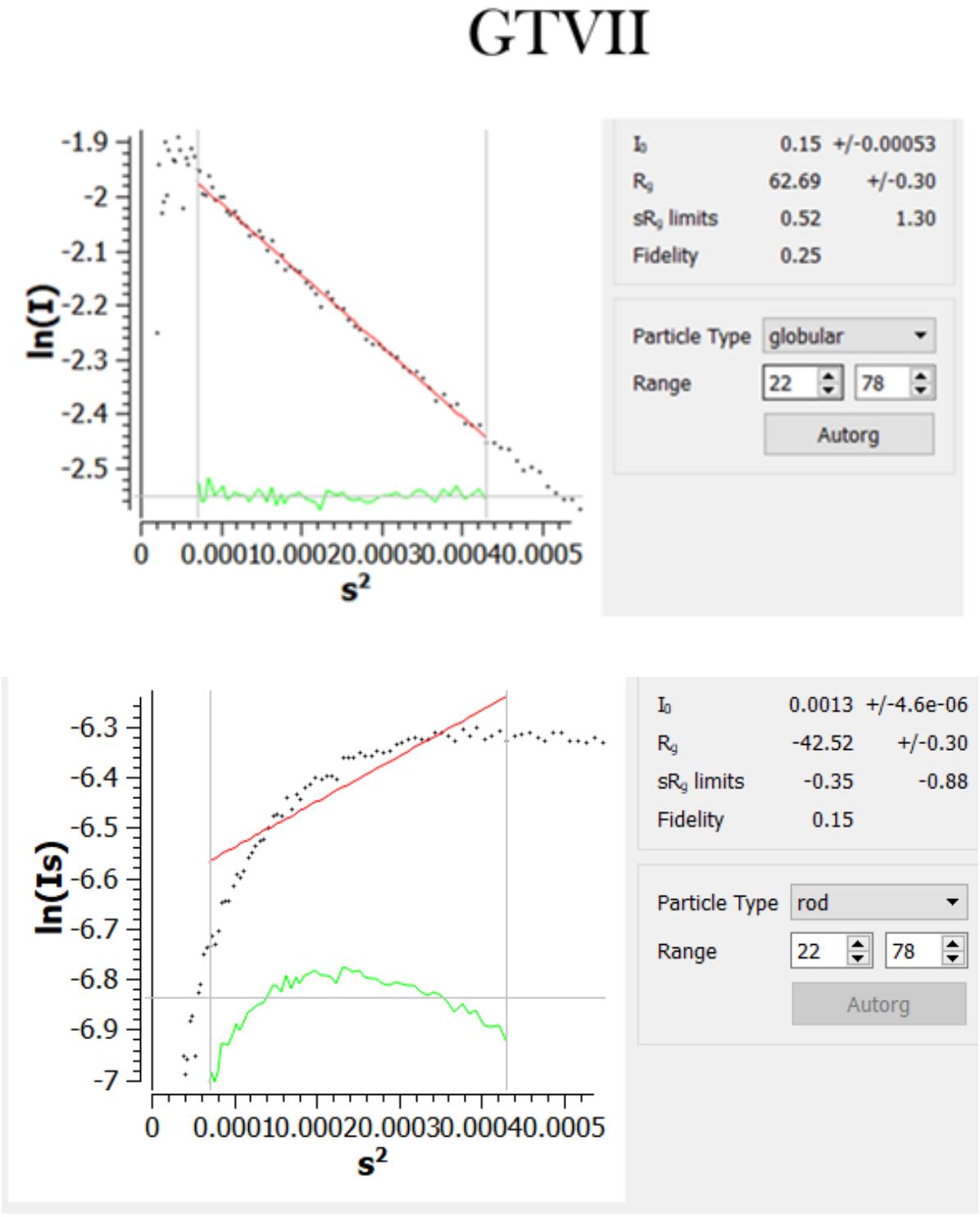


Figure 29: The Guinier plots of La Sota and GTVII at pH8, comparing best fit lines of globular and rod conformations to the data. These suggest that they are more likely to be in a globular conformation.

Finally the Primus tool allows known PDB structures (and models) to be compared to the SAXs data, by simulating the scattering curves these structures would generate. The structure of NDV F in the post

fusion conformation (3MAW) and PIV5 in the pre-fusion conformation (4WSG) were used for this comparison (see figure 30). There was a convincing fit between the simulated PIV5 trace and the two NDV constructs at pH8. Contrarywise, there was no fit between the simulated data for the post-fusion NDV F and the data generated, suggesting that the protein is in the pre-fusion conformation.

The plots of the protein run at pH5 indicated a more linear protein, with an extended conformation. When compared to the simulated data from the PIV5 Pre-F structure, there was a less convincing fit. However, in comparing the simulated data from Post-F NDV, there is a more convincing fit.

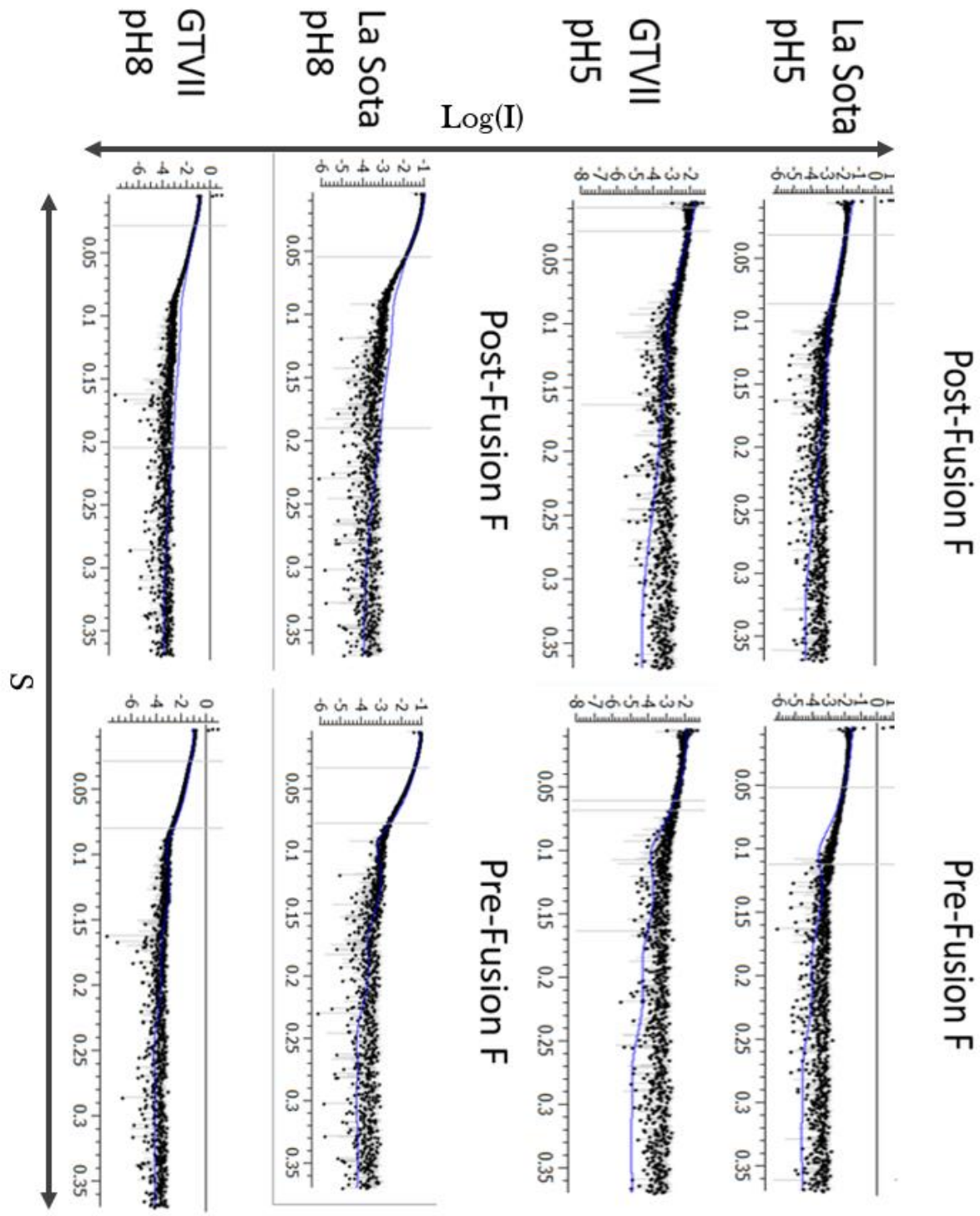


Figure 30: A comparison of the SAXs Data (Black dots) with simulated data generated using the known structures of post-fusion NDV F (3MAW) and pre-fusion PIV5 F (4WSG) (Blue line).

Following the use of the Swiss-Model and Alpha-Fold, we have a model of the NDV F protein to compare the SAXs-derived data to, using the primus tool as before (see figure 29).

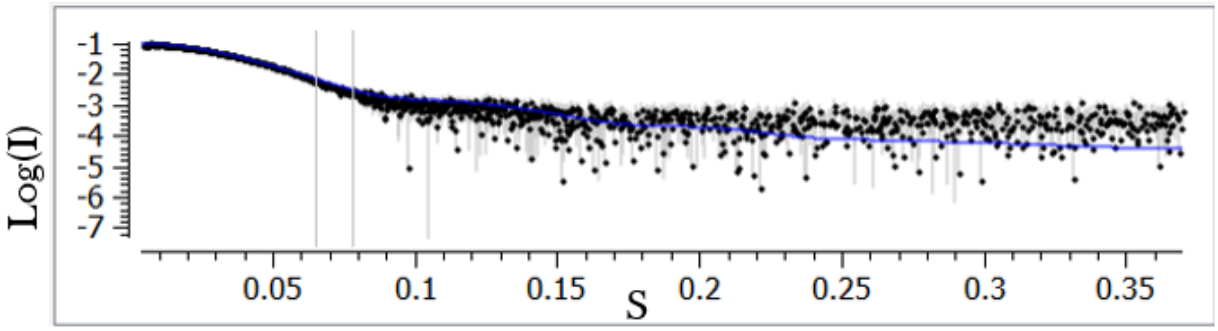


Figure 31: A comparison of the pH8 La Sota SAXs Data (Black dots) with simulated data generated using the model of pre-fusion NDV F generated by AlphaFold.

It appears to quite closely fit with the data generated for La Sota by SAXs, suggesting that there is some validity to the model. We can also infer that the La Sota construct has a structure which makes sense within the context of other Class-1 fusion proteins. Fitting the GTVII model to the F protein achieved a similar alignment. With all of these comparisons, a measure of the quality of the fit should have been taken but was not at the time.

Discussion:

The models produced by Swiss-Model and Alpha-Fold are useful in that they show how the trimeric F protein could look. However, the Swiss-modelled F proteins could exhibit model bias, where the residues are arranges so as to better fit the template, rather than how they might actually be arranged. The De-novo structures created by AlphaFold2, less the curious splaying of the HRB domain, are potentially more useful for homology modelling, gaining an insight into what the protein looks like in the pre-fusion conformation and to assist in mapping across any stabilising mutations. AlphaFold2's method of generating structures and the similarity of models produced to the solved structures make it a good benchmark to which we can compare any solved structure. The query raised by the mis-positioning of the HRB linker domain does imply a question as to the veracity of the rest of the model. However, as the software identified these areas as being of low confidence, there appears little reason to discount the remainder of the model. The disorder of the domain, linking the HRB domain to DI, is present in the post-fusion conformation of the protein, though not bent at such an angle. It could allow the protein to re-fold during the transition from pre-to the post fusion conformation (8, 49). The

models of Class I fusion proteins transitioning from pre-fusion to post fusion, in order to fuse the vial envelope with a cell membrane, do suggest a transition state where the HRA domain pivots while HRB remains fixed (see figure 32). Perhaps the AlphaFold2 software picked up on this transitory state when it computed the relative placement of HRB and the linker domain.

The formation of crystals on the first attempt to crystallise the protein appeared to set up a successful aspect of the project, but was a false portent of what was to come. Despite all best efforts, all attempts to replicate the formation of crystals failed. Several fusion proteins were only crystalised with neutralising antibody (47), even then reproducibility of the crystals was problematic. Most notably, Chen et al failed to get crystals of pre-Fusion NDV F (8). It remains a very stochastic process.

The lack of reproducibility of the crystals remained a conundrum. The construct and vector remained constant (though for completeness, crystal trays were set up with both GTVII and La Sota). In retrospect, it would have been sensible to begin re-working the construct at an earlier stage in the project. By the time work began on this, it was halted due to the pandemic, then constrained timelines. Suggested re-designs of the constructs included lengthening the HRB domain to decrease solubility, also making point mutations to replace larger, more flexible residues on the protein surface with shorter ones. This would have to be modelled first, to ensure that these do not interfere with the wild-type structure of the protein.

If the crystals that grew were actually NDV F protein, it raises the question of why La Sota and GTVII did not form crystals as they were subjected to identical conditions. Otherwise, if what crystallised was a one-time contaminant, that would explain why nothing else crystalised. BSA contamination would have come from the cell media and could be eliminated by more rigorous wash-steps during purification. Anecdotally, colleagues did not report issues with reagents at that time.

The in-ability of Phenix to solve the structure has also raised a number of questions. Phenix is a very well-regarded program used routinely to process data from XRC. However it is very challenging to determine a structure from low resolution data with multiple molecules in the asymmetric unit (43). With more experience and many attempts to find an appropriate search strategy, it could be possible to determine the structure. Use of pre-fusion F protein structures of PIV5 and RSV as models for

molecular replacement could have been unsuccessful due to the low homology (33% by amino-acid sequence) between NDV F and the two models used. Use of the pre-fusion NDV F structure, as a model, was also unsuccessful.

As a relative novice to both the technique and the methods of processing it, it is very possible that mistakes were made during analysis. However, when attempted by other members in the group, the same results were achieved. This further suggests that there is something at the sample level that went awry. There is a possibility that the crystals could have been contaminant. If so, it did not match any unit-cell in the PDB database, suggesting that whatever crystallised did so in a different space group or not a contaminant (48). A contaminant in protein samples used to set-up crystals subsequently could have been the reason why crystals did not form again. There are more ways to solve the structure of a protein, but the relative inexpensiveness of X-Ray crystallography made it a good starting point.

SAXs offered a chance to learn more about the structure of the protein in solution. At pH8 both constructs appear largely folded and globular, which fits generally with the literature on the structures of Class I Fusion proteins in the pre-fusion conformation (12-15). The unfolding that occurred at the lower pH also fits with the idea that the change of conformation can occur at a lower pH (48). This was further supported by the similarity of the pH 5 traces to post-fusion F.

The similarity of the pH8 traces of both La Sota and GTVII, to PIV5 further suggested that the protein was in the pre-fusion conformation. This was further supported by the difference between the plots of the SAXs data and the post-fusion structure of NDV (3MAW).

SAXs also offered the opportunity to validate portions of the models generated by Alpha-Fold, with the trace generated by Chromixs for the AlphaFold2 model quite closely matching the data from La Sota. This adds to the confidence with which the AlphaFold2 model can be viewed. However the splaying of the legs is much less credible and could be an artefact of the protein only being submitted to the AlphaFold2 server as a monomer.

The similarity of the samples to the structures the protein is predicted to have, and due to the failure to reproduce crystals and the problematic nature of the structure solution of X-Ray Crystallography to

determine the structure, it was decided to attempt to solve the structure of NDV F using single particle cryo-EM analysis described in Chapter IV.

Chapter IV:

A-Tissue, A-Tissue, we all fall-apart!

Introduction

Following on from the unsuccessful attempts to solve the structure of the protein by crystallography, the decision was taken to shift focus towards using single particle Cryo-electron microscopy (Cryo-EM).

Electron microscopy first emerged around the same time as Doyle and Minnet were characterising NDV, with the first Transmission Electron Microscope (TEM) being built in 1931 by Ernst Ruska (1-3). While early TEMs were limited in resolution, by the 1980s they could achieve a resolution of 100Å. Improvements to the microscopes, particularly to lenses and electron detectors, while increasing the stability of the stages played a large part in this (4). In 1981 vitrified sample preparations began to be used, allowing for an increase in the amount of data that could be collected as the vitrification gave the sample a measure of protection from the radiation. Furthermore, it would "Freeze the sample in time," which is potentially very useful for meta-stable proteins (5).

As a technique, it has grown in both sophistication and popularity, in recent years. The technique not require crystals for structural determination, plus the so-called 'Resolution Revolution' have contributed massively to its popularity. Indeed, in the last 12 years alone, the average Cryo-EM structure resolution has improved from 15Å to 6 Å, with one structure in 2021 achieving a resolution of 1.7Å (6-7).

Other advantages of Cryo-EM identified include the ability to screen more dynamic protein structures, to screen larger structures and it requires smaller amounts of sample (8). This is especially important for proteins that cannot be readily expressed or purified. At lower-to-moderate resolutions, Cryo-EM is able to produce more accurate maps of the 3D structures as phase information is not lost – unlike in X-ray crystallography (9).

Negative stain EM (neg-EM) pre-dates Cryo-EM which can be used to characterise biological samples through and producing high contrast images of protein on grids using radioactive salts, such as uranyl formate, to apply the contrast (10). Often it is used as a first step in characterising sample for Cryo-EM and X-ray crystallography and to give a point of comparison for any structures seen later. It has featured

in most of the structural work carried out on the Class I Fusion proteins, including NDV, PIV5, RSV and Nipah (11-14).

Although the structure of the pre–Fusion Nipah F protein was solved using X-ray crystallography, use was made of negative stain EM. In this, six-unit assemblies of the F protein were discovered, mirroring both the six-unit structure that they managed to crystallise and the multi-unit structure seen in the CDV micrographs, see figure 30 (14).

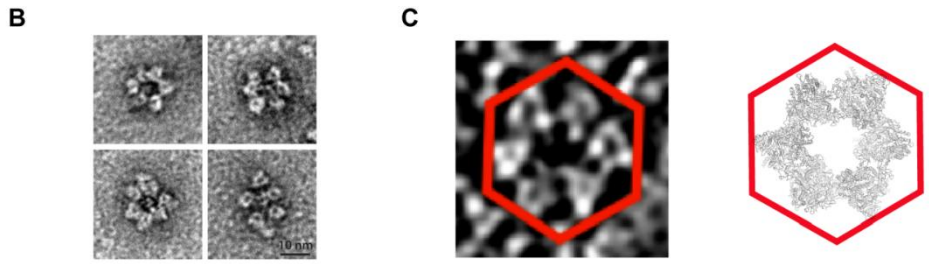


Figure 30: Some of the multi-unit structures formed by NIV F during EM. B: Micrographs, C: 2D classes, D: The model generated. Reproduced from (14). See Appendix C

Negative stain EM formed part of the structural analysis on RSV F, see figure 31, where it was used both as a preliminary technique prior to setting up crystal trays and to explore the binding of neutralising FABs specific to the pre-fusion conformation of the virus (13).

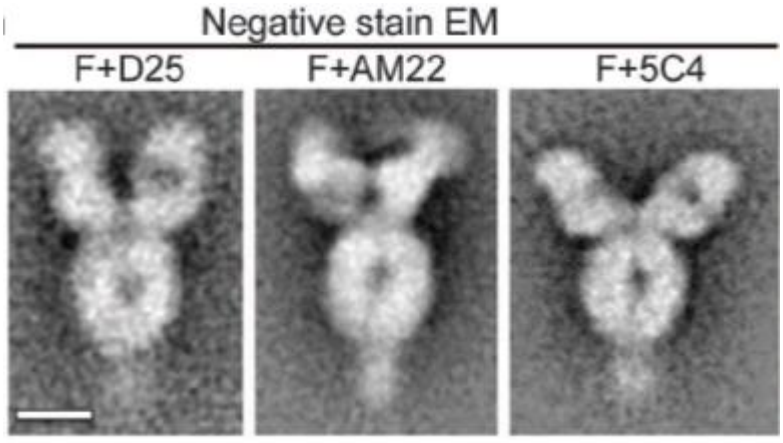


Figure 31: Negative stain EM micrograph RSV F protein bound with FABs. Reproduced from (13).

In 2006 Swanson et al used cryo-EM to further investigate the properties of a pre-fusion conformation of NDV F. Images of the protein suggested it was in conformation similar to known pre-fusion structures, however the technology was not there to collect and process data to the high resolution that is now possible (11). Furthermore there was no guarantee that the grids made would have permitted a high resolution structure to be built.

Cryo-EM has been of use in solving several Class I Fusion proteins, most notably, Canine Distemper Virus (CDV). This was carried out using F protein, stabilised with a GCN(t) domain, expressed in HEK293 cells (15). As such it would imply that transferring the construct already in use for XRC would also work for cryo-EM. In looking at the micrographs, it was noted that the CDV F trimers were often arranged in multi-unit aggregations, resembling pizza slices:

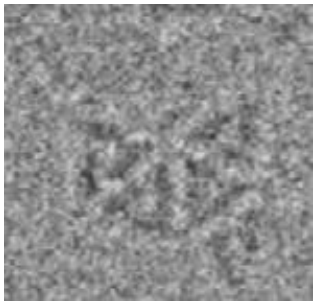


Figure 32: CDV F protein on cryo-EM grids, reproduced from (15).

Despite this, data were collected from cryo-cooled grids where the trimer F protein was sufficiently ordered that the structure could be reconstructed to 4.3 Å using Relion (15).

Chameleon:

Diamond Light Source and the RFI have in recent years purchased the Chameleon system (16) for preparing grids for cryo-EM. It is an automated grid preparation system where the sample is dispensed from a nozzle onto the grid, then cryo-cooled in under a second (16). This was utilised later in the project as it was hypothesised the shorter freezing time could reduce the disruption of the trimeric protein.

Pegylation:

Pegylation is the conjugation of polyethylene glycol (PEG) to a protein. Both N and C-terminal regions of a protein, and lysine residues, are possible targets for pegylation (17). Non-specific pegylation was first demonstrated 40 years ago and has since been shown to reduce protein aggregation, increase protein stability and half-life (18). Furthermore, the fact that it has been shown not to alter the structure of proteins makes it an ideal candidate to for cryo-EM applications (19).

Results:

Negative stain EM

Performing negative stain EM (NEM) was a logical first step in preparing the sample for Cryo-EM. Grids with F protein at pH8 and pH5 were prepared with 0.05mg/ml of protein and stained using uranyl fumarate (See Chapter II).

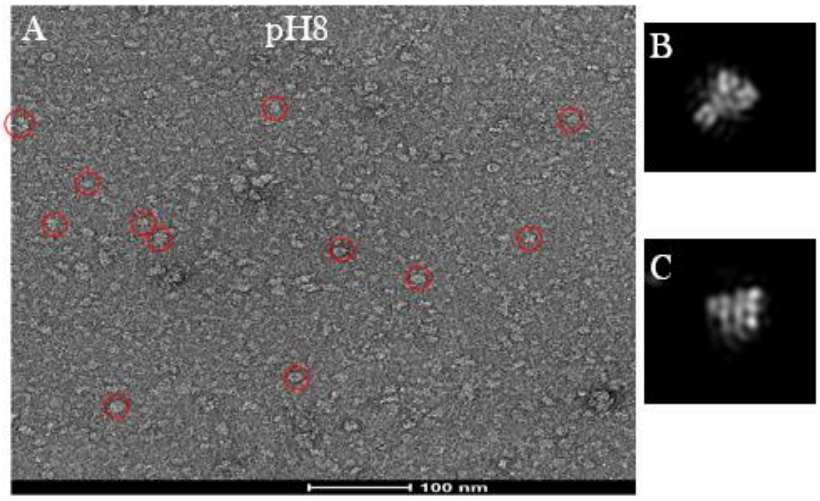


Figure 33A: A negative stain EM micrograph of La Sota F protein at pH8 with example particles circled in red. 33B and 33C: 2 of the 2D classes generated from the micrographs using CryoSPARC (20), from particles isolated from micrographs.

Looking at the micrographs (see figure 33), what looks like a number of trimeric F particles are visible, appearing to be in the pre-fusion conformation. Rudimentary processing in CryoSPARC (20) revealed what could be 2D classes that look like the pre-fusion F trimers.

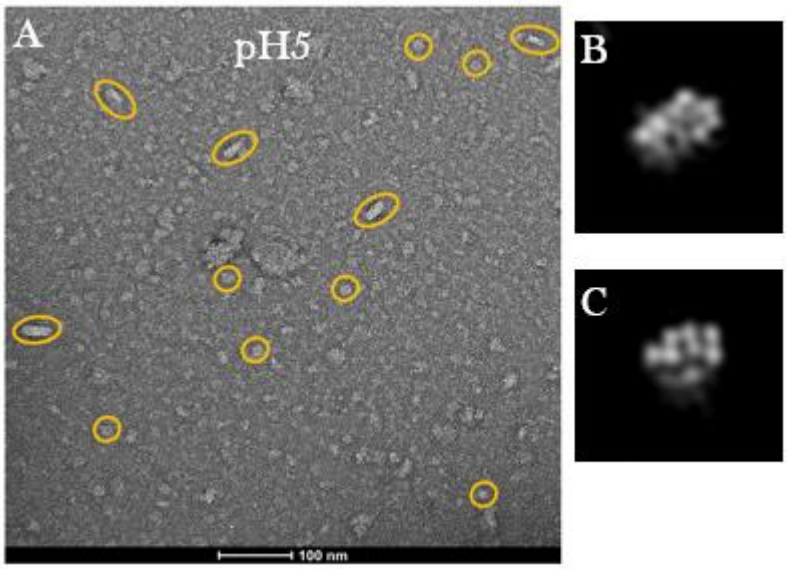
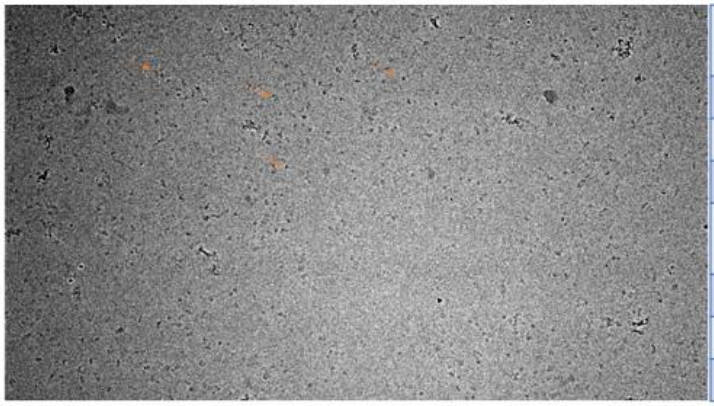


Figure 34A: A negative stain EM micrograph of LaS F protein at pH5 with example particles circled. 34B and 34C: 2 of the 2D classes generated from the micrographs using CryoSPARC (20), from particles isolated from micrographs.

The pH5 micrographs (figure 34) display a mixed population of more elongated particles with other, smaller particles that could be the semi-cylindical particles top down. Here the CryoSPARC 2D classification was less clear. The more cylindrical shape resembles the post-fusion conformation, but not conclusively.

Initial attempts at Cryo-EM

Once the NEM had been carried out, and the sample appeared to be of sufficient quality, cryo-Em could be attempted. The first round of Cryo-EM was very encouraging, using 0.6 mg/ml of La Sota F protein, see Chapter II the grid preparation. The table accompanying figure 35 shows the collection parameters. Micrographs, if somewhat sparse in particles, were still useable for extracting useable particles.



Protein Concentration (mg/ml)	0.6
Raw Pixel Size (A)	2
Accelerating Voltage (kV)	200
Spherical Abberation (mm)	2.7
Total Exposure dose (e/A^2)	124
Grids	Quantifoil Gold
Mesh	200
Particles	11,425

Figure 35: An example micrograph from the initial round of Cryo-EM along with the data collection parameters Examples of the particles that were selected for sorting into 2D classes are indicated with orange arrows.

Micrographs, when processed in CryoSPARC yielded 11425 particles, sorted into 2D classes. 36 of these were used for further processing, see figure 36.

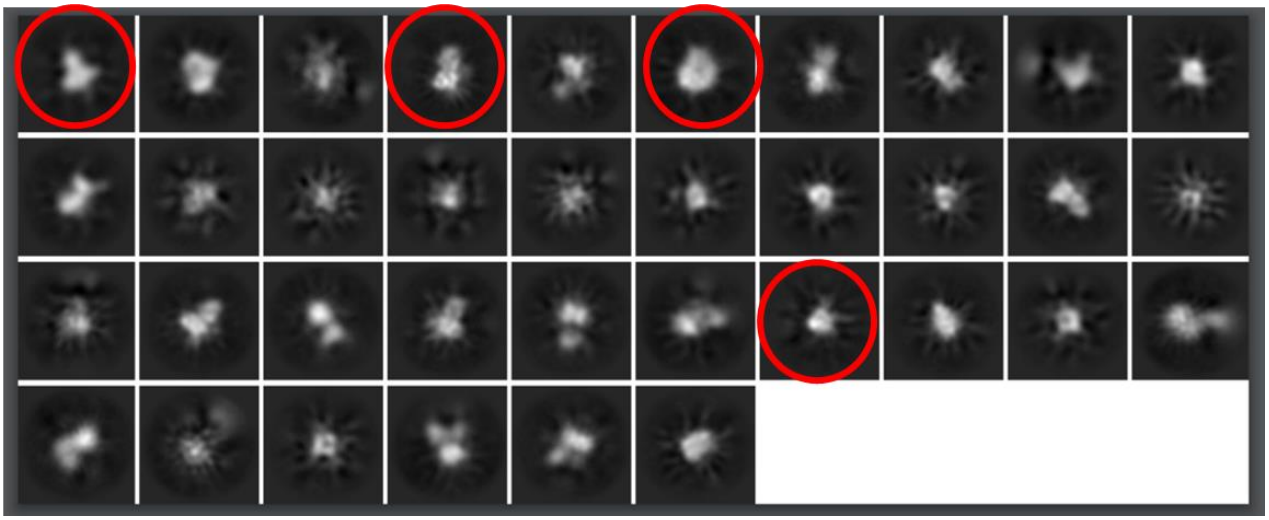


Figure 36: The 2D classes that particles from the micrographs were sorted into (by CryoSPARC) during the first round of Cryo-EM screening, with some of the more promising classes marked with red circles.

Several of the 2D classes displayed the globular, triangular motif, similar to the known structure of PIV5, many with a stalk. Some of the other classes could too be F protein in different orientations.

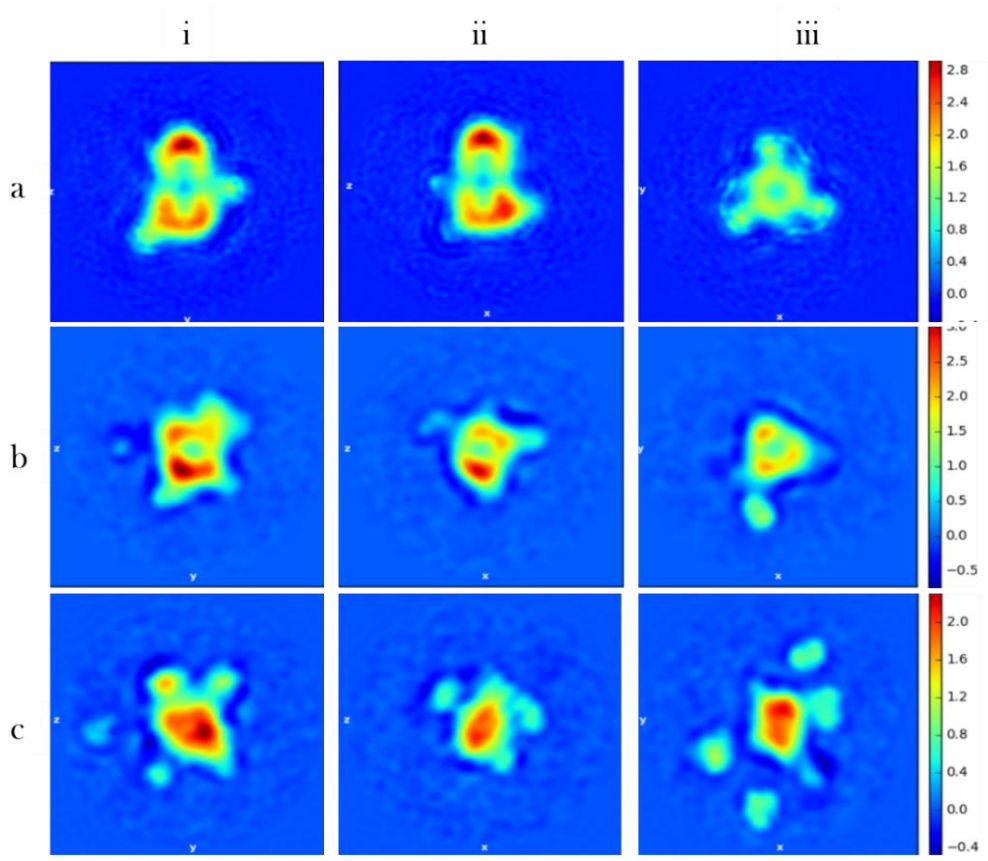


Figure 37: The three 3D class averages generated by CryoSPARC from the 2D classes shown in figure 36 the initial round of Cryo-EM. A, B and C refer to the different classes. Images aiii and biii closely resemble the pre-fusion structure of PIV5 from the top-down aspect.

These were then used to generate three 3D models, which are shown in figure 37. The three models generated from the 2D classes suggested that of trimeric F protein on the grid, indeed the panel at the top right of the above figure appeared to very closely resemble a top-down view the PIV5 F protein.

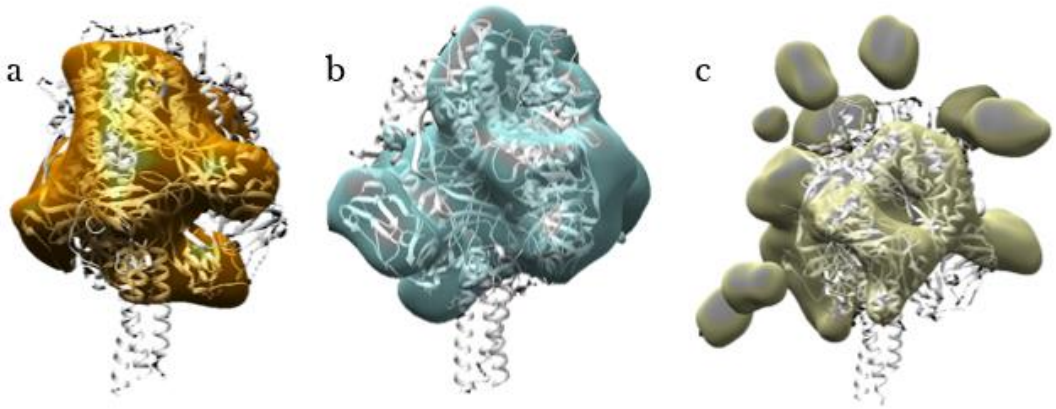


Figure 38: Three 3D-reconstructions, each generated from the 3D class structures shown in figure 37, fitted against the pre-fusion structure of PIV5.

The models generated were then fitted against the known structure of PIV5, however further processing did not lead to high resolution structures. See Figure 38. It was thought that a larger data-set and further grid optimisation would allow for a higher resolution structure.

A failed attempt at data-processing:

Following a number of screenings to optimise the grid preparation, a larger data-set ($1.7x10^6$ particles) was collected on the Krios at 300kV. This was undertaken at the Oxford Particle Imaging Centre.

Protein Concentration (mg/ml)	*5*
Raw Pixel Size (A)	0.8455
Accelerating Voltage (kV)	300
Spherical Abberation (mm)	2.7
Total Exposure dose (e/A^2)	63.34
Grids	Quantifoil Gold
Mesh	200
Particles	1766318

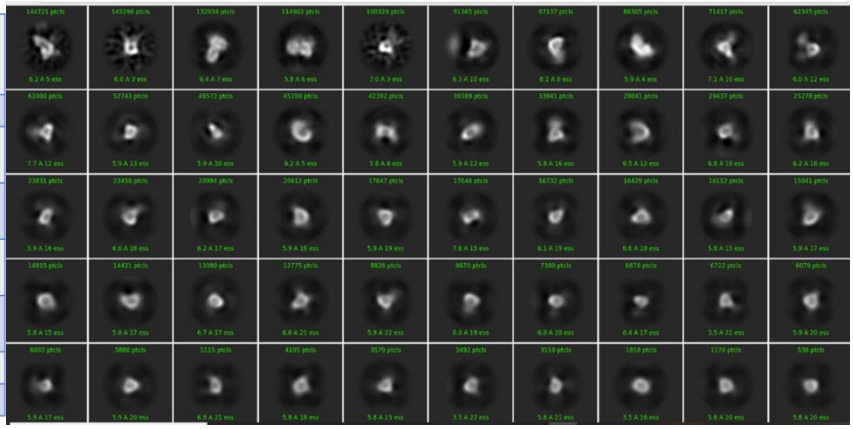


Figure 39: Some of the 2D classes generated, in CryoSPARC, from a second cryo-EM dataset. This led to it being suggested that the protein was monomeric. 5mg/ml of protein is correct.

Looking at figure 39, the 2D classification suggested a large number of potential globular proteins, but also a number of classes where the protein had a 'C' shaped profile. When these were used to generate 3D models of the protein, with further rounds of 3D reclassification and refinement (figure 40), it became apparent, from the images of the 3D model, that F was not trimeric.

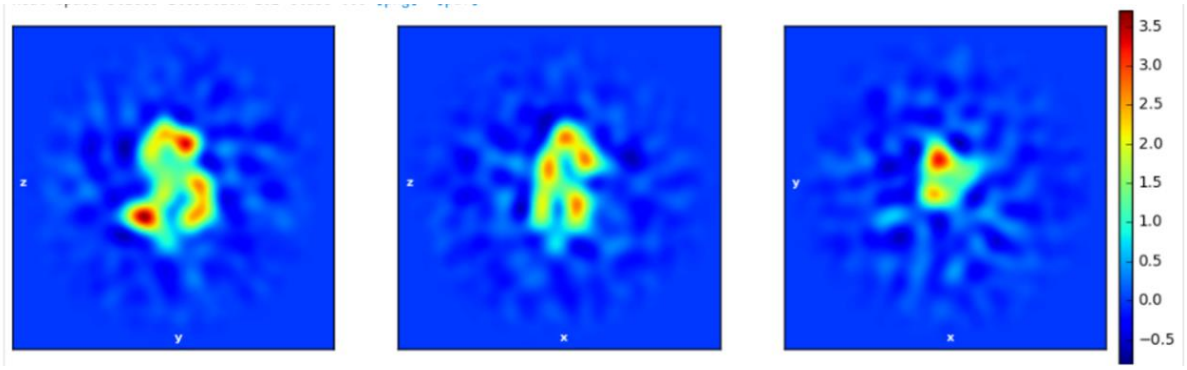


Figure 40: one class of 3D projections, modelled by CryoSPARC, from the selected 2D classes shown in figure 39.

Examining the model in 3D (see figure 41) it was even more obvious that most of the reconstructions were not trimeric.

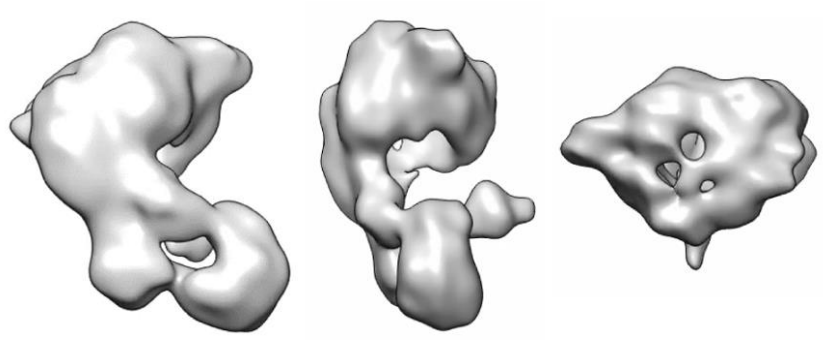


Figure 41: The 3d model generated from the cryo-EM generated models shown as a 3D volume in figure 40.

When fitted against a monomer of PIV5 (figure 42), the calculated 3D maps suggested that the particles were monomeric fragments of the F protein. However, as confidence in these reconstructions was low and they could constitute the averaging of noise, this is only speculation. The 'C' shape lines up with the arc of DIII, DI, DII of the PIV prefusion monomer.

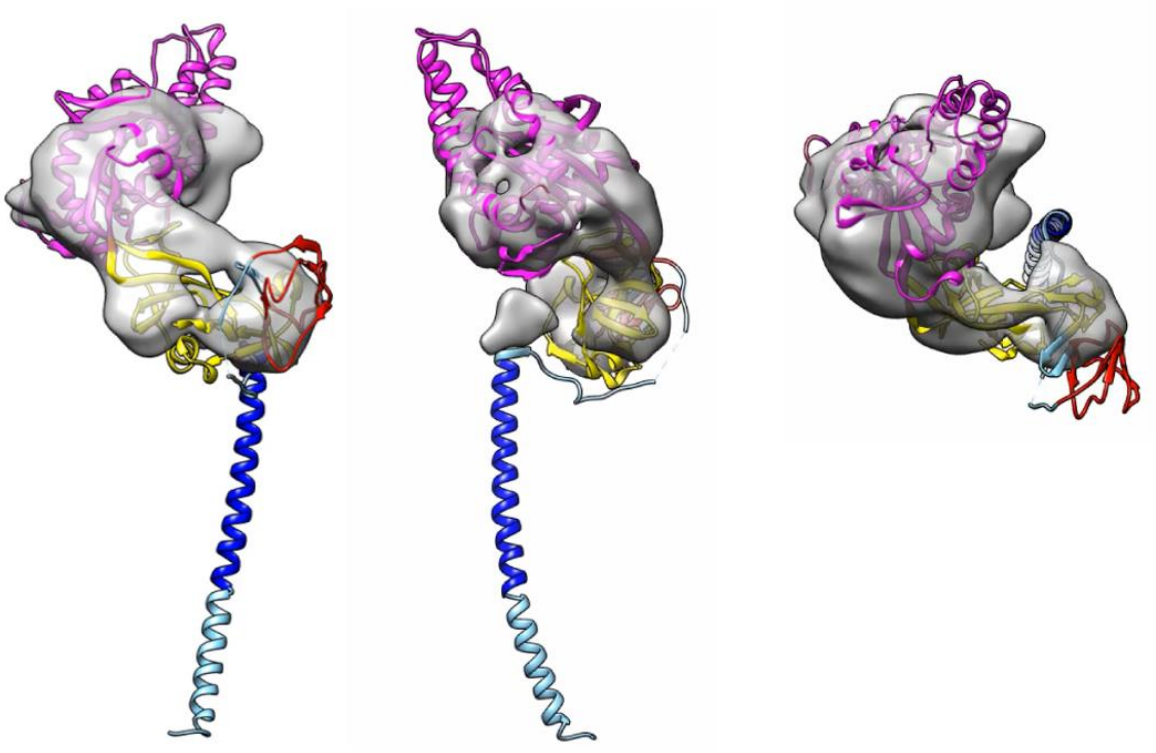


Figure 42: The model generated from the cryo-EM data (grey) fitted against the monomeric PIV5 F model.

Cross-Linking

With the protein potentially forming monomers under the conditions of preparing Cryo-EM grids, cross-linking was the natural next step to solve this problem. Cross-linking with Dimethyl pimelimidate (DMP) was attempted on 0.1Mg/ml of F protein but did not show much success (see figure 43).

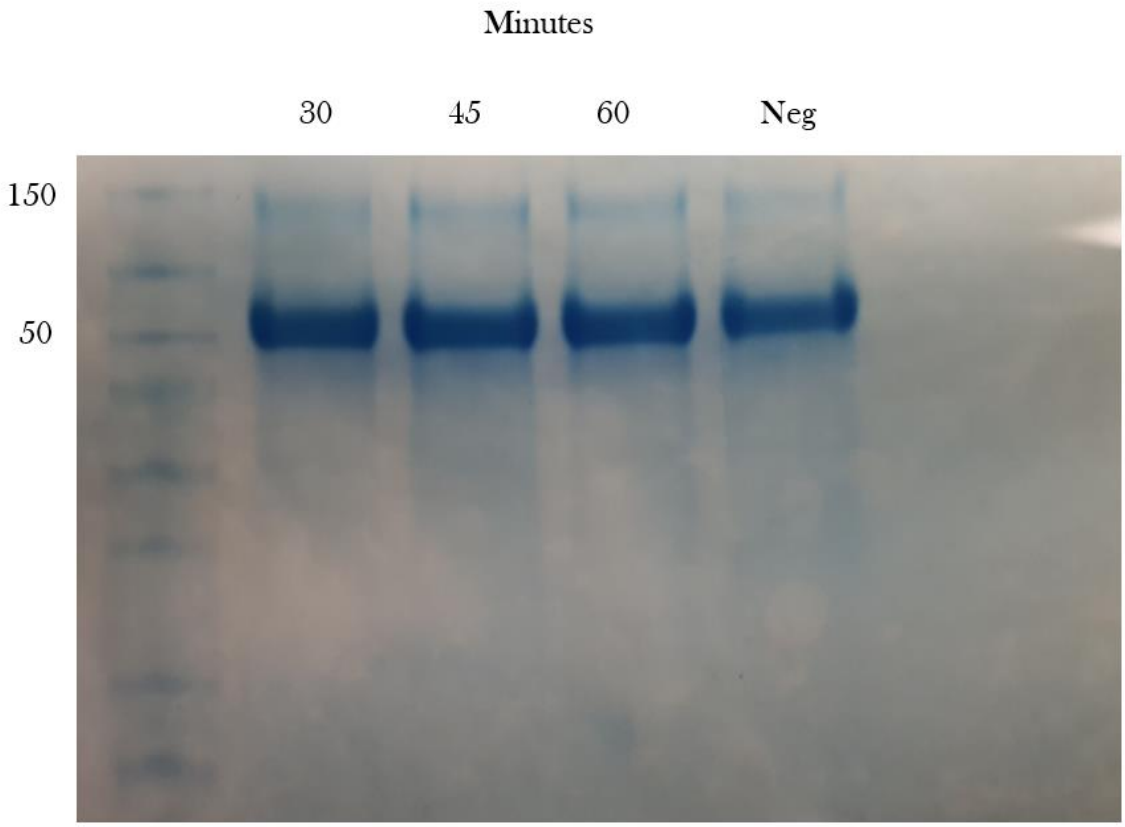


Figure 43: LaS F protein at 0.5mg/ml cross linked with 10-fold molar excess DMP for the time indicated. Quenched with Tris, run on an SDS-PAGE gel at 100V.

Over the recommended time course, the crosslinker had very little impact on the proportion of trimeric protein. DMP consists of two imidoesters, which bind amine groups, separated by 7-atom spacer arms (21). It is possible that the spacer was too short to span the space between monomers. Perhaps related compounds with variable spacers could be explored, or a higher ratio of DMP could have been used.

Glutaraldehyde:

With DMP ineffective, it was decided to try and use glutaraldehyde. An preliminary experiment was carried out to determine the best concentration of Gluteraldehyde to use. Concentrations of 0.03% to

1% glutaraldehyde were created by serial dilution, and cross-linked for four hours. 0.12% Gluteraldehyde solution with 0.1mg/ml of F protein was found to be optimal.

A concentration of 0.12% glutaraldehyde was chosen to cross-link the protein, but it was necessary to establish the optimum duration of cross-linking. Thus, a time course was run, see figure 44.

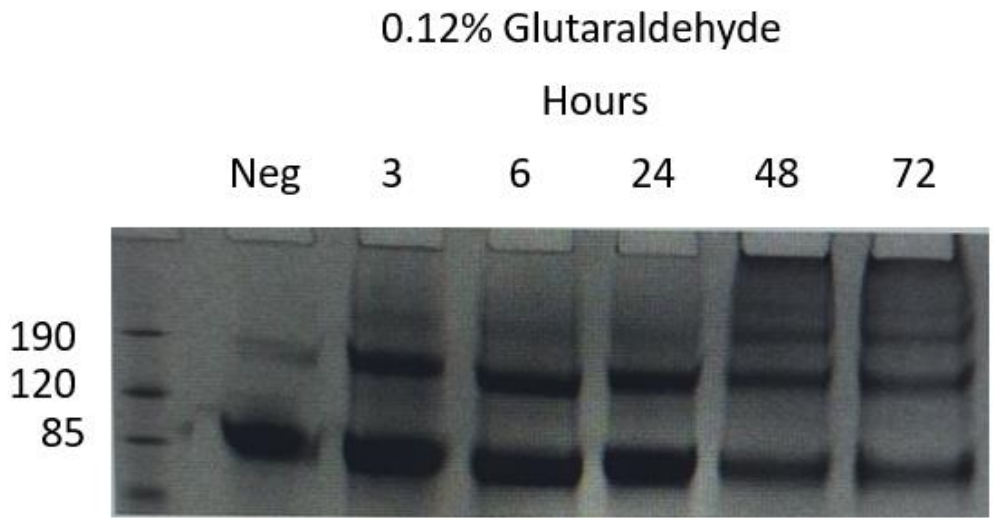


Figure 44: A gel showing the effect of 0.12% glutaraldehyde crosslinking F protein at 0.1mg/ml over time. Note: Molecular markers between gels vary due to supply issues at the time.

A proportion of un-crosslinked protein remained after even 72 hours of cross-linking. This would need removal by size exclusion. By the 48 hour time point, over-crosslinking was starting to appear. The minor differences between 6 hour and 12 hour incubation suggested that the former would be the more efficient.

With the optimum duration established, a larger batch of protein was crosslinked for 6 hours at room temperature before being concentrated and subjected to SEC. See figure 45 for the trace from the SEC and the results of a reducing protein gel using samples taken from the fractions of the SEC.

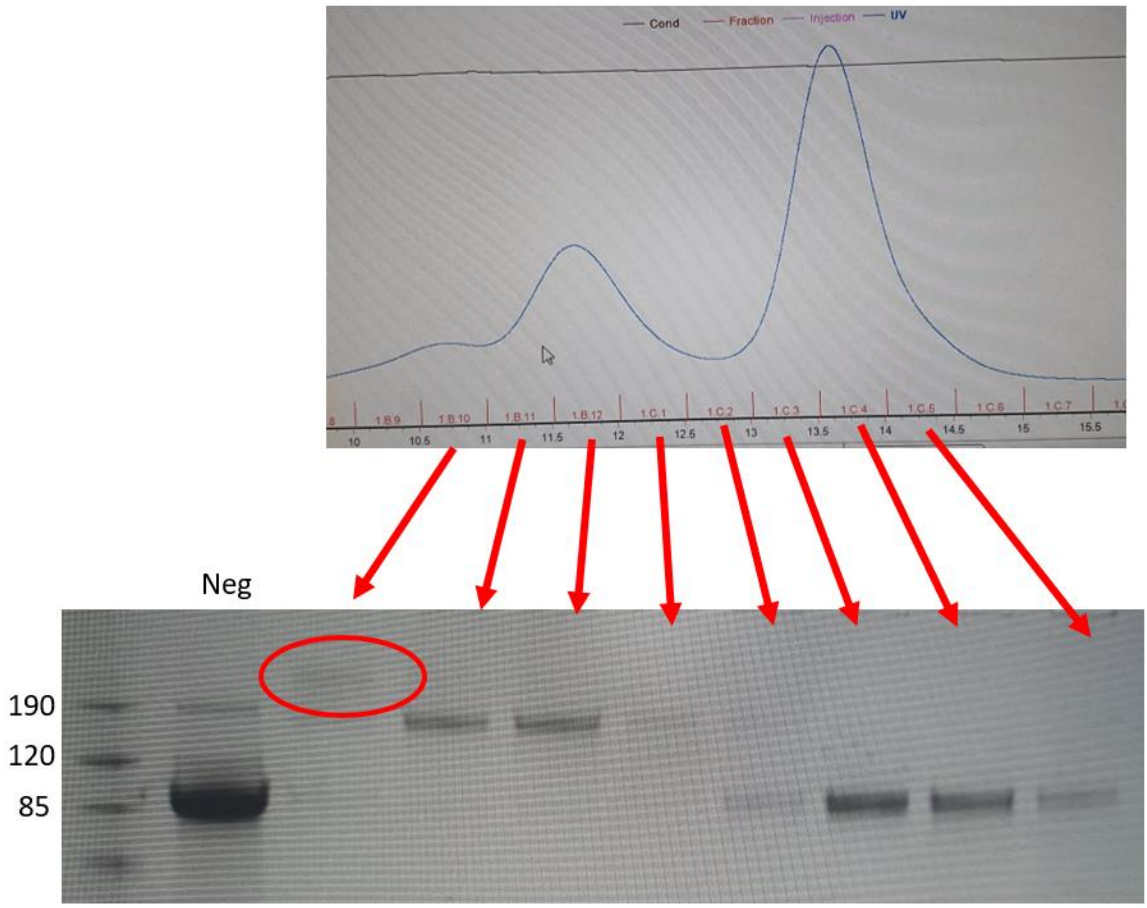


Figure 45: Putative F protein cross-linked with 0.12% Gluteraldehyde for 6 hours before the reaction was quenched. This was concentrated to 1mg/ml and put through SEC on a Superdex S200 10/300. The fractions were collected and loaded onto a SDS-PAGE gel. Un-crosslinked F was used as the negative control. In retrospect, NiV F could have been used as a positive control.

The fractions from size exclusion, when run on a gel suggested that fractions B11 and B12 were trimeric, while fractions C3 to C5 were monomeric. Further attempts at cryo-EM with 'crosslinked' protein were less successful. During screening, all visible particles were too small to be trimers.

Grid optimisation:

Part of the problem could have been grid preparation: the material from which the grids were made, the addition of carbon or the process of loading, blotting and plunge freezing the grid could all have

contributed to the formation of monomers. An extensive array of grids were prepared and screened in order to optimise conditions for data collection with the cross-linked protein, see table 4.

Grid type	Protein concentration (mg/ml)	Other conditions:	Result/comment
Gold 300 Mesh	0.5	10um ZnCl present	Data-set not collected-particles too small
Gold 300 Mesh	0.5	Cross-linked F	Data-set not collected-particles too small
Gold 200 Mesh	0.6	/	Data-set not collected-particles too small
Copper 200 Mesh	0.5	/	Data-set not collected-particles too small
Lacey's Continuous Carbon ultra-thin 400 mesh	0.05 (sic)	Continuous Carbon grids	Poor contrast, too few particles
Lacey's Continuous Carbon ultra-thin 400 mesh	0.1	Continuous Carbon grids	Poor contrast, particles too small

Table 4: A summary of the conditions trialled to optimise the preparation of cryo-EM grids with trimeric F protein.

The inability to find a pairing of grid and protein that might be workable suggested that there was an issue with the sample.

Switch to SF9 cells:

A switch to SF9 cells was made as they were easier to express, and it might have had an impact on the stability of the protein. It quickly became apparent that the purification exhibited different size-exclusion profiles. See figure 46.

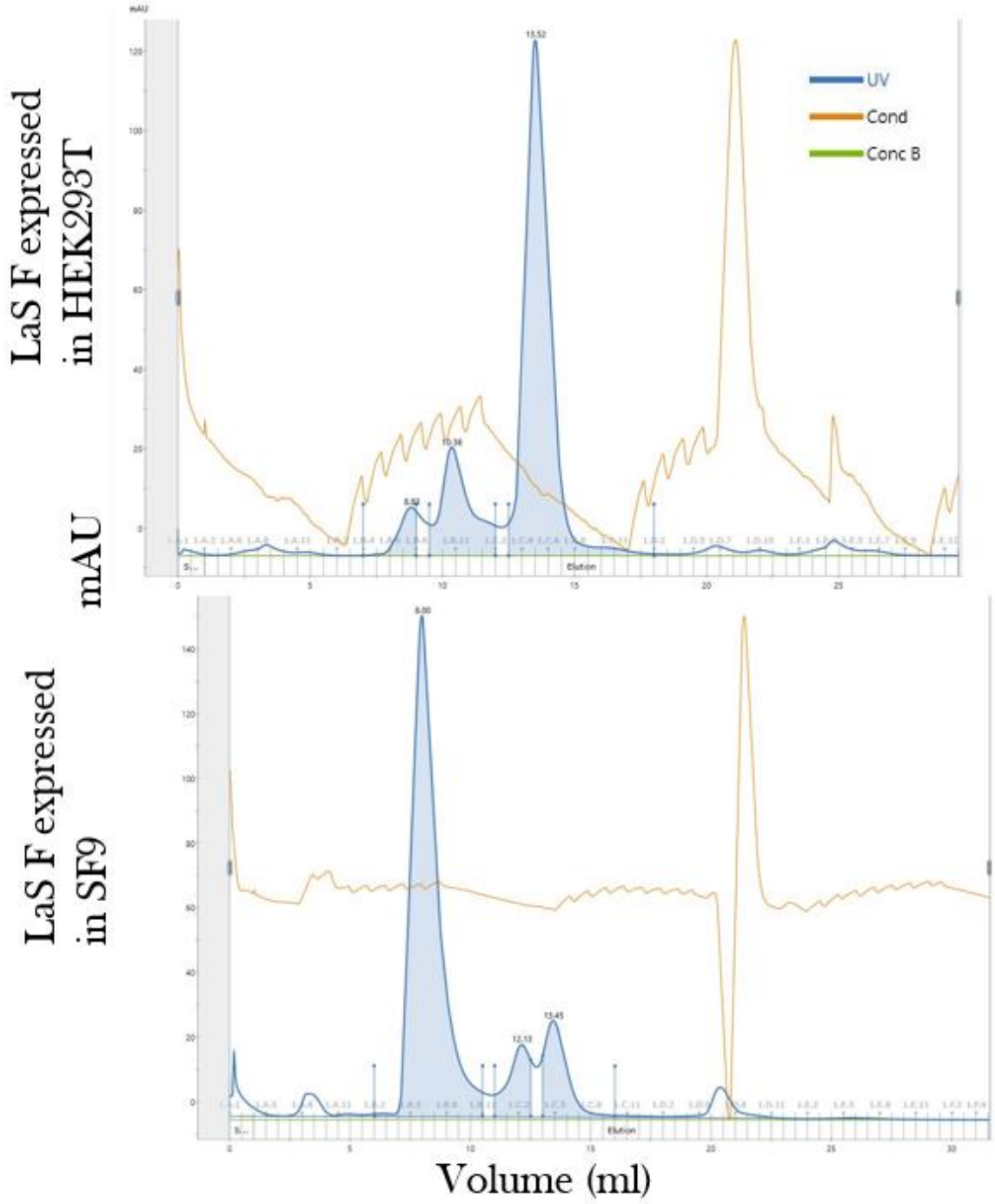


Figure 46: Comparison of SEC traces of LaS F protein expressed in HEK293T (mammalian) and SF9 (Insect) cells, loaded on to a Superose S200 10/30.

SF9 cells were grown at 26°C, for 2 days, where as HEK283T cells were frown at 37°C for 4 days. Perhaps this was less likely to degrade the expressed protein.

Chameleon:

Following the failure of traditionally prepared cryo-EM samples and cross-linking to yield the structure of a trimeric F protein, it was decided to try Diamond Light Source's Chameleon sample preparation system, which is able to load and freeze samples in as-short a time as 100ms (21).



Figure 47A: A screen-capture showing the aggregation of purified LaS F protein at 4 mg/ml upon the nozzle of the Chameleon nozzle during grid preparation.

At the higher concentrations, the F protein was aggregating in the nozzle of the protein Gun that was used to project the sample onto the gids (see figure 47A). Despite this, there was sufficient coverage of the protein to enable data collection at 400ms, 200ms and 100ms, with the protein at the latter concentration appearing to be the most useful.

Initial processing and 2D classification revealed that 200ms and 400ms was not sufficiently fast, most of the protein appeared monomeric. 100ms appeared more favourable, so a larger dataset was collected (see figure 47B).

Protein Concentration (mg/ml)	2
Raw Pixel Size (A)	1.049
Accelerating Voltage (kV)	200
Spherical Aberration (mm)	2.7
Total Exposure Dose (e/A^2)	45
Grids	Quantifoil Gold
Mesh	300
Particles	405,132

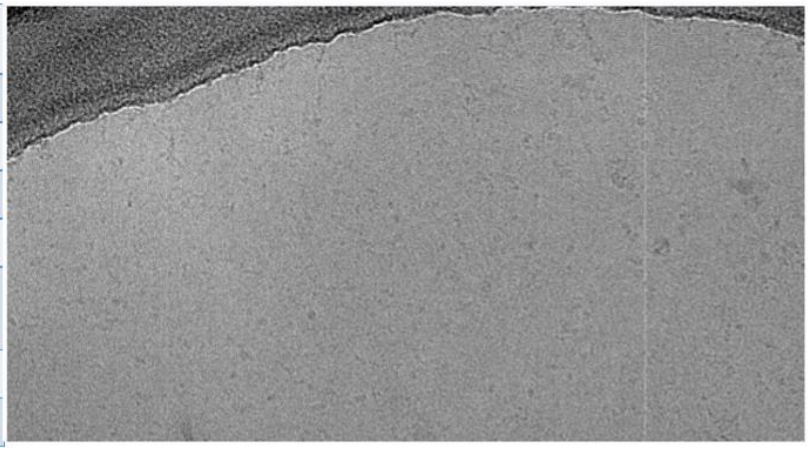


Figure 47B: An example of a micrograph from a grid prepared using the Chameleon, collected during the 100ms data collection along with the parameters used.

The micrographs initially appeared quite encouraging, with globular structures of approximately the correct size for trimers. Using the automated sorting into 2D classes was where things began to go less well.

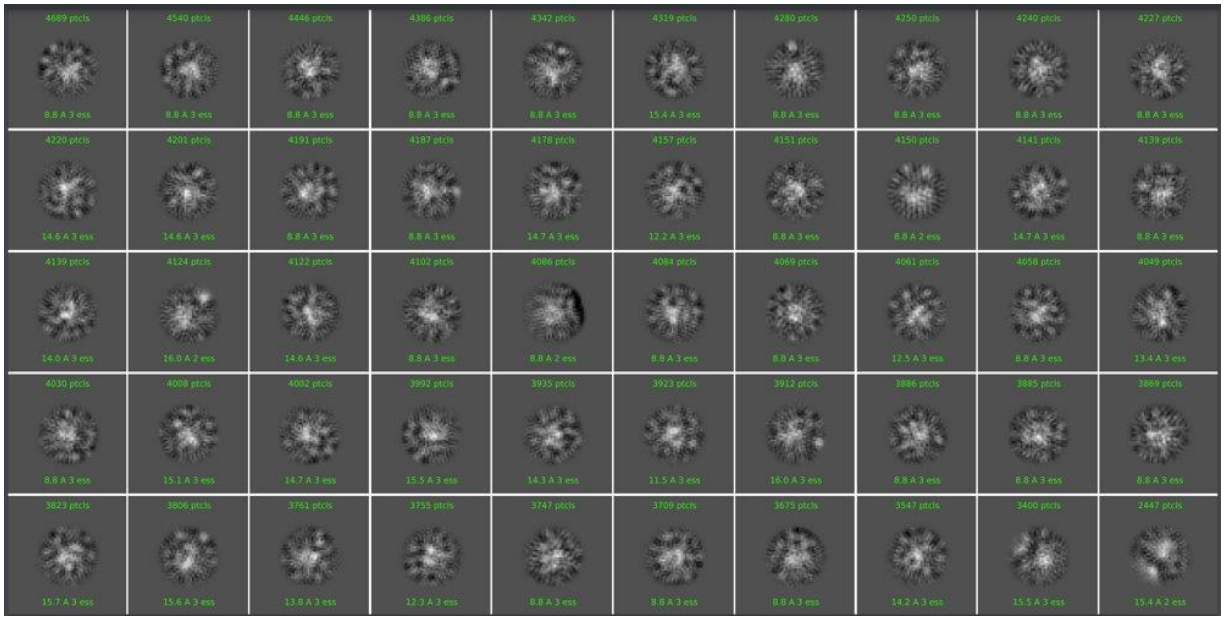


Figure 48: Particles extracted from micrographs from the 100ms Chameleon data set sorted into 2D classes using the CryoSPARC (20).

The 2D classes were somewhat more difficult to determine and did not make for easy processing. A round of manual particle picking was used to generate a further series of 2D classes, see figures 48 and 49.

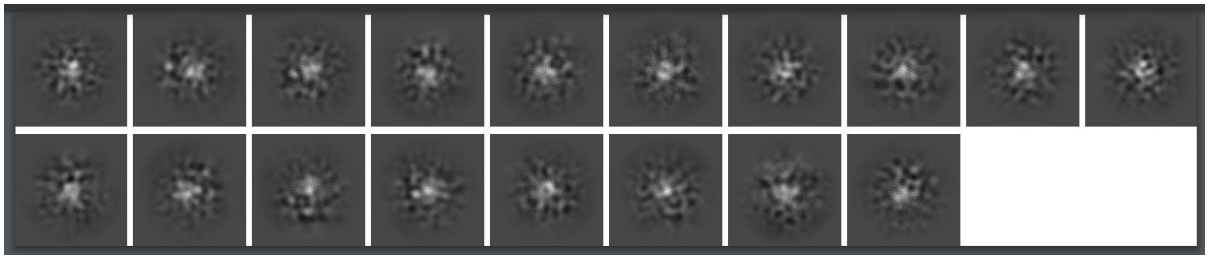


Figure 49: The 2D classes from figure 48 that were selected to generate the 3D models seen in figure 50.

Three models were created form these particles in CryoSPARC, see figure 50. Class 2 could be discounted as it appeared to be an aberration. Class 0 and Class 1 appear to show a stalk with folded domains, which superficially resemble the pre-fusion structure of monomeric F protein.

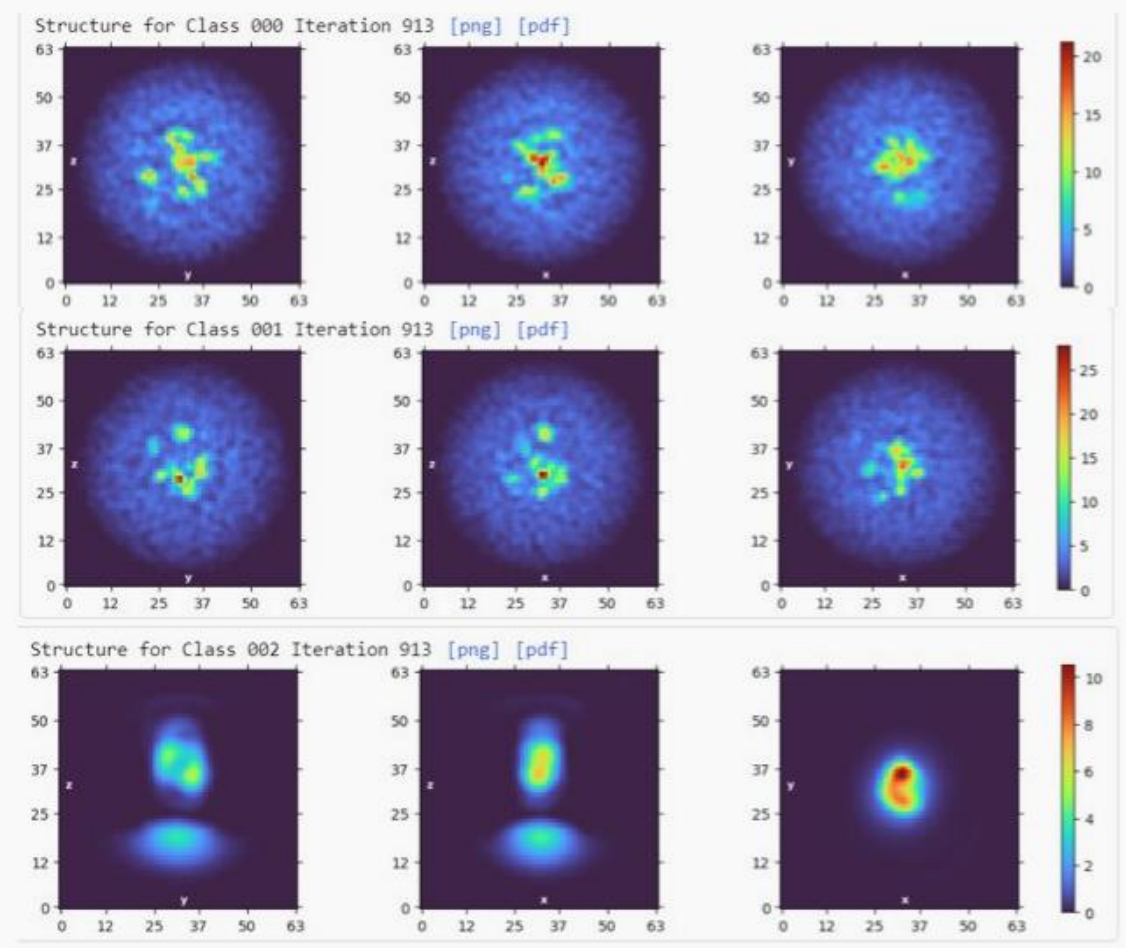


Figure 50: The classes modelled in CryoSPARC from the selected 2D classes in figure 49.

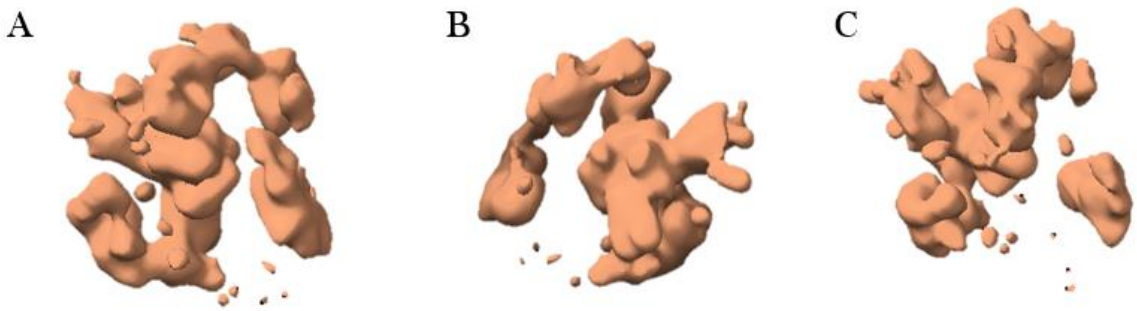


Figure 51: The volume mask generated in CryoSPARC from the class 0 particles shown in figure 50.

The model generated for class 0 (figure 51), unfortunately appeared to be a discrete series of "limbs". Conceivably, each of the limbs could be formed by a monomer that aggregated in the nozzle of the Chameleon, but this is not convincing. It is more likely that this is 'noise'.

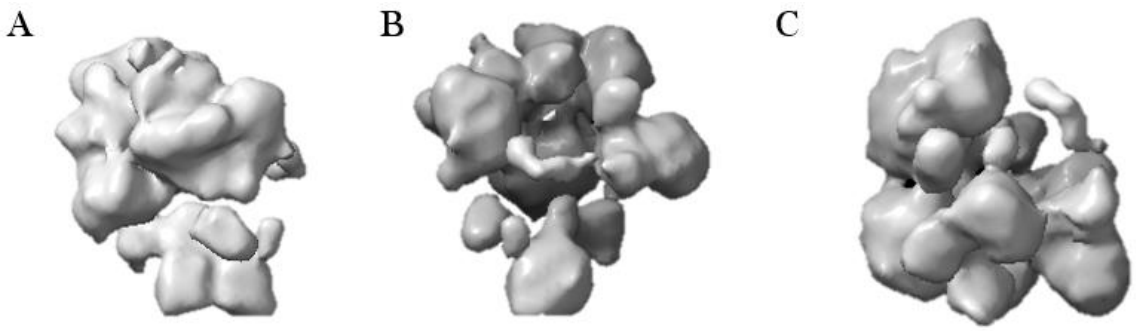


Figure 51: The volume mask generated from the class 1 particles in CryoSPARC shown in figure 50.

The model generated from class 1 (figure 51), is much more globular and shares much of the curvature seen in the monomeric form of the F protein. However this idea loses traction when compared with what a monomer in the pre-fusion conformation is supposed to look like: See figure 52 for an alignment of the La Sota F monomer (Generated by AlphaFold in Chapter III), with the volumes from figure 51.

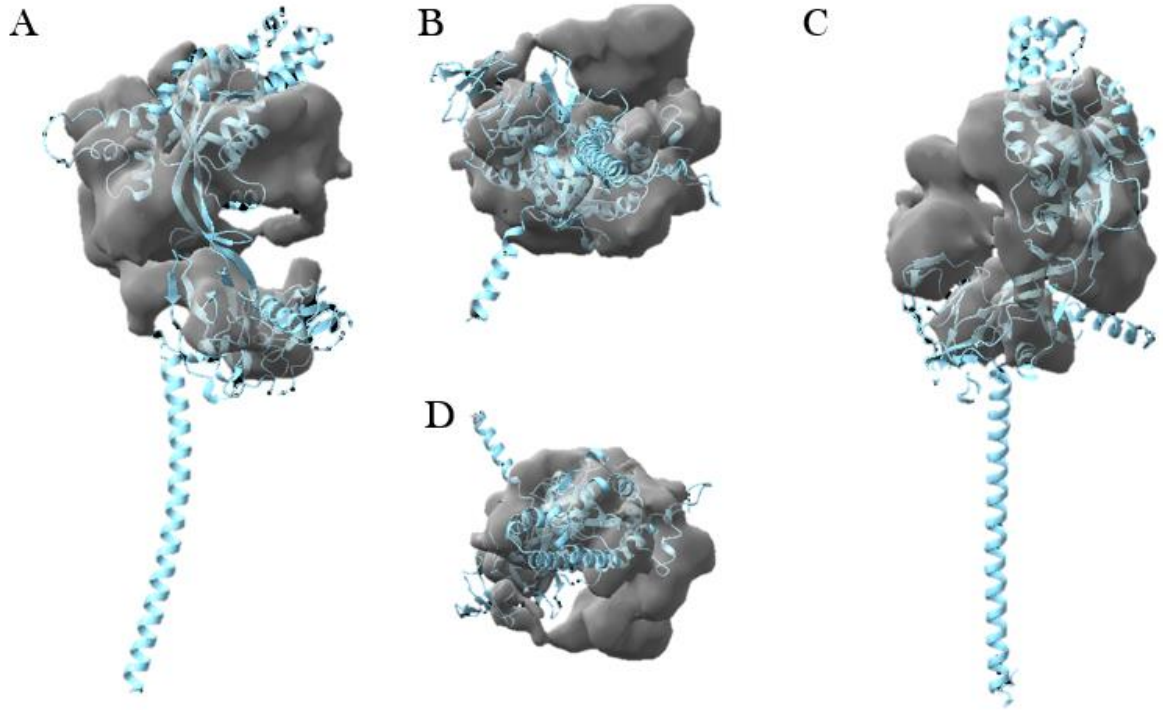


Figure 52: The volume mask generated from the class 1 particles (figure 51) mapped against the Alpha-Fold generated model of La Sota F in the monomeric pre-fusion conformation. Alignment performed in Pymol.

There are significant differences between the reconstruction generated by CryoSPARC and the Alpha-Fold construct, namely how the DIII region extends beyond the top of the CryoSPARC model. Also there is too much mass in the moon-shaped cleft between DIII, DI and DII.

UNcle:

We sought to increase the solubility of the protein, and thus needed to better understand the protein's melting temperature (TM) and temperature of aggregation (TAgg). Alternatively, the aggregation seen in the Chameleon nozzle it could have been due to the harsh nature of the aspiration in the nozzle.

The UNcle allows for the calculation of TM by measuring the intrinsic fluorescence of tryptophan, tyrosine, and phenylalanine residues, while the Static Light Scattering (SLS) reveals the TAgg (22). These points are measured at the start of the exponential phase of figures 54-59. While his was the method of calculating the TM and TAgg recommended by the member of the department who delivered instruction on the machine, in retrospect, a more mathematical method should have been used.

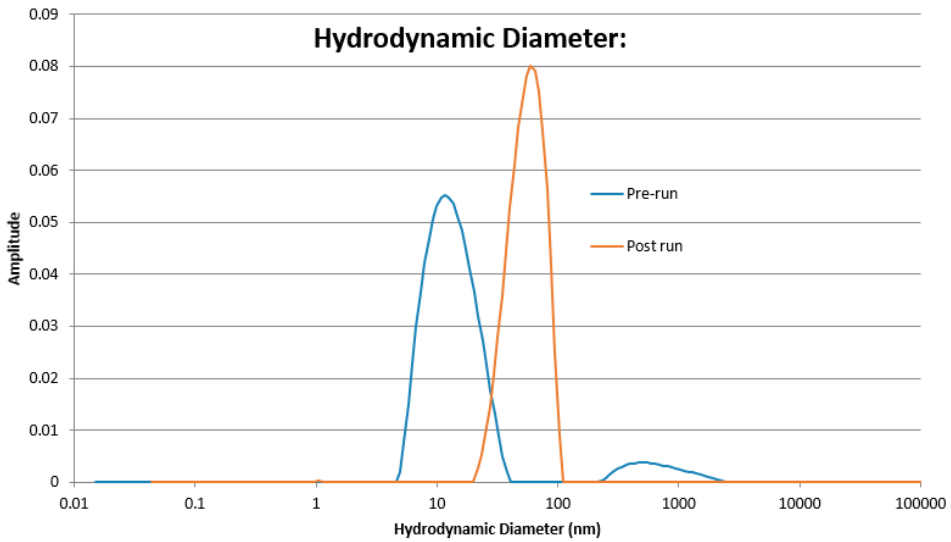


Figure 53: A chart showing the relative hydrodynamic diameter of the La Sota F protein before and after stepwise heating to 93 °C in the UNCLE.

Among the features of the UNCLE is the ability to do thermostability ramps to see how the protein changes in conformation using fluorescence and allows us to follow changes in aggregation using static

light scattering. The measure of SLS before and after the temperature ramp shows the average Hydrodynamic diameter of the protein increasing from 10nm to 90nm as the protein structure unfolds and possibly aggregates upon increase in temperature (see figure 53).

It further allowed the investigation of the effects of various putative stabilizing agents such as non-detergent sulfobetaine (NDSB) 195, 205 and 254, L-arginine, trehalose and L-glutamate. It was hypothesised that if these agents could reduce the disassembly of the trimer due to increased heat, they could do the same for plunge freezing.

NDSB195:

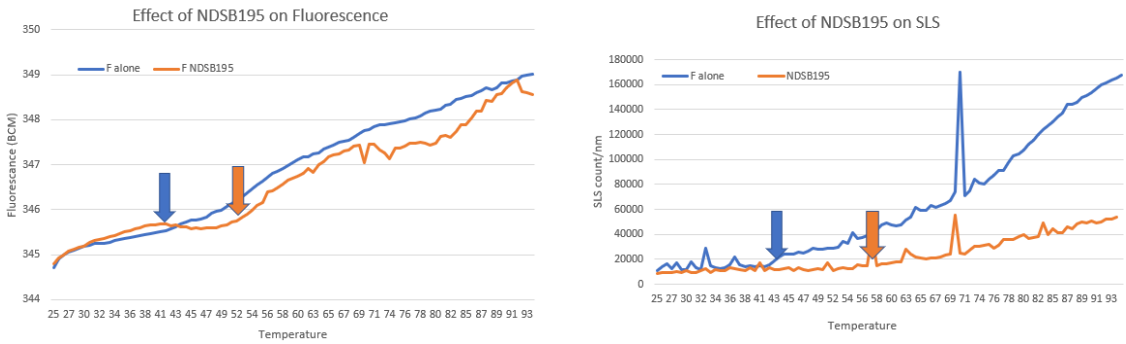


Figure 54: Charts showing the effect of NDSB195 on the TM and TAgg of NDV F protein at 2mg/ml. The plotted values show the mean of the experiment performed in triplicate.

The first additive investigated was 0.5mM NDSB195. Looking at figure 54, the change in fluorescence and the SLS at 266nm, it became very clear that the NDSB 195 had a stabilising effect on the protein and delayed the onset of aggregation. The two inflection points on the fluorescence curves, like the pH and additive screens shown in Chapter III, imply that as the protein is heated, the trimer first disassembles, then the monomers unfold. There is also the possibility that the F protein transitions from the pre-fusion to the post-fusion conformation prior to unfolding. NDSB195 increased the TM of the F protein from approximately 40°C to approximately 50°C.

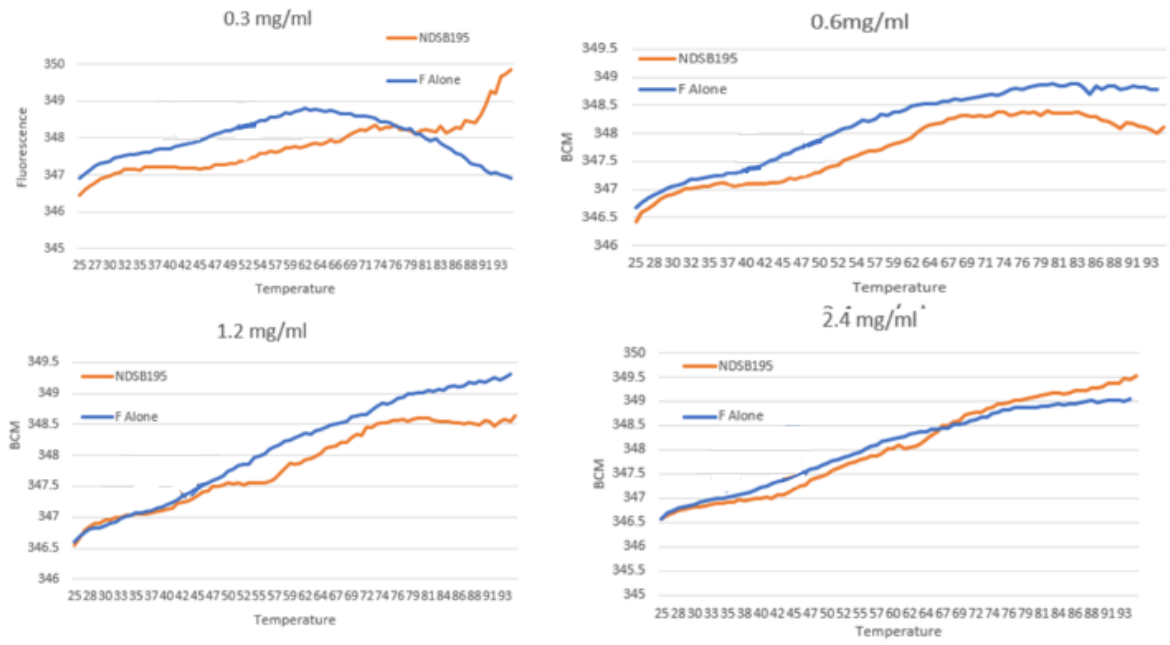


Figure 55: Charts showing the effect of 0.5mM NDSB195 on the TM of NDV F protein at different concentrations of F protein. The arrows indicate the start of the exponential phase, where protein begins to unfold. Fluorescence is measured in Barycentric mean (BCM).

It was necessary to see if the protection given by the NDSB was dependant on the concentration of F protein. Subsequent experimentation showed that the NDSB195 to be effective at stabilising the F protein at multiple concentrations (see figure 55), though it only provided so much protection as the F protein became more concentrated.

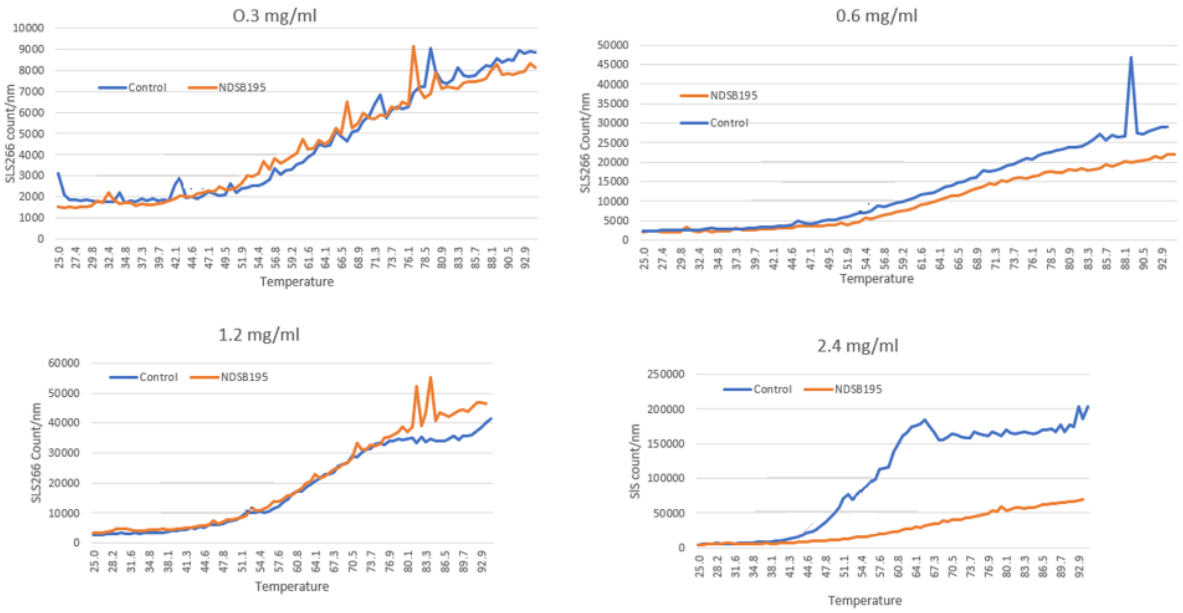


Figure 56: Charts showing the effect of NDSB195 on the TAgg of NDV F protein, at various concentrations. The arrows indicate the start of the exponential phase, where protein begins to aggregate.

Looking at the effect on the protein (figure 56), NDSB195 has a stabilising effect of approximately 5°C at the lower concentrations of F, with its effect becoming more pronounced at higher concentrations, 10 °C by 2.4 mg/ml.

Other additives:

The effects of other thermo-protective additives were examined, including L-Arganine, Trehalose and L-glutamate, see figure 57.

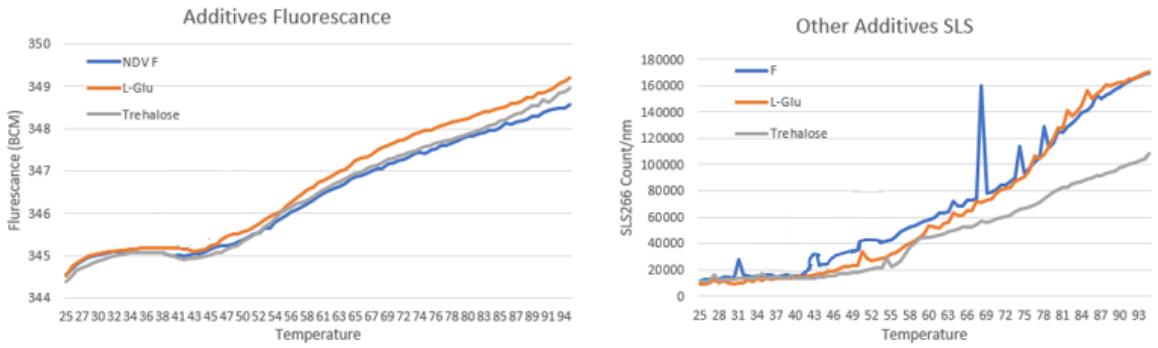


Figure 57: Charts showing the effect of 0.5mM L-Glutamate and Trehelose on the TM and TAgg of NDV F protein at 2mg/ml.

Of these L-glutamate offered a moderate increase of 5°C in the TM and Tagg, while Trehalose offered a more substantial increase of 10°C.

Pegylation:

Pegylation offered another possibility for protecting the F protein under the freezing conditions of cryo-EM, with 3ul of Methyl-PEG4-NHS per 250ul of F protein being used to cross-link F protein at a concentration of 1 mg/ml on ice for 60 minutes (see Chapter II).

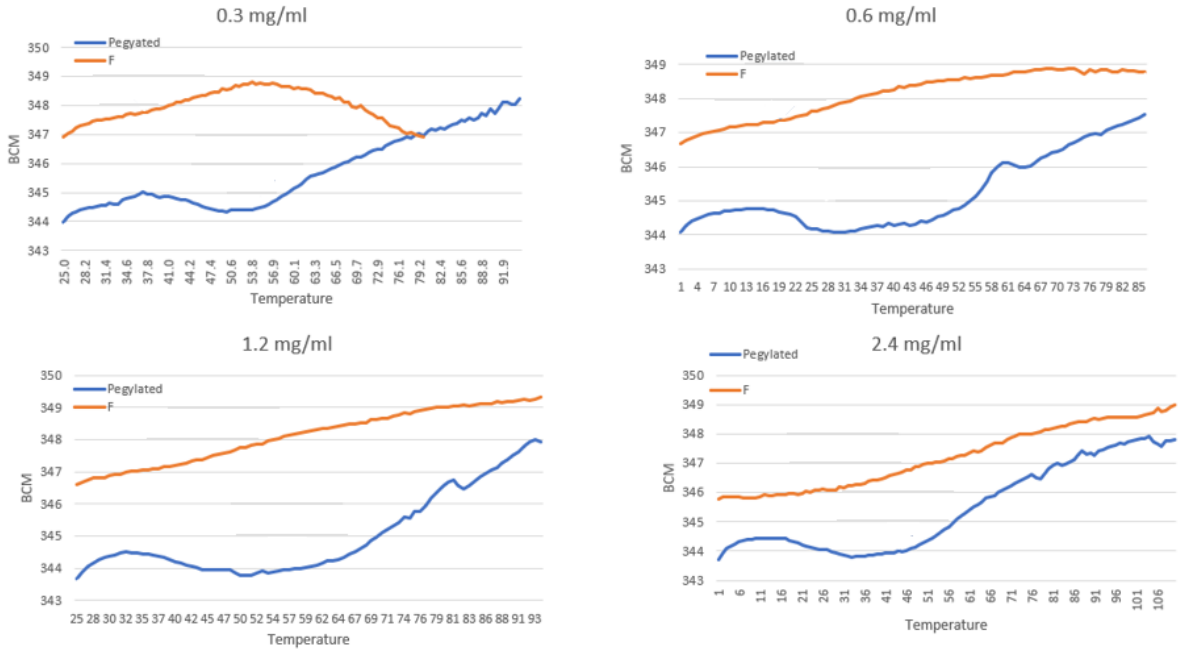


Figure 58: A series of charts showing the effect of pegylation on the Fluorescence (in BCM) and TM of F protein at various concentrations.

The above fluorescence plots (figure 58) show quite a substantial increase in the TM, by as much as 20 °C in some cases. The effect does appear partly dependant on the starting concentration of the F protein. However, since some of the charts do not follow the classical form, there are, in hindsight, confidence in their conclusions is limited.

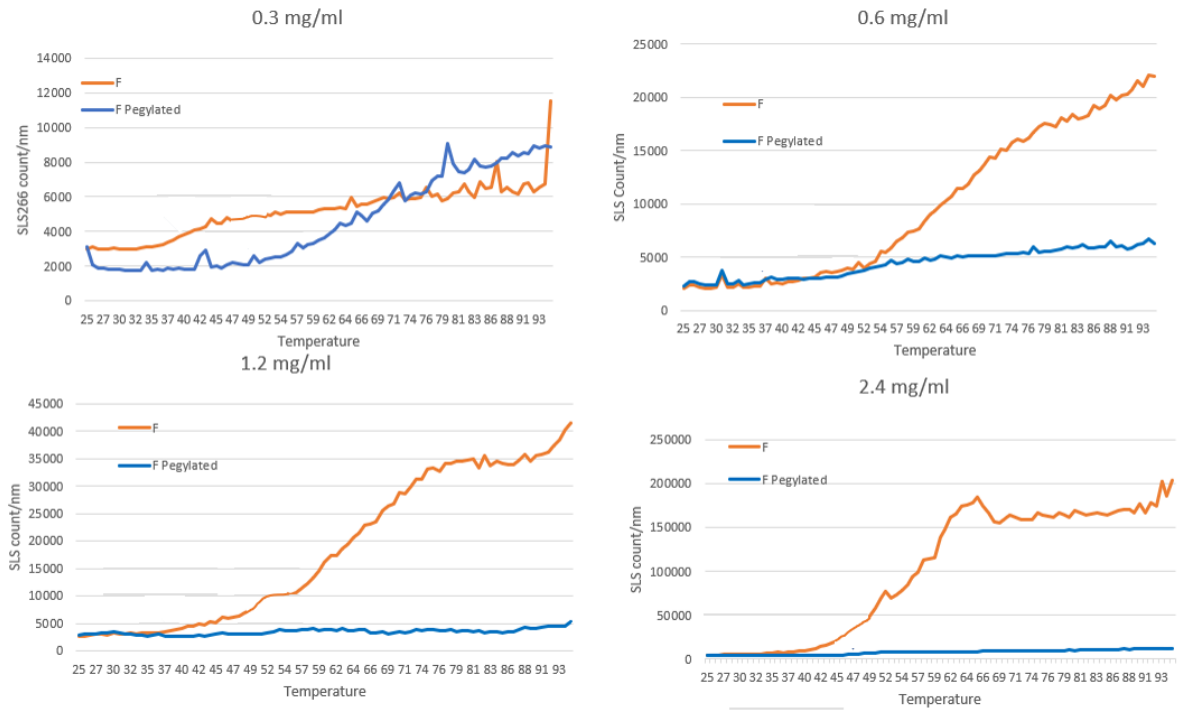


Figure 59: A series of charts showing the effect of pegylation on the SLS and thus the TAgg of F protein at various concentrations.

The pegylation of F appeared to increase the TAgg (Figure 59) by 10 °C through several concentrations of F. The protection against aggregation appears to be less dependent on the starting concentration. See Table 5 for a summary of the thermostability findings.

Summary of the effects of the additives:

Additive	Mean TM (°C)	Δ TM (°C)	P	TAgg (°C)	Δ TM (°C)	P
F alone	40	/	/	39	/	/
NDSB195	51	11	0.0004	52	13	0.0003
NDSB201	42	2	0.0569	42	3	0.1656
NDSB256	40	0	0.923	43	4	0.0246
L-Arg	43	3	0.0249	42	2	0.1717
L-Glu	44	4	0.0103	45	6	0.0057
Trehalose	46	6	0.025	52	13	0.0011

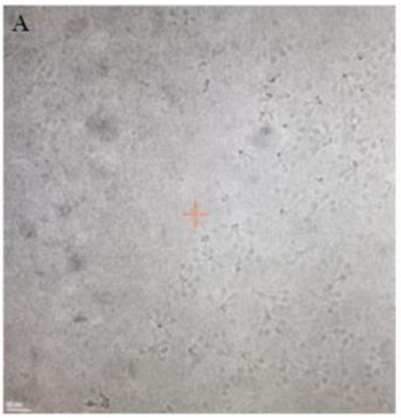
Pegylation	57	17	0.0001	51	12	0.0007

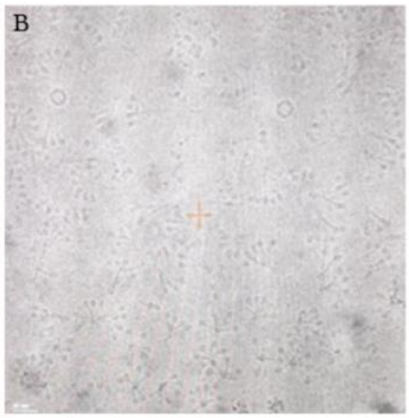
Table 5: A table summarising the effects of the various additives on the TM and TAgg of F protein.

Examining the P values: the difference in both TM and TAgg caused by pegylation or the addition of 50mM NDSB195 can be considered statistically significant, and could be useful for cryo-EM. Furthermore Trehalose was also protective and could also be used.

Pegylation and cryo-EM

Pegylation appeared to offer an effective means by which to reduce the tendency of the F protein to aggregate, which would be very useful for future work on the Chameleon. As a preliminary, it was decided to prepare and screen some grids, made using the departmental vitribots. See figure 60.





Protein Concentration (mg/ml)	0.3
Raw Pixel Size (A)	0.7
Accelerating Voltage (kV)	300
Spherical Aberration (mm)	2.7
Total Exposure Dose (e/A^2)	50
Grids	Quantifoil Gold
Mesh	300
Particles	3,190,420

Figure 60: Two micrographs taken of cryo-EM grids prepared with Pegylated La Sota NDV F protein and the collection parameters. Of special note are the micelle-like arrangements of particles. Collected on a Glacios 300kV microscope.

Examining these images from the microscope, it appears that there is a mixed population of what could be F in both pre and post fusion conformations. Furthermore, there appeared to be multi-unit structures forming.

Examining the 2D classifications shown in figure 61 the idea that there was a mixed population gained traction:

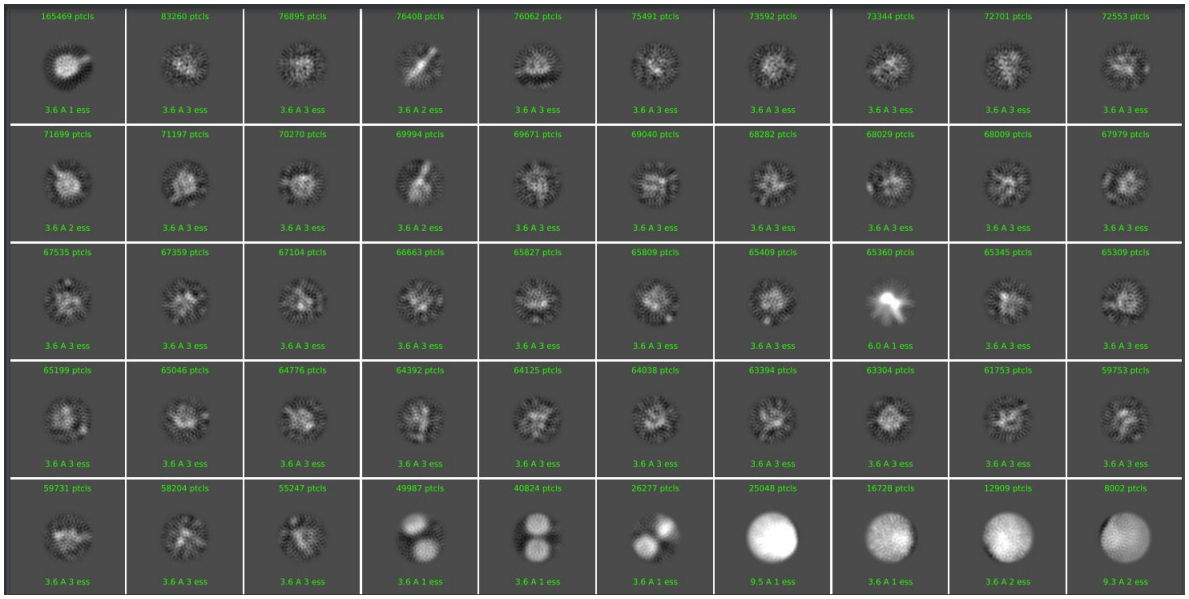


Figure 61: The 2D classes generated in CryoSPARC from the particles extracted from the micrographs of Pegylated La Sota F.

The 2D classes show some very globular proteins with stalks, suggestive of a prefusion conformation of F. Also visible are some more triangular proteins that could be trimeric F in a top-down view. These were then used to generate 3D models in CryoSPARC, see figure 62.

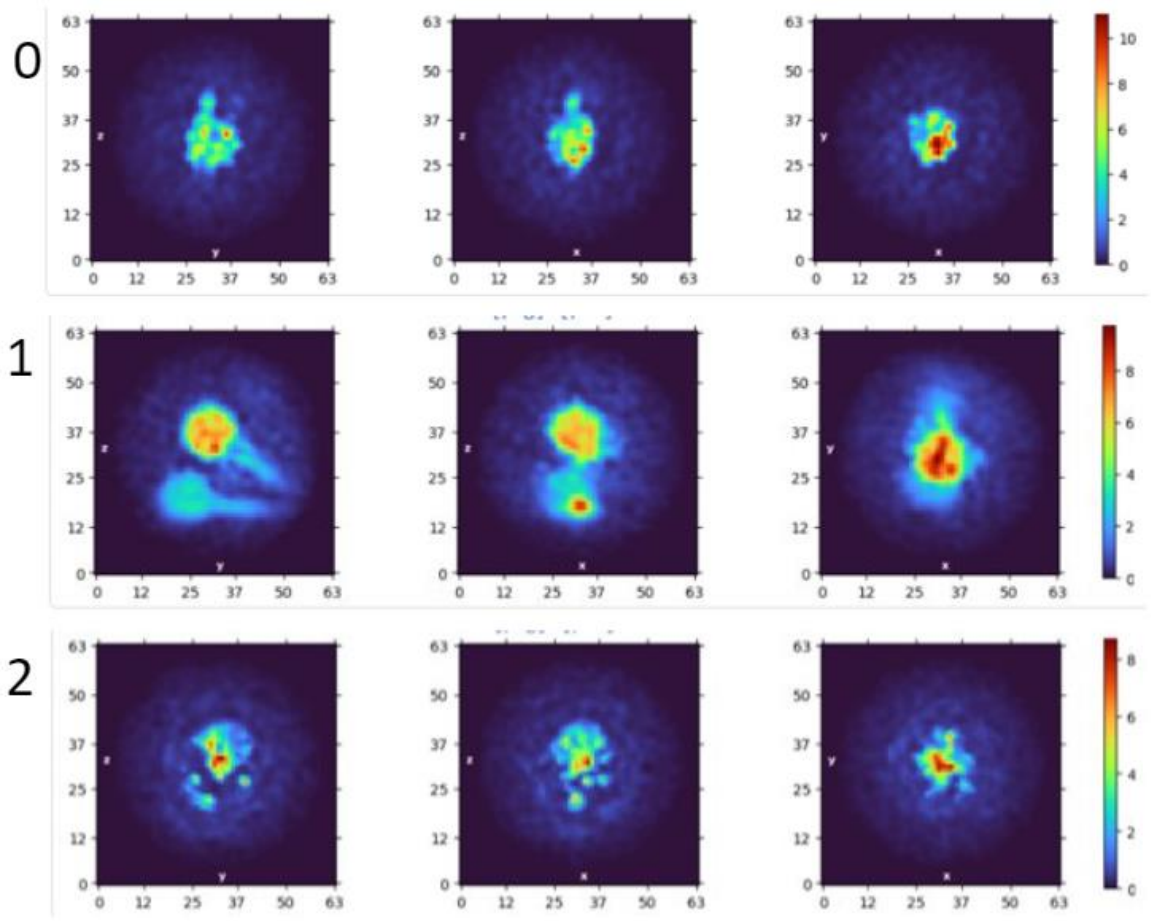


Figure 62: The 3 classes of model generated by CryoSPARC from the selected 2D classes, of Pegylated F protein, seen in figure 61. Class 1 are of particular note.

Once classes had been picked and 3D models generated there was cause for optimism: Class 0 was globular with a tail, while class 1 was very globular with a pronounced tail region. Additionally, there appears to be multiple units in class 1, which looks credible given the multi-unit association in the micrographs. The extended tail and small head in the class 1mask also closely resemble the post-fusion structure of NDV.

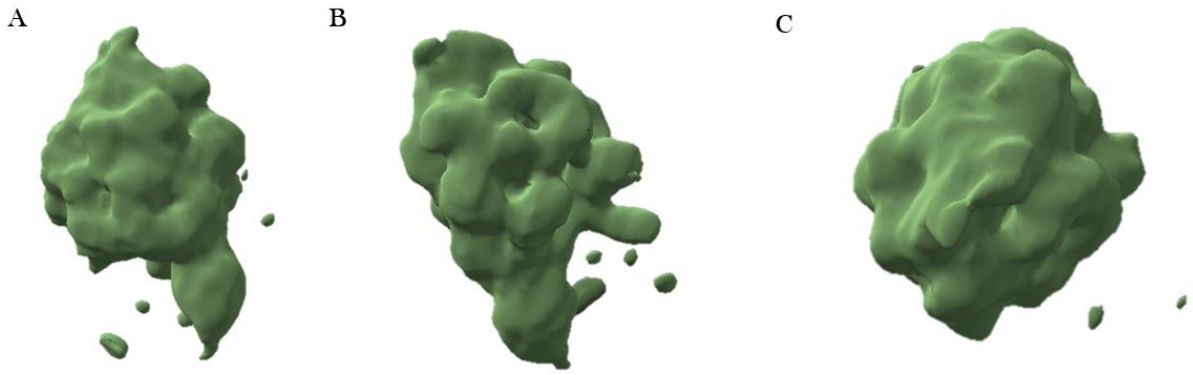


Figure 63: the volume mask generated in CryoSPARC from class 0 seen in figure 62.

The class 0 model, shown in figure 63 is indeed globular, though the stalk is not central, suggesting it is either not trimeric, or an artefact of the 3D rendering. It could be another monomer, but again has too much density in the centre.

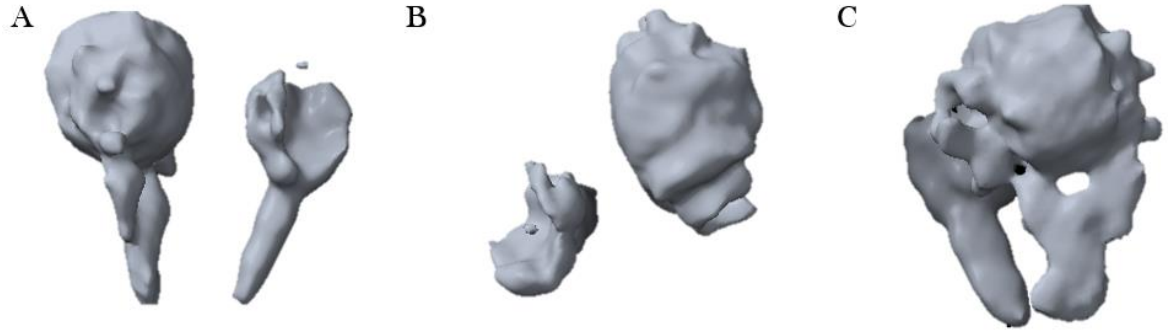


Figure 64: the volume mask generated from class 1 seen in figure 62.

Class 1 (figure 64), however showed a particle with a globular head and a longer, slender tail, in proximity to what could be a second particle. When the class 1 model is compared to the structure of monomeric post- fusion NDV F (3MAW) figure 65 is the result:

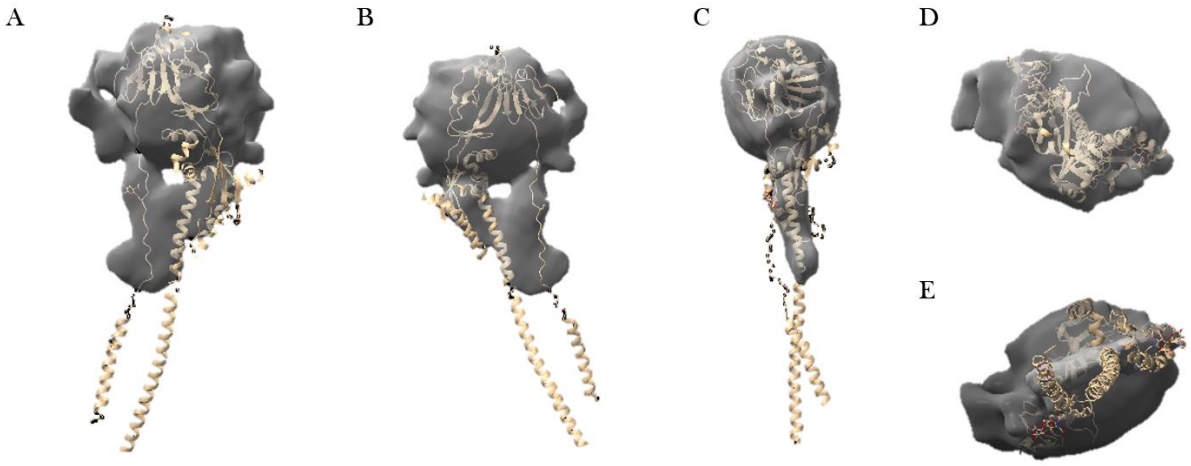


Figure 65: the volume mask generated from class 1, paired down to only one unit, fitted to the monomeric NDV F in the post fusion conformation.

The comparison is quite striking as the globular head fits with what would be the head of the post-F protein, as does the denser region of the tail. However a post-fusion monomer has not been reported. Incompleteness of the model due to low resolution and heterogenous density are the likely determinates of this mask.

Multi-unit classes:

The micrographs from the first pegylated data-set included a number of multi-unit structures, which could be used to solve a multi-unit structure of the protein. Figure 66 shows the 2D classes generated from manually picking multi-unit particles from the micrographs (figure 60).

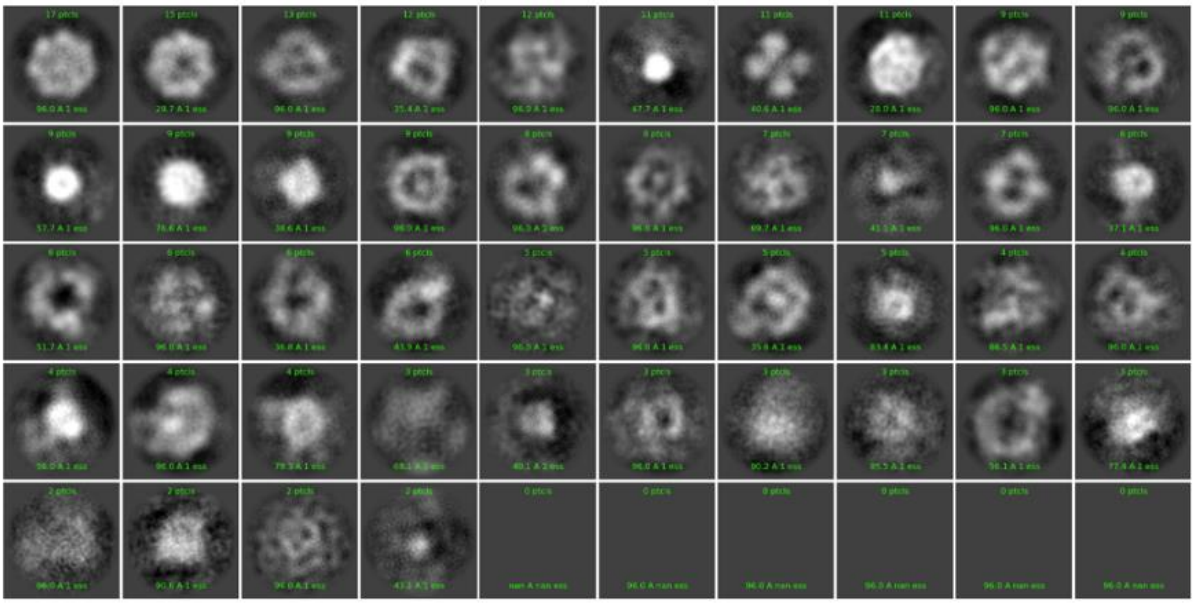


Figure 66: The 2D classes generated in CryoSPARC from the pegylated F data-set by selecting multi-unit structures from micrographs of the pegylated F data-set, see figure 60 for the collection details.

The 2D classes are a very heterogenous collection and would be difficult to use to generate a 3D model from. The number of particles in this format numbered in the low hundreds, and was too few to draw conclusions from.

Further processing:

While the subsequent experiments were being carried out, it was decided to go back and have another attempt at reprocessing the data collected with Pegylated F protein (see figure 60 for the parameters). Upon further processing, it became apparent that the data-set contained some possibilities for generating a structure. Through a more selective choice of particle picking parameters, it was possible to generate more 2D classes, see figure 67.

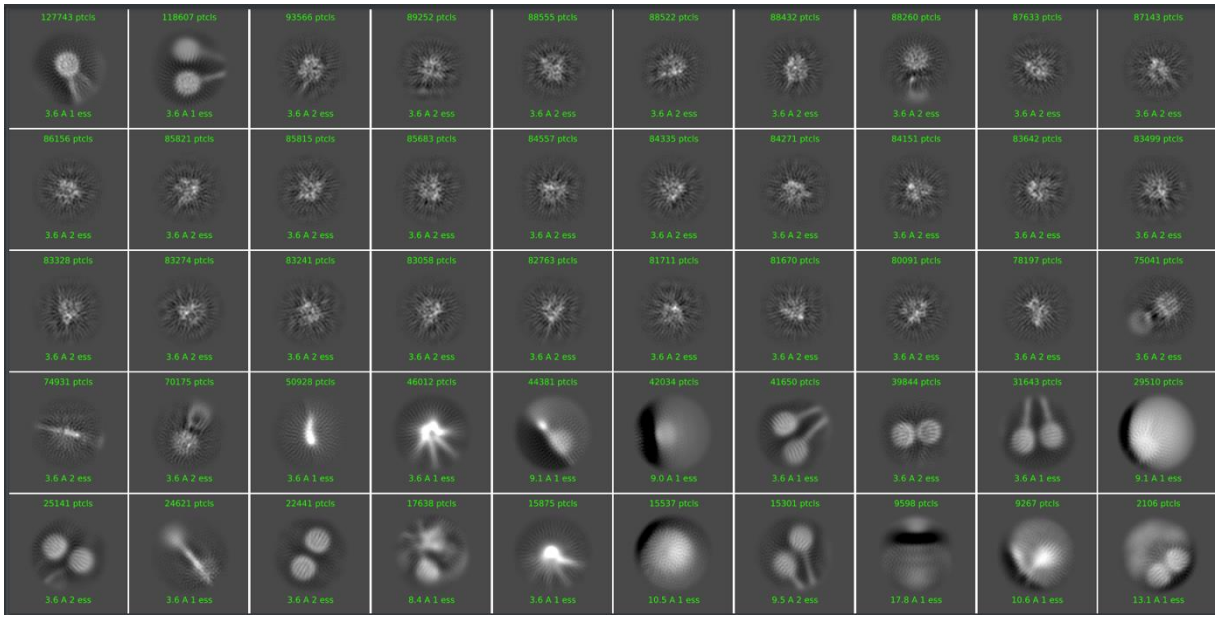


Figure 67: More 2D classes from the pegylated F dataset (parameters in figure 60). The particles selected from the micrographs were chosen more stringently and used to find similar particles using the function of CryoSPARC where the 'particle picker' function can be trained. These particles were then sorted into these classes.

Of the picks, several appeared triangular. These and the globular particles were chosen and used to generate the 3D models. Of these, one seemed to have promise, so was further refined, see figure 68.

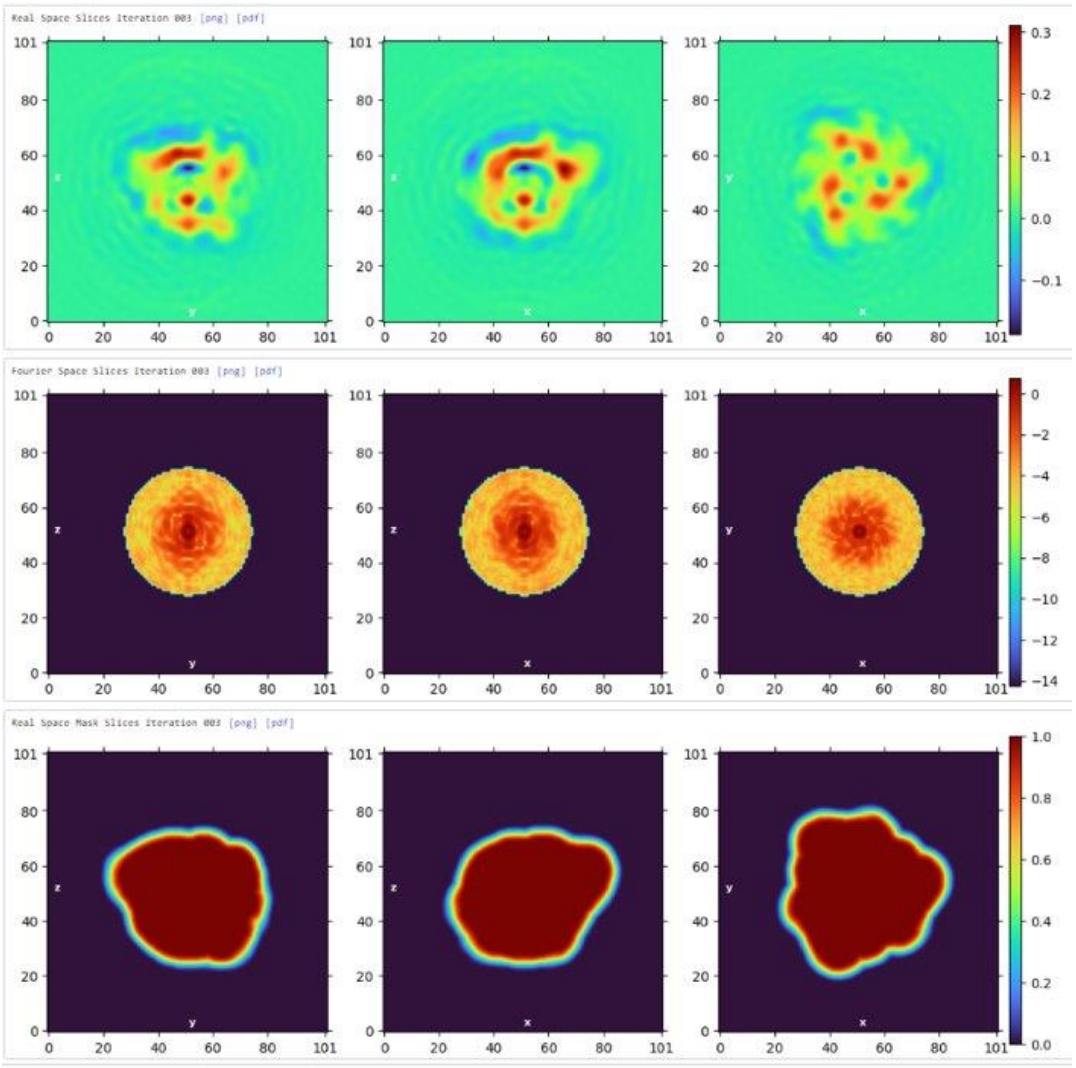


Figure 68: A model generated in CryoSPARC ab-initio from the 2D classes shown in figure 67.

At first glance the structure looks like a trimer from the top down, Especially in the top right frame of the panel that looks like the interlocking arms of the trimeric F protein in the pre-fusion conformation. However the sides appear to lack the length of the pre-fusion structure. This worry was confirmed in figure 69 when the model compared to the known post-fusion NDV F protein structure (Figure 69A and 69B) and the Pre-Fusion F from PIV5 (figure 69 C and 69D).

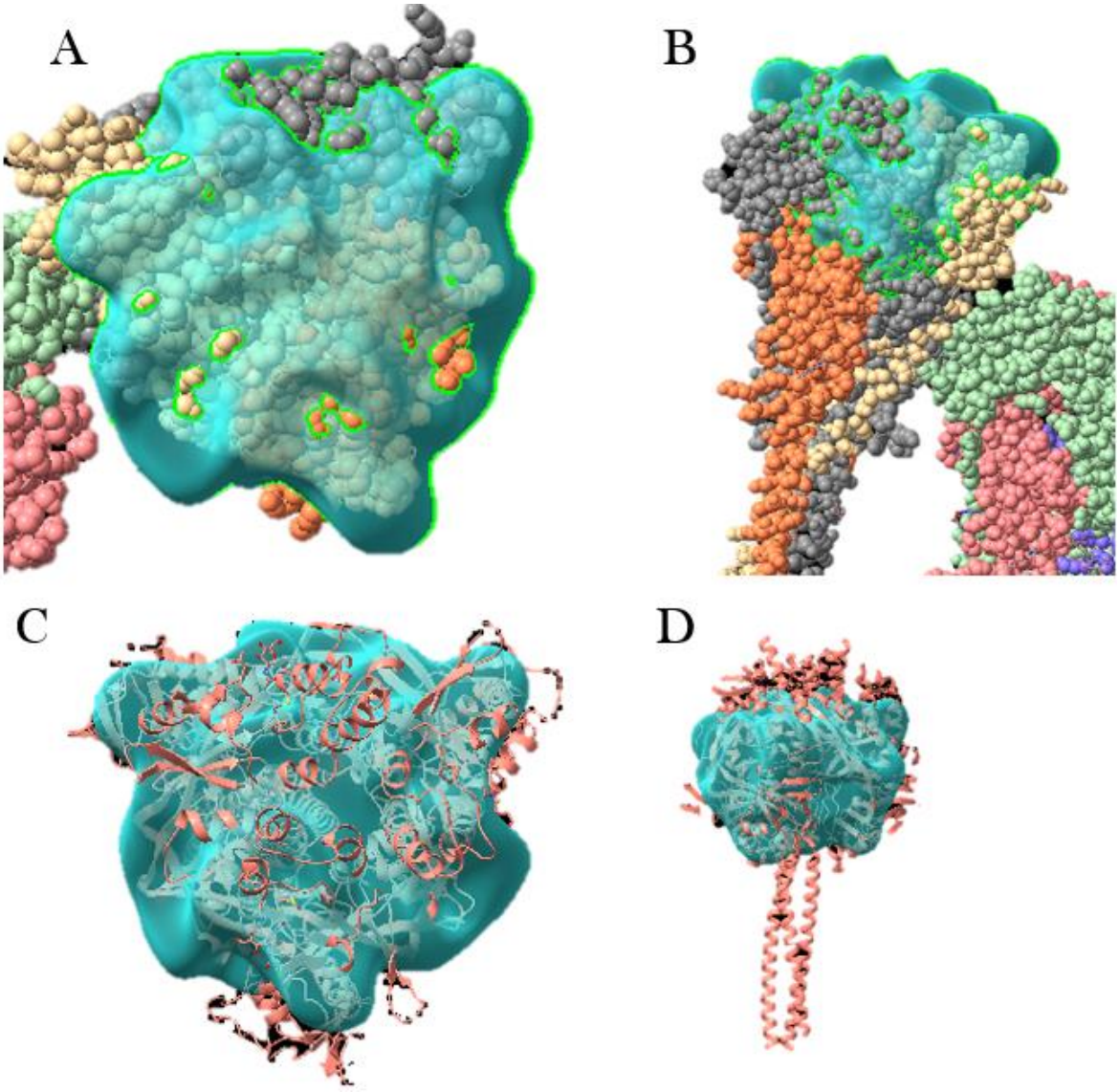


Figure 69A and 69B: The mask generated from the ab-initio model (figure 58) fitted against the NDV F protein in the post fusion conformation (PDB 3MAW). 69C and 69D: The same model fitted against the trimeric PIVF in the pre-fusion conformation (PDB 4WSG).

The fit with the pre-Fusion F could be viable, from the top-down but there are a couple of regions that don't fit quite as well. Side on it becomes more apparent that the fit is from one plane only. When fitted against the pre-fusion F, the structure appears to fit well in the top-down view. Side on, the fit is not convincing, but not as divergent as with post-fusion F.

This dataset had yielded as much as it was going to, so it was necessary to optimise the preparation of grids prepared with pegylated F protein and to see if the effects of the pegylation were reproducible. There followed some more grids prepared with Pegylated F protein. A data-set was collected on the

Glacious with freshly pegylated protein, see figure 70 for the parameters. A lower accelerating voltage (kV) was used for the following data collections to improve the quality of the image by reducing electron penetration and thus increasing contrast.

Protein Concentration (mg/ml)	0.25
Raw Pixel Size (A)	1.2
Accelerating Voltage (kV)	120
Spherical Aberration (mm)	50
Total Exposure Dose (e/A^2)	60.17
Grids	Quantifoil Gold
Mesh	300
Particles	2,123,606

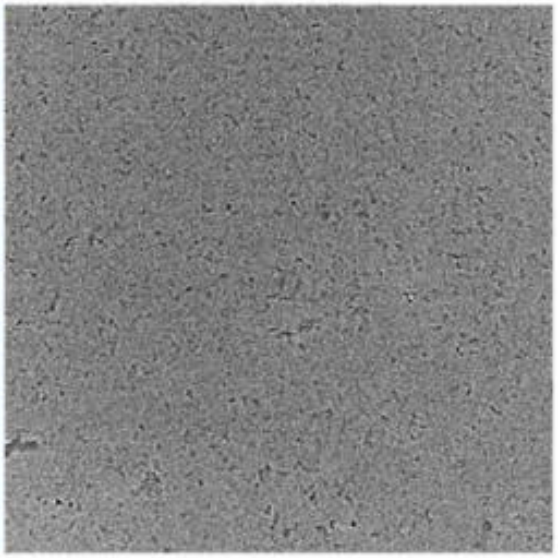


Figure 70: A cryo-EM micrograph and collection parameters of a further data set collected with pegylated La Sota NDV F using the Glacious 300kV microscope.

This preparation of sample lacked the micelle-like structures of the original pegylated data-set (figure 60), though an apparent mix of conformations were also present. When the particles were extracted from the micrographs and sorted and sorted into 2D classes (figure 71), this idea was supported.

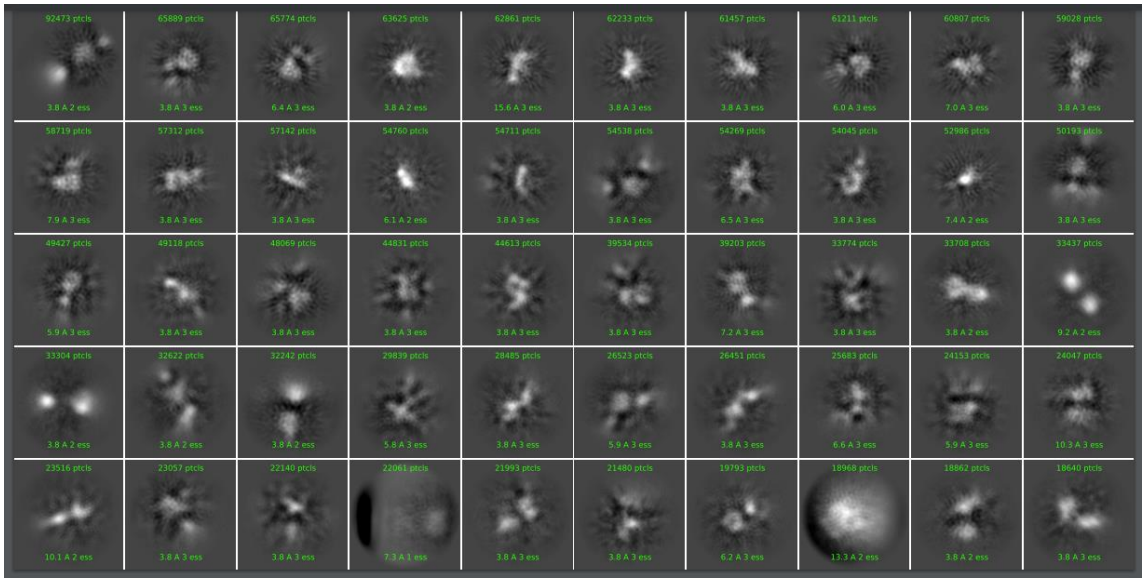


Figure 71: The 2D classes that particles in the micrographs were sorted into, using CryoSPARC (from the second dataset collected with pegylated F protein).

The 2D classes were encouraging as a globular form is visible, with a distinct stalk forming part of the arrangement. However there were quite a few classes that did not resemble the desired conformation. The results of the ab-initio modelling, seen in figure 72, were more questionable.

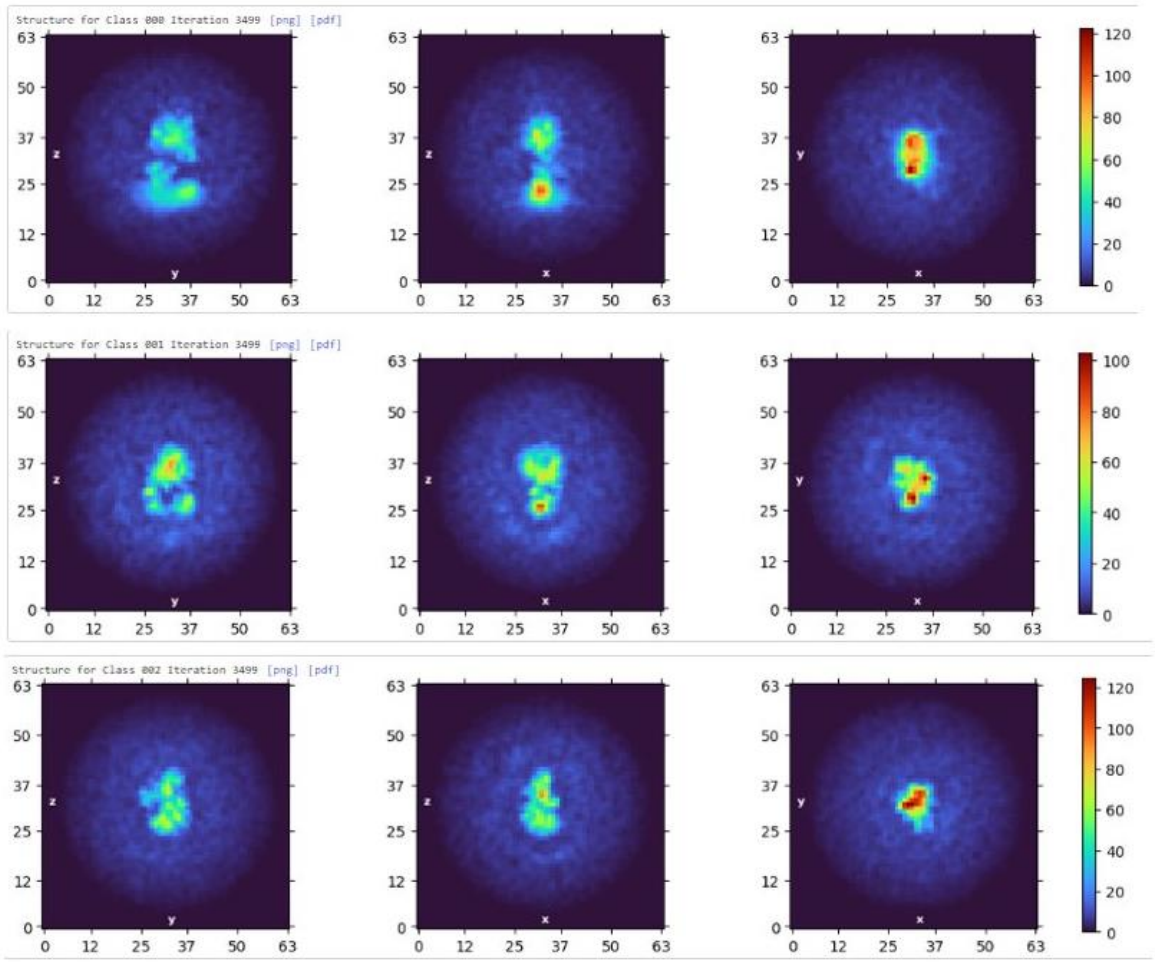


Figure 72: The three classes generated by CryoSPARC's ab-initio modelling from the selected 2D classes in figure 71

The three models generated all suggested that the protein was also monomeric. However, examining figures 73 and 74, the shape of the protein on the grid was somewhat different to what would be expected by both AlphaFold2 and the structures of PIV5.

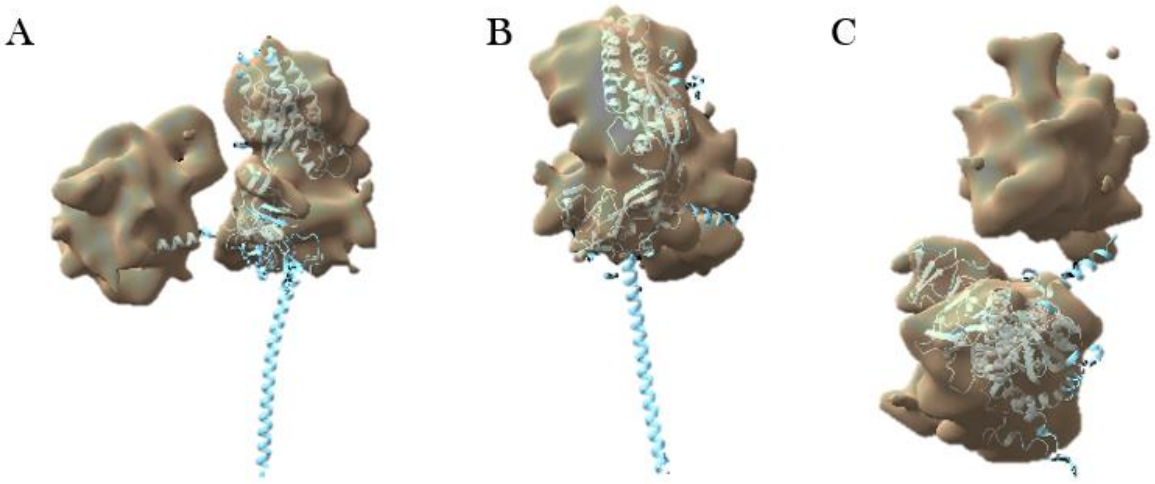


Figure 73: The class 0 mask from figure 72, fitted (Using PyMol) against the model of monomeric NDV F in the pre-fusion conformation (generated by Alpha-Fold).

The class 0 structure in figure 73 appears to include two particles in close proximity and possesses a greater level of density than could be expected of the monomer, though this could be an artefact of both the sugars on the molecule, or the pegylation.

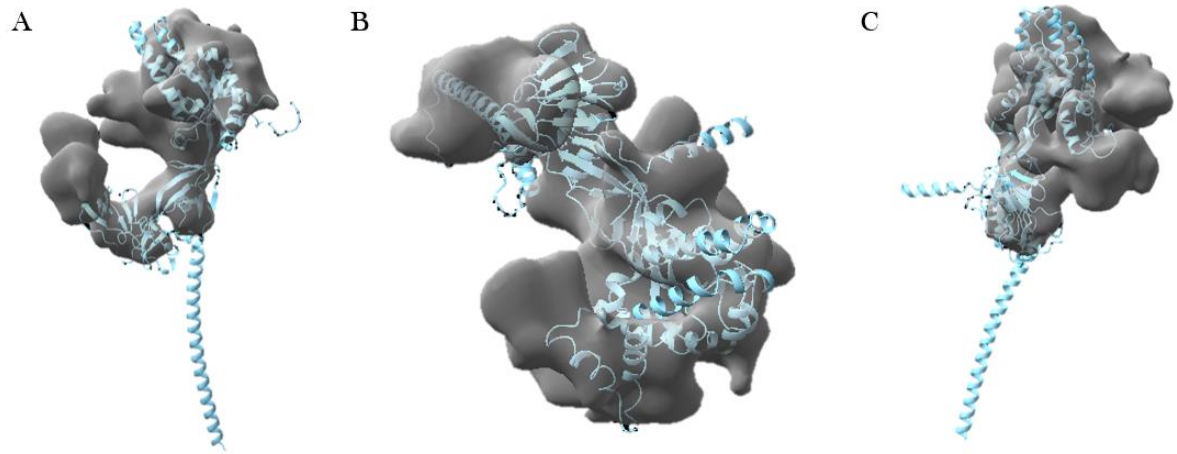
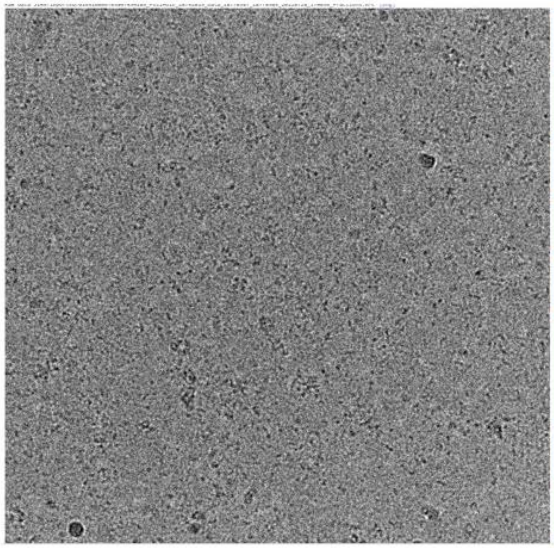


Figure 74: The class 1 mask from figure 72, fitted against the model of monomeric NDV F in the pre-fusion conformation (generated by Alpha-Fold).

The class 1 product in figure 74 does look more like the c shape of the modelled monomeric pre-fusion F protein, though this is not a perfect fit and there are some regions that lie outside the ribbon structure.

A final attempt:

A final cryo-EM dataset with pegylated F protein was collected, using grids that had been made using back-blotting only, see figure 75 for the parameters. The use of back-blotting would also remove the chance that the traditional blotting was causing the detrimerisation of the F protein.



Protein Concentration (mg/ml)	0.3
Raw Pixel Size (A)	1.2
Accelerating Voltage (kV)	120
Spherical Aberration (mm)	50
Total Exposure Dose (e/A^2)	60
Grids	Quantifoil Gold
Mesh	200
Particles	430925

Figure 75: A cryo-EM micrograph and collection parameters of a further data set collected with pegylated La Sota NDV F that had been back-blotted instead of dual-sided blotting. The Glacios 300kV microscope was used at a lower power setting of 120 kV.

The back-blotting appeared to leave more protein on the grid than the traditional front and back blotting. A good level of what appears to be trimeric pre-Fusion protein was visible, as was protein that appeared to be in the post fusion conformation. The 2D classes for this dataset, seen in figure 76, were also sub-optimal.

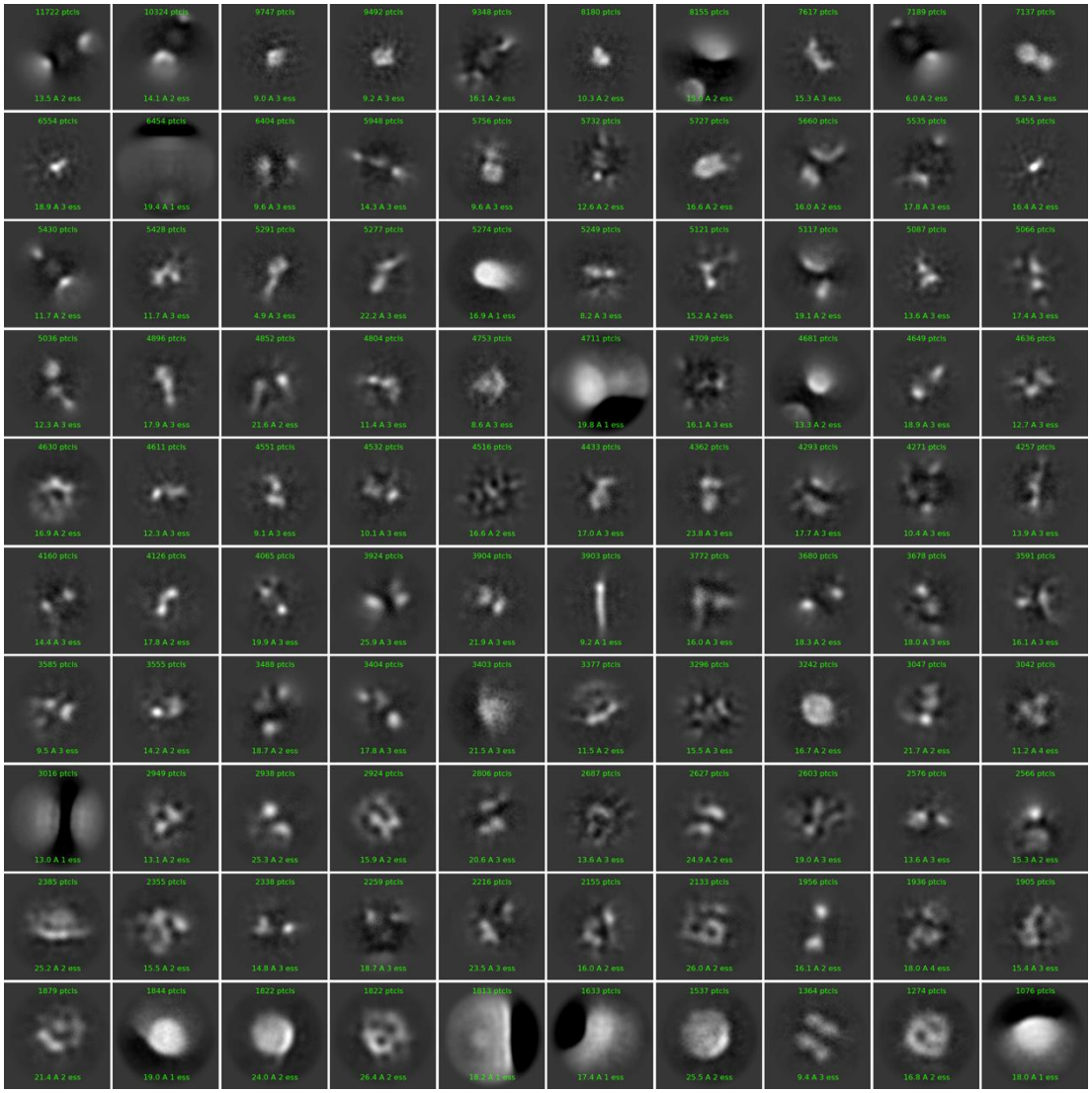


Figure 76: The 2D classes generated in CryoSPARC from the particles in the back-blot only data-set. The number of classes that the particles were to be sorted into was set at 100, to be more specific in the collection of the data.

The 2D classes did not give much cause for hope with this dataset but did identify that there could be some more multi-unit structures, like those seen in figures 60 and 66. The lollipop-like shapes seen in the 2D classes carry over to the 3D models generated.

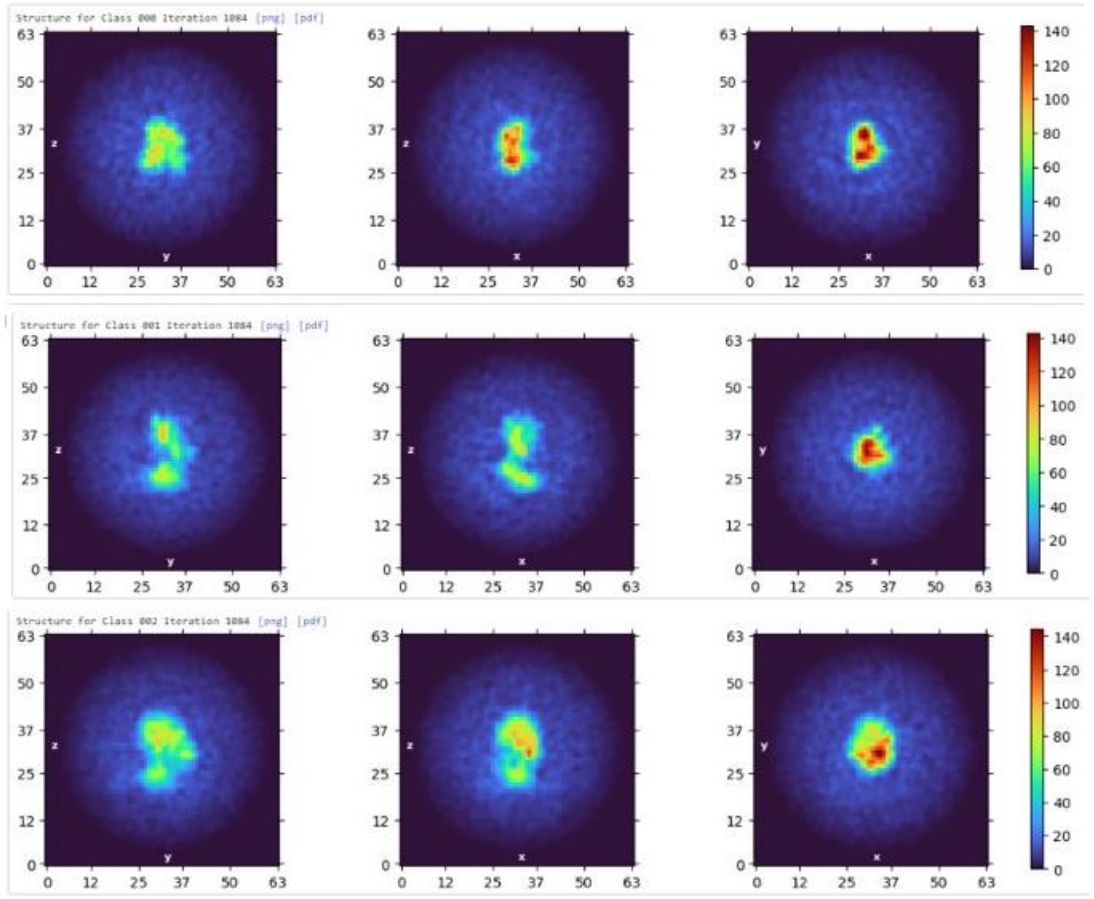


Figure 77: The three classes generated by ab-initio modelling from the selected 2D classes seen in figure 76, in CryoSPARC.

These 3D models in figure 77 all resemble the post-fusion monomers seen in the first pegylation experiment. It also appears that back-blotting does leave more protein on the grid. As both back-blotted and traditionally blotted F protein tends to result in the same data quality, the mechanical force of the blotting is probably not relevant.

Discussion:

The results of the negative stain EM imply that trimeric F protein is being expressed and purified, and appears to be in the pre-fusion conformation at pH8. These resemble the micrographs produced in previous work on class I fusion proteins, reproduced in figure 78 (11-13). For example: these electron micrographs taken from Lamb and Jardetsky's review showing GCNt stabilised PIV5 F resemble the GCNt stabilised NDV F. The more extended conformation seen in the pH5 negative stain EM of NDV

does not to e in the 'Golf tee' conformation of PIV5 (24). However, it is worth noting that while PIV5 is not pH dependant, NDV F is (25).

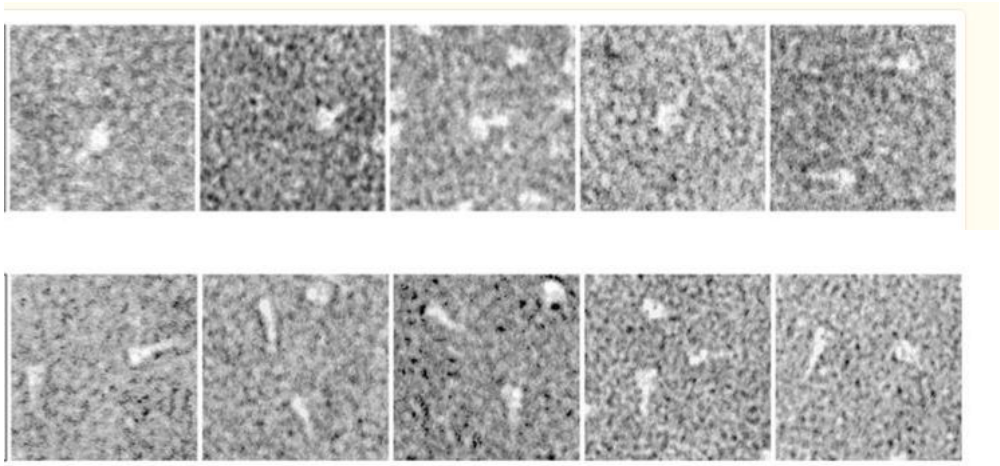


Figure 78: Top row The Micrographs of GCNt stabilised PIV5 F. Bottom row: PIV5 F after heating to 50°c for 30 minutes. reproduced from (24).

The inability to derive a credible reconstruction during the initial runs of cryo-EM was somewhat disappointing, but also not something reported in the literature of many of the class I fusion proteins. Seow et al reported a cryo-EM structure of a patient-derived antibody fragment, bound to SARS-CoV-2 Spike protein. This was bound to an epitope not accessible unless Spike was monomeric, suggesting that dissociation is common (26).

It also prompted a closer look at the purification process. It was suggested that the Superose-6 column lacked the necessary precision to separate proteins of 50kDa and 150kDa, I.e. monomeric and trimeric F protein, both appeared in the same peak. The switch to a S200 column revealed that the amount of trimeric F protein in solution was exceedingly low, possibly as the HEK293 cells were incubated for four days at 37°C, degrading the protein. The shift to Baculovirus expressed F protein was not only more time-efficient but the increased proportion of trimeric F (at the time of purification) was a distinct advantage. Possibly this increase was due to a combination SF9 cells being incubated at 26°C, and the fact that the incubation was only for 2 days, rather than 4. See figure 46.

The failure of crosslinking to prevent the trimer disassembling under cryo-cooling was also disappointing. However, this coincided with the discovery of how SF9 grown protein was more stable

than HEK-expressed protein and the use of pegylation in grid preparation. With time and resources becoming increasingly tight, it was decided to put the cross-linking to one side as the initial pegylation discovery appeared to solve the potential problem of de-trimerisation.

The Chameleon experiment revealed that the sheer forces causing the trimer to disassemble remained, but to a lesser degree with much more rapid freezing. Initially, with the live-processing, it appeared that Chameleon prepared grids had a large proportion of putatively trimeric protein, but with subsequent processing the 2D classification revealed a large proportion of monomeric F. The protein's tendency to aggregate demonstrated the need for more biophysical characterisation of the protein.

UNCLE offered the opportunity to carry out said biophysical characterisation of the F protein and several agents were identified that could reduce the aggregation of the F protein and also help protect it in the stress of grid preparation.

Pegylation had been useful in the work of a college on the Class I fusion protein SARS_CoV-2 Spike, and opened a different approach for this project (27). It held promise in this aspect of this NDV project, with both pre and post-F conformations being putatively visible on the micrographs made with pegylated F.

Like with the chameleon, initial processing and 2D class selection suggested that the particles could be trimeric F in the pre-fusion conformation. However, further processing struggled to build a higher resolution structure.

The formation of multi-unit structures in both negative stain and cryo-EM conditions resembles very closely the multi-unit structures formed by both NiV and CDV (14, 15). These offer an insight into how the structure of the F protein could be solved. Such structures could conceivably be formed, especially in the post-fusion conformation as the hydrophobic areas of the DII region usually sit towards the apex of the post fusion structure. That would explain the formation of the structures with an almost lipid micelle look. 1G5G, the structure proposed by Chen et al (reproduced in figure 79) also contains two post-Fusion F proteins in close proximity (27):

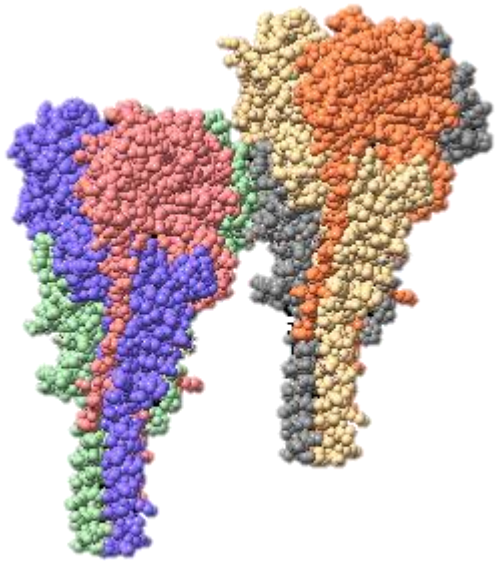


Figure 79: PDB item 1G5G, the structure of NDV F in the post fusion conformation determined by Chen et al (19)

An alternative explanation is again offered by Lamb and Jardetsky, who took GCNt stabillised PIV5 F protein, performed a trypsin digest and incubated it at 50°C for 30 minutes. Perhaps the energetics of the pegylation, or some part of the post pegylation purification caused something similar to the NDV F protein (24). See figure 80. The similarities to the star-burst pattens seen in figure 60 are quite striking.

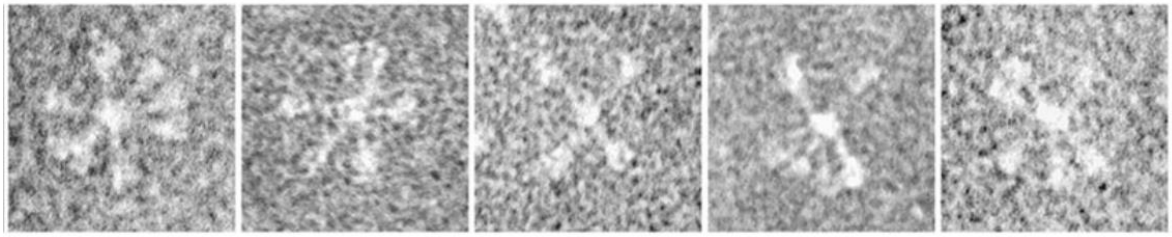


Figure 80: PIV5 F-GCNt digested with trypsin and then heated to 50°C for 30 min, reproduced from (24).

Subsequent re-processing of the initial round of pegylated-F data suggested that the protein on the grids could be preferentially oriented. While this was not reported with the other Class I fusion proteins discussed above, it was reported that the 150kDa trimer 'HA' from influenza orientates also in the 'top-down' plane (28). One possibility going forward would be to collect a tilted dataset, which would

overcome the effects of preferential orientations and aid in the successful structure determination of NDV F protein.

The difference visible in the micrographs between the first data-set with pegylated protein and subsequent ones could well be due to the first data-set being made using pegylated sample stored overnight before grids were made, due to the first batch of grids being too icy to use and having to be re-made. Subsequent grids of pegylated F were set-up immediately after size exclusion chromatography. Another possible explanation for the differences could be that the PEGylating agent became less effective due to time or storage issues.

A difference with the pegylated protein and the un-pegylated protein is that many of the monomers formed are post-fusion in appearance, whereas the non-pegylated datasets all appear to be a slightly mis-shaped pre-fusion conformation. PEG has been recorded as being a mediator of process by-which viruses fuse with the membranes of their hosts (30). Perhaps the pegylation is triggering a conformational change in the sample. This could either be through the presence of the PEG as a trigger, or by the energetic changes within the solution. Such a conformational change could also explain why the protein appears to be more stable when pegylated, it is in a more energetically favourable conformation. This would be problematic for further work carried out on the F protein, but if a suitably stabilised F protein could be generated, then pegylation may be worth revisiting.

One of the most effective ways to stabilise a protein structure is to induce some stabilising mutations within the protein (12, 13). Several fusion proteins have already been subject to this and have notedly been more temperature resistant than un-stabilised protein. Notably PIV5 has been quite well stabilised (12). of these, some could well be translated over to NDV.

However this would take considerable time to design and test all potential stabilising mutations before any grids could be made. The stabilising mutations of PIV5 could be marked on a structure, and be compared with the AlphaFold model of NDV. With the mutations mapped across locally, these could be introduced to the NDV constructs using site-directed mutagenesis.

Work did begin on this endeavour, but was halted to focus on the pegylation, which was judged to be the most promising lead, at the time. Another consideration was that any mutated variants would have to be properly cloned, expressed and characterised, to see which were the most effective. These stabilising mutations could be a useful counter to the issues of the PEG triggering, or whose addition triggers, the change of conformation. Given the lack of time and the difficulties of accessing the labs due to the Pandemic this approach was not pursued.

All of this EM-based work shows that this would be the method by which the structure of the F protein is solved. The fact that it has yet to yield a structure is due to a factor of time constraints and the inherent structural instabilities of this meta-stable fusion protein. Further design work on the construct, introducing stabilising mutations or replacing the GCNt domain, may also prove necessary for the protein to survive pegylation and the freezing process and yield grids where the NDV F protein can be imaged to high resolution.

Chapter V:

HVT Vectors and Nanobodies: or things go even more pear-shaped...

Modern NDV Vaccines have started to move away from the use of live lentogenic or attenuated virus due to the risk of the virus reverting to virulence (1, 2). The former would still cause the milder symptoms of the disease, most crucially a reduction in eggs and stunted growth, making it an unattractive prospect (3). An additional problem faced is the cold-chain transportation and storage, which may not be available where the vaccine is most needed.

Table 1
Notable vaccine types used in Newcastle disease control.

Vaccine type	Composition	Advantages	Limitations	References
Live	Attenuated organism to make it less virulent/harmless	Mucosal and humoral immunity	Proper dose administration Cold chain maintenance	(Vrdoljak, 2018)
Inactivated	Whole killed organism (through physical and chemical means)	High humoral immunity	Individual administration Require withdrawal period before use as a food Negligible cell mediated response Virus shedding	(Schijns et al., 2014)
Recombinant herpes virus vector vaccine (rHVT-ND)	Made by insertion of F gene coding protein into thymidine kinase site	i)No clinical disease and mortality ii) Long-term immunity iii) In *ovo* and hatchery administration	Maternal antibody interference Cold chain maintenance Strict biosecurity measures needed	(Armour and García, 2014)
Recombinant NDV vector vaccines	Foreign gene insertion into the intergenic space	i)Acts as a bivalent vaccine ii)More hatchability and survival	Evaluation of virus efficacy (vector and insert) with maternal antibodies is difficult	• Monitoring (Ge et al., 2014)
Antigenically matched vaccines	Modified to be matched with the challenged virus (due to excessive genomic mutations)	Reduced viral shedding Higher production of antibodies	Require higher biosecurity	(Cornax et al., 2013)
Nanoparticle vaccine	Nanoparticles acting as a carrier	i)High antigen uptake ii)Controlled release iii)Longer duration of immune response	Regulation of antigen concentration Number and dose of vaccination	(Gourapura and Renu, 2020)
Virus-like particles (VLPs)	Virus protein without genome	i)Least infectious, ii)Least chances of reversal or recombination iii)Highly immunogenic iv)Stimulate humoral and cellular immunity	Mutations may render vaccine ineffective Require higher scale-up and purification VLPs delivery system with lesser cost and time	(Dimitrov et al., 2017)

Figure 78: A summary of the strengths and weaknesses of the major vaccine types. Reproduced from Naz et al. (3).

Inactivated vaccines are an alternative, providing good immunity, all-be-it primarily humoral, with more limited cellular based immunity (4-5). Unfortunately they do not stop birds shedding the virus if

subsequently infected with the disease (6). Furthermore, food safety regulations mandate a withdrawal period before the hen can enter the human food chain (7).

Chickens require a suitably trained person to administer the vaccine at 9 days of age, then there is a minimum withdrawal time of 21 to 48 days post-vaccination (depending on vaccine and local laws) (8). Since Broilers are commonly slaughtered from 28 to 56 days (9), the essential mathematics is that inactivated vaccines can delay the point at which birds are slaughtered.

Crucially, while inactivated NDV vaccines showed efficacy in SPF flocks, when deployed in the field there were reports of vaccine failure (10, 11). The exact reasons for these reported failures are fully not known, but the complexity of dose and administration does not help.

Unsurprisingly, other methods of vaccination have been developed, including using vectored vaccines. The technology for vectored vaccines first emerged in 1972 with Jackson et al. inserting fragments of various genes into circularised SV40 (12). Since then, vectored vaccines have been developed and adopted quite widely, as they offer the opportunity for multi-valent vaccination if an attenuated species-specific pathogen is used (13). Additionally, they offer an opportunity to insert the antigen into, say a human, using a self-replicating organism that does not cause clinical disease (14). This would obviate the principal risk of using live attenuated viruses as vaccines, namely the risk of reversion to virulence, and in the case of the live NDV vaccines, shedding of live NDV (15).

Turkey Herpesvirus (HVT) was first used as a vector for NDV in the early 1990s, following work carried out on Fowlpox (16, 17). HVT is a Serotype 3 Marek's disease virus, which is commonly used to vaccinate chickens against Marek's disease in the same manner that Edward Jenner observed Cow Pox protected against Smallpox (18, 19).

Morgan et al, followed by Reynolds et al inserted the NDV F protein gene into HVT's thymidine kinase gene and confirmed that the protein was expressed by the HVT vector (17, 20). This has the advantage of creating a bivalent vaccine, which generates immunity to both pathogens, with no mortality or morbidity associated with NDV. In later years Innovax also added an antigen from Infectious Bursal

Disease Virus (IBDV), making it tri-valent, while other poultry viruses, such as Infectious Laryngotracheitis are having vaccines developed in a similar manner (21-23).

While providing long-lasting levels of immunity against the virus and able to be administered in-ovo, HVT-based vaccines are not without their problems. The HVT virus itself is cell-dependant, necessitating liquid nitrogen storage, which will be very difficult to maintain in the sub-Saharan climate where the virus is most prevalent (24, 25). Timing of the dose is also problematic as maternal antibodies can interfere with the development of natural immunity (26).

As such the HVT-vectored vaccines may not be suitable for deployment to developing countries, but they do supply a niche that provides the poultry flocks with immunity to the pathogen without having to introduce live vaccine that still causes some morbidity or could even mutate or revert to virulence. Recently a combination of HVT-vectored vaccines, followed by a live or inactivated NDV booster has proven successful, allowing the benefits of both types of vaccine to complement one-another (27).

The advent of CRISPR/Cas9 caused a veritable stir when it emerged as a viable tool for gene editing. It realised the possibility of performing site-specific gene insertions into more complex genomes, with greater efficiency, ease and at lower cost (28). As such it has been used in the development of African Swine Fever, Avian Influenza and even SARS-CoV-2 vaccines amongst others (29-31).

A system for preparing gene fragments for insertion into HVT as a way of rapidly devising vectored vaccines was recently developed within the Nair group at the Pirbright institute (32). This was chosen to create a vectored vaccine with stabilised F protein. It makes use of a plasmid containing an insertion cassette, into which the gene of interest is cloned. The cassette is then transfected into chicken embryonic fibroblasts (CEFs), along with plasmids containing the CRISPR/Cas9 gene and also the specific guide RNAs. These CEFs were then infected with HVT and the cassette with the foreign gene was inserted into the HVT genome. The GFP reporter added to the insertion cassette, once expression of HVT was stable, was removed by Cre-Lox excision.

Other viruses are being trialled as vectors for NDV vaccines, namely duck enteritis virus, another avian herpesvirus (33). This is likely to suffer from similar requirements for cold chain storage and distribution as HVT, limiting its usefulness (25).

In more recent years, scientists have begun experimenting with inserting strain-matched F proteins into these vaccine vectors with mixed efficacy. In an interesting turn of fate, NVD itself has been considered for use as a vaccine vector for human pathogens, most notably, SARS-CoV-2 (34).

Nanobodies and antibody fragments (FABs)

Nanobodies are the heavy chain of camelid antibodies, usually obtained from llamas (35). In recent years they have found use two-fold in structural biology. Firstly they have been used to map epitopes of where antibodies bind to antigens, and determine which areas, when targeted by antibodies, are neutralising. Their second function is to increase the size of proteins, allowing sub-100kDa proteins to be imaged by Cryo-EM, indeed combinations of protein specific nanobodies and anti-nanobody antibody fragments (FABs) have been used to increase the size of proteins (36).

While both functions could and are fulfilled by traditional FABs, nanobodies have a less flexible structure and are able to 'Lock' proteins in specific conformations, thereby allowing structural biologists to capture data on conformations that are not normally conducive towards cryo-EM or too flexible for a technique like X-ray Crystallography (37).

It would not be too much of a leap to suggest that viable nanobodies may hold the key to solving the structure of the NDV F protein in its pre-fusion conformation.

Furthermore, any FABs or nanobodies raised against F, which are bound in complex during crystal formation or the preparation of grids for cryo-EM would offer insights into the mechanism by which different epitopes are neutralising, or not as the case may be (37).

In the case of the VHH nanobodies received from our collaborators in China, we were informed that the three strains of nanobodies raised were non-neutralising. While this is disappointing, it offers an opportunity to assist in imaging the F protein in both pre and post-fusion conformations, if the binding does not inhibit the conformational change.

Results:

A note on a most embarrassing setback:

The results contained in this section pertain to the second time that the insertion constructs were prepared for insertion into the HVT. The first time round mirrored the methodology used here, but a very substantial oversight was made. Expression of the F protein in the PCDA3.1 and PGEMTeasy vectors, in cell culture was not checked due to a combination of inexperience and the dependency on a GFP reporter in the PGEMTeasy vector, assuming that if the GFP is present, then the gene is inserted. This assumption prevailed right up until the stage where the four constructs had been 'inserted' into the HVT vector. Upon immunostaining, it became very clear that while GFP was being expressed, the F protein was not.

Upon sequencing, it appeared that a rogue insert had found its way into the construct. As such the process began again from scratch with checks performed at every stage.

Creating the insert for HVT

In order to create the prototype HVT-NDV vaccine, the NDV insert had to be prepared for insertion. The genes of La Sota and GTVII NDV, both with (G) and without (/) the GCNt domain, were amplified successfully by PCR, before being run on agarose gels to isolate the amplified fragment see figure 79. Details of the primers used can be found in Chapter II, which added motifs for digestion by the restriction enzyme NotI. The bands were excised from the gel and the DNA extracted.

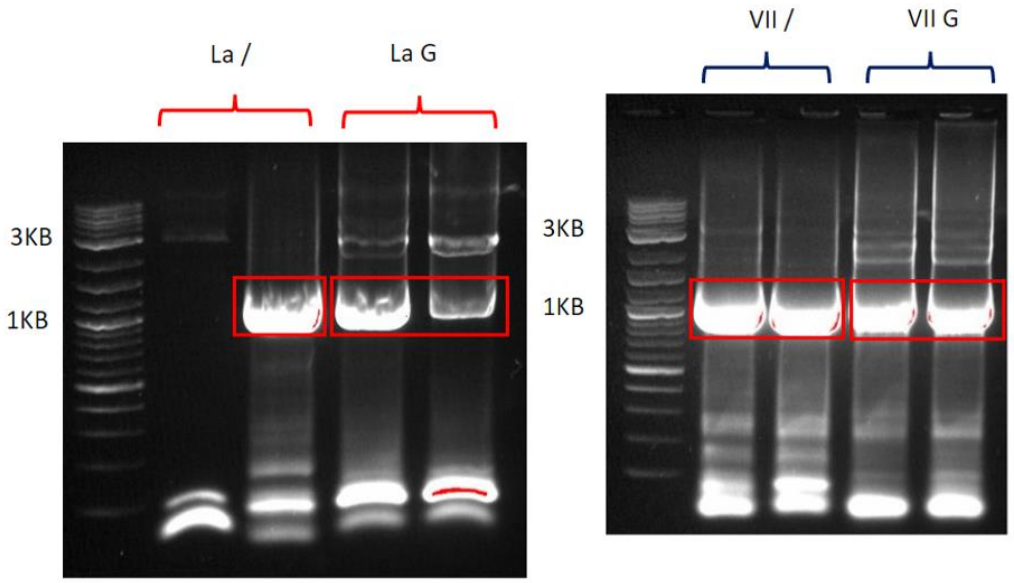


Figure 79: The agarose gels showing the amplification of the four F constructs and addition of NotI restriction enzyme sites by PCR.

The amplified NDV F gene and PCDA3.1 were both digested with NotI and ligated. The ligated DNA being used to transform bacteria in SOC media before being transferred to agar plates. The colonies that emerged on the plates were tested for presence of the insert by PCR, see figure 80.

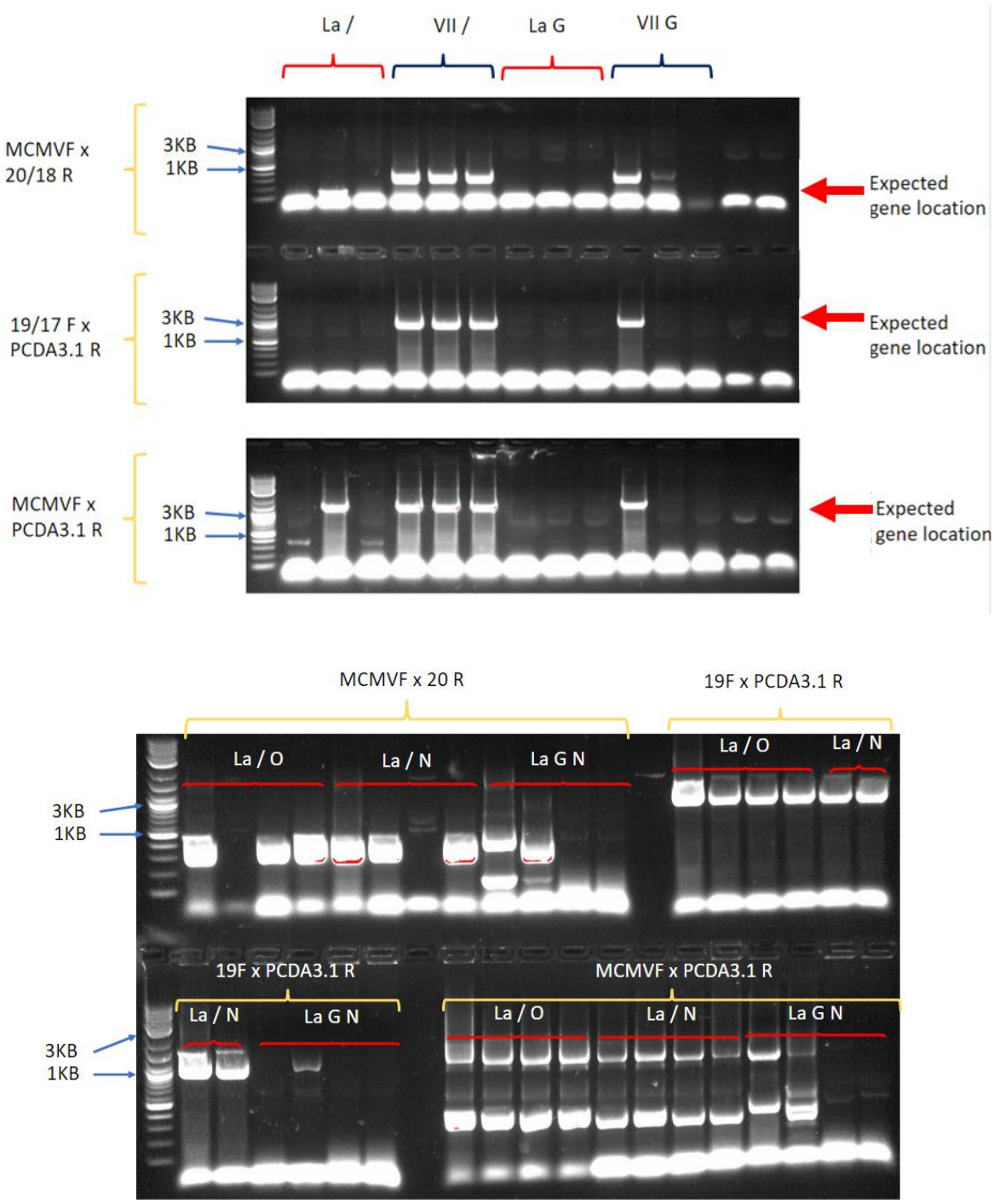


Figure 80: Agarose gels run to visualise the PCRs run to confirm the presence of the F protein genes in the PCDA3.1 plasmid.

The panel of confirmatory PCRs were able to identify which cultures had the F genes successfully inserted. These could then be used for further construct creation.

Insertion of NDV F into PGEMTeasy was the second stage of this process. The PCDA3.1 plasmid with the inserted NDV F gene was digested SfiI to extract the gene and the flanking sequence necessary to insert into HVT. This was then ligated into a similarly digested PGEMteasy vector and used to transform cells, as described above, and in detail in Chapter II. These clones were screened by PCR, the results being seen in Figure 81.

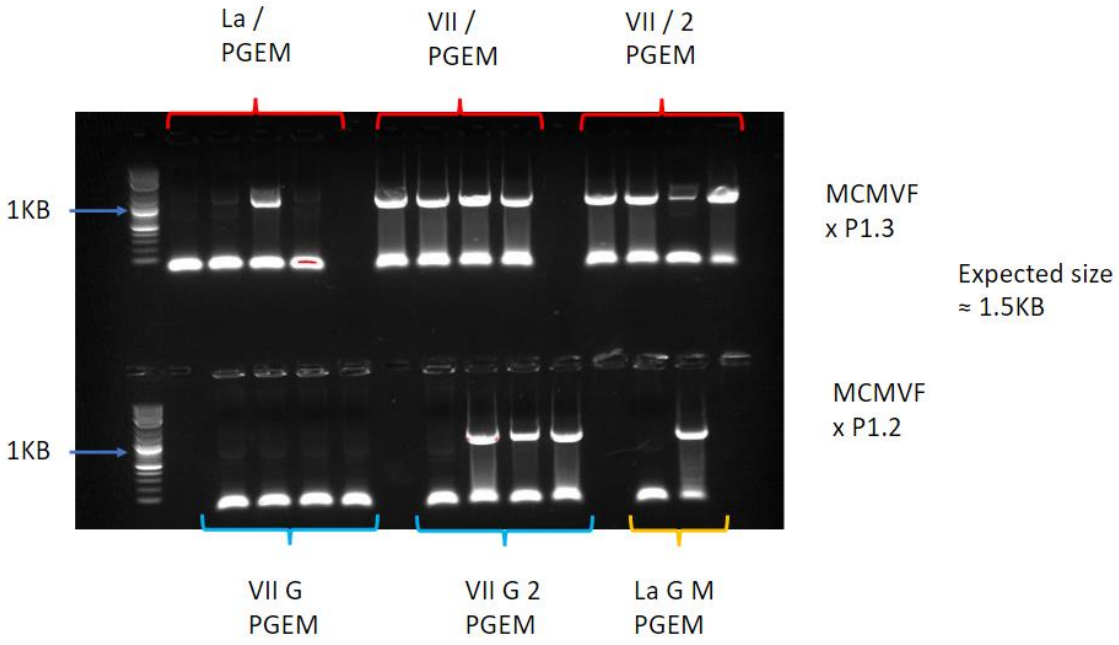


Figure 81: The Agarose gels run to visualise the PCRs run to confirm the presence of the F protein genes in the PGEMteasy plasmid.

The presence of the band indicates the presence of the insert in the clone. The clones with the gene present could then be amplified and the DNA harvested. Once the constructs had been created, it was time to insert them into the HVT vector, though this was not without difficulty.

<u>Confirmation of the insertion being able to express:</u>

After the incident detailed above, it became very important to ensure that the gene not-only was present but expressed within the cells. Firstly, an immunofluorescence was carried out on the F protein in the PCDA3.1 plasmid to confirm expression, in figure 82 the cells transfected with the PCDA3.1 with the La Sota gene were stained with anti-6-his primary antibodies and GFP tagged secondary antibodies.

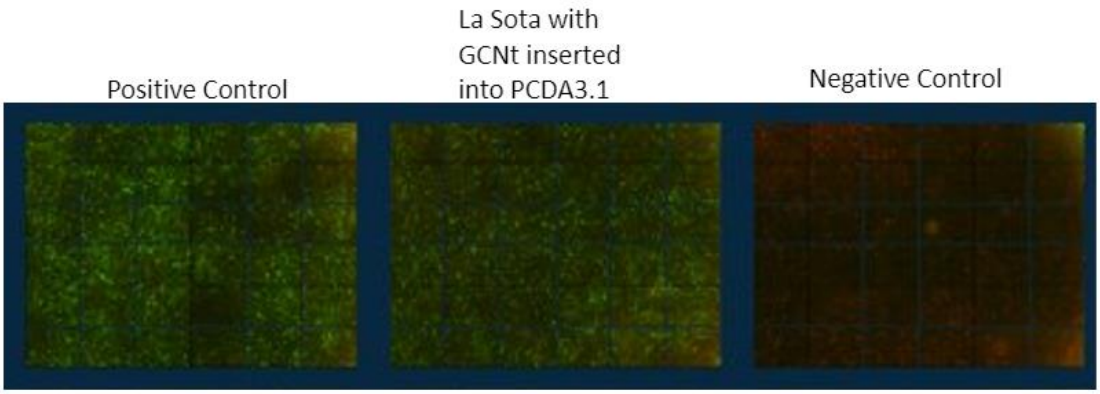


Figure 82: Fluorescent micrographs showing cells transfected with the La Sota F protein in the PCDA3.1 plasmid stained with anti-6-His tag primary antibodies and GFP tagged secondary antibodies. GFP excitation at 488nm and detection at 510nm.

Afterwards, when the gene insert was further inserted into the PGEMTeasy vector, it was essential that the F protein was also expressed in this construct. The PGEMTeasy vector includes a GFP tag that expresses when a foreign gene is inserted into the SfiI restriction site. In figure 83, the construct was stained with anti-6-His primary antibodies and RFP tagged secondary antibodies. It was necessary to find a population of cells where the GFP and the RFP, when merged, aligned.

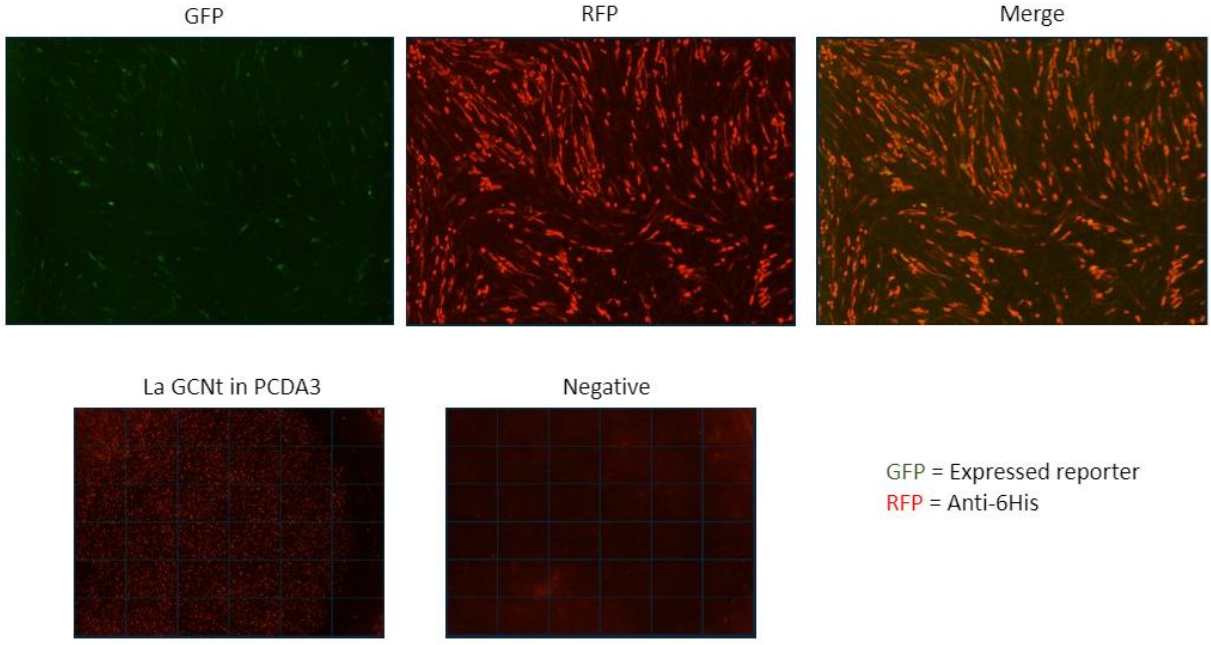


Figure 83: Fluorescent micrographs showing cells transfected with the La Sota F construct, in the GFP expressing PGEMTeasy plasmid, stained with anti-6-His tag primary antibodies and RFP tagged

secondary antibodies. GFP excitation at 488nm and detected at 510nm, RFP at 532 and 588, respectively.

As seen, the GFP levels in the LaS(with GCNt) insertion were much lower than expected, however the expression of the protein was acceptable. As time was limited and resources within the lab were stretched, due to supply chain issues, only the LaS (with the GCNt stabilisation) was carried forwards for insertion into HVT. Once the parameters for successful insertion of this F construct were established, insertion of the others would be possible.

In a 'Belt and braces' approach, the construct in PGEM-Teasy was sent for sequencing, see figure 84.

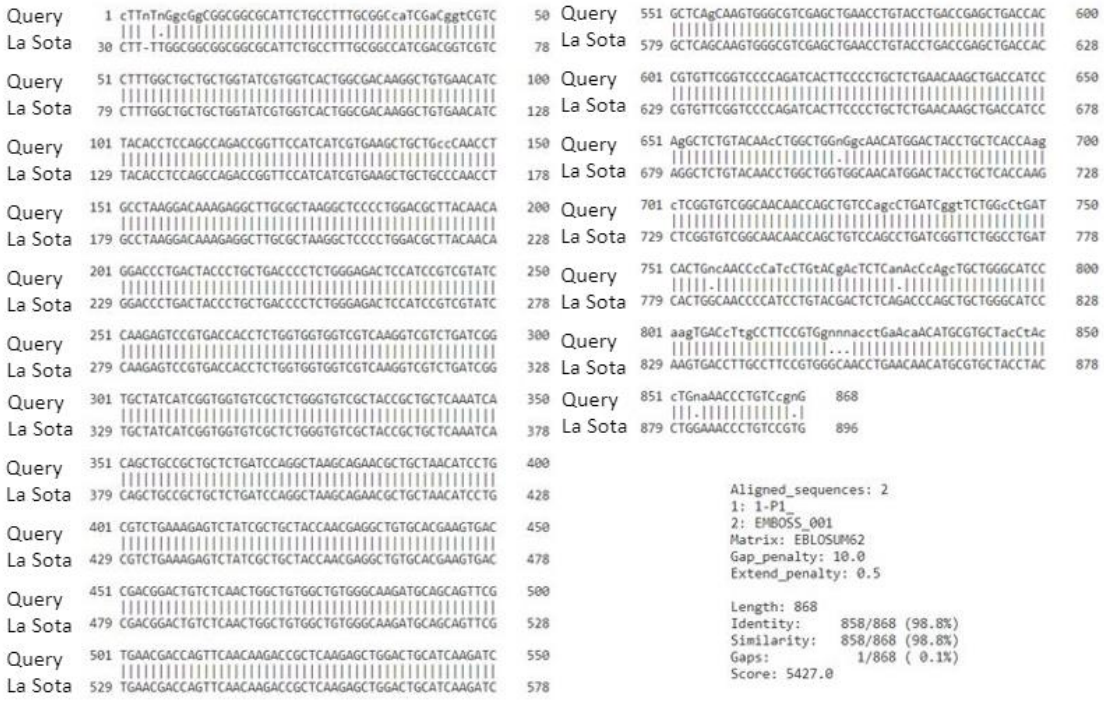


Figure 84: A comparison of the F gene inserted into the PGEMTeasy plasmid, sequenced by Eurofins, to the original LaS F sequence, carried out by WATER from EMBOSS.

Comparing the sequence returned from Eurofins to the sequence of the F protein construct used, there appears to be a good match, with all the variation clustered at either end of the read.

Insertion of the fragment into HVT:

Various attempts were made to try and insert the construct containing the stabilised F protein into HVT, including varying the ratios of reagents within the insertion, varying the titre of virus used, using fresh aliquots of the plasmid containing vectors, but none were effective. For reference a number of HVT plaques 48 hours after being picked from the initial transfection, are shown in figure 85.

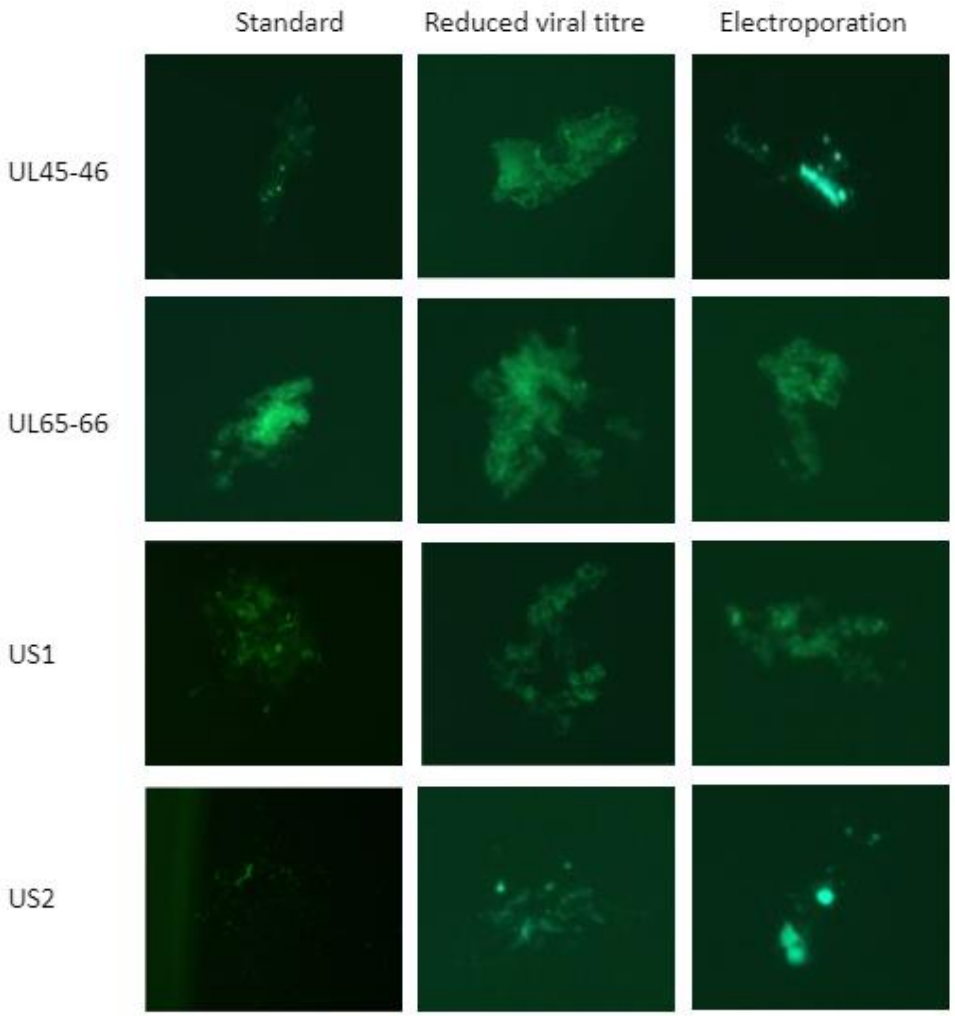


Figure 85: Fluorescent micrographs of HVT plaques with NDV F constructs inserted into different regions of the HVT genome under different conditions. GFP excitation at 488nm and detection at 510nm

While not being especially convincing, these plaques were transferred to 24-well plates seeded with CEF cells. After a further 48 hours incubation, the wells were imaged to look for plaques. No GFP positive plaques were seen.

Confirmation of the positive insertion (Hypothetical):

After plaques would have been picked from the 24-well plates they could be passed to 6-well plates to proliferate. FACS would then sort cells infected with GFP positive HVT, from cells infected with HVT without the insert. This was performed during the abortive first attempt at producing the prototype vaccine, with the results of this shown in figure 86.

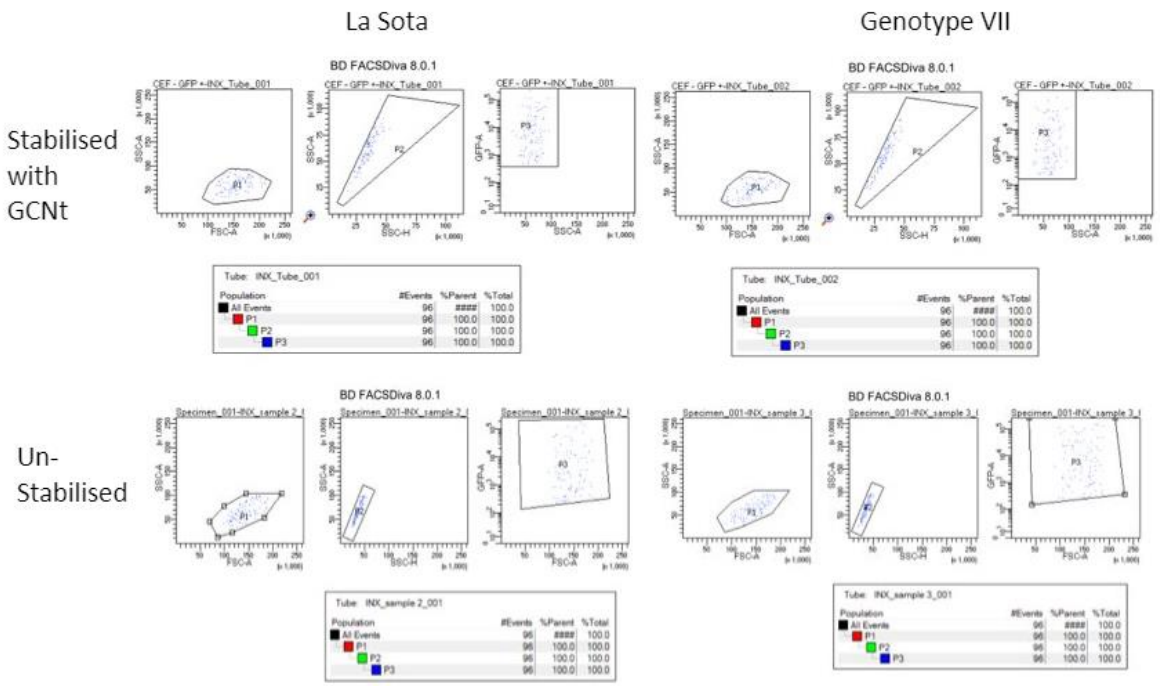


Figure 86: Example FACS plots used to separate HVT infected cells that have had the insert successfully inserted from those that have not, using the GFP cassette within the plasmid.

Whilst this set of FACS data illustrates nothing in the grand scheme of things, it illustrates what the population of GFP positive HVT infected cells would look like, had the insertion been successful.

The GFP positive population was sorted into 96 well plates. These would be imaged for GFP positive plaques and grown in duplicate. Half would be used to confirm, through immunofluorescence, whether the F protein was expressing. If expressing the F protein, the cre-lox system could be used to knock-out

the GFP reporter protein and a number of passages carried out to ensure construct stability. However that eventual goal was to remain elusive.

The animal experiment:

While at the Pirbright institute, the group took part in an animal study investigating the immunogenicity of several different avian vaccines, with 3 chicks allocated per group. The birds were immunised on day 0 and boosted on days 7 and 14. An ELISA (carried out by Dr Zhang) was used to quantify the humoral response to the antigens, see figure 87.

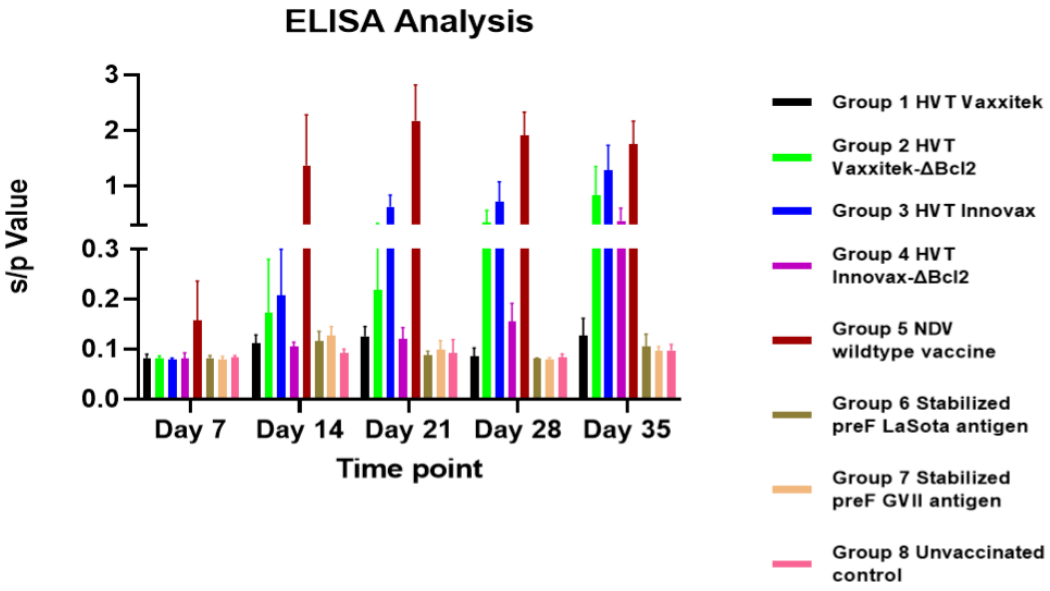


Figure 87: The results of an indirect ELISA analysis (caried out by Dr Zhang) of the effect of the 'stabilised' NDV antigens (synthesised from the constructs in use) on the generation of antibodies, compared with other antigens. NDV F antigen (supplied as part of the Innovative Diagnostics NDV detection kit NDVS-10P), was bound to the plate. This was hybridised with nanobodies from the chick blood samples. Anti-chicken conjugated with HRP was the secondary, emitting at 428nm. An s/p value above 0.3 is considered positive.

Unfortunately, the animal experiment showed that both proteins on their own were likely, not very immunogenic, with the s/p values recorded for the La Sota and GTVII antigens resembling those of the un-vaccinated control group, being far inferior to the La Sota based vaccine in commercial use.

Preparation of the Nanobody constructs:

Initial expression of the nanobodies

With the receipt of the nanobody sequences from our Chinese collaborators, the genes could be synthesised into expression vectors. If they bound the F protein, they could have aided with the formation of crystals for X-ray crystallography or assisted in stabilising the trimer for Cryo-EM. They were expressed by transfecting CEF cells with the plasmid. CEF were used as they were the only expression system available at the time. A Western blot was carried out to confirm Nanobody expression. See figure 88.

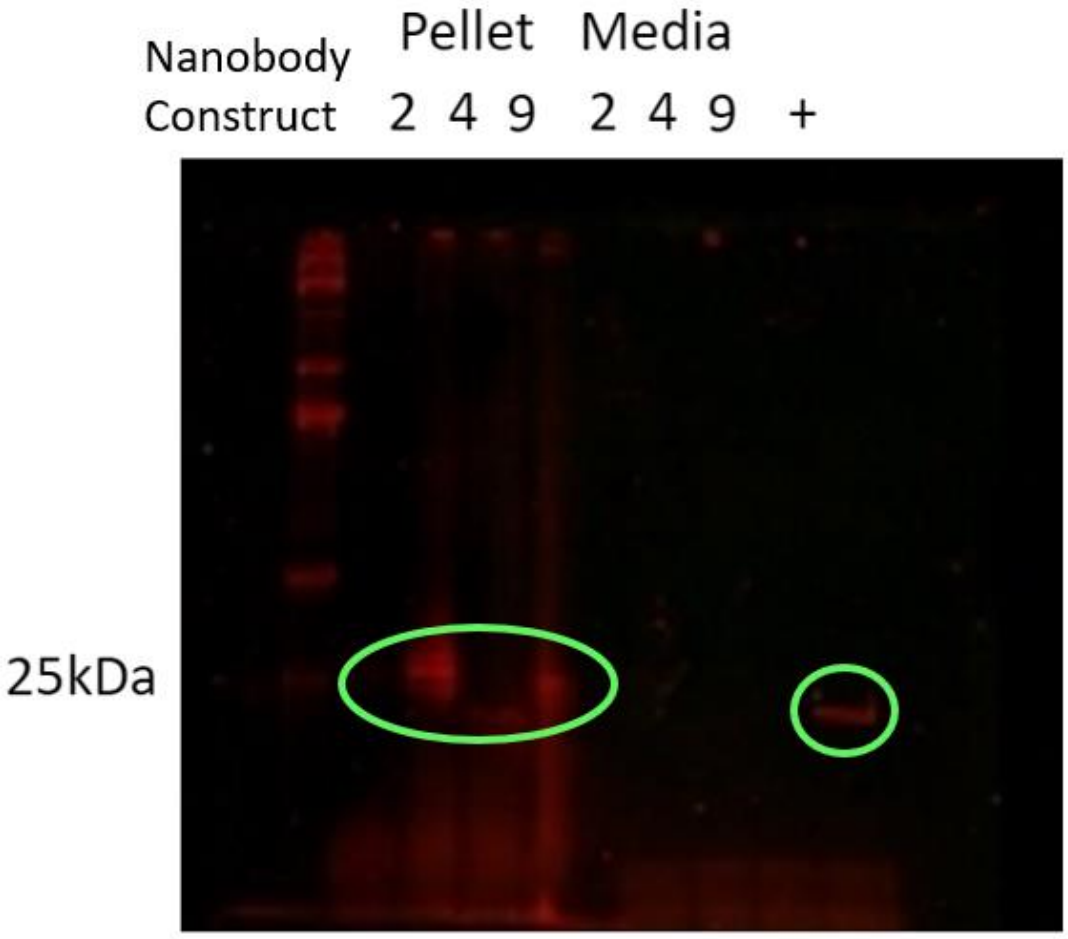


Figure 88: A western blot showing the expression of the three nanobody constructs in pellet and in the media. Circled are the bands for present nanobody in the cell pellet and the V5-tagged positive control

(A nanobody supplied by Dr Zhang. Mouse anti-V5 was the primary antibody. The secondary was HRP-tagged anti-mouse.

The blot suggests that all three nanobodies expressed, to some degree in the cell pellet and are approximately 25 kDa in size. The slight variation in size could well be due to bubbles within the gel.

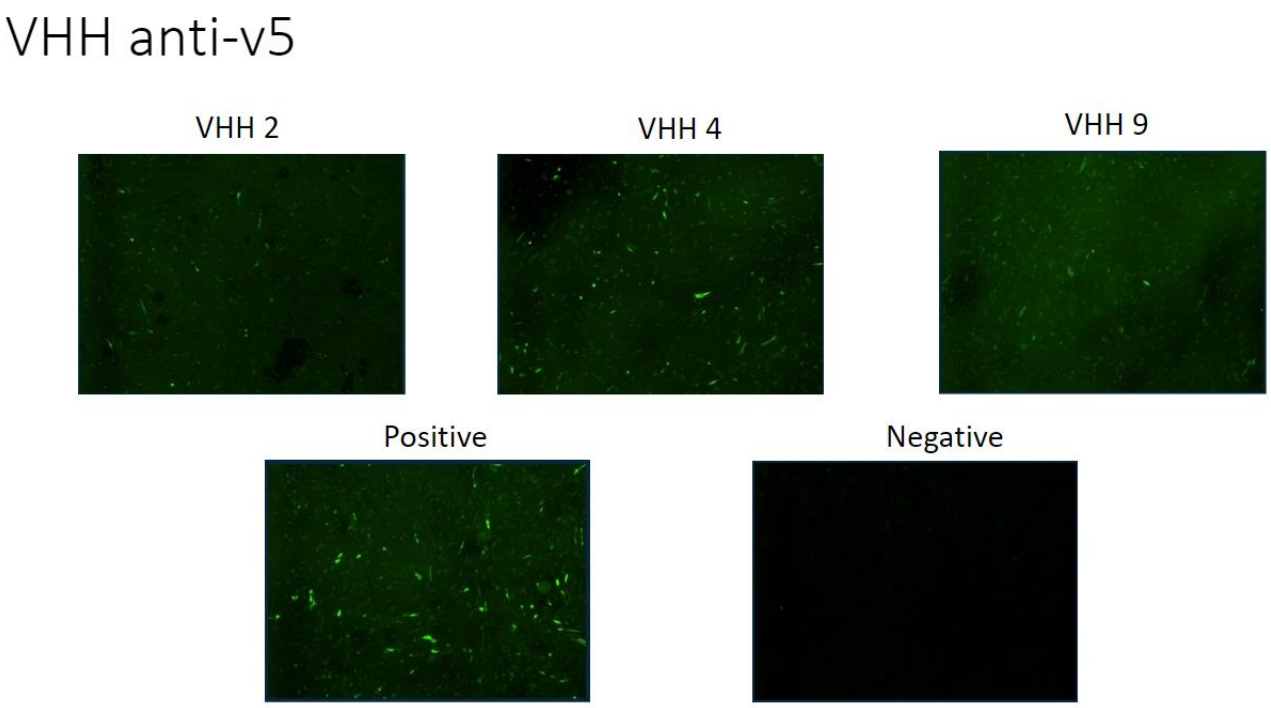


Figure 89: Fluorescent micrographs showing cells transfected with the nanobody constructs in the PCDA3.1 plasmid, stained with anti-V5 tag primary antibodies and GFP tagged secondary antibodies. GFP excitation at 488nm and detection at 510nm.

The second test of expression was an immunofluorescence assay. Cells transfected with the plasmid were fixed and stained with an anti-V5 primary antibody and a GFP tagged secondary. See figure 89. Of the three constructs expressed, VHH 4 appeared to be the most strongly expressing, though the other two appeared to express V5 tagged protein, I.e. the nanobody construct.

Initial binding assays:

Once the nanobodies were being expressed, a number of cross-binding assays were attempted to confirm whether the nanobodies bound the F protein. The first of these was carried out with two

populations of cells, half expressing the nanobodies, half expressing the F protein having their media collected. The media from one population was incubated with the other, the cells were then washed, fixed and stained. See chapter II for the specifics.

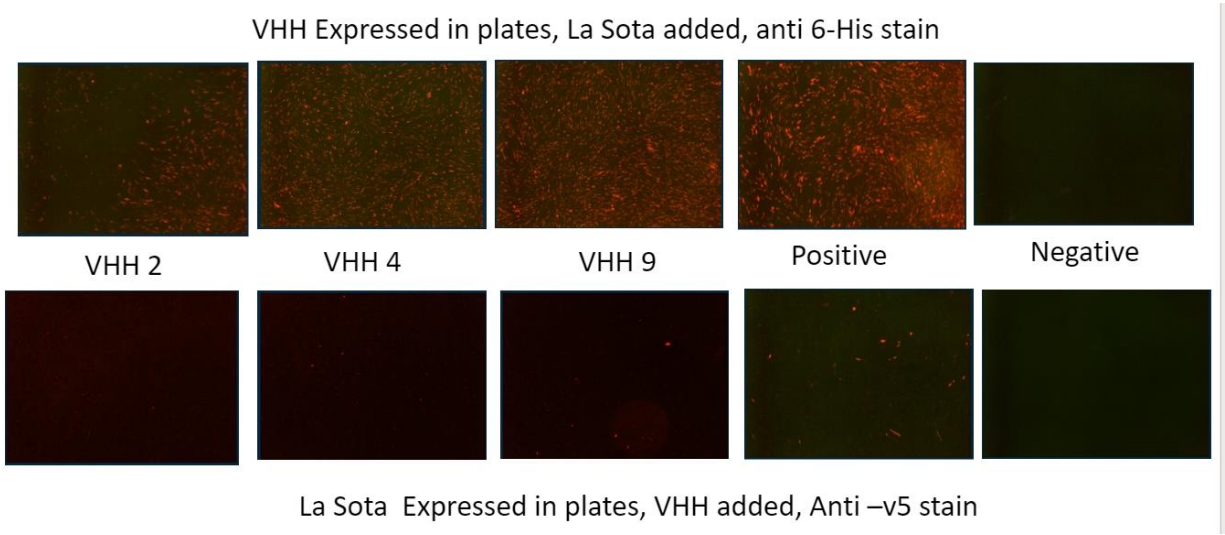


Figure 90: Fluorescent micrographs showing two groups of cells transfected with the nanobody constructs in the PCDA3.1 plasmid in one set and La Sota in the other. The media from these were collected, the cells fixed, and the media applied to the other group, incubated then washed stained with anti-V5 tag and anti-6-Hisprimary antibodies, with RFP tagged secondary antibodies. The same wavelengths for GFP and RFP were used.

In figure 90, the nanobody was expressed in cells, the cells fixed and the media from wells transfected with the F protein construct incubated in those plates. These were washed and stained for immuno-fluorescence. There appeared to be F protein bound to the nanobodies. However, this did not occur the other way around. A more concrete way of confirming expression was required.

Immunoprecipitation of nanobodies:

Performing an immunoprecipitation would confirm whether the nanobodies bind the F protein. Cells transfected with the nanobody plasmid were harvested, pelleted and lysed, as described in Chapter II. The lysate was then incubated overnight with purified F protein, before the immunoprecipitations were carried out. With two lots of antigens, the immunoprecipitations were carried out both ways. Firstly, anti V5 antibodies were bound to the Dynabeads, allowing the Nanobodies to bind. If any F was bound to

the nanobodies, it would have been extracted by the beads. A Western Blot was then run with the lysates, using an anti 6-His antibody primary to confirm the presence of the F protein, see figure 91.

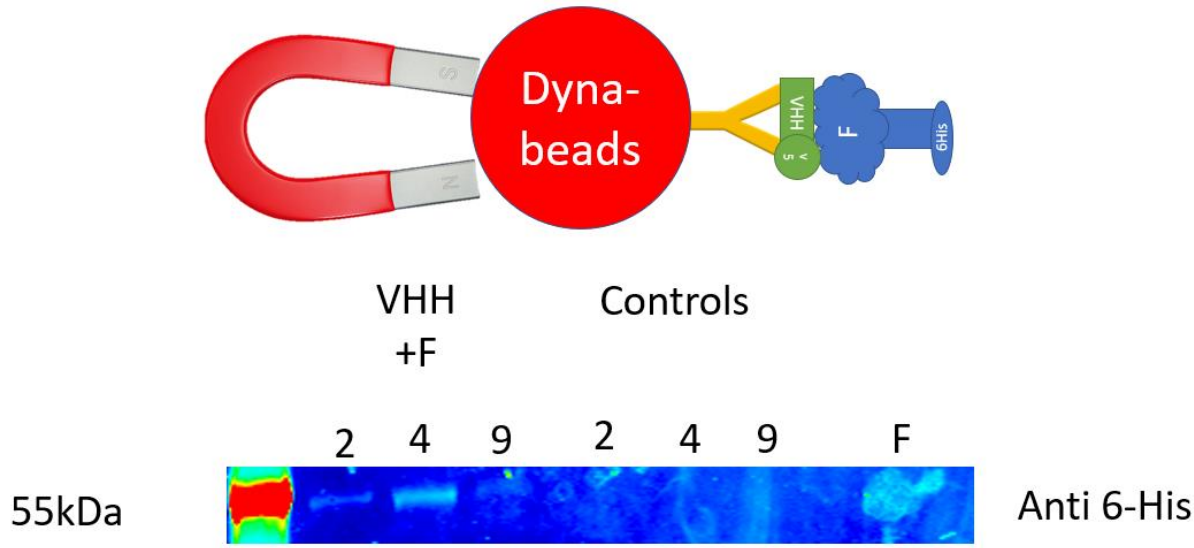


Figure 91: A western blot showing the presence of anti-6-His tagged protein (La Sota F protein) pulled from solution by immunoprecipitation with bead-bound nanobodies.

The first round, where the nanobody was bound to the bead via an anti-V5 nanobody, suggested that all 3 nanobodies bind the La Sota F protein. This arrangement was then reversed, with anti 6-His antibodies being bound to the beads. These could then pull-down any nanobodies bound to the F protein and visualised with a western Blot. See figure 92.

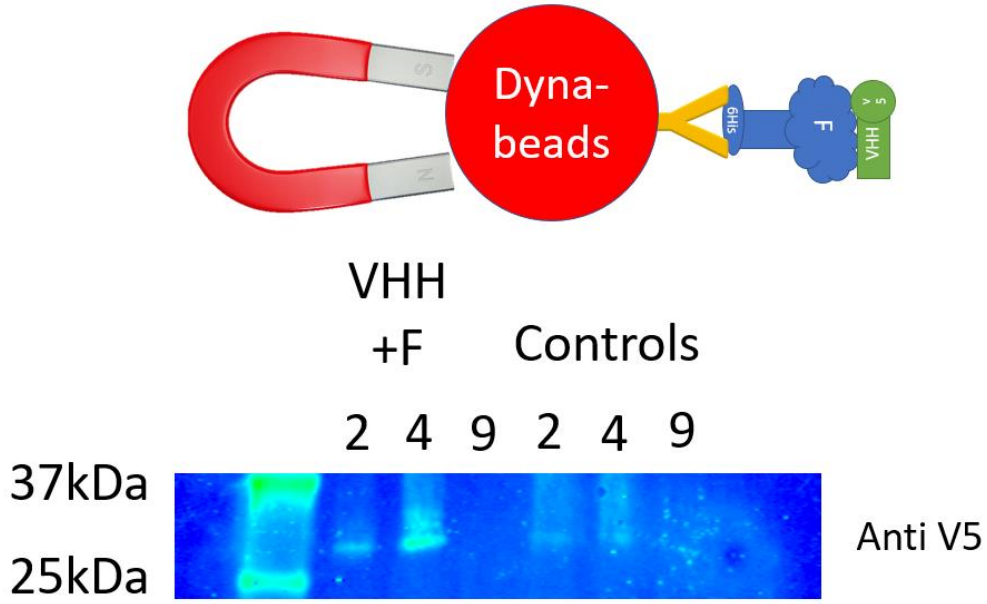


Figure 92: A western blot showing the presence of anti-V5 tagged protein (nanobodies) pulled from solution by immunoprecipitation with bead-bound La Sota F protein.

With the Anti-His antibodies binding the F protein to the beads, it was possible to stain for the presence of each nanobody. Of the three, only nanobodies 2 and 4 bound, whereas number 9 appeared to have expression issues. This is where the foray to the Pirbright institute ended, and a return to Oxford was possible.

Preparing nanobodies for large scale expression:

The first task for expressing the nanobodies for use in structural biology was to clone them into a more suitable vector for expression and purification. Once the nanobodies had successfully been cloned into the pMESy4 vector, it was necessary to see if they expressed, so a small-scale expression run was carried out using competent Rosetta cells, see Chapter II. The lysate from the cell transfected with the new plasmids were loaded onto a His-trap and eluted, see figure 93.

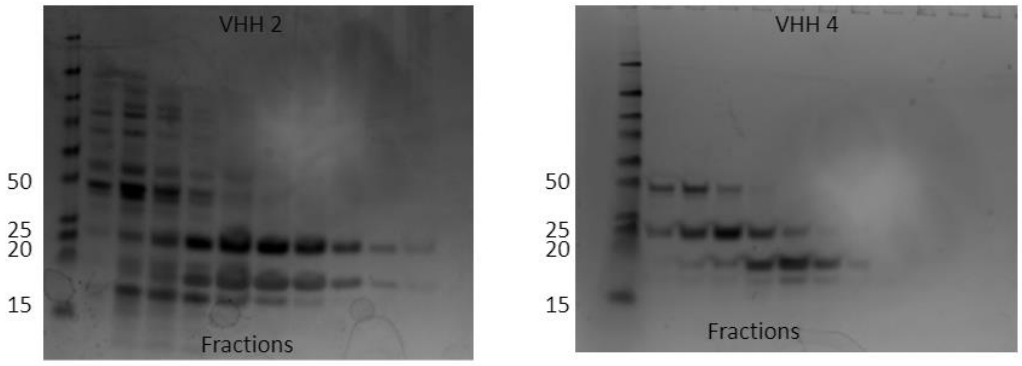
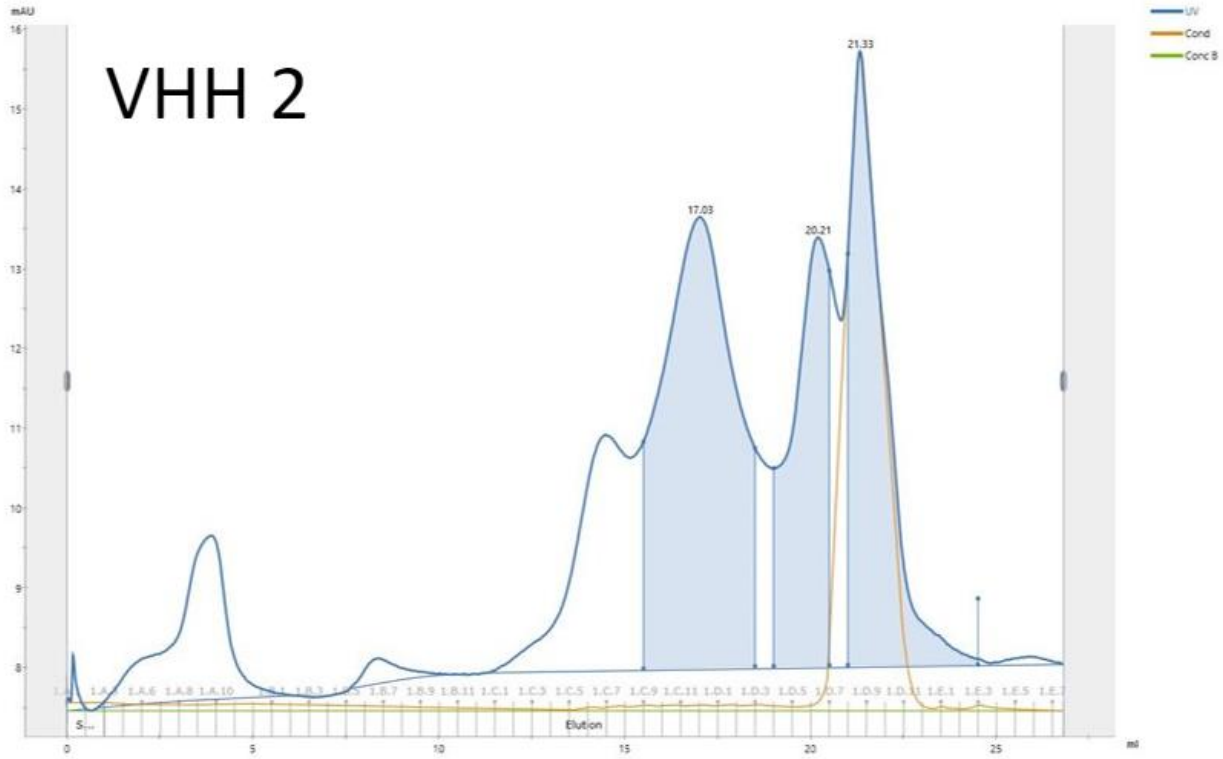


Figure 93: The contents of the fractions eluted from a His Trap run on SDS PAGE gels.

The elute was somewhat dirty, though the fractions containing protein of the 25 to 15 kDa range was collected. The elute was subjected to size exclusion, to attain a homogenous sample. See figure 94



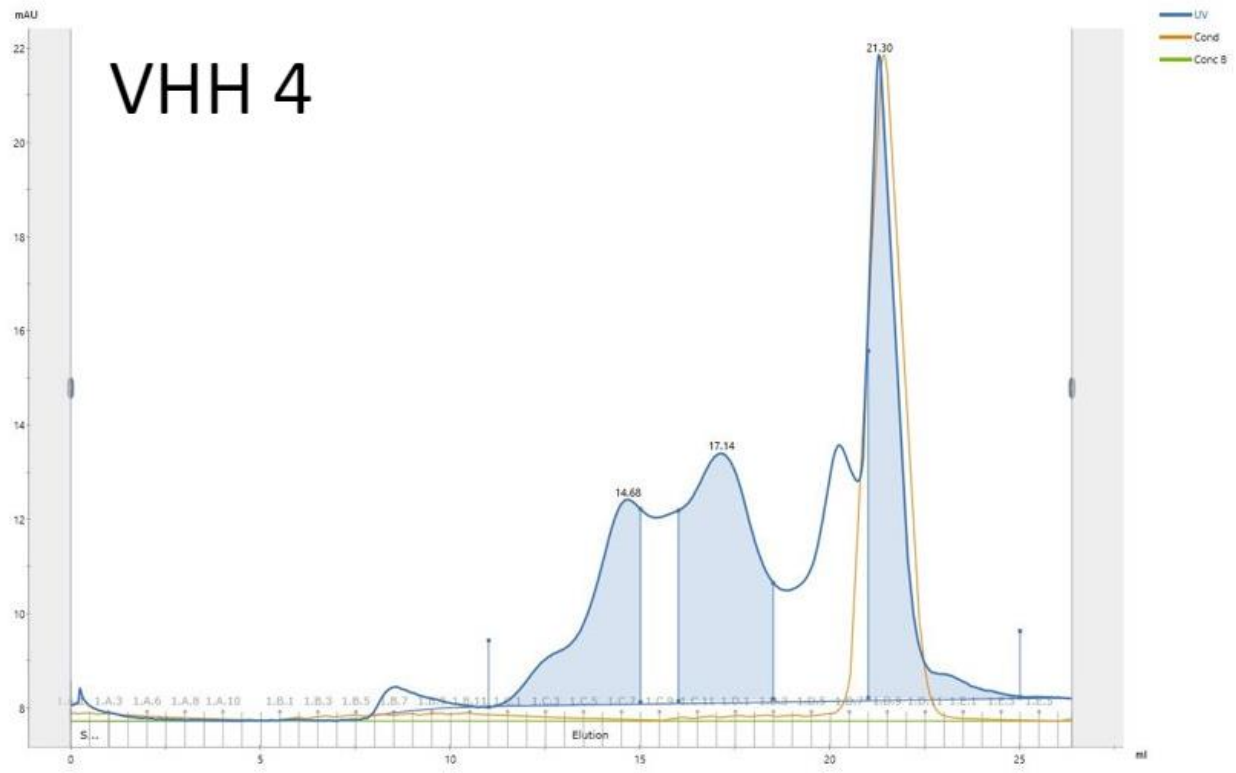


Figure 94: The traces generated by size-exclusion of the nanobodies expressed in Rosetta competent cells on AKTA machines for 2 of the nanobody constructs. SEC was carried out using a Superdex S75 10/300.

The size exclusion revealed that there were multiple sized protein within the elute from the his-Trap, so a western blot was performed on the fractions collected to determine which was the correct one, See figure 95. A yield of approximately 0.5mg was obtained from a 2L culture.

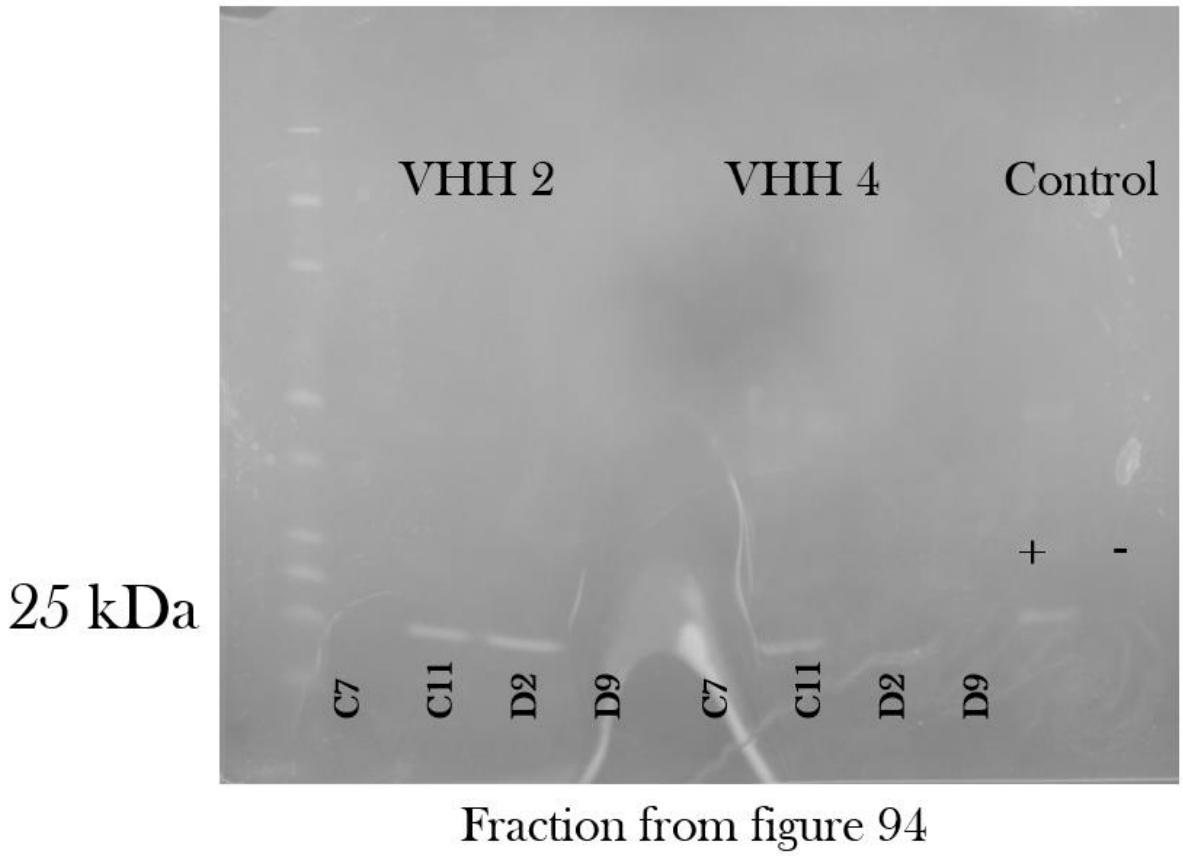


Figure 95: A western blot showing the presence of 6-His tagged proteins (Nanobodies) in the fractions separated during size exclusion shown in the previous figure. The positive control was 6-His tagged nanobody, the negative control was a strep-tagged nanobody, both supplied by Dr Jeremy Keown.

Examining the blot, the nanobodies appear to be expressing in the C11-D2 peak on size exclusion, corresponding with roughly 30-20kDa. This allowed for the purification of the nanobodies.

Binding assays:

Once a suitable expression system had been developed, it was possible to express the nanobodies at sufficient scale, binding assays could be attempted. Purified La Sota F protein was incubated with Purified nanobody, with the latter at 3-fold molar excess for 1 hour. The mixture was then separated by size exclusion chromatography, with the sample from the F protein peak and the Nanobody peak being subjected to a Western Blot. As both proteins have an anti-6-His antigen, the same anti-6-His antibody could be used.

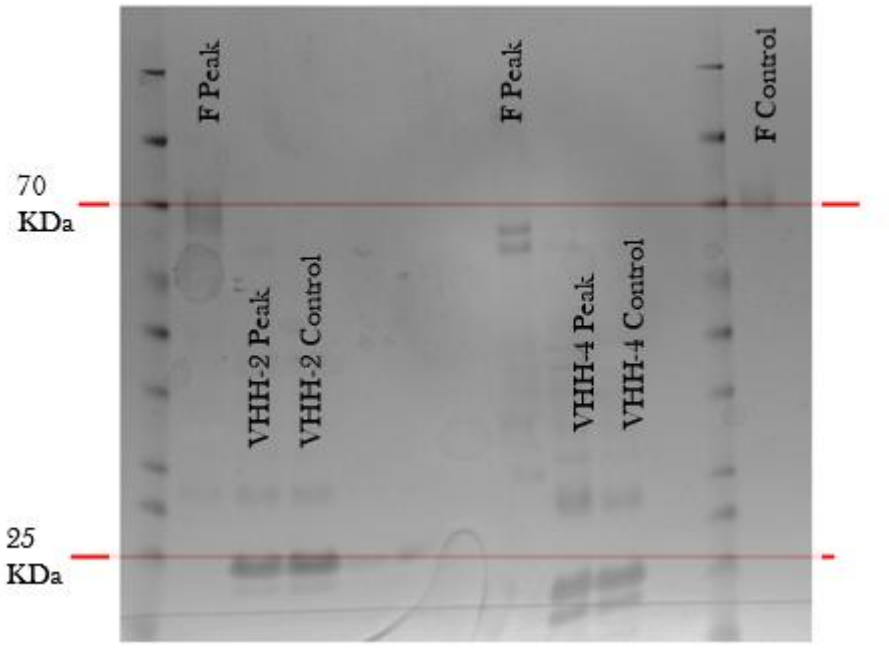


Figure 96: A western blot showing F protein incubated with 3-fold molar excess of nanobody, run through size exclusion chromatography, then subjected to a SDS PAGE under reducing conditions, stained with anti-6-His primary antibodies.

Under the reducing conditions of the Western blot, any nanobodies bound to the F protein would have dissociated and be visible on the western blot, as a band beneath the F protein band. It appears that the F protein did not bind to either of the nanobodies. The same experiment was attempted, incubating F protein which had been cross-linked with the nanobodies, as seen in figure 97. This was carried out contemporaneously with the cross-linking experiments outlined in chapter IV.

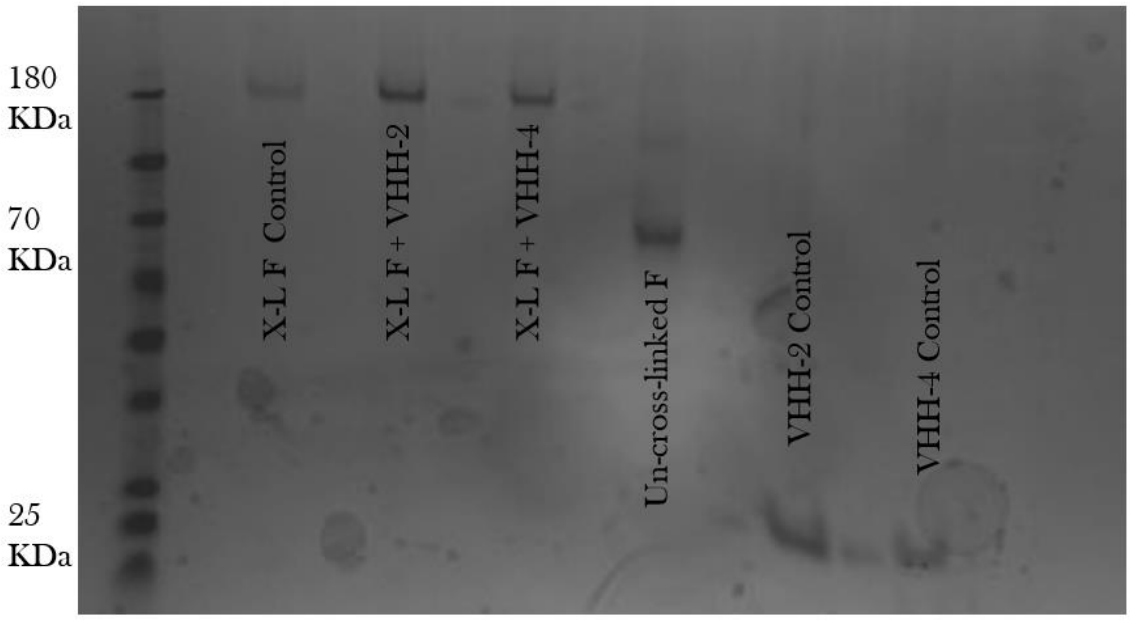


Figure 97: A western blot showing cross-linked F protein incubated with 3-fold molar excess of nanobody, run through size exclusion chromatography, then subjected to a SDS PAGE under reducing conditions, stained with anti-6-His primary antibodies. 'X-L F' refers to cross linked F protein.

The cross-linked F protein appeared to not to bind to the nanobodies either, as evidenced by the absence of bands of the nanobody's size under the F protein column.

Crystal trays:

It was possible that the nanobodies would assist the formation of crystals. So, crystal trays of the F protein incubated with each of the nanobodies, were set up (as per chapter III). None of the trays saw the formation of crystals.

The absence of VHH 9:

Upon close examination, the VHH 9 construct was expressing at a much lower level than the other two nanobodies and was described by the collaborator as "Being of lower affinity binding to the antigen." Information on how this was measured was not supplied. Coupled with low levels of expression, time limitations determined that the other two constructs were prioritised.

Discussion:

HVT has been used for years as a vector for many other viral antigens and the method used in this work was recently developed within the group (32). However, the inability of being able to insert the fragment of F into HVT does raise questions.

The first round of insertions, which appeared to work, until it was realised that the F gene had not been inserted, was interesting because the system was designed that the GFP reporter would only express if a gene was inserted into the cassette. The fact that it began expressing GFP suggests that something was inserted. It does explain why the first attempts to perform IF on the HVT constructs were unsuccessful and gave a lesson on needing to perform IF at each stage of the cassette synthesis.

During the second attempt to create the insertion cassette for the F protein, there was another issue. While the F gene in the donor cassette was shown to express by western blot and IF, when insertion was attempted, it would appear to succeed temporarily. HVT plaques expressing GFP, when picked initially would continue to fluoresce, though on secondary picking, these would diminish. Perhaps the cassette had not inserted properly, or the HVT genome was able to remove the insertion through the proof-reading mechanism built into the viral polymerase.

The trial of different guide RNAs to allow the use of different insertion sites for HVT implies that it was not the location of the insertion that was at fault, perhaps it was the methodology or the or the materials used to attempt the insertion. The use of electroporation in lieu of transfecting with lipofectamine was also unsuccessful. It is likely that the cell mortality form electroporation was too high.

With multiple trials of each aspect of the insertion process being trialled, the ratios of each plasmid, time before infection, the use of different aliquots of HVT etcetera, the conduct of the methodology was not at fault, but likely the constructed cassette. What must also be remembered is that CRISPR/Cas9, though increasingly utilised, is still quite a young technique. While the understanding of the system's complexities is ever growing, not every facet is completely understood. Time constraints and the need to return to structural studies lead to this aspect of the project being discontinued.

In the animal experiment, the inferiority of the stabilised F protein to the commercial NDV vaccine was not unexpected, as a full NDV vaccine has been especially formulated to raise antibodies, with effective combinations of adjuvant and other factors to encourage such production. Perhaps the age of the sample and the freezing process had degraded it. Differences in the adjuvant could also be factor in the limited effect of the protein alone. During subsequent work it emerged that much of the F protein used to immunise the chicks had been purified using a Superose 6 column. It was suggested that this column could not distinguish between monomeric and trimeric protein. Perhaps this was also a factor in the low immunogenicity of the protein. This was not considered credible by the examiners.

The Nanobody sequences offered a major opportunity to solve the structure of the F protein through forming complexes that stabilised the F protein, and also another source of disappointment. The early successes in expressing the antibodies and the results of the immuno-precipitation implied that the nanobodies would be suitable for future structural work.

For structural work, it was necessary to clone the gene into an expression vector with a suitable tag to allow purification. It then required extensive optimisation to achieve suitable yields of nanobody. It was disappointing that when the nanobodies were incubated with F protein, run down a size exclusion column and assayed with western blot, appeared not to bind. There could be many reasons behind this:

Firstly, the F protein the nanobody was raised against could have a fundamentally different conformation to those with which the size exclusion was performed. The F protein sample used in the initial IPs was the same batch as was used to raise the nanobodies, where-as the F used for the binding and size-exclusion was freshly purified. In order to preserve the protein for shipment to China, the protein was flash frozen in liquid nitrogen and stored at –80°c. Given the theory as to how plunge-freezing caused the F trimer to disassemble (See chapter IV), it is possible that the F that the nanobodies were raised against was in-fact some form of monomer, hence the frozen samples of F used for the IP worked, but the binding assay with the freshly purified trimeric protein did not. The principal flaw in this theory is that monomeric F protein was purified and incubated with the nanobodies, and also failed to bind. The same is true of F protein that was repeatedly flash frozen and thawed before

incubating with nanobody, as was post-fusion F. As such it could well be an issue with the nanobody itself.

Secondly, the cloning of the nanobody genes into a vector for large scale expression, adding a c-terminal his-tag could have fundamentally altered something in the structure of the nanobody and the tag could interfere with the binding to the ligand. The latter has been reported in numerous cases and could have an impact here. His-tags are notoriously flexible and could well be folding round to obscure part of the binding site (38). One way to test this theory would be to swap the His-tag for another tag, or change its placement, though this could not be done, due to time constraints. This could well be the most likely scenario.

Thirdly, the nanobodies could be denatured at some stage during expression or purification. Using osmotic shock to lyse *E. coli* has been reported to de-nature some proteins, as can chemical lysis or sonication. Indeed, factors like temperature can also have an impact on protein folding. It is possible that expression in a bacterial system led to different folding than when expressed in avian cells. It would however have taken a lot of time to test each of these potential factors to identify the problem.

In retrospect, expressing a known nanobody that was known to express well, in parallel with the anti-F constructs, would have provided a useful positive control. Chemical and sonication-based cell lysis was not used because they resulted in relatively low yields of nanobody, while osmotic shock was the most efficient. It could well be that in solving one problem, another was created.

It is unfortunate that both attempts to move the project forward by creating the vectored vaccine and through raising nanobodies came to naught. But despite this both struggles were major learning points moving forwards. Ideally the segment of work carried out at the Pirbright Institute would have been carried out following any structural discoveries. The pandemic and restrictions on working necessitated that the work at Pirbright had to be brought forwards.

A second tranche of nanobodies was set to be raised in the winter of 2019, though the resources for this were diverted to work on SARS-CoV-2, meaning that the three epitopes gained were the only ones available. The genes for these were only made available in the spring of 2021.

Between more potential epitopes of anti F nanobodies, and with the knowledge of hindsight on expressing and purifying the nanobodies, it may be possible to use them for more effective structural studies. More-so than any other aspect of the project, this segment was affected by the pandemic, including in limited access to the labs and other disruptions to supply chains. It was an exceedingly difficult time, but still an important learning experience.

Chapter VI

Overall conclusions and future direction:

Avian Ortho Avulavirus remains a persistent threat to poultry two decades after the first protein structure of the post-Fusion structure of the F protein was published (1, 2). Subsistence farmers in developing countries depend heavily on poultry and even a mild 'Lentogenic' infection of a backyard flock would significantly impact the prosperity of these farmers (3). On the other end of the spectrum, the mere presence of NDV in many western countries would have draconian and expensive impacts on the poultry industry (4).

The variation and evolution into over 20 distinct genotypes make creating a specific and broadly effective vaccine difficult. La Sota and the B1 are live vaccines generate humoral and cell-based immunity to the F and HN proteins, while Innovax only presents the F protein (for NDV) as an antigen (5, 6). The F antigens used by these vaccines are genotype II but provide limited protection against different genotypes (7). Without an accurate structure the structural elements of these changes to the Fusion protein of different genotypes cannot be calculated with a degree of confidence. While software, such as AlphaFold2 could be used to model the effect of these changes on the surface of the F protein, no structure of the F protein in the pre-fusion conformation yet exists (8). This makes validation of these models more difficult.

It is hypothesised that the pre-fusion conformation of F is more immunogenic than the post fusion conformation, based on work on other class I fusion proteins (9). Thus a pre-fusion structure of the F protein would be of critical importance. Membrane fusion proteins, and class I fusion proteins in particular, are non-trivial to work with due to their meta-stable nature. The meta-stable pre-fusion form of F rapidly transitions to the post-fusion form, which has hindered efforts to analyse pre-fusion structures (8). An accurate pre-fusion structure would allow for a more accurate assessment of the effect of which regions of the protein are constant and which are subject to antigenic drift. It would also allow

for the design of more effective vaccines through the design and testing of putative stabilising mutations (10).

It was believed that the adoption of GCNt stabilisation, as in the cases of RSV and PIV5, would allow this project to continue the work of Chen et al and Swanson et al (1, 8, 11, 12). The primary aim of the project was to solve the structure of the pre-fusion conformation of the F protein of NDV. This structure was then to be used to be used to map the effects of genotype variance and calculate the effect of stabilising mutations. The effects of these could then be tested in vitro, before being inserted into a vector to measure the humoral response. From there, these putative antigens could be further tested for industry applications. This all depended on solving the structure of the Pre-fusion F protein.

Chapter III:

The first output of chapter III were the AlphaFold2 3D models of the NDV F protein in the pre-fusion conformation (8). They accurately predict the structure of the head domains, but the orientation of the helices which connect the head of the protein to the membrane did not match the known structures of pre-fusion conformation F proteins. This could have been because the protein is meta-stable and is thought to pivot about this region during conformational change (13). Alternatively, it could be an artefact of only submitting a monomer to the server. The server could not process the trimer. It is interesting to note that the models presented were in the pre-fusion conformation. Perhaps if the settings were altered to produce more models, one might be the post fusion conformation. These models have been quite useful through the project in giving an idea of what to look for, in terms of F protein structure.

The SAXs data also provides some insight into the purified samples of F protein at pH8 it is globular, like a pre-fusion F protein and it has the same scattering properties as pre-fusion PIV5 F protein. This suggests that the recombinant protein in use is representative of the other class I fusion proteins. It also showed that the conformation of the F protein changed at pH 5, becoming less globular and demonstrated the same scattering properties as the post-fusion NDV F protein. This is in-keeping with the conformational change of NDV F protein being pH-dependant (14).

The SAXs data for these models also served to provide limited validation for the models generated by Alpha-Fold and vice-versa. When the Alpha-Fold models were used to generate SAXs traces, and this was compared with the data generated from the two constructs at pH8, there was a lot of similarity.

Chapter III confirms that growing crystals of F protein in the pre-fusion conformation is very difficult. A list of conditions that failed to produce crystals, of F protein, is presented in Chapter III. Both Chen and Swanson's publications made reference to their own failed attempts to produce crystals of pre-fusion F protein, but the successes of PIV5 and RSV suggested that it was possible (1, 8, 10-12). Glycosylation of the F protein was one of the factors investigated. While it was suggested that approximately 35 kDa of each monomer was glycan, radical deglycosylation of the F protein did not result in the formation of F protein crystals. Each monomer of F is thought to have 4 N-Linked glycosylation sites (see Appendix B), allowing 2kDa per glycosylation site, suggesting 24kDa of glycan per trimer. The difference could be due to variations in the gels and how they are read. Perhaps the removal of the glycans reduced the stability of the F trimer, causing it to disassemble.

During the course of the project crystals were grown from one protein preparation. This could not be repeated. Low resolution data collected from these crystals and a relatively high number of asymmetric units made structural determination difficult. Use of Pre-fusion PIV5, RSV F, and post-fusion NDV F as models for homologous recombination was not successful. Using different numbers of trimers and space group enantiomers also failed to give solutions that were convincing or maps that could be interpreted. In 2006 Yin et al. reported growing crystals of PIV5 F protein stabilised with the GCNt domain but being un-able to solve them using NDV or PIV3 models. Isomorphous Replacement and three-fold non-crystallographic symmetry were used to determine the structure instead (15).

As X-ray crystallography had failed to produce a high-resolution structure of pre-fusion NDV F, cryo-EM was then attempted.

X ray crystallography could be used successfully if the nanobodies produced by Professor Nair's collaborators (see chapter V) could be successfully expressed and these bind to the F protein. These might then improve the likelihood of crystal growth. More information from these is not available.

Chapter IV:

The first major finding from the cryo-EM screening of NDV F protein is that the F-protein Trimer is very sensitive to something in the process of grid preparation. Perhaps it was the temperature change of plunge freezing grids in liquid ethane. The attempts to solve the structure of the protein floundered on the inability to keep the protein trimeric during the preparation of the grids. The tendency of the fusion protein to disassemble under cryo-EM conditions mirrored the disassembly of the SARS-CoV-2 Spike protein but less so in the Nipah F glycoprotein, in work carried out by colleagues in the department (16). In the case of the latter, perhaps it was the opportunity to re-work the construct at the genetic level, introducing some stabilising mutations, and re-validate expression that made the difference.

The work outlined in Chapter IV was able to highlight some of the ways in which the F protein could be made more thermally resilient. The experiments using the UNCLE allowed for the identification of compounds that could help stabilise the F protein for cryo-EM work. Pegylation was the most promising of the leads; micrographs taken using the microscope closely resembled the grids seen with CDV (17). This suggested that if these could be optimised, it may be possible to solve the structure using cryo-EM. Unfortunately, due to time limitations, the preparation of pegylated sample could not be optimised to its full potential.

Additionally, it was learnt that cross-linking the F protein trimer with glutaraldehyde was not sufficiently stabilised to remain trimeric while being cryogenically cooled. Also, while several other compounds were identified as being thermo-protective, they also failed to prevent trimer disassembly. However these additives, especially NDSB195 and pegylation reduced the aggregation of the F protein at higher concentrations. This would be of prime importance if more time was allocated on the CHAMELEON at Diamond Light Source (18), where it was noted that the protein was aggregating at 4mg/ml in the nozzle of the protein gun.

Chapter V:

Following the work carried out at the Pirbright Institute, outlined in chapter V, much would have to be done to move this aspect of the project forward. It was premature to start cloning the recombinant F

protein gene into potential vectors before knowing it was stabilised. This was an artefact of the pandemic caused by restrictions on the number of staff working at each site.

Inserting the gene from this viral protein into MDV was not trivial. It was successfully cloned into the donor vector, but the gene could not be transferred from the plasmid to the viral DNA. DNA sequencing confirmed the F protein had been inserted correctly and expression tests confirmed the expression of the F protein from cells transfected with the vector. Perhaps the sequences of the plasmid that recognise the CRISPR-cas-9 machinery were compromised. Perhaps the plasmid containing the CRISPR machinery was also defective, or even the guide RNA. Each would have to be rigorously tested. The immunisation experiment also showed that the F protein, with adjuvant, does not elicit a humoral response when chickens are vaccinated. The freezing and thawing of the purified F protein could have caused it to degrade.

The nanobodies could have been the factor that allowed for the successful structural determination of the F protein in the pre-fusion conformation. But frustratingly in the format they were, they were of limited use to the project. The reasons why they did not appear to bind to the F protein after being cloned into a vector for large scale expression are not known. To speculate:

- the freezing of the protein could have caused the protein to degrade.
- The samples could have been subjected to freeze-thaw or other mishandling.
- The F protein sent could itself have been monomeric as it was purified on a size exclusion column that could not distinguish between monomeric and trimeric protein, or in the post-fusion conformation.
- The cloning process introduced a mutation that fundamentally altered the nanobody.
- The position of the c-terminal 6-His tag could block binding to the F protein (19).

Future work and lines of inquiry:

First, the nanobodies would be re cloned into a different vector or have the His-tag replaced with another tag for purification: possibly a strep tag or a biotin tag. The initial choice of a His-tag was

expedient but made other techniques more difficult. From there, these re-cloned nanobodies would have to be re-characterised and binding assays with the F protein carried out.

If these re-cloned nanobodies bound the F protein, the shift away from using a His-Tag would allow for the carrying out of Surface Plasmon Resonance (SPR) to gain a better understanding of the relative affinity and stoichiometry of each of the nanobody constructs. If this was achieved, the nanobodies could be complexed with the F protein for use in both x-ray crystallography and cryo-EM.

The sensitivity of the F protein to cryo-cooling and tendency of the trimer to break down during purification would also have to be addressed. The use of various stabilising mutations that could more effectively lock the protein in the pre-fusion conformation. Other Class I fusion proteins (RSV and SARS-CoV-2) have already been subject to this work (10, 20). Given the relative structural homology of the Class I fusion proteins, it would be possible to map such mutations across to the NDV F protein. The AlphaFold2 model of NDV in the pre-fusion conformation could be useful in identifying the corresponding domains of NDV. AlphaFold could be used to model the effects of the stabilising mutations, in order to ensure that the mutations could be mapped over from one species to another, without altering the structure drastically. These stabilising mutations were only reported recently. Time constraints of the project meant that cloning multiple mutant constructs was not pursued.

If such modifications successfully stabilised the construct, they could be used for X-ray crystallography to increase the proportion of trimeric F protein expressed and improve sample homogeneity. Furthermore it might help prevent the detrimerisation of the protein during the cryo-cooling of F protein on grids for cryo-EM. The stabilisation by point mutation of the F protein could also lock the F protein in the pre-fusion conformation, which could reduce the mixed population of pre-and post-fusion conformation F seen in Pegylated F protein during cryo-EM.

With a high-resolution structure, it would be possible to start investigating the properties of, other potentially stabilising mutations which may further stabilise the pre-fusion conformation without altering the immunogenic or structural properties of the protein. Such mutations include the replacement of the cleavage site, preventing the conformational change, introducing di-sulphide bonds and disrupting the

helices that form during the transition from pre-Fusion to post-Fusion F. Such work has already been carried out on RSV F, whose pre-fusion structure has already been solved (10, 20, 21).

With a known structure of the pre-fusion conformation, the effect of the mutations that make certain strains of NDV more pathogenic could be studied and mapped to the structure (22). While this could be carried out using AlphaFold, it may not hold the credibility of a model based on crystallography or EM. Part of the optimisation of an antigen would be investigating antibodies drawn from chickens vaccinated against NDV and also those infected with more pathogenic strains. Neutralisation assays would identify which antibodies are neutralising and which are cross reactive. These could then be used for epitope mapping to identify immunogenic regions of the protein that are common to multiple key genotypes.

Once future antigen development has been completed, insertion into a vector, for testing would be possible. It would require resources to be allocated for a trial in chickens, assessing cell-based immunity and humoral responses. Antibodies from this could also be used for epitope mapping. A challenge study would also be necessary: this would have to take place under SAPO-4 conditions (23).

<u>In summary:</u>

This project aimed to solve the structure of the NDV F protein in the pre-fusion conformation. This was not achieved but factors that could be used in future work have been identified, especially that pegylation improved the stability of the protein. The data collected during cryo-EM with pegylated protein could form a starting point to optimise future preparations. Additionally, the nanobodies generated could be re-cloned into a vector with a different tag to allow for more pull-down experiments, ELISAs, and SPR. The models of the F protein generated by AlphaFold are also useful in the short term for designing mutations which could stabilise the F protein for structural studies, especially as the SAXs data offers limited validation. NDV remains a threat to poultry and will probably remain so for the foreseeable future.

Appendix A: Bibliography

Chapter I:

1. Ganar K, Das M, Sinha S, Kumar S. Newcastle disease virus: Current status and our understanding. Virus Research. 2014;184:71–81
2. Dimitrov KM, Abolnik C, Afonso CL, Albina E, Bahl J, Berg M, et al. Updated unified phylogenetic classification system and revised nomenclature for newcastle disease virus. Infection, Genetics and Evolution. 2019;74:103917.
3. Cattoli G, Susta L, Terregino C, Brown C. Newcastle disease. Journal of Veterinary Diagnostic Investigation. 2011;23(4):637–56. doi:10.1177/1040638711407887
4. Muhammad Munir. Mononegaviruses of Veterinary Importance, Volume 2. 2nd ed.: CABI; 2017.
5. [Internet]. [cited 2023 Apr 30]. Available from: https://www.facebook.com/JPSFarming/
6. Thomas NJ, Hunter DB, Atkinson CT. Infectious diseases of wild birds. Ames, IA: Blackwell Pub.; 2007.
7. Caimbeul, Seonaidh, et al. Orain Ghaidhlig / le Seonaidh Caimbeul (Seonaidh Mac Dhomnaill 'ic Iain Bhàin) ; air an toirt sìos le Iain Mac Aonghuis; air an deasachadh le Iain Latharna Caimbeul (fear-deasachaidh "Duain Ghàidhealach mu Bhliadhna Thearlaich"). I.B. Mac Aoidh agus a Chuideachd, 1936.
8. MacPherson LW [Internet]. The University of Edinburgh; 1970 [cited 2023 Apr 30]. Available from: http://hdl.handle.net/1842/20660
9. Soñora M, Moreno P, Echeverría N, Fischer S, Comas V, Fajardo A, et al. An evolutionary insight into Newcastle disease viruses isolated in Antarctica. Archives of Virology. 2015;160(8):1893–900. doi:10.1007/s00705-015-2434-y
10. Doyle TM. A hitherto unrecognized disease of fowls due to a filter-passing virus. J. Com. Pathol. Ther.. 1927;40:144-69.
11. Kraneveld F. Over een in Ned-Indie heerschende Ziete under het pluimves. Ned. Indisch Bl. Diergeneesk. 1926;38:448-50.
12. Kanno T, Ochi Y, Hashimoto K. Neue Geflugelscuche in Korea. Dtsch. Tier-arzhl. Wshr. 1929.
13. Asperlin FD. Newcastle disease in ducks and geese. The Veterinary Record. 1947 Nov;59(45):621-624
14. Zargar SL, Pomeroy BS. The effects of commercial living Newcastle disease virus vaccines. American Journal of Veterinary Research. 1950;11:272-7.
15. Dimitrov KM, Ramey AM, Qiu X, Bahl J, Afonso CL. Temporal, geographic, and host distribution of avian paramyxovirus 1 (Newcastle disease virus). Infection, genetics and evolution. 2016 Apr 1;39:22-34.
16. Hitchner SB. Guest editorial: Serendipity in science: Discovery of the B-1 Strain of Newcastle disease virus. Avian Diseases. 1975 Apr 1;19(2):215-23.
17. Dimitrov KM, Afonso CL, Yu Q, Miller PJ. Newcastle disease vaccines—a solved problem or a continuous challenge? Veterinary Microbiology. 2017;206:126–36.
18. Gorbalenya AE, Lauber C. Phylogeny of viruses ☆. Reference Module in Biomedical Sciences. 2017;

19. Rima B, Balkema-Buschmann A, Dundon WG, Duprex P, Easton A, Fouchier R, et al. ICTV Virus Taxonomy Profile: Paramyxoviridae. Journal of General Virology. 2019;100(12):1593–4.
20. Khadzhiev G. (1987). Tipizirane na izolati na Niukiasŭlskiia virus po srednoto vreme za umirane na pileta v kloachen test [Typing of Newcastle virus isolates by mean death time of chickens in the cloacal test]. Veterinarno-meditsinski nauki, 24(5), 16–26.
21. Diseases of Poultry (England) Order, 2003, (United Kingdom) No1087
22. Newcastle disease - woah - world organisation for animal health [Internet]. 2023 [cited 2023 May 30]. Available from: https://www.woah.org/en/disease/newcastle-disease/
23. Halvorson DA. Prevention and management of avian influenza outbreaks: experiences from the United States of America. Revue scientifique et technique (International Office of Epizootics). 2009 Apr 1;28(1):359-69.
24. Goldhaft TM. Guest editorial: Historical note on the origin of the Lasota strain of Newcastle Disease Virus. Avian Diseases. 1980;24(2):297.
25. Dimitrov KM, Ferreira HL, Pantin-Jackwood MJ, Taylor TL, Goraichuk IV, Crossley BM, et al. Pathogenicity and transmission of virulent Newcastle Disease Virus from the 2018–2019 California outbreak and related viruses in young and adult chickens. Virology. 2019;531:203–18.
26. Lomniczi B, Wehmann E, Herczeg J, Ballagi-Pordány A, Kaleta EF, Werner O, et al. Newcastle disease outbreaks in recent years in Western Europe were caused by an old (VI) and a novel genotype (VII). Archives of Virology. 1998;143(1):49–64.
27. Steensels M, Van Borm S, Mertens I, Houdart P, Rauw F, Roupie V, et al. Molecular and virological characterization of the first poultry outbreaks of genotype VII.2 velogenic avian orthoavulavirus type 1 (NDV) in north-west Europe, Benelux, 2018. Transboundary and Emerging Diseases. 2020;68(4):2147–60.
28. Miller PJ, Haddas R, Simanov L, Lublin A, Rehmani SF, Wajid A, Bibi T, Khan TA, Yaqub T, Setiyaningsih S, Afonso CL. Identification of new sub-genotypes of virulent Newcastle disease virus with potential panzootic features. Infection, genetics and evolution. 2015 Jan 1;29:216-29.
29. Global Poultry Industry and trends [Internet]. 2021 [cited 2022 Apr 30]. Available from: https://www.feedandadditive.com/global-poultry-industry-and-trends/#:~:text=The%20poultry%20sector%2C%20which%20has%20a%20market%20value%20of%20%24%20310.7,with%20a%20CAGR%20of%207%25.
30. Global animal farming, meat production and meat consumption [Internet]. 2023 [cited 2023 April 30]. Available from: https://www.destatis.de/EN/Themes/Countries-Regions/International-Statistics/Data Topic/AgricultureForestryFisheries/livestock_meat.html
31. Uzundumlu AS, Dilli M. Estimating chicken meat productions of leader countries for 2019-2025 years. Ciência Rural. 2023;53(2). doi:10.1590/0103-8478cr20210477
32. Erian I, Phillips CJ. Public understanding and attitudes towards meat chicken production and relations to consumption. Animals. 2017 Mar 9;7(3):20.
33. Kalhor T, Rajabipour A, Akram A, Sharifi M. Environmental impact assessment of chicken meat production using life cycle assessment. Information processing in agriculture. 2016 Dec 1;3(4):262-71.
34. UK Egg Industry Data [Internet]. [cited 2023 Mar 30]. Available from: https://www.egginfo.co.uk/egg-facts-and-figures/industry-information/data
35. Poultry meat Quality Guide, Department for Environment, Food and Rural Affairs, 2011, PB13457
36. Berg LR, Bearse GE, Hamilton CM. The effect of Newcastle disease on egg production and egg quality. Poultry Science. 1947 Nov 1;26(6):614-22.
37. Notifiable Avian Disease Control Strategy for Great Britain, Animal Health Policy and Implementation, Defra, Sept 2019

38. Agency APHA. Poultry (including game birds): Registration Rules and Forms [Internet]. GOV.UK; 2022 [cited 2023 Apr 30]. Available from: https://www.gov.uk/government/publications/poultry-including-game-birds-registration-rules-and-forms
39. Miller PJ, Afonso CL, Spackman E, Scott MA, Pedersen JC, Senne DA, Brown JD, Fuller CM, Uhart MM, Karesh WB, Brown IH. Evidence for a new avian paramyxovirus serotype 10 detected in rockhopper penguins from the Falkland Islands. Journal of virology. 2010 Nov 1;84(21):11496-504.
40. Birhanu MY, Osei-Amponsah R, Yeboah Obese F, Dessie T. Smallholder poultry production in the context of increasing global food prices: Roles in Poverty Reduction and Food Security. Animal Frontiers. 2023;13(1):17–25.
41. Hu Z, He X, Deng J, Hu J, Liu X. Current situation and future direction of Newcastle disease vaccines. Veterinary Research. 2022 Dec;53(1):1-3.
42. Walker JW, Heron BR, Mixson MA. Exotic Newcastle disease eradication program in the United States. Avian Diseases. 1973 Jul 1:486-503.
43. Alexander DJ. Newcastle disease in the European Union 2000 to 2009. Avian pathology. 2011 Dec 1;40(6):547-58.
44. Gallili GE, Ben-Nathan D. Newcastle disease vaccines. Biotechnology advances. 1998 Mar 1;16(2):343-66.
45. Campbell ZA, Thumbi SM, Marsh TL, Quinlan MB, Shirima GM, Palmer GH. Why isn't everyone using the thermotolerant vaccine? preferences for Newcastle disease vaccines by chicken-owning households in Tanzania. PLOS ONE. 2019;14(8).
46. Hipra. Live vaccine against Newcastle disease, La Sota strain, in oral freeze-dried tablet [Internet]. Hipraviar. 2018. https://www.hipra.com/portal/en/hipra/animalhealth/products/detail/hipraviar-s
47. Okwor EC, Eze DC, Umeh M. Newcastle disease virus shedding among healthy commercial chickens and its epidemiological importance.
48. Innovax-nd-ilt [Internet]. 2022 [cited 2023 Apr 30]. Available from: https://www.ema.europa.eu/en/medicines/veterinary/EPAR/innovax-nd-ilt
49. Yang CY, Shieh HK, Lin YL, Chang PC. Newcastle disease virus isolated from recent outbreaks in Taiwan phylogenetically related to viruses (genotype VII) from recent outbreaks in western Europe. Avian diseases. 1999 Jan 1:125-30.
50. Roohani K, Tan SW, Yeap SK, Ideris A, Bejo MH, Omar AR. Characterisation of genotype VII Newcastle disease virus (NDV) isolated from NDV vaccinated chickens, and the efficacy of LaSota and recombinant genotype VII vaccines against challenge with velogenic NDV. Journal of veterinary science. 2015 Dec 1;16(4):447-57.
51. APHA National Surveillance Network [Internet]. [cited 2023 Mar 30]. Available from: http://apha.defra.gov.uk/vet-gateway/surveillance/diagnostic/national-network.htm
52. Cox JE, Cheng TL. Egg-based vaccines. Pediatrics in Review. 2006;27(3):118–9.
53. Matthews JT. Egg-based production of influenza vaccine: 30 years of commercial experience. BRIDGE-WASHINGTON-NATIONAL ACADEMY OF ENGINEERING-. 2006 Apr 15;36(3):17.
54. Cosgrove J. To stop a virus, California has euthanized more than 1.2 million birds. is it reckless or necessary? [Internet]. Los Angeles Times; 2019 [cited 2023 Feb 30]. Available from: https://www.latimes.com/local/lanow/la-me-ln-virulent-newcastle-disease-outbreak-in-southern-california-20190607-story.html
55. Department for Environment F& RA. Newcastle disease in Belgium [Internet]. GOV.UK; 2018 [cited 2023 Jan 30]. Available from: https://www.gov.uk/government/publications/newcastle-disease-in-belgium
56. Steensels M, Van Borm S, Mertens I, Houdart P, Rauw F, Roupie V, Snoeck CJ, Bourg M, Losch S, Beerens N, van den Berg T. Molecular and virological characterization of the first

poultry outbreaks of Genotype VII. 2 velogenic avian orthoavulavirus type 1 (NDV) in North-West Europe, BeNeLux, 2018. Transboundary and Emerging Diseases. 2021 Jul;68(4):2147-60.
57. Takakuwa H, ITO T, Takada A, Okazaki K, Kida H. Potentially virulent Newcastle disease viruses are maintained in migratory waterfowl populations. Japanese Journal of Veterinary Research. 1998 Feb 27;45(4):207-15.
58. Irvine RM, Aldous EW, Manvell RJ, Cox WJ, Ceeraz V, Fuller CM, Alexander DJ, Brown IH, Wood AM, Milne JC, Wilson M. Outbreak of New castle disease due to pigeon paramyxovirus type 1 in grey partridges (Perdix perdix) in Scotland in October 2006. Veterinary Record. 2009 Oct;165(18):531-5.
59. Alexander DJ. Newcastle disease and other avian paramyxoviruses. Revue Scientifique et Technique-Office International des Epizooties. 2000 Aug 1;19(2):443-55.
60. Nath B, Sharma K, Ahire K, Goyal A, Kumar S. Structure analysis of the nucleoprotein of Newcastle disease virus: An insight towards its multimeric form in solution. International journal of biological macromolecules. 2020 May 15;151:402-11.
61. Battisti AJ, Meng G, Winkler DC, McGinnes LW, Plevka P, Steven AC, Morrison TG, Rossmann MG. Structure and assembly of a paramyxovirus matrix protein. Proceedings of the National Academy of Sciences. 2012 Aug 28;109(35):13996-4000.
62. Cong J, Feng X, Kang H, Fu W, Wang L, Wang C, Li X, Chen Y, Rao Z. Structure of the Newcastle Disease Virus L protein in complex with tetrameric phosphoprotein. Nature Communications. 2023 Mar 10;14(1):1324.
63. Gravel KA, Morrison TG. Interacting domains of the HN and F proteins of Newcastle disease virus. Journal of Virology. 2003 Oct 15;77(20):11040-9.
64. Jardetzky TS, Lamb RA. Activation of paramyxovirus membrane fusion and virus entry. Current opinion in virology. 2014 Apr 1;5:24-33.
65. Sánchez-Felipe L, Villar E, Muñoz-Barroso I. Entry of Newcastle Disease Virus into the host cell: role of acidic pH and endocytosis. Biochimica et Biophysica Acta (BBA)-Biomembranes. 2014 Jan 1;1838(1):300-9.
66. Park MS, Shaw ML, Munoz-Jordan J, Cros JF, Nakaya T, Bouvier N, Palese P, García-Sastre A, Basler CF. Newcastle disease virus (NDV)-based assay demonstrates interferon-antagonist activity for the NDV V protein and the Nipah virus V, W, and C proteins. Journal of virology. 2003 Jan 15;77(2):1501-11.
67. Qiu X, Fu Q, Meng C, Yu S, Zhan Y, Dong L, Song C, Sun Y, Tan L, Hu S, Wang X. Newcastle disease virus V protein targets phosphorylated STAT1 to block IFN-I signaling. PloS one. 2016 Feb 9;11(2):e0148560.
68. Nagai Y, Hamaguchi M, Toyoda T. Molecular biology of Newcastle disease virus. Prog Vet Microbiol Immunol. 1989 Jan 1;5:16-64.
69. Lamb RA, Kolakofsky D (1996) Paramyxoviridae: the viruses and their replication. In: Fields BN, Knipe DM, Howley PM (eds) Virology. Lipincott-Raven, Philadelphia, pp 1 177–1 203
70. Kim SH, Samal SK. Newcastle disease virus as a vaccine vector for development of human and veterinary vaccines. Viruses. 2016 Jul 4;8(7):183.
71. Iorio RM, Field GM, Sauvron JM, Mirza AM, Deng R, Mahon PJ, Langedijk JP. Structural and functional relationship between the receptor recognition and neuraminidase activities of the Newcastle disease virus hemagglutinin-neuraminidase protein: receptor recognition is dependent on neuraminidase activity. Journal of virology. 2001 Feb 15;75(4):1918-27.
72. Gravel KA, Morrison TG. Interacting domains of the HN and F proteins of Newcastle disease virus. Journal of Virology. 2003 Oct 15;77(20):11040-9.
73. Mahon PJ, Mirza AM, Iorio RM. Role of the two sialic acid binding sites on the newcastle disease virus HN protein in triggering the interaction with the F protein required for the promotion of fusion. Journal of virology. 2011 Nov 15;85(22):12079-82.

74. White JM, Delos SE, Brecher M, Schornberg K. Structures and mechanisms of viral membrane fusion proteins: multiple variations on a common theme. Critical reviews in biochemistry and molecular biology. 2008 Jan 1;43(3):189-219.
75. Morrison TG. Structure and function of a paramyxovirus fusion protein. Biochimica et Biophysica Acta (BBA)-Biomembranes. 2003 Jul 11;1614(1):73-84.
76. Pager CT, Craft Jr WW, Patch J, Dutch RE. A mature and fusogenic form of the Nipah virus fusion protein requires proteolytic processing by cathepsin L. Virology. 2006 Mar 15;346(2):251-7.
77. Russell CJ, Hu M, Okda FA. Influenza hemagglutinin protein stability, activation, and pandemic risk. Trends in microbiology. 2018 Oct 1;26(10):841-53.
78. Connolly SA, Leser GP, Yin HS, Jardetzky TS, Lamb RA. Refolding of a paramyxovirus F protein from prefusion to postfusion conformations observed by liposome binding and electron microscopy. Proceedings of the National Academy of Sciences. 2006 Nov 21;103(47):17903-8.
79. Modis Y. Class II fusion proteins. Viral Entry into Host Cells. 2013:150-66.
80. Rey FA. Dengue virus envelope glycoprotein structure: new insight into its interactions during viral entry. Proceedings of the National Academy of Sciences. 2003 Jun 10;100(12):6899-901.
81. Lindenbach BD, Rice CM. Molecular biology of flaviviruses. Advances in virus research. 2003 Jan 1;59:23-62.
82. Reske A, Pollara G, Krummenacher C, Chain BM, Katz DR. Understanding HSV-1 entry glycoproteins. Reviews in medical virology. 2007 May;17(3):205-15.
83. Sun X, Belouzard S, Whittaker GR. Molecular architecture of the bipartite fusion loops of vesicular stomatitis virus glycoprotein G, a class III viral fusion protein. Journal of Biological Chemistry. 2008 Mar 7;283(10):6418-27.
84. Backovic M, Leser GP, Lamb RA, Longnecker R, Jardetzky TS. Characterization of EBV gB indicates properties of both class I and class II viral fusion proteins. Virology. 2007 Nov 10;368(1):102-13.
85. Backovic M, Jardetzky TS. Class III viral membrane fusion proteins. Current opinion in structural biology. 2009 Apr 1;19(2):189-96.
86. Boursnell M, Green P, et al. Insertion of the fusion gene from Newcastle disease virus into a non-essential region in the terminal repeats of fowlpox virus and demonstration of protective immunity induced by the recombinant, Journal of General Virology 1990;71 (Pt 3):621-8
87. Samson AC, Chambers P, Dickinson JH. Location of post-translational cleavage events within F and HN glycoproteins of Newcastle disease virus. Journal of General Virology. 1980;47(1):19–27.
88. Melanson VR, Iorio RM. Addition of N-glycans in the stalk of the Newcastle disease virus HN protein blocks its interaction with the F protein and prevents fusion. Journal of virology. 2006 Jan 15;80(2):623-33.
89. Aguilar HC, Matreyek KA, Filone CM, Hashimi ST, Levroney EL, Negrete OA, Bertolotti-Ciarlet A, Choi DY, McHardy I, Fulcher JA, Su SV. N-glycans on Nipah virus fusion protein protect against neutralization but reduce membrane fusion and viral entry. Journal of virology. 2006 May 15;80(10):4878-89.
90. Plemper RK, Hammond AL, Gerlier D, Fielding AK, Cattaneo R. Strength of envelope protein interaction modulates cytopathicity of measles virus. Journal of virology. 2002 May 15;76(10):5051-61.
91. Cullen LM, Boukhvalova MS, Blanco JC, Morrison TG. Comparisons of antibody populations in different pre-fusion f VLP-immunized cotton rat dams and their offspring. Vaccines. 2020 Mar 18;8(1):133.
92. Plattet P, Alves L, Herren M, Aguilar HC. Measles virus fusion protein: structure, function and inhibition. Viruses. 2016 Apr 21;8(4):112.
93. Chang LA, Phung E, Crank MC, Morabito KM, Villafana T, Dubovsky F, Falloon J, Esser MT, Lin BC, Chen GL, Graham BS. A prefusion-stabilized RSV F subunit vaccine elicits B cell

responses with greater breadth and potency than a postfusion F vaccine. Science Translational Medicine. 2022 Dec 21;14(676):eade0424.
94. Bolt G, Pedersen LØ, Birkeslund HH. Cleavage of the respiratory syncytial virus fusion protein is required for its surface expression: role of furin. Virus research. 2000 Jun 1;68(1):25-33.
95. Sergel-Germano T, McQUAIN CA, Morrison T. Mutations in the fusion peptide and heptad repeat regions of the Newcastle disease virus fusion protein block fusion. Journal of virology. 1994 Nov;68(11):7654-8.
96. Panda A, Huang Z, Elankumaran S, Rockemann DD, Samal SK. Role of fusion protein cleavage site in the virulence of Newcastle disease virus. Microbial pathogenesis. 2004 Jan 1;36(1):1-0.
97. Hu S, Ma H, Wu Y, Liu W, Wang X, Liu Y, Liu X. A vaccine candidate of attenuated genotype VII Newcastle disease virus generated by reverse genetics. Vaccine. 2009 Feb 5;27(6):904-10.
98. Fan W, Xu Y, Zhang P, Chen P, Zhu Y, Cheng Z, Zhao X, Liu Y, Liu J. Analysis of molecular evolution of nucleocapsid protein in Newcastle disease virus. Oncotarget. 2017 Nov 11;
Wawina-Bokalanga T, Martí-Carreras J, Vanmechelen B, Bloemen M, Wollants E, Laenen L,
99. Cuypers L, Beuselinck K, Lagrou K, André E, Van Ranst M. Genetic diversity and evolution of SARS-CoV-2 in Belgium during the first wave outbreak. bioRxiv. 2021 Jun 29:2021-06.8(57):97127.
100. Parvin JD, Moscona A, Pan WT, Leider JM, Palese P. Measurement of the mutation rates of animal viruses: influenza A virus and poliovirus type 1. Journal of virology. 1986 Aug;59(2):377-83.
101. Glick B. Bursa of fabricius. Avian biology. 1983 Jan 1;7:443-500.
102. Glick B, Chang TS, Jaap RG. The bursa of Fabricius and antibody production. Poultry Science. 1956 Jan 1;35(1):224-5.
103. da Silva WD, Tambourgi DV. IgY: a promising antibody for use in immunodiagnostic and in immunotherapy. Veterinary immunology and immunopathology. 2010 Jun 15;135(3-4):173-80.
104. Zhang X, Calvert RA, Sutton BJ, Doré KA. IgY: a key isotype in antibody evolution. Biological Reviews. 2017 Nov;92(4):2144-56.
105. Sesarman A, Mihai S, Chiriac MT, Olaru F, Sitaru AG, Thurman JM, Zillikens D, Sitaru C. Binding of avian IgY to type VII collagen does not activate complement and leucocytes and fails to induce subepidermal blistering in mice. British Journal of Dermatology. 2008 Mar 1;158(3):463-71.
106. Pereira EP, Van Tilburg MF, Florean EO, Guedes MI. Egg yolk antibodies (IgY) and their applications in human and veterinary health: A review. International immunopharmacology. 2019 Aug 1;73:293-303.
107. Yasuda M, Taura Y, Yokomizo Y, Ekino S. A comparative study of germinal center: fowls and mammals. Comparative immunology, microbiology and infectious diseases. 1998 Jul 1;21(3):
108. Six A, Rast JP, McCormack WT, Dunon D, Courtois D, Li Y, Chen CL, Cooper MD. Characterization of avian T-cell receptor γ genes. Proceedings of the National Academy of Sciences. 1996 Dec 24;93(26):
109. Genovese KJ, He H, Swaggerty CL, Kogut MH. The avian heterophil. Developmental & Comparative Immunology. 2013 Nov 1;41(3):334-40.
110. [Internet]. [cited 2023 Jan 1]. Available from: https://www.mybiosource.com/learn/polyclonal-antibodies
111. Kim S-H, Wanasen N, Paldurai A, Xiao S, Collins PL, Samal SK. Newcastle disease virus fusion protein is the major contributor to protective immunity of genotype-matched vaccine. PLoS ONE. 2013;8(8).

112. Meulemans G, Letellier C, Gonze M, Carlier MC, Burny A. Newcastle disease virus f glycoprotein expressed from a recombinant vaccinia virus vector protects chickens against live-virus challenge. Avian Pathology. 1988 Jan 1;17(4):821-7.
113. Xu T, Xiong T, Xie W, Wu J, Liu X, Li G, Lv Y, Li L, Yang Z, Wang H, Liu D. Construction and Evaluation of the Immunogenicity and Protective Efficacy of Recombinant Replication-Deficient Human Adenovirus-5 Expressing Genotype VII Newcastle Disease Virus F Protein and Infectious Bursal Disease Virus VP2 Protein. Vaccines. 2023 May 31;11(6):1051.
114. Zhang X, Rosenthal PB, Formanowski F, Fitz W, Wong CH, Meier-Ewert H, Skehel JJ, Wiley DC. X-ray crystallographic determination of the structure of the influenza C virus haemagglutinin-esterase-fusion glycoprotein. Acta Crystallographica Section D: Biological Crystallography. 1999 May 1;55(5):945-61.
115. Wu KP, Wu CW, Tsao YP, Kuo TW, Lou YC, Lin CW, Wu SC, Cheng JW. Structural basis of a flavivirus recognized by its neutralizing antibody: solution structure of the domain III of the Japanese encephalitis virus envelope protein. Journal of Biological Chemistry. 2003 Nov 14;278(46):46007-13.
116. Chen L, Gorman JJ, McKimm-Breschkin J, Lawrence LJ, Tulloch PA, Smith BJ, et al. The structure of the fusion glycoprotein of Newcastle Disease Virus suggests a novel paradigm for the molecular mechanism of membrane fusion. Structure. 2001;9(3):255–66.
117. Swanson K, Wen X, Leser GP, Paterson RG, Lamb RA, Jardetzky TS. Structure of the Newcastle disease virus F protein in the post-fusion conformation. Virology. 2010 Jul 5;402(2):372-9.
118. Welch BD, Liu Y, Kors CA, Leser GP, Jardetzky TS, Lamb RA. Structure of the cleavage-activated prefusion form of the parainfluenza virus 5 fusion protein. Proceedings of the National Academy of Sciences. 2012 Oct 9;109(41):16672-7.
119. McLellan JS, Chen M, Leung S, Graepel KW, Du X, Yang Y, Zhou T, Baxa U, Yasuda E, Beaumont T, Kumar A. Structure of RSV fusion glycoprotein trimer bound to a prefusion-specific neutralizing antibody. Science. 2013 May 31;340(6136):1113-7.
120. Avanzato VA, Oguntuyo KY, Escalera-Zamudio M, Gutierrez B, Golden M, Kosakovsky Pond SL, et al. A structural basis for antibody-mediated neutralization of Nipah virus reveals a site of vulnerability at the fusion glycoprotein apex. Proceedings of the National Academy of Sciences. 2019;116(50):25057–67. doi:10.1073/pnas.1912503116
121. Wong JJ, Paterson RG, Lamb RA, Jardetzky TS. Structure and stabilization of the Hendra virus F glycoprotein in its prefusion form. Proceedings of the National Academy of Sciences. 2016 Jan 26;113(4):1056-61.
122. Zhang J, Xiao T, Cai Y, Chen B. Structure of SARS-CoV-2 spike protein. Current opinion in virology. 2021 Oct 1;50:173-82.
123. Yin HS, Wen X, Paterson RG, Lamb RA, Jardetzky TS. Structure of the parainfluenza virus 5 F protein in its metastable, prefusion conformation. Nature. 2006 Jan 5;439(7072):38-44.
124. Weissenhorn W, Dessen A, Calder LJ, Harrison SC, Skehel JJ, Wiley DC. Structural basis for membrane fusion by enveloped viruses. Molecular membrane biology. 1999 Jan 1;16(1):3-9.
125. Harrison SC. Mechanism of membrane fusion by viral envelope proteins. Advances in virus research. 2005 Jan 1;64:231-61.
126. Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, Zhang Q, Shi X, Wang Q, Zhang L, Wang X. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. nature. 2020 May 14;581(7807):215-20.
127. Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. Cell. 2020 Apr 16;181(2):281-92.

128. Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, Graham BS, McLellan JS. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science. 2020 Mar 13;367(6483):1260-3.
129. Crank MC, Ruckwardt TJ, Chen M, Morabito KM, Phung E, Costner PJ, et al. A proof of concept for structure-based vaccine design targeting RSV in humans. Science. 2019;365(6452):505–9.

Chapter II:

1. Wong JJ, Paterson RG, Lamb RA, Jardetzky TS. Structure and stabilization of the Hendra virus F glycoprotein in its prefusion form. Proceedings of the National Academy of Sciences. 2016 Jan 26;113(4):1056-61.
2. Yin HS, Wen X, Paterson RG, Lamb RA, Jardetzky TS. Structure of the parainfluenza virus 5 F protein in its metastable, prefusion conformation. Nature. 2006 Jan 5;439(7072):38-44.
3. Welch BD, Liu Y, Kors CA, Leser GP, Jardetzky TS, Lamb RA. Structure of the cleavage-activated prefusion form of the parainfluenza virus 5 fusion protein. Proceedings of the National Academy of Sciences. 2012 Oct 9;109(41):16672-7.
4. Owens R. Poping (plasmid #26046) [Internet]. [cited 2023 Feb 1]. Available from: https://www.addgene.org/26046/
5. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., Schwede, T. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 46, W296-W303 (2018)
6. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with alphafold. Nature. 2021;596(7873):583–9. doi:10.1038/s41586-021-03819-2
7. Elbein, A D et al. "Kifunensine, a potent inhibitor of the glycoprotein processing mannosidase I." The Journal of biological chemistry vol. 265,26 (1990): 15599-605m
8. Veronica T. Chang, Max Crispin, A. Radu Aricescu, David J. Harvey, Joanne E. Nettleship, Janet A. Fennelly, Chao Yu, Kent S. Boles, Edward J. Evans, David I. Stuart, Raymond A. Dwek, E. Yvonne Jones, Raymond J. Owens, Simon J. Davis, Glycoprotein Structural Genomics: Solving the Glycosylation Problem, Structure,
9. Volume 15, Issue 3, 2007, Pages 267-273,
10. Panjkovich A. and Svergun D.I. CHROMIXS: automatic and interactive analysis of chromatography-coupled small angle X-ray scattering data. Bioinformatics, (2017)
11. Winter G. xia2: an expert system for macromolecular crystallography data reduction. Journal of applied crystallography. 2010 Feb 1;43(1):186-90.
12. D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkóczi, V. B. Chen, T. I. Croll, B. Hintze, L.-W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams and P. D. Adams. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Cryst. (2019). D75, 861-877
13. Uncle - Protein Stability Screening Platform: Unchained Labs [Internet]. 2023 [cited 2023 Oct 1]. Available from: https://www.unchainedlabs.com/uncle/#resources

14. Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. CryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nature methods. 2017 Mar;14(3):290-6.
15. Source DL. Chameleon [Internet]. [cited 2023 Jan 1]. Available from: https://www.diamond.ac.uk/Instruments/Biological-Cryo-Imaging/eBIC/Instruments/Chameleon.html
16. Tang N, Zhang Y, Sadigh Y, Moffat K, Shen Z, Nair V, Yao Y. Generation of a triple insert live avian herpesvirus vectored vaccine using CRISPR/Cas9-based gene editing. Vaccines. 2020 Feb21;8(1):97

Chapter III:

1. Hodgkin, D. C. X Rays And The Structure Of Insulin. The British Medical Journal, 1971. 4(5785), 447–451
2. HODGKIN, D C. "The X-ray analysis of the structure of penicillin." Advancement of science 1949vol. 6,22: 85-9.
3. Greer J, Erickson JW, Baldwin JJ, Varney MD. Application of the three-dimensional structures of protein target molecules in structure-based drug design. Journal of Medicinal Chemistry. 1994;37(8):
4. D'Argenio DA, Wilson CB. A decade of vaccines: Integrating immunology and vaccinology for rational vaccine design. Immunity. 2010 Oct 29;33(4):437-40.
5. Hand E. Cheap shots. Science. 2020;367(6476):354–8.
6. Khoshouei M, Radjainia M, Baumeister W, Danev R. Cryo-EM structure of haemoglobin at 3.2 Å determined with the Volta phase plate. Nature communications. 2017 Jun 30;8(1):16099.
7. Kojić-Prodić B. A century of X-ray crystallography and 2014 international year of X-ray crystallography. Macedonian Journal of Chemistry and Chemical Engineering. 2015 Jun 2;34(1):19-32.
8. Chen L, Gorman JJ, McKimm-Breschkin J, Lawrence LJ, Tulloch PA, Smith BJ, et al. The structure of the fusion glycoprotein of Newcastle Disease Virus suggests a novel paradigm for the molecular mechanism of membrane fusion. Structure. 2001;9(3):255–66.
9. Swanson K, Wen X, Leser GP, Paterson RG, Lamb RA, Jardetzky TS. Structure of the Newcastle disease virus F protein in the post-fusion conformation. Virology. 2010 Jul 5;402(2):372-9.
10. Yin HS, Paterson RG, Wen X, Lamb RA, Jardetzky TS. Structure of the uncleaved ectodomain of the paramyxovirus (hPIV3) fusion protein. Proceedings of the National Academy of Sciences. 2005 Jun 28;102(26):9288-93.
11. Harbury PB, Zhang T, Kim PS, Alber T. A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. Science. 1993 Nov 26;262(5138):1401-7.
12. Yin HS, Wen X, Paterson RG, Lamb RA, Jardetzky TS. Structure of the parainfluenza virus 5 F protein in its metastable, prefusion conformation. Nature. 2006 Jan 5;439(7072):38-44.
13. Welch BD, Liu Y, Kors CA, Leser GP, Jardetzky TS, Lamb RA. Structure of the cleavage-activated prefusion form of the parainfluenza virus 5 fusion protein. Proceedings of the National Academy of Sciences. 2012 Oct 9;109(41):16672-7.
14. McLellan JS, Chen M, Leung S, Graepel KW, Du X, Yang Y, Zhou T, Baxa U, Yasuda E, Beaumont T, Kumar A. Structure of RSV fusion glycoprotein trimer bound to a prefusion-specific neutralizing antibody. Science. 2013 May 31;340(6136):1113-7.
15. Xu K, Chan YP, Bradel-Tretheway B, Akyol-Ataman Z, Zhu Y, Dutta S, Yan L, Feng Y, Wang LF, Skiniotis G, Lee B. Crystal structure of the pre-fusion Nipah virus fusion glycoprotein reveals a novel hexamer-of-trimers assembly. PLoS pathogens. 2015 Dec 8;11(12):e1005322.
16. Yang X, Lee J, Mahony EM, Kwong PD, Wyatt R, Sodroski J. Highly stable trimers formed by human immunodeficiency virus type 1 envelope glycoproteins fused with the trimeric motif of T4 bacteriophage fibritin. Journal of virology. 2002 May 1;76(9):4634-42.

17. Morrison TG. Structure and function of a paramyxovirus fusion protein. Biochimica et Biophysica Acta (BBA)-Biomembranes. 2003 Jul 11;1614(1):73-84.
18. Cavasotto CN, Phatak SS. Homology modeling in drug discovery: current trends and applications. Drug discovery today. 2009 Jul 1;14(13-14):676-83.
19. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., Schwede, T. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 46, W296-W303 (2018)
20. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with alphafold. Nature. 2021;596(7873):583–9. doi:10.1038/s41586-021-03819-2
21. Yang Z, Zeng X, Zhao Y, Chen R. AlphaFold2 and its applications in the fields of biology and medicine. Signal Transduction and Targeted Therapy. 2023 Mar 14;8(1):115.
22. TS Analysis : Group performance based on combined z-scores [Internet]. Protein Structure prediction centre; [cited 2023 Jan 1]. Available from: https://predictioncenter.org/casp14/zscores_final.cgi
23. Zemla A. LGA: A method for finding 3D similarities in protein structures. Nucleic Acids Research. 2003;31(13):3370–4.
24. de Leeuw O, Peeters B. Complete nucleotide sequence of Newcastle disease virus: Evidence for the existence of a new genus within the subfamily Paramyxovirinae. Journal of General Virology. 1999;80(1):131–6.
25. Marshall RD. Glycoproteins. Annual review of biochemistry. 1972 Jul;41(1):673-702.
26. Gavel Y, Heijne GV. Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering. Protein Engineering, Design and Selection. 1990 Apr 1;3(5):433-42.
27. Yamaguchi Y, Nagae M. Function and 3D Structure of the N-Glycans on Glycoproteins. International Journal of Molecular Sciences. 2012.
28. Mitra N, Sinha S, Ramya TN, Surolia A. N-linked oligosaccharides as outfitters for glycoprotein folding, form and function. Trends in biochemical sciences. 2006 Mar 1;31(3):156-63.
29. Takeuchi T, Nishiyama K, Sugiura K -i., Takahashi M, Yamada A, Kobayashi S, et al. Caenorhabditis elegans galectins LEC-6 and lec-1 recognize a chemically synthesized gal 1-4fuc disaccharide unit which is present in Protostomia glycoconjugates. Glycobiology. 2009;19(12):1503–10.
30. Vigerust DJ, Shepherd VL. Virus glycosylation: role in virulence and immune interactions. Trends in microbiology. 2007 May 1;15(5):211-8.
31. Reitter JN, Means RE, Desrosiers RC. A role for carbohydrates in immune evasion in AIDS. Nature Medicine. 1998;4(6):679–84.
32. Kwong PD, Wyatt R, Desjardins E, Robinson J, Culp JS, Hellmig BD, Sweet RW, Sodroski J, Hendrickson WA. Probability analysis of variational crystallization and its application to gp120, the exterior envelope glycoprotein of type 1 human immunodeficiency virus (HIV-1). Journal of Biological Chemistry. 1999 Feb 12;274(7):4115-23.
33. Chang VT, Crispin M, Aricescu AR, Harvey DJ, Nettleship JE, Fennelly JA, Yu C, Boles KS, Evans EJ, Stuart DI, Dwek RA. Glycoprotein structural genomics: solving the glycosylation problem. Structure. 2007 Mar 1;15(3):267-73.
34. Elbein AD, Tropea JE, Mitchell M, Kaushal GP. Kifunensine, a potent inhibitor of the glycoprotein processing mannosidase I. Journal of Biological Chemistry. 1990 Sep 15;265(26):15599-605.
35. Thomas P, Smart TG. HEK293 cell line: a vehicle for the expression of recombinant proteins. Journal of pharmacological and toxicological methods. 2005 May 1;51(3):187-200.
36. Skou S, Gillilan RE, Ando N. Synchrotron-based small-angle X-ray scattering of proteins in solution. Nature protocols. 2014 Jul;9(7):1727-39.

37. Cowieson NP, Edwards-Gayle CJ, Inoue K, Khunti NS, Doutch J, Williams E, Daniels S, Preece G, Krumpa NA, Sutter JP, Tully MD. Beamline B21: high-throughput small-angle X-ray scattering at Diamond Light Source. Journal of Synchrotron Radiation. 2020 Sep 1;27(5):1438-46.
38. Bernado P, Svergun DI. Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering. Molecular biosystems. 2012;8(1):151-67.
39. Source DL [Internet]. [cited 2023 Jan 1]. Available from: https://www.diamond.ac.uk/Home.html
40. Boivin S, Kozak S, Meijers R. Optimization of protein purification and characterization using Thermofluor screens. Protein expression and purification. 2013 Oct 1;91(2):192-206.
41. Gildea RJ, Beilsten-Edmands J, Axford D, Horrell S, Aller P, Sandy J, Sanchez-Weatherby J, Owen CD, Lukacik P, Strain-Damerell C, Owen RL. xia2. multiplex: a multi-crystal data-analysis pipeline. Acta Crystallographica Section D: Structural Biology. 2022 Jun 1;78(6):752-69.
42. Avanzato VA, Oguntuyo KY, Escalera-Zamudio M, Gutierrez B, Golden M, Kosakovsky Pond SL, et al. A structural basis for antibody-mediated neutralization of Nipah virus reveals a site of vulnerability at the fusion glycoprotein apex. Proceedings of the National Academy of Sciences. 2019 Nov 25;116(50):25057–67
43. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallographica Section D: Biological Crystallography. 2010 Feb 1;66(2):213-21.
44. Vagin A, Teplyakov A. Molecular replacement with MOLREP. Acta Crystallographica Section D: Biological Crystallography. 2010 Jan 1;66(1):22-5.
45. Panjkovich A. and Svergun D.I. CHROMIXS: automatic and interactive analysis of chromatography-coupled small angle X-ray scattering data. Bioinformatics, (2017) Winter G. xia2: an expert system for macromolecular crystallography data reduction. Journal of applied crystallography. 2010 Feb 1;43(1):186-90.
46. Hopkins J. Kratky analysis [Internet]. 2022 [cited 2023 Jan 1]. Available from: https://bioxtas-raw.readthedocs.io/en/latest/tutorial/s1_kratky.html#:~:text=A%20Kratky%20plot%20is%20a,%2Dshaped%20(Gaussian)%20peak.
47. Lee JE, Fusco ML, Hessell AJ, Oswald WB, Burton DR, Saphire EO. Structure of the ebola virus glycoprotein bound to an antibody from a human survivor. Nature. 2008;454(7201):177–82.
48. Sánchez-Felipe L, Villar E, Muñoz-Barroso I. Entry of Newcastle Disease Virus into the host cell: role of acidic pH and endocytosis. Biochimica et Biophysica Acta (BBA)-Biomembranes. 2014 Jan 1;1838(1):300-9.

Chapter IV:

1. Kausche GA, Pfankuch E, Ruska H. Die Sichtbarmachung von pflanzlichem virus im Übermikroskop. Naturwissenschaften. 1939 May;27:292-9.
2. Von Borries B, Ruska E, Ruska H. Bakterien und Virus in Übermikroskopischer Aufnahme: Mit einer Einführung in die Technik des Übermikroskops. Klinische Wochenschrift. 1938 Jul;17:921-5.
3. Renaud J-P, Chari A, Ciferri C, Liu W, Rémigy H-W, Stark H, et al. Cryo-EM in drug discovery: Achievements, limitations and prospects. Nature Reviews Drug Discovery. 2018;17(7):471–92.

4. Faruqi AR, Henderson R. Electronic detectors for electron microscopy. Current opinion in structural biology. 2007 Oct 1;17(5):549-55.
5. Dubochet J, McDowall AW. Vitrification of pure water for electron microscopy. Journal of Microscopy. 1981;124(3):3-4.
6. Kühlbrandt W. The resolution revolution. Science. 2014 Mar 28;343(6178):1443-4.
7. Yip KM, Fischer N, Paknia E, Chari A, Stark H. Breaking the next Cryo-EM resolution barrier–Atomic resolution determination of proteins!. BioRxiv. 2020 May 22:2020-05.
8. García-Nafría J, Lee Y, Bai X, Carpenter B, Tate CG. Cryo-EM structure of the adenosine A2A receptor coupled to an engineered heterotrimeric G protein. elife. 2018 May 4;7:e35946.
9. Sayre D. X-ray crystallography: The past and present of the phase problem. Structural Chemistry. 2002 Feb;13:81-96.
10. Ohi M, Li Y, Cheng Y, Walz T. Negative staining and image classification—powerful tools in modern electron microscopy. Biological procedures online. 2004 Jan;6:23-34.
11. Swanson K, Wen X, Leser GP, Paterson RG, Lamb RA, Jardetzky TS. Structure of the Newcastle disease virus F protein in the post-fusion conformation. Virology. 2010 Jul 5;402(2):372-9.
12. Welch BD, Liu Y, Kors CA, Leser GP, Jardetzky TS, Lamb RA. Structure of the cleavage-activated prefusion form of the parainfluenza virus 5 fusion protein. Proceedings of the National Academy of Sciences. 2012 Oct 9;109(41):16672-7.
13. McLellan JS, Chen M, Leung S, Graepel KW, Du X, Yang Y, Zhou T, Baxa U, Yasuda E, Beaumont T, Kumar A. Structure of RSV fusion glycoprotein trimer bound to a prefusion-specific neutralizing antibody. Science. 2013 May 31;340(6136):1113-7.
14. Xu K, Chan Y-P, Bradel-Tretheway B, Akyol-Ataman Z, Zhu Y, Dutta S, et al. Crystal structure of the pre-fusion Nipah virus fusion glycoprotein reveals a novel hexamer-of-trimers assembly. PLOS Pathogens. 2015;11(12).
15. Kalbermatter D, Shrestha N, Gall FM, Wyss M, Riedl R, Plattet P, et al. Cryo-EM structure of the Prefusion State of canine distemper virus fusion protein ectodomain. Journal of Structural Biology: X. 2020;4:100021.
16. Source DL. Chameleon [Internet]. [cited 2023 Jan 1]. Available from: https://www.diamond.ac.uk/Instruments/Biological-Cryo-Imaging/eBIC/Instruments/Chameleon.html
17. Zhang Z, Shigematsu H, Shimizu T, Ohto U. Improving particle quality in cryo-EM analysis using a pegylation method. Structure. 2021;29(10).
18. Jain A, Jain SK. PEGylation: an approach for drug delivery. A review. Critical Reviews™ in Therapeutic Drug Carrier Systems. 2008;25(5).
19. Jain A, Jain SK. PEGylation: an approach for drug delivery. A review. Critical Reviews™ in Therapeutic Drug Carrier Systems. 2008;25(5).
20. Mattson, G et al. “A practical approach to crosslinking.” Molecular biology reports vol. 17,3 (1993): 167-83.
21. Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. CryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nature methods. 2017 Mar;14(3):290-6
22. Source DL. Chameleon [Internet]. [cited 2023 Jan 1]. Available from: https://www.diamond.ac.uk/Instruments/Biological-Cryo-Imaging/eBIC/Instruments/Chameleon.html
23. Uncle - Protein Stability Screening Platform: Unchained Labs [Internet]. 2023 [cited 2023 Jan 1]. Available from: https://www.unchainedlabs.com/uncle/#resources
24. Lamb RA, Jardetzky TS. Structural basis of viral invasion: lessons from paramyxovirus F. Current opinion in structural biology. 2007 Aug 1;17(4):427-36.
25. Sánchez-Felipe L, Villar E, Muñoz-Barroso I. Entry of Newcastle Disease Virus into the host cell: role of acidic pH and endocytosis. Biochimica et Biophysica Acta (BBA)-Biomembranes. 2014 Jan 1;1838(1):300-9.

26. Seow, Jeffrey et al. “A neutralizing epitope on the SD1 domain of SARS-CoV-2 spike targeted following infection and vaccination.” Cell reports vol. 40,8 (2022): 111276.
27. Informal conversation
28. Chen L, Gorman JJ, McKimm-Breschkin J, Lawrence LJ, Tulloch PA, Smith BJ, et al. The structure of the fusion glycoprotein of Newcastle Disease Virus suggests a novel paradigm for the molecular mechanism of membrane fusion. Structure. 2001;9(3):255–66
29. Tan YZ, Baldwin PR, Davis JH, Williamson JR, Potter CS, Carragher B, et al. Addressing preferred specimen orientation in single-particle cryo-em through tilting. Nature Methods. 2017;14(8):793–6.
30. Ahkong QF, Howell JI, Lucy JA, Safwat F, Davey MR, Cocking EC. Fusion of hen erythrocytes with yeast protoplasts induced by polyethylene glycol. Nature. 1975 May 1;255(5503):66-7.

Chapter V:

1. Dimitrov KM, Afonso CL, Yu Q, Miller PJ. Newcastle disease vaccines—a solved problem or a continuous challenge? Veterinary Microbiology. 2017;206:126–36.
2. VICH G. 41: Guideline on target animal safety: examination of live veterinary vaccines in target animals for absence of reversion to virulence. European Medicines Agency. EMEA/CVMP/VICH/1052/2004. 2007.
3. Naz D, ur Rahman S, Aslam MA, Muhammad F. Newcastle disease virus in poultry with an interface as a human vector. Veterinary Vaccine. 2022 Sep 13:100003.
4. Fawzy M, Ali RR, Elfeil WK, Saleh AA, El-Tarabilli MM. Efficacy of inactivated velogenic Newcastle disease virus genotype VII vaccine in broiler chickens. InVeterinary Research Forum 2020 (Vol. 11, No. 2, p. 113). Faculty of Veterinary Medicine, Urmia University, Urmia, Iran.
5. Schijns VE, van de Zande S, Lupiani B, Reddy SM. Practical aspects of poultry vaccination. InAvian immunology 2014 Jan 1 (pp. 345-362). Academic Press.
6. Miller PJ, Afonso CL, El Attrache J, Dorsey KM, Courtney SC, Guo Z, Kapczynski DR. Effects of Newcastle disease virus vaccine antibodies on the shedding and transmission of challenge viruses. Developmental & Comparative Immunology. 2013 Dec 1;41(4):505-13.
7. Ben Peeters, Guus Koch, Newcastle Disease Virus (Paramyxoviridae), Encyclopedia of Virology (Fourth Edition), Academic Press,2021, Pages 648-653,
8. Stewart-Brown B. Vaccination programs for poultry - poultry [Internet]. MSD Veterinary Manual; 2023 [cited 2023 Jan 2]. Available from: https://www.msdvetmanual.com/poultry/nutrition-and-management-poultry/vaccination-programs-for-poultry
9. Cicek H, Tandogan M. Estimation of optimum slaughter age in broiler chicks. Indian Journal of Animal Research. 2016;50(4):621-3.
10. Rehmani SF, Wajid A, Bibi T, Nazir B, Mukhtar N, Hussain A, Lone NA, Yaqub T, Afonso CL. Presence of virulent Newcastle disease virus in vaccinated chickens in farms in Pakistan. Journal of clinical microbiology. 2015 May;53(5):1715-8.
11. Perozo F, Marcano R, Afonso CL. Biological and phylogenetic characterization of a genotype VII Newcastle disease virus from Venezuela: efficacy of field vaccination. Journal of clinical microbiology. 2012 Apr;50(4):1204-8.
12. Jackson DA, Symons RH, Berg P. Biochemical method for inserting new genetic information into DNA of Simian Virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of Escherichia coli. Proceedings of the National Academy of Sciences. 1972 Oct;69(10):2904-9.
13. Karaca K, Sharma JM, Winslow BJ, Junker DE, Reddy S, Cochran M, McMillen J. Recombinant fowlpox viruses coexpressing chicken type I IFN and Newcastle disease virus HN and F genes: influence of IFN on protective efficacy and humoral responses of chickens

following in ovo or post-hatch administration of recombinant viruses. Vaccine. 1998 Oct 1;16(16):1496-503.
14. Folegatti PM, Ewer KJ, Aley PK, Angus B, Becker S, Belij-Rammerstorfer S, Bellamy D, Bibi S, Bittaye M, Clutterbuck EA, Dold C. Safety and immunogenicity of the ChAdOx1 nCoV-19 vaccine against SARS-CoV-2: a preliminary report of a phase 1/2, single-blind, randomised controlled trial. The Lancet. 2020 Aug 15;396(10249):467-78.
15. Roohani K, Tan SW, Yeap SK, Ideris A, Bejo MH, Omar AR. Characterisation of genotype VII Newcastle disease virus (NDV) isolated from NDV vaccinated chickens, and the efficacy of LaSota and recombinant genotype VII vaccines against challenge with velogenic NDV. Journal of veterinary science. 2015 Dec 1;16(4):447-57.
16. Taylor J, Christensen L, Gettig R, Goebel J, Bouquet JF, Mickle TR, Paoletti E. Efficacy of a recombinant fowl pox-based Newcastle disease virus vaccine candidate against velogenic and respiratory challenge. Avian diseases. 1996 Jan 1:173-80.
17. Morgan RW, Gelb Jr J, Schreurs CS, Lütticken D, Rosenberger JK, Sondermeijer PJ. Protection of chickens from Newcastle and Marek's diseases with a recombinant herpesvirus of turkeys vaccine expressing the Newcastle disease virus fusion protein. Avian diseases. 1992 Oct 1:858-70.
18. Nair V. Successful control of Marek's disease by vaccination. Developments in biologicals. 2004 Jan 1;119:147-54.
19. Riedel S. Edward Jenner and the history of smallpox and vaccination. InBaylor University medical center proceedings 2005 Jan 1 (Vol. 18, No. 1, pp. 21-25). Taylor & Francis.
20. Reynolds D, McMillen J, Cook S, Schwartz R, Sharma J. A recombinant HVT vaccine expressing Newcastle disease virus antigens protects chicks against a lethal Newcastle disease challenge. InWestern Poultry Disease Conference (USA) 1993.
21. Dey S, Chellappa MM, Pathak DC, Gaikwad S, Yadav K, Ramakrishnan S, Vakharia VN. Newcastle disease virus vectored bivalent vaccine against virulent infectious bursal disease and Newcastle disease of chickens. Vaccines. 2017 Sep 26;5(4):31.
22. Innovax-nd-ilt [Internet]. 2022 [cited 2023 Apr 30]. Available from: https://www.ema.europa.eu/en/medicines/veterinary/EPAR/innovax-nd-ilt
23. Maekawa D, Riblet SM, Newman L, Koopman R, Barbosa T, García M. Evaluation of vaccination against infectious laryngotracheitis (ILT) with recombinant herpesvirus of turkey (rHVT-LT) and chicken embryo origin (CEO) vaccines applied alone or in combination. Avian Pathology. 2019 Nov 2;48(6):573-81.
24. Abdul-Cader MS, Palomino-Tapia V, Amarasinghe A, Ahmed-Hassan H, De Silva Senapathi U, Abdul-Careem MF. Hatchery vaccination against poultry viral diseases: potential mechanisms and limitations. Viral immunology. 2018 Jan 1;31(1):23-33.
25. Garrido C, Okazaki W, Purchase HG, Burmester BR. Storage of a cell-associated herpesvirus of turkeys (HVT Strain FC126) vaccine. Avian diseases. 1972 Apr 1:45-51.
26. Bublot M, Pritchard N, Le Gros FX, Goutebroze S. Use of a vectored vaccine against infectious bursal disease of chickens in the face of high-titred maternally derived antibody. Journal of comparative pathology. 2007 Jul 1;137:S81-4.
27. Ghanem IA, Abdullatif TM, Hassanin O. The protection conferred against virulent Newcastle disease virus (vNDV) genotype VII by commercial double recombinant HVT vaccines and NDV live-attenuated vaccine as prime/boost vaccination regimens in commercial broiler chickens carrying maternally-derived antibodies (MDAs) against NDV. Avian Pathology. 2023 Jun 6:1-3.
28. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. Cell. 2014 Jun 5;157(6):1262-78.
29. Zhu J, Ananthaswamy N, Jain S, Batra H, Tang WC, Rao VB. CRISPR engineering of bacteriophage T4 to design vaccines against SARS-CoV-2 and emerging pathogens. Vaccine Design: Methods and Protocols, Volume 1. Vaccines for Human Diseases. 2022:209-28.

30. Zou Z, Huang K, Wei Y, Chen H, Liu Z, Jin M. Construction of a highly efficient CRISPR/Cas9-mediated duck enteritis virus-based vaccine against H5N1 avian influenza virus and duck Tembusu virus infection. Scientific Reports. 2017 May 3;7(1):1478.
31. Borca MV, Holinka LG, Berggren KA, Gladue DP. CRISPR-Cas9, a tool to efficiently increase the development of recombinant African swine fever viruses. Scientific reports. 2018 Feb 16;8(1):3154.
32. Tang N, Zhang Y, Sadigh Y, Moffat K, Shen Z, Nair V, Yao Y. Generation of a triple insert live avian herpesvirus vectored vaccine using CRISPR/Cas9-based gene editing. Vaccines. 2020 Feb 21;8(1):97.
33. Ding L, Chen P, Bao X, Li A, Jiang Y, Hu Y, Ge J, Zhao Y, Wang B, Liu J, Chen H. Recombinant duck enteritis viruses expressing the Newcastle disease virus (NDV) F gene protects chickens from lethal NDV challenge. Veterinary microbiology. 2019 May 1;232:146-50.
34. Nakaya T, Cros J, Park MS, Nakaya Y, Zheng H, Sagrera A, Villar E, García-Sastre A, Palese P. Recombinant Newcastle disease virus as a vaccine vector. Journal of virology. 2001 Dec 1;75(23):11868-73.
35. Hamers-Casterman CT, Atarhouch T, Muyldermans SA, Robinson G, Hammers C, Songa EB, Bendahman N, Hammers R. Naturally occurring antibodies devoid of light chains. Nature. 1993 Jun 3;363(6428):446-8.
36. Roosild TP, Castronovo S, Choe S. Structure of anti-FLAG M2 Fab domain and its use in the stabilization of engineered membrane proteins. Acta Crystallographica Section F: Structural Biology and Crystallization Communications. 2006 Sep 1;62(9):835-9.
37. Desmyter A, Spinelli S, Roussel A, Cambillau C. Camelid nanobodies: killing two birds with one stone. Current opinion in structural biology. 2015 Jun 1;32:1-8.
38. Song J, Markley JL. Cautionary tail: the presence of an N-terminal tag on dynein light-chain Roadblock/LC7 affects its interaction with a functional partner. Protein and Peptide Letters. 2007 Mar 1;14(3):265-8.

Chapter VI:

1. Chen L, Gorman JJ, McKimm-Breschkin J, Lawrence LJ, Tulloch PA, Smith BJ, et al. The structure of the fusion glycoprotein of Newcastle Disease Virus suggests a novel paradigm for the molecular mechanism of membrane fusion. Structure. 2001;9(3):255–66.
2. Dimitrov KM, Afonso CL, Yu Q, Miller PJ. Newcastle disease vaccines—a solved problem or a continuous challenge? Veterinary Microbiology. 2017;206:126–36.
3. Roy P. Diagnosis and control of Newcastle disease in developing countries. World's Poultry Science Journal. 2012 Dec;68(4):693-706.
4. Kinde H, Hullinger PJ, Charlton B, McFarland M, Hietala SK, Velez V, Case JT, Garber L, Wainwright SH, Mikolon AB, Breitmeyer RE. The isolation of exotic Newcastle disease (END) virus from nonpoultry avian species associated with the epidemic of END in chickens in southern California: 2002–2003. Avian diseases. 2005 Jun 1;49(2):195-8.
5. Ibrahim AL, Lai MC, Aini I. Spray vaccination with an improved F Newcastle disease vaccine. A comparison of efficacy with the B1 and La Sota vaccines. British Veterinary Journal. 1983 May 1;139(3):213-9.
6. Innovax-nd-ilt [Internet]. 2022 [cited 2023 Apr 30]. Available from: https://www.ema.europa.eu/en/medicines/veterinary/EPAR/innovax-nd-ilt
7. Ferreira HL, Reilley AM, Goldenberg D, Ortiz IR, Gallardo RA, Suarez DL. Protection conferred by commercial NDV live attenuated and double recombinant HVT vaccines against

virulent California 2018 Newcastle disease virus (NDV) in chickens. Vaccine. 2020 Jul 22;38(34):5507-15.
8. Swanson K, Wen X, Leser GP, Paterson RG, Lamb RA, Jardetzky TS. Structure of the Newcastle disease virus F protein in the post-fusion conformation. Virology. 2010 Jul 5;402(2):372-9.
9. Chang LA, Phung E, Crank MC, Morabito KM, Villafana T, Dubovsky F, Falloon J, Esser MT, Lin BC, Chen GL, Graham BS. A prefusion-stabilized RSV F subunit vaccine elicits B cell responses with greater breadth and potency than a postfusion F vaccine. Science Translational Medicine. 2022 Dec 21;14(676):eade0424.
10. Ebel H, Benecke T, Vollmer B. Stabilisation of viral membrane fusion proteins in prefusion conformation by structure-based design for structure determination and vaccine development. Viruses. 2022 Aug 18;14(8):1816.
11. McLellan JS, Chen M, Leung S, Graepel KW, Du X, Yang Y, Zhou T, Baxa U, Yasuda E, Beaumont T, Kumar A. Structure of RSV fusion glycoprotein trimer bound to a prefusion-specific neutralizing antibody. Science. 2013 May 31;340(6136):1113-7.
12. Welch BD, Liu Y, Kors CA, Leser GP, Jardetzky TS, Lamb RA. Structure of the cleavage-activated prefusion form of the parainfluenza virus 5 fusion protein. Proceedings of the National Academy of Sciences. 2012 Oct 9;109(41):16672-7.
13. Lamb RA, Jardetzky TS. Structural basis of viral invasion: lessons from paramyxovirus F. Current opinion in structural biology. 2007 Aug 1;17(4):427-36.
14. Sánchez-Felipe L, Villar E, Muñoz-Barroso I. Entry of Newcastle Disease Virus into the host cell: role of acidic pH and endocytosis. Biochimica et Biophysica Acta (BBA)-Biomembranes. 2014 Jan 1;1838(1):300-9.
15. Yin HS, Wen X, Paterson RG, Lamb RA, Jardetzky TS. Structure of the parainfluenza virus 5 F protein in its metastable, prefusion conformation. Nature. 2006 Jan 5;439(7072):38-44.
16. Avanzato VA, Oguntuyo KY, Escalera-Zamudio M, Gutierrez B, Golden M, Kosakovsky Pond SL, et al. A structural basis for antibody-mediated neutralization of Nipah virus reveals a site of vulnerability at the fusion glycoprotein apex. Proceedings of the National Academy of Sciences. 2019;116(50):25057–67. doi:10.1073/pnas.1912503116
17. Kalbermatter D, Shrestha N, Gall FM, Wyss M, Riedl R, Plattet P, et al. Cryo-EM structure of the Prefusion State of canine distemper virus fusion protein ectodomain. Journal of Structural Biology: X. 2020;4:100021.
18. Source DL. Chameleon [Internet]. [cited 2023 Jan 1]. Available from: https://www.diamond.ac.uk/Instruments/Biological-Cryo-Imaging/eBIC/Instruments/Chameleon.html
19. Song J, Markley JL. Cautionary tail: the presence of an N-terminal tag on dynein light-chain Roadblock/LC7 affects its interaction with a functional partner. Protein and Peptide Letters. 2007 Mar 1;14(3):265-8.
20. Hsieh CL, Goldsmith JA, Schaub JM, DiVenere AM, Kuo HC, Javanmardi K, Le KC, Wrapp D, Lee AG, Liu Y, Chou CW. Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. Science. 2020 Sep 18;369(6510):1501-5.
21. Crank MC, Ruckwardt TJ, Chen M, Morabito KM, Phung E, Costner PJ, Holman LA, Hickman SP, Berkowitz NM, Gordon IJ, Yamshchikov GV. A proof of concept for structure-based vaccine design targeting RSV in humans. Science. 2019 Aug 2;365(6452):505-9.
22. Gershoni, J. M., Roitburd-Berman, A., Siman-Tov, D. D., Tarnovitski Freund, N., & Weiss, Y. (2007). Epitope mapping: the first step in developing epitope-based vaccines. BioDrugs : clinical immunotherapeutics, biopharmaceuticals and gene therapy, 21(3), 145–156.
23. The regulation of specified animal pathogens [Internet]. Health and Safety Executive; [cited 2023 Jan 3]. Available from: https://www.hse.gov.uk/biosafety/sapo.htm

Appendix B: Gene Sequences

B1: DNA Constructs used for Protein Expression: Highlighted yellow DNA residues indicate NcoI and NotI restriction enzyme sites. The N residues in GTVII are highlighted to show potential targets for N-linked glycosylation.

La Sota DNA Sequence:

CCATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCGGCCATCGACGGTC
GTCCTTTGGCTGCTGCTGGTATCGTGGTCACTGGCGACAAGGCTGTGAACATCTACACCTCCAGCCAGACCGGTT
CCATCATCGTGAAGCTGCTGCCCAACCTGCCTAAGGACAAAGAGGCTTGCGCTAAGGCTCCCCTGGACGCTTACA
ACAGGACCCTGACTACCCTGCTGACCCCTCTGGGAGACTCCATCCGTCGTATCCAAGAGTCCGTGACCACCTCTG
GTGGTGGTCGTCAAGGTCGTCTGATCGGTGCTATCATCGGTGGTGTCGCTCTGGGTGTCGCTACCGCTGCTCAAA
TCACAGCTGCCGCTGCTCTGATCCAGGCTAAGCAGAACGCTGCTAACATCCTGCGTCTGAAAGAGTCTATCGCTG
CTACCAACGAGGCTGTGCACGAAGTGACCGACGGACTGTCTCAACTGGCTGTGGCTGTGGGCAAGATGCAGCAGT
TCGTGAACGACCAGTTCAACAAGACCGCTCAAGAGCTGGACTGCATCAAGATCGCTCAGCAAGTGGGCGTCGAGC
TGAACCTGTACCTGACCGAGCTGACCACCGTGTTCGGTCCCCAGATCACTTCCCCTGCTCTGAACAAGCTGACCA
TCCAGGCTCTGTACAACCTGGCTGGTGGCAACATGGACTACCTGCTCACCAAGCTCGGTGTCGGCAACAACCAGC
TGTCCAGCCTGATCGGTTCTGGCCTGATCACTGGCAACCCCATCCTGTACGACTCTCAGACCCAGCTGCTGGGCA
TCCAAGTGACCTTGCCTTCCGTGGGCAACCTGAACAACATGCGTGCTACCTACCTGGAAACCCTGTCCGTGTCTA
CCACCAGGGGTTTCGCTTCAGCTCTGGTGCCTAAGGTCGTGACCCAAGTGGGATCCGTGATCGAGGAACTGGACA
CCTCCTACTGCATCGAGACTGACCTGGACCTGTACTGCACCCGTATCGTGACTTTCCCCATGTCTCCCGGTATCT
ACTCCTGCCTGTCCGGCAACACCTCCGCTTGCATGTACTCCAAGACCGAGGGTGCTCTGACCACTCCTTACATGA
CCATCAAGGGTAGCGTGATCGCCAACTGCAAGATGACCACTTGCCGTTGCGTGAACCCTCCTGGTATCATCTCCC
AGAACTACGGCGAGGCTGTCTCCCTGATCGACAAGCAGTCTTGCAACGTGCTGTCCCTCGGTGGTATCACCCTGC
GTTTGTCTGGCGAGTTCGACGTGACCTACCAGAAGAACATCTCCATCCAGGACTCCCAAGTGATCATCACCGGCA
ACCTGGACATCTCCACCGAGCTGGGAAACGTGAACAACTCCATCTCCAACGCTCTCAACAAGTTGGAGGAATCCA
ACCGCAAGCTGGACAAGATCGAGGATAAGATCGAAGAGATCCTGTCCAAGATCTACCACATCGAGAACGAGATCG
CCCGCATCAAGAAGCTGATCGGAGAGGCTCCTGGCAAGCACCACCATCACCATCATTGATAGCGGCCGC

La Sota amino acid sequence:

>GP64 (signal sequence)_LaS_GCNt_his6

MVSAIVLYVLLAAAAHSAFAAIDGRPLAAAGIVVTGDKAVNIYTSSQTGSIIVKLLPNLP
KDKEACAKAPLDAYNRTLTTLLTPLGDSIRRIQESVTTSGGGRQGRLIGAIIGGVALGVA
TAAQITAAAALIQAKQNAANILRLKESIAATNEAVHEVTDGLSQLAVAVGKMQQFVNDQF
NKTAQELDCIKIAQQVGVELNLYLTELTTVFGPQITSPALNKLTIQALYNLAGGNMDYLL
TKLGVGNNQLSSLIGSGLITGNPILYDSQTQLLGIQVTLPSVGNLNNMRATYLETLSVST
TRGFASALVPKVVTQVGSVIEELDTSYCIETDLDLYCTRIVTFPMSPGIYSCLSGNTSAC
MYSKTEGALTTPYMTIKGSVIANCKMTTCRCVNPPGIISQNYGEAVSLIDKQSCNVLSLG
GITLRLSGEFDVTYQKNISIQDSQVIITGNLDISTELGNVNNSISNALNKLEESNRKLDK
IEDKIEEILSKIYHIENEIARIKKLIGEAPGKHHHHHH
(A was added)

GTVII DNA sequence:

CCATGGTAAGCGCTATTGTTTTATATGTGCTTTTGGCGGCGGCGGCGCATTCTGCCTTTGCTGCCTTGGACGGTC
GTCCTTTGGCTGCTGCTGGTATCGTGGTCACTGGCGACAAGGCTGTGAACATCTACACCTCCAGCCAGACCGGTT
CCATCATCGTGAAGCTGCTGCCCAACATGCCCAAGGACAAAGAGGCTTGCGCTCGCGCTCCTCTCGAGGCTTACA
ACCGTACTCTGACTACCCTGCTGACCCCTCTGGGAGACTCCATCAGAAAGATCCAGGGTTCCGTGTCCACCTCCG
GTGGTCGTAGACGTTTCATCGGTGCTATCATCGGTTCCGTGGCTCTGGGTGTCGCTACCGCTGCTCAAATCACAG
CTGCCGCTGCTCTGATCCAGGCTAACCAGAACGCTGCTAACATCCTGCGTCTGAAAGAGTCTATCGCTGCTACCA
ACGAGGCTGTGCACGAAGTGACCAACGGACTGTCCCAGCTGTCTGTGGCTGTGGGCAAGATGCAGCAGTTCGTGA
ACGACCAGTTCAACAACACCGCTCGTGAACTGGACTGCATCAAGATCACCCAGCAAGTGGGCGTCGAGCTGAACC
TGTACCTGACCGAGCTGACCACCGTGTTCGGTCCCCAGATCACTTCTCCAGCTCTGACTCAGCTGACCATCCAGG
CTCTGTACAACCTGGCTGGTGGCAACATGGACTACCTGCTCACCAAGCTCGGTGTCGGCAACAACCAGCTGTCCT
CTCTGATCGGTTCTGGCCTGATCACTGGTTACCCCATCCTGTACGACTCTCAGACCCAGCTGCTGGGCATCCAAG
TGAACCTGCCTTCCGTGGGCAACCTGAACAACAAGCGTGCTACCTACCTGGAAACCCTGTCCGTGTCTACCACCA
AGGGTTTCGCTTCCGCTCTGGTGCCTAAGGTGGTCACCCAAGTGGGATCCGTGATCGAGGAACTGGACACCTCCT
ACTGCATCGAGTCCGACCTGGACCTGTACTGCACCCGTATCGTGACTTTCCCCATGTCTCCCGGTATCTACTCCT
GCCTGTCCGGCAACACCTCCGCTTGCATGTACTCCAAGACCGAGGGTGCTCTGACCACTCCTTACATGACCCTGA
AGGGTTCTGTGATCGCCAACTGCAAGATCACTACCTGCCGTTGCGCTGACCCTCCTGGTATCATCTCCCAGAACT
ACGGCGAGGCTGTCTCCCTGATCGACCGTCGTTCTTGCAACGTGCTGTCCCTGGACGGTATCACCCTGCGTTTGT
CTGGCGAGTTCGACGCCACCTACCAGAAGAACATCTCCATCCTGGACTCCCAAGTGATCGTGACCGGCAACCTGG
ACATCTCCACCGAGCTGGGAAACGTGAACAACTCCATCTCCAACGCTCTGGACAAGCTGGCCGAGTCCAACTCCA
AGCTGGACAAGATCGAGGATAAGATCGAAGAGATCCTGTCCAAGATCTACCACATCGAGAACGAGATCGCCCGCA
TCAAGAAGCTGATCGGAGAGGCTCCTGGCAAGCACCACCACCATCACCACTGATAGCGGCCGC

GTVII amino-acid sequence:

>GP64 (signal sequence)_GTVII_GCNt_his6

MVSAIVLYVLLAAAAHSAFAALDGRPLAAAGIVVTGDKAVNIYTSSQTGS
IIVKLLPNMPKDKEACARAPLEAYNRTLTTLLTPLGDSIRKIQGSVSTSGGRRRFIGA
IIGSVALGVATAAQITAAAALIQANQNAANILRLKESIAATNEAVHEVTNGLSQLSVAVG
KMQQFVNDQFNNTARELDCIKITQQVGVELNLYLTELTTVFGPQITSPALTQLTIQALYN
LAGGNMDYLLTKLGVGNNQLSSLIGSGLITGYPILYDSQTQLLGIQVNLPSVGNLNNKRA
TYLETLSVSTTKGFASALVPKVVTQVGSVIEELDTSYCIESDLDLYCTRIVTFPMSPGIY
SCLSGNTSACMYSKTEGALTTPYMTLKGSVIANCKITTCRCADPPGIISQNYGEAVSLID
RRSCNVLSLDGITLRLSGEFDATYQKNISILDSQVIVTGNLDISTELGNVNNSISNALDK
LAESNSKLDKVNVRLTSTSAIEDKIEEILSKIYHIENEIARIKKLIGEAPGGKHHHHHH

modified proteolytic cleavage site

B2: Primers for the cloning of NDV into the PCDA3.1 vector, prior to further cloning into PGEMTeasy, for insertion into HVT:

Primers for cloning F protein into PCDA3.1:

Primer P1-1:
Fwd: 5'-Protection->Not1->SigSeq (universal)
→5'- ataagaatGCGGCCGCATGGTAAGCGCTATTGTTTT
Primer P1-2:
Rev: 3'-Protection->Not1->Term->6His->GCNT (LaS & GTVII)
5'-CTGGCAAGCACCACCACCATCACCACTGATAGCGGCCGCataagaat
3'- GACCGTTCGTGGTGGTGGTAGTGGTGACTATCGCCGGCGtattctta
→5'- attcttatGCGGCCGCTATCAGTGGTGATGGTGGTGGTGCTTGCCAG
Primer P1-3:
Rev: 3'-Protection->Not1->Term->6His->Stalk (LaS & GTVII)
5'- TCTCACCACCACCATCACCACTGATAGCGGCCGCataagaat
3'- AGAGTGGTGGTGGTAGTGGTGACTATCGCCGGCGtattctta
→5'- attcttatGCGGCCGCTATCAGTGGTGATGGTGGTGGTGAGA

Primers for cloning the F from PCDA3.1 into PGemTeasy:

mCMV Forward	5'-TCTATTGGCTGAGCTGCGTT
pCDNA3.1- Reverse	5'-AGCGGGTTTAAACGGCCATA

:

B3: DNA VHH sequences sent by collaborators, information is not available. Yellow indicates the PstI restriction site and green indicates the restriction site for KpnI.

VHH2

atgcgaCTGCAGATGGGATGGAGCTGTATCATCCTCTTCTTGGTAGCAACAGCTACAGGTGTCCACTCCGCTCA
GGTGCAGCTCGTGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGGGGGTCTCTGAAGCTCTCCTGTACAGCC
TCTGGATTCACTTTGGATCAATACGCCACAGCCTGGTTTCGCCAGGCCCCAGGGAAGGAGCGTGAGGGGGT
CGCATGTATTAGTGCTGGTGATGGCAGCACAGACTATGCAGATTTCGTGGAGGGCCGATTCACCATGGCCA
GAGACAAGACCAAGAACTCGGTATATTTGCAAATGGGCAGCCTGGAACCTGAGGACACAGCCGTTTATTAC
TGTGCCCCACTCTTCAGTTTCTACTGTAAAAATTATGAACGCGACGTCCAAGACATAGACTACTGGGGCAA
AGGGACCCTGGTCACCGTCTCCTTAGAACCCAAGACACCAAAACCACAACCAGGCCATTATCACCACCACCA
TCACCACTAGGTCACCatatcg

VHH4

atgcgaCTGCAGATGGGATGGAGCTGTATCATCCTCTTCTTGGTAGCAACAGCTACAGGTGTCCACTCCGCTCA
GTTGCAGCTCGTGGAGTCTGGGGGAGACTTGGTCCAGCCTGGGGGGTCTCTGACACTCTCTTGTGCAGCC
TCTGGCTTCGCTTTGGGTAATACTGCCGTAGGCTGGTTCCGCCAGGCCCCAGGGAAGGAGCGTGAGGGGC
TCTCATGTATCAGTAGTATTACGGAAATCACAAAGTATACAGACTCCGTGAAGGGCCGATTCACCGTCTCC
AGAGACAACGCCAAGAACGCGGTGTATCTGCAAATGAACAGCCTGAAACCTGAGGACACGGCCGTTTATTA
CTGTGCAGCAGAAGGAGACCCGGTAGATGGTCTATGTACAACAGTAGGCCGCTATTATACGTACTGGGGCC
AGGGGACCCAGGTCACCGTGTCCTCAGAACCCAAGACACCAAAACCACAACCAGGCCATTATCACCACCACC
ATCACCACTAGGTCACCatatcg

VHH 9 STOP39

atgcgaCTGCAGATGGGATGGAGCTGTATCATCCTCTTCTTGGTAGCAACAGCTACAGGTGTCCACTCCGCTCA
GGTGCAGCTCGTGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGGGGGTCTCTGTGACTCTCCTGTGCAGCC
TCTGGAGGCTCCCTTAGAGAATATGCGATGGGCTGGTTCCGCCAGCGTCCAGGGAAGGCGCGTGAGTTTG
TCGCAGTTATTAGTTGGAGTGCTGATATGAGGTACTATTCAGACTCCGTGAAGGGCCGATTCACCATCTCC
AGAGACAACGGCGCGACCACGGTGTATCTGCAAATGAACAGCCTGAAACCAGAAGACACGGCCGTTTATTA
TTGTTATGCGGATGAAGTTGGGGGTTGGAGTCCATCGCGGGTTGACTACTGGGGTCAGGGGACCCAGGTC
ACCGTCTCAACAGAACCCAAGACACCAAAACCACAACCAGGCCATTATCACCACCACCATCACCACTAGGTC
ACCatatcg

B4: pMESy4 vector and Primers used to clone VHH sequences into it.

Key:
PstI
BsteII(Off site BsteII in blue)
CseI (Recognition site, NNNNN before cut in black)
STOP
6-His
CUT segment

PMEASY Vector:

agcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactgg
aaagcgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttcc
ggctcgtatgttgtgtggaattgtgagcggataacaatttcacaagcttaaggagacagtacatatgaaatacctattgc
ctacggcagccgctggattgttattactcgcggcccagccggccatggcccaggtgcagctgcaggagtctagaggggac
ccaggtcaccgtctcctcacaccaccatcaccatcacgaacctgaagcctagtacccgtacgacgttccggactacggtt
ccactgttgaaagttgtttagcaaaacctcatacagaaaattcatttactaacgtctggaaagacgacaaaactttagat
cgttacgctaactatgagggctgtctgtggaatgctacaggcgttgtcgtttgtactggtgacgaaactcagtgttacgg
tacatgggttcctattgggcttgctatccctgaaaatgagggtggtggctctgagggtggcggttctgagggtggcggtt
ctgagggtggcggtactaaacctcctgagtacggtgatacacctattccgggctatacttatatcaaccctctcgacggc
acttatccgcctggtactgagcaaaaccccgctaatcctaatccttctcttgaggagtctcagcctcttaatactttcat
gtttcagaataataggttccgaaataggcagggtgcattaactgtttatacgggcactgttactcaaggcactgaccccg
ttaaaacttattaccagtacactcctgtatcatcaaaagccatgtatgacgcttactggaacggtaaattcagagactgc
gctttccattctggctttaatgaggatccattcgtttgtgaatatcaaggccaatcgtctgacctgcctcaacctcctgt
caatgctggcggcggctctggtggtggttctggtggcggctctgagggtggcggctctgagggtggcggctctgagggtg
gcggttctgagggtggcggctctgagggtggcggttccggtggcggctccggttccggtgattttgattatgaaaaaatg
gcaaacgctaataaggggggctatgaccgaaaatgccgatgaaaacgcgctacagtctgacgctaaaggcaaacttgattc
tgtcgctactgattacggtgctgctatcgatggtttcattggtgacgtttccggccttgctaatggtaatggtgctactg
gtgattttgctggctctaattcccaaatggctcaagtcggtgacggtgataattcacctttaatgaataatttccgtcaa
tatttaccttctttacctcagtcggttgaatgtcgcccttatgtctttggcgctggtaaaccatatgaattttctattga
ttgtgacaaaataaacttattccgtggtgtctttgcgtttcttttatatgttgccacctttatgtatgtattttcgacgt
ttgctaacatactgcgtaataaggagtcttaataagaattcactggccgtcgttttacaacgtcgtgactgggaaaaccc
tggcgttacccaacttaatcgccttgcagcacatccccctttcgccagctggcgtaatagcgaagaggcccgcaccgatc
gcccttcccaacagttgcgcagcctgaatggcgaatggcgcctgatgcggtattttctccttacgcatctgtgcggtatt
tcacaccgcatacgtcaaagcaaccatagtacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggttacgcgc
agcgtgaccgctacacttgccagcgccctagcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccgg
ctttccccgtcaagctctaaatcgggggctccctttagggttccgatttagtgctttacggcacctcgaccccaaaaaac
ttgatttgggtgatggttcacgtagtgggccatcgccctgatagacggtttttcgccctttgacgttggagtccacgttc
tttaatagtggactcttgttccaaactggaacaacactcaaccctatctcgggctattcttttgatttataagggatttt
gccgatttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaattttaacaaaatattaacgttta
caattttatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagccccgacacccgccaacacccgct
gacgcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtca
gaggttttcaccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatgtcatga
taataatggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaata
cattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattgaaaaaggaagagtatgagtat
tcaacatttccgtgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtga
aagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttgag
agttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattga
cgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccagtcacagaaaagc
atcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttactt
ctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttg
ggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgca
aactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcagga
ccacttctgcgctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtat
cattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatg
aacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatata
ctttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaat
cccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttttttc

tgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaact
ctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccacca
cttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagt
cgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggggttcgtgcaca
cagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccga
agggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacg
cctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcgg
agcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttctttcc
tgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgacc

Forward Primer: S11.1

5'- atgtaaCTGCAGATGGGATGGAGCT

Reverse Primer: S11.2

*5'-*GGCCATTATCACCACCACCATCACCACTAG

Reverse Primer: S11.3

5'- CCACTAGACGCATATGGTCACCatatcg

Appendix C: Permissions for the reproduction of Figures:

C1: Permission for figure 32,



Cryo-EM structure of the prefusion state of canine distemper virus fusion protein ectodomain

Author: David Kalbermatter,Neeta Shrestha,Flavio M. Gall,Marianne Wyss,Rainer Riedl,Philippe Plattet,Dimitrios Fotiadis
Publication: Journal of Structural Biology: X
Publisher: Elsevier
Date: 2020

© 2020 The Authors. Published by Elsevier Inc.

Creative Commons

This is an open access article distributed under the terms of the Creative Commons CC-BY license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

You are not required to obtain permission to reuse this article.

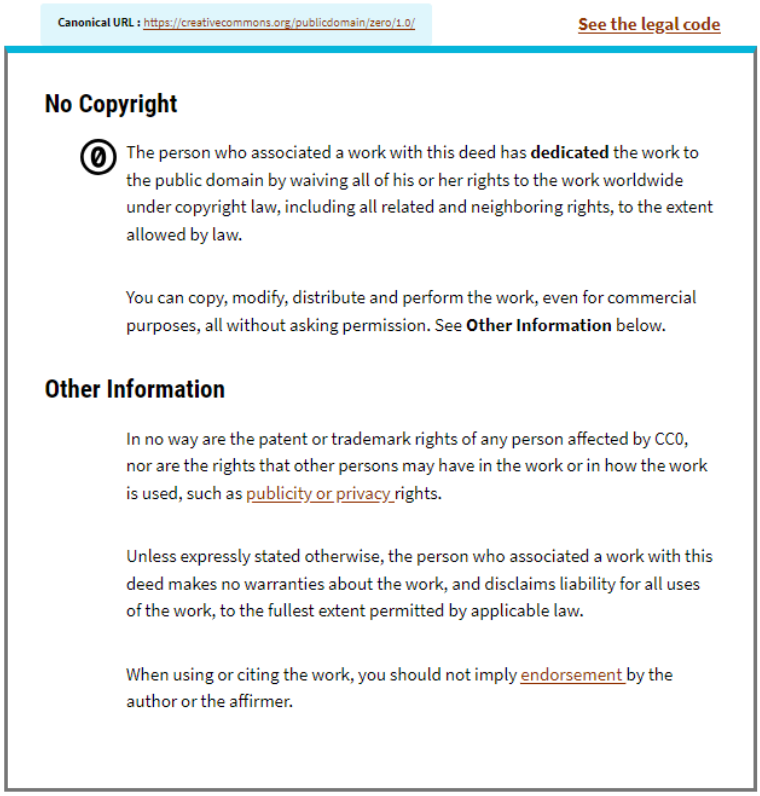
To request permission for a type of use not listed, please contact Elsevier Global Rights Department.

Are you the author of this Elsevier journal article?

C2: Open access statement for reproducing figure 30 (Xu K, Chan Y-P, Bradel-Tretheway B, Akyol-Ataman Z, Zhu Y, Dutta S, et al. Crystal structure

of the pre-fusion Nipah virus fusion glycoprotein reveals a novel hexamer-of-trimers assembly.

PLOS Pathogens. 2015;11(12)



C3: Permission for figures 78 and 80

Your confirmation email will contain your order number for future reference.

Printable Details

License Number	5762081439184
License date	Apr 04, 2024

Licensed Content

Licensed Content Publisher	Elsevier
Licensed Content Publication	Current Opinion in Structural Biology
Licensed Content Title	Structural basis of viral invasion: lessons from paramyxovirus F
Licensed Content Author	Robert A Lamb,Theodore S Jardetzky
Licensed Content Date	Aug 1, 2007
Licensed Content Volume	17
Licensed Content Issue	4
Licensed Content Pages	10

Order Details

Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No

About Your Work

Title of new work	Antibody Engineering and Antigen Discovery against Avian Paramyxovirus
Institution name	University of Oxford
Expected presentation date	Apr 2024

Additional Data

Portions	Figure 5

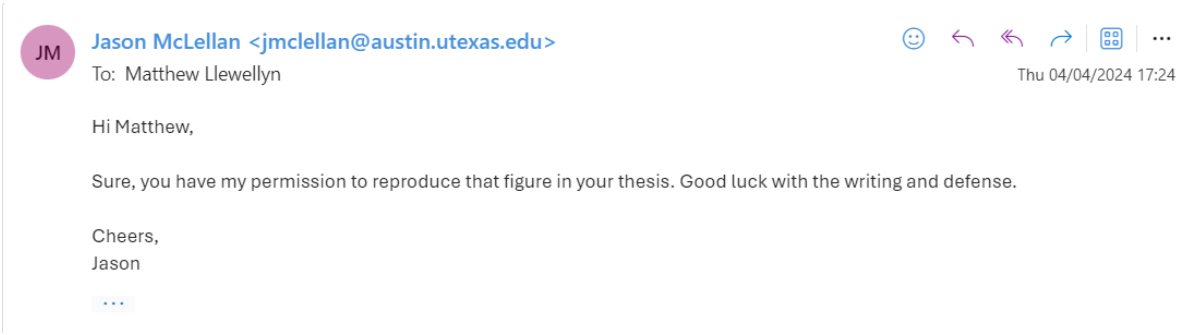
Requestor Location

Requestor Location	University of Oxford Center for Human Genetics Roosevelt Drive Oxford, OX3 7BN United Kingdom Attn: University of Oxford

Tax Details

Publisher Tax ID	GB 494 6272 12

C4: Permission for figure 31



JM

Jason McLellan <jmclellan@austin.utexas.edu>

To: Matthew Llewellyn

Thu 04/04/2024 17:24

Hi Matthew,

Sure, you have my permission to reproduce that figure in your thesis. Good luck with the writing and defense.

Cheers,
Jason