Knots and crosses: The roles of RNA structure and the 3B protein in foot-and-mouth disease virus replication

Joseph Charles James Ward

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds

School of Molecular and Cellular Biology

September 2018

The candidate confirms that the work submitted is their own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

The work in Chapter 3: 3B or not 3B, has appeared in publication as follows Genetic economy in picornaviruses: Foot-and-mouth disease virus replication exploits alternative precursor cleavage pathways, October 2017, Morgan R. Herod, Sarah Gold, Lidia Lasecka-Dykes, Caroline Wright, Joseph C. Ward, Thomas C. McLean, Sophie Forrest, Terry Jackson, Tobias J. Tuthill, David J. Rowlands , Nicola J. Stonehouse

I was responsible for several experiments within this publication as described in the text. Morgan Herod, as primary author, drafted the manuscript and designed the experiments.

Acknowledgements

It is not an understatement to say that my finishing the past four years could not have been possible without the help of a consortium of friends and colleagues pushing me along the way. Thank you to everyone in Garstang 8.64 and 8.71 for making the past four years so enjoyable!

My never-ending thanks goes to my excellent supervisors Prof. Nicola Stonehouse and Prof. Dave Rowlands for their constant patience, advice and understanding! Likewise, this thesis would be a lot shorter if it wasn't for the efforts of Dr Morgan Herod who pushed, bullied and dared me to do more and be better.

The virus work presented in this thesis would not have been possible without the collaboration with the Pirbright Institute. Thank you to Toby Tuthill, Lidia Laseka-Dykes and Sarah Gold for all your hard work.

To the entire 'Stonelands' group, members past and present, for keeping me sane and offering your thoughts on discussions from relevant science, to whether bats have an anus and whether 'Pelumis throat norovirus' is a thing. To this end, I could not forget to mention Jamie Kelly, who is responsible for me even applying for a PhD and is the constant source of fun in the lab, from 'hide the Pelumi' to the notorious Stonelands 'hot-pink' colour scheme. Likewise thanks to Kee and Leith for your brains and the hours spent in front of a beer on a Friday night. Nat.... you are the sole cause of any weight gain in the past year, however, your food is delicious and you have kept me alive! Pelumi! You have been around my entire time in the Stonehouse group, thank you for your constant support, never ending enthusiasm, and questionably useful figures! Finally to Dani, while you are the newest, it feels like you've been around forever, thank you for endless help and company in the past year!

Thank you to all my friends outside of the group; Becky, the past four years couldn't have been done without you, from shopping to drinks you have always been there to help! To Sammy, thank you for putting up with living with me for two years, and sitting next to me for four, your organisation and talent will always be an inspiration. Thanks Kat, for always being around, giving me tasty treats and being a constant cinema companion! Of course, thanks to Dan for being there whenever I want to go climbing... except if it's a Saturday.

This thesis could not have been completed without the efforts of Suki and Kate, who were probably more stressed about it then I was. Without you constantly telling me to go do some work this likely would not be written. Likewise, thank you Captain Latex for always running over when my citations decide to disappear or my error count hits over the hundreds.

To my non lab friends, Sam, Tim and Tom, you have all had to deal with living with me for numerous years. Thank you for your patience and for keeping the complaining to a minimum! Having you around to de-stress has kept me sane.

Last, but not least, thank you to my parents for understanding that I have become a communication hermit and for always supporting me!

Abstract

Foot-and-mouth disease virus (FMDV) is a single stranded RNA virus in the picornavirus family. It is the causative agent of foot-and-mouth disease, globally the most important affliction of cloven hoofed animals. The FMDV genome has several features that are not found amongst other viruses within the Picornaviridae. These include a large $5'$ untranslated region (UTR), almost twice the length of that found in enteroviruses, containing highly structured RNA elements unique to FMDV, such as the S-fragment and several tandemly repeated pseudoknots. Unique aspects are also found within the coding region, where FMDV is the only picornavirus reported to contain multiple copies of the 3B gene. The reasons behind possession of these unusual deviations from the dogma of the picornaviral genome is so far unknown and therefore poses an attractive target for further research.

The S-fragment is a predicted 360 nucleotide stem-loop at the $5'$ end of the FMDV genome. The better studied poliovirus (PV) has a well characterised 80 nucleotide clover leaf structure in the equivalent position which is used as a model for potential S-fragment function. In close proximity to the S-fragment are a series of tandemly repeated pseudoknots, these pseudoknots are maintained across all sequenced isolates of FMDV, however, the number of pseudoknots varies between two and four depending on the specific isolate. As with the S-fragment, no function is yet assigned to these pseudoknots and the reason behind the variation in number remains unknown. Here, biochemical probing of the S-fragment and pseudoknots has been used to confirm the secondary structures of these elements and helped inform mutagenesis experiments to disrupt conformations. Sub-genomic FMDV replicons are used in tandem with infectious virus to help dissect the roles of these elements in distinct areas of the viral lifecycle.

Alongside RNA structure in the $5'$ UTR, this thesis also investigates the non-structural P3 region and how the 3B protein plays a key role in correct polyprotein processing. The complex dynamics of picornavirus polyprotein processing generates intermediate products which expand the repertoire of functions encoded in the genome. Dissecting these processes is therefore key for understanding the details of virus replication. As the polyprotein of FMDV uniquely includes three tandem functional copies of the 3B primer peptide, the manipulation of P3 cleavage sites is used here to further uncover the roles of the 3B protein in replication. The importance of the C-terminal region of 3B3 for correct polyprotein processing is described as well as the capabilities for 3B proteins to accept C-terminal fusions, providing a novel set of molecular tools and approaches to further analyse the complexities of picornavirus genome replication in the future.

Contents

List of Figures

List of Tables

Abbreviations

3B3*2	C-terminal Swap of 3B3 and 3B2
3D-GNN	3Dpol Active site mutation
A1G	*cre* mutation
AR	Androgen receptor
B-	Prefix, Bottom truncation
BHK-21	Baby-hamster kidney cells
BSA	Bovine serum albumin
CL	CLover leaf
CPE	Cytopathic effect
DIVA	Differentiating infection in vaccinated animals
DMEM	Dulbecco's modified eagle's media
EBP1	ErbB3-binding protein
EMCV	Encephalomyocarditis virus
ER	Endoplasmic reticulum
ERAV	Equine Rhinitis A virus
eRF1	Eukaryotic release factor 1
FBS	Foetal bovine serum
FISH	Fluorescent in situ hybridisation
FMD	Foot and mouth disease
FMDV	Foot and mouth disease virus
GARs	Glycyl-tRNA synthetase
GCU	Average mean GFP intensity
H	Hoogsten
HIV-1	Human immunodeficiency virus 1
HPT	Hours post transfection

IF	Immunofluorescence
IRE	Iron response element
lncRNA	Long noncoding RNA
Lpro	Leader protease
MDBK	Madin Derby Bovine Kidney cells
MEM	Minimal essential media
MHC Class I	Major histocompatability complex class I
miRNA	Micro RNA
MNV	Murine norovirus
mRNA	Messenger RNA
ncRNA	Noncoding RNA
NEMO	NF-kB essential modulating protein
NMR	Nuclear magnetic resonance
OIE	World organisation for animal health
P3*	Uncleavable P3 region
PABP	Poly A binding protein
PARS	Parallel analysis of RNA structure
PB	Phosphate buffer
PBS	Phosphate buffered saline
piRNA	Piwi RNA
PV	Polio virus
RHA	RNA helicase A
SAT	South african territories
SHAPE	Selective 2'Hydroxyl Acetylation Analysed by Primer Extension
SRA	Steroid receptor RNA activator
T-	Prefix, Top truncation
TBE	Tris-Borate-EDTA
TIA-1	Cytotoxic granule associated RNA binding protein

TNT *in vitro* Transcription and Translation

uORF Upstream open reading frame

UTR Untranslated region

VLP Virus like particle

VPg Virus protein genome linked

WB Western blot

WC Watson Crick

WT Wild type

Y3F 3B inactivating protein

Chapter 1

Introduction

1.1 RNA Structure

1.1.1 How RNA can fold

Nucleic acids, like proteins, have primary, secondary and tertiary structures. This is true for both DNA and RNA. Primary structure relates to the linear sequence of nucleotides linked together by the phosphodiester bonds, in coding regions this primary structure relates to the production of specific proteins (Fig 1.1A). In untranslated regions, particular nucleic acid motifs can represent sites of interaction with protein or other nucleic acid sequences.

The stabilisation of secondary structure involves base pairing. Watson-Crick base pairing is typically represented in double-stranded DNA, however, single-stranded RNA can fold back on itself and interact with nucleotides in the $5'$ or $3'$ direction on the same strand creating secondary structure. Secondary structures take on a variety of appearances, hairpin loops are formed when an RNA strand folds back upon itself, creating a double-stranded stem and usually a loop of varying size[36,16,202]. Bulges can appear in stem loops when unpaired nucleotides are present on only one side of the stem causing a small break in the double stranded nature of the stem before it reforms again (Fig 1.1B)[218]. Multi-branched hairpin structures can occur when bulges divide stem loops creating an endless supply of possible complex secondary structure organisations[211]. RNA sequences can also engage in non-Watson-Crick

interactions which are thought to be essential for building complex three dimensional structures of large RNAs. Nucleobases consist of three edges capable of forming interactions, the Watson-Crick (WC) edge, the Hoogsteen (H) edge and the sugar edge (Fig 1.2). Two nucleotides are capable of interacting with each other through any of the three edges found in the nitrogenous base in both cis and trans orientation, creating 168 possible interaction conformities[105]. Non-WC base pairing has been studied extensively using experimental approaches such as X-ray crystallography and nuclear magnetic resonance (NMR), as well as computational approaches such as molecular dynamic simulations and it is thought that these interactions, although typically creating fewer hydrogen bonds, are thought to substantially contribute to the RNA structure[83,178,201].

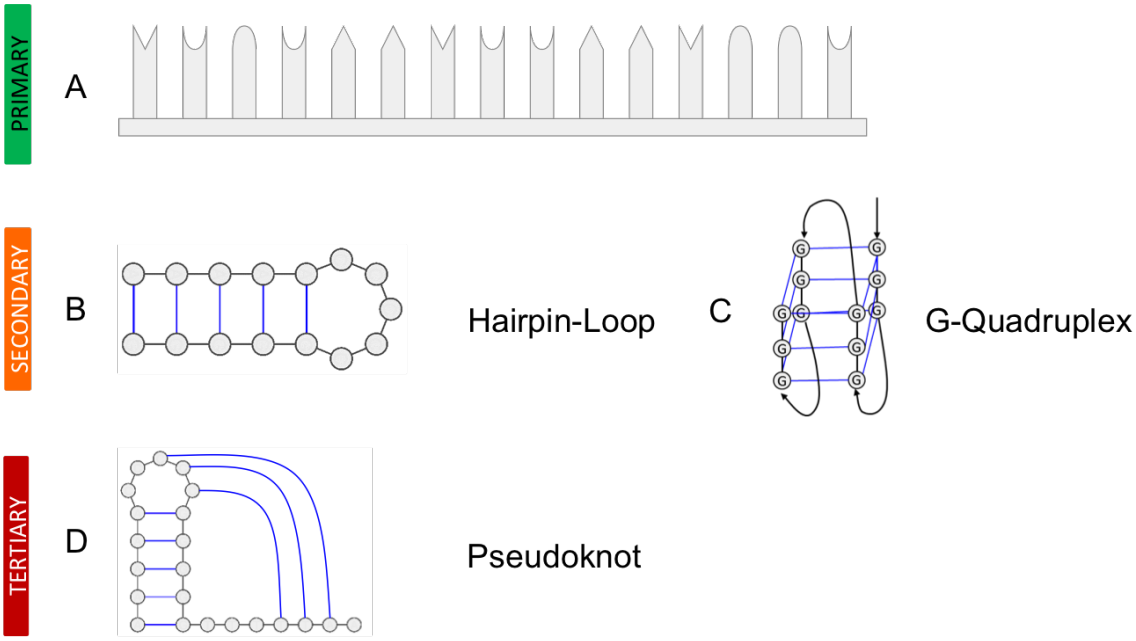

Figure 1.1: Cartoon representation of RNA structures.
Primary structure is shown as the RNA nucleotide sequence (A), secondary structure as a hair-pin loop (B) and G-quadruplex (C), where nucleotide interaction causes base pairing. Tertiary structure is exemplified as a pseudoknot (D) where the loop of a hairpin interacts with nucleotides down/up stream to create additional folding.

Nucleotide sequences containing regions of repetitive guanosines are capable of forming a different structure, reliant on the ability of guanosines to self-associate using

hoogsteen interactions. These structures known as G-quadruplexes can form four stranded helical structures most known as structures found at the end of telomeres (Fig 1C) [28] [137].

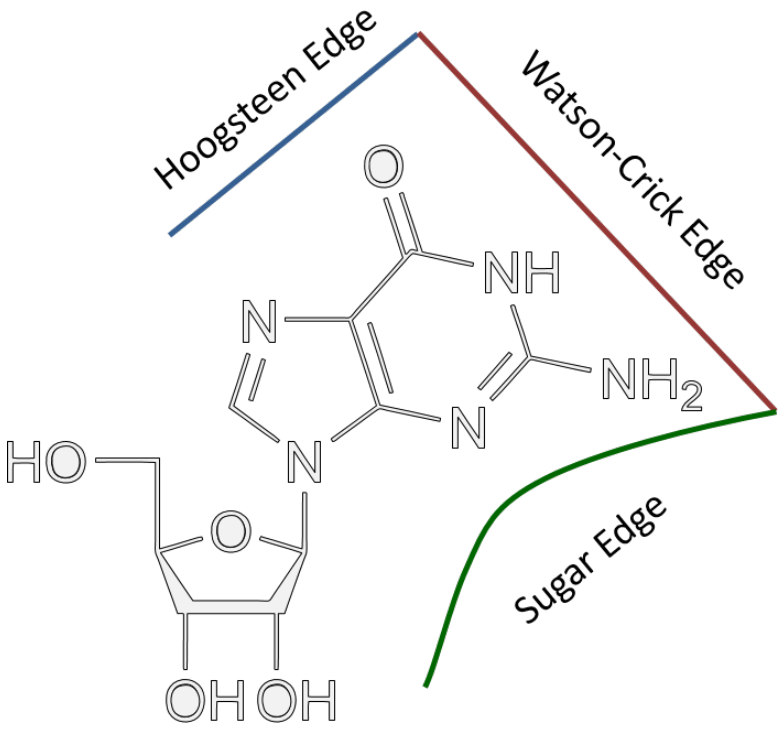

Figure 1.2: Interactive edges of nucleosides.
The three edges of a nucleoside that can perform interactions. Guanosine shown as an example. [35].

Tertiary structures add to the level of complexity of RNA structure, they can include interactions involving two helices, two single-stranded regions or one single-stranded region and a double-stranded helix. Interactions between two helices can occur in two ways, either by a sequential strand stacking on each other or two distant helices position themselves so their shallow grooves fit, creating a helix of helices [206]. Interactions between two single-stranded regions found on loops of hairpins or bulges of stems can lead to the formation of a pseudoknot where a traditional hairpin loop interacts with nucleotides elsewhere causing a complex structure (Fig 1.1D) [26]. Classification of pseudoknots is challenging due to sequence and structural diversity.

Simple classification has been described, categorising pseudoknots into three typical structures, depending on how their tertiary interactions are made. H-type pseudoknots are potentially the simplest type and is also the most widely used term to describe known pseudoknots. These are formed when nucleotides within the loop of a hairpin interacting with nucleotides either down or upstream, causing a folding of the hairpin (discussed further in section 5.1). B-type pseudoknots form similarly to an H-type, except the nucleotide interaction comes from a bulge within a stem, here the unpaired nucleotides interact with single stranded nucleotides up or down stream. Finally I type pseudoknots which have a similar structure to B-type knots, with the folding interaction coming from an internal loop within the stem. As with the other pseudoknots, the nucleotides in the loop can interact with nucleotides either up or down stream (Fig 1.3)[184].

Finally interactions between a single stranded region and a double stranded helix can occur via pairing of the single stranded nucleotides within the deep or narrow groove of a double helix creating a triple helix like structure.

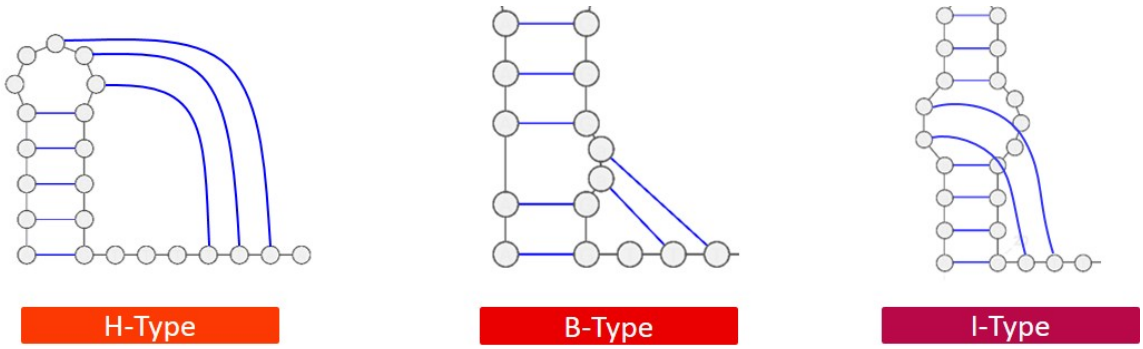

Figure 1.3: Cartoon representation of pseudoknot structures.
Pseudoknot structures are classified here into 3 categories. The H-type pseudoknot, where nucleotide interactions occur in the loop region, B-type where interactions occur in a bulge with nucleotides on the same side and I-type where nucleotides within an internal loop interact with nucleotides on the either side of the stem.

1.1.2 RNA stability

The stability of RNA structures can vary based on the number of interactions involved, this can be key to their function, with conformational flexibility altering their functional state. Predicted stability of RNA structures can be calculated by the free energy difference between folded and unfolded states, often described as the ΔG. The lower the ΔG score the more structurally stable the RNA secondary structure is believed to be. This is due to the amount of energy required to break the interactions and dissolve the secondary structure. Throughout this thesis ΔG will be referred to when talking about the stability of particular RNA structures[122].

1.1.3 How we can visualise RNA folding

RNA structure can be predicted using computational techniques, requiring little else except the sequence of the region of interest. The most common program to do this is called mFold which was developed in the 1980s[224]. This software predicts RNA structure by energy minimization using nearest neighbour energy parameters. Essentially this means that the mFold program utilises algorithms to calculate which nucleotide interactions can form using the lowest amount of energy, i.e the most energetically favourable. To do this, the algorithm compares all possible base pairing outcomes and the energy requirements for this to occur, it then selects the structures requiring the least energy as the most likely. RNA predictions made in this way are becoming more frequently criticised as less informative due to its reductionist approach as it cannot take into account other components that may alter RNA structure, such as protein-RNA interactions and the conditions of the cell, as well as exogenous RNA-RNA interactions.

Biochemical approaches are now favoured more heavily over that of mFold predictions, providing the ability to closely examine nucleotide interactions and model potential RNA structures. Whilst initially these were laborious experiments involving

radioactive nucleotides and large sequencing gels, advances in molecular biology techniques means it has been possible to do so at large scales in relatively short periods of time. Most importantly the development and falling costs of high-through-put sequencing has allowed for the modernisation of the traditional biochemical techniques, with the replacement of slow and tedious analysis of small segments of target RNAs by radioactive sequencing gels with the rapid sequencing of entire genomes with next generation technologies.

Strategies to investigate RNA structure biochemically typically use either specific nucleases or chemical probes, now combined with NGS sequencing[44]. One such method, parallel analysis of RNA structure (PARS), utilises two nucleases, a V1 RNase, which cleaves the phosphate-sugar backbone of RNA at dsRNA points and an S1 RNase which cleaves at single stranded RNA[91]. Target RNA is folded and digested by these RNAses in parallel experiments and following nuclease digestion, RNA is fragmented and libraries prepared for analysis by deep sequencing. By comparing V1 and S1 profiles it is possible to obtain nucleotide resolution of the probability of individual nucleotides being double or single stranded. This data can then be used to model potential conformations of RNA.

Selective 2' hydroxyl acetylation analysed by primer extension (SHAPE) experiments work in a similar fashion, except the nucleases are replaced with a biochemical probe. Many probes exist such as 1M7, DMS or NMIA, these probes act as acetylation agents which react preferentially with the 2'-hydroxyl groups of flexible (unpaired) nucleotides to form a 2'-O-adduct. Once labelled, target RNA is used in a reverse transcription reaction and when the reverse transcriptase reaches a 2'-O-adduct extension is terminated, creating a pool of cDNA fragments[210]. These fragments can then be analysed using high throughput sequencing with the termination nucleotide being recorded as single-stranded. Using these protocols can provide data of single stranded regions to the nucleotide resolution, however, when compiling these with predicted folding algorithms it is not uncommon to see nucleotides which have high single stranded reactivity but are predicted to be base paired. This is

most commonly seen at the proximal regions of bulges where interactions can dissociate allowing nucleotides to become transiently accessible for modification in a mechanism known as "breathing" [84].

The above techniques use biochemical probing or digestion to aid in the prediction of RNA structure, other techniques to investigate this could use the 3D structural techniques such as crystallography and nuclear magnetic resonance (NMR). Crystallography is commonly used in the discovery of protein structure, however, the study of RNA in crystallography can be extremely challenging while, the size of RNA complexes can often be troublesome for NMR. NMR typically has size limitations of around 80kDa, roughly relating to an RNA size of about 350 nucleotides (depending on sequence) [178]. If however, the target RNA is small enough to permit investigation by NMR it is capable of producing incredibly high resolution structures. New techniques of NMR are being developed, pushing the limitations of RNA size, there is still however a long way to go before analysing targets such as whole genomes.

1.1.4 Where RNA structures are found

While the above techniques focus on resolving RNA structure, a key question is where to look for these structures to begin with. Structured RNA has been reported in nearly all forms of life, particularly, in regions of RNA that do not code for protein. These regions can be divided into whole RNAs that are non-coding (ncRNA) and regions that are attached proximally to coding RNAs [149] [86] [130]. These proximal non-coding regions are seen in particular in messenger RNA (mRNA) which possess a tripartite structure (consisting of a protein coding region flanked by $5'$ and $3'$ untranslated regions (UTRs)), as well as in viral genomic RNA (discussed section 1.1.6) [216]. RNA structure is also a key element in ribosomal RNA, comprising approximately 60% of the ribosome its structure is key in the interaction of the protein constituents of the ribosome and ultimately ribosome function.

Hunting for RNA structure at proximal ends of mRNA in untranslated regions or in ncRNA is a logical step, however, essential RNA structure can also exist within coding regions, such as the cis-active replicative elements of picornaviruses (discussed 1.2.8). Whilst protein-coding sequences might be more highly conserved across isolates, ensuring the proteins can maintain their function, the nucleotide sequence of the coding region can vary, while still permitting an unchanged protein sequence. Analysing sequences for regions of high nucleotide homology can suggest an important RNA structure. This has been used to investigate the coding regions of RNA viruses to high effect with several regions identified in viruses such as picornaviruses and hepatitis E virus[58] [52].

1.1.5 Non-coding RNAs

RNA was originally thought to be a "middle man", acting between genomic DNA and protein translation. However, RNA is now known to be involved in many aspects of cellular life, particularly in the regulation of as translational control, splicing, localisation and the control of RNA turnover. High throughput transcriptomic analyses have revealed that eukaryotic genomes transcribe up to 90 % of their genomic DNA. However, only 1-2 % of these transcripts encode for protein, the vast majority instead are transcribed as non-coding RNAs (ncRNA)[11,172,128]. These ncRNAs can be roughly sorted in two groups based on their size, the first group contains short RNAs less than 200 nucleotides in length such as the microRNAs (miRNAs) and piwi-interacting RNAs (piRNA). The second group contains the long ncRNAs (lncRNAs), at a size ranging between 200 nucleotides to over 100 kb in length[149]. Dysregulation of lncRNAs can have extensive downstream effects on cellular functions such as cell proliferation, apoptosis resistance and promotion of metastasis as well as evasion of tumour suppressors. As such, dysregulation of many lncRNAs have been associated with human disease, with particular regard to cancers, such as the up-regulated HULC in hepatocellular carcinoma and the downregulation of

PTCSC3 in thyroid cancers[214]. Although as in the case of these examples the mechanism is not always fully understood[85].

Long noncoding RNAs represent a new field in molecular biology, yet remain the least characterised of all different RNAs. Numerous new lncRNAs are reported every year but the structural data on these is still limited. This is partially due to their generally low expression levels within cells, making obtaining sufficient quantities for structural analysis difficult[132]. Interestingly, unlike their protein counterparts, the primary structure, or nucleotide sequence, displays a large amount of diversity. This is likely due to nucleotide substitutions in coding sequences being more deleterious whilst substitutions in noncoding regions can be tolerated for and secondary or structure maintained by compensatory mutations elsewhere[161]. Indeed, despite sequence diversity, evidence of secondary and tertiary structure conservation is being evermore evident. Unlike the smaller RNA species such as miRNAs whose function mostly derives from their primary structure and base-pairing to their targets, the length of lncRNA provides opportunity for large complicated structures[123]. The first entire human lncRNA structural characterisation was performed on the 870 nt steroid receptor RNA activator (SRA) lncRNA using a complement of chemical and enzymatic approaches[145]. SRA has been shown to participate in nuclear co-activation for many hormone-related systems, including the oestrogen, androgen and progesterone receptors[99,100]. Binding studies of the SRA lncRNA has shown direct binding to nuclear receptor coactivator proteins, RNA helicases and the pseudouridylases Pus1p and Pus3p, although how these interactions take place is unknown[219]. Structural analysis revealed four major domains that were folded into complex secondary structure. Mutation to disrupt this secondary structure reduced activity of SRA by 40 %, supporting the hypothesis that the structure was key to its function and these secondary structures likely play a role in the interactions with proteins[145].

1.1.6 Untranslated Regions

It has long been thought that RNA structures present in the UTRs of mRNA help to regulate protein expression. This control can be through many pathways such as modulation of transport of mRNAs out of the nucleus, control of translation, subcellular localisation and RNA stability[200,10,88]. With the advances in molecular biology techniques, it is now possible to closely examine relationships between RNA structure and the pathways mentioned above.

Both $5'$ and $3'$ untranslated regions of human mRNA have been shown to have RNA structure which is crucial to the regulation of protein expression. The $5'$ UTRs of eukaryotic mRNA has a median length of around 53-218 nucleotides, with the longest length occurring in humans, ranging from a few nucleotides to a few thousand[106].

Translation starts at the $5'$ end of mRNA, where the cap structure and UTR act as the entry point for the ribosome (Fig 1.4). As mentioned previously some mRNAs possess extremely short UTRs and can therefore undergo scanning-free initiation, others with longer UTRs must have the ribosome scan down the UTR until it finds the initiation site[73]. Highly structured regions in the $5'$ UTR have been shown to stop this ribosome scanning by blocking the entry of the 43S preinitiation complex. The best studied example of this is the iron responsive element (IRE), this comprises a small (30 nucleotide) RNA hairpin-loop found in the $5'$ UTR of a subset of mRNAs important for the storage or transport of iron. The IRE structure acts in a feedback mechanism and in low iron conditions can be seen bound by iron-regulator protein 1 and 2[138]. Once these regulatory proteins are bound, translation initiation is prevented by steric inhibition of the 43S complex, in an iron rich environment the protein-RNA interaction is lifted and translation allowed to continue[77].

Structures in the $3'$ UTR have been shown to regulate RNA stability, and indeed IRE structures in the $3'$ UTR can enhance the RNA stability of mRNAs mentioned above in times of iron surplus, permitting the mRNA to persist longer and produce more

proteins. In a similar mechanism, one cause of the blood disorder, β-Thalassemia, is mutation of the terminating codon in the β-globin gene, this results in the reduction of haemoglobin production. This reduction is caused by the prevention of ribosome termination, causing the ribosome to progress into the 3′ UTR. Here it 'masks' secondary structures with determining the stability of the globin mRNA. With these structures masked mRNA stability is reduced and globin levels decrease, causing the onset of disease symptoms[154].

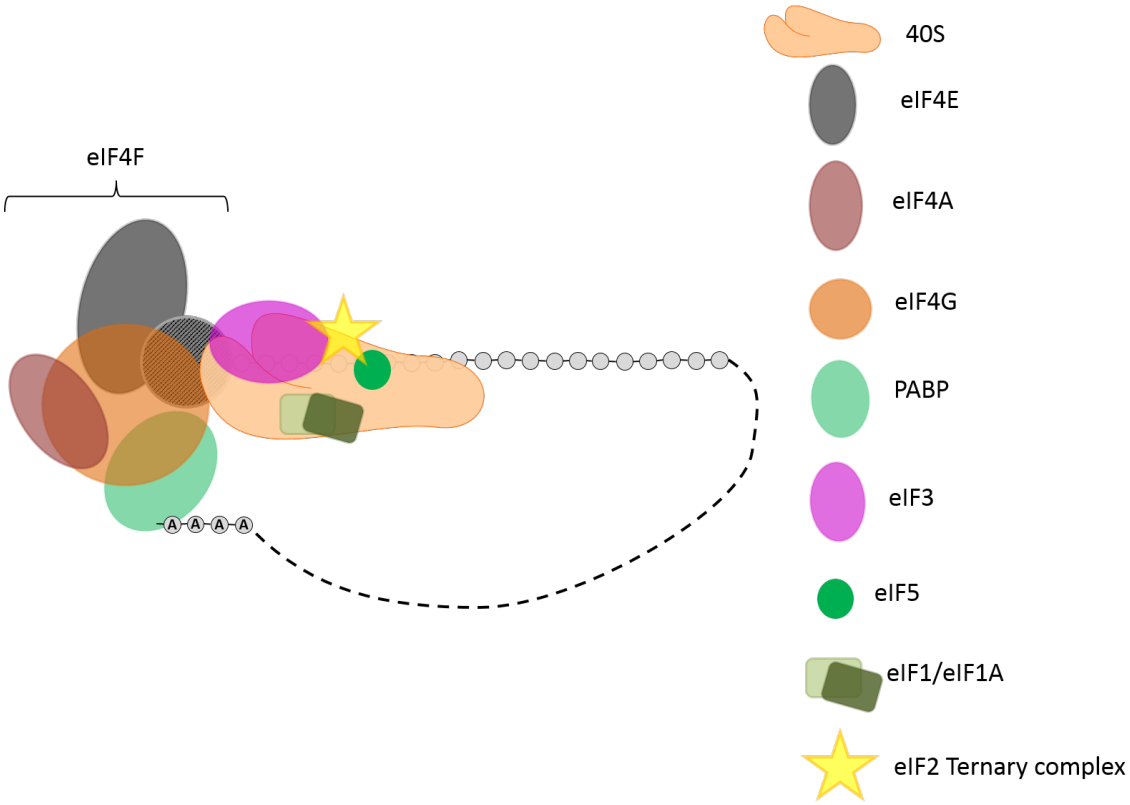

Figure 1.4: Initiation of cap-dependent mRNA translation.
Cap-dependent translation of mRNAs occurs via interaction of several cellular proteins. The eIF4F complex, consisting of eIF4G, 4A and 4E recognise the m7G cap of mRNA. The poly-A-binding protein (PABP) interacts with the poly-A-tail of mRNA and with eIF4G, causing circularisation of the mRNA. The 43S pre-initiation complex (40S ribosomal subunit, eIF2 ternary complex, eIF1, eIF1A and eIF5) is recruited to the mRNA by eIF3. Upon recognition of the AUG start codon assembly of the 80S ribosome occurs and elongation begins[129].

1.1.7 Untranslated regions in viral RNA

Cellular RNAs are not the only UTR-containing RNAs. It has long been known that viral RNAs possess non-coding regions, many of which are highly structured. Like those of cellular mRNAs the size of these can vary with some being only a few nucleotides, to others being kb in length, taking up significant lengths of the genome[31,117]. This has been well studied in human immunodeficiency virus -1 (HIV-1), the 5′ UTR of HIV has been shown to be highly structured, one such structure the named TAR interacts with both viral and cellular proteins, mutations to the TAR structure have shown to severely impair HIV translation and viral production, showing the essential nature of this RNA structure to the viral lifecycle[74]. RNA structure plays a particularly important and well defined role in the translation of uncapped RNA viruses. Eukaryotic mRNA usually possess a 5′ cap recognised by eIF4E , recognition by eIF4E permits recruitment of the rest of the eukaryotic translation machinery to the mRNA allowing ribosome scanning and the initiation of translation. Uncapped RNA viruses can not initiate translation by this method, but use instead a collection of highly structured RNA hairpin loops to recruit translation machinery. These are known as internal ribosome entry sites (IRES) and have been reported in both viral UTRs as well as more recently in cellular mRNA[98,120]. During infection, viral and cellular mRNAs are always competing for the attention of the cellular translation machinery, viruses often circumvent this competition by modifying eIFs to reduce the translation of mRNAs. Such modification often comes in the form of targeted cleavage of one or several eIFs by virally expressed proteases[14]. For many years it was a mystery to how viruses could inhibit the translation of mRNA while still permitting the expression of their own proteins. The discovery of IRES elements in the 5′ UTR of the human picornavirus, poliovirus (PV) revealed that the PV 5′ UTR could recruit ribosomal 40S subunits involving a reduced set of eIFs, importantly disregarding the need for the cap binding protein eIF-4E[155]. Finding this alternative route of translation has provided many uncapped RNA viruses with the competitive advantage over their host cell. Not all IRESs are the same, with mul-

tiple different conformations been described even within a single virus family such as the picornaviruses. Four IRES types are described here with the enteroviruses using a 'type I' IRES, the animal cardio and apthoviruses using a 'type II' structure, Hepatitis A virus using a 'type III' structure and the remaining picornaviruses use a mixture of type I, type II and hepatitis C virus-like IRESs. All types differ in their primary and secondary structure, with differences in the number, appearance and sequence of stem loops. While all these structures ultimately allow the initiation of translation, they may bring this about by differential interaction with host cell factors[120]. For example type I IRES structures require the glycyl-tRNA synthetase (GARs) and La protein, whilst type II IRES require neither but do interact with gemin5 and the ErbB3-binding protein (EBP1) (Fig 1.5). Whilst IRES – protein interactions are key for the initiation of translation, it is reported that IRES elements do not solely interact with proteins, long range RNA-RNA interactions have also been reported which can modulate IRES activity[64]. For instance, the FMDV IRES has been shown by radiolabelled electro-mobility shift assays to interact specifically with both the FMDV 3′untranslated region and poly-A-tail, both of which independently enhance IRES activity[115].

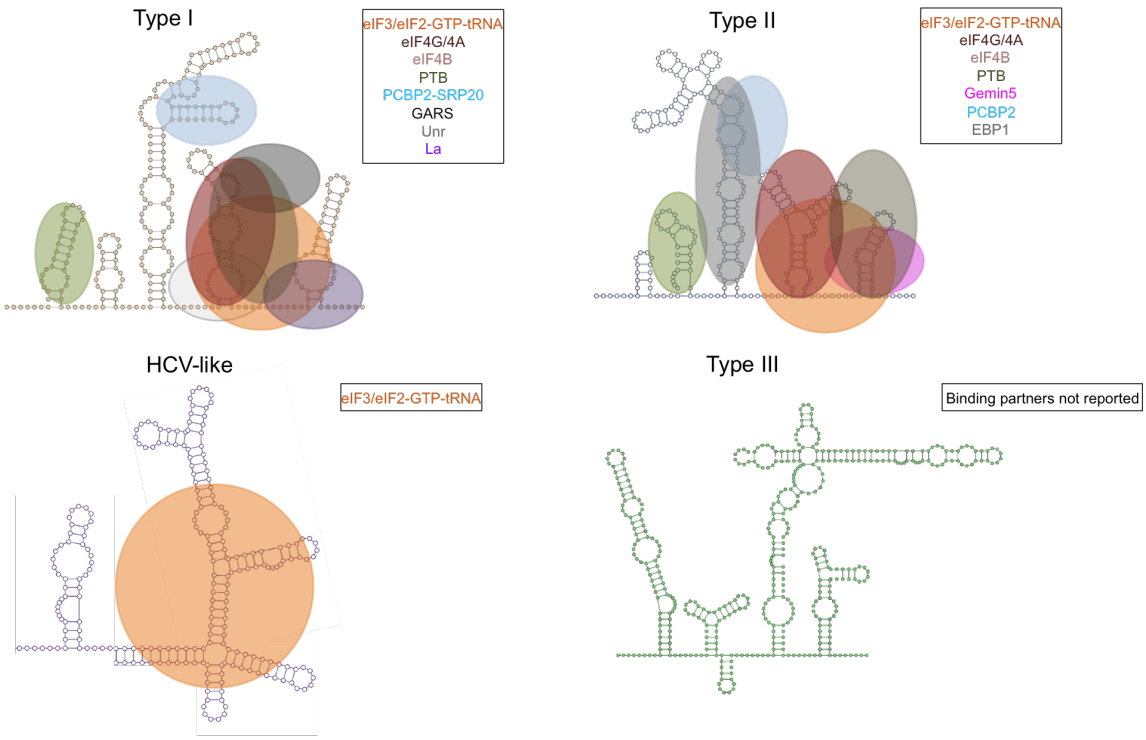

Figure 1.5: Cap-independent translation by viral IRES motifs.
Four types of picornavirus IRES are shown, type I, II, III and HCV-like. Host-cell proteins interacting with IRES components are depicted by coloured shapes and listed to the right hand side of each IRES. Type I and II are well characterised with many cellular proteins known. HCV-like and Type III are not as well investigated, leaving few to none experimentally shown interacting proteins [121]

1.1.8 uORFs

Untranslated regions do not solely contain RNA structural elements, ribosome profiling of many UTRs has revealed translational activity upstream of main mRNA open reading frames in as much as 40 % of mammalian mRNAs [183,150]. It is thought that some of these upstream open reading frames (uORFs) as well as encoding proteins, can also play a role in inhibiting the translation of downstream ORFs by causing ribosome stalling due to RNA structure, codon usage or structure of peptide encoded by the uORF [217,135]. Ribosome termination can also occur by uORFs positioning out of frame with the downstream coding region causing early termination after start

codons, likewise, even inframe uORFs can cause termination of downstream translation by presence of the a stopcodon at the end of their reading frame. Upstream ORFs have also been reported within viral genomes, such as in the Ebola virus $5'$ UTR, where the presence of an uORF has been shown to modulate the translation of the viral polymerase in response to cellular stress[180]. Recently, discovery of an uORF has also come to light within the picornavirus genome, dramatically changing the long-thought dogma of picornavirus virology, with experimental evidence of a novel protein produced in the late stages of poliovirus infection in organoid cell lines (personal communication, Andrew Firth, University of Cambridge). These recent discoveries show that we should not dismiss untranslated regions to be, as their name suggests, untranslated. While secondary structure does dominate these regions, hidden reading frames encoding proteins can exist with critical functions to the viral lifecycle.

1.2 RNA structures and Viruses

As described above, RNA structure has been shown to be essential for many aspects of viral lifecycles, from translation to genome replication as well as functions in genome packaging and mature virus production. One virus family that has been particularly well studied is the Picornaviridae a family of positive-sense single stranded RNA viruses.

1.2.1 Picornaviruses

As of February 2017 the Picornaviridae consists of 94 species grouped into as many as 40 genera as stated by the international committee on taxonomy of viruses. This diverse family of viruses comprises significant animal and human viruses, causing a wider range of illnesses than most, if not all other virus families.

The term picornavirus is derived from the word 'pico' meaning small, and RNA, referring to the genetic material making up its genome. All picornaviruses comprise a single-stranded positive sense RNA genome of between 7.5 and 8.5 kb in length enclosed within a classically non-enveloped icosahedral capsid measuring between 18 and 30 nm in diameter[197]. The most well studied members of the Picornaviridae include poliovirus (PV), human rhinovirus, hepatitis A virus and foot-and-mouth-disease virus (FMDV). Despite being such a large and diverse family, picornaviruses share a similar genome organisation, with a single open reading frame encoding the viral structural and non-structural genes flanked by $5'$ and $3'$ untranslated regions (UTR) (Fig 1.6). While the general genome organisation might be similar, there are many differences at both the protein. Examples include the inclusion of extra or duplicated proteins seen in some viruses, and different RNA structures being present as well as different locations of conserved structures.

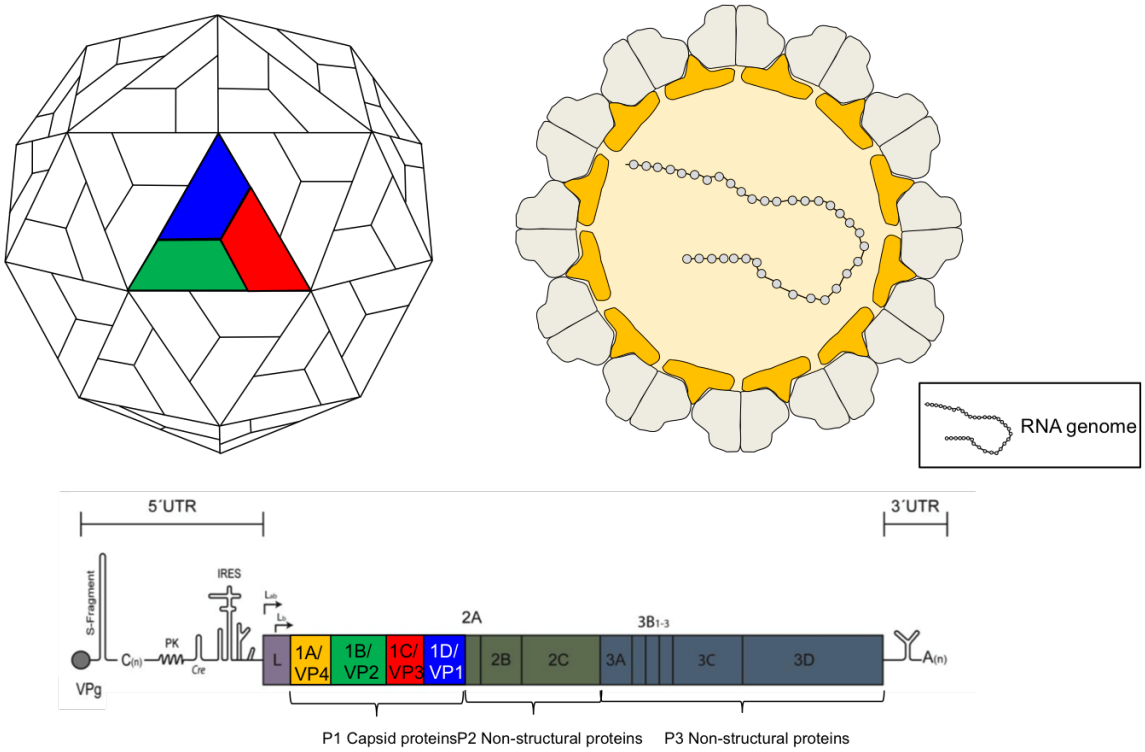

Figure 1.6: Picornaviral virus capsid and genome.
Picornavirus possess icosahedral capsids, which are arrangements of 60 protomers, each consisting of one copy of VP1, VP2 and VP3. VP4 forms an internal lattice providing capsid stability. The picornavirus genome, exemplified here by FMDV, consists of a single open reading frame, encoding P1 capsid proteins and P2 and P3 non-structural proteins. The open reading frame is flanked by 5′ and 3′ UTRs[197].

1.2.2 Aphthoviruses

Aphthoviruses are a genus of animal picornaviruses comprising FMDV, bovine rhinitis A and B virus and equine rhinitis A virus (ERAV). FMDV, is arguably the most well-known virus of this genus, with outbreaks having occurred in every country worldwide with the exception of New Zealand (Fig 1.7). FMDV is a highly infectious virus causing disease of the same name (foot and mouth disease, FMD) in both domestic and wild cloven hoofed animals. The list of susceptible animals is long, including cattle, swine, sheep and deer as well as non-domestic animals such as wildebeest[96]. Outbreaks of FMDV can have severe economic consequences, with

the 2001 outbreak in the UK leading to the death of approximately 10 million animals and an overall economic cost of at least £8.6 billion[97]. Consequences of FMDV outbreaks are so severe that the world organisation of animal health has placed FMDV in its list A, reserved for diseases which can have worldwide implications upon outbreak. Despite the vast publicity FMDV gets it can still be found endemic in many parts of world, particularly in lesser economically developed countries such as India.

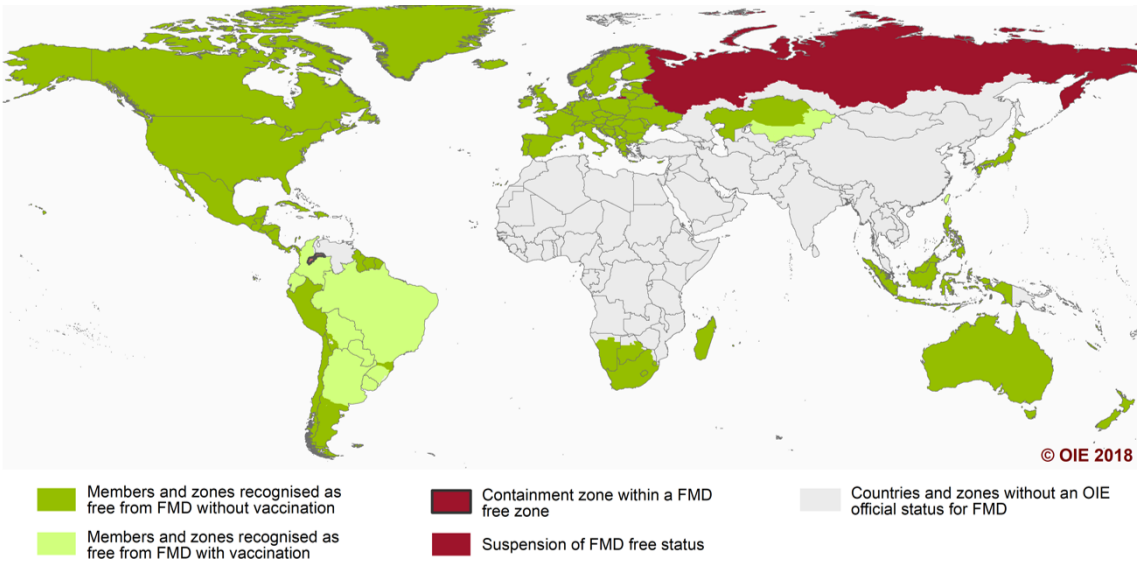

Figure 1.7: FMDV official status map.
Map of countries defined as, free of FMD with or without vaccination, contained FMD disease, or no official status, as determined by the world organisation for animal health (OIE) - 2018. Figure adapted from the OIE website.

1.2.3 Clinical aspects of FMDV

Infection with FMDV was first described by an Italian monk in Venice in 1514. It was noted that infected animals would not eat and showed vesicles in their oral cavity and feet. The clinical symptoms of FMD are now well defined by fever, shivering, excessive saliva production and the formation of blisters or vesicles on the oral mucosa and epithelium of the tongue, nose and coronary bands[70]. FMDV is

capable of rapidly replicating and some infected animals can show symptoms as early as 24-48 hours post-infection. For most animals FMD is not fatal, with a mortality rate of approximately 5 %, however, infection can result in weight loss and a decrease in milk production resulting in a loss of agricultural productivity. Furthermore, the mortality rate is significantly higher in young animals, where mortality rate of approximately 50 % are seen. Due to potential losses to the agricultural industry that an outbreak of FMDV can have, general practice is for disease control, resulting in the slaughter of both infected and possibly infected animals to limit the spread of the disease.

Cases of FMDV infection in humans has been reported, however, these cases are extremely rare and most reports are attributed to laboratory accidents. Due to the sensitivity of FMDV to stomach acid, it is unlikely to be transmitted through the consumption of infected meat.

1.2.4 FMDV Outbreaks

Due to the high infectivity of FMDV, outbreaks are common amongst unprotected animals and spread rapidly. Interestingly the spread, indicated as the number of animals the disease can spread to from a single infected animal (R0) of FMDV can vary amongst susceptible animals as well as between animals of different species. Sheep to sheep transmission has been reported to have an R0 of around 1.1, which is relatively low, with a similar score seen with sheep to cattle spread, cattle to cattle transmission, however, has been shown to have R0 scores of between infinity and 2.4. While studies have shown the susceptibility of these animals to FMDV is similar, comparing the R0 of cattle – cattle and sheep to cattle indicates that cattle are more infectious than sheep, and as such likely play a major role in the transmission of FMDV in a mixed population[24]. While not much data exists on the transmission dynamics of FMDV, investigation into this would be interesting to observe if any differences between serotypes and outbreaks show differences in the

rate of transmission, this data could be used to help inform FMD spread models, advising control strategies for different animal species and virus serotypes.

Outbreaks of FMDV can have major global consequences, high impact cases include the outbreak in the UK in 2001 as well as a large outbreak in Japan and Korea in 2010. Epidemics are recurrent in areas such as South Korea which, despite being FMD free for 66 years, since 2000 are now reporting cases of FMD almost annually[204,151]. South Korea had its last large outbreak in 2015, which led to the slaughter of over 50,000 pigs on infected farms. This outbreak has occurred despite a government policy to vaccinate against FMD since 2011. While this policy is no doubt preventing a widespread outbreak of FMD it does reveal a problem with the current vaccine procedure. All vaccines currently in use are composed of inactivated FMDV particles. An issue with these vaccines is that should vaccinated animals fail to develop proper immunity they could develop sub-clinical infections and later develop FMD as well as excrete large amounts of virus that can overcome the immunity of other animals. This has proven to be a particular problem in swine, where it has been hard to provide complete protection, especially if they have been in contact with infected animals. This causes a repetitive cycle where they can then be a source of disease and infect other animals[29].

Whilst the inactivated virus vaccines are available, many countries choose to forgo vaccination for several reasons, including cost as well as the logistical difficulties of organising the vaccination of animals dispersed over the countryside as well as the near impossible task of vaccinating wild animals who may roam freely across entire continents acting as a reservoirs of infection.

1.2.5 FMDV Serotypes

An additional challenge to FMD vaccination schemes is the multiple serotypes of FMDV, some of which can be found co-circulating within geographical locations. FMDV is divided into seven serotypes, O and A, which are widely distributed worldwide, South African territories (SAT) 1, 2 and 3, normally restricted to Africa and Asia 1 and C serotypes, commonly found in Asia[70]. Each of these serotypes has numerous subtypes and infection or vaccination of an animal with one serotype does not provide immunity against any of the others, indeed vaccination may even be only partially effective against some subtypes within the same serotype. Therefore making vaccination more troublesome, as to be fully protective vaccines must be cross-protective against multiple serotypes.

1.2.6 FMDV Vaccines

Whilst inactivated virus vaccines are available for the prevention of FMDV infection they are only economically viable in regions and farms endemic to FMDV. Combined with the reasons stated above, as well as the short lived immunity provided by these vaccines, it is not a surprise that many countries choose to forgo vaccination.

Other issues surrounding current vaccines is the need for high containment facilities in the production process. As current vaccines are inactivated, large amounts of virus must be produced by cell culture before being inactivated and formulated into a vaccine. This is problematic at two levels, firstly, the use of live infectious virus in vaccine production possesses its own risks, as accidental release of live virus from these facilities can lead to devastating outbreaks. Secondly despite recent improvements, it is difficult to get complete purification of virus from cell culture, often resulting in high concentrations of contaminating non-structural proteins[29]. This becomes an issue as vaccinated animals can raise antibodies against both structural and non-structural proteins, making differentiation of vaccinated and infected ani-

mals difficult. Differentiation is especially important for FMDV because, as stated above, the potential for subclinical carriers of virus in vaccinated animals. This problem is being targeted by the production of new viral strains with specific markers to determine between vaccinated or infection, these strategies for differentiating infection in vaccinated animals (DIVA) can include the introduction of specific sequences within the genome, or exclusion of non-essential non-structural genes, like for example one of the multiple copies of 3B.

New vaccines that do not require the use of infectious virus are in development, including a protein vaccine using the VP1 capsid protein. However, when tested in cattle, the vaccine only produced limited protection against viral challenge[21]. Virus like particle (VLP) vaccines are also in development with capsid proteins being expressed in both bacterial and insect expression systems before assembly into a capsid-like structure and show promise in the production of FMDV-specific antibodies[215,190]. Live-attenuated vaccines are also in development and have shown evidence of good protection in initial trials. However, there is always a concern when using attenuated virus vaccines in the risk of reversion to virulence and the cause of disease, as well as the fact that similarly to inactivated virus vaccines, large amounts of cell culture must be used.

For the reasons described here improvement in FMDV vaccine strategy is crucial for future protection against outbreaks as well as to help control currently circulating FMDV.

1.2.7 FMDV Genome Structure

While FMDV is not the most studied picornavirus, the overall genome organisation of all picornaviruses is fairly conserved. This allows for any regions not investigated specifically in FMDV can have functions hypothesised by looking at better studied picornaviruses such as PV.

FMDV has a standard genome split into three regions, a single open reading frame flanked by 5′ and 3′ UTRs preceding a large poly-A-tail (Fig 1.6 and Fig 1.8). The FMDV genome is slightly larger than usual for a picornavirus at 8.5 kb, this is partially due to its unusually large 5′ UTR, covering greater than 1/7 of the entire genome[61].

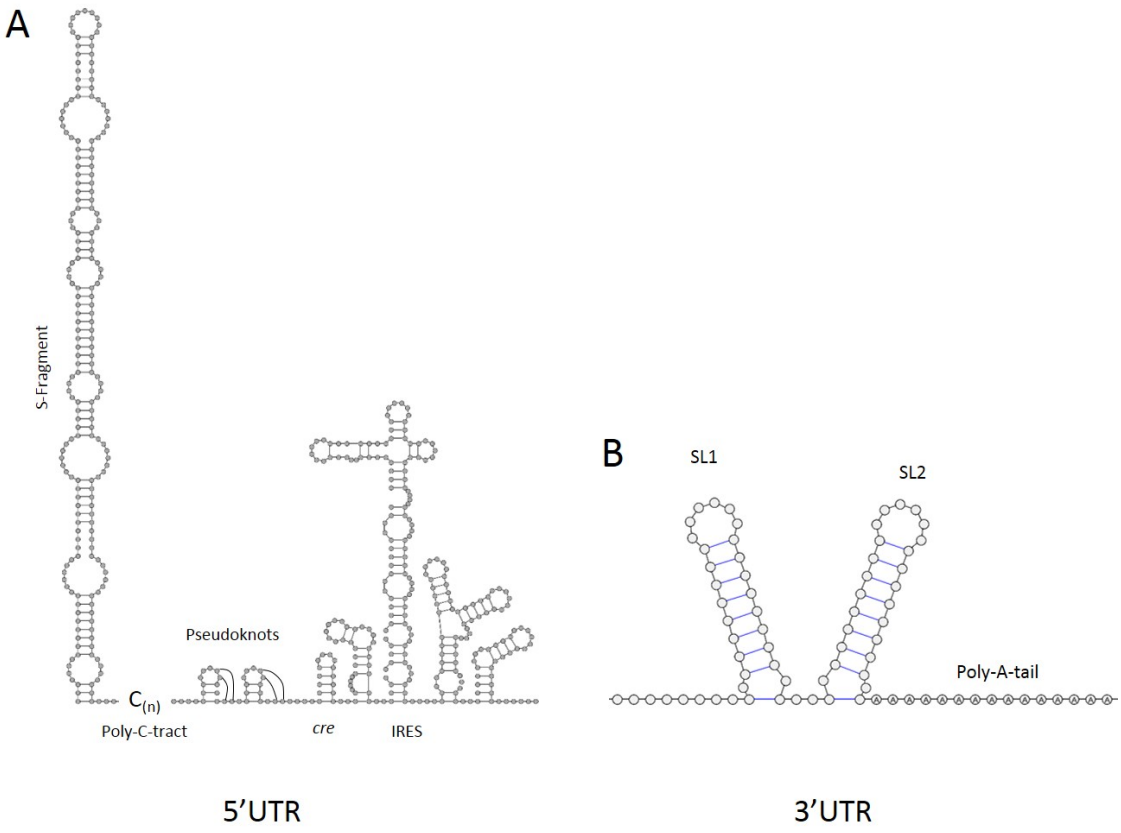

Figure 1.8: FMDV untranslated regions.
(A) The 5′ UTR of FMDV contains a large stem-loop S-fragment followed by a poly-C-tract of variable length, 2-4 repeated pseudoknots, *cre* and type II IRES. (B) The 3′ UTR contains two small stem loops (SL1 and SL2), as well as a poly-A-tail[61]. Structures not shown to scale for clarity.

1.2.8 UTR of FMDV

As described above untranslated regions have been shown to have high degrees of RNA structure in both eukaryotic mRNA and viral genomes. This thesis will discuss the UTR of the picornavirus, FMDV. The $5'$ UTR of FMDV is unique amongst its picornavirus relatives, being approximately 1.3 kb it takes up 1/7 of the entire genome. Within this large region are many highly structured RNA elements, several of which are unique to FMDV. At the $5'$ most end is a 360 nucleotide region predicted to fold into a single large hair-pin loop termed the S-fragment. Although this structure is largely conserved within FMDV isolates, there is a surprising amount of sequence variability, with only 12 % of the S-fragment sequence being conserved [31]. Due to its position in the genome investigation into the function of the S-fragment can be problematic, indeed simply obtaining the full sequence has proven challenging. Although known to be essential for viral replication, no defined function has been attributed to this RNA structure. Through protein-RNA interaction experiments, one cellular protein has been found to bind specifically to the S-fragment, RNA helicase A (RHA). It possess two RNA binding domains and contains a classical DEAD box helicase domain [101]. RHA is typically observed shuttling between the nucleus and cytoplasm permitted through the action of an encompassed nuclear transport domain [62]. RHA functions as a helicase and was first reported to unwind double stranded DNA, althought activity in both DNA and RNA unwinding is now reported [104]. Investigation into RHA in FMDV infection shows a re-localisation into the cytoplasm as well as specific binding activity to the S-fragment [102]. Due to the largely double stranded nature of the S-fragment, it is likely that RHA plays an essential role in the unwinding of this region to allow for the viral polymerase to carry through this region.

Immediately $3'$ to the S-fragment lies the remaining $5'$ UTR, starting with a large poly-C-tract of variable length, due to the repetitive nature of the sequence it is difficult to get a reliable figure, but is generally considered to be around 200 nucleotides

of mostly cytosine residues[133]. Whilst not the only picornavirus to possess a poly-C-tract with both encephalomyocarditis virus (EMCV) and equine rhinitis A virus (ERAV) also reported to possess a similar region, no function has yet been described to the poly-C-tract of FMDV[30]. Interestingly, in animal studies deletion of the EMCV poly-C-tract showed complete viral attenuation in mice, whilst truncation of the FMDV poly-C-tract showed only a slight decrease in viral replication with disease still evident in infected mouse models[50].

After the poly-C-tract are a series of two to four tandemly repeated small pseudoknots, the number of pseudoknots found in the $5'$ UTR depends upon isolates, but no virus to date has been reported to consist of less than two or greater than four pseudoknots[38]. Like the poly-C-tract, no function has been attributed to the pseudoknots. Following the pseudoknots is the well-studied *cis*-acting replicative element, *cre*. The *cre* forms a small stem loop structure containing a critical conserved AAACA motif in its top loop[68]. The *cre* has been reported in nearly all picornaviruses reported so far and has been well characterised to be essential in the priming of new strand synthesis and genome replication. The position of the *cre* within picornavirus genomes is fluid, with reports of this stem loop structure being found in the 2C gene of PV, the capsid proteins of cardioviruses and within the 2A protein of rhinovirus[112,68,165]. Studies into the FMDV *cre* have confirmed that the location of this RNA structure is flexible, with replication tolerant to re-localisation of the *cre* to the $3'$ UTR, however, deletion of the *cre* or mutation of its conserved AAACA motif is always fatal[140]. The final structure in the $5'$ UTR is the IRES, FMDV belongs to the category of viruses described earlier, possessing no $5'$ cap and instead relying on the IRES RNA structure to permit viral translation[13] [49]. Possession of the IRES allows for the continuation of the translation of FMDV proteins after the cleavage of eIF4G by viral proteases, preventing its interaction with the cap binding protein eIF4E[89].

1.2.9 FMDV structural and non-structural proteins

1.2.9.1 Leader protease

Following the IRES the first ATG start site for translation initiation can be found. Unusually amongst picornaviruses FMDV contains two AUG start sites, producing two variations in the first protein, the Leader protease (L^{pro})[168,111]. In defiance of the well-studied enterovirus genome structure, the first protein found in the FMDV genome is not a structural protein, but a non-structural protease unique to the apthoviruses. The leader protease cleaves itself away from the viral poly-protein upon translation and plays a role in the promotion of translation by inhibition of host cell protein synthesis by cleavage of the cellular eukaryotic initiation factor eIF4G, preventing its interaction to the cap-binding protein eIF4E[46]. The two start sites noted within this protein produce two different sizes of L^{pro}, the full length Lab^{pro} and a truncated form 28 amino acids shorter 'Lb^{pro}'[39]. Investigation into L^{pro} has shown that deletion of the first AUG is lethal to virus replication, but mutation of the second AUG produced viable virus, although a slower growth rates of this virus have been reported. Apart from Lab^{pro} and Lb^{pro}, another form of L^{pro} has been observed, named sLb^{pro}, this protein represents a truncated form of Lb^{pro} with six or seven residues removed from the C terminus[71]. The role, of sLb^{pro} has not been elucidated, although initial observations suggest it may have an altered substrate specificity compared to Lb^{pro}. Investigation into the virulence of viruses lacking the entire leader protein revealed remarkable attenuation in cattle, with no clinical symptoms becoming apparent and no sign of animal-animal spread, however, replication in tissue cultures revealed only a moderate reduction replication efficiency[27]. This attenuation achieved by the removal of L^{pro} has been used in the generation of new-age attenuated virus vaccines.

1.2.9.2 P1 Structural proteins

Downstream of the Leader protease is the P1 region. This houses the genes (1A-1D) that encode the 4 structural proteins (VP1-4) required to construct the mature icosahedral viral capsid. These proteins are auto-catalytically cleaved from the viral polyprotein by the action of the 2A protein. The capsid is composed of 60 copies of each of the four structural proteins. However, upon translation, the P1 proteins form into protomers containing equal amounts of VP1, VP3 and VP0 (the precursor to VP2 and VP4)[70]. Five protomers assemble into a pentamer and 12 pentamers into the icosahedral capsid. VP0 cleavage into VP2 and VP4 is only reported to occur upon the encapsidation of the genome, although the exact mechanism of this is currently unknown (Fig 1.6)[42].

1.2.9.3 Non-structural proteins: P2 Region

Proceeding P1 are the P2 and P3 regions containing the remaining non-structural proteins. The P2 region contains three proteins, 2A, 2B and 2C. While 2A is considered part of the P2 region, during translation it induces ribosome skipping to allow detachment of the P1 region away from the rest of the translated polyprotein, as this occurs the 2A protein remains attached to the separate P1 region[176]. It is this action that is thought to account for the excess of translational products upstream of 2A. As The 2A of FMDV is much smaller than that of that found in the enteroviruses, at only 18 amino acids in length compared to the 150 amino acids in its enterovirus analogues. The 2A of PV has been well characterised, and apart from the release of P1 has been shown to have more roles in the viral replication cycle, with functions akin to that of the FMDV leader protein[32]. The proteolytic function of PV 2A has been defined, with a catalytic active site key to its function, however, due to the small size of the FMDV 2A, an alternative mechanism of cleavage has been adopted. Investigation into the activity of FMDV 2A has revealed that this alternate cleavage is operated through a translational event termed "stop-carry on"[48]. This is driven

by the C-terminus of the 2A protein between the last amino acid of 2A (gly) and the first of 2B (pro)[48]. Once the ribosome reaches this point, the prolyl-tRNA enters the ribosome, but is un-able to form the peptide bond and subsequently exits, this allows the access of eukaryotic release factor 1 (eRF1) which releases the nascent peptide by hydrolysis of the ester linkage. Two outcomes can continue from this point, the termination of translation, or the re-entry of prolyl-tRNA allowing the continuation of translation of the 2B protein. Exploitation of this mechanism has been used in many aspects of molecular biology, particular where the expression of two proteins is desired from one expression construct[175].

Following 2A are the 2B and 2C proteins, during catalytic processing of the translated polyprotein, before the final mature proteins are formed, a variety of processing intermediate precursors are produced. Some of these have known functions whilst the roles, if any, of others are not known. The P2 precursor protein, 2BC, has been shown by immunofluorescence microscopy and co-localisation studies to block host protein endoplasmic reticulum (ER) – Golgi transport, resulting in cytopathic events within FMDV infected cells, a similar role is prescribed to the PV 3A protein[186] [134].

Upon complete processing the mature 2B is produced, this protein has been shown to enhance membrane permeability and is thought to function as a viroporin. It contains two putative transmembrane domains characteristic of viroporins, supporting the predicted function[3]. 2B proteins of the better studied PV have been shown to localise to the ER, thought to be the site of replication of PV[174]. It is possible that 2B plays an essential role in the reorganisation of cellular membranes for the creation of sites of viral replication. Recent observations have shown an interaction between 2B and the immune modulating protein RIG-I[222]. This interaction was specific to the FMDV 2B and was not seen in investigation of enteroviruses such as enterovirus 71 and coxsackievirus A16. Suggesting that the FMDV 2B has additional functions in the downregulating of RIG-I, complementing the similar actions of both the leader and 3C proteins.

The final protein in the P2 region is 2C, like 2B, the functions of 2C are not well defined. The architecture of FMDV 2C exemplifies the common format of picornaviral 2C proteins, with a predicted N-terminal amphiphatic helix, a feature shown for other picornaviruses to be essential to bind to cellular membranes and thought to be required for membrane rearrangement and formation of membrane bound replication complexes[191,51,17]. Following this is the ATPase domain, 2C has 3 Walker motifs resulting in its classification as an AAA+ (ATPases associated with various cellular activities) protein, alongside other viral helicases such as the SV40 large T antigen and bovine papillomavirus E1[69]. The classification of FMDV 2C in this category implies a putative function as an RNA helicase, and while ATPase and RNA binding functions have been reported, no helicase activity has yet been shown for any picornavirus 2C. PV 2C has been shown by UV crosslinking to bind to viral RNA (although specificity of 2C RNA binding is still unclear), and an RNA binding domain has been annotated onto its structure, however, no such region has yet been described for the FMDV 2C[8]. RNA binding studies are complicated by the poor solubility of recombinant protein and only recently has FMDV 2C been reliably produced by truncation of the N-terminal amphiphatic helix allowing an increase in solubility[188]. As well as putative RNA binding actions, the FMDV 2C protein has been shown to interact with host cell proteins such as regulators of the autophagy pathway such as Beclin 1. Overexpression studies of Beclin1 in virally infected cells shows a strong decrease in viral titre, alongside evidence showing that the fusion of lysosomes to autophagosomes is not seen during FMDV infection, suggesting an activity of 2C to bind to Beclin 1 and prevent the fusion of lysosomes[66].

1.2.9.4 Non-structural proteins: P3 region

The P3 region contains the final series of non-structural proteins, 3A, 3B, 3C and 3D. There are two suggested processing pathway for this region, the major and minor pathways (Fig 1.9)[110]. The major pathway cleaves 3ABCD precursors into 3AB and 3CD, whilst the minor pathway procures 3A and 3BCD, 3BCD is then further processed into 3BC and 3D and then finally into 3B, 3C and 3D[103][152]. As mentioned before for the P2 region, some of the precursor protein products have suggested functions, such as 3CD, while others are less well characterised. These pathways have been suggested through analysis of the poliovirus genome, and while these products can be seen upon western blot analysis of the P3 region during viral replication, it is unclear if this pathway is followed precisely. It is likely that some level of disparity occurs, especially due to the unique features to FMDV in this region not seen in PV.

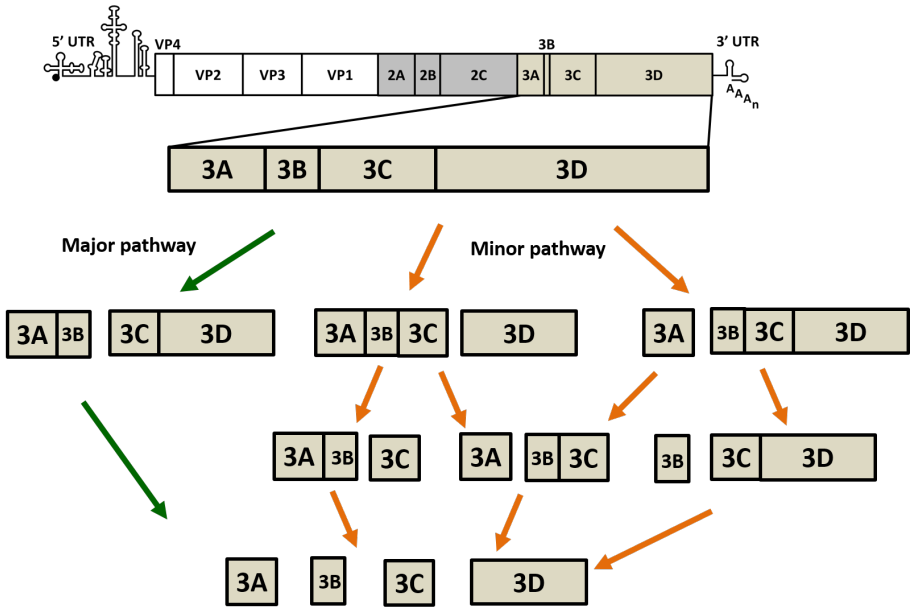

Figure 1.9: Poliovirus P3 processing pathway.
P3 non-structural proteins are processed by two pathways, a major (green arrows) and minor (orange arrows). Processing of proteins is performed by the virally expressed protease 3C.[110].

The 3A protein marks the first unusual aspect of the FMDV P3 region as it is the largest recorded picornavirus 3A protein at 153 amino acids in length, compared to the 87 amino acid PV 3A. This large difference in size is partially due to the inclusion of a large C-terminal flexible domain. Several natural deletions have been reported in this C-terminal region, which resulted in a change in virulence and host range, with deletions associated with increased virulence in swine but reduction in the ability to infect bovine populations[37]. This suggests that the extension of the FMDV 3A may play a key role in different host-pathogen interactions[148,65]. As well as this C-terminal flexible domain is a hydrophobic α-helix domain, thought to be involved in 3A homodimerisation[61]. Interference of this domain results in a reduction of virus viability. At the N-terminus of 3A lies a highly conserved hydrophobic, membrane associated domain. Immunofluorescent microscopy has revealed that this trans-membrane domain causes polarity of the 3A protein towards the ER and Golgi-derived membranes, similarly to the 2B and 2BC proteins mentioned above[146]. The N-terminus of 3A has also been shown to be important in modulating host cell immune responses. Experiments complementing studies investigating the immune modulating effects of $3C^{pro}$ have identified the 3A protein as a negative regulator of IFNβ signalling. Co-immunprecipitation experiments suggest 3A N-terminus dependent interactions between the FMDV 3A protein and host cell immune factors RIG-I, MDA5 and VISA[108].

Following 3A lies the 3B gene. This produces the virus protein genome linked (VPg), also called 3B for ease. FMDV uniquely carries three non-identical, copies of this 3B gene, 3B1, 3B2 and 3B3[60,93]. These small, 23-24 amino acid, proteins are essential in the priming of replication and the initiation of new strand synthesis. All copies of 3B possess an essential conserved tyrosine at the third amino acid position. For 3B function to occur, it must be modified by uridylylation of this tyrosine residue to create the VPg-pUpU[41]. This modification is performed by the viral polymerase, $3D^{pol}$, in coordination with 3CD and the previously described *cre*, using the *cre* AAACA motif as a template[140]. New mechanisms show that the poly-A tail of the

genome can also be used as a template for 3B uridylylation, suggesting a mechanism for differentially made VPg-pUpU, with poly-A-tail templates being used for negative strand synthesis and *cre* templates used for positive strand synthesis, however, the use of a poly-A-tail template has only been shown in *in vitro* uridylylation experiments and is yet to be confirmed[165]. FMDV is unusual in possessing multiple copies of 3B, however, more picornaviruses are being reported carrying multiple copies, although these have only been identified by sequence analysis and have not been tested functionally (personal communication Nick Knowles, the Pirbright Institute). Studies into the multiplicity of the 3B protein have suggested that removal of the first two copies of 3B (3B1 and 3B2) still permits the formation of viable virus, with little impact to replication seen when infected into BHK-21 cells, interestingly this deletion has shown impaired viral replication in porcine cell culture and showed attenuated disease in pigs, suggesting like many other unique aspects of FMDV discussed so far, a role in determining host range and pathogenic potential[148,5]. Further experimentation into this region revealed that upon deletion of 3B3 no infectious virus could be recovered, possibly due to disruption of intergral sequences at the 3B3/3C boundary[54]. These experiments suggest that although viral replication appears to only need one copy of the 3B gene, the multiplicity of this region is clearly important with additional roles in viral replication not yet elucidated.

One of the best described precursor proteins is 3CD, this has been shown to significantly enhance 3B uridylylation in *in vitro* experiments[140]. This complements functional experiments of the 3CD protein in other picornaviruses such as PV. Interestingly, while both PV and FMDV 3CD proteins lack any polymerase activity, the PV 3CD has been shown to maintain the proteolytic capacities of the 3C protein, a function not yet described for the FMDV 3CD protein. Since 3CD lacks any polymerase activity it's thought the function it is providing in uridylylation experiments must be through the RNA binding action of the $3C^{pro}$[75]. Indeed in *in vitro* 3B uridylylation experiments 3CD can be substituted for mature $3C^{pro}$, however this substitution does cause a significant drop in the amount of VPg-pUpU produced[142].

Following 3B in the P3 region is the major picornaviral protease, $3C^{pro}$. This protein is responsible for the majority of the processing of the viral polyprotein with the exception of the autocatalytic cleavage between the leader protein and P1 as well as the 2A induced cleavage between 2A and 2B. The proteolytic activity of $3C^{pro}$ is well defined and is described as a cysteine protease with a Cys-His-Asp catalytic triad within its active site, much like that of other described picornaviral 3C proteins[189] [18]. As well as the roles in processing viral proteins, much like the leader protease, $3C^{pro}$ has shown affinity for cleaving host cell proteins as well. Evidence exists showing the $3C^{pro}$ induced cleavage of translation initiation factor eIF4A, histone protein H3 as well as showing protease dependent fragmentation of Golgi compartments, inhibiting intra-Golgi transport, potentially via degradation of key host proteins[53] [14] [220]. These functions of $3C^{pro}$ are thought to be unique to FMDV to subvert host cell translation and have not been reported in other picornavirus systems. As mentioned previously, $3C^{pro}$ is also implicated in immune evasion by regulating RIG-I mediated signalling through the cleavage of the NF-kB essential modulating protein (NEMO), providing a similar role to the Leader, 2B and 2C proteins[205].

The final protein to be encoded by FMDV is the 3D viral RNA-dependent-RNA-polymerase (RdRp/$3D^{pol}$). Initial observation of the $3D^{pol}$ protein was reported in 1967 where it was originally named the FMDV infection-associated antigen, as antibodies could be easily detected in the sera of infected animals[40] [15]. It was later discovered that this antigen was the viral polymerase. Most picornaviral $3D^{pol}$ proteins share a similar structure and a highly conserved sequence is seen amongst FMDV serotypes, $3D^{pol}$ like many RdRps possesses the almost universally conserved polymerase GDD active site motif[87,203]. The tertiary structure of $3D^{pol}$ resembles a standard right hand, consisting of palm, fingers and thumb domains with the catalytic GDD motif located in the palm[181,57]. The major role of $3D^{pol}$ is in the replication of the viral genome, but early roles also include the uridylylation of the 3B protein as described above. Once uridylylated the 3B forms a primer for genome

replication by interaction of the newly synthesised pUpU with the poly-A-tail, creating a small double stranded RNA region, recognised by the polymerase as a primer for negative strand synthesis. It is for this reason that the first two nucleotides of all picornaviral genomes is a TT, representing the pUpU of the uridylylated 3B, it also serves as the new binding partner for positive strand synthesis, as in the negative strand this TT becomes an AA, providing a binding site for the uridylylated 3B, allowing the formation of the double stranded primer for positive strand synthesis initiation. Picornaviral $3D^{pol}$ proteins are non-proof reading polymerases, allowing for the high level of mutation rates typical of RNA viruses[47]. Specific mutations have been discovered and introduced into the $3D^{pol}$ sequence which can create artificially high-fidelity and low-recombination viral mutants, useful in the study of polymerase kinetics and recombination in the viral lifecycle[125,4,158].

1.2.10 $3'$ UTR and Poly-A-tail

The $3'$ UTR is smaller than the one found at the $5'$ end of the genome, at around 90 nucleotides in length[70]. Within this region are structured cis-acting RNA elements, shown to be required for efficient genome replication[20].

These structures are predicted to form two RNA hairpin loops, named SL1 and SL2. Both S-fragment and IRES elements have been shown to interact specifically with the $3'$ UTR suggesting a role for these stem loops in translation and replication of the viral genome by encouraging circularisation of the genome[179,64]. Deletion of SL2 has shown to inhibit viral replication, whilst deletion of SL1 results in reduced growth kinetics and a reduction in viral genome synthesis[171,177]. A natural deletion of SL2 has been reported in a O type virus, although the mechanisms behind this deletion have not been functionally analysed[20].

Proceeding the $3'$ UTR is the poly-A-tail, as with the poly-C-tract, determining the length of such repetitive sequences is difficult, but lengths of around 40 A residues

are commonly used in recombinant virus systems. The length of the poly-A-tail in PV has downstream effects on infectivity, this could be due to modulation of RNA stability or the maintenance of essential protein-RNA interactions, but no data of investigation in the length of the poly-A-tail on FMDV replication could be found at the time of writing this thesis. Like the structural elements in the $3'$ UTR have been shown to interact with the IRES in the $5'$ UTR via long range RNA-RNA interactions[116]. It is also known that the poly-a-binding protein (PABP) interacts with the poly-A-tail in a specific manner[187]. The interaction of PABP with eIF4G suggests one mechanism for the circularisation of the genome[23].

1.3 Replication phenomenon

1.3.1 Transcomplementation

Translation of viral proteins are known to cause cellular membrane rearrangements as described above with the FMDV 2C and 3A proteins. These rearrangements are thought to provide an environment enriched in the multiple viral and host-cellular factors required for viral replication whilst also protecting viral replication from detection by cellular factors, such as immune modulating proteins. This process of creating viral replication complexes or 'factories' has been well described for other RNA viruses including dengue and hepatitis C virus[45,107,212].

Many of the proteins found within the replication complex possess multiple roles in viral RNA replication. It has been described by both ourselves and previous research that some of these proteins can be provided in trans (i.e, from a separate RNA molecule), while others are required in cis (i.e, expressed from the same RNA molecule)[147,78,82]. We have shown that a single protein might play multiple roles within viral replication, some of which can be provided in trans whilst other are required in cis. For example, we have previously published that the FMDV $3D^{pol}$ can provide its catalytic polymerase function in trans to permit the replication of another viral genome[195]. However, mutations made to non-catalytic residues could not be recovered by a wild-type polymerase provided in trans, suggesting specific cis-acting roles of the viral polymerase (Fig 1.10)[78]. The investigation of protein functions that can be provided in trans will be covered in detail within chapters 3 and 4 of this thesis.

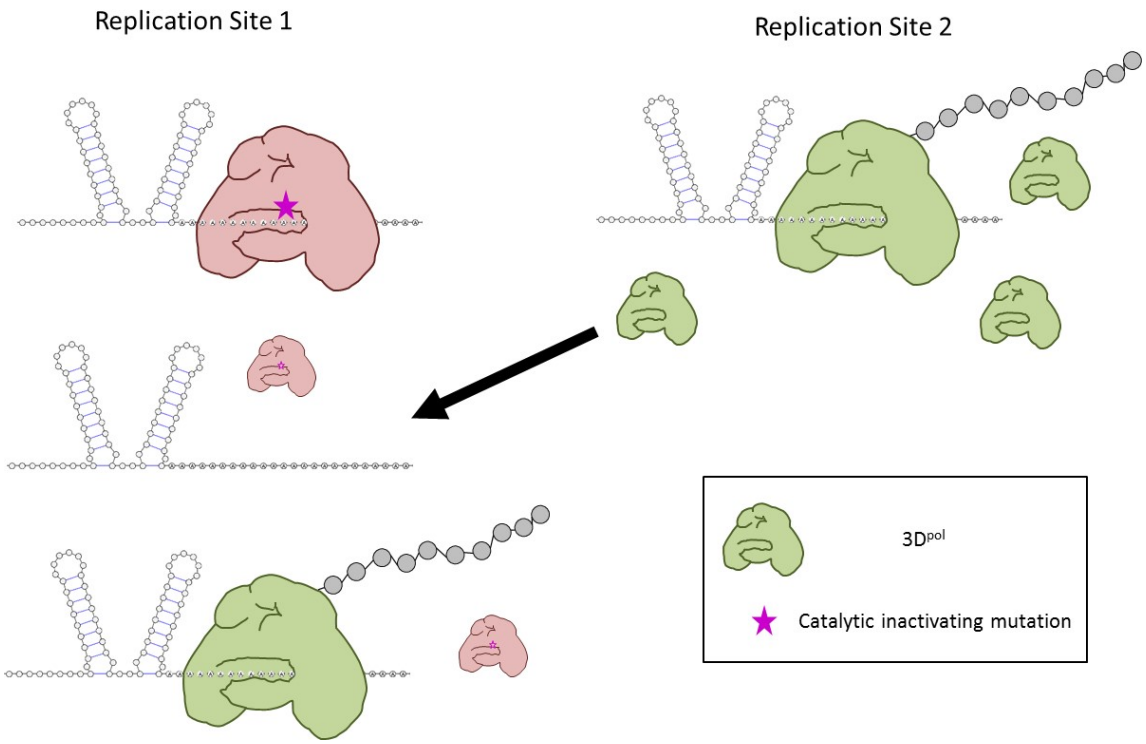

Figure 1.10: Complementation of viral proteins.
Complementation of viral proteins can provide rescue of null-replicative virus genomes. Here a catalytically inactivated 3D polymerase can not replicate the viral genome found in replication site 1. The functional polymerase from replication site 2, can provide its polymerisation function in trans to permit the replication of viral genome in replication site 1.

1.3.2 Recombination

Complementation of viral proteins is one method to recover non-functional mutations to proteins that can stall replication complexes, another way is for the recombination of viral genomes, allowing the replacement of fatal mutations within a sequence. Recombination can occur in both replicative and non-replicative fashions [208]. Replicative recombination is described as traditional strand switching events, where the viral polymerase stalls in new-strand elongation, this can occur by misincorporation of a nucleotide or by elements such as RNA secondary structure, the polymerase and new strand then disassociate from the initial template and restart elongation using

a separate template strand[43]. This form of recombination has been described in numerous virus systems and is thought to account for a majority of recombinatory events within the viral lifecycle[119,182]. The second mechanism of recombination is not reliant on functional replicative elements, but instead the ligation of RNA fragments to create a functional genome. Non-replicative recombination has been demonstrated in poliovirus experiments where transfection of cells with overlapping RNA fragments produced replicative virus. Although all studies of non-replicative recombination have been performed *in vitro*, as such it is hard to ascertain whether this occurs within in vivo infection (Fig 1.11)[67].

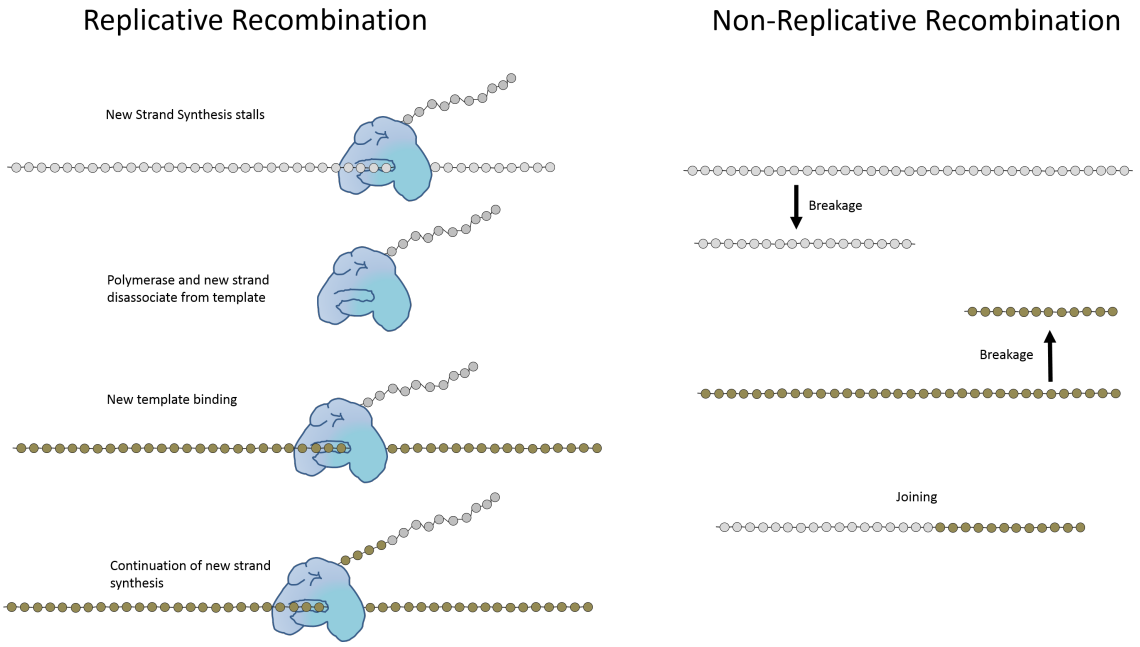

Figure 1.11: Recombination strategies of RNA genomes.
Two forms of recombination are well documented. Firstly replicative recombination, where polymerase new strand elongation is stalled, the polymerase and new strand are disassociated from the template and bind to a second template where elongation can resume, creating a new hybrid genome. Non-replicative recombination requires no replicative elements. Here RNA fragments are simply ligated together to form one complete genome.

1.4 FMDV Growth Systems and the sLoLa consortium

The work described in this thesis forms part of a large effort of a BBSRC sLoLa funded consortium, involving the Universities of St Andrews, Edinburgh, Dundee, Leeds and the Pirbright institute. Due to the highly infectious nature of FMDV, work with full length infectious virus is limited to key, high containment facilities, such as the Pirbright institute. As such investigation into viral replication in Leeds has used a sub-genomic replicon system. Construction of this replicon was first described by our collaborators in St Andrews University and involves the replacement of the FMDV structural proteins with a fluorescent reporter (Fig 1.12) [193,59]. Replication can then be monitored by observation of the expression of this reporter. Part of the successful nature of the FMDV replicon has been the utilisation of an Incucyte Zoom, an automated live-cell fluorescent microscope situated within an incubator, allowing the collection of vast amounts of real-time data regarding the replication of the FMDV replicon.

Due to the limitations of working with infectious virus, initial investigation into FMDV replication is carried out in replicon, before the selection of interesting phenotypes which could then be sent to collaborators in the Pirbright Institute for introduction into infectious virus.

Also described in this thesis is proteomic investigation of RNA binding proteins to the $5'$ UTR of FMDV. Although all selection and experimentation was performed here in Leeds, as part of the consortium the University of Dundee offered their substantial mass spectrometry facility for the analysis of these experiments.

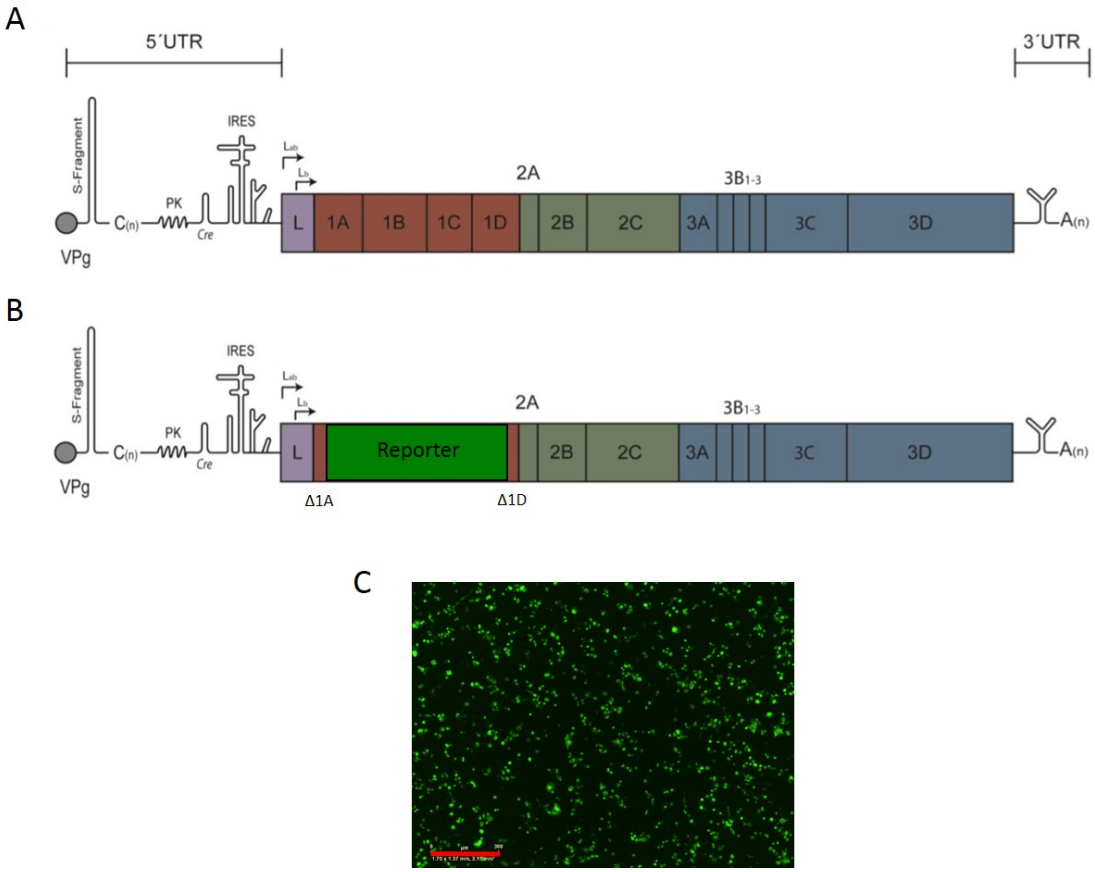

Figure 1.12: FMDV replicon.
The FMDV sub-genomic replicon was produced by removal of the structural proteins from the viral infectious clone (A) and replacement of a fluorescent reporter (B). Remnants of the 1A and 1D genes remain to permit proper poly-protein processing. GFP containing replicons are transfected into tissue culture cell lines and monitored for GFP expression using the IncuCyte Zoom fluorescent microscope (C). Scale bar (red) represents 300 μM.

1.5 Aims of the project

The aim of the sLola consortium was to investigate the replication of FMDV, with the aim of developing novel ways of disease control in the future. This includes the development of helper cell lines, minimising the risks of potential facility outbreaks, as well as developing attenuating strains of FMDV which could be used as new vaccine producer strains, or implemented in research facilities as safer alternatives to fully infectious virus. My contribution to this research topic was initially to investigate the role of the unusual RNA structure found in the $5'$ UTR, the S-fragment. The focus was to generate a better understanding of S-fragment function in genome replication by investigating potential protein interactions and how modification of the structure altered viral replication kinetics. The project continued, encompassing the exploration of other RNA structures such as the pseudoknots, also found in the $5'$ UTR, as well as the unique aspect of possessing multiple copies of the 3B gene.

Chapter 2

Methods

2.1 Mammalian Cell Culture

2.1.1 Cell Lines

Cell line	Derivation	Cell type	Source
BHK-21	Hamster Kidney	Fibroblast	ATCC
MDBK	Bovine Kidney	Epithelial	ATCC
SK-RST	Porcine Kidney	Epithelial	ATCC
HEK 293-T (AUF shRNA)	Human Kidney	Epithelial	Prof. Bert Semler, UCI
HEK 293-T (Scrambled shRNA)	Human Kidney	Epithelial	Prof. Bert Semler, UCI

Table 2.1: Mammalian Cell Types Used in this Thesis
List of mammalian cell types used within this thesis, HEK 293-T cell lines expressing shRNA provided by Professor Bert Semler at the University of California Irvine (UCI).

2.1.2 Cell Maintenance

The cell lines used during the project were maintained in Dulbecco's modified Eagle's medium (DMEM; Sigma), supplemented with 10 % (v/v) heat inactivated foetal bovine serum (FBS; Sigma) and 500 units/mL Penicillin and Streptomycin (Sigma) in 75 cm^2 or 175 cm^2 flasks (Corning) with a Nuncon Delta surface for adherent cell lines. Cells were incubated at 37 ^oC in a humidified incubator at 5 % CO_2 and

routinely passaged using trypsin/EDTA (Sigma).

2.1.3 Treatment of HEK 293-T shRNA cell lines

HEK293-T cells contained doxycycline inducible shRNA expression constructs. 48 hours prior to transfection, cells were treated with 1 μg/mL of doxycycline (Sigma) [198].

2.1.4 Preparation of Cells for Incucyte Analyses

IncuCyte analyses to measure levels of fluorescence in replicon-transfected cells were undertaken in cells seeded out into 24-well plates at a density of $1x10^5$ cells per well and maintained in 500 μl DMEM for 24 hours.

2.2 Nucleic Acid Manipulation

2.2.1 Plasmid Constructs

The pcDNA3 based expression plasmids used for coupled *in vitro* transcription and translation assays expressed FMDV WT and mutant P3 regions were constructed by PCR. Briefly, the appropriate P3 DNA sequence was amplified from the relevant replicon using primers FMDV-3A-ATG-fwd and FMDV-3D-NotI-rvs. Primers included addition of a Kozak modified translational start site and flanking unique NotI restriction sites. The PCR product was digested with NotI and cloned into NotI digest pcDNA3.1(+) (Thermo Fisher Scientific). The sequences of all plasmids were verified by DNA sequencing.

2.2.2 FMDV Replicon Constructs

DNA copies of FMDV replicons (FMDV pRep) were generated by Tulloch et al by replacement of the capsid region of the FMDV infectious clone plasmid with a fluorescent reporter[194]. Mutations made to this construct were performed by PCR mutagenesis or introduction by synthesised gene string (Genewiz) before sequence conformation by DNA sequencing. *In vitro* transcription of the FMDV replicons is powered by an upstream T7 promoter site (Materials and Methods 2.2.8).

2.2.3 PV Replicon Constructs

PV replicons (PV pRep) were made in a similar fashion to the FMDV replicon constructs, with replacement of the PV infectious clone capsid proteins with fluorescent reporters, performed by Dr Lee Sherry (University of Leeds). Mutagenesis of PV replicon construct was performed by PCR mutagenesis and sequences confirmed by DNA sequencing. As with the FMDV replicon, transcription of RNA copies is powered by an upstream T7 promoter.

2.2.4 Transformation of DH5α Bacteria

Plasmids were transformed into DH5α competent E. coli bacterial cells by adding 1 ng DNA to 50 μl thawed DH5α competent cells and incubated on ice for 30 mins. Subsequently, the bacteria and DNA were subjected to a heat shock at 42 ^{o}C for one minute, followed by rapid cooling on ice for two minutes. 500 μl LB media was added and the transformed bacteria were recovered at 37 ^{o}C for one hour prior to spreading on LB agar plates containing 100 μg/ml ampicillin. The plates were incubated for 16 hours at 32 ^{o}C.

2.2.5 Preparation of Plasmid DNA

A single colony from the LB agar plates was picked and grown for 16 hours (overnight) in a 32 ^{o}C shaking incubator in 10 ml LB containing 100 µg/ml ampicillin. Following the overnight culture growth of the colony in the media, 300 µl of the culture was transferred to a microcentrifuge tube and mixed with 600 µl 30 % (v/v) glycerol. Glycerol stocks were stored at -80 ^{o}C for long term storage.

DNA was purified from overnight cultures using a Qiagen Miniprep kit following manufacturer's instructions. Purified DNA was eluted from the column using nuclease free water and the concentration of the eluted plasmid DNA measured using a nanodrop spectrophotometer.

2.2.6 Replicon Linearisation

FMDV pRep plasmid DNA was linearised with restriction enzyme *Asc*I (New England Biolabs – NEB). PV pRep plasmid DNA was linearised with restriction enzyme *EcoR*I (NEB). Linearisation reactions were composed by incubation of 10 units (*Asc*I) or 20 units (*EcoR*I) of enzyme with 5 µg DNA in a 10 µl reaction. Linearisation was confirmed by resolving linearised DNA by 1% agarose (1% (w/v) agarose in Tris-Borate-EDTA (TBE) (89 mM Tris, 89 mM boric acid and 2 mM EDTA)) gel electrophoresis.

2.2.7 DNA Purification

Linearised plasmid DNA (5 µg) was phenol-chloroform extracted and ethanol-precipitated by mixing equal volumes of phenol-chloroform with linearised plasmid DNA and centrifuging the sample at approximately 16,000 x g (or 13,300 RPM) for 10 minutes at 16 ^{o}C. The upper aqueous phase was carefully withdrawn and dispensed into a

fresh nuclease-free microcentrifuge tube.

The aqueous phase was then mixed with an equal volume of chloroform and centrifuged at 16,000 x g (or 13,300 RPM) for a further 10 minutes at 16 ^{o}C. The upper aqueous phase was once again removed and dispensed into a fresh nuclease-free microcentrifuge tube.

The subsequent aqueous phase was subjected to ethanol precipitation by mixing in 1/10 volume 3M sodium acetate and 2.5 x volume 100 % ethanol. The mixture was agitated and placed at -20 ^{o}C for 20 mins. After precipitation, the mixture was centrifuged at 16,000 x g for 30 mins.

The precipitation supernatent was removed and discarded. Pellets were washed with an equal volume of 70 % ethanol in nuclease free water and centrifuged once more at 16,000 x g for 30 minutes at 16 ^{o}C. The supernatant was removed and discarded and the pellet of purified DNA resuspended in 18 µl nuclease-free dH_2O. The concentration of the purified DNA was measured using a nanodrop spectrophotometer.

2.2.8 T7 *in vitro* Transcription

Purified DNA containing FMDV and PV replicons was used as a template for *in vitro* transcription in a T7 transcription reaction. 2.5 µl 10 x T7 transcription buffer (final concentration 40 mM Tris-HCl, 6 mM $MgCl_2$, 1 mM DTT, 2 mM spermidine, pH 7.9), 0.5 µl bovine serum albumin (BSA) (10 mg/ml) and 1.25 µl 20 x RNA Secure (Invitrogen) were added to the DNA and incubated at 60 ^{o}C for 10 minutes, followed by cooling on ice. Subsequently, 30 units RNase OUT inhibitor (Invitrogen), 2 µl ribonucleotide tri-phosphates mix (rNTPs) from 100 mM stock (Roche) and 50 units T7 RNA polymerase (NEB) was added prior to incubation at 32 ^{o}C for four hours.

Reactions were treated with 1.25 µl RQ1 RNase-free DNase (Promega) for 30 minutes at 37 ^{o}C. RNA was recovered using RNA Clean and Concentrator (Zymo Research)

according to manufacturer's instructions and eluted nuclease-free water. The RNA concentration was measured by nanodrop spectrophotometry and stored at -80 ^{o}C.

The integrity and concentration of the RNA was tested by electrophoresis on a 1 % MOPS-formaldehyde gel (1% w/v agarose, 0.5% formaldehyde solution (Sigma), MOPS running buffer (20 mM MOPS pH 7.0, 8 mM Sodium acetate, 1 mM EDTA (VWR)). 500 ng RNA was combined with 2 x RNA loading dye (Fermentas) heated to 85 ^{o}C for 10 minutes and electrophoresed at 80 V for 60 minutes.

2.3 Transfection of RNA

2.3.1 Standard Transfections

Cells prepared in 24-well plates (as described in Materials and Methods 2.1.4) were used for transfection, DMEM was removed and cells washed with PBS before addition of 250 µl of minimal essential media (MEM) supplemented with 10 % (v/v) FBS and 500 units/mL Penicillin and Streptomycin. 1 µg of RNA per well, prepared as described section 2.2.8, was combined with 25 µl MEM and incubated for 20 minutes. 5 µl (per well) of Lipofectin transfection reagent (Life Technologies) was incubated with 20 µl MEM and incubated at room temperature for 20 minutes. After incubation RNA and Lipofectin were combined and incubated for a further 10 minutes. Lipofectin-RNA complexes were then added to cells in a drop-wise fashion as previously described[194,195].

For complementation and competition studies, 0.5 µg of both ptGFP and mCherry replicon RNAs were simultaneously co-transfected per well using Lipofectin (Life Technologies) as previously described. Evidence of replication as a result of transfection was monitored using IncuCyte Dual Colour ZOOM® FLR for 24 hours.

2.3.2 Passaging of FMDV Replicons

BHK-21 cells co-transfected with *in vitro* transcribed replicon transcripts were harvested at 8 hours post transfection using Trizol reagent (Invitrogen). Total RNA was purified using Direct-zol RNA purification kit (Zymogen) according to manufacturer's instructions and eluted in 30 µl Nuclease free water and quantified using a nanodrop.

Naïve BHK-21 cells seeded into 24-well tissue cultures plates were allowed to adhere for 16 hours before transfection with 1 µg of harvested total cell RNA per well

as described above. Fluorescent reporter protein expression was monitored using an IncuCyte Zoom Dual Colour FLR (Essen BioSciences) and analysed using the associated software for fluorescent protein expression as described above.

2.4 Protein Analysis

2.4.1 Immunofluorescence

BHK-21 cells were seeded on to glass coverslips in 24-well plates and transfected with FMDV replicon RNA as described above. Transfected cells were incubated for a range of times (2-8 hours). Following incubation, the media was removed and the cells were gently washed once with phosphate buffered saline (PBS), (137 mM NaCl, 12 mM Na2HPO4, 2.7mM KCl $_{\text{p}}$H 7.4) and fixed using 4 % paraformaldehyde for 10 minutes at room temperature. Once fixed, the cells were stored in PBS at room temperature.

2.4.2 Permeabilisation and Antibody Probing

Cells were permeabilised and blocked in saponin buffer (0.1 % saponin, 10 %FCS, 0.1 % sodium azide) for 2 hours at room temperature on a rocking platform. Permeabilised cells were subsequently probed with the appropriate primary antibody diluted in saponin buffer for 2 hours at room temperature, or for 16 hours at 4 $^{\text{o}}$C. Cells treated with the appropriate primary antibody (as listed in Table 2.2).

Following primary antibody incubation, coverslips were washed 3 times in saponin buffer and probed with the appropriate secondary fluorescent antibodies (as listed in section 2.2) diluted in saponin buffer. Cells were incubated with secondary fluorescent antibodies for 2 hours at room temperature on a rocking platform in the dark.

Coverslips were washed 3 times in saponin buffer and twice in PBS and mounted on microscope slides after blotting dry with 5 µl ProLong Gold antifade mounting agent with DAPI (Life Technologies). Mounted coverslips were sealed before visualisation using a Zeiss LSM 880 confocal microscope.

2.4.3 List of Antibodies

Antibody	Species	Use	Dilution	Source
Anti-3A (2C2) (Primary)	Mouse	WB	1:2000	Gift from Francisco Sobrino (University of Madrid)
Anti-3B (1F8) (Primary)	Mouse	WB	1:2000	Gift from Francisco Sobrino (University of Madrid)
Anti-3D (397) (Primary)	Rabbit	WB	1:5000	Gift from Francisco Sobrino (University of Madrid)
Anti-His (HRP)(Primary)	Mouse	WB	1:5000	Sigma-Aldrich
Anti-HA (Primary)	Mouse	WB	1:5000	Sigma-Aldrich
Anti-V5 (Primary)	Rabbit	IF	1:5000	Gift from Ade Whitehouse (University of Leeds)
Anti-TIA-1 (Primary)	Mouse	IF	1:50	Santa Cruz
Anti-HLA (Primary)	Mouse	IF	1:30	Santa Cruz
Anti-Rabbit Alexa Fluor 568 (Secondary)	Goat	IF	1:2000	Life Technologies
Anti-Mouse Alexa Fluor 647 (Secondary)	Donkey	IF	1:2000	Life Technologies
Anti-Mouse Alexa Fluor 568 (Secondary)	Rabbit	IF	1:2000	Life Technologies
Anti-Mouse (HRP) (Secondary)	Goat	WB	1:2000	Sigma-Aldrich
Anti-Rabbit (HRP) (Secondary)	Goat	WB	1:5000	Life Technologies
Anti-HuR (Primary)	Mouse	WB	1:1000	Thermo-Fisher
Anti-eIF4A (Primary)	Rabbit	WB	1:2000	Gift from Nicolas Locker (University of Surrey)

Table 2.2: List of Antibodies
List of antibodies used throughout this thesis, including species, dilution and source. Application of antibody as a primary or secondary is indicated next to the name. Use indicates western blot (WB) or immunofluorescence (IF).

2.4.4 SDS Polyacrylamide Gel Electrophoresis

Protein samples were prepared for SDS-PAGE analysis by dilution in 2 x Laemmeli sample buffer (Sigma) and heating to 100 ^{o}C for 5 minutes before loading onto SDS-PAGE gels. Proteins were separated on 10, 12 or 15 % polyacrylamide gels (30:1; Severn Biotech) at 180 V using Tris-Glycine running buffer (25 mM Tris, 190 mM glycine, 0.1 % SDS) until necessary separation was achieved. Gels were then either stained with Coomassie stain (0.5 % (w/v) G250 Coomassie Blue (BDH Chemicals), 40 % (v/v) methanol, 10 % acetic acid) or the presence of specific proteins determined through immunoblotting.

2.4.5 Western Blotting

After electrophoresis, proteins were transferred onto Immobilon®-P polyvinylidene difluoride (PVDF) membrane (Merck) (activated in 100 % methanol prior to transfer) on XCell SureLock Mini-Cell wet transfer apparatus (Life Technologies) in Tris-Glycine transfer buffer (25 mM Tris, 192 mM Glycine, 0.1 % SDS, 15 % (v/v) methanol) for one hour 30 minutes at 25 mA. The membrane was subsequently washed in TBS-T (25 mM Tris $_{p}$H 7.5, 137 mM NaCl, 0.1 % tween-20) and blocked in blocking buffer (10 % (w/v) skimmed milk powder in TBS-T) for one hour at room temperature on a rocking platform.

Membranes were incubated for 16 hours at 4 ^{o}C on a rocking platform with the primary antibodies (Table 2.2), diluted to the appropriate concentration in 5 % milk (5 % (w/v) milk powder in TBS-T). Following incubation, the membranes were washed 3 times in TBS-T and incubated for 2 hours at room temperature on a rocking platform with the appropriate secondary antibody (Table 2.2) diluted to the appropriate concentration in 5 % w/v skimmed milk powder in TBS-T.

After secondary antibody incubation, membranes were washed three times in TBS-T

before incubation in combined ECL I and II enhanced chemiluminescence reagents (Thermofisher) and exposed onto CL-Xposure Film (Life technologies). Films were developed using an Xograph Compact X4 Automatic Processor.

2.4.6 Immunoprecipitation

2.4.6.1 Sample Preparation

BHK-21 cells were seeded in 6 well plates at the relevant density and left to adhere for 16 hours at 37 ^{o}C. Cells were transfected with WT or 3B1/2 HA 3B3 Y3F replicons or were untransfected. After 8 hours, whole cell lysates were harvested using IP lysis buffer (Thermo Fisher) following manufacturer's instructions. Whole cell lysates were centrifuged for 10 minutes at 16,000 x g and supernatant moved to a new micro-centrifuge tube. Protein quantification was performed by BCA Assay (Thermo Fisher).

2.4.6.2 Immunoprecipitation

Immunoprecipitations were performed using the Pierce Direct IP Kit (Thermo Fisher), coupling anti-HA antibodies to the resin provided. 5 µg of HA anti-body was used per column and incubated for 2 hours rotating at room-temperature. 400 µg of prepared MDBK lysate was added to the antibody-coupled beads and incubated rotating overnight at 4 ^{o}C. Flow through samples were harvested and kept for analysis by western blot alongside three washes and the eluate.

2.5 Expression and Purification of His-Tagged $3D^{pol}$

WT his-tagged $3D^{pol}$ (a kind gift from Esteban Domingo) and I189L mutant (constructed by PCR mutagenesis) were expressed in Codon plus (+) RiPL competent E. coli. For protein production, a 12 ml LB medium containing kanamycin (100 μg/ml) and chloramphenicol (34 μg/ml in 100 % ethanol) grown for 16 hours at 37 ^{o}C was used as a starter culture. 10 ml of starter culture was used to inoculate 1 L LB broth with 100 μg/ml kanamycin.

The cultures were grown at 37 ^{o}C in a shaking incubator until optical density at 600 nm (OD600) was in the range of 0.6-0.8. Cultures were then induced with 1 mM IPTG and grown for a further 3 hours at 37 ^{o}C. The cultures were centrifuged at approximately 6,000 x g (or 5,000 RPM) for 10 minutes and the supernatant was discarded. The bacterial pellet was resuspended in 1 ml storage buffer (20 mM Tris-HCl, $_{p}$H 8) and stored at -80 ^{o}C.

Pellets were lysed with the addition of 30 ml lysis buffer (20 mM Tris-HCl, 500 mM NaCl, $_{p}$H 8), 200 μl 100 mg/ml lysozyme (Sigma), 15 μl 100 mg/ml DNaseI (Life Technologies), and 1 cOmplete EDTA-free protease inhibitor cocktail tablet (Roche) and placed on a roller at 4 ^{o}C for 30 minutes. Lysed pellets were subsequently sonicated alternating 10 seconds on and 40 seconds off for 6 minutes. The sonicated lysate was centrifuged at approximately 27,000 x g (or 15,000 RPM) for 20 minutes and the supernatant filtered through 0.22 μm filter.

The filtered supernatant (load) was subjected to nickel affinity chromatography using 1 ml HisTrap HP columns (GE Healthcare) attached to a peristaltic pump. HisTrap columns were prepared by washing with five column volumes (CV) of dH_20, and equilibrating with five CV buffer A (wash)(50 mM Tris HCl $_{p}$H 8, 500 mM NaCl, 25 mM imidazole). The load was added and collected as flow through. The column was washed three times in 3 CV buffer A (wash). Each wash was collected separately. The column was then washed with 5 CV buffer B (elution) (50 mM Tris $_{p}$H8, 500

mM NaCl, 500 mM imidazole). The flow through was collected in eight 250 µl fractions and the final 3 ml was collected in fraction 9.

The flow through, washes and eluted fractions were analysed by SDS-PAGE and Coomassie (brilliant blue) stain.

2.6 Sym/Sub $3D^{pol}$ polymerisation Assays

2.6.1 ^{32}P UTP $5'$ End Labelling of Sym/Sub RNA

Sym/Sub RNA (Sigma) was $5'$ end labelled using γ^{32}P UTP (Perkin Elmer) and T4 polynucleotide kinase (NEB) as specified by the manufacturer. Reactions, typically 20 µl, contained 25 µCi of ^{32}P UTP, 60 µM sym/sub RNA, 2 µl 10x kinase buffer (final concentration 70 mM Tris-HCl, 10 mM MgCl2, 5mM DTT $_{p}$H 7.6), and 20 units of T4 PNK. Reactions were incubated at 37 ^{o}C for 60 minutes. Labelled RNA was purified by ethanol precipitation and resuspended in 50 µl of Nuclease free water.

2.6.2 Polymerisation

0.5 µM of ^{32}P labelled sym/sub RNA (Sigma) in polymerisation buffer (30 mM MOPs $_{p}$H 7, 33 mM NaCl, 5 mM MgAc) was heated to 95 ^{o}C for 2 minutes before incubation on ice for a further 2 minutes. 2 µM of recombinant 3D polymerase was added to the RNA and incubated at 37 ^{o}C for 10 minutes to promote annealing. After incubation 50 µM of rNTPs were added to the reaction, addition of nucleotide varied depending on experiment and is outlined in the text. Aliquots were taken periodically and reactions stopped by addition of 2 x TBE-UREA RNA loading dye (Thermo Fisher).

2.6.3 Visualisation of Extension Products

Extension products in loading dye were heated to 70 ^{o}C for 5 minutes prior to loading on a 23 % denaturing polyacrylamide gel containing 7M urea in TBE buffer. Samples were electrophoresed at a constant 30 mA until sufficiently separated. After electrophoresis, gels were fixed for 30 minutes in fixative solution and exposed onto CL-Xposure Film or a phosphoscreen.

2.7 Selective $2'$ Hydroxyl Acetylation Analysed by Primer Extension (SHAPE)

2.7.1 NMIA Treatment

12 pmol of RNA, transcribed using the protocol above (2.2.8), were folded by heating to 95 ^{o}C for 2 minutes followed by incubation on ice for 2 minutes, followed by incubation with 3x folding buffer (333 mM HEPES $_{p}$H8, 20 mM MgCl2, 333 mM NaCl), RNA was incubated at 37 ^{o}C for 30 minutes. Folded RNA was subsequently incubated with 5 mM NMIA or DMSO and incubated at 37 ^{o}C for 50 minutes.

Treated RNA was ethanol precipitated by addition of 5 % v/v 5 M NaCl, 2.5 % v/v 100 mM EDTA (making a final concentration of 25 mM NaCl, 250 μM EDTA), 20 μg glycogen, and made up to a total concentration of 70 % ethanol in Nuclease free water. Precipitations were incubated at -80 ^{o}C for 30 minutes before pelleting by centrifugation for 30 minutes at 16,000 x g for 30 minutes. Pelleted RNA was resuspended in 0.5x TE.

2.7.2 Primer Extensions

NMIA treated RNA was reverse transcribed using superscript III using manufacturer's protocol, NMIA and DMSO reactions were extended using 0.84 μM 6-FAM fluorescein labelled primers. Upon reverse transcription, cDNA was treated with 4M NaOH and heated at 95 ^{o}C for 3 minutes, followed by addition of 2 M HCl and incubation on ice. Newly made cDNA was ethanol precipitated as described before with the replacement of 5 M NaCl with 3 M NaAc. Pelleted cDNA was resuspended in deionized formamide (Applied Biosystems) by heating to 65 ^{o}C for 10 minutes and vortexing.

Sequencing reactions were performed as above, with the following changes; 0.84 μM of Hexachloro-fluorescein labelled primer was used, as well as the addition of 106 nM ddCTP.

2.7.3 Capillary Sequencing of SHAPE Reactions

Sequencing reactions were mixed with NMIA and DMSO prepared samples and submitted to Dundee DNA sequencing for analysis by capillary electrophoresis.

2.7.4 Analysis of SHAPE Reactions

In vitro SHAPE fragment analysis was conducted by capillary electrophoresis by DNA sequencing and Services, Dundee University. SHAPE data was analysed in the program QuSHAPE and RNA structure prediction was carried out using RNAstructure software using the generated SHAPE reactivity profile as a pseudo free energy constraint[90,162]. All figures were created using the RNA structure applet VARNA.

2.8 The Pirbright Institute Recovery of Infectious Virus

2.8.1 Virus Recovery

The coding region for the reporter and flanking FMDV sequence were excised from DNA copies of mutant replicons by DNA digestion with PsiI and XmaI restriction sites and replaced with the corresponding fragment from pT7S3 FMDV infectious clone encoding the capsid region[94]. The resulting infectious clone plasmids were

verified by restriction digest analysis.

Infectious clone plasmids were linearized by digestion with *Asc*I and full-length viral RNA transcribed using a T7 MEGAscript kit (Thermo Fisher Scientific) and subjected to DNase digestion using TurboDNase (Thermo Fisher Scientific) followed by purification using a MEGAclear Transcription Clean-Up kit (Thermo Fisher Scientific).

RNA quality and concentration were determined by RNA denaturing agarose gel electrophoresis and nanodrop spectrophotometry.

BHK-21 cells were transfected in a 25 cm^2 flask with 8 μg of infectious clone RNA using TransIT transfection reagent (Mirus) according to the manufactures protocol. At 24 hours post-transfection or at appearance of cytopathic effect (CPE), cell lysates were prepared by freeze/thawing overnight and the lysate clarified by centrifugation.

The clarified lysate (1 ml) was passaged blind onto subsequent BHK-21 cell monolayers (70 % confluent) in 25 cm^2 flasks in 5 ml of virus growth medium (reduced serum medium with 1 % FCS) for a maximum of 48 hours.

Flasks were scored for the appearance of CPE indicating recovery of infectious virus. The cells were freeze/thawed and the resulting lysate clarified by centrifugation and stored at -80 ^{o}C.

2.8.2 Sequencing of Recovered Virus

Recovered viruses (at BHK-21 p4), were sequenced on the Illumina Miseq (Illumina) platform using a modified version of a previously described PCR-free protocol[113]. Total RNA was extracted from clarified infected cell lysates using TRIzol Reagent (Thermo Fisher Scientific) as per manufacturer's instructions. Residual genomic DNA was removed using DNA-free DNA Removal Kit (Thermo Fisher Scientific) following manufacturer's protocol.

After precipitation with 3 M sodium acetate and ethanol, 10 µl (containing from 1 pg to 5 µg) of RNA was used in a reverse transcription (RT) reaction as previously described with the exception that, in addition to Random Hexamers (Bioline Reagents Ltd), two primers (Rev6 and NK72, previously described[113]) were included in the first incubation step, and the final incubation step at 42 ^{o}C was carried out for 40 minutes[1].

Second strand synthesis was carried out using the NEBNext mRNA Second Strand Synthesis Module (NEB) following the manufacturer's protocol and subsequent cDNA extracted by the addition of an equal volume of phenol:chloroform:isoamyl alcohol (Thermo Fisher Scientific) followed by 3 M sodium acetate/ethanol precipitation as described[1].

cDNA was quantified using the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific) as per manufacturer's instructions and a cDNA library was prepared using the Nextera XT DNA Sample Preparation Kit (Illumina) following manufacturer's recommendations. Sequencing was carried out on the MiSeq platform using Miseq Reagent Kit v2 (300 cycles) chemistry (Illumina).

The quality of subsequent FastQ files was checked using FastQC and poor quality reads were filtered out using the Sickle algorithm.

Host cell reads were removed using the FastQ Screen algorithm and FMDV reads assembled de novo into contigs with IDBA-UDz[156].

Only contigs which matched the FMDV library after running a Basic Local AlignmentSearch Tool (BLAST) algorithm were assembled into consensus sequences using SeqMan Pro software implemented in the DNA STAR Lasergene 13 package (DNASTAR)[2]

To confirm the consensus, reads were mapped using Burrows-Wheeler Aligner (BWA)[109]. Mapped reads were visualised using Integrative Genomics Viewer (IGV)[169].

2.8.3 Plaque Assays

Confluent BHK-21 cell monolayers were infected with 10-fold serial dilutions of virus stock, overlaid with Eagle overlay media (from the Central Services Unit at the Pirbright Institute) supplemented by 5 % tryptose phosphate broth solution (Sigma), penicillin (100 units/ml) and streptomycin (100 µg/ml) (Sigma) and 0.6 % Indubiose (MP Biomedicals), and incubated for 48 h at 37 ^{o}C. Cells were fixed and stained with 1 % (w/v) methylene blue in 10 % (v/v) ethanol and 4 % formaldehyde in PBS.

2.8.4 Cell Killing Assays

Virus titres were determined by plaque assays. BHK-21 cells were seeded with initial seeding density of 3 x 10^4 cells/well in 96 well plate and allowed to settle overnight. Cell monolayers were inoculated with each rescued virus (WT, Aerie, B-48, B-97 A, B-97 B, B-97 C) at MOI of 0.01 PFU for 1 h, inoculum was removed and 150 µl of fresh media (supplemented with 1 % FCS) was added to each well. Appearance of CPE was monitored every 30 minutes using IncuCyte S3 for a total time of 50 hrs.

2.9 ^{3}H Labelling of Newly Generate RNA Strands

2.9.1 RNA Transfection of BHK-21 Cells

5 x 10^5 BHK-21 cells were seeded per well of a 6 well plate and left 16 hours to adhere. Two hours pre-transfection, cells were treated with 10 µg/ml per well of Actinomycin D (Sigma). RNA transfection as described above except RNA and Lipofectin reagent were appropriately scaled up in proportion with the higher number of cells seeded.

2.9.2 ^{3}H Uridine Addition

One hour post transfection 18 µCi of tritiated uridine (^{3}H) (Perkin Elmer) was added to each well.

2.9.3 RNA Harvest

Total RNA was harvested at 4 hours post-transfection in 1ml of Trizol reagent (Invitrogen) per well. Harvested RNA was purified using a purelink RNA mini kit (Invitrogen), following manufacturer's instructions, and eluted in 50 µl of Nuclease free water and quantified using a nanodrop.

2.9.4 Gradient Separation

14 mL, 5-25 % sucrose RNA gradients were prepared in 100 mM NaAc and 0.1 % SDS, 30 µg of total RNA was added to the gradient and centrifuged at 150,000 RCF in a SW40 rotor (Beckman) for 5 hours at room temperature.

Gradients were hand fractionated into 350 µl fractions directly into 3ml of scintillation fluid (Perkin Elmer). ^{3}H incorporation was monitored by scintillation counting, counting for on minute per tube[173]. Purified transcribed RNA was used as a control to determine where the replicon RNA would lie on the gradient. Here, 50 µg of purified transcribed replicon RNA was separated and fractionated as above, however, fractions were instead analysed by absorbance at 260 nm reading using a nanodrop.

2.10 Coupled Transcription/Translation Assays

Coupled *in vitro* transcription and translation assays were performed using the TNT Quick Coupled Transcription/Translation System (Promega) following manufacturer's protocols. Reactions were assembled on ice containing 500 ng of pcDNA T7 expression plasmid and 0.4 mCi / ml of ^{35} S methionine (PerkinElmer). Reactions were incubated at 30 ^{o}C for 40 minutes before addition of 1 µl of 50 mg / ml unlabelled methionine and cysteine. At 20 minute intervals samples of each reaction were stopped by the addition of 2 x Laemmelli buffer. The radiolabelled products were separated by SDS-PAGE before incubation in fixing solution (12 % methanol, 10 % acetic acid) for 30 minutes. Fixative was then replaced with amplify fluorographic reagent (GE Healthcare) and incubated for another 30 minutes. Fixed and enhanced gels were dried using a vacuum gel dryer for two hours before exposure to a phosphoscreen.

2.11 Electron Microscopy of BHK-21 Cells

2.11.1 Preparation of BHK-21 cells

BHK-21 cells were seeded into 6 well plates as described above. Duplicate wells were transfected with FMDV replicon and fixed using 2.5 % glutaraldehyde for 2.5 hours at 0, 4, 6 and 8 hours post-transfection. Fixed cells were collected into a 15 ml falcon tube using a cell scraper and washed twice with phosphate buffer (PB) (0.1M sodium phosphate), pelting the fixed cells by centrifugation (5 minutes at 4000 x g) between each wash.

2.11.2 Resin Embedding

Fixed cells were incubated in 1 % osmium tetroxide for 60 minutes before four washes in PB. Cells were dehydrated by incubation for 20 minutes in an ascending alcohol series (40 %, 60 %, 80 % 2 x 100 %) with cells pelleted by centrifugation between each step.

Following dehydration, cells were incubated in propylene oxide for 20 minutes twice with pelleting by centrifugation between each step.

Propylene oxide – Araldite mixture was made at 50 % of each reagent and added to the cells. Samples were then incubated overnight on a rotator at room temperature. Cells were pelleted by centrifugation and propylene oxide-araldite mixture replaced with neat araldite mixture (53 % araldite Cy212, 45 % DDSA, 2 % DMP30) for 8 hours rotating before pelleting of cells and replacement with fresh araldite mixture for overnight incubation on a rotator.

Cells were pelleted by centrifugation before removal of araldite and replacement with fresh araldite mixture and centrifuged to remove bubbles and cell pellet disrupted using a pipette. Araldite mixture was polymerised by incubation in an oven at 60 ^{o}C for 24 to 36 hours.

2.11.3 Sectioning of Resin Embedded Cells

Resin embedded cells were sectioned using a glass knife and ultra-microtome. Gold and silver sections were isolated and loaded onto EM grids.

2.11.4 Uranyl Acetate Staining of Sections

Grids were incubated in 8 % uranyl acetate for 1 hour before washing five times for 5 minutes each in water. Grid were then stained with lead citrate for 10 minutes before washing with 0.02 M NaOH three times for 1 minute and water five times.

2.12 RNA-Protein Pull Down

2.12.1 Preparation of MDBK Cell Lysate

A T75 flask of MDBK cells (approximatley 8 x 10^6 cells) was harvested by lysis with RIPA buffer (25 mM Tris $_p$H 7, 150 mM NaCl, 0.1 % SDS 0.5 % sodium deoxycholate, 1 % triton x-100). After cell lysis, debris was pelleted by centrifugation at 16,000 x g for 10 minutes. Supernatant was removed and protein concentration calculated using a BCA assay according to manufacturer's instructions (Thermo Fisher).

2.12.2 RNA $3'$ Biotinylation

RNA fragments representing the $5'$ UTR were produced by transcription as described above (2.2.8). Once purified, 50 pmol of RNA was used in biotinylation reactions according to the manufacturer's instructions (Pierce), with the expansion of incubation to overnight at 16 ^{o}C. Once biotinylated, RNA was purified by ethanol precipitation and resuspended in nuclease free water.

2.12.3 RNA-Protein Pull Down

RNA – protein pull downs were performed using the Pierce magnetic RNA- protein pull down kit (Thermo Fisher) using 200 µg of MDBK lysate. Successful pull-down

was checked by western blot of control AR RNA interaction with HuR protein.

2.12.4 Mass Spectrometry

Samples were sent to the University of Dundee for mass spectrometry and protein identification. Identified proteins were annotated using the Bos Taurus protein library. Raw data containing annotated proteins was given to myself for data analysis. Data was analysed using the Perseus software.

Chapter 3

3B or not 3B: Studying the multiplicity of the FMDV 3B gene

3.1 Introduction

FMDV, as described previously, is a single stranded, positive sense RNA virus. The genome contains one open reading frame encoding both viral structural and non-structural proteins. This chapter is concerned with the P3 non-structural region, encoding the 3A protein, three non-identical copies of 3B, the $3C^{pro}$ and RNA-dependent RNA polymerase (RdRp/$3D^{pol}$).

Translation of the FMDV genome produces a single large polyprotein, this is efficiently cleaved by viral proteases to generate the mature viral proteins. As the genome is translated the leader protein self-cleaves from the rest of the viral polyprotein, the structural capsid proteins are then removed by the action of the small 2A protein. The discovery of the 2A mechanism (described in Section 1.2.9.3), first proposed by Professor Martin Ryan at the University of St Andrews has proven to be useful for many areas of molecular biology and is now widely used as a tool in many different formats[131]. The 2A protein is commonly used as a linker region inserted between two or more proteins during co-expression studies, this allows the production of distinct mature proteins whilst only having to use one expression construct.

The FMDV P2 and P3 regions are cleaved by the P3 encoded viral protease, $3C^{pro}$. Processing of these regions produces not only mature viral proteins, but also immature precursor products (Fig 3.1)[63]. Some of these have been shown to provide essential functions, such as that of the 3CD protein, involved in the uridylylation of the FMDV 3B protein as described (Introduction 1.2.9.4), whilst others have no prescribed function[140].

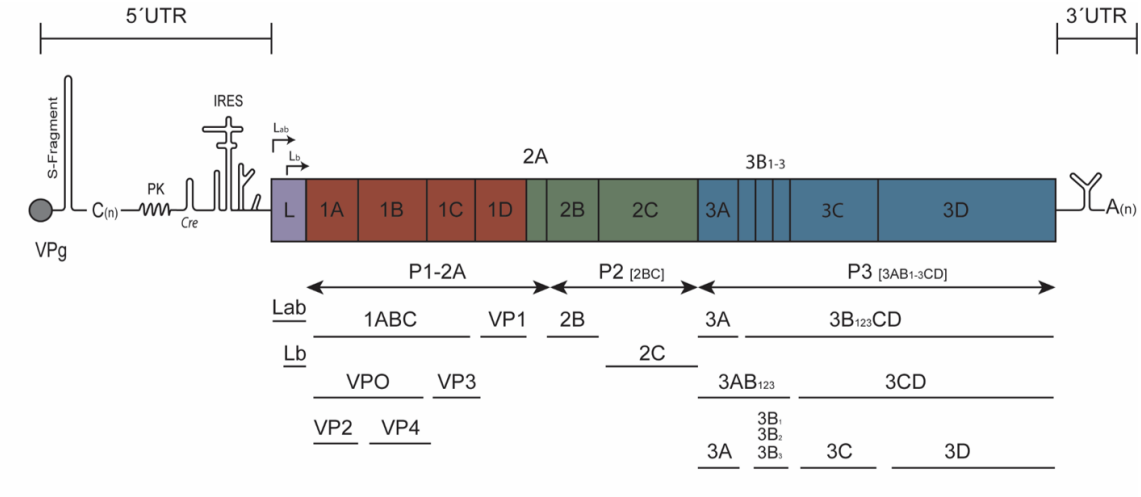

Figure 3.1: Schematic of the FMDV genome showing the proposed processing pathways representation of the FMDV Genome.
The FMDV genome contains a single open reading frame flanked by 5′ and 3′ UTRs. Two start sites are shown by arrows, labelled Lab and Lb. Also shown is the cleavage of the FMDV polyprotein creating a series of precursor proteins as well as mature products. Figured adapted from Forrest et al[59].

One of the many unique features about the FMDV genome (compared to other picornaviruses) is the 3B gene. Most picornaviruses possess one copy of this gene, FMDV, however, possesses three, non-identical copies (Fig 3.2). The presence of three copies of 3B is reported to be conserved in all but one natural isolate to date, suggesting a high selective pressure for maintaining all copies of 3B, however, the driving force behind this selection is so far unknown. The three 3B genes encode three non-identical copies of the virus-protein genome-linked (VPg/3B). The 3B protein is a small (23 to 24 amino acids) protein involved in the initiation of new strand RNA synthesis where it acts as a primer for the viral encoded polymerase $3D^{pol}$[60,153]. Critical to

this function is the conserved tyrosine found in the third residue of all picornaviral isolates reported so far. Sequence analysis of new picornaviruses has unearthed viral isolates possessing more than one copy of 3B, however, these have yet to be functionally characterised leaving FMDV as the only virus functionally characterised to possess three active copies (personal communication, Nick Knowles, the Pirbright Institute).

```
3B1  GPYAGPLERQKPLKVRAKLPQQE- 23
3B2  GPYAGPMERQKPLKVKAKAPVVKE 24
3B3  GPYEGPVKKPVALKVKAKNLIVTE 24
     *** **       *** **
```

Figure 3.2: Sequence alignment of O1K 3B proteins
Amino acid sequence alignment of the 3 copies of 3B from the FMDV O1K genome. Stars indicate conserved residues, the essential conserved tyrosine is highlighted in hot pink.

The essential tyrosine in 3B is modified by the addition of two uridines in an uridylylation reaction performed by a complex of $3D^{pol}$, the 3CD precursor and a viral RNA structure termed the cis-active replicative element, or cre. This is somewhat ironically named as it has been shown that the cre can act in both cis and trans, so the term 3B uridylylation site or 'bus' is also used by some researchers. The cre forms a stem loop structure found in the $5'$ untranslated region immediately prior to the IRES[192]. The loop of this hairpin includes a conserved AAACA motif. Where with the help of $3D^{pol}$ and 3CD, this motif is used to uridylylate the tyrosine of the 3B proteins to create the product VPg-pUpU. This uridylylated product can then be transferred to the $3'$ UTR where the newly added uridine residues can interact with the adenosines of the poly-A-tail, forming a double stranded primer. The viral polymerase, $3D^{pol}$, can then use this primer as an initiation site for new strand synthesis. As discussed in the introduction to this thesis (Introduction 1.2.9.4), investigation into new strand synthesis with other picornaviruses has revealed the possibility of

alternative routes of 3B uridylylation, using the poly-A tail as a template instead of the *cre*[187,139,140]. This is hypothesised to be the case for negative strand synthesis in the replication of PV, but so far has not be shown to be the case for other picornaviruses[141].

Due to the small size of 3B (approximatley 2.6 kDa), working with this protein can be challenging. Therefore, to date, efforts to investigate the uridylyation of 3B have been limited to *in vitro* techniques using purified recombinant proteins and ^{32}P labelled UTP. These methods have been useful for investigating the context in which 3B is uridylylated, for example, to identify whether uridylylation occurs using a mature 3B protein or a larger 3B-containing precursor protein such as 3AB or 3BC, (which is later cleaved to produce the mature VPg-pUpU). To answer this question, recombinant precursor proteins were created where 3B1, 3B2 and 3B3 and 3C^{pro} proteins were expressed together to create 3B123C. Processing of this precursor was inhibited by catalytic inactivation of the 3C^{pro} active site. The 3B123C protein was then used in *in vitro* uridylylation experiments, where this protein, RNA fragments representing the cre, recombinant 3CD protein and ^{32}P labelled UTP were co-incubated. The addition of ^{32}P UTP residues to the 3B of 3B123C was then analysed by SDS-PAGE electrophoresis and autoradiography. The results showed that both FMDV 3B123C and poliovirus (PV) 3BC precursor proteins could be uridylylated, whereas further PV experiments showed that the 3AB precursor could not[153,141,152]. Difficulties in producing recombinant FMDV 3A containing proteins means that this experiment has not been performed for FMDV. These experiments suggest that 3B can be uridylylated in the format of a precursor. However, difficulties in investigating 3B uridylylation in a cellular environment means that these observations are yet to be confirmed during natural virus infection.

Interestingly, although only 3BC precursors have been shown to be capable of uridylylation in *in vitro* uridylylation assays, investigation into the requirements of providing 3B to a 3B inactivated PV in cis (i.e expressed from a different IRES on the same RNA), revealed that only when linked to 3A (3AB) could 3B be provided[110].

This means that there are currently two compelling bodies of evidence, one suggesting 3BC is a precursor for uridylylation, the other saying 3AB is essential to be able to provide 3B, making the exact pathway hard to elucidate.

It is thought that the mechanism of 3B uridylylation is conserved throughout all picornaviruses, with the cre element being reported in all genera with the exception of the Kobu, Erbo and Teschoviruses. The mechanism of 3B uridylylation can be disrupted in two different ways to prevent its function and ultimately disrupt replication. One way is to mutate the cre AAACA motif, for example to GAACA, (commonly described as an A1G mutation), this prevents its ability to act as a template for uridylylation[185]. Alternatively, the 3B protein can be mutated to replace the essential tyrosine, with, for example a phenylalanine, 'Y3F'. Both mutations act to inhibit viral replication and will be referred to in several chapters throughout this work[110].

Why FMDV has maintained three copies of 3B remains unclear. Creation of recombinant viruses containing only one copy of 3B has shown attenuation in porcine cells, suggesting a potential role in host-dependent interactions[148]. Similar experiments showed that removing 3B1 and 3B2 allowed production of viable virus in BHK-21 cells, with no reduction in RNA replication[148,6]. However, it was also reported that viruses created containing only one copy of 3B (3B3) showed a compromised ability in the release of virus particles[6]. It was further reported that removal of 3B3 produced non-viable viruses, with suggestion that alongside its role as a primer for replication, 3B3 was important in some other aspect of viral replication, although what that role is has not yet been elucidated[54].

This chapter aims to investigate the unusual aspect of FMDV to possess multiple copies of the 3B gene, alongside the reports of some aspect of 3B3 being essential for replication. This is explored by deletion and inactivation of 3B proteins as well as investigation into the requirements of P3 cleavage by mutation of cleavage boundaries between the proteins of 3B.

3.2 Results

3.2.1 Inactivation and deletion of individual 3B proteins

As described above, previous experimentation has shown no specific copy of 3B is essential for viability. This was verified using FMDV replicons with the introduction of inactivating tyrosine (Y3F) mutations and deletions to individual copies of 3B. The data comprising this part of this chapter formed a large body of work published by Herod et al[79]. The results presented here were obtained by a post-doctoral scientist in the Stonehouse/Rowlands group, Dr Morgan Herod, and myself, with credit for individual experiments described in the text.

The replicon used throughout this thesis was originally created by Tulloch et al, and has been modified, including the introduction of a new, GFP reporter (ptGFP), which is not cleaved by the FMDV leader protein (an issue which caused reduced sensitivity with the original GFP used), or a red fluorescent, mCherry reporter[193]. Ribozymes were also introduced, with hammerhead and hepatitis D ribozymes inserted at the $5'$ and $3'$ ends of the FMDV sequence respectively. Introduction of these ribozymes allows the production of precise ends upon the transcription of FMDV RNA (Fig 3.3A). Maintenance of these precise ends is known to increase both viral titre and genome replication, as such the introduction of the ribozymes ensure the optimal genome replication from the replicon possible.

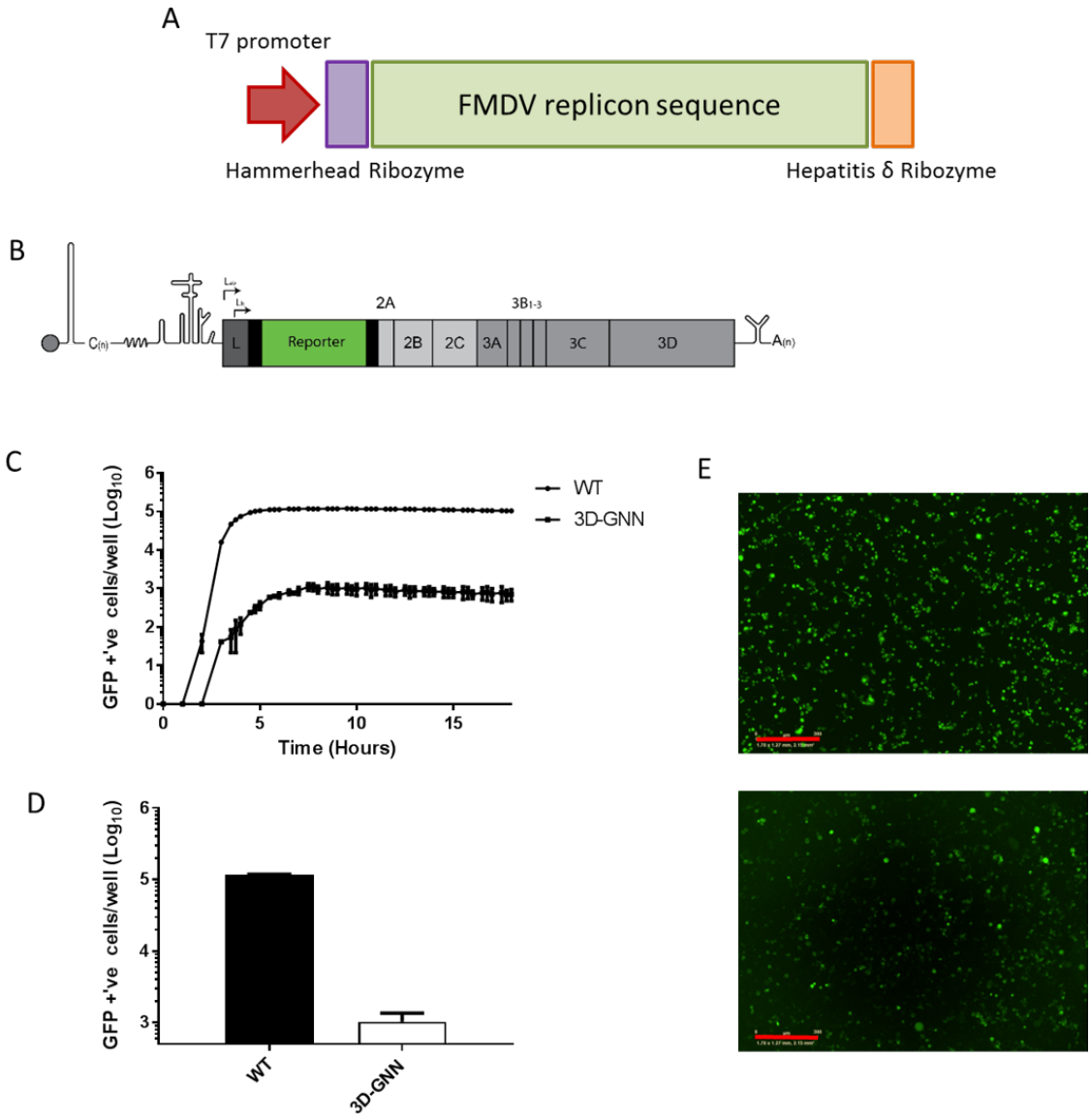

Figure 3.3: Schematics and replication of FMDV O1K replicon
A. Representation of the O1K FMDV DNA plasmid, transcription is performed using a T7 polymerase with a T7 promoter inserted present before the replicon sequence. Replicon sequence is flanked at each end by hammerhead and hepatitis D ribozymes. B. Schematic of the FMDV replicon, with structural proteins replaced by a fluorescent reporter. Replication of WT and null-replicative control (3D-GNN) replicons in BHK-21 cells, GFP expression shown over 18 hours (C) or at 8 hours post transfection (D), as measured by the Incucyte Zoom. E. GFP images taken at 8 hours post-transfection by the Incucyte, showing WT replicon (upper) and 3D-GNN (lower). Scale bar (Red) represents 300 μM

Genome replication of the replicon is monitored by observing expression of the fluorescent reporter, with fluorescing cells counted hourly by an Incucyte Zoom, an automated fluorescent microscope situated inside an incubator. A $3D^{pol}$ mutation, inactivating the polymerase active site (GDD¿GNN), is used as a negative control for genome replication, helping the distinction between input IRES-mediated translation and genome replication (Fig 3.3). Baby hamster kidney cells (BHK-21) are used as the cell line for monitoring replication for most of these studies. This is both due to the permissive nature of this cell line for FMDV and to allow comparison with other research in the field as most research performed on FMDV is undertaken in BHK-21 cells. Other cells are used throughout this thesis include Madin-Darby Bovine kidney cells (MDBKs) and a porcine cell line SK-RST. The rationale behind the use of these cell lines will be explained in the relevant chapters.

Initially, six replicons were generated, with each of the individual copies of 3B either inactivated by introduction of tyrosine mutations (3B1 Y3F, 3B2 Y3F and 3B3 Y3F), or deleted (Δ3B1, Δ3B2, Δ3B3). Mutations were introduced into the DNA Replicon clone containing a mCherry reporter (pRep mCherry). These DNA constructs were then transcribed to produce replicon RNA and transfected into BHK-21 cells alongside WT and 3D-GNN controls.

Inactivating tyrosine mutations showed, as expected, that there was no preference to which copy of 3B was active. However, deleting whole copies of 3B showed a change in phenotype. While no change in replication was observed when deleting 3B1 (Δ3B1) or 3B2 (Δ3B2), upon deletion of 3B3 (Δ3B3), replication was reduced to the level of input translation (GFP produced by translation of the transfected replicon RNA), as shown by the 3D-GNN (Fig 3.4). Therefore showing a clear need for presence of 3B3, even if it was not active, for replicon replication to take place.

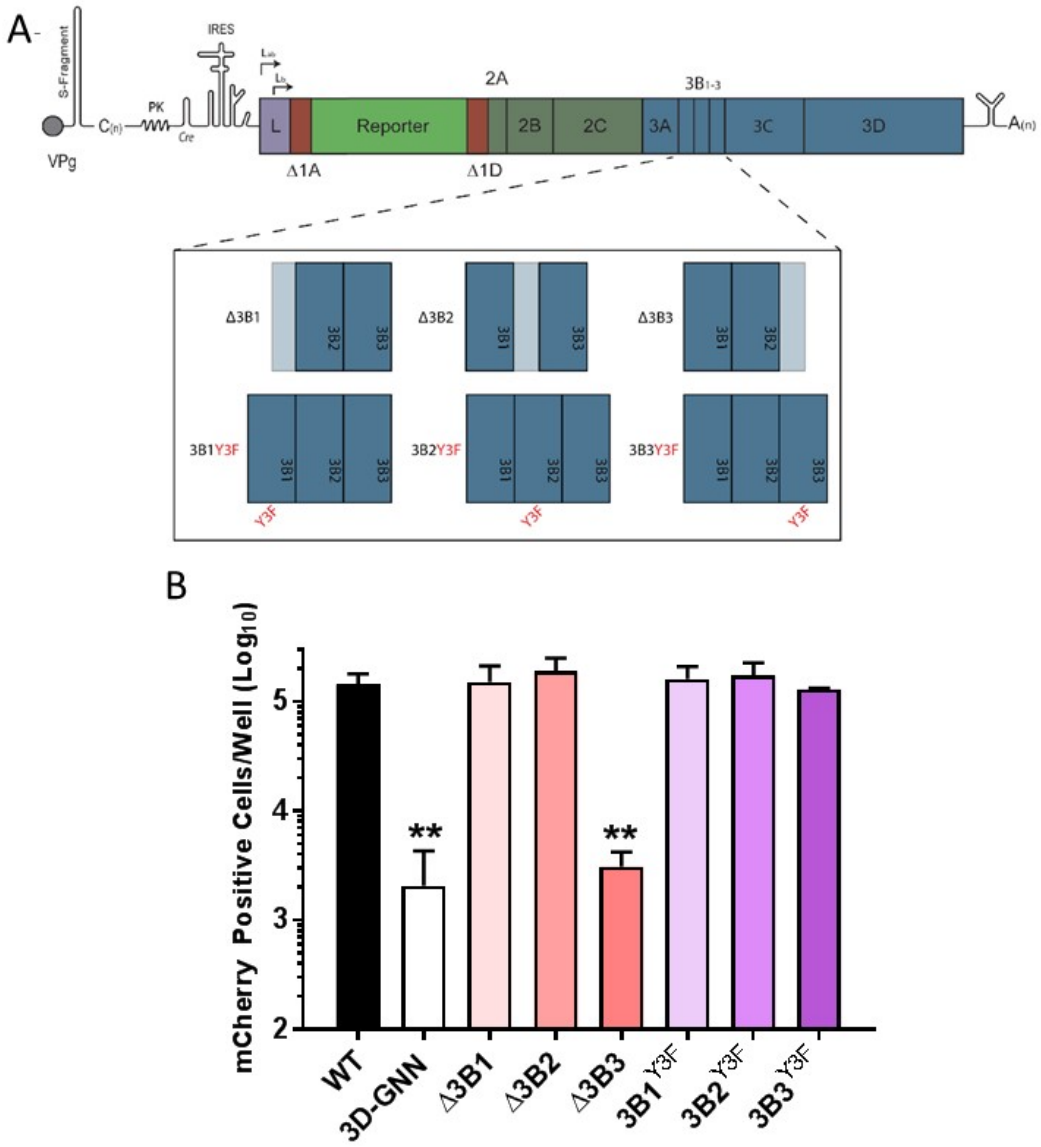

Figure 3.4: Sequences within 3B3 are essential for FMDV RNA replication

(A) Schematic of the FMDV sub-genomic replicons containing 3B deletions and tyrosine mutations, with the 3B region expanded for clarity. (B) Replicons containing 3B deletions or mutations were transfected in to BHK-21 cells alongside WT and 3D-GNN controls. Replication was monitored hourly and shown here at 8 hours post-transfection. Error bars represent SEM and significance is shown when compared to the WT control (n=3 error bars represent SEM, * = $p < 0.05$, ** = $p < 0.01$, statistical analysis performed by paired T-test). This work was performed by Dr Morgan Herod and myself.

3.2.2 The C-terminus of 3B3 is essential for replication

Since replication seemed dependent on the presence of 3B3, rather than its active state, it was hypothesised that the junction between 3B3 and 3C was the critical element. This was explored by swapping the C-terminus of 3B3 with 3B2 (3B3*2), and deleting 3B3, but replacing the C-terminus of 3B2 with 3B3 (3B2*3) (Fig 3.5A). These constructs were transcribed to make replicon RNA and transfected into BHK-21 cells alongside WT, 3D-GNN and Δ3B3 controls as before.

Replacement of the C-terminus of 3B3 with that of 3B2 (3B3*2), caused a reduction in replication equivalent to that of 3D-GNN and Δ3B3 negative controls, showing a clear need for the C-terminus of 3B3 which cannot be provided by the same region from 3B2. Confirming this, the replacement of the 3B2 C-terminus with that of 3B3 (3B2*3), while having 3B3 deleted, recovered the otherwise fatal deletion of 3B3 (Fig 3.5B).

Since the C-terminal of 3B3 was shown to be essential for replication, it was thought that conservation of the cleavage boundary between 3B3 and 3C might be the important factor to permit replication. To test this pcDNA expression constructs were made containing the WT or 3B3*2 P3 protein region. These constructs were used in *in vitro* coupled transcription-translation (TnT) systems to observe the processing of the P3 region when the replication impairing 3B3*2 mutation was introduced. P3 expression constructs were incubated in the TnT reagent along with ^{35}S cysteine/methionine radiolabel at 37^{o}C for 40 minutes, before addition of a cold cysteine/methionine chase. Initial incubation allows for both transcription and translation of the pcDNA constructs with incorporation of the ^{35}S cysteine/methionine label into the newly translated protein. Once chased with cold label, no further translated protein would incorporate label, allowing the visualisation of ^{35}S labelled P3 protein processing. Aliquots of labelled protein were taken every 20 minutes and electrophoresed on a 12 % SDS-PAGE gel before fixation, drying and visualisation by autoradiography as described in materials and methods 2.10 (Fig 3.6A).

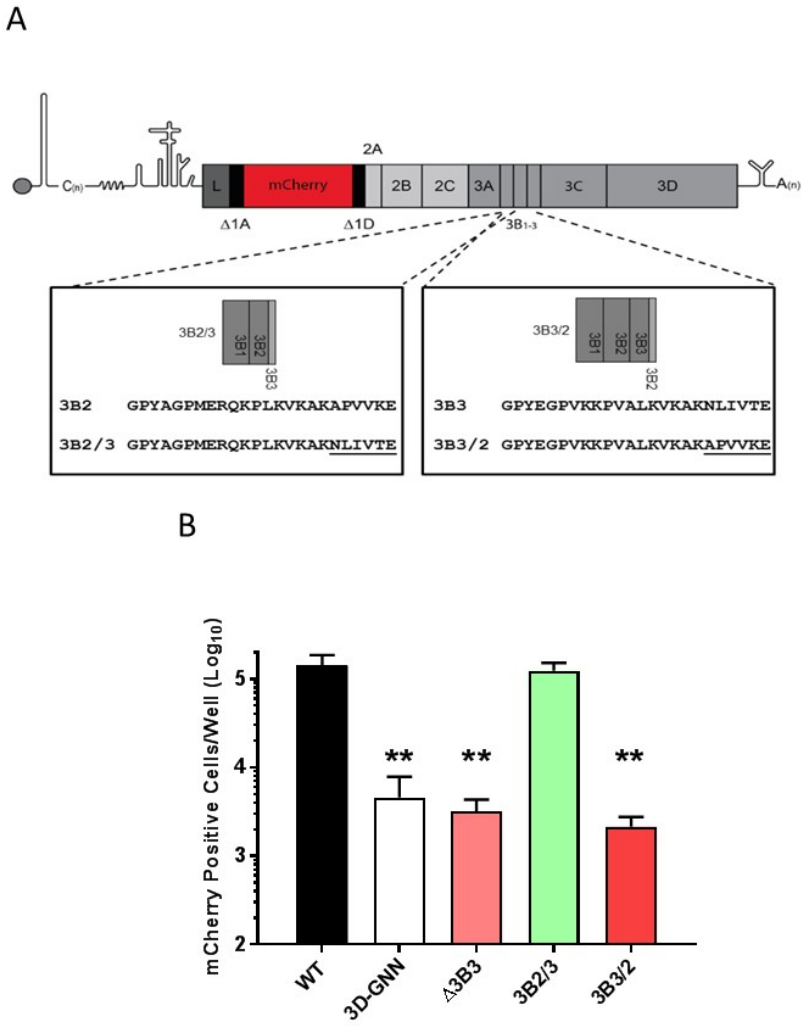

Figure 3.5: Residues at the 3BC boundary are important for replicon replication.

(A) Schematic of the FMDV replicon with chimeric 3B boundary mutations (3B2/3 and 3B3/2) showing the specific amino acid sequences of each mutation. (B) Replicons containing mutant 3B C-terminal regions where transfected into BHK-21 cells alongside WT, $3D^{pol}$-GNN and Δ3B3 constructs controls. Replication is shown at 8 hours post-transfection. Error bars represent SEM with, significance shown when compared to WT control (n=3 error bars represent SEM, * = $p < 0.05$, ** = $p < 0.01$). This work was performed by Dr Morgan Herod and myself.

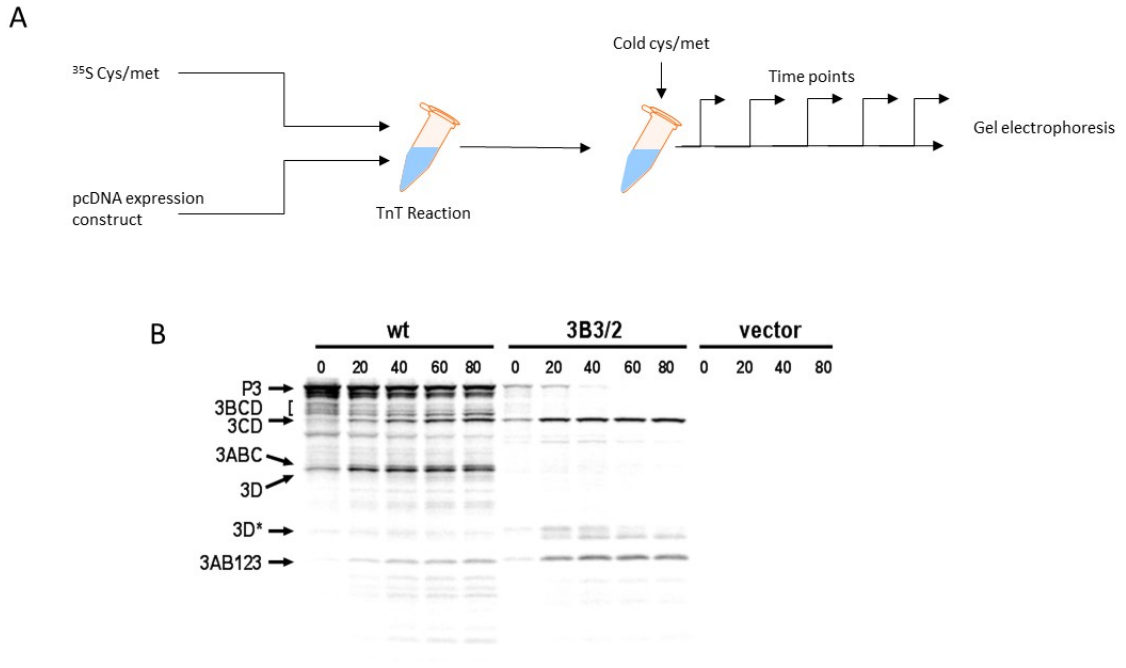

Figure 3.6: Mutations at the 3BC boundary disrupt P3 polyprotein processing.
Plasmid pcDNA expression constructs containing WT FMDV P3 or the P3 containing the 3B3/2 mutant polyprotein were used in transcription-translation reactions with ^{35}S labelled methionine. (A) Reactions were incubated for 40 minutes before addition of excess unlabelled methionine/cysteine as a chase, samples were taken at regular intervals and reaction stopped by the addition of 2 x Laemmli buffer. (B) Proteins were separated on 12 % SDS-PAGE and visualised by autoradiography. The positions of FMDV protein products are indicated by arrows. The 3D* product represents a degradation or cleavage product of 3D^{pol} as confirmed by Western blot (Appendix 5.9). Control reactions were assembled using empty expression vector alone. This work was performed by Dr Morgan Herod.

Proteins separated by SDS-PAGE electrophoresis were identified by western blot (Fig 3.7). For the WT P3 control, full length P3 precursor could be seen at 0 minutes after chase, as well as 3ABC, 3BCD, 3CD and fully processed $3D^{pol}$. Over time increased amounts of mature $3D^{pol}$, 3CD and 3AB123 was observed, alongside reductions in 3BCD and full length P3. In comparison to the WT, P3 precursors containing the 3B3*2 mutation produced very little full length P3, instead large amounts of 3CD and 3AB123 could be detected at early time points. Interestingly, little mature $3D^{pol}$ was observed, suggesting a reason why replicons containing this mutation were replication incompetent. These experiments suggested that the replacement of the 3B3 C-terminus with that of 3B2 changed the processing efficiency at the 3B-3C boundary, increasing the rate of proteolytic cleavage. The increased processing at this boundary in turn resulted in a defect in 3CD processing preventing the generation of fully processed mature $3C^{pro}$ and $3D^{pol}$ (Fig 3.6B). This highlights a role of 3B proteins, not only in the priming of replication, but also in ensuring the correct polyprotein processing of the P3 region, revealing a role in correct cleavage boundaries for generation of mature proteins.

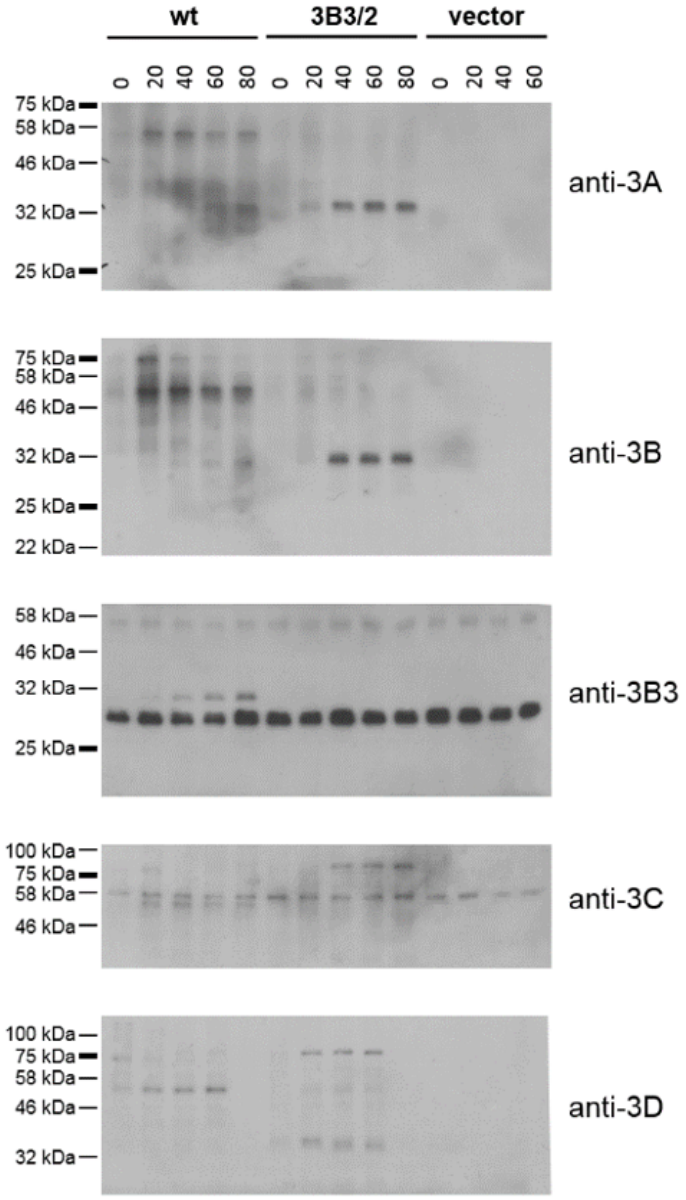

Figure 3.7: Confirmation of P3 proteins by western blot
Plasmid constructs expressing wild-type FMDV P3 or the 3B3/2 mutant polyprotein were used to assemble coupled transcription/translation reactions with 35S labelled methionine. At 20 minute intervals samples were taken and the reaction stopped by the addition of 2 x Laemmli buffer. Protein products were iseparated on 12% SDS-PAGE and probed by Western blot for expression of FMDV non-structural protein expression using different antibodies (labelled).

3.2.3 Disruption of the cleavage boundaries between P3 proteins

To investigate the role of cleavage boundaries between P3 proteins, individual cleavage sites were mutated between each protein in the P3 region, i.e between, 3A and 3B1, 3B1 and 3B2, 3B2 and 3B3, 3B3 and 3C and 3C and 3D.

The cleavage boundaries were chosen considering the literature on $3C^{pro}$ induced cleavage. Thanks to the work of Curry et al, the specificity of $3C^{pro}$ for cleavage sites has been well documented[142]. Peptide cleavage assays revealed that the $3C^{pro}$ recognition site sequence spans at least four residues either side of the cleavage boundary, residues up stream of this boundary are named P5-1, and residues downstream P1′-P5′ as seen in Figure 3.8. The substitution of wild type residues for alanine residues in the P1 and P2 positions (P1 and P4 for the cleavage site between 3A and 3B1) were thought to be sufficient to disrupt $3C^{pro}$ recognition and inhibit cleavage between these proteins (Fig 3.8).

To ensure cleavage boundary mutations would effectively inhibit $3C^{pro}$ cleavage, pcDNA expression constructs were made as described in section 3.2.2. Here P3 regions containing individual cleavage mutations were made and introduced into the TnT system as described before. Due to the small size of the 3B proteins (2.6 kDa), fusion proteins containing 3B would only cause subtle differences to the overall polyprotein processing, which might not be discernible on an SDS-PAGE gel (i.e +/- 2.6kDa). Consequently a construct was made including all cleavage mutations into one P3 region, creating a construct containing a completely uncleavable P3 (P3*). This was used alongside individual cleavage mutations to see if any processing could take place.

TnT reactions were separated on 12 % SDS-PAGE gels and visualised by autoradiography. Samples containing the cleavage mutation between 3B3 and 3C (3B3/C), were electrophoresed in parallel to the WT P3 and P3 3B3*2 samples, to allow easy

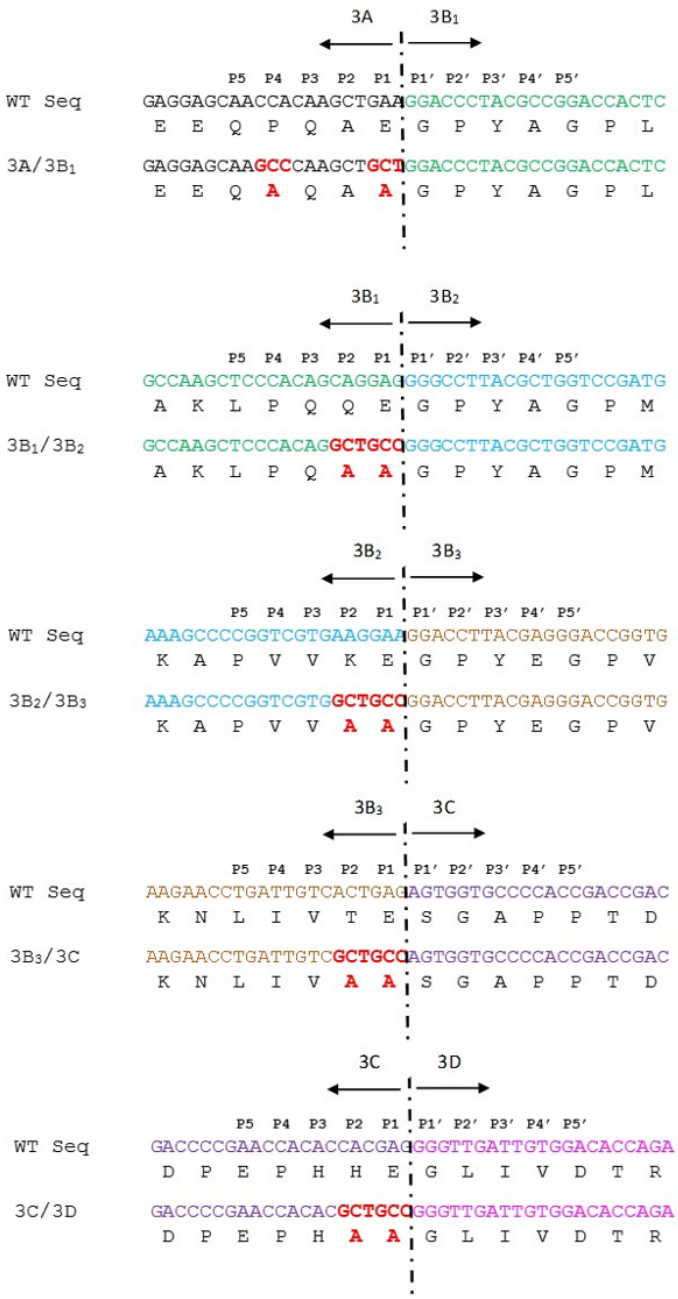

Figure 3.8: Comparative protein boundaries between P3 proteins, with mutations to prevent 3C induced cleavage.

Nucleotide and amino acid sequences shown of the boundaries between O1K P3 proteins. P1-P5 indicate positions 5′ of the cleavage site and P1′-P5′ indicate positions 3′ to the cleavage site. Nucleotides and residues in bold red represent mutation to disrupt the cleavage boundary.

comparison of mutations in this region. It was feared that the 3B3/C mutation would cause a similar defect in processing as 3B3*2, however, processing looked similar to that of the WT. As described above, it is unlikely that, unless major processing abnormalities occurred, to visualise any differences in protein size due to the small size of 3B3. Noticeably however, the band representing 3AB123 could not be seen (Fig 3.9A).

In the 3A/B1 TnT reaction, all major bands such as 3D and 3B123 could be seen. However, a slight variation in the higher molecular weight bands was observed (represented by the red star). This could refer to the protein precursor 3BCD and due to the disruption of the 3A/B1 boundary likely refers to 3B23CD. Reactions corresponding to 3B1/2 and 3B2/3 showed little difference to that of the WT with all major bands seen. It is likely that differences would be seen with these constructs with the ability to visualise individual 3B proteins. The expression construct containing P3 3C/D as expected showed no sign of mature $3D^{pol}$, and as such would be predicted to be non-replicative (Fig 3.9B and C).

Compiling all mutations into a single P3 region (P3*), showed a unique protein processing pattern, not seen in any of the other conditions. Many protein bands were produced, making it difficult to ascertain the identity of the proteins. It is possible, that since the normal cleavage boundaries were disrupted, but $3C^{pro}$ was still active that cleavage occurred at other sites within P3 to create many aberrant proteins. There was, however, noticeable enrichment of a collection of proteins of a similar size to a protein seen in the 3A/B1 construct (indicated by *). This could be a variation of 3BCD (i.e 3B123CD, 3B23CD, 3B3CD), suggesting that in the P3* construct $3C^{pro}$ could still be cleaving at the 3A/B1 boundary or somewhere within 3A. When compared to inactivation of the $3C^{pro}$ active site (3C C163A), a different protein pattern was observed, where 3C C163A showed little protein processing. This supports the conclusion that within the P3* construct the $3C^{pro}$ is still active, although it may not be cleaving at authentic cleavage sites (Fig 3.9B).

In combination, the TnT experimentation suggested there was sufficient difference in processing patterns to suggest disruption of $3C^{pro}$ induced cleavage. As such it was decided to take these mutations forward into replicons to see any changed in replication .

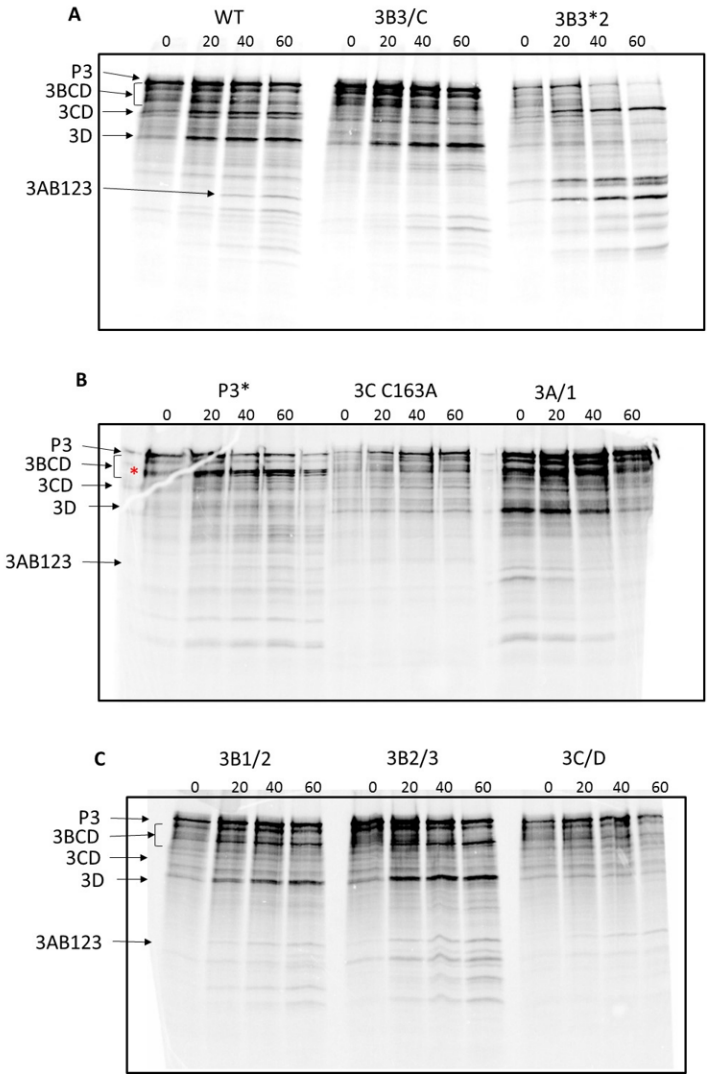

Figure 3.9: Investigating P3 cleavage mutations.
Plasmid pcDNA expression constructs containing different P3 mutations; A: WT, 3B3/C, 3B3*2; B: P3*, 3C C163A, 3A/1; C: 3B1/2, 3B2/3, 3C/D, were used in transcription-translation reactions with ^{35}S labelled methionine. Reactions were incubated for 40 minutes before addition of excess unlabelled methionine/cysteine as a chase, samples were taken every 20 minutes and reaction stopped by the addition of 2 x Laemmli buffer. Proteins were separated on 12 % SDS-PAGE and visualised by autoradiography. The positions of FMDV protein products are indicated by arrows. * indicates position of proteins related in the main text.

3.2.4 Effects of disruption of cleavage boundaries on replicon replication

Once the cleavage mutations were shown to be functioning as expected the individual mutations were introduced into the DNA copy of the FMDV GFP containing replicon (pRep-ptGFP). The pRep-ptGFP plasmid containing the cleavage boundary mutations could then be transcribed creating replicon RNA possessing cleavage boundary mutations. These replicon RNAs were then transfected into BHK-21 cells.

Replication of replicons containing mutant cleavage boundaries (shown in Figure 3.8) was analysed alongside WT and 3D-GNN controls, as well as a negative control for 3B function, the 3B123 Y3F replicon, where all copies of 3B are inactivated with the tyrosine to phenylalanine mutation. Replication is shown at 8 hours post-transfection when GFP expression is at its peak, as measured by the Incucyte. It was observed that while there was a noticeable difference between our WT and replication deficient controls, introduction of the cleavage mutations between 3A and B1, B1 and B2 and B2 and B3 caused no reduction in replication. Fusing the 3B3 and $3C^{pro}$ proteins together caused a slight decrease in replication, but the level was still approximately 70 times higher than that of the negative controls, suggesting replication is still occurring. Unsurprisingly, fusing $3C^{pro}$ and $3D^{pol}$ together caused a replicon that was incapable of replicating its genome, as no mature $3D^{pol}$ could be released (Fig 3.10).

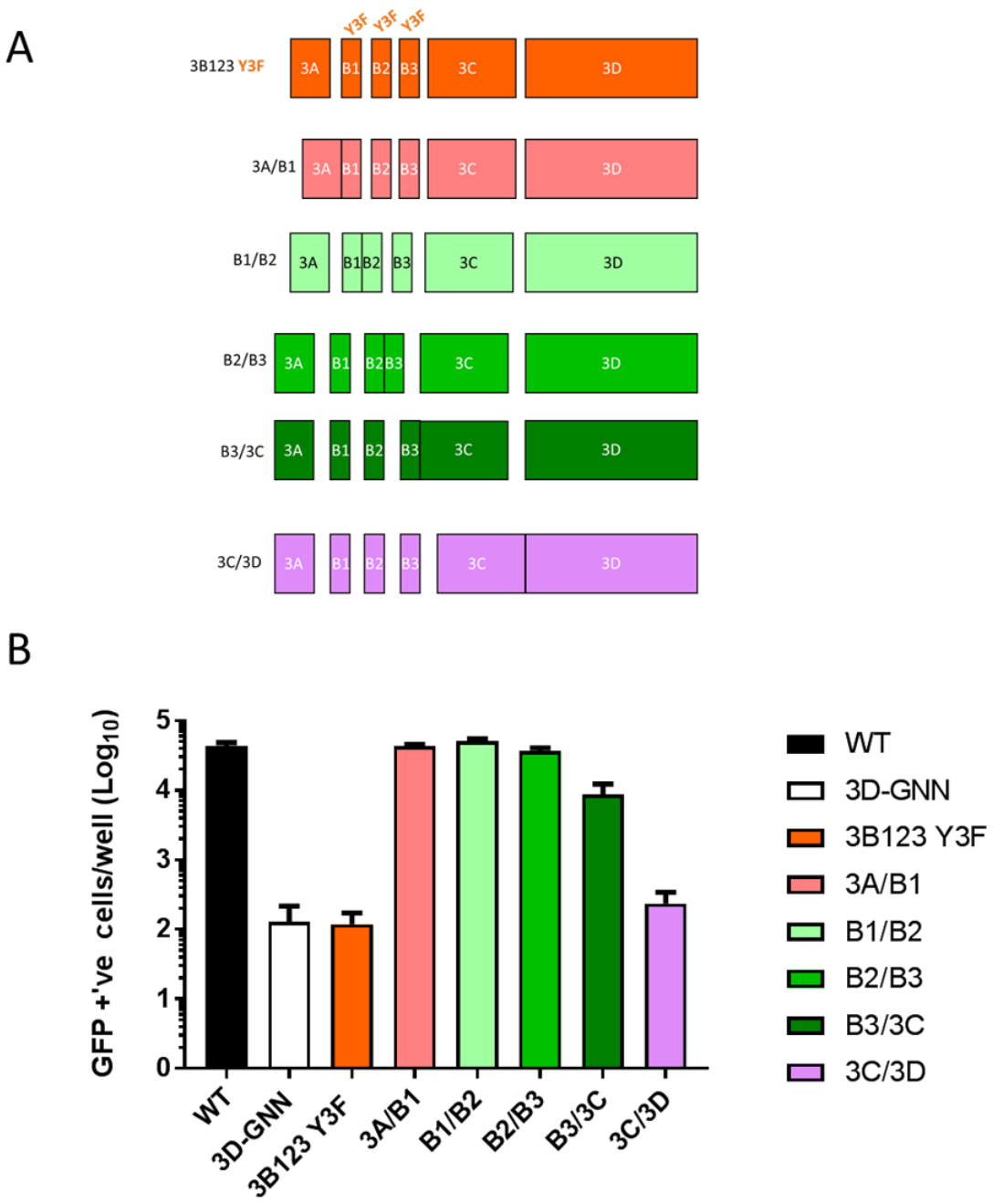

Figure 3.10: Mutation of cleavage sites is tolerated between many P3 Proteins.

A. Schematics of modified cleavage boundaries between individual proteins in the P3 region. B. Mutant replicon RNA was transfected into BHK-21 cells alongside wild-type (wt), 3D-GNN and 3B123Y3F constructs as positive and negative controls. Replication shown here at 8 hours post transfection. Error bars represent SEM n=2.

As the constructs described above have additional 3Bs which could be used to prime genome replication, it is impossible to say if the fused 3Bs are functional or not. To investigate if fused 3B proteins still function as primers of replication, the cleavage mutations were introduced alongside inactivating 3B tyrosine mutations. These constructs were produced so that the only 3B with a functional tyrosine was the one fused to another protein.

These additional inactivating mutations were introduced into the pRep-ptGFP plasmid and then transcribed to produce replicon RNA, replicon RNA was then transfected into BHK-21 cells as described before. Replication of these mutant replicons is shown again at 8 hours post-transfection alongside WT, 3D-GNN and 3B123 Y3F controls (Fig 3.11). A dramatic change in the performance of some of the fusion mutants was observed. The replicon containing the 3A/B1 fusion, which showed wild type levels of replication in the initial experiment (Fig 3.10) was reduced to a similar level as the negative controls upon introduction of the Y3F mutation to 3B2 and 3B3. This suggests that when 3B2 and 3B3 were inactivated, fusing 3A to 3B1 prevents the use of 3B1 as a functional primer of replication.

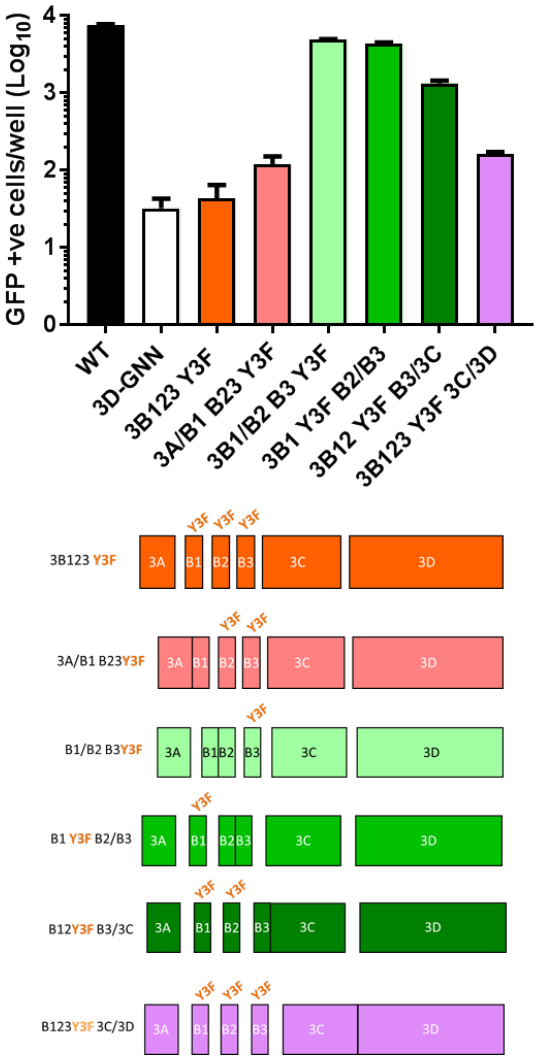

Figure 3.11: Combining cleavage boundary mutations with 3B tyrosine mutations reveals 3B proteins can tolerate fusion.
Replicons with mutated cleavage boundaries between P3 proteins were combined with inactivating tyrosine mutations in non-fused 3B proteins. Replicons were transfected into BHK-21 cells alongside wild-type (WT), 3D-GNN and 3B123Y3F constructs as positive and negative controls. Replication shown here at 8 hours post transfection. Error bars represent SEM, n=2.

In contrast, the fusion of 3B1 to 3B2 while inactivating 3B3 (3B1/2 3B3 Y3F) as well as inactivating 3B1 and fusing 3B2 and 3B3 (3B1 Y3F 3B2/3) showed replication at similar levels to the wild type replicon. As these constructs only produce one functional 3B molecule, albeit in the context of a fusion protein, the data supports previous findings that only one molecule is necessary for function. Interestingly unlike the 3A/B1 fusion, these findings suggest that fusion to a 3B protein does not inactivate its function. Although it can be argued that 3A is a much larger protein than 3B therefore this fusion could have a much larger deleterious effect than the fusion of two 3Bs together.

Surprisingly fusing 3B3 to $3C^{pro}$, whilst inactivating 3B1 and 3B2 (3B12 Y3F 3B3/C) resulted in similar phenotype as when 3B1 and 3B2 were not inactivated. These data suggest that 3B3/C can act as a functional primer for replication. This is particularly interesting as 3A, a similar sized protein, when fused to 3B1, inhibited its function as a primer of replication. Of course, size is just one aspect of the effect that a fusion protein could be having on 3B functionality, the actual protein fused could have a significant impact. The difference we are observing here could be due, for example, to the 3A protein having a trans-membrane domain, targeting 3B1 away from its function as a primer and embedding it in a membrane, or the difference in position of fusion with 3A being N-terminal and $3C^{pro}$ being C-terminal.

Due to the replication deficiency of the 3C/D fusion replicon, as predicted, inactivation of the 3Bs did not cause a change in the phenotype observed.

The difference between the fusion of 3A and $3C^{pro}$ on the functionality of 3B proteins was investigated further. We have previously published that removal of 3B1 has no obvious effect on replicon replication in BHK-21 cells, so could be removed with no downstream effects. Therefore, for simplicity, 3B1 was removed from the constructs described here (Δ3B1)[79].

The Δ3B1 construct was used as the template to investigate if the difference seen with the 3A vs $3C^{pro}$ fusion was due to the position of the fusion protein being on

the N or C-terminus of the 3B protein. To test this hypothesis, 3B2 was fused to 3B3, creating the Δ3B1 3B2/3 replicon. Inactivating tyrosine, Y3F mutations, were then introduced to either the 3B2 (Δ3B1 3B2 Y3F/3B3) or the 3B3 (Δ3B1 3B2 Y3F/3B3) or to both 3B2 and 3B3 (Δ3B1 3B2 Y3F/ 3B3 Y3F). These constructs were transcribed and replicon RNA transfected into BHK-21 cells, alongside WT and 3D-GNN controls.

As seen in the previous experiments, fusion of 3B2 and 3B3 created a replicon that performed to a similar level as the WT. Similarly, fusing 3B2 and 3B3 whilst inactivating the 3B3 with a tyrosine mutation, showed wildtype levels of replication. In contrast, when inactivating the 3B2 in the 3B2/3 fusion, replication decreased to that of the 3D-GNN negative control. This phenotype was also observed when inactivating both 3B2 and 3B33.12.

These results suggested that while the 3Bs could still perform their function as primers of replication with C-terminal fusions, fusions to their N-terminus resulted in an inhibition of function. These findings explain the result seen in the previous experiments, where 3A fusion prevented replication while replication was still observed in the $3C^{pro}$ fusion. As such it is essential for the N-terminus of 3B to be available for its function to be carried out. It is unknown however, if availability of the N-terminus is essential for the uridylylation of a 3B molecule or for its role in priming replication.

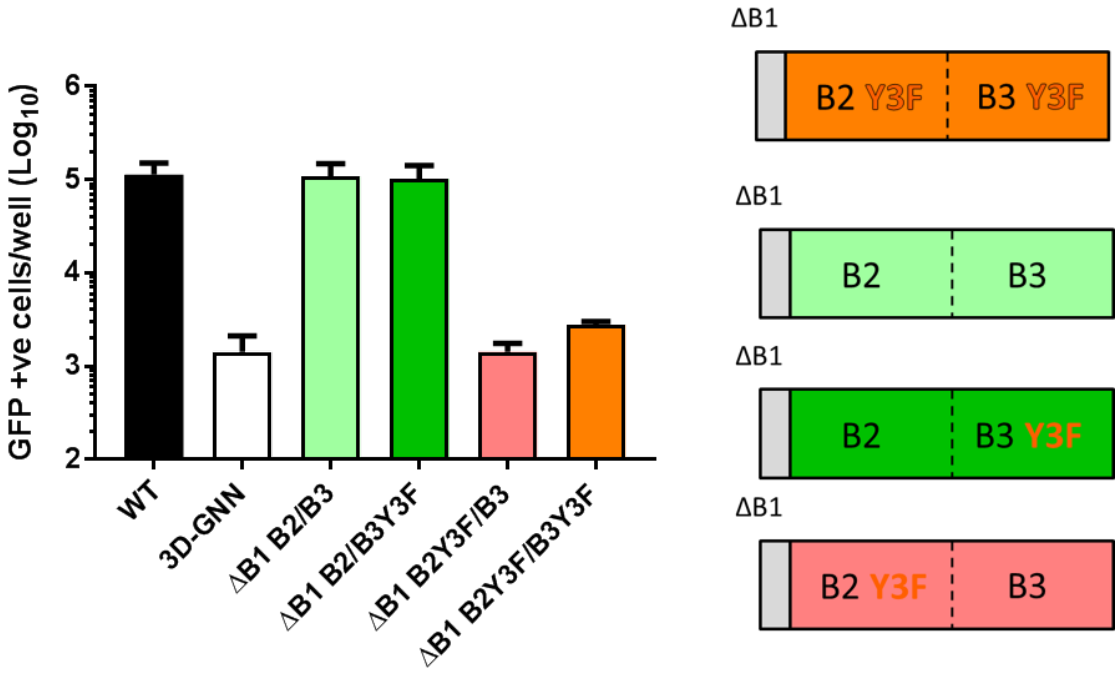

Figure 3.12: 3B proteins can tolerate C but not N-terminal fusion. Constructs containing a deleted 3B1 (Δ3B1) and cleavage boundary mutation between 3B2 and 3B3, were combined with inactivating tyrosine mutations to either the N or C-terminal 3B, or both. Mutant replicons were transfected into BHK-21 cells along side WT and 3D-GNN positive and negative controls. Error bars represent SEM, n=2.

3.2.5 Can fusion products be used in trans-complementation experiments?

A benefit of using a replicon system is the relative ease with which certain experiments can be performed. One of these is co-transfection experiments, by using two replicons with differential reporters, such as GFP and mCherry, it is easy to distinguish replication of two constructs within the same experiment and see how the replication of one might affect the replication of the other (Fig 3.13).

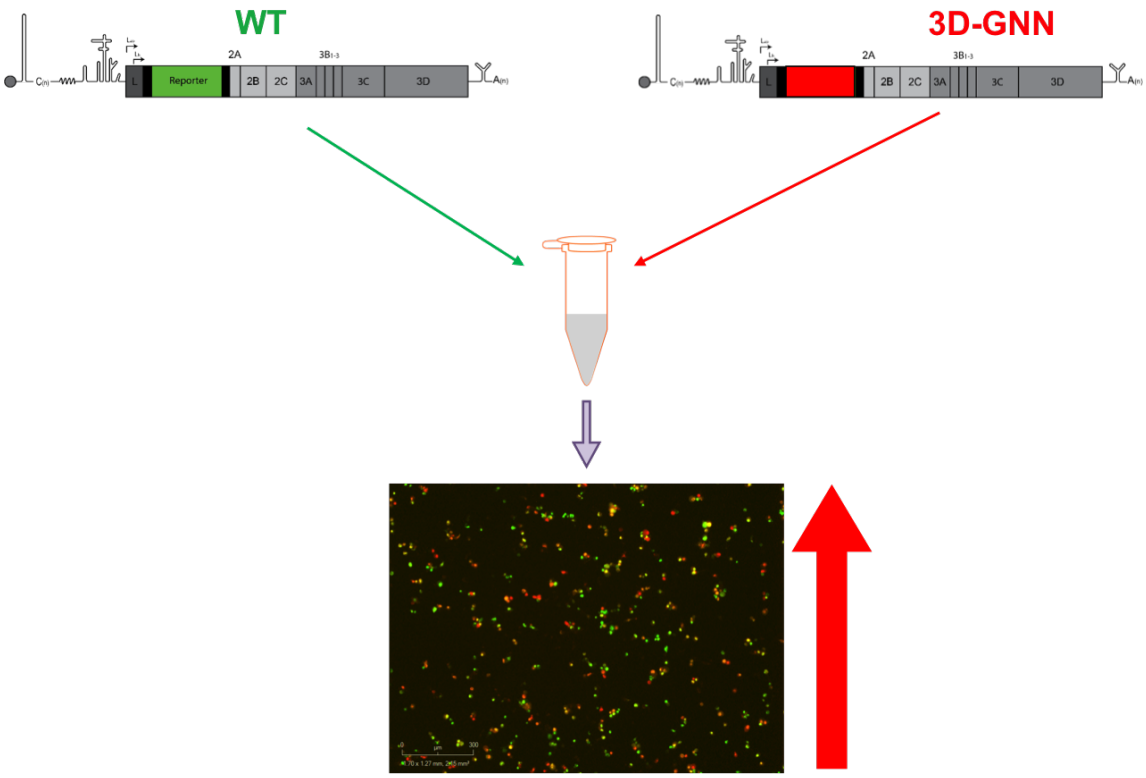

Figure 3.13: Schematic of replicon co-transfection experiments.
Co-transfection experiments are performed by mixing of GFP and mCherry containing replicons before transfection into BHK-21 cells. Complementation can be observed by the detection of an increase in fluorescent protein signal from the mutant replicon. In this example the WT GFP replicon would cause an increase in mCherry GNN signal as the 3D protein is provided in trans.

In previous publications by the Stonehouse/Rowlands group we have used this technique to great effect to show mutations in $3D^{pol}$, such as those in the polymerase active site, could be provided in trans by another replicon (Fig 3.14)[78]. Therefore, employment of this strategy here could elucidate if fusion proteins could be used to reveal functions of proteins or precursor proteins by using them to complement replicon deficient constructs.

GFP replicons containing fusion proteins where used in co-transfection experiments to test if they could provide help (or 'rescue') to mCherry replicons containing non-active 3D polymerases (3D-GNN) or no functional 3Bs (3B123 Y3F). Initial ex-

periments used GFP replicons containing functional fusion proteins. These were co-transfected alongside WT, 3D-GNN or 3B123 Y3F mCherry replicons, or yeast tRNA as a control.

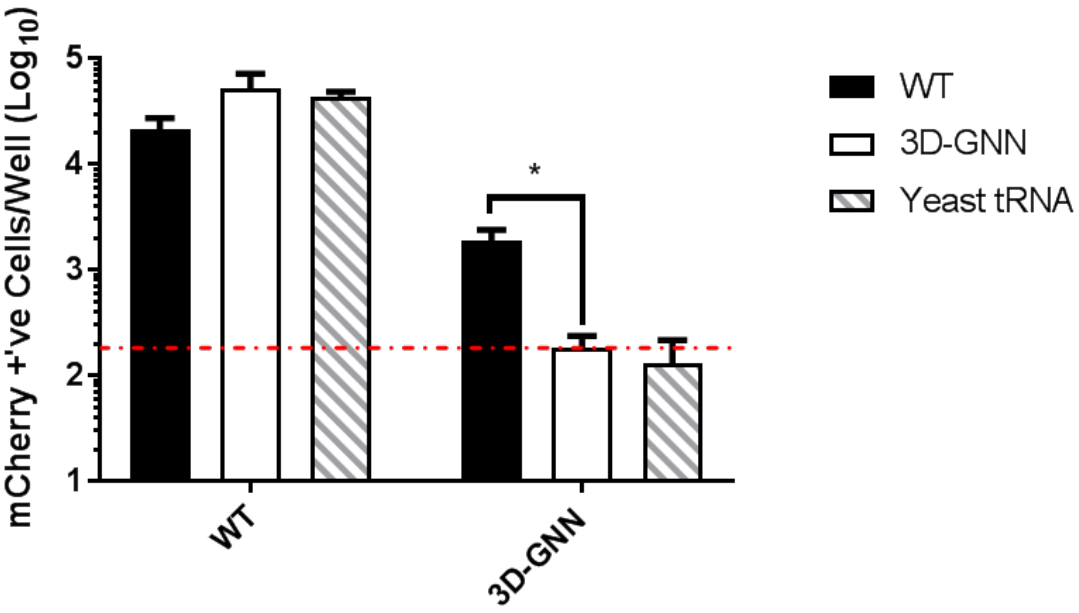

Figure 3.14: Co-transfection of WT and 3D-GNN replicons causes an increase in 3D-GNN signal.
WT and 3D-GNN GFP replicons were co-transfected with WT and 3D-GNN mCherry replicons. Yeast tRNA was used as a negative control for co-transfection, showing the replicative ability of replicons in the absence of complementation. An increase in mCherry 3D-GNN signal was observed when co-transfected with the WT GFP replicon, but not when co-transfected with 3D-GNN GFP replicon or yeast tRNA. Red dotted line indicates the threshold for complementation, error bars indicated SEM, n=3, * = $p < 0.05$, statistical analysis performed by paired T-test. [78]

The threshold for replication was set by observing reporter expression of the negative controls, for example co-transfection of a GFP 3D-GNN and mCherry 3D-GNN should produce no replicative complexes as no functional $3D^{pol}$ is present, yeast tRNA is also used as a control for replicon-replicon effects and shows the fluorescent protein expression generated by just one replicon whilst keeping the amount of RNA

transfected equivalent. The threshold for replication is represented by the red dashed line overlaid on the replication graphs.

Figure 3.15A shows the replication of these mCherry replicons when co-transfected with the GFP fusion mutants, as well as GFP WT, 3D-GNN and 3B123 Y3F controls. It is evident that none of the GFP replicons had any effect on the mCherry WT replicon, showing they possess no dominant negative effects. However, when the GFP replicons were co-transfected alongside the 3D-GNN mCherry replicon a large increase in mCherry expression was observed. The WT, 3A/B1, 3B1/2, 3B2/3 and 3B3/C replicons all increased the mCherry expression of the 3D-GNN replicon, raising this above that of the threshold for replication. This recovery of 3D-GNN replicon is not surprising as previous data infers that all the GFP P3 fusion replicons can produce an active $3D^{pol}$, with the exception of the 3C/D fusion. However, this does reveal that these fusions do not alter the ability of the replicon to be able to provide 3D in trans. Figure 3.15B, shows the other side of this experiment, i.e. looking at the change in the GFP expression of the fusion replicons when co-transfected with mCherry replicons. Here, as most of the fusion replicons are replication competent, there is little change when co-transfected with the mCherry constructs. The only exception to this is the 3C/D fusion replicon which is effectively rescued by the mCherry 3B123 Y3F and mCherry WT replicons, presumably due to the ability of these constructs to produce an active $3D^{pol}$ and provide this in trans.

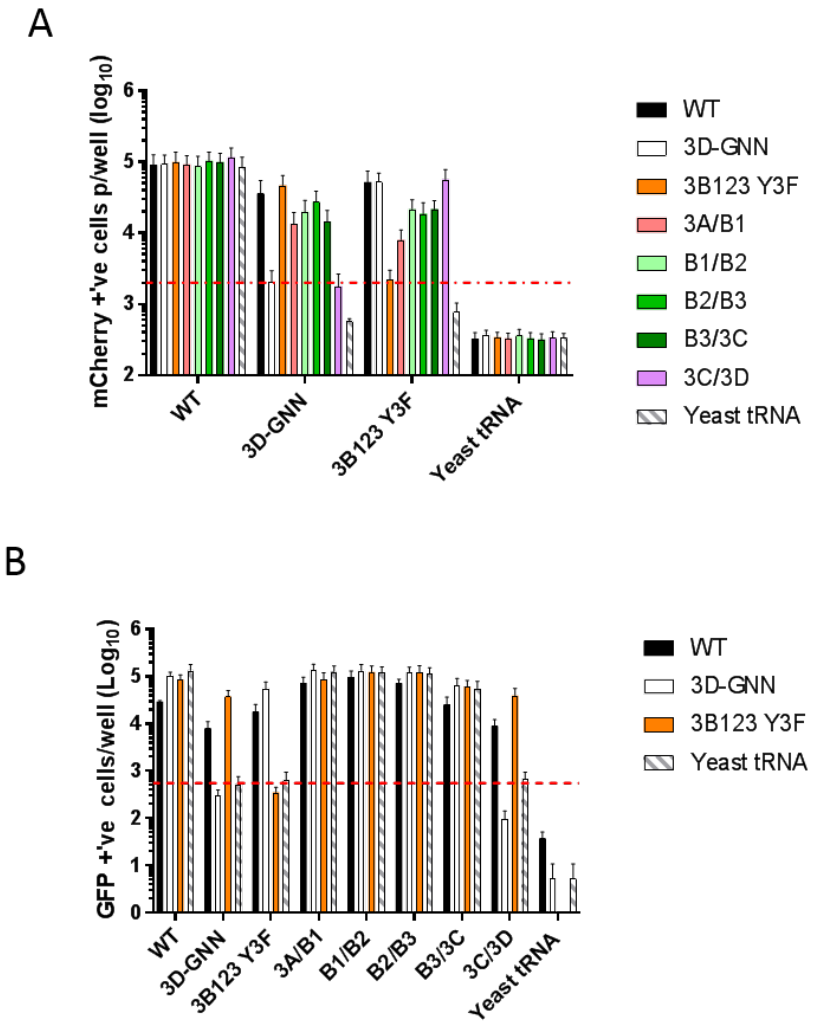

Figure 3.15: Complementation of replicons containing fusion proteins could rescue 3D-GNN and 3B123Y3F replicons.
Complementation experiments involved co-transfecting mCherry replicons containing replication-defective 3D-GNN or 3B123Y3F mutations, with wild-type GFP helper replicons containing mutated cleavage boundaries between P3 proteins or yeast tRNA. BHK-21 cells were co-transfected with mCherry and GFP replicons with yeast tRNA used as a negative control. Fluorescent protein production was monitored hourly and shown here at 8 h post-transfection, mCherry (A) and GFP (B). Red dotted line indicates the threshold for complementation, error bars represent SEM, n=2.

The complementation experiments were continued using the fusion mutations in combination with the 3B Y3F mutations seen in Fig 3.11. Constructs were generated such that the only 3Bs with functional tyrosines were in the context of a larger 3B fusion, i.e 3B-B, 3A-3B or 3B-3C.

When co-transfecting these replicons with the mCherry 3D-GNN construct, a similar increase was observed to that seen previously, with all fusion GFP replicons causing an increase in the mCherry expression except the 3B123 Y3F 3C/D replicon which as described previously should not be able to produce any mature $3D^{pol}$ (Fig 3.16A).

However, when co-transfecting these fusion replicons alongside the mCherry 3B123 Y3F replicon a different phenotype was observed. Previously all fusion products were seen to increase the mCherry expression, but, with the inactivation of non-fused 3Bs an increase in mCherry expression was no longer observed. This was surprising as some of the constructs could replicate alone, as previously shown (Fig 3.11), but could not provide their 3Bs in trans to another replicon (Fig 3.16A).

Looking at the GFP expression in this experiment it can be noted that, as seen before, there is no discernible difference with regards to replication of the fusion constructs. The only difference observed was with absence of the recovery of the 3B123 Y3F 3C/D replicon by the mCherry 3B123 Y3F replicon. This was as expected due to the fact the 3B123 Y3F replicon could not provide both functional 3D and 3B to this construct (Fig 3.16B).

3.2.6 Two copies of 3B are required for the sharing of 3B proteins in trans

It was unclear why constructs containing active 3B proteins and which could replicate themselves, could not provide 3B in trans to another replicon. The initial hypothesis (from the data shown in figure 3.16), was that recovery of the mCherry 3B123 Y3F replicon was prevented because the number of functional copies of 3B has been

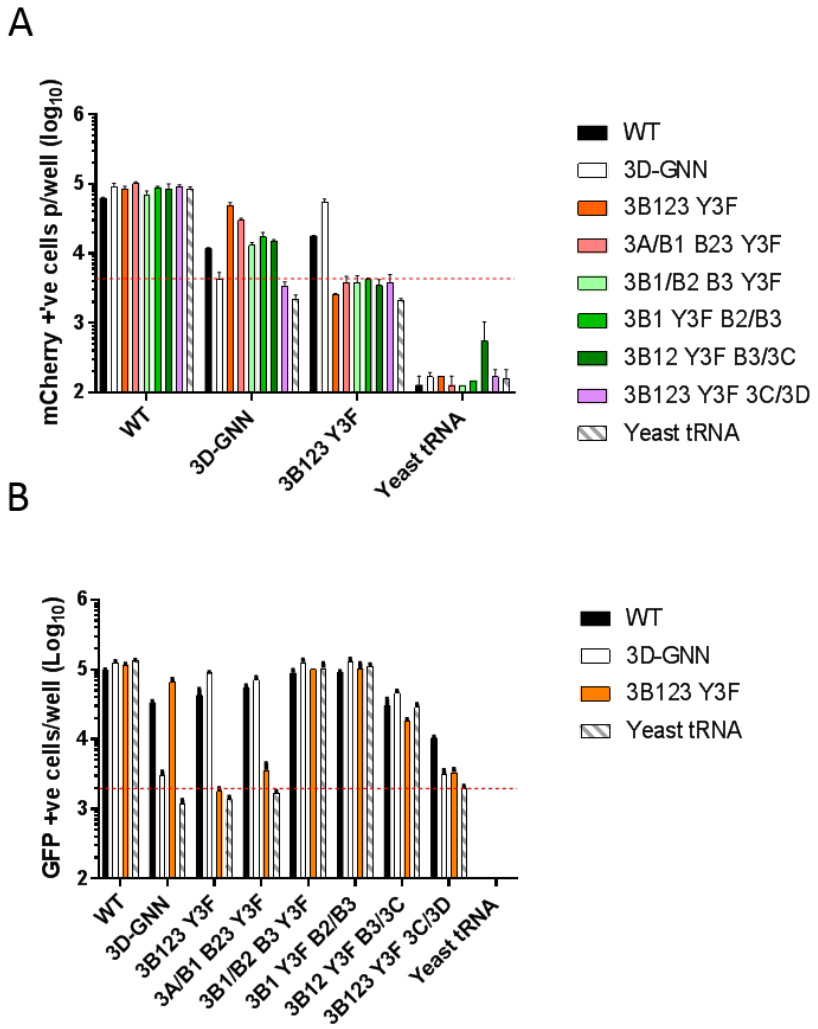

Figure 3.16: Introduction of Y3F mutations to 'free' 3B proteins prevented the recovery of 3B123 Y3F replicons.
Co-transfection of GFP replicons containing disrupted cleavage boundaries between individual P3 proteins in combination with tyrosine, Y3F knock-out mutations to non-fused 3B protein was performed with mCherry WT, 3D-GNN and 3B123Y3F replicons. Yeast tRNA was used as a negative control for complementation. Co-transfections were performed in BHK-21 cells and fluorescent protein production monitored over 24 hours. GFP (B) and mCherry (A) expression shown here at 8 hours post-transfection. Red dotted line indicates the threshold for complementation, error bars represent SEM, n=2.

reduced to just one. While we know that one 3B is sufficient for replication, it is not known if there is a minimal requirement for providing these in trans.

To test this hypothesis, mCherry expressing replicons with inactivated 3Bs by mutation of the essential tyrosine residues were used. This was performed on each individual copy of 3B; 3B1 Y3F, 3B2 Y3F and 3B3 Y3F, as well as in combination, 3B1+2 Y3F, 3B1+3 Y3F and 3B2+3 Y3F, the latter constructs leave just one functional copy of 3B.

To check that mutating 3B tyrosines still permitted replication, mCherry replicons containing 3B tyrosine mutations were transfected individually alongside WT, 3D-GNN and 3B123 Y3F controls. Expression of the mCherry reporter, shown at 8 hours post transfection confirms that reducing the number of active 3B copies to one still permitted wild type levels of replication (Fig 3.17A). These mCherry replicons were then used in similar complementation experiments to those described above, co-transfecting them alongside, WT, 3D-GNN and 3B123 Y3F GFP replicons. As expected all mCherry replicons recovered the 3D-GNN GFP replicon, except for the mCherry 3D-GNN replicon. However, when co-transfected alongside the GFP 3B123 Y3F replicon, only those constructs with more than one functional copy of 3B could provide help in trans to the GFP construct and result in an increase in GFP expression (Fig 3.17B). This result complements the experiments described previously, supporting the idea that although one copy of 3B is sufficient for replication in cis, more than one is required to be able to provide that 3B trans. This experiment also shows that there was no preference as to which 3B was provided in trans, suggesting they all perform the same role.

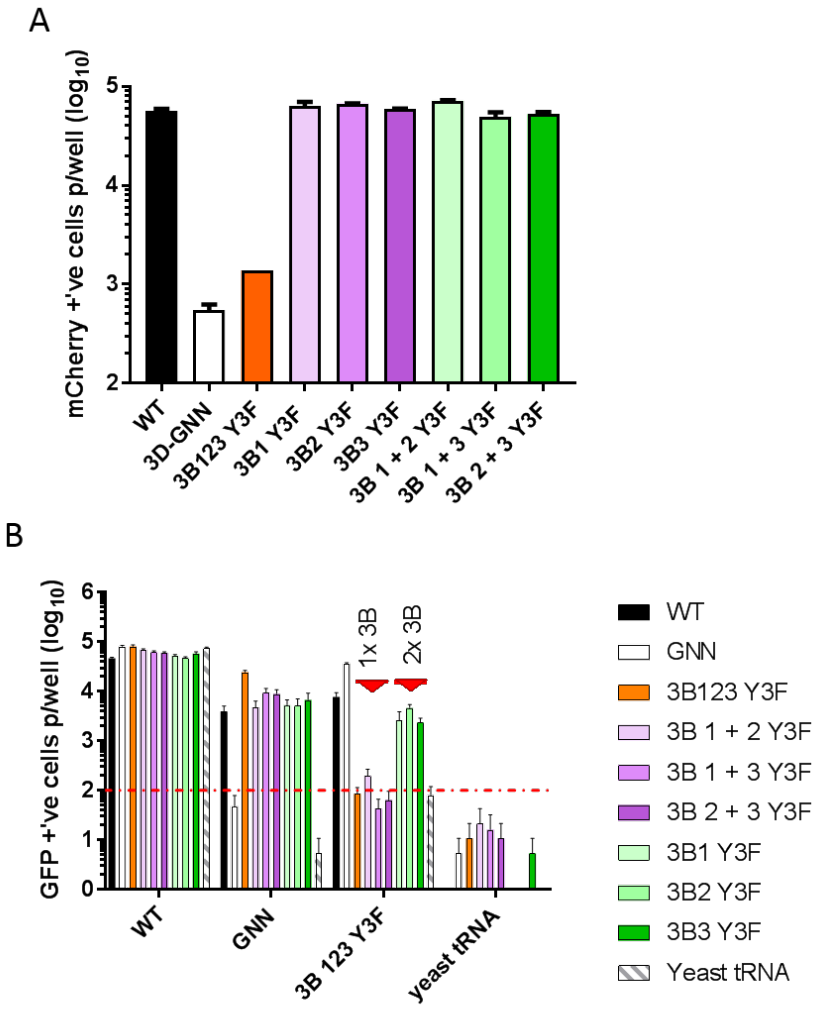

Figure 3.17: More than one copy of 3B is required to enable complementation in trans.

A. BHK-21 cells were transfected with mCherry replicons containing 3B tyrosine mutations (Y3F) as well as wild-type (wt) and replication-defective polymerase mutant (3D-GNN) controls. B. Co-transfection experiments were performed to determine the minimal requirements for 3B molecules to be shared between replicons in trans. An mCherry replicon with no functional copies of 3B (3B123Y3F) was co-transfected with either a WT GFP replicon or 3B123Y3F GFP replicon (as positive and negative controls) or with replicons containing either 1 or 2 active 3Bs (represented by '1x 3B' or '2x3B' labels on the graph). Replication shown at 8 hours post transfection, red dotted line indicates the threshold for complementation, error bars represent SEM, n=2.

While this work supports the initial hypothesis that for FMDV a single functional 3B cannot provide help in trans, it begs the question of how this relates to picornaviruses with only one copy of 3B. Can these provide their 3Bs in trans or is this an ability that FMDV has evolved alongside its acquisition of multiple copies of 3B? To investigate the capacity of other picornaviruses to provide their 3Bs in trans we used a PV replicon. This work was made possible due Dr Lee Sherry, a post-doctoral researcher in the group, who produced an efficient poliovirus replicon systems to complement the FMDV work (Fig 3.18).

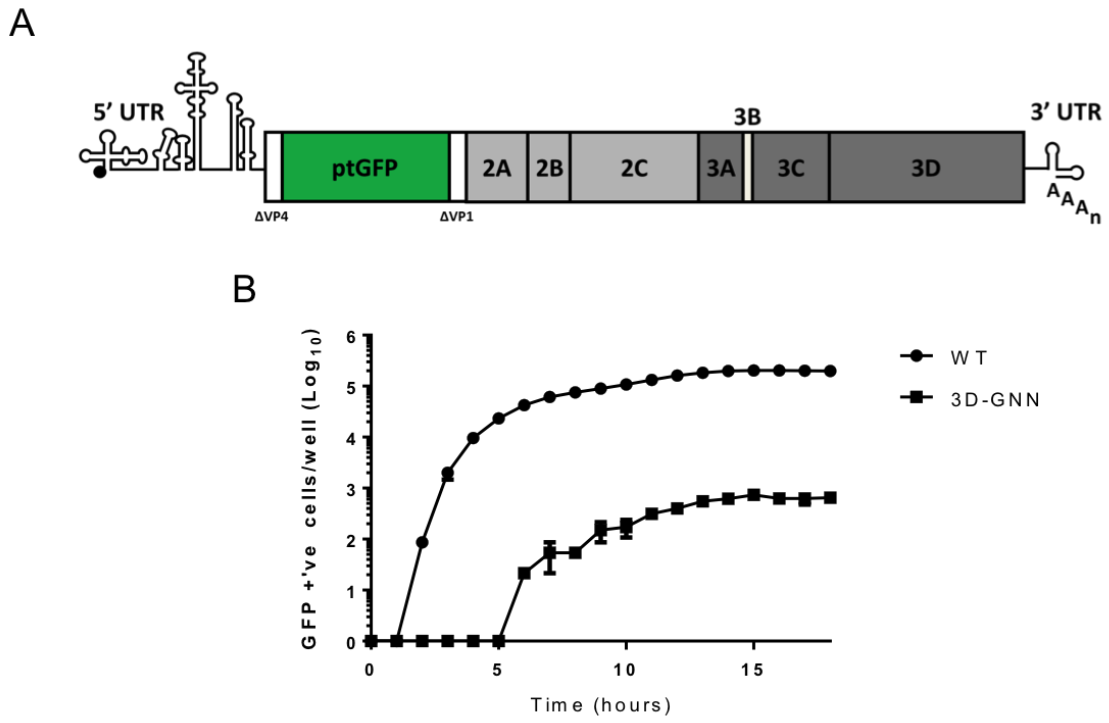

Figure 3.18: Schematic of PV replicon genome.
A. Cartoon representation of the PV replicon genome showing the single open reading frame flanked by 5′ and 3′ UTRs. The replicon was produced by replacement of the structural proteins with a GFP or mCherry fluorescent reporter. The single 3B protein encoded by PV is shown in the P3 non-structural region. B. Replication curve of PV WT and 3D-GNN replicon in HeLa cells, monitored by expression of GFP reporter, n=2.

3.2.7 PV can share 3B in trans to another replicon

A simpler version of the same trans-complementation experiments described above was designed. WT or 3BY3F GFP poliovirus replicons were co-transfected with WT or 3D-GNN mCherry replicons, to evaluate if these WT or 3D-GNN mCherry replicons could provide 3B to the 3BY3F GFP poliovirus replicon. These experiments were performed in BHK-21 cells, as described for FMDV.

Co-transfection of both the WT and GNN mCherry replicons showed an increase in the expression of the GFP from the 3BY3F GFP replicon when compared to the 3BY3F replicon being co-transfected with yeast tRNA (Fig 3.19). This was surprising as these findings suggest that poliovirus has an efficient mechanism of providing 3B in trans to other replicons despite encoding only a single copy. It could be that poliovirus and FMDV have evolved different mechanisms to permit the sharing of 3B molecules, in FMDV this would involve the acquisition of multiple copies of 3B, whilst in poliovirus this must be by some other unknown mechanism.

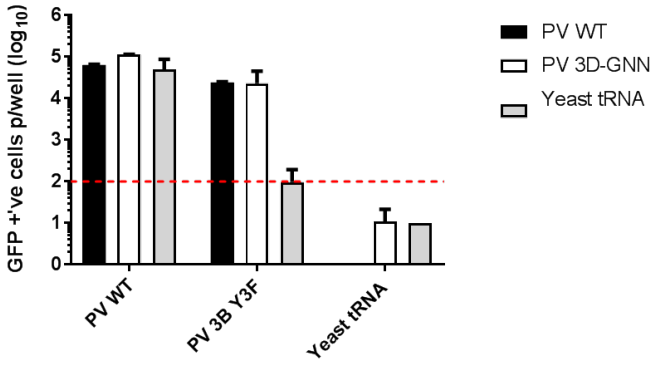

Figure 3.19: Poliovirus replicons can share their single 3B in trans. Complementation experiments of poliovirus replicons was performed by co-transfection of GFP containing WT and 3BY3F replicons, with mCherry WT or 3D-GNN helper replicons. Co-transfections were performed in BHK-21 cells and fluorescent protein production monitored over 24 hours. GFP expression shown here at 10 hours post-transfection. Red dotted line indicates the threshold for complementation, error bars represent SEM of a single experiment with three technical repeats, n=1.

3.2.8 Investigating processing by using a $3C^{pro}$ cleavage mutant.

All the experiments mentioned above were described using mutations which disrupted the cleavage sites between proteins. To independently confirm these findings, the $3C^{pro}$ protein was catalytically inactivated, inhibiting its ability to cleave between the P3 proteins. Picornaviral $3C^{pro}$ proteins function as cysteine proteases with an active site invariably containing Cys-His-Asp/Glu catalytic triad. Here, the His residue primes the Cys for nucleophilic attack on the scissile bond between proteins in the polyprotein chain[18]. Mutagenesis of the $3C^{pro}$ active site altered a residue in the catalytic triad from a cystine to an alanine (C163A), previously reported to abolish the catalytic activity of $3C^{pro}$[19].

The mutation was introduced into the pRep-ptGFP plasmid and transcribed to make replicon RNA containing the C163A mutation, termed 3C C163A. Replicon RNA was transfected into BHK-21 cells, alongside WT and 3D-GNN controls to complement the earlier TnT experiment (Fig 3.9) and ensure that the mutation functioned appropriately and no replication could take place (Fig 3.20).

The 3C C163A replicon was then used in a co-transfection experiment as described previously. The aim was to see if the C163A mutant, which could not process itself, could provide any help and recover the 3D-GNN or 3B123Y3F mCherry replicons in the form of functional precursors.

When co-transfecting the $3C^{pro}$ mutant with 3D-GNN and 3B123 Y3F mCherry replicons, a clear enhancement of mCherry expression from both mCherry replicons was observed. This suggests that some part of the C163A translated polyprotein was being used to provide 3D to the 3D-GNN and 3B to the 3B123 Y3F (Fig 3.21A).

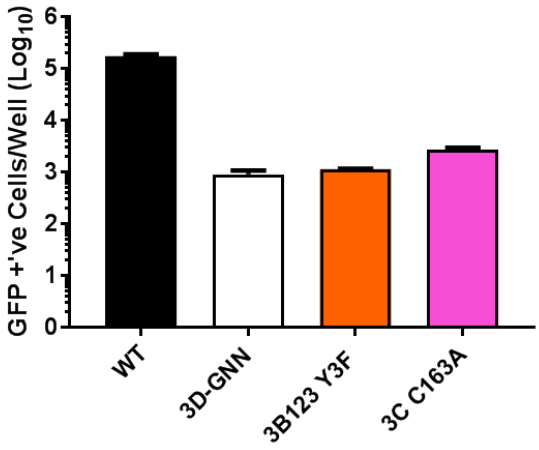

Figure 3.20: Inactivation of the FMDV 3C protein inhibits replication. Mutation of the 3C catalytic triad cytosine to an alanine (3C C163A) was introduced into a GFP containing FMDV replicon. Transfection of this replicon into BHK-21 cells was performed along side WT, 3D-GNN and 3B123Y3F positive and negative controls. GFP expression was monitored hourly and shown here at 8 hours post-transfection. Error bars represent SEM, n=2.

Interestingly, when looking at the reverse side of the experiment, no recovery could be seen of the 3C C163A replicon. When co-transfected alongside, WT, 3D-GNN and 3B123Y3F mCherry replicons no increase in GFP fluorescence could be seen. This is despite being able to see the C163A help recover mCherry replicons (Fig 3.21B). This data suggests an essential cis role of the 3C protease, unrelated to the processing of either 3B or 3D proteins.

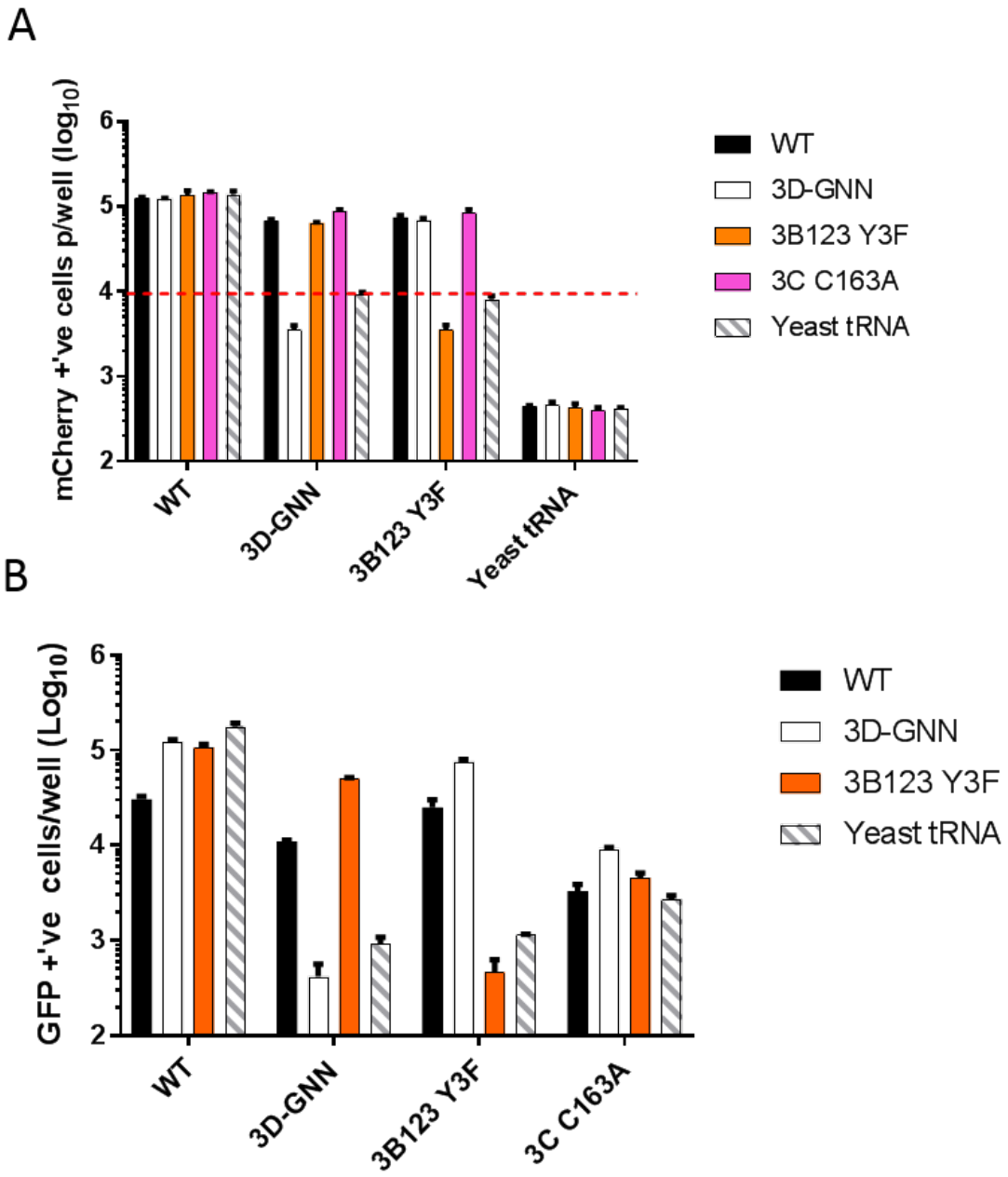

Figure 3.21: Complementation of a $3C^{pro}$ inactivated replicon reveals independent cis and trans activities of $3C^{pro}$.
Co-transfection of WT, 3D-GNN, 3B123Y3F and 3C C163A GFP replicons with WT, 3D-GNN and 3B123Y3F mCherry replicons was performed in BHK-21 cells. Observation of fluorescent protein production was monitored and shown here at 8 hours post—transfection, (A) mCherry and (B) GFP. Error bars represent SEM of three technical repeats of this n=1 experiment.

3.2.9 Exploiting the 3B fusion proteins as tools for FMDV research

With the discovery that a 3B molecule was tolerant to s C-terminal fusions, it was proposed that this observation could be used to design new tools to help the future study of the 3B proteins.

With the help of two undergraduate students under my supervision, Marissa Arfan and Rhys Kingston, we investigated whether the 3B fusion proteins could be used as carriers for epitope tags. The epitope tag could be used to detect 3B via techniques such as immunofluorescent microscopy, immunoprecipitation and electron microscopy.

Using the 3B1/2 3B3 Y3F construct, HA and V5 epitope tags, derived from hemagglutinin of human influenza virus and the V protein of paramyxovirus respectively, were introduced into the middle of 3B2. This site of insertion was chosen based on our previously published work investigating sites tolerant to insertion using transposon mediated mutagenesis (Fig 3.22)[196]. Making constructs with two different epitope tags allows for future work where replicons containing different tags can be co-transfected and their 3B proteins tracked independently via immunoprecipitation or immunofluorescence. As well as these small epitope tags, a much larger metallothionin tag, was also inserted. The metallothionin (MT) tag represents a fragment of the metallothionin protein which retains the metal binding domain of the larger protein. The hope is to use this tag alongside electron microscopy to identify sites of replication within cells, as has been achieved with the Rubella virus P150 protein[167]. This can be achieved by introducing gold salts to replicon transfected cells. The MT tag can then bind the electron dense gold particles and be visualised using electron microscopy. This could be incorporated alongside a tool such as correlative light microscopy, which could help identify which cells are transfected by observing the GFP produced by the replicon.

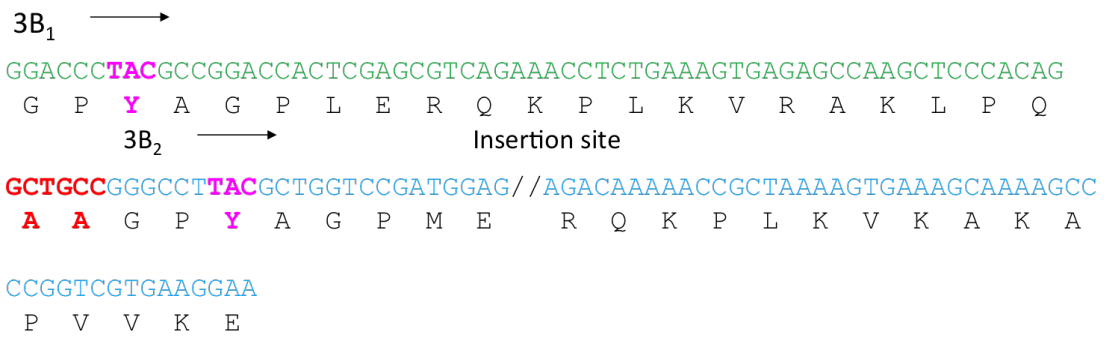

Figure 3.22: Insertion site for additional sequences within 3B2.
Nucleotide and amino acid sequence of 3B1 (green) and 3B2 (blue). Insertion site is indicated by "//" allowing in-frame insertion of new sequences. Essential tyrosine residues shown in hot pink and cleavage boundary mutation shown in red.

To permit antibody free visualisation of 3B containing proteins, split GFP and mCherry tags where also inserted into the same region. GFP and mCherry fluorescent proteins are comprised of 11 beta-strands which fold to make a barrel structure. The mechanism behind split fluorescent proteins is to divide the protein into two regions, a small tag and the rest of the protein. These split fluorescent tags contained the 11th beta-strand whilst the remaining fragment contains strands 1-10 [55]. When in close proximity, these fragments can self-complement and generate a complete fluorescent unit. Although never tested in this system, the small fluorescent tags where inserted into the fusion 3Bs with the thought that they could be a useful tool for the future study of FMDV (Fig 3.23).

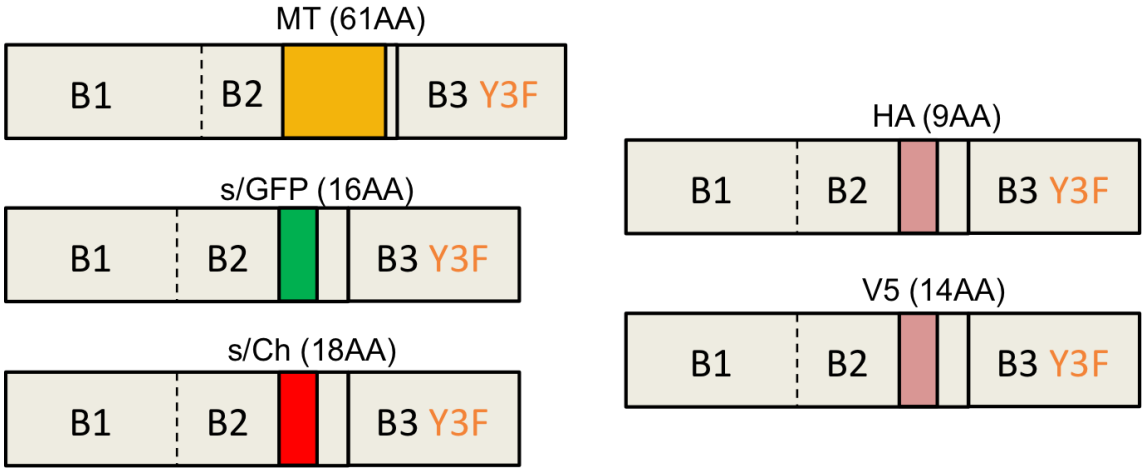

Figure 3.23: Cartoon schematic of insertions to 3B1/2 fusion proteins. Constructs with 3B1 and 3B2 fused together with 3B3 inactivated by tyrosine mutation. Coloured labelled boxes within 3B2 indicated additional inserted sequences along with their size in amino acids (AA).

All epitopes and split fluorescent tags were well tolerated when replicon RNA was transfected into cells, with no decrease in replication from wildtype levels observed (Fig 3.24). The V5 tagged replicon was then investigated as a target epitope for use in immunofluorescence observation of replicon transfected BHK-21 cells. Here, V5 signal (red), showed the diffuse localisation of 3B1/2-HA proteins within cells shown to be transfected with FMDV replicon (as seen by GFP expression) (Fig 3.25). The distribution of 3B1/2-HA within the transfected cell (diffuse cytoplasmic localisation), was reminiscent of experimentations observing the distribution of the $3D^{pol}$ within BHK-21 cells[194].

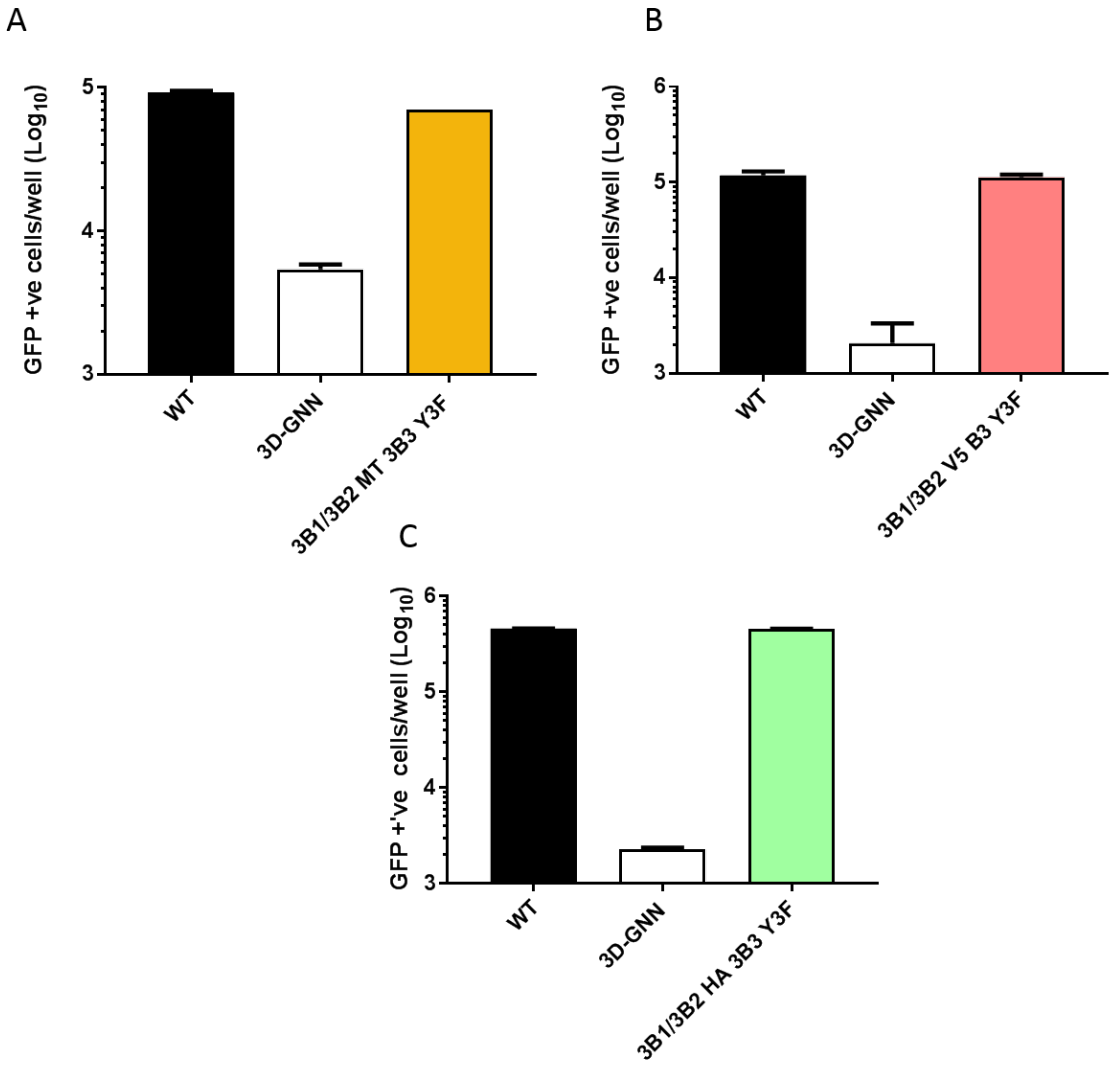

Figure 3.24: Insertion of both large and small sequences into 3B2 is tolerated in replication.

3B1/2 fusion proteins with either an MT tag (A), V5 tag (B) or HA tag (C) were introduced to GFP containing replicons and transfected into BHK-21 cells alongside WT and 3D-GNN controls. Replication shown here at 8 hours post transfection. This work was performed by undergraduate students, Rhys Kingston and Marissa Arfan, under my supervision Error bars indicate SEM, n=2.

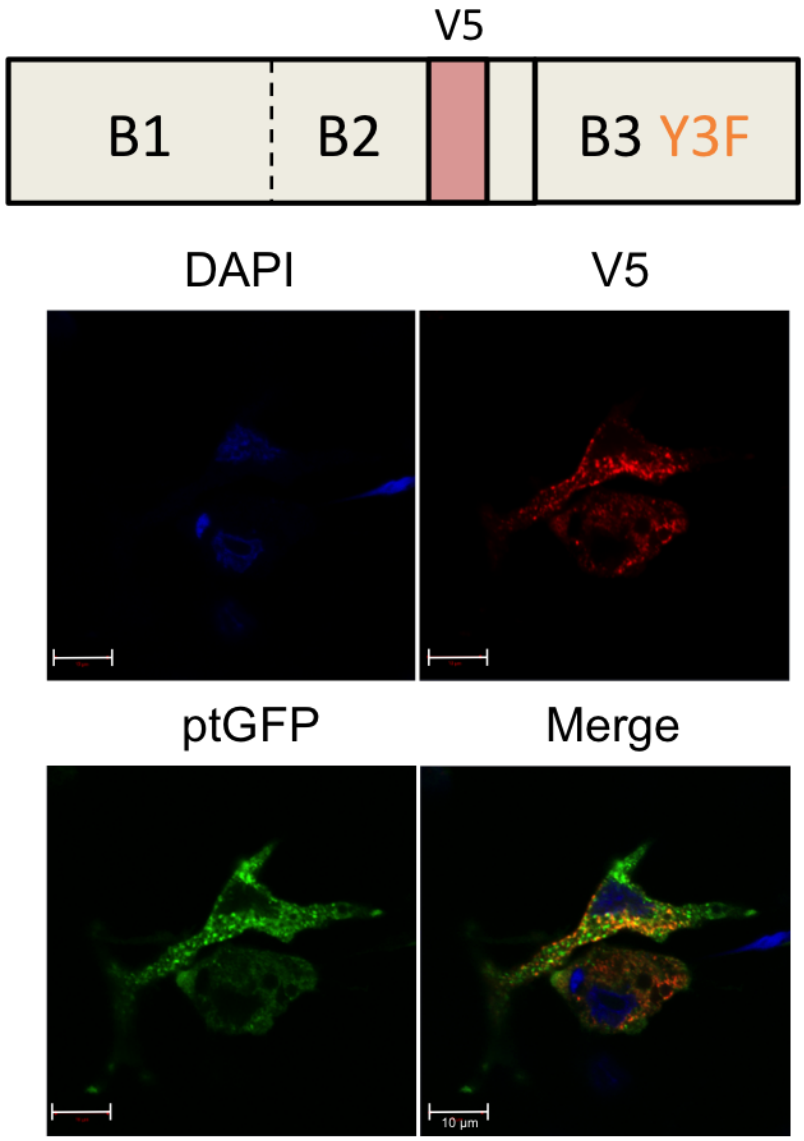

Figure 3.25: V5 epitope tag can be effectively used for visualisation of 3B1/2 fusion proteins.
Replicons containing a V5 tag inserted into the 3B1/2 fusion protein were transfected into BHK-21 cells. At 4 hours post—transfection BHK-21 cells were fixed and stained using a mouse anti-V5 antibody followed by an Alexa 568 anti-mouse fluorescent secondary antibody. Cells were stained using DAPI before visualisation under an LSM880 fluorescent microscope. This work was carried about by an undergraduate student, Marissa Arfan, under my supervision.

The HA tag was used to test immunoprecipitation protocols. Cell lysates were harvested from WT or B1/B2-HA replicon transfected BHK-21 cells, or untransfected BHK-21 cells as a control. Upon completion of immunoprecipitation, flow-through and elutes were electrophoresed on a 15 % acrylamide SDS-PAGE gel and transferred on PVDF membrane. Membranes were probed with a variety of antibodies raised against different FMDV P3 proteins. Western blotting showed anti-HA reactive bands in the flow through of both WT and B1/2-HA transfected cells, with no signal in the untransfected control. However, reactive bands were only detected in the eluate of the B1/2-HA transfected cells. The results show that this approach could be used to specifically pull out replicon-HA tagged proteins (Fig 3.26).

The bands in the flow through of immunoprecipitations are thought to be 3B containing precursors, whilst those seen in the eluate showed to be 3AB12-HA containing bands by sequential probing with anti-3B, anti-3A and anti-HA antibodies. It was interesting that the majority band was 3AB12 as it is thought that uridylylation of 3B occurs in the context of a $3C^{pro}$ containing precursor, and precursors containing 3A have been shown not to be capable of uridylylation in experiments involving poliovirus 3B.

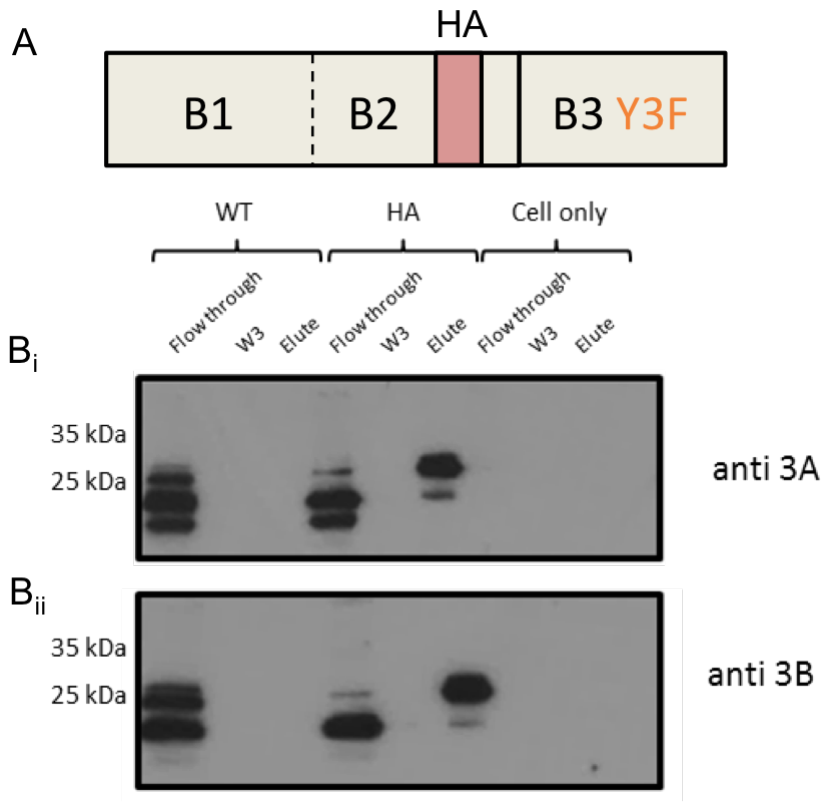

Figure 3.26: HA tags inserted into 3B1/2 can be used effectively in immunoprecipitation experiments.
(A) WT replicon and a replicon containing a HA tag within the 3B1/2 fusion protein were transfected into BHK-21 cells, alongside a non-transfected BHK-21 control ('Cell only'). 8 Hours post-transfection protein was harvested from both transfected cells and untransfected cells as a control. Protein lysates were harvested using RIPA buffer. Immunoprecipitations were performed using Pierce IP-direct kit (Invitrogen) with an anti-HA antibody. Flow through, final wash and eluate were separated on 15 % SDS-PAGE gels before visualisation of proteins by western blot. Antibodies against 3A (Bi) and 3B (Bii) were used to confirm protein identity as a 3AB containing precursor.

Initial optimisation was undertaken to prepare for the use of the MT tag in electron microscopy experimentation. BHK-21 cells, transfected with the MT tag containing replicon, were fixed with glutaraldehyde at 0, 6 or 8 hours post-transfection, due to the cytopathic nature of FMDV infection this was an important preliminary experiment to establish the best time post-transfection to visualise the cells. Visualisation too early would mean replication may not be occurring and too late would mean all cells are in the late stages of apoptosis. Fixed cells were stained with 1 % osmium

tetroxide, dehydrated in an ascending alcohol series and embedding in araldite resin. 50 micron sections were taken from the resin using a glass knife and ultra-microtome. Sections were isolated and placed onto EM grids before staining with uranyl acetate and visualisation on an electron microscope (Fig 3.27). At 0 hours post-transfection BHK-21 cells could be seen resembling a healthy cell with a clear defined nucleus and nucleolus as well as organised membrane arrangements in the cytoplasm (Fig 3.27A). At four hours post-transfection membrane rearrangements can start to be visualised with a collection of electron dense particles polarising to one end of the cell, as well as massive rearrangements of double membranes resembling the endoplasmic reticulum. At this point it is also possible to see spherical vesicles forming within the cytoplasm (black triangles) (Fig 3.27B). By 8 hours post-transfection the cell has undergone what appears to be complete autophagy with disruption of the nuclear membrane clearly visible as well as an electron dense cytoplasm (Fig 3.27C)

3.3 Discussion

3.3.1 3B3 C-terminus is important for correct release of $3D^{pol}$

Here we confirm that FMDV is permissive of replication when containing only 1 functional copy of 3B. However, the presence of 3B3 is essential for proper processing of the P3 region and release of mature $3D^{pol}$, specifically the presence of the 3B3 C-terminus. Interestingly, replacement of the 3B3 C-terminus with that of 3B2 appeared to increase cleavage at the 3B3/C boundary, while mutating the natural 3B3/C boundary to inhibit cleavage still permitted correct processing and release of mature $3D^{pol}$. These data suggests an alternative route of processing than previously suggested. P3 processing is implied from the literature by studying the processing of PV P3[110]. This is described as the major and minor routes with the major

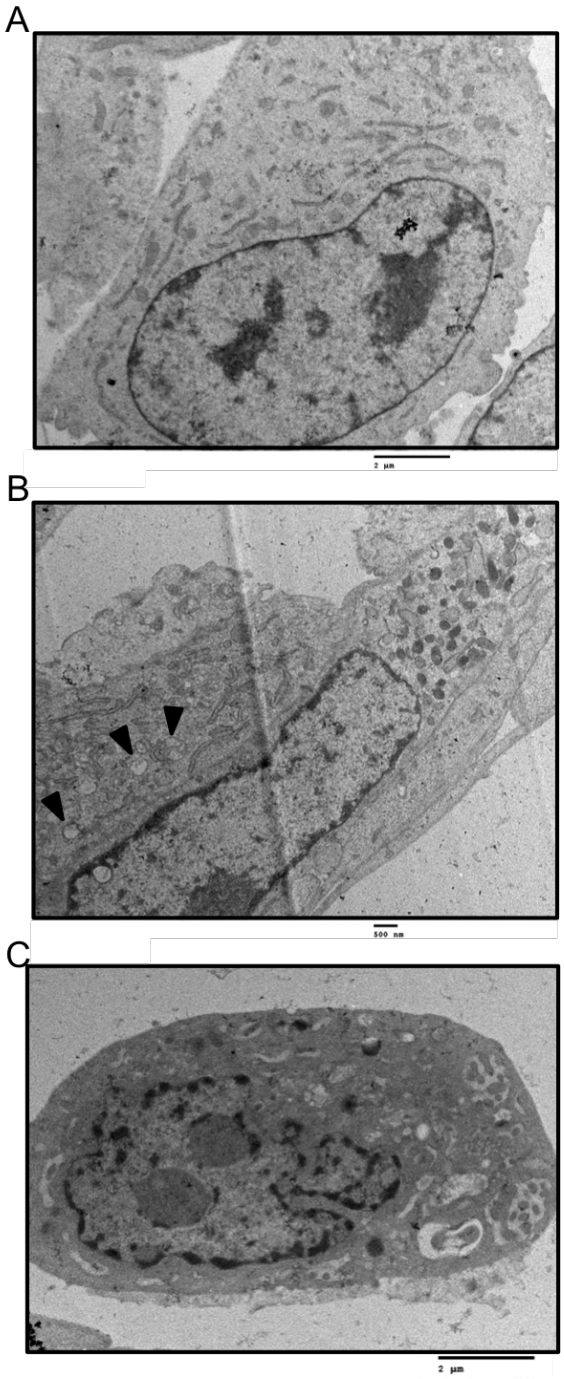

Figure 3.27: Electron Micrographs of FMDV transfected BHK-21 cells. BHK-21 cells transfected with FMDV replicons were harvested and fixed at 0 (A), 6 (B) or 8 (C) hours post transfection. Fixed cells were embedded in resin and sectioned using a glass knife and ultra-microtome. Sections were stained using uranyl acetate and visualised on a Jeol electron microscope. Triangles indicate membrane bound vesicles forming in the cytoplasm.

route creating 3AB123 and 3CD. Due to the rapid cleavage observed between 3B3*2 and $3C^{pro}$, and how slowing or stopping cleavage between 3B3 and $3C^{pro}$ had no major effect on processing, it could be that this major pathway is slightly different in FMDV, instead creating 3AB12 and 3B3CD. Due to the difficulties in distinguishing between individual 3B additions, combined with the similarity between the 3B1 and 3B2 C-termini, it is also possible that 3AB1 3B23CD is also produced.

3.3.2 New tools for the study of FMDV

Cleavage mutations between proteins in P3 have been further dissected in this chapter to investigate the processing and requirements of this region to produce functional proteins. By mutating cleavage boundaries and creation of fusion proteins it was observed that 3B can function as a primer for replication if it has an available N-terminus. C-terminal fusions indicated a tolerance of large additions including the entire $3C^{pro}$ protein.

These 3B fusions have been developed into tools to help further study the processing of the P3 region and the paths taken by 3B in its role as a primer for replication. These tools will allow for easy biochemical manipulation of many aspects of the lifecycle, including the extraction of newly synthesised RNA molecules through the incorporation of epitope tags into the 3B. As the epitope tagged 3B is translated and used as a primer, it means all newly generated RNA strands will have, for example, a VpG-HA at the $5'$ end of the genome. This allows the selection of replicated RNAs from other RNAs in the cell, including the transfected replicon RNA. Quantification and observation of newly generated strands has been a difficult aspect of replicon work, as the large amounts of RNA used in initial transfections causes a large amount of background signal. Incorporation of epitope tags also allows the labelling of 3B molecules by immunofluorescence, and specific labelling of individual 3Bs allowing techniques such as proximal ligation assays to be performed.

It would be interesting to use these epitope tags in combination with tritiated uridine (^{3}H), to investigate the uridylylation of P3 precursors in the cellular environment, an experiment that has so far proved elusive. It is generally thought that 3B proteins may not be uridylylated as a mature protein, but instead provided in the format of a precursor protein, experiments using recombinantly expressed protein suggest this may be a 3BC precursor, with no uridylylation observed with 3AB precursors used in poliovirus experiments. Of course, for FMDV this may be different and potentially 3AB, 3BC or both are used. The large amounts of 3AB seen by western blot when pulling down with the 3B1/B2-HA mutant suggests that 3AB must have a role in the replication of FMDV. However, it could also be argued that the products observed in the highest quantity are not the ones being used, and it could be that 3BC precursors are created but rapidly used.

Tritiated uridine could be used to label 3B during uridylylation and subsequent products observed via SDS-PAGE. Identification of proteins encompassing the radioactive label could then be confirmed using western blot to identify uridylylatabale precursor proteins.

The insertion of the MT tag into 3B leads to an exciting development with the potential to visualise where translation and genome replication are taking place at the minute scale using electron microscopy and gold labelling. Initial experimentation has revealed the optimum time for visualisation of BHK-21 cells transfected with FMDV replicons, at around 6 hours post transfection. Further optimisation of this experiment must be undertaken with the time, amount and route of gold introduction to cells elucidated. A difficult aspect of this experiment is deciphering which cells are FMDV transfected. This is easy at later time points when cells are undergoing apoptosis, but by the time this occurs all significant data will have been lost. Combining the gold labelling with correlative light approaches using the fluorescent reporter created by the replicon could prove a powerful tool for identifying transfected cells. If used together, these techniques will allow the detailed study of FMDV replication complexes at the RNA and protein level, something which to

date has not been reported.

The use of split GFP and mCherry tags, although never functionally tested, to 3B also has great potential in the general study of FMDV. Collaborators in the Pirbright institute have shown great interest in these constructs, which, when converted into infectious virus by addition of the capsid proteins, could be used in animal's recombinantly expressing the 1-10 segment of the split fluorescent molecule. Allowing easy tracking of virus throughout the cycle of infection.

3.3.3 Why does FMDV has multiple copies of 3B?

Whilst initially attempting to decipher requirements for mature 3B function and polyprotein processing, performing combinations of P3 cleavage mutations and inactivating 3B mutations revealed unexpected results. It was discovered that with less than two functional 3Bs no complementation could occur and recovery of a co-transfected replicon containing no functional 3Bs was inhibited. This result potentially helps shed light on why FMDV has need for more copies of 3B. By using a poliovirus replicon as a tool for the study of picornaviruses with naturally one copy of 3B, it was observed that although only one copy of 3B is present, it could still share 3B with another replicon.

As poliovirus contains a single copy of 3B, PV replicons were exploited to determine the ability of picornaviruses with naturally one copy of 3B to share these in trans. Complementation assays revealed that even possessing a single copy, PV replicons were capable of sharing their 3B proteins in trans to a 3B inactivated replicon. Supporting previous work by Wimmer et al, who investigated the complementation of 3AB in dicistronic polioviruses[110].

The consistent observation is that although the requirements are different, both wild-type poliovirus and FMDV can share their 3Bs in trans. It is plausible that the sharing of 3B molecules in trans to other constructs is a technique used by

picornaviruses to aid rapid evolution of a diverse population. For example, sharing a 3B molecule could overcome RdRp error, which may cause inactivation of a virus-encoded 3B. If like poliovirus, this virus only encodes one 3B, the sharing of 3B from another construct would permit maintenance in the population until reversion or compensatory mutations are established. If this is so it is a mystery why FMDV requires a minimum of two to permit sharing of a 3B. The answer to this could be within the other aspects of FMDV polyprotein processing and ability to share other P3 proteins.

3.3.4 The cis function of $3C^{pro}$

In performing the experiments described above, the use of an inactivated $3C^{pro}$ replicon was used to investigate the processing from a different angle. While these experiments did not show the expected results, an insight into the importance of $3C^{pro}$ was revealed. Co-transfections showed that whilst a replicon containing the inactivating 3C C163A mutation could rescue both $3D^{pol}$ and 3B deficient replicons. None of the replicons used could rescue the C163A replicon. This reveals an insight into the cis and trans mediated aspects of the FMDV processing pathway. For example, the release of $3D^{pol}$ from the polyprotein chain can be performed in trans, likewise functional copies of 3B could be released from a trans-provided $3C^{pro}$. However, the lack of recovery of the C163A replicon, suggests that at some point in the replication cycle of this construct there is a requirement for a cis $3C^{pro}$ cleavage activity. Experiments allowing sensitive detection of the processing of the polyprotein could be used to investigate this. For example, TNT experiments where a construct possessing the C163A mutation is translated in the presence of ^{35}S labelled cys/met is then chased with large amounts of cold cys/met. The addition of constructs containing functional $3C^{pro}$ could be added after this chase to prevent labelling and the processing of C163A observed to see if a clear processing step is altered or missed.

Chapter 4

The Role of the S-fragment in viral replication

4.1 Introduction

Positive strand RNA viruses exhibit a range of genome organisations, this is particularly evident in the appearances of their 5′ untranslated regions (UTR), which can vary from very small as in Hepatitis E virus (26 nucleotides (nts)) to a significant proportion of the genome as shown in FMDV.

The 1300 nts of the FMDV 5′ UTR represent one seventh of the entire genome, comparatively, the 5′ UTRs of other picornaviruses are much shorter. Typical enteroviruses such as poliovirus (PV) possess a 5′ UTR of approximately 740 nts and the cardiovirus, EMCV, a 5′ UTR of around 830 nts. Along with its length, the FMDV 5′ UTR includes several unique RNA structures (Fig 4.1). At the 5′ end, there is an approximate 360 nt region predicted to fold into one large hairpin loop, termed the S-fragment (Fig 4.2). Based on energy favourability (ΔG), it has been predicted that the FMDV S-fragment has a ΔG of approximately -161.64 kcal/mol, suggesting that the predicted structure is stable and energetically favourable. The term S-fragment arose during early investigation into the FMDV genome when digestion of the poly-C tract using RNase H in the presence of oligo(dG) left two segments of RNA, a short 'S-fragment' and a large 'L-fragment'[143]. While the term S-fragment has persisted, over time the L-fragment has been identified as the

remaining portion of the FMDV genome. Sequence variability analysis of the S-fragment has revealed a surprisingly low amount of conservation with only 12 % invariant nucleotides found across the 103 isolates examined[31].

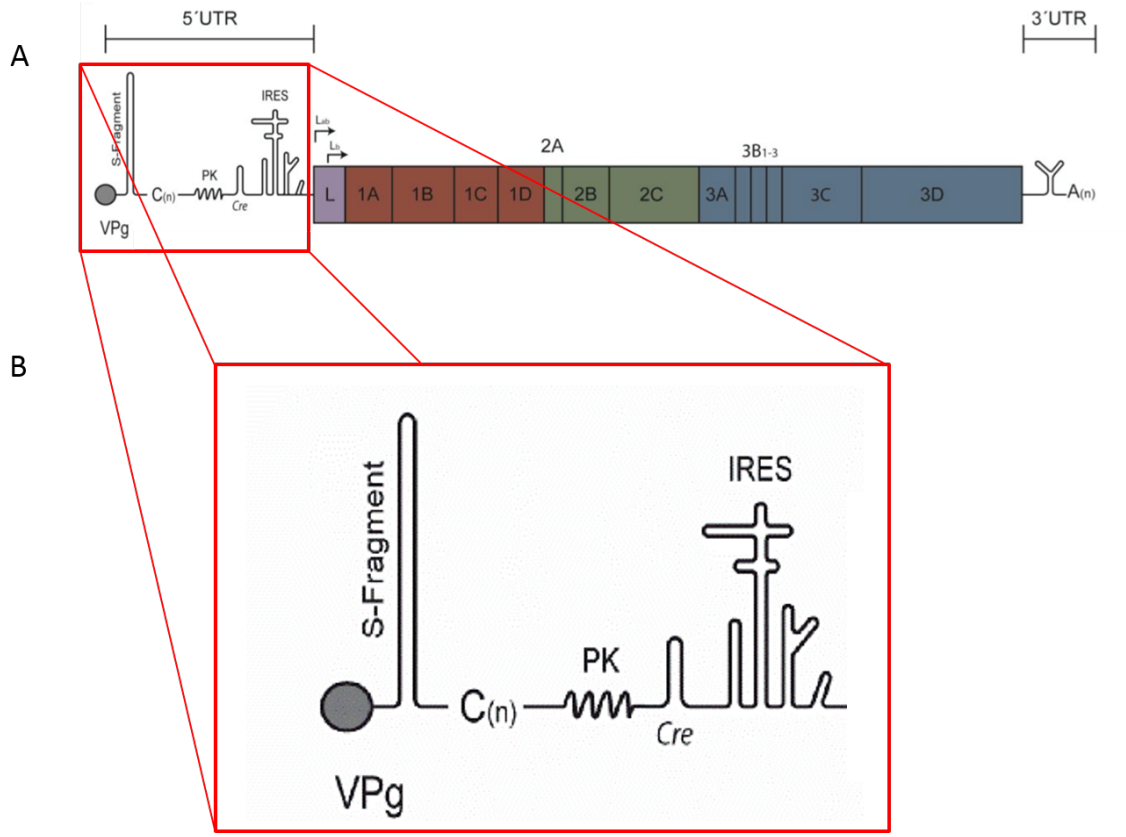

Figure 4.1: Cartoon representation of the FMDV Genome
A.Cartoon schematic of the FMDV genome, including 5′ UTR, structural and non structural coding regions and 3′ UTR. B. An enlargment of the 5′ UTR is shown, displaying the S-fragment, poly-C-tract (C(n)), pseudoknots (PK), cis-active replicative element (cre) and internal ribosome entry site (IRES).

The poly-C tract, which separates the S-fragment from the rest of the genome, is another unusual feature of the 5′ UTR. It is of variable length, typically around 150 to 200 nts in field strains, although it has been noted to be shorter in laboratory strains[126,173]. Following the poly-C-tract lies a repeat of two to four predicted pseudoknots (to be discussed in chapter 5), a cis-active replicative element 'cre', essential for the uridylylation of the 3B protein, discussed in chapter 3, and the internal ribosome entry site or 'IRES', critical for the translation of viral proteins[192].

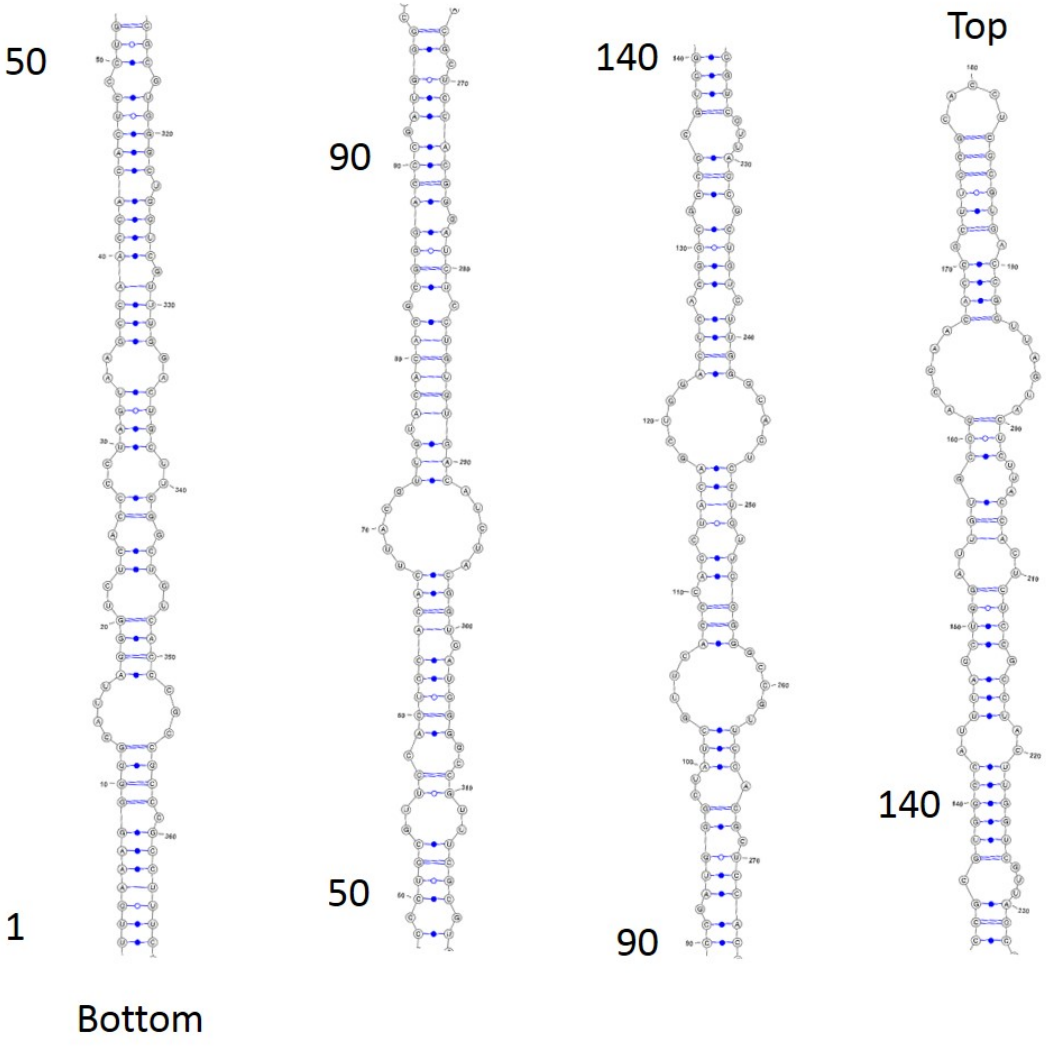

Figure 4.2: mFold prediction of FMDV O1K S-fragment structure.
RNA sequence of FMDV O1K S-fragment input into the Vienna mFold server. Folding based on minimal free energy requirements. S-fragment has been broken into three segments for visualisation with the top and bottom marked as well as nucleotide number for reference.

In a similar position to the FMDV S-fragment, PV possesses a 5′ cloverleaf (CL) structure of approximately 80 nts[9]. As PV is arguably the most studied virus within the Picornaviridae, the PV cloverleaf has been well characterised and is understood to have several different functions, including being required for negative strand synthesis[9].

Understanding the functions of the PV CL structure can help shed light on the potential roles of the FMDV S-fragment. One of the more notable functions of the CL is its role in both host-protein and viral-protein interactions where it plays a key part in the circularisation and replication of the viral genome. Circularisation of the genome is carried out through a complex series of interactions[81]. Upon translation of the poliovirus genome, the 3CD protein, amongst others, is produced. The 3CD of PV is catalytically active and cleaves the host protein poly-C-binding protein 2 (PCBP2), causing its dissociation from the IRES (where it normally plays a role in translation initiation) and allows re-localisation of the cleaved form to the CL, where truncated forms of the PCBP2 protein binds to the right arm of the CL[34,22,157]. This re-localisation of PCBP2 causes a dual effect – reducing translation initiation and clearing the genome of ribosomes as well as allowing an interaction with the cellular poly-A binding protein (PABP). PABP can be found bound to the poly-A tail of the viral genome to interacts with PCBP2 that is now bound to the CL, allowing circularisation of the genome and efficient negative strand synthesis.

Alongside the important roles in genome replication, the PV CL has also been shown to have a direct effect on the overall stability of the RNA genome. When mutant poliovirus RNA lacking the 5′ CL is produced and incubated in HeLa cell-free extracts, alongside full length poliovirus RNA, a rapid increase in RNA degradation is seen. This observation suggests it plays a similar role to the 5′ cap of cellular mRNAs, protecting PV RNA from rapid nuclease degradation. Adding a 5′ cap to poliovirus RNA lacking the CL restores the RNA stability, thereby supporting this hypothesis[9].

Other picornaviruses, such as the well-known enteroviruses EV71 and HRV16, also have CLs with predicted functions similar to that of PV. Interestingly, EMCV (cardiovirus), possesses a stem-loop resembling that of FMDV, but with a size comparable to the enteroviruses CL at approximately 80 nts[30]. It would be interesting to compare other viruses in the Aphthoviridae, bovine rhinitis A and B viruses and equine rhinitis A virus, to FMDV, however due to the complex UTRs of these viruses full sequences of these are yet to be evaluated.

In comparison to the PV CL, the role of FMDV S-fragment is poorly understood. Indeed, the structure has only been predicted by m-fold computational analysis. Whilst mfold can be an accurate tool for predictions, particularly when it is used as a model for designing mutations to confirm predicted structures, it lacks the ability to consider other factors that may affect RNA folding. This includes the presence of cellular or viral proteins and the composition of folding conditions, which can be affected by the presence of metal ions and salts.

It is remarkable that while there is a notable amount of sequence variation amongst FMDV sub-types and strains, the size of the S-fragment is usually retained at around 360 nts. However, there are a few exceptions to this rule; some SAT type viruses have small truncations thus reducing the S-fragment length. Reported deletions involve the removal of a segment from the middle of the S-fragment sequence (a reduction of around 70 nts), which correlates to the tip of the stem loop[199]. The presence of small truncating mutations found in field isolate suggests that the full length S-fragment is not essential for virus replication, although these mutations have not been functionally tested. It is also unknown whether there is a specific region of the S-fragment that plays a more significant role than others.

This chapter aims to address aspects of S-fragment biology through investigation into its structure using SHAPE analysis and to better define potential functions in the viral lifecycle. Using both replicon and viral systems as well as multiple cell-lines allows for easy and targeted dissection of different stages of the viral lifecycle and

determine any potential host-related differences.

4.2 Results

4.2.1 Throwing SHAPEs: Refining the predicted structure of the S-fragment using NMIA chemical probing

Predictions of the S-fragment structure have previously been investigated, often using the computational mFold approach. Here, biochemical methods were used to confirm or refute the predicted large hairpin loop structure[144].

We used a fluorescent variant of the selective 2' hydroxylation analysed by primer extension (SHAPE) protocol, with NMIA compound to label non-base paired nucleotides. Pre-folded RNA was incubated with NMIA which binds to non-base paired nucleotides to form a $2'$-O-adduct. DMSO was used in the place of NMIA as a negative control. Once labelled, fluorescent primers were used to perform reverse transcription reactions, when the reverse transcriptase reached a NMIA adduct, extension terminated. The pool of DNA fragments created from the reverse transcription were purified and analysed by capillary electrophoresis (full protocol described in Materials and Methods 2.7). The full workflow of SHAPE analysis can be seen in figure 4.3. Sequencing data, provided by Dundee University DNA sequencing and services, were analysed using QuSHAPE. Outputs from QuSHAPE was entered into the Vienna RNAFold server where folding could be predicted using the SHAPE data as pseudo energy constraints, providing a more reliable prediction of the S-fragment structure. The individual nucleotide reactivity to the NMIA compound where any negative reactivity, i.e the terminations seen in the DMSO control was higher than that in the NMIA, were set to 0 and any 'extremely reactive' nucleotides limited to a maximum of 3 (Fig 4.4A).

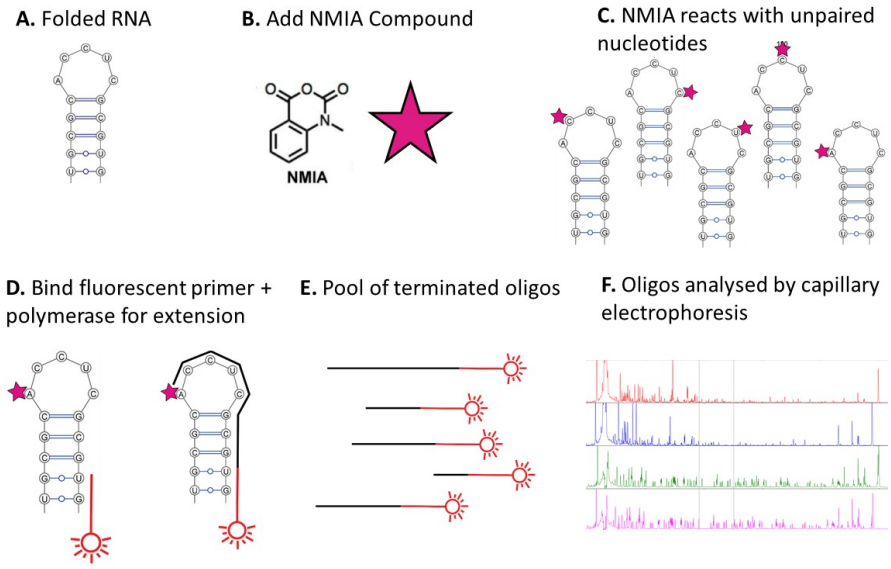

Figure 4.3: Experimental flow of SHAPE reactions.
In vitro transcribed RNA is folded in folding buffer (A). NMIA (purple star) compound is added to the RNA (B). NMIA reacts with unpaired nucleotides (C). Fluorescent primer is bound to labelled RNA and reverse transcriptase susperscript III is added for primer extension (D). Primer extension is terminated by NMIA adduct, generating a pool of different length oligos (E). Oligos are purified and analysed by capillary electrophoresis alongside negative controls (DMSO) and sequencing ladders (F).

Unfortunately, due to the position of the poly-C-tract, immediately $3'$ to the S-fragment, there were limited options in the binding location of the reverse transcription primers. Therefore, some information at the $3'$ end of the S-fragment is missing due to initial polymerase disassociation causing the datum in this area to be unreliable.

It was clear after analysis that the SHAPE data generally agreed with the overall mFold predications of a single large stem-loop. However, there are subtle variations in the locations and sizes of bulges within the stem-loop (Fig 4.4B). It could also be argued that these bulges may be plastic in their appearance due to 'breathing' of nucleic acid interactions in the residues at the proximal ends of bulges. It is also important to note that this was performed in a cell-free environment so protein-RNA interactions could drastically change the structure of the S-fragment [84].

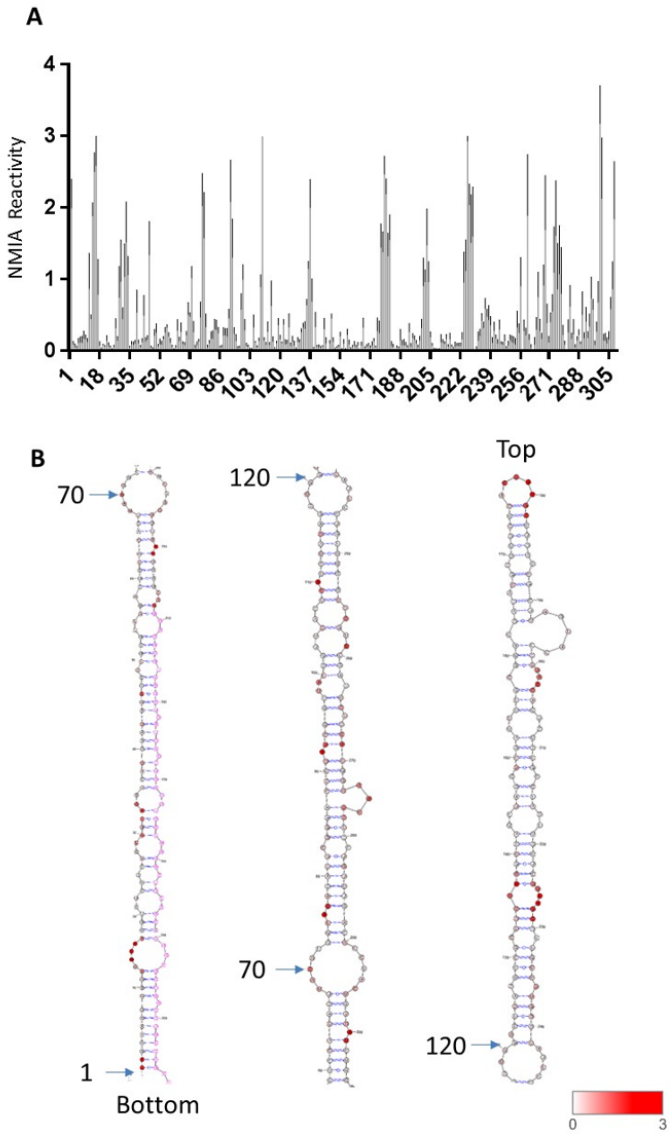

Figure 4.4: NMIA reactivity of the wild type S-fragment.
(A) Read out of individual nucleotide NMIA reactivity as analysed by SHAPE reactions and capillary electrophoresis. High reactivity indicates high probability of un-base paired nucleotides. Data was analysed using QuSHAPE where negative reactivates were set to 0 and 'super reactive' nucleotides set to a maximum of 3. (B) Predicted structure of a wild-type S-fragment when SHAPE reactivity was used to help predict secondary structure. Structure was predicted using the Vienna RNA probing package and visualised using VARNA. NMIA reactivity's are overlaid and represented on a colour scale from low (white) to high (red). Nucleotides with no data are represented as grey. (n=6, error bars represent SEM).

4.2.2 Birds of a feather: Addition of two EagI sites to produce an easily modifiable S-fragment 'Aerie' construct

Chemical probing of the S-fragment using the SHAPE analysis above (4.2.1) provided greater confidence in the predicted RNA structure. Using this information, we designed truncating mutations to the stem-loop to determine the absolute requirements for replication. Due to the position of the poly-C-tract, generating mutations in the S-fragment was troublesome. To solve this, a cassette system was created by introducing unique EagI restriction sites at both $5'$ and $3'$ ends to allow easy digestion of the replicon backbone and replacement with desired mutations. EagI restriction sites were chosen as they required minimal modification to the native sequence of the S-fragment, and the EagI sequence allowed maintenance of base pairing in the chosen region (Fig 4.5A).

The EagI sites were introduced into the DNA copy of the FMDV O1K replicon (pRep-ptGFP), initially in isolation ($5'$ EagI, and $3'$ EagI) and subsequently in combination (Aerie). To ensure there was no negative impact on replication, the pRep-ptGFP plasmid containing EagI insertions was transcribed to make the mutant replicon RNA and transfected into BHK-21 cells. Replication was monitored using the Incucyte Zoom to measure GFP reporter expression as previously described. Overall, insertions of both restriction sites were tolerated. However, introduction of the $5'$ EagI site alone caused a larger decrease in GFP expression (50 %) than that of the $3'$ site (25 % reduction). Combining both $5'$ and $3'$ EagI sites, and therefore restoring the base pairing of the stem of the S-fragment (Aerie), increased GFP expression to approximately 80 % of that of the WT replicon. Although the replication assay did show a reduction in the Aerie replicon replication, it was not significant. Furthermore, GFP production was well above that observed with the 3D polymerase active site knock-out '3D-GNN' control (Fig 4.5B). Therefore, it was decided that the fitness cost was small enough to continue using the Aerie replicon as a template for further S-fragment mutations.

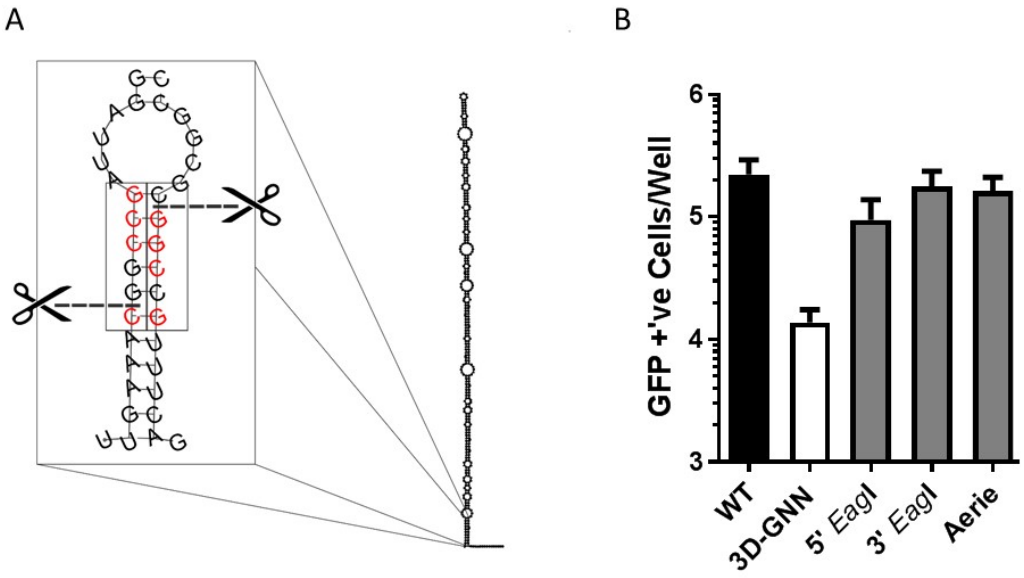

Figure 4.5: Modification of the FMDV 01K genome to allow easy mutagenesis.
(A) Cartoon schematic of site directed mutagenesis to the bottom of the S-fragment. Altered nucleotides are indicated in red to create two EagI restriction sites, while maintaining secondary structure. (B) Replicon replication of S-fragment mutants, shown as the number of GFP positive cells per well at 8 hours post transfection in BHK-21 cells. WT and 3D-GNN act as positive and negative controls for replication. $5'$ and $3'$ EagI constructs represent the addition of the $5'$ and $3'$ sites respectively whilst Aerie contains both, n=3, error bars represent SEM.

The Aerie replicon was used initially to permit efficient deletion of the S-fragment from the genome and confirm historic findings that the S-fragment is essential for replication. Digestions of the pRep-ptGFP-Aerie DNA clone with EagI was performed prior to re-ligation, to exclude the S-fragment region. This construct, termed 'ΔS' hereafter, was transcribed into RNA and transfected into BHK-21 cells. Replication of this replicon was monitored using methods described previously. As expected, a substantial decrease in GFP-expression was observed, reducing levels similar to that observed in the 3D-GNN negative control (Fig 4.6), corroborating previous work that the S-fragment is essential for the replication of FMDV.

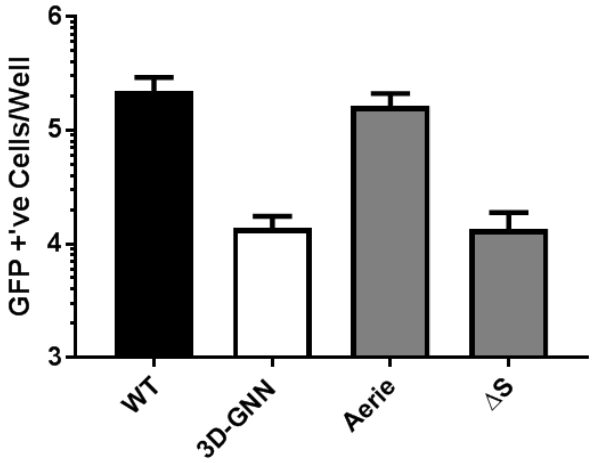

Figure 4.6: Removal of the S-fragment inhibits replication.
Replicon replication of FMDV O1K replicon with S-fragment deleted (ΔS). ΔS replicon transfected alongside, WT, 3D-GNN and Aerie positive and negative controls. Replication expressed as the number of GFP positive cells per well at 8 hours post transfection in BHK-21 cells. n=2, error bars represent SEM.

Interestingly, mFold predictions showed that the reverse complement of the S-fragment sequence creates a highly ordered branched structure, notably different from that of the large stem loop (Fig 4.7A). However, this structure maintains a low ΔG of 152.36 kcal/mol similar to that observed in the WT structure. Thus we were also interested to see whether re-orientation of the S-fragment had any effects on replication. Manipulation of the pRep-ptGFP-Aerie DNA plasmid was performed to remove the S-fragment by digestion with EagI. The DNA fragment relating to the S-fragment could then be isolated by agarose gel electrophoresis, purified, and used in a ligation reaction to be re-ligated into the replicon DNA plasmid. In this case, we looked for constructs where the S-fragment was inserted into the incorrect orientation. Creating a construct with a reverse complement of the S-fragment (S Rvs). As before, this was transcribed and replicon RNA transfected alongside Aerie, 3D-GNN and ΔS controls. Considering the difference in predicted structure, it was not surprising that insertion of the reverse complement S-fragment abrogated replication (Fig 4.7B).

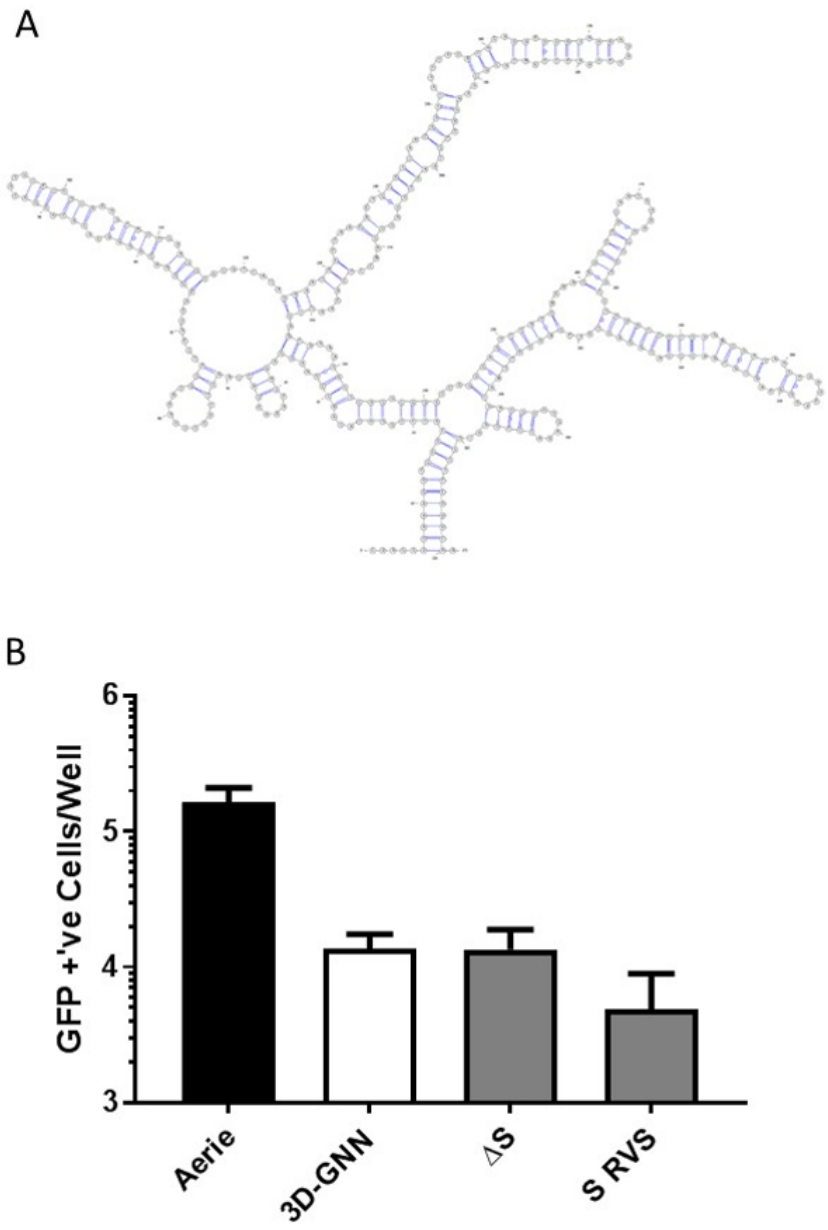

Figure 4.7: Reversing the S-fragment inhibited replicon replication.
(A). RNA sequence of the reverse complement of the FMDV O1K S-fragment was input into the Vienna mFold server. Folding based on minimal free energy requirements. (B) Replicon replication of FMDV O1K replicon with S-fragment reversed (S RVS). S Rvs replicon was transfected alongside Aerie, 3D-GNN and ΔS replicons as positive and negative controls. Replication expressed as the number of GFP positive cells per well at 8 hours post transfection in BHK-21 cells. n=2, error bars represent SEM.

4.2.3 Investigating S-fragment functions using picornaviral 5′UTR elements to create FMDV UTR chimeras.

Once created, the Aerie replicon could be used as a cassette system for inserting mutations into the S-fragment. As mentioned earlier, there are considerable variations in S-fragment nucleotide composition between different FMDV isolates. To investigate the importance of S-fragment nucleotide composition (whilst maintaining a similar structure), the sequence of a SAT2 S-fragment was removed by PCR amplification using primers with terminal EagI sites. This fragment was then digested and ligated into the O1K FMDV replicon plasmid. The SAT2 sequence was taken from an Egyptian isolate EGY/9/2012, alignment to the O1K sequence revealed some areas of homology, as well as nucleotide variation (Fig 4.8A). Once produced, the SAT2 S-fragment containing plasmid, was transcribed and transfected into BHK-21 cells as described previously.

Comparing the SAT2 S-fragment containing replicon to that of the O1K WT showed no significant difference, suggesting that the sequence might not play a key role in the function of the S-fragment (Fig 4.8B). However, it is possible that the areas of homology seen in the sequence alignment may hide key motifs for interactions such as S-fragment – RNA, or S-fragment-protein interactions.

In a previous study investigating the PV CL, chimeras were made replacing the PV CL with the cloverleaves of human rhinoviruses (HRV)[166]. The resulting chimeras were capable of producing functional viruses. Therefore, it is plausible to hypothesize that chimeras generated with 5′ UTR elements from other picornaviruses may maintain function when introduced to the FMDV genome.

To investigate this, further chimeras were made using RNA structures found in similar positions from other picornaviruses, including the PV, EV71 and HRV16 cloverleaf structures as well as the EMCV S-fragment. Predicted structures of these were generated via mFold (shown in Fig 4.9).

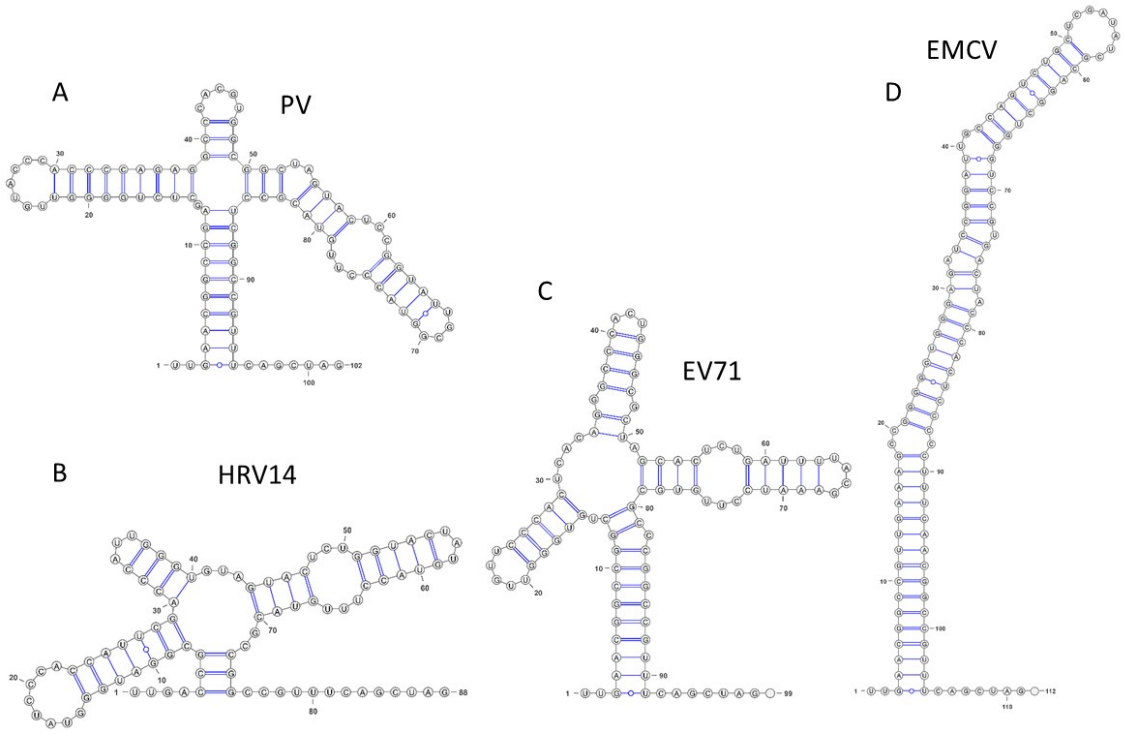

Figure 4.9: mFold predicted structures of picornavirus 5′ RNA elements. Predicted folding of structures found at the 5′ end of different picornaviruses. (A) poliovirus, 95 nucleotides. (B) HRV14, 75 nucleotides. (C) EV71, 91 nucleotides. (D) EMCV, 104 nucleotides. Folding predicted using mFold Vienna server.

Chimeric replicons were generated by the addition of EagI sites at the bottom of the 5′ RNA elements of PV, EV71, HRV16 and EMCV (shown in Fig 4.9) and production of DNA fragments by gene art gene strings (Thermofisher). These were then introduced into the pRep-ptGFP-Aerie plasmid by digestion with EagI followed by ligation into the FMDV replicon plasmid. The plasmid containing replicon mutants was then transcribed and transfected into BHK-21 cells. None of the mutant chimeras showed any replication above that of the 3D-GNN negative control, suggesting the inserted RNA structures could not provide the necessary functions for

A

```
01K       TTGAAAGGGGGCATTAGGGTCTCACCCCTAGTAAGC--CAACGACAGTCCCTGCGTTGCA   58
SAT2      TTGAAAGGGGGCGCTAGGGTCTCGCCCT-AGTTCGCCGCAACGACCGTCCCTGCGATTCA   59
          ************. *********.***  ***:.**  *******.*********:* **

01K       CTCCACACTTACGTTG--TACACACGCGGGACCCGATGGGCTATCGTTCACCCACCTACA   116
SAT2      CTTCGCACCCAGGCCACTACTCTTCGGTGGCCCTGTTGGACTGTCGCCCACCCACCTTTG   119
          ** *.***  * *  .  :. . :**  **.** *:***.**.***  *********: .

01K       GCTGGACTCACGGCGCCGCG---TGGCCATTTAGCTGGATTGTGCGGACGAACACCGCTT   173
SAT2      GCTGGACCTACTACACGGCACCCGGCTCGGATGGCTGGATTGTGCGGACGAACGTCACCC   179
          *******  ** .*.* **.    *  *. :*.********************. *.*

01K       GCGCACCTCGCGTGACCGGTTAGTACTCTTACCACTCTCCGCCTACTTGGTCGTTA-GCG   232
SAT2      GCGTATCTCGCGTGGTTGGCTAGTACTCTTACCACCCTCCGCCTACCTAGTCGTTTCGCT   239
          *** * ********.  ** *************** ********** *.******: **

01K       CTGT-CTTGGGCACTCCTGTTGGGGGCCGTTCGACGCTCCACGGGATCTCCTGTGT----   287
SAT2      CTGTTGACAGGCACTCCCATTGGGGGCTGTTCGGCGCTCCACAGAGGCGTCTACGTTTAG   299
          ****  : .******** .******** *****.********.*.. *  **. **

01K       GACATCTACGGTGATGGGGCCGTTTCGCGTGGGCTGGTCGTTTGGACTGCTT-CGGCTGT   346
SAT2      TATGGCTACGGTGTTGGAGCCGTCTCGCGCGGCCGACTCGCTTGCGCGTGTCTAGCTCGT   359
           * . ********:***.***** ***** ** * . *** *** .*   *  .*   **

01K       CACCCG-GCGCCCGCCTTTCAGCTAGC                                   372
SAT2      CGCCCGAAAGCCTGCCTTTCA------                                   380
          *.**** ..*** ********
```

B

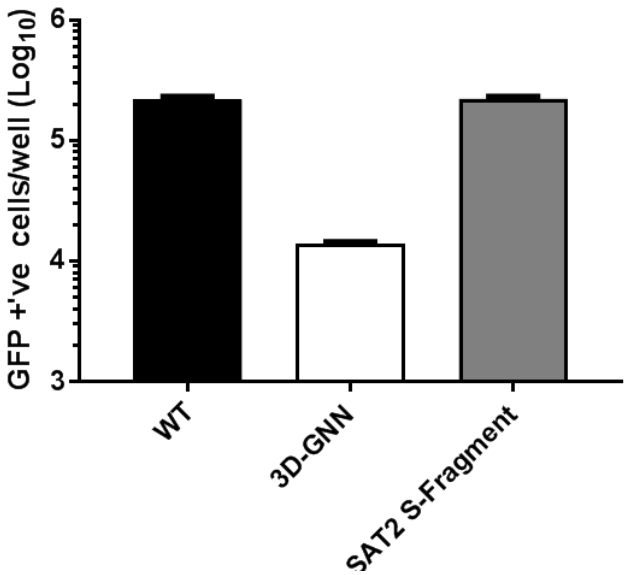

Figure 4.8: Replicative ability of the SAT2/01K chimeric replicon.
(A) Sequence alignment of the SAT2 S-fragment with that of the 01K. Stars indicate conserved nucleotides. Alignment performed using clustal omega. (B) Incucyte analysis of replicon replication, shown 8 hours post transfection in BHK-21 cells. (n=2, error bars represent SEM)

FMDV replication (Fig 4.10). This may be due to the lack of consistency in RNA structure and/or sequence meaning that whilst they could potentially provide a similar function, any essential FMDV or cellular proteins or long range RNA interactions that may happen in this region are not fulfilled. Although the PV, EV71 and HRV structures had the CL structure, the EMCV S-fragment forms a similar structure to that of the FMDV S-fragment. It may be that this structure was simply too small to replace the function of the FMDV S-fragment.

Notably, all of the tested chimeric replicons showed GFP production below that of the 3D-GNN negative control, suggesting a negative impact on the translating behaviour of the FMDV IRES.

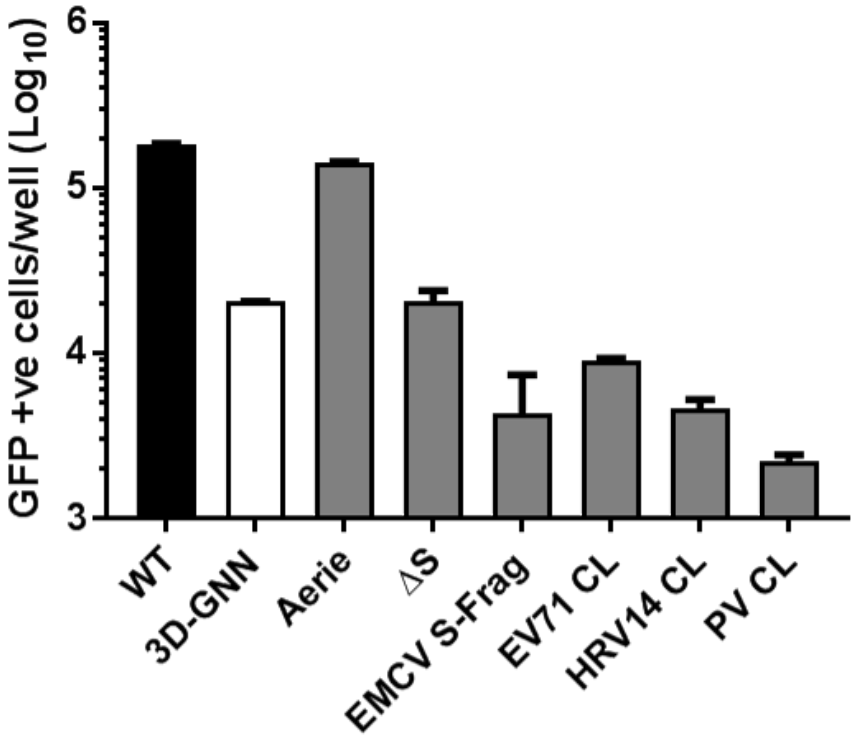

Figure 4.10: Replicon replication of different $5'$ UTR chimeras.
Replicon replication shown at 8 hours post transfection. WT and Aerie replicons act as positive controls with 3D-GNN and ΔS serving as negative controls. Chimeras represent the replacement of the FMDV S-fragment with either the EMCV S-fragment, or the EV71, HRV14 or PV cloverleaves. (n=2, error bars represent SEM).

4.2.4 Does size matter? Truncation of the FMDV replicon S-fragment

As the EMCV S-fragment could not provide functions necessitated by FMDV, we sought to see whether varying the size of the FMDV S-fragment had impacts on replicon replication.

Analysis into the requirements of the S-fragment was continued using the Aerie replicon as a backbone for the introduction of truncating mutations from both the top (T-) and the bottom (B-) of the stem loop. Deletions were designed using the SHAPE predicted S-fragment structure, with the aim to finish truncation at a stable region of the S-fragment, i.e not in the middle of bulges. Removal of the nucleotides from the top of the S-fragment started with deletion of the top 70 nucleotides (replicating the mutation found in a SAT isolate[199]) to generate mutant T-70. Further truncation included removal of 148 (T-148), 246 (T-246) and 295 (T-295) nucleotides from the top of the S-fragment, as well as 48 (B-48), 97 (B-97), 195 (B-195) and 271 (B-271) nucleotides from the bottom (Fig 4.11). Mutant S-fragments were created by cloning in EagI digested Thermofisher gene art gene strings into a similar digested pRep-ptGFP-Aerie replicon plasmid as described for the picornavirus chimeras.

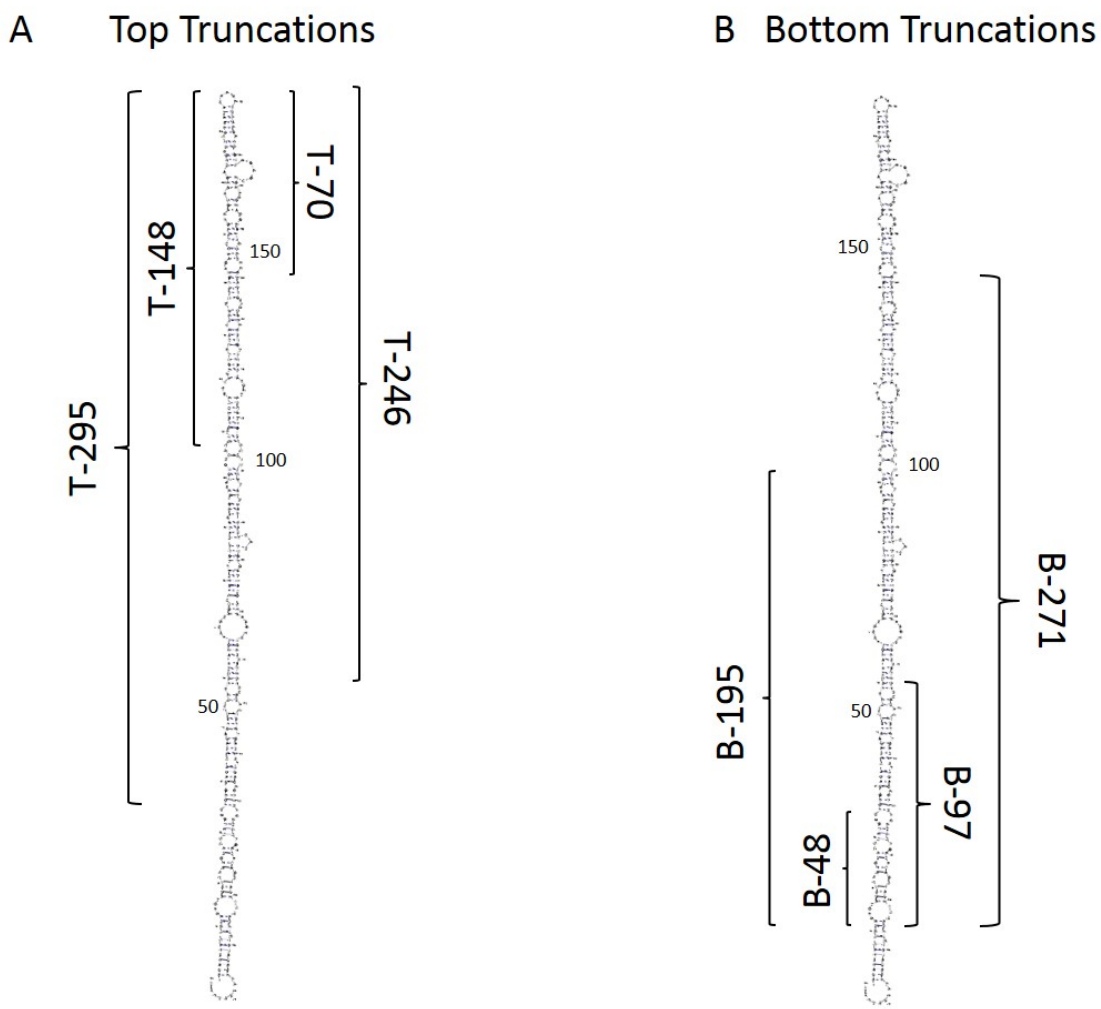

Figure 4.11: Schematic representing S-fragment top and bottom truncations.

Truncating mutations were designed using the SHAPE predicted structure. A. Truncations from the top of the S-fragment, number representing total removed nucleotides, T-70, T-148, T-246 and T-295. B. Truncations from the bottom of the S-fragment, starting after the EagI sites, B-48, B-97, B-195, B-271. Numbers indicate nucleotide position.

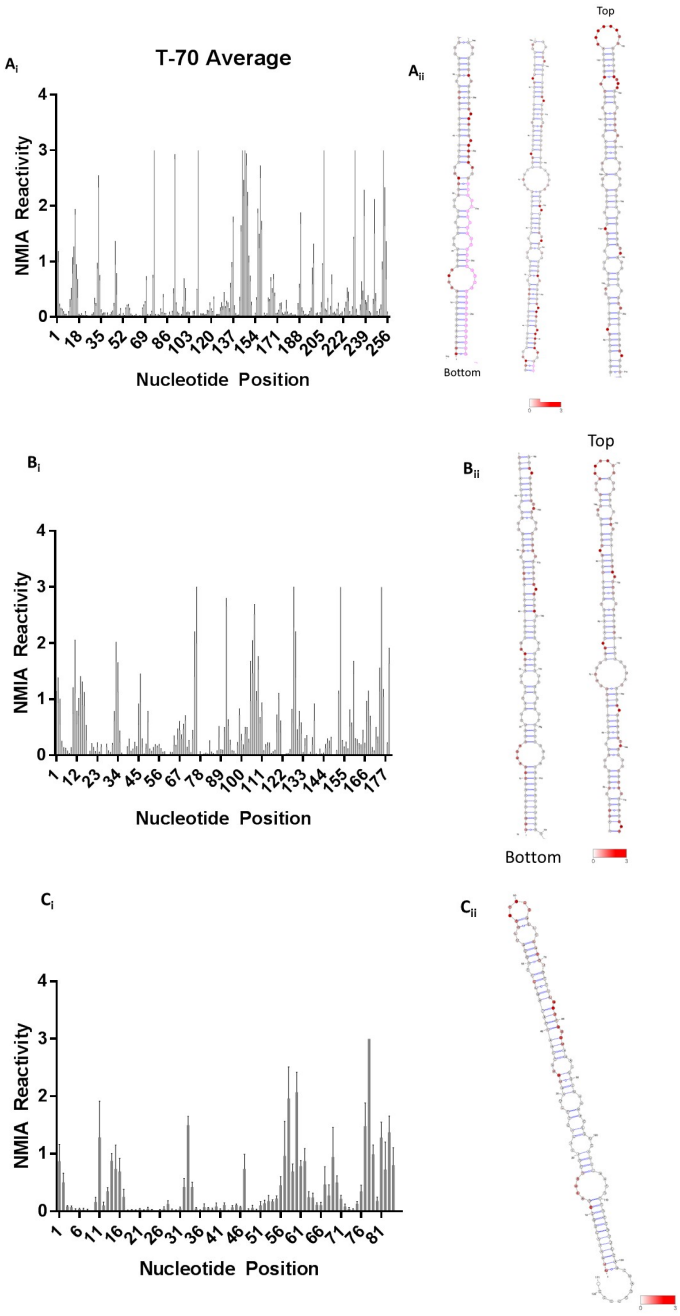

Figure 4.12: NMIA reactivity and predicted structure S-fragment truncations.

SHAPE experiments were performed on the mutant truncated S-fragment with 70 (A), 148 (B) or 246 (C) nucleotides removed from the top. (i) individual nucleotide reactivity's of mutant S-fragments as calculated using QuSHAPE. Non reactive nucleotides were set to 0 and 'super reactive' set to a maximum of 3. (ii) A new predicted structure was generated using the SHAPE reactivity's using the mFold Vienna server and visualised in VARNA. NMIA reactivity's were overlaid and represented on a colour scale from 0 (low) in white to 3 (high) in red. (n=3 T-70, n=3 T-148, n=4 (T-246), error bars represent SEM).

SHAPE reactions were undertaken with replicon constructs which has truncations to the top of the stem-loop (T-70, T-148, T-246, T-295) to ensure that the truncating deletions did not affect the stem-loop structure, as described before for the WT S-fragment. All truncating mutants tested kept the expected structure and were carried forward into replication experiments (Fig 4.12).

Replicon RNA was created by transcription of replicon plasmid DNA containing the truncation mutations. Purified replicon RNA was then transfected into BHK-21 cells. T-70 was tolerated with no significant reduction in replication, which was expected as viruses containing this truncation exist in the wild. T-148 was likewise tolerated, however, a small, non-significant reduction in replication was observed. The larger truncations T-246 and T-295 caused a significant reduction in replication, with GFP expression reduced to the level of the 3D-GNN negative control (Fig 4.13A). The bottom truncations showed similar results with the smaller truncations B-48 and B-97 being tolerated with no significant reduction in replication whereas B-271 appeared to reduce replication to the level of input translation. The B-195 deletion showed a significant decrease when compared to the WT, but was still significantly above that of the 3D-GNN negative control (Fig 4.13B).

Interestingly the T-148 and B-195 deletions overlap meaning there are no conserved nucleotides between them, yet both can perform as functional S-fragments. This suggests that it is not necessarily the sequence of the S-fragment, but the size, which is essential for function possibly providing an explanation to how it can tolerate such sequence diversity as described earlier.

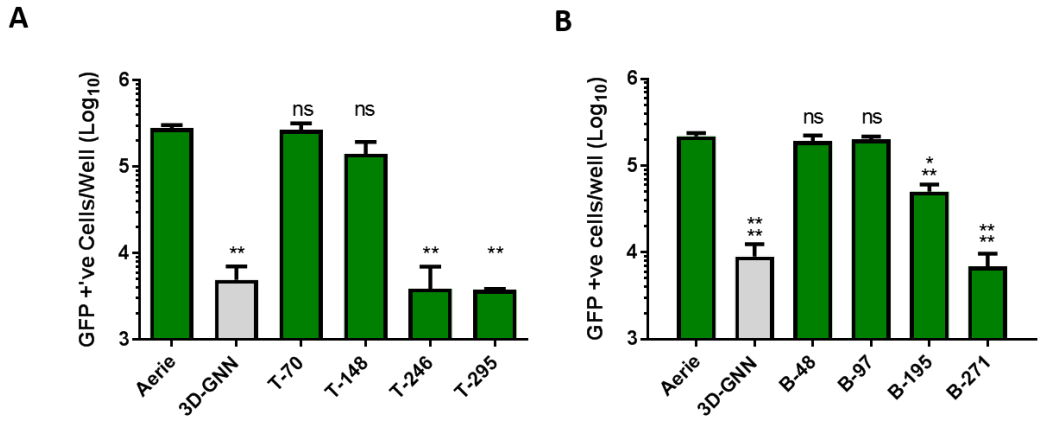

Figure 4.13: Replication of top and bottom S-fragment truncated replicons.
(A) Replicons with 70, 148, 246 and 295 nucleotides removed from the top of the S-fragment. (B) Replication of replicons with 49, 97, 195 and 271 nucleotides removed from the bottom of the S-fragment. Replication shown at 8 hours post-transfection of BHK-21 cells alongside Aerie and Aerie 3D-GNN positive and negative controls. (n=3, error bars represent SEM) Significance shown when comparing mutant replicons to the Aerie replicon, ns $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, $p < 0.001$, **** $p < 0.0001$.

Replicon RNA containing the truncated S-fragment mutations was also transfected in a similar manner into MDBK cells. These are a bovine cell line which, unlike BHK-21 cells, have a fully functioning immune response. The deficiency in BHK-21 cells is not fully known, however it is believed that they can react to IFNs but do not have the ability to produce these intrinsically.

The replicon containing the T-70 truncation performed similarly to the WT Aerie replicon in MDBK cells as it did in BHK-21 cells, however, T-148 which showed replication in BHK-21 cells at levels replicating the wildtype did not show any replication in MDBK-21 cells, while T-246 and T-295 replicated in neither BHK-21 nor MDBK cells. The bottom truncations showed a similar phenotype with B-48 and 97 replicating in both cells types while B-195, which replicated in BHK-21 cells showed no replication in MDBK cells (Fig 4.14).

This difference in the replicative ability when transfected into the different cell lines

was surprising. Further investigation into the effects of these truncations were carried out by inserting the mutations into the full length viral genome.

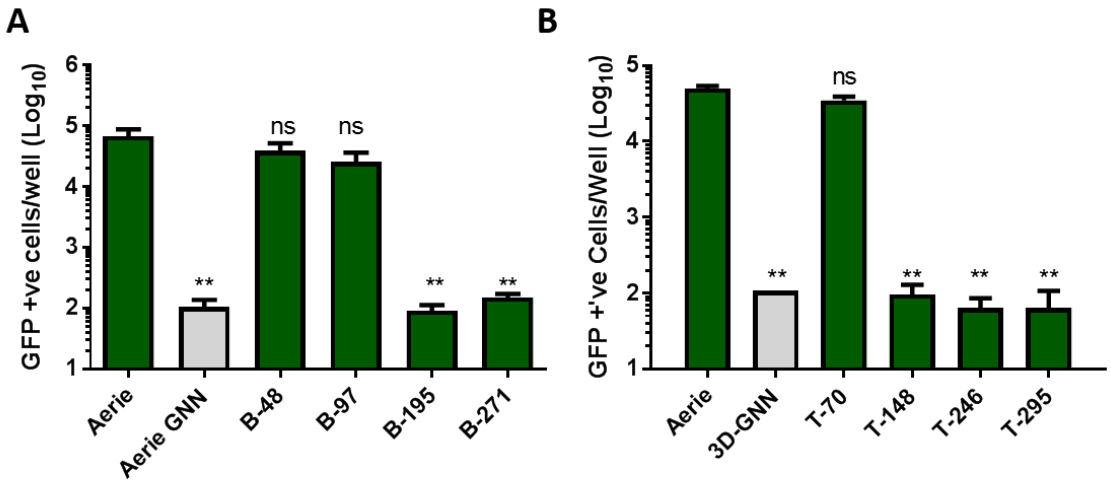

Figure 4.14: Replication of top and bottom S-fragment truncated replicons in MDBK cells.
(A) Replicons with 70, 148, 246 and 295 nucleotides removed from the top of the S-fragment. (B) Replication of replicons with 49, 97, 195 and 271 nucleotides removed from the bottom of the S-fragment. Replication shown at 8 hours post-transfection of MDBK cells alongside Aerie and Aerie 3D-GNN positive and negative controls. (n=3, error bars represent SEM). Significance shown when compare to the Aerie control, ** $p < 0.01$.

4.2.5 Does size matter? Truncation of the FMDV viral S-fragment

In collaboration with the Pirbright Institute, and the efforts of post-doctoral researchers in the research group headed by Toby Tuthill, (Lidia Laseka-Dykes and Sarah Gold, the truncated S-fragment mutations were introduced into infectious virus, to observe any effects these mutations may have on elements of the viral life cycle not visible in the replicon system.

Replicons were converted into infectious virus by the removal of the fluorescent reporter and insertion of the O1K capsid proteins. Infectious clone RNA was then transcribed in a similar fashion to that of replicons and transfected into 25 cm^2 flasks of BHK-21 cells and blind passaged five times, full method described in materials and methods 2.8. The appearance of CPE was noted and recorded at each passage as an indication of recovered virus (Fig 4.15).

Recovered virus was sequenced at passage 4 to confirm the input sequence and look for additional mutations or reversions as described in the materials and methods section 2.8.

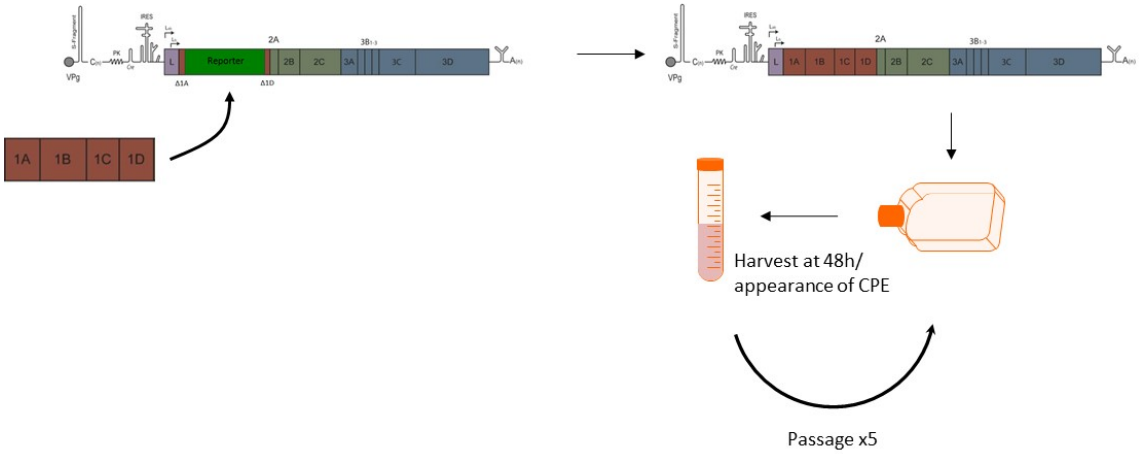

Figure 4.15: Recovery of mutant replicons in infectious virus.
DNA copies of replicons harbouring S-fragment mutations were converted into infectious virus clones by removal of the reporter gene and replacement with capsid proteins. Infectious clones were transcribed and RNA transfected into BHK-21 cells. At 48 hours post-transfection or when CPE was observed, cell lysates were prepared by freeze/thawing overnight and lysate clarified by centrifugation. Clarified lysate was passaged blind onto naïve BHK-21 cells. Viruses harvested after five passages were sequenced to check for additional mutations or reversions.

4.2.5.1 Recovery of WT and Aerie infectious clones

Initially WT and Aerie infectious clones were transcribed and transfected into BHK-21 cells. Once transcribed the 'recovered virus' was blind passaged five times as described above. The WT infectious clone was used as a positive control and was undertaken only once, producing infectious virus in this single attempt. The Aerie infectious clone was transfected three times, producing infectious virus two out of the three attempts. Sequencing of recovered virus indicated no change in sequence, suggesting the addition of EagI sites was tolerated, as seen in replicon assays (Table 4.1).

Virus	Number of Transfections	Virus Recovered	Virus not recovered	Sequence of Recovered virus
WT	1	1	0/1	No Change
Aerie	3	2/3	1/3	No Change
T-70	1	1/1	0/1	No Change
T-148	6	1/6	5/6	No Change
T-246	6	4/6	2/6	No Change
T-295	3	0/3	3/3	N/A
B-48	3	3/3	0/3	No Change
B-97	3	3/3	0/3	**Change**
B-195	3	0/3	3/3	N/A
B-271	3	0/3	3/3	N/A

Table 4.1: Recovery of S-fragment mutations in virus.
Recombinant viruses were transfected into BHK-21 cells and blind passaged 5 times. Biological replicates are indicated by number of transfections and subsequent virus recovery is shown. At passage 4, recovered viruses were sequenced byMiSeq high-through put sequencing to identify additional mutations or reversions. *Data provided by Lidia Laseka-Dykes, the Pirbright Institute.*

4.2.5.2 Recovery of Top truncations

Top-truncations representing T-70 was similarly undertaken once, with recovered virus produced and no change in sequence reported. The T-148 infectious clone was transfected 6 times into BHK-21 cells, only one attempt out of the 6 produced infectious virus, again. Sequencing revealed no change when compared to the input sequence. T-246 was similarly transfected 6 times, with infectious virus produced on four occasions, and no sequence change observed from recovered virus. This is a striking difference to replicon assays, where the T-246 showed no signs of replication, in either BHK-21 or MDBK cells. Likewise, T-148 which showed high levels of

replication in BHK-21 cells in replicon struggled to produce infectious virus. As seen in the replicon experiments the T-295 infectious clone produced no infectious virus in any of the three repeats, as no virus was produced sequencing could be performed (Table 4.1).

4.2.5.3 Recovery of bottom truncations

Infectious clones were also made from the replicons containing truncations from the bottom of the S-fragment (B-48, B-97, B-195 and B-271). Infectious clones were transcribed as above and transfected into BHK-21 cells. B-48 and B-97 infectious clones both recovered infectious virus in all three repeat transfections, showing similar results to the replicon experiments. Sequencing of recovered virus showed no change to the input sequence for B-48, however, when recovered virus for B-97 was sequenced several mutations in the non-structural proteins were detected. Interestingly, neither B-195 nor B-271 showed recovery of infectious virus, despite B-195 showing evidence of replication when introduced into replicon and transfected into BHK-21 cells (Table 4.1.)

4.2.5.4 Investigation of B-97 recovered virus

Recovery of virus containing the B-97 mutation was repeated three times, with infectious virus recovered successfully each time, sequencing of each of the recovery attempts was performed to look for mutations from the input sequence. Transfection 1 revealed mutations within the viral 2C non-structural protein. Three mutations were observed each appearing at approximately 70 % within the viral population (70 % of sequenced RNA copies containing mutations). A synonymous mutation was reported in the first amino acid of 2C, encoding a cytosine to adenosine change, whilst not changing the amino acid sequence. Other mutations reported included changes to the amino acid sequence with a mutation of isoleucine to valine at residue 249 (I249) caused by a single nucleotide change from an adenosine to a guanosine

(A5115G). The third mutation from these recovered viruses included a serine to arginine (S318R) caused by a uracil to adenosine change (U5309A).

The second repeat of recovering B-97 infectious virus revealed mutations within the viral $3D^{pol}$, only one mutation was observed, with the subtle change of an isoleucine to a leucine (I189L). With this recovery, additional sequencing reactions were performed in an attempt to trace when the mutation arose. Sequencing at passage three indicated the I189L mutation was prevalent at 40 % within the population, at passage four this had increased to 51 % and at passage five increased again to 61 %. Indicating a preference and enrichment of this mutation in the recovered viruses.

The third recovery of B-97 revealed no mutations in the sequence of recovered virus at passage 5, therefore no further sequencing of this sample was performed (Table 4.2).

		Mutation prevealence				
B-97 Transfections	Location of mutation	Passage 1	Passage 2	Passage 3	Passage 4	Passage 5
1	2C	?	?	?	I249V (ATC>GTC) 69% S318R (AGT >AGA) 70% (S) L1L (CTC >CTA) 72%	?
2	$3D^{pol}$	?	?	I189L (AUC>CUC) 40%	I189L (AUC>CUC)51%	I189L (AUC>CUC) 61%
3	None	?	?	?	?	✖

Table 4.2: Analysis of B-97 mutations.
MiSeq high-throughput sequencing of recovered viruses at different passages revealed enrichment of mutations in viral populations. Red cross indicates no identified mutation. Question mark represents absence of sequencing data (sequencing in progress). *Data provided by Lidia Laseka-Dykes, the Pirbright Institute.*

4.2.5.5 B-97 virus showed increased rates of cell killing when compared to B-48.

Comparison of the time taken to reach CPE appeared to vary when comparing the recovery of B-48 and B-97 infectious clones. At passages 1 and 2, no CPE was seen in either B-48 or B-97 experiments. However, at passage 3, transfection 1 of B-97 appeared to be slower than the recovery of B-48, transfection 2 of B-97 was faster and transfection 3 was noted to have a similar rate of CPE as B-48 (Fig 4.16A).

At passage 4, both transfections 2 and 3 showed an increase in time to CPE, being fast than the B-48 virus, while transfection 1 was still reported to be slower. At passage 5 however, all three transfections were shown to reach CPE faster than the attempts to recover virus containing the B-48 truncation.

The rate of CPE was further confirmed using cell killing assays. BHK-21 cells were infected with passage 5 recovered virus and cell confluency monitored hourly using the Incucyte zoom. Previously, the Incucyte has been described to count fluorescent cells for the monitoring of replicon replication, in this experiment the phase setting was used to monitor BHK-21 cell confluency. Each of the three recovery transfections of B-97 (B-97 1, 2 and 3) were used to infect the BHK-21 cells alongside WT, Aerie and B-48 recovered virus. Mock transfected cells were used as a negative control (Fig 4.16B). WT virus showed to reduce the cell confluency at the fastest rate, followed by virus containing the Aerie mutations. Passage 5 recovered virus from each of the B-97 transfections showed a similar rate of reduction in cell confluency, with B-48 being the slowest to cause CPE.

It is remarkable that, despite having no reported mutations within the viral genome, B-97 (3), still showed a faster rate of CPE than B-48, while B-97 (1) which showed presence of mutations in the 2C protein at passage 4, produced CPE more slowly until an increase in speed at passage 5. These data suggest the mutations were not necessary to result in a faster rate of CPE than B-48 virus.

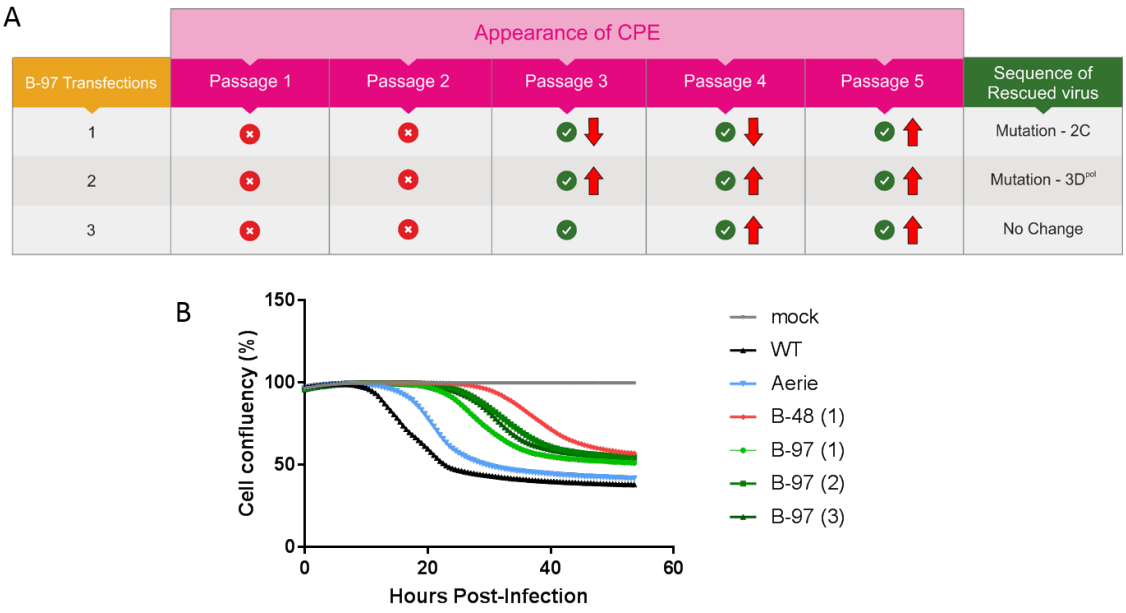

Figure 4.16: Analysis of B-97 CPE.
(A) Biological replicates are represented by transfection number. CPE was monitored over 5 passages. Red crosses indicate no observable CPE. Green ticks represent visible CPE. Red arrows indicate rate of CPE compared to B-48 virus recovery (up arrows represent faster rate of CPE, down arrows represent slower rate of CPE). Sequence of passage 5 virus highlighted mutations in specific non-structural proteins. (B) Passage 5 virus was infected onto BHK-21 cells and cell confluency monitored using the Incucyte Zoom. Virus recovered from one transfection of B-48 (B-48 (1)) was compared to the three of B-97, Aerie and WT virus. Uninfected mock cells were used as a negative control. n=1. *Data provided by Lidia Laseka-Dykes, the Pirbright Institute.*

4.2.6 Exploration of the novel mutation in the 3D polymerase

With the identification of a novel mutation in the polymerase of viruses containing the B-97 mutation, further investigation was undertaken to see any potential role this mutation has on the viral replication cycle. The I189L mutation was mapped to the crystal structure of the FMDV type C $3D^{pol}$ when complexed with RNA, Structural predictions were made using pymol to visualise the change in residue (Fig 4.17)[57].

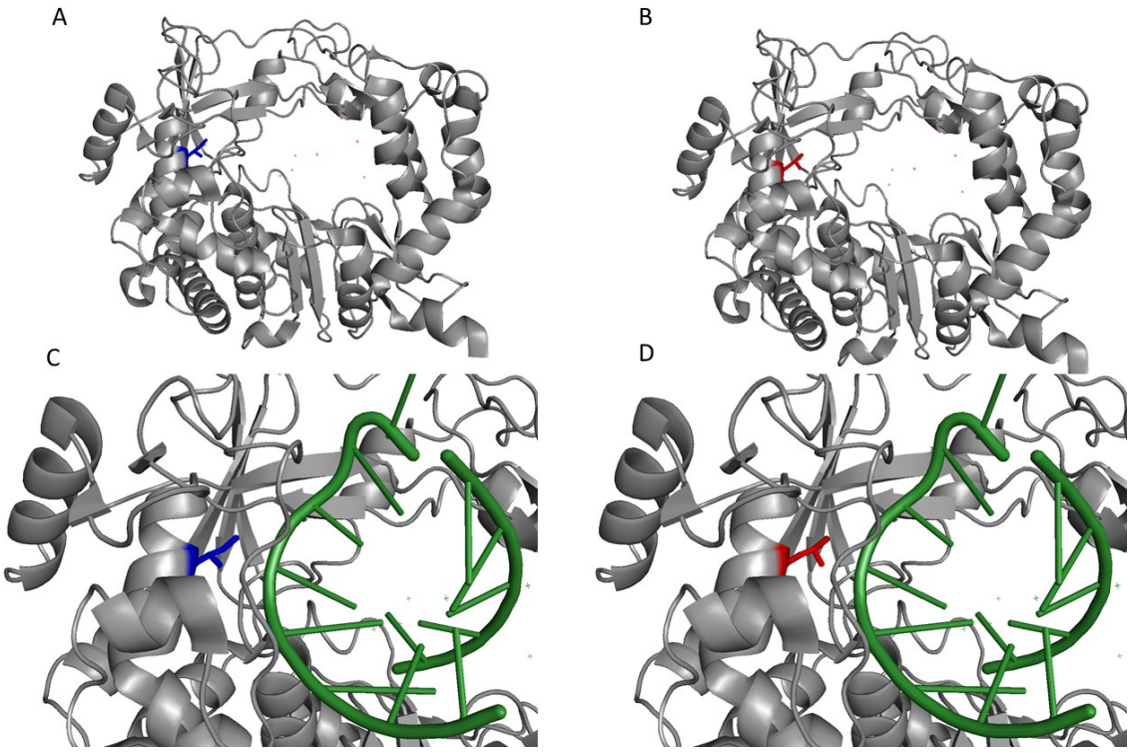

Figure 4.17: Structural prediction of 3D I189L mutation.
Ribbon diagram of the 3D protein (grey) with I189 shown in blue (A) and substitution of I189L in red (B). Visualisation of the 3D polymerase when complex with RNA is shown, with WT I189 residue (C) and I189L (D). Mutations were mapped to the type C polymerase crystal structure (PDB 1WNE)[57].

Previous experimentation on picornavirus 3D polymerases has used sensitive biochemical approaches to great effect. A technique first developed by the Cameron group (university of Pennsylvania) uses a small 10mer of RNA which can anneal to other copies of itself[7]. Once annealed a small region of 6 nucleotides is base paired, acting as a primer for 3D polymerase activity. Depending on the addition of other factors such as nucleotides or other fragments of RNA, changes in 3D polymerisation can be observed at the nucleotide level. Figure 4.18 provides a basic overview of the sym/sub experimental protocol, briefly, the 10 mer of RNA, known as "sym/sub", is 5′ end labelled using 32P UTP and T4 polynucleotide kinase. Once labelled and purified, this can be incubated to promote self-annealing to act as both primer and template. The addition of polymerase and nucleotide can di-

rect the extension of sym/sub in a specific manner. This extension can be observed over time by electrophoresis of harvested time points on a high percentage (23 %) acrylamide-urea-TBE gel, followed by exposure of the electrophoresed gel onto film or a phosphoscreen.

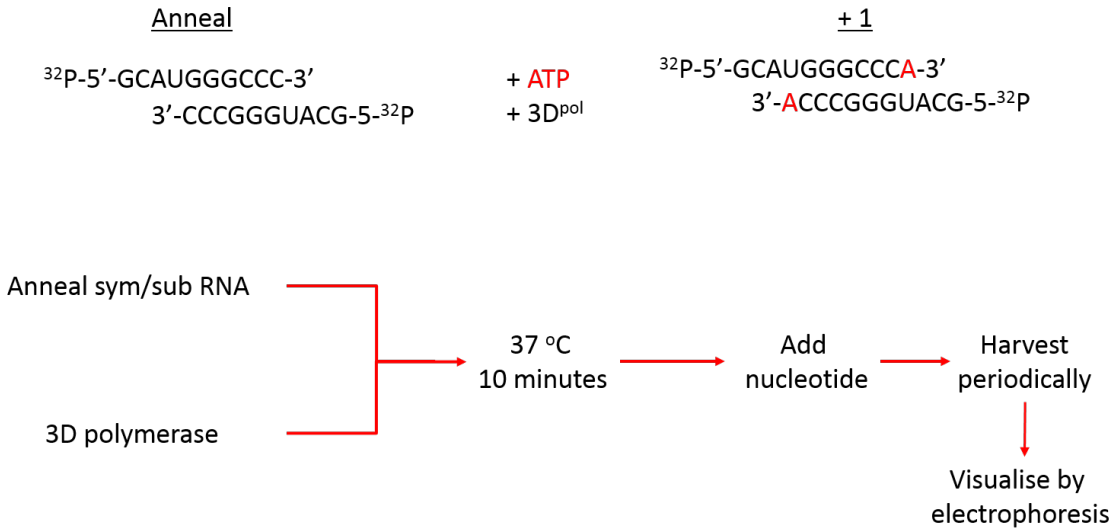

Figure 4.18: Pipeline of sym/sub experimental protocol.
A. Representation of the 5′ 32P end labelled sym/sub annealing with another copy, after addition of ATP and the 3D polymerase ($3D^{pol}$), elongation of both sym/sub strands can be seen creating the +1 RNA. B. Pipeline of the sym/sub experimental protocol.

The sym/sub assay was initially tested using bacterially expressed recombinant wild-type 3D polymerase alongside recombinant 3D-GNN polymerase (both produced by a previous PhD student, Eleni-Anna Loundras). Wildtype and GNN polymerases were used as positive and negative controls to confirm the validity of the sym/sub protocol. With the addition of ATP as the sole nucleotide, a clear extension of the sym/sub template can be seen, creating the +1 band. This occurs rapidly, with most of the extensions taking place within the first 20 seconds. As expected, the +1 band is not observed with the GNN polymerase as polymerase activity is knocked out by the GNN mutation (Fig 4.19).

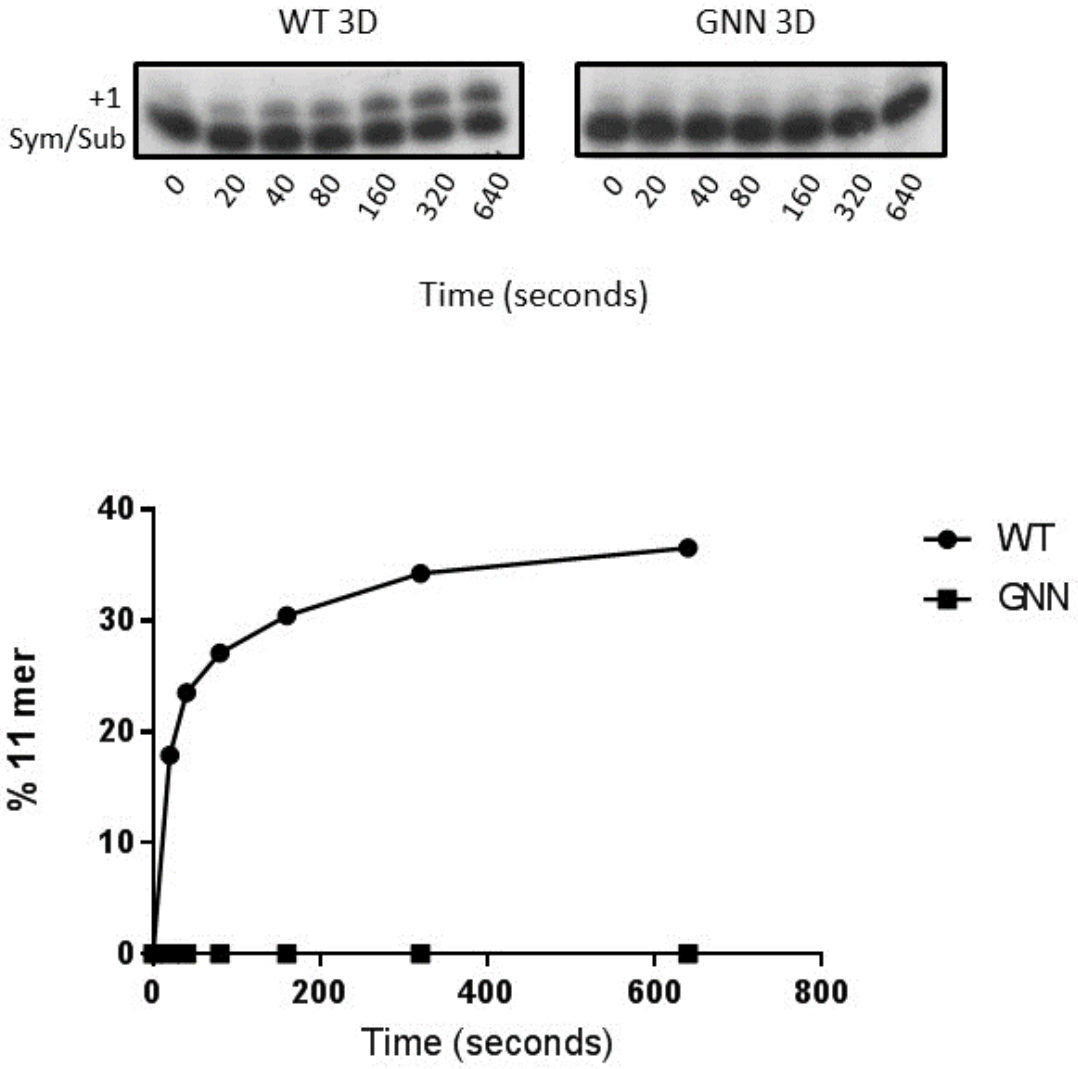

Figure 4.19: Comparing the polymerase activity of WT and GNN 3D polymerases

A. WT and GNN polymerases were added to sym/sub reactions alongside ATP, aliquots taken periodically reveal the extension of the sym/sub template in the WT and not in the GNN. B. Proportion of total sym/sub template being elongated by one nucleotide to create the 11mer in both WT and GNN conditions, n=1.

The I189L mutation was cloned into a 3D expression construct (a kind gift from Esteban Domingo) using PCR mutagenesis. This construct is composed of a pet28 expression plasmid encompassing the FMDV type C 3D polymerase with a C terminal his-tag for efficient purification. Bacterial expression of the I189L 3D was performed by Natalie Kingston, a post-doctoral researcher in the Stonehouse/Rowlands group using protocols outlined in materials and methods 2.5. Upon elution from the his-trap affinity column, samples were electrophoresed and visualised by coomassie blue staining and western blot. Coomassie blue staining of the samples showed large amounts of material in elution 2, consistent with the size of the 3D protein (53 kDa). This band was also present in the 3D-His positive control, as well as both lysate and pre-filtered lysate samples. Decreased amounts of this band were also observed in elutions 1 and 3-9. The identity of this majority product was confirmed by western blot, using antibodies against both the his-tag and 3D specific polyclonal antibodies (Fig 4.20). Following expression, the bacterially expressed protein was dialysed and quantified using a BCA assay.

The recombinantly produced I189L 3D was used in a sym/sub assay alongside the WT 3D polymerase. Differences in the kinetics of elongation could provide an insight to how the I189L mutation facilitates increased CPE rates observed in the virus recovery. Initially the sym/sub assay was performed in a similar manner to comparing the WT and GNN polymerases. WT and I189L polymerases were incubated with 32P labelled sym/sub RNA for 10 minutes at 37 ^{o}C before the addition of ATP. As ATP is the only nucleotide added, polymerase extension was limited to a single nucleotide (+1) to create the 11 mer product.

Comparison of the WT and I189L polymerases revealed no significant differences in single nucleotide addition (Fig 4.21) with the reaction occurring rapidly and most of the 11 mers were made within the first 20 seconds.

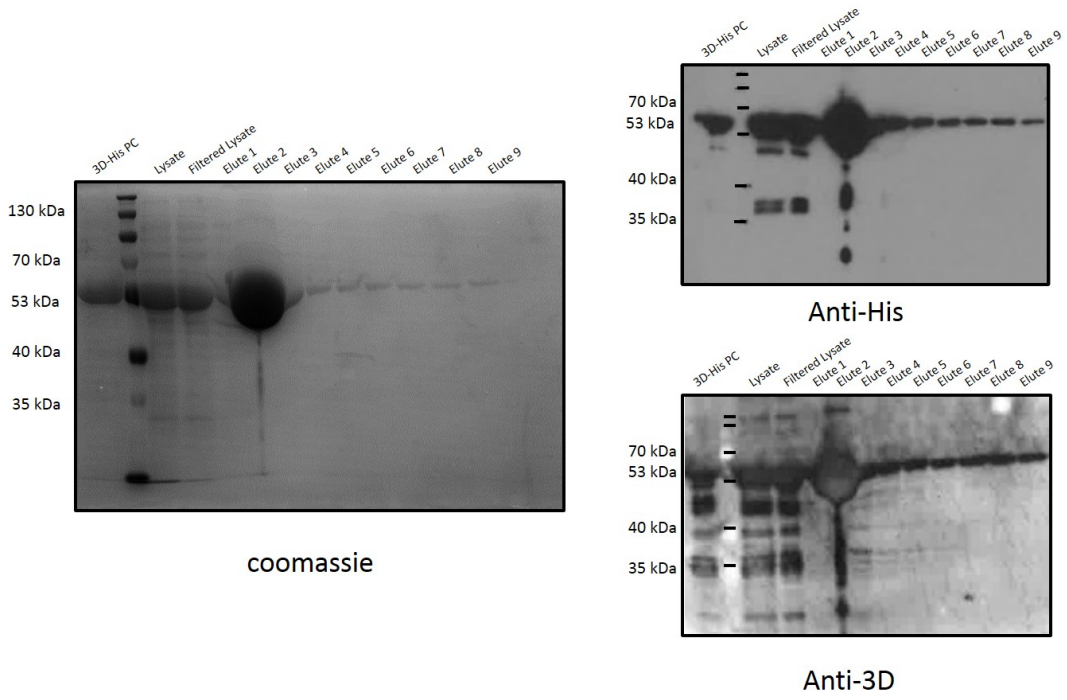

Figure 4.20: Production of recombinant I189L 3D polymerase.
Bacterially expressed 3D polymerase encompassing the I189L mutation was prepared. A. Coomassie blue stained gel of crude bacterial lysate, filtered lysate and elutions from his-trap column. Left hand side shows the 3D-his positive control (PC). B. Western blot imagery of recombinantly produced 3D I189L probed with an anti-His antibody (B) and anti-3D antibody (C)

To see if any difference in further nucleotide addition could be detected, the same experiment was performed to include all four nucleotides which were added after the incubation of polymerase and sym/sub. No changes were observed when looking at the band corresponding to the addition of four nucleotides (+4) to generate a 14mer; both polymerases produced equal amounts at the same rate, with approximately 30 % of the total input sym/sub RNA extended by 20 seconds (Fig 4.22A). Surprisingly, a much larger band was also produced by both polymerases, longer than the +4 extension (Fig 4.22B). Moreover, this band was produced more rapidly and to a higher proportion by the mutant I189L polymerase than the WT (Fig 4.22C). Due to the design of sym/sub RNA, polymerase extension should be limited at four nucleotides (incorporating one of each nucleotide) therefore, how this larger product, named the "super product" was synthesized remains a mystery.

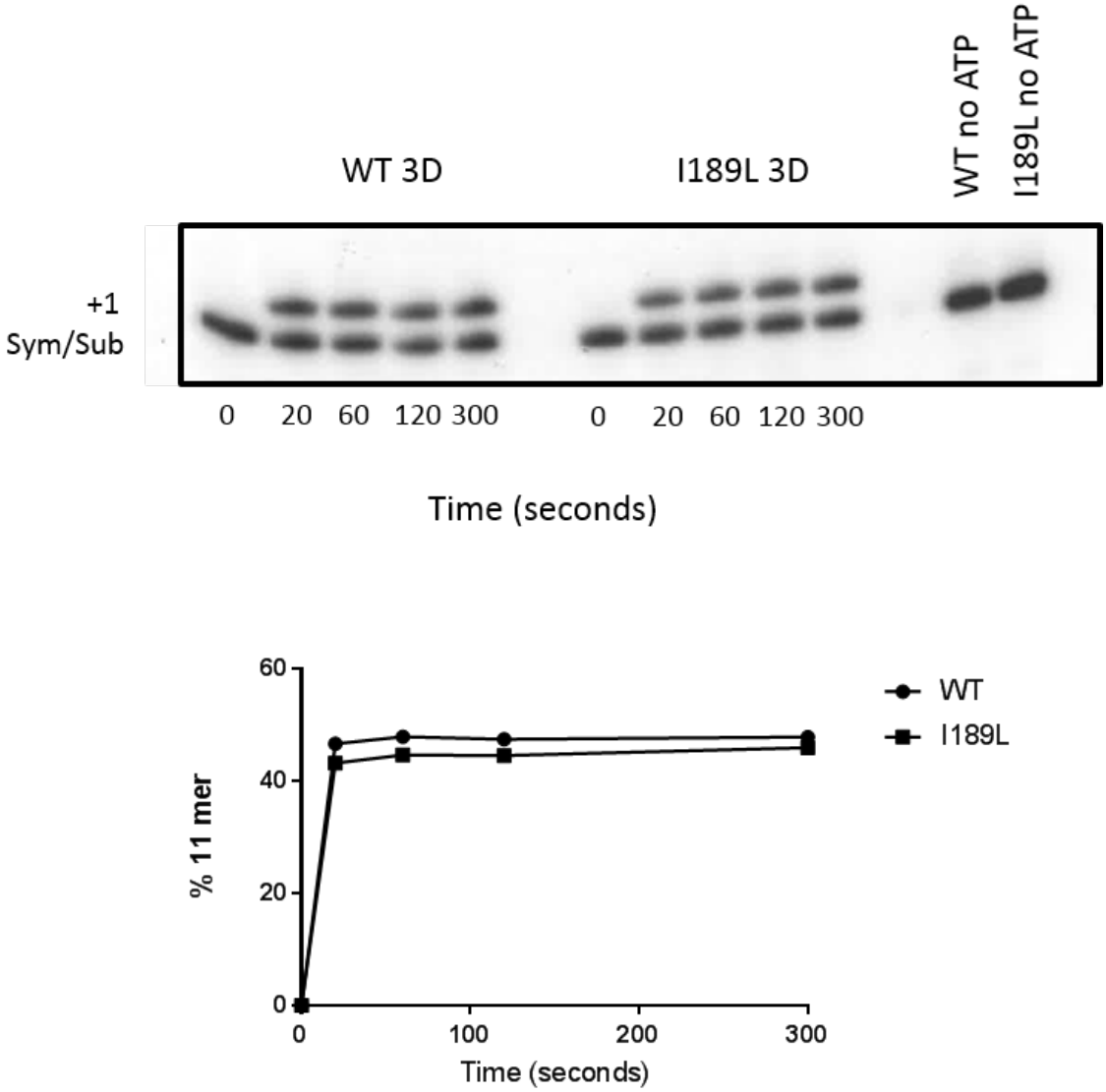

Figure 4.21: Polymerase activity of WT and I189L 3D polymerases. Comparison between WT and I189L 3D polymerase to incorporate ATP in the elongation of sym/sub. A. Shows radiography of elongation over time. B. Densitometry performed on radiograph shows the percentage of total sym/sub RNA extended by one nucleotide to make the 11mer, n=1.

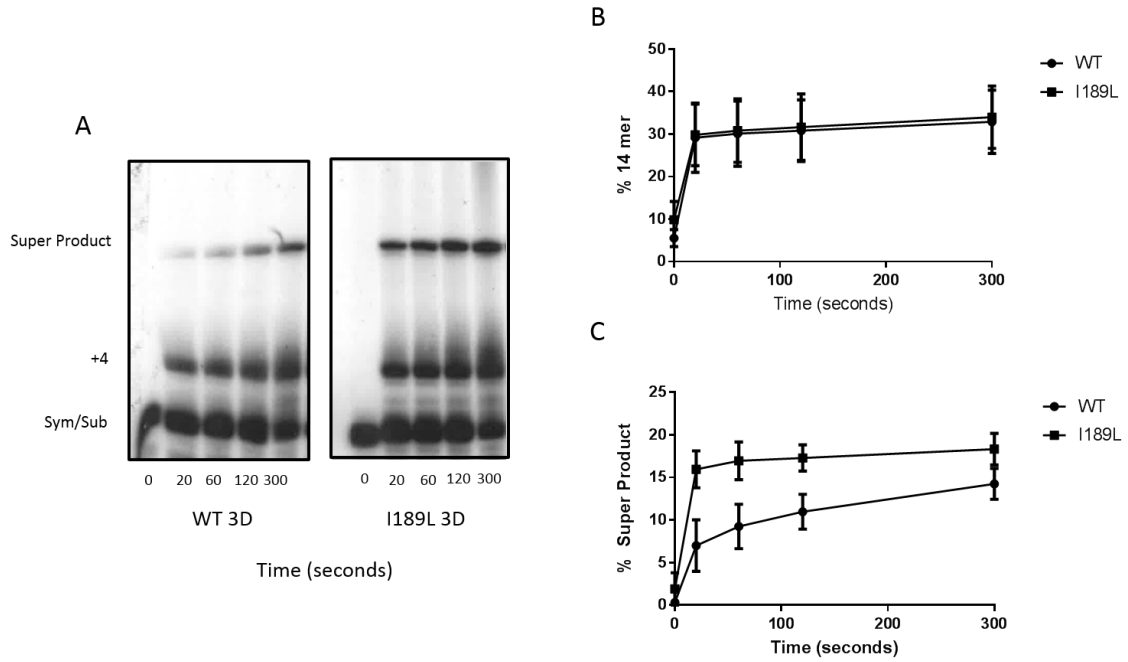

Figure 4.22: Polymerase activity of WT and I189L 3D polymerases in the presence of all nucleotides.
A. Extension radiographs of WT and I189L, visible are the sym/sub template, +4 nucleotide 14mer, and super product. Densitometry of radiographs was performed, displaying the percentage of total sym/sub RNA extended to make the 14mer (B) or super product (C), n=2.

Investigation into the super product was initiated by sequential addition of each nucleotide to try and ascertain when and how it appears. Sym/sub reactions were assembled using the WT polymerase, with the addition of ATP, ATP and UTP, ATP and UTP and GTP, or all four nucleotides after the incubation of RNA and polymerase. Following nucleotide addition, the reaction was incubated at 37 ^{o}C for 10 minutes, before an aliquot was taken and electrophoresed as before. With the addition of ATP, the standard +1 band could be seen as before and no further extension products were observed. When adding both ATP and UTP a slightly larger product could be seen, presumably representing +2 nucleotides. Notably, in this control, faint bands higher up corresponding to +3 and +4 extensions are observed, suggesting possible misincorporation of ATP or UTP nucleotides during elongation. Addition of ATP, UTP and GTP results in the production of +3 and

+4 bands in equal amounts, suggesting misincorporation of nucleotides to produce the +4 band. Finally, when adding all four nucleotides, a distinct band at +4 can be seen, as well as the larger super product (Fig 4.23A).

The observation of potential misincorporation events could hold the key to how the super product is produced. If misincorporation of a GTP occurs at the final nucleotide of the +4 extension, it is possible that a new template could be bound providing 6 nucleotides of complementation. This new template would permit the addition of a further 8 nucleotides making a total product of 12 nucleotides, approximately matching the size of the super product (Fig 4.23B). The questions remain: why does this occur more rapidly in the I189L polymerase than in the WT, and does this explain the phenotype of increased CPE observed in virus recovery?

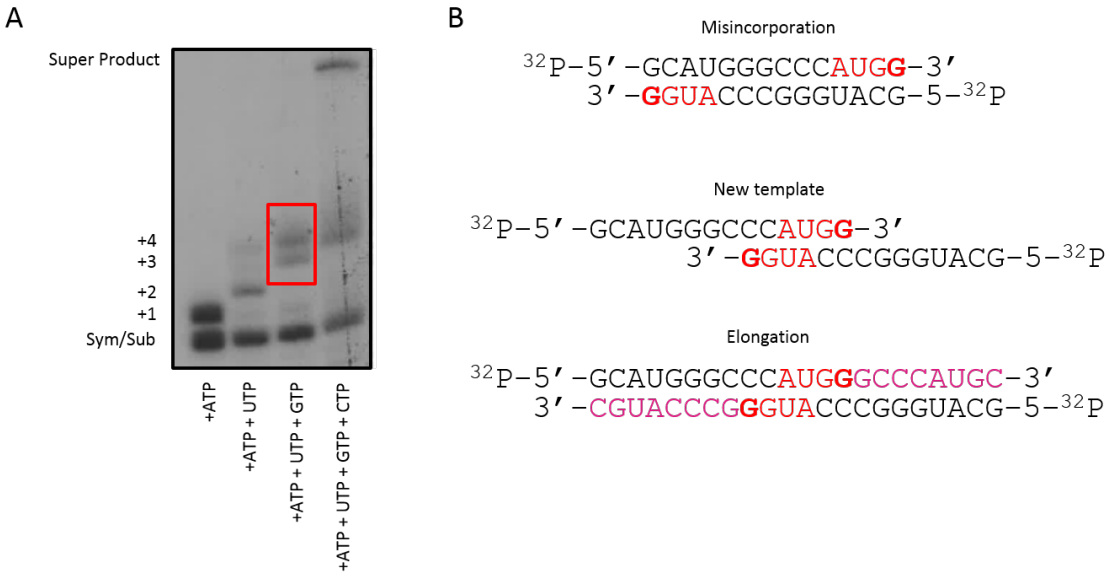

Figure 4.23: Sequential addition of nucleotides to sym/sub reactions.
A. Radiograph showing the products of sequential addition of nucleotides to a sym/sub reaction using WT 3D polymerase. Misincorporation events can be seen highlighted in the red box. B. Hypothetical consequences of misincorporation leading to the production of the super product. Initial elongation shown in red with misincorporation events shown in bold and additional elongation shown in hot pink.

4.2.7 Paying complements: Transcomplementation of the S-fragment

Many functions of the FMDV genome can be provided in trans, where protein functions can be provided by other proteins produced from different constructs (Introducion 1.3). One example of this can be seen in the case of the 3D polymerase; we have shown that a 3D molecule from a separate replicon can induce replication of a replicon carrying the replication defective GNN mutation, as described in Section 3.2.5[78]. 'Function donation' is not limited to protein activities as it has also been reported that some RNA structures, including the cis-active replicative element 'cre'. The essential RNA structure for the uridylylation of the 3B protein, can also provide their function in trans[192].

To investigate whether the S-fragment could also provide functions in trans, we co-transfected two replicons containing either a GFP or mCherry reporter into BHK-21 cells. Wild-type or ΔS GFP replicons were co-transfected with either WT, ΔS or cre A1G (a lethal mutation disabling the template function of cre for uridylylation of 3B) mCherry replicons. The ΔS replicon was also co-transfected with yeast-tRNA serving as a substitute for the helper replicon to maintain equal amounts of transfected RNA, thereby functioning as a negative control.

The results showed that co-transfection of the ΔS replicon with either 3D-GNN or cre A1G replicons led to an increase in GFP expression, compared to yeast tRNA (Fig 4.24). This suggests that functions of the S-fragment could indeed be provided by another construct. Interestingly, the WT mCherry replicon was not able to provide S-fragment function(s) in trans. This phenomenon has been previously documented; where 3D-GNN and 3B123 Y3F mutants could enhance replication of a mutant cre more efficiently than a WT replicon[78]. One explanation of this could be due to rapid product turnover in WT replicon replication, limiting the chance that these products are able to provide function in trans.

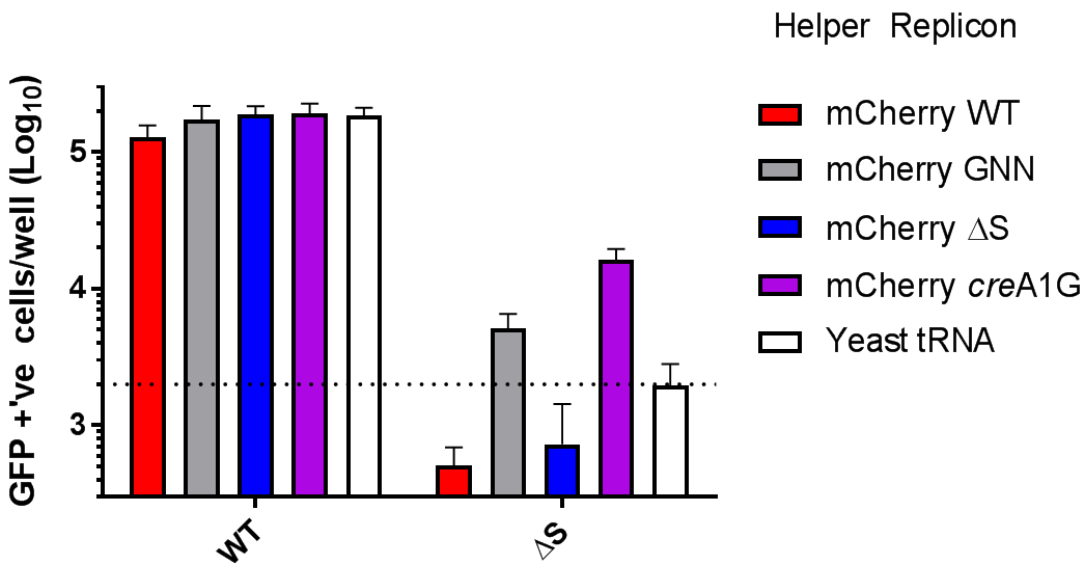

Figure 4.24: Co-transfection of ΔS replicon with helper mCherry replicons.
WT of ΔS ptGFP expressing replicons were co-transfected with either WT, 3D-GNN, ΔS or creA1G mCherry replicons, or yeast tRNA as a negative control. Replication shown 8 hours post transfection with y-axis line indicating the cut-off for complementation. (n=2, error bars represent SEM).

The S-fragment complementation described above was further explored by investigating the minimum requirements for complementation. For this, the ΔS replicon was co-transfected in combination with one of the following RNA fragments: full length 5′ UTR, 5′ UTR without the IRES (UTR ΔIRES), 5′ UTR without the IRES or pseudoknots (UTR ΔIRES ΔPK), 5′ UTR without the S-fragment (UTR ΔS), the S-fragment by itself, or yeast tRNA as a negative control. Interestingly co-transfection of the ΔS replicon with full length 5′ UTR RNA was sufficient to generate an increase in GFP expression. However, GFP expression was reduced in transfections with truncated forms of the UTR, specifically, when the IRES was removed, complementation was reduced (Fig 4.25).

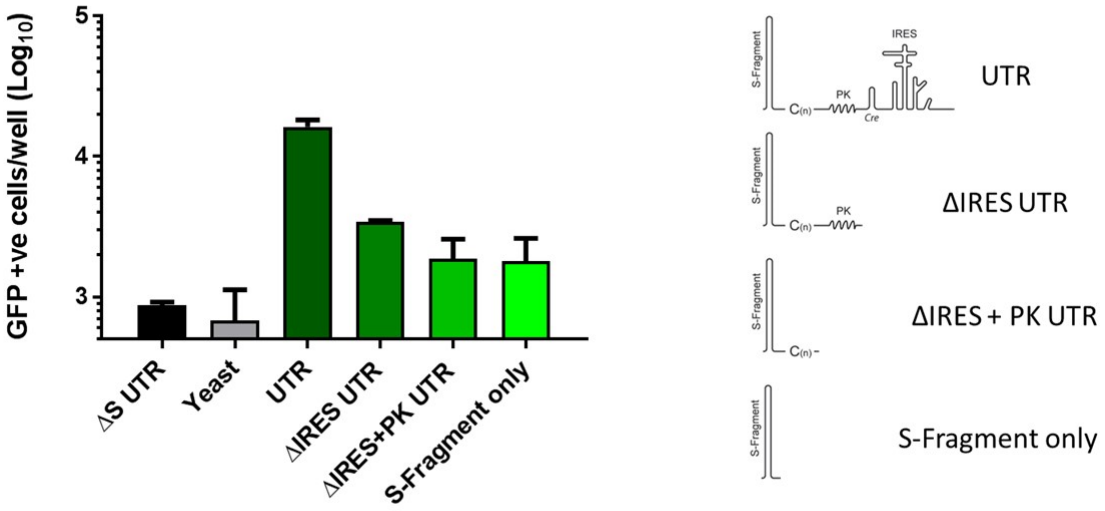

Figure 4.25: Co-transfection of ΔS replicon with 5′ UTR RNA.
A ptGFP expressing ΔS replicon was co-transfected with RNA representing either the 5′ UTR, 5′ UTR ΔIRES, 5′ UTR ΔIRES + ΔPKs, 5′ UTR ΔIRES, ΔPKs and ΔPCT. As well as ΔS UTR and yeast tRNA as negative controls. Replication shown at 8 hours post-transfection. (n=2, error bars represent SEM).

The observed results suggests that the IRES could play an important role in the complementation of S-fragment function. To see whether this was specific to the FMDV IRES, we created constructs where the FMDV IRES was replaced with that of PV and the cardiovirus, Mengo virus. Non-functional scrambled IRES sequence was used as a control. These mutated 5′ UTR RNAs were co-transfected with the ΔS replicon and GFP expression was measured. All of the tested IRES sequences, including the scrambled non-functional sequence were able to complement the S-fragment, suggesting that it is the size of the overall UTR, rather than the IRES sequence, that is important for complementation (Fig 4.26A). This could be caused by rapid degradation of small RNA fragments whereas larger RNAs containing complex secondary structures are comparatively more resistant from degradation and can persist long enough to provide complementation. To answer this question, artificial UTRs RNAs were created consisting of the S-fragment followed by 918 or 783 nts

of scrambled UTR sequence which represent a full length UTR and a UTR without the IRES, respectively. These scrambled sequences were co-transfected with the ΔS replicon and GFP expression observed as previously described. Indeed, the results showed that sequence specificity did not seem to matter as the larger sequence was able to provide function better than the smaller sequence (Fig 4.26B).

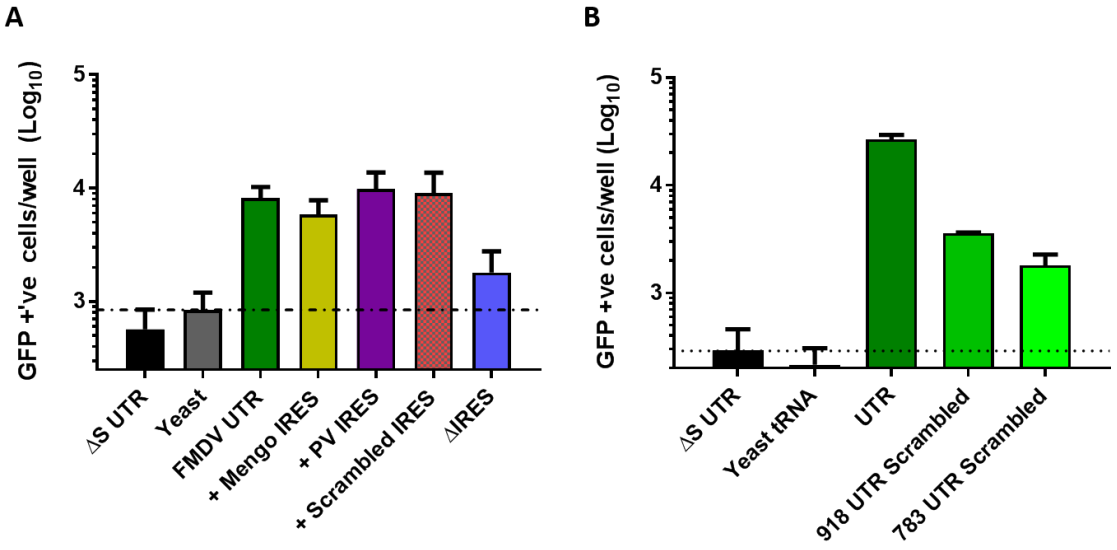

Figure 4.26: Co-transfection of ΔS replicon with 5′ UTR RNA containing mutant IRES's or scrambled sequences.
(A) ptGFP expressing ΔS replicon was co-transfected with RNA representing either the full FMDV 5′ UTR or the 5′ UTR ΔIRES as controls, as well as FMDV 5′ UTR RNA were the FMDV IRES has been replaced for the Mengo virus IRES (+ Mengo IRES), PV IRES (+PV IRES), or a scrambled FMDV IRES sequence (+ Scrambled IRES). (B) ΔS ptGFP replicon co transfected with either UTR RNA or an S-fragment followed by 918 nts or 783 nts of scrambled sequence. Both experiments where also transfected with ΔS UTR and yeast tRNA as negative controls. Replication shown at 8 hours post-transfection. (n=2, error bars represent SEM).

Recovery of replication after co-transfection can also result from recombination between the two transfected constructs to produce one fully replicative replicon. To discover if the increase in ΔS replicon replication was due to recombination with the transfected helper RNAs, co-transfections were performed and total RNA was harvested at 8 hours post transfection by the addition of Trizol®. Total RNA was

then purified and re-transfected into BHK-21 cells and GFP expression observed. If recombination has occurred then the same increase in replication of the ΔS replicon might be observed, whereas if complementation alone has taken place the increase might not be visible. This has been proven to be a reliable measurement of recombination when investigating complementation of the 3D polymerase, where passaging of the RNA showed no increase in GFP production upon the second transfection[78]. When this was performed using co-transfections of the ΔS replicon with RNA representing either the FMDV UTR, or FMDV UTR Δ IRES or the WT replicon, an increase in replication could still be observed, suggesting that recombination of ΔS replicon and helper RNAs has occurred (Fig 4.27).

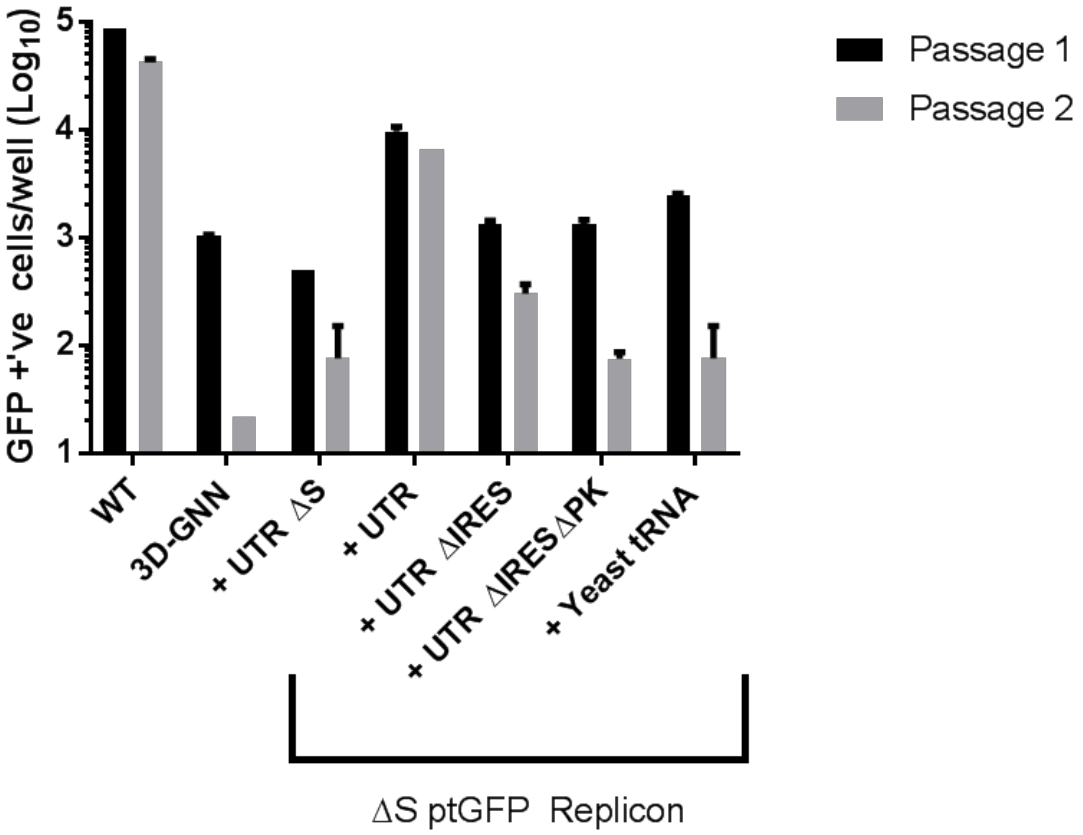

Figure 4.27: Passaging of trans-complementation experiments suggests recombination of the S-fragment.
Replicon replication of ΔS ptGFP replicon when co-transfected with various RNA representing truncated forms of the UTR. RNA has was harvested at 8 hours post transfection after initial transfection (Passage 0). RNA was then purified and re-transfected onto naïve BHK-21 cells (Passage 2). WT and 3D-GNN controls are shown as passaging controls. Replication expressed as the number of GFP positive cells per well at 8 hours post transfection in BHK-21 cells. n=1, error bars represent SEM of three technical repeats of one biological experiment.

4.2.8 "S-earching" for S-fragment protein interactors

Although it is speculated that both cellular and viral proteins bind to the S-fragment, to date, only specific interactions with RNA helicase A (RHA) have been experimen-

tally demonstrated. Here, RNA-protein pull down experiments were performed to identify other potential S-fragment binding proteins, with the aim of elucidating the functions of the S-fragment in replication. RNA representing different regions of the 5′UTR, (S-fragment alone, full length 5′ UTR, UTR ΔIRES, UTR ΔS-Fragment) were modified by the addition of a 3′ biotin molecule using the Thermo Fisher RNA-Protein pull down kit as described in materials and methods 2.12. The 3′UTR of the androgen receptor RNA (AR) and a 25-nucleotide poly-A RNA were used as positive and negative controls, respectively, and were also labelled. Biotinylated RNA was first immobilised on streptavidin conjugated magnetic beads then incubated with MDBK whole cell protein lysate in binding buffer to encourage protein:RNA binding. Following incubation, unbound proteins were washed off and any bound proteins were harvested by elution with lammeli buffer. Eluted samples were then analysed by mass spectrometry, performed by collaborators at the University of Dundee (Fig 4.28).

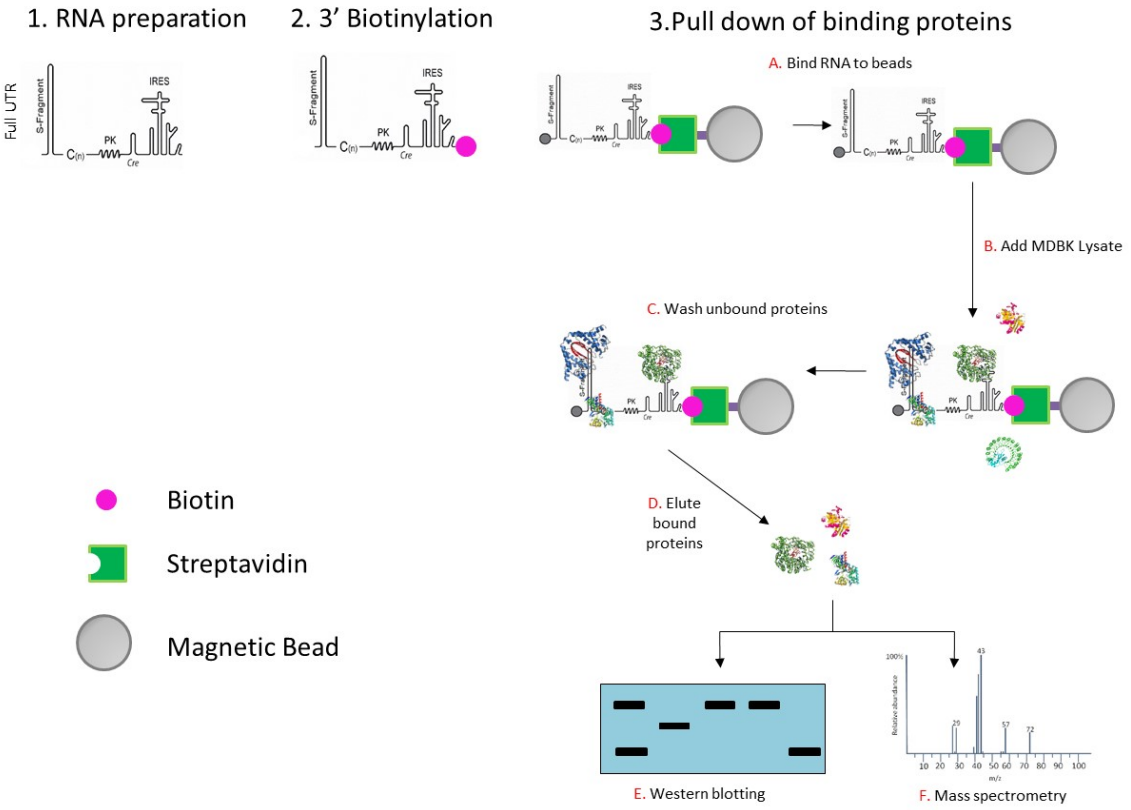

Figure 4.28: Pipeline of isolation and identification of RNA binding proteins.
1. RNA fragments representing different regions of the 5′ UTR of FMDV (S-fragment alone, full length 5′ UTR, UTR ΔIRES, UTR ΔS-Fragment) were biotinylated along side control RNA fragment representing the 3′ UTR of AR RNA and a 25 nucleotide poly-A RNA. Biotinylated RNAs were immobilised with streptavidin magnetic beads and incubated with harvested MDBK protein lysate. RNA and protein were incubated for 1 hour to permit protein-RNA binding before washing away unbound proteins. Bound proteins were eluted using laemelli sample buffer and successful isolation of binding proteins detected by western blot and mass spectrometry.

Prior to mass spectrometry analysis, conformation of successful pull-down was checked using positive and negative controls. The 3′ UTR of the androgen receptor mRNA (AR) was used as a positive control and has been shown to bind to the HuR protein, thus, upon successful labelling and pull-down, this protein should be detectable [209]. Western blot analysis of collected flow through, washes and elutions was performed with an anti-HuR antibody, it was observed that HuR was present only in the elution of the positive control, demonstrating successful purification (Fig 4.29). Confirma-

tion of successful RNA-protein pull down was also monitored using RNA representing the entire 5′ UTR, as this RNA fragment included the IRES, it was possible to western blot against known IRES binders such as eIF4A. A positive western blot showing specific elution of a band representing eIF4A could be seen in the 5′ UTR control and not in the control (AR) RNA (Fig 4.29).

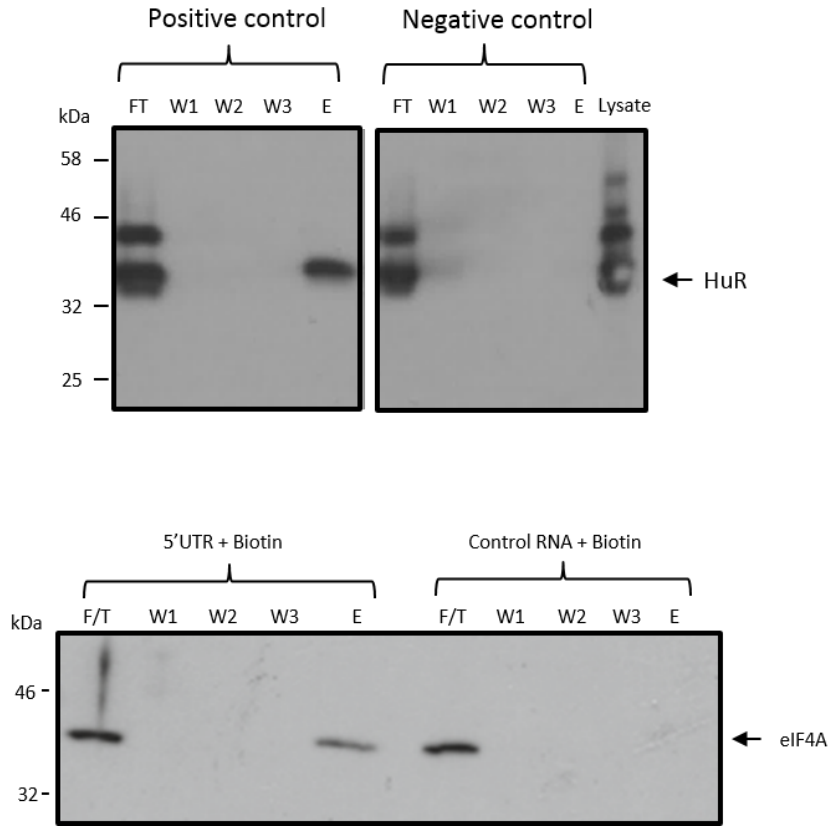

Figure 4.29: Confirmation of successful protein pull down.
(A) Samples from pull downs using Positive (AR) and negative (poly-A) RNA were electrophoresed, transferred and western blotted against the HuR protein. MDBK lysate was prepared alongside as a positive control for HuR identification. (B) Protein pull down samples from experimentation using biotinylated 5′ UTR and control RNA (AR) were electrophoresed, transferred and western blotted using an anti eIF4A antibody. Flow through (FT), washes 1-3 (W1-3) and elution (E) were compared. Size marker is indicated in kDa on the left hand side of each western blot.

Mass spectrometry identified a total of 765 proteins which were annotated against the Bos Taurus protein library. Proteins that interacted specifically with the S-fragment were isolated by comparison of protein iBAQ scores in the different pull-down conditions using the Perseus software. Specific binders were isolated by comparing proteins bound to the AR RNA (non-specific control), UTR ΔS and full UTR, with the theory that the proteins binding specifically to the S-fragment would appear only in the full UTR control (Fig 4.30). These potential binders were then also cross-referenced with the proteins found in the S-fragment only RNA pull down. Using strict criteria, allowing no protein hits in any of the other controls revealed 10 unique binders representing members from a variety of cellular pathways (Table 4.3). As it was possible that the criteria for selecting potential S-fragment binding proteins was too strict, the analysis was performed again, this time allowing inclusion of proteins which also bound to the control RNAs. However, to be carried forward as a potential S-fragment interacting protein, a minimum enrichment in the iBAQ scores of 1 log10 must be observed in the S-fragment containing RNA controls. Re-analysis revealed an additional 10 potential binders seen in table 4.4.

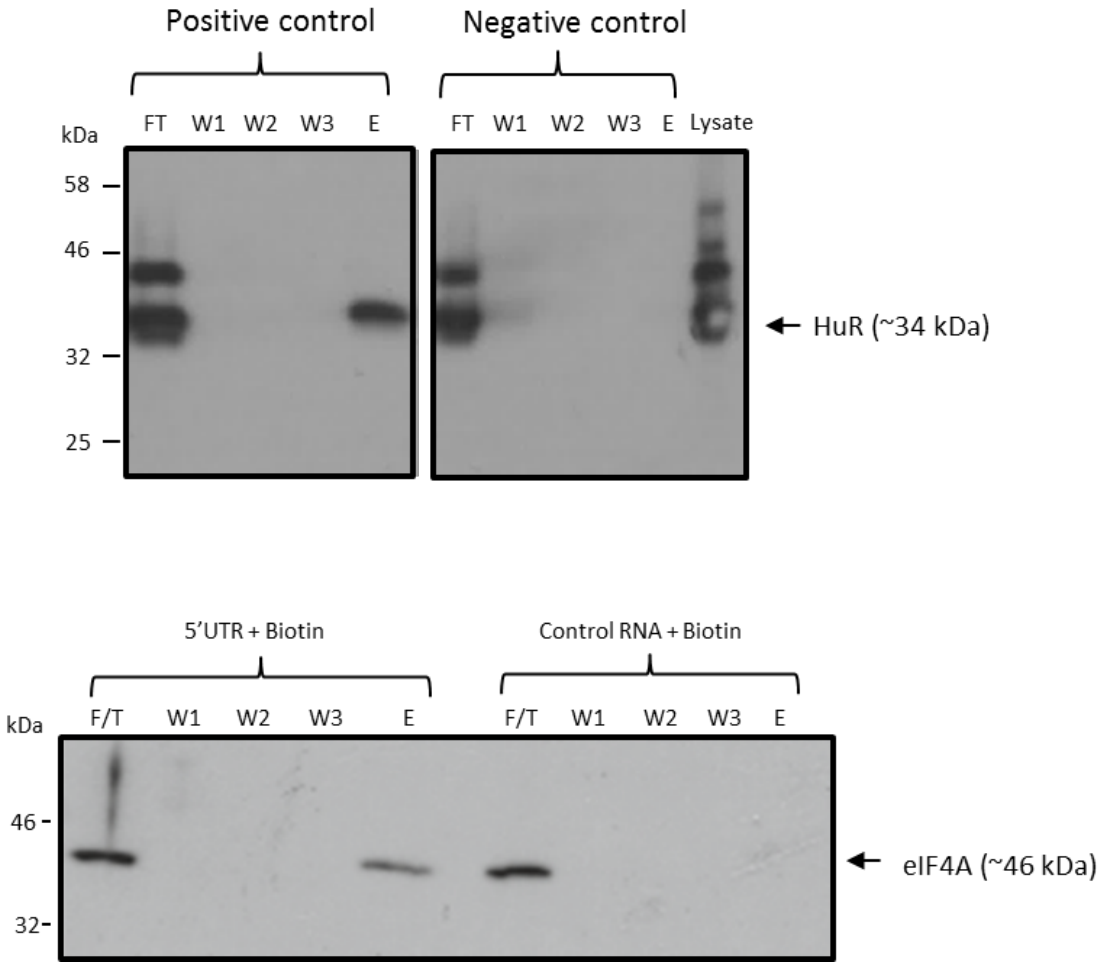

Figure 4.30: Isolation of potential S-fragment binding partners.
Following mass spectrometry analysis of RNA-protein pull downs, protein hits were analysed using Perseus to determine hits specific to the S-fragment. Initially, proteins binding to the non-specific RNA (AR RNA), ΔS UTR and whole $5'$ UTR were compared. Proteins that were found bound to the full size $5'$ UTR but not the ΔS UTR or non-specific RNA controls were isolated and identified (indicated by blue filled region).

Protein	Protein ID	Function
MHC I antigen	Q3YFH7	Immune-response
HSP70 Binding Protein	A7E3S8	Stress-response
Actin related protein 2	F1MRG7	Cytoskeletal
SPOUT domain containing methyltransferase 1	F1MXD0	miRNA binding
26s proteasome non-ATPase regulatory subunit 6	F1MXE4	Enxyme regulator activity
Afadin	F1N2X7	Cell adhesion: adherens junction
Hsp70	K8FK38	Stress-response
TIA-1	Q0VBZ6	Stress granule formation
Nuclear Basket protein	E1BF47	RNA transport
Collagen alpha-1	F1N0K0	Extracellular Matrix

Table 4.3: S-Fragment specific binding proteins.
Identified proteins, specific to the S-fragment, were isolated using the Perseus software. Protein IDs were entered into uniProt and protein identification listed alongside a general function of that protein.

Protein	Protein ID	Function
Poly-A-binding protein	A4IFC3	Poly-A- Binding
DDPG Subunit 1	A3KN04	Glycosyltransferase
Cytoskeleton associated protein 5	E1B7K5	Cytoskeletal
Filamin C	E1BE25	Cytoskeletal
Septin-2	E1BKU2	Cell cycle
AHNAK nucleoprotein	F1MCK2	Diverse, inc. RNA splicing, Calcium channel activity
Asparagine tRNA ligase	F1MKH6	tRNA production
Catenin alpha-1	F1MM34	Cytoskeleton/Cell Adhesion/Motility
EEF1D	L7R5X3	Translation
Peptidylprolyl Isomerase	F1MU79	Isomerase activity

Table 4.4: Less strict conditions isolated addition potential S-Fragment binding proteins.
To permit the discovery of additional proteins, reanalysis of the mass spectrometry data was performed using Perseus software. This time, proteins found bound to either ΔS UTR or non-specific RNA were not excluded, but to be considered a potential S-fragment binding protein enrichment of protein abundance must be enriched by 1 log10. Additional potential binding proteins are listed here alongside a brief indication of their functions.

4.2.8.1 TIA-1

Of these, the major histocompatability complex class I antigen (MHC class I antigen) and cytotoxic granule associated RNA binding protein (TIA-1) were chosen as interesting candidates for further investigation. TIA-1 is a stress granule associated protein involved in protecting host cell translational machinery and mRNA during cell stress caused by events such as nutrient deprivation or viral infection. Here, expression of TIA-1 was observed in BHK-21 cells by immunofluorescence, initial confirmation of antibody signal was visualised by treating cells with (200 μM) sodium arsenate, a known inducer of stress granule formation. Comparison of naïve BHK-21 cells and sodium arsenate treated cells showed an upregulation of distinct TIA-1 punctate in the cytoplasm upon treatment (Fig 4.31). With the antibody verified, cells were transfected with RNA representing either WT, ΔS or 3D-GNN replicons, with TIA-1 expression observed at two, four and 8 hours post-transfection (hpt). As early as two hpt, TIA-1 puncta could be observed in all three conditions, which become more distinct at four and 8 hpt in cells transfected with either the 3D-GNN or ΔS replicons. In contrast, although initial TIA-1 puncta could be observed at 2 hours (Fig 4.32) in cells transfected with the WT replicon, these became more diffused at four hours (Fig 4.33) and are completely dispersed by 8 hours with no observable puncta detected (Fig 4.34).

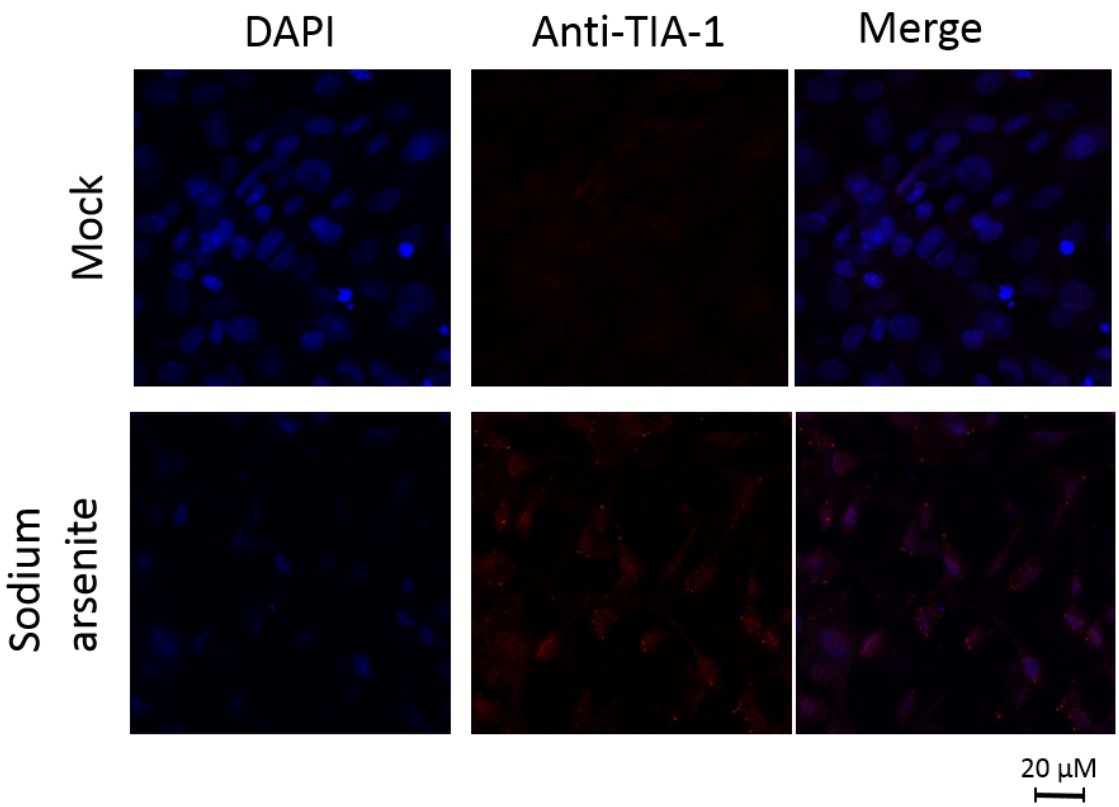

Figure 4.31: Sodium Arsenate treatment of BHK-21 cells.
BHK-21 cells were treated either untreated, or treated with 200 µM sodium arsenate for 2 hours. Cells were then fixed using 4 % PFA and permeabilised overnight in saponin buffer at 4 ^{o}C. Both mock and treated cells were subjected to immunofluorescent analysis using an anti-TIA1 antibody for the detection of stress granule markers. Nuclei were stained using DAPI.

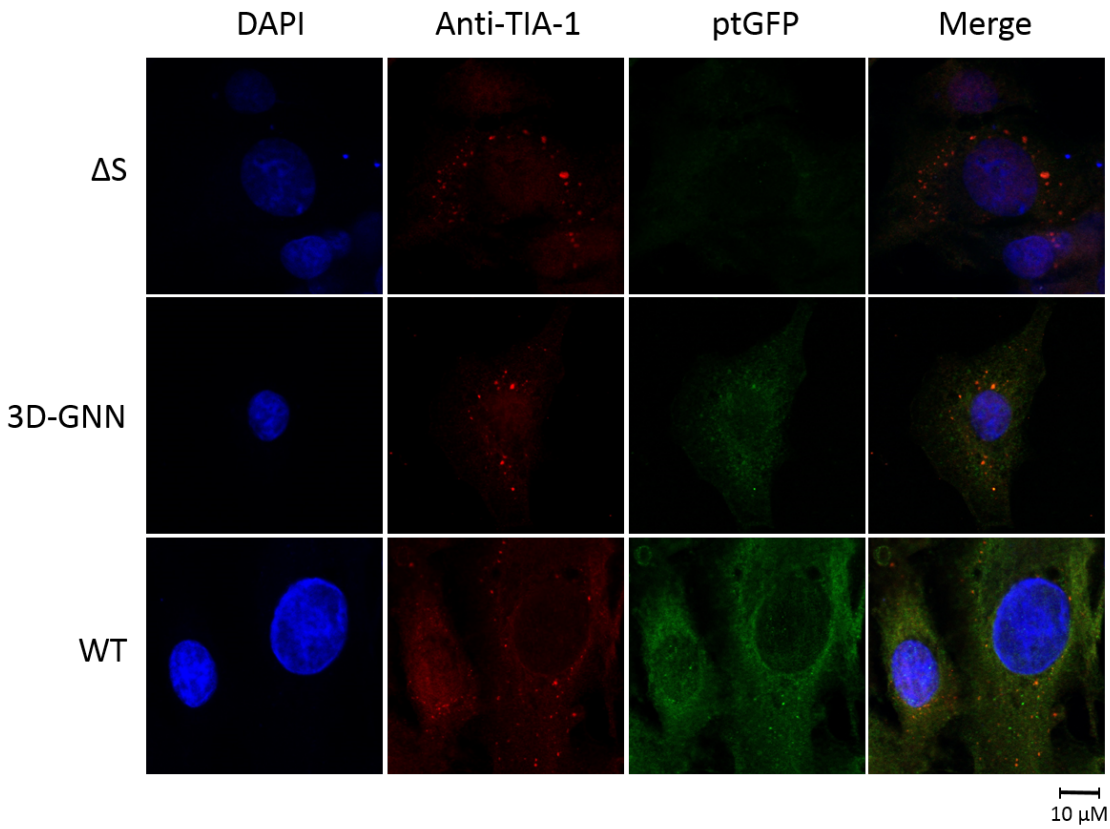

Figure 4.32: Visualisation of TIA-1 in FMDV transfected BHK-21 cells 2 hours post-transfection.
BHK-21 cells were transfected with wither WT, 3D-GNN or ΔS GFP expressing replicons. Cells were fixed using 4 % PFA at two hours post-transfection and permeabilised overnight in saponin buffer. TIA-1 was visualised using an anti-TIA1 antibody. Replicon transfected cells were visualised by GFP expression and nuclei stained using DAPI. Pictures are representative images of 3 experiments.

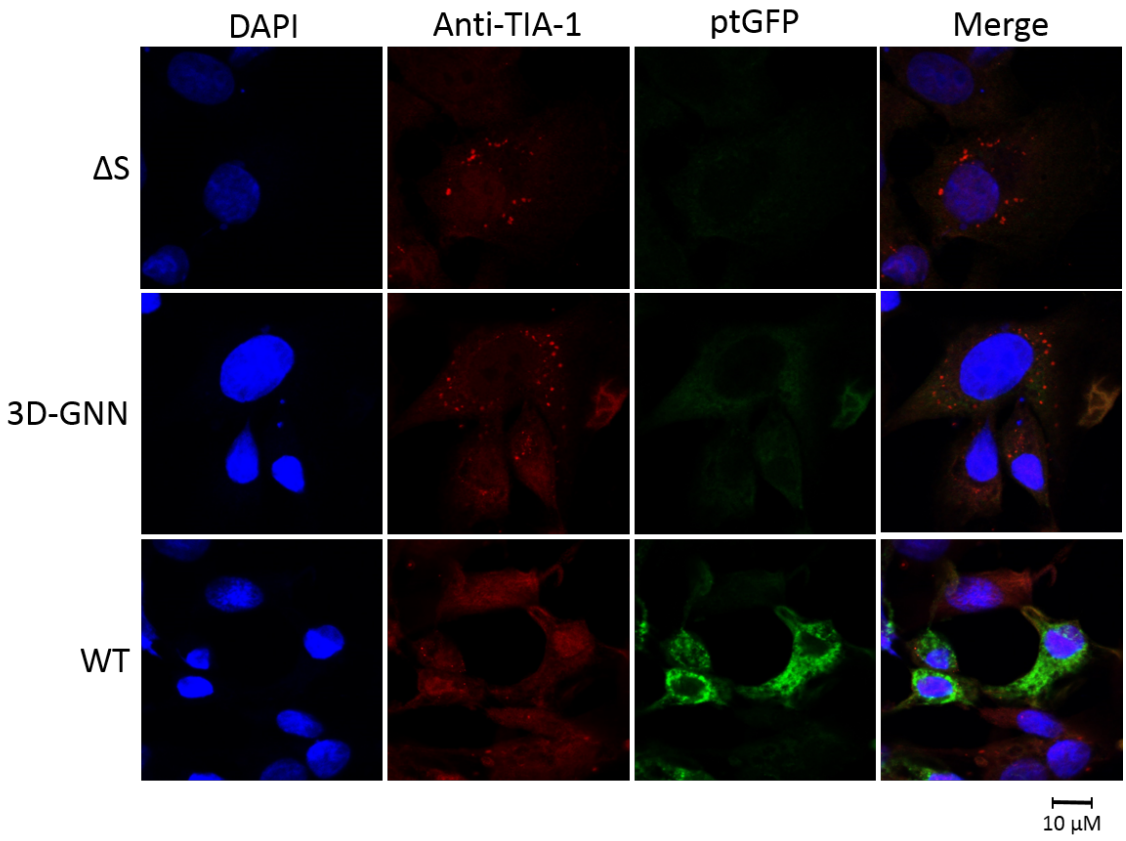

Figure 4.33: Visualisation of TIA-1 in FMDV transfected BHK-21 cells 4 hours post-transfection.
BHK-21 cells were transfected with wither WT, 3D-GNN or ΔS GFP expressing replicons. Cells were fixed using 4 % PFA at four hours post-transfection and permeabilised overnight in saponin buffer. TIA-1 was visualised using an anti-TIA1 antibody. Replicon transfected cells were visualised by GFP expression and nuclei stained using DAPI. Pictures are representative images of 3 experiments.

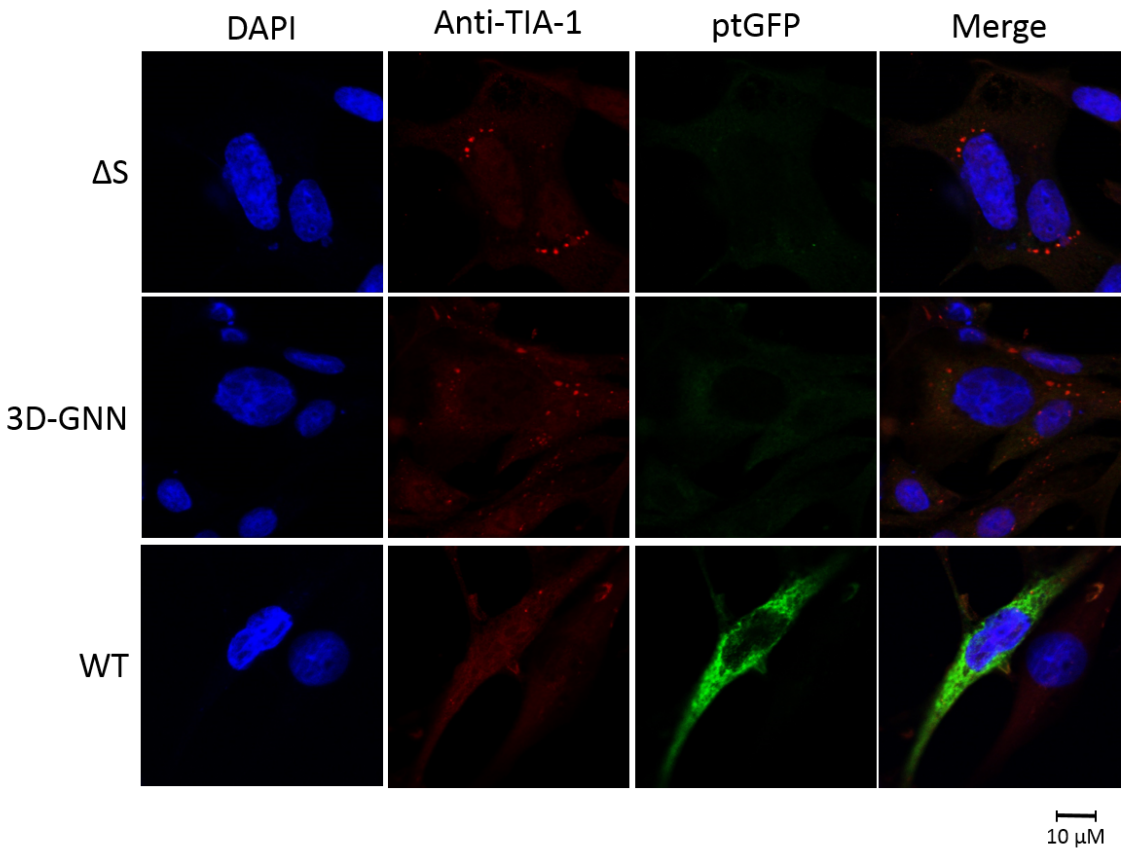

Figure 4.34: Visualisation of TIA-1 in FMDV transfected BHK-21 cells 8 hours post-transfection.
BHK-21 cells were transfected with wither WT, 3D-GNN or ΔS GFP expressing replicons. Cells were fixed using 4 % PFA at 8 hours post-transfection and permeabilised overnight in saponin buffer. TIA-1 was visualised using an anti-TIA1 antibody. Replicon transfected cells were visualised by GFP expression and nuclei stained using DAPI. Pictures are representative images of 3 experiments.

4.2.8.2 MHC Class I antigen

The other protein of interest selected for further study, MHC class I antigen, forms part of the MHC class I pathway, and was the most abundant in mass spectrometry analysis. An antibody against MHC class I (anti-HLA) was used for immunofluorescent visualisation of MHC class I protein during replicon transfection in BHK-21 cells. Initial investigation was performed using WT or 3D-GNN replicon transfected cells visualised at four hpt, where cells could be easily identified (by the replicon reporter), but still appear healthy. Surprisingly, MHC class I signal could not be detected in the mock or 3D-GNN transfected controls, and caused only a general increase in cytoplasmic signal upon transfection with WT replicon (Fig 4.35).

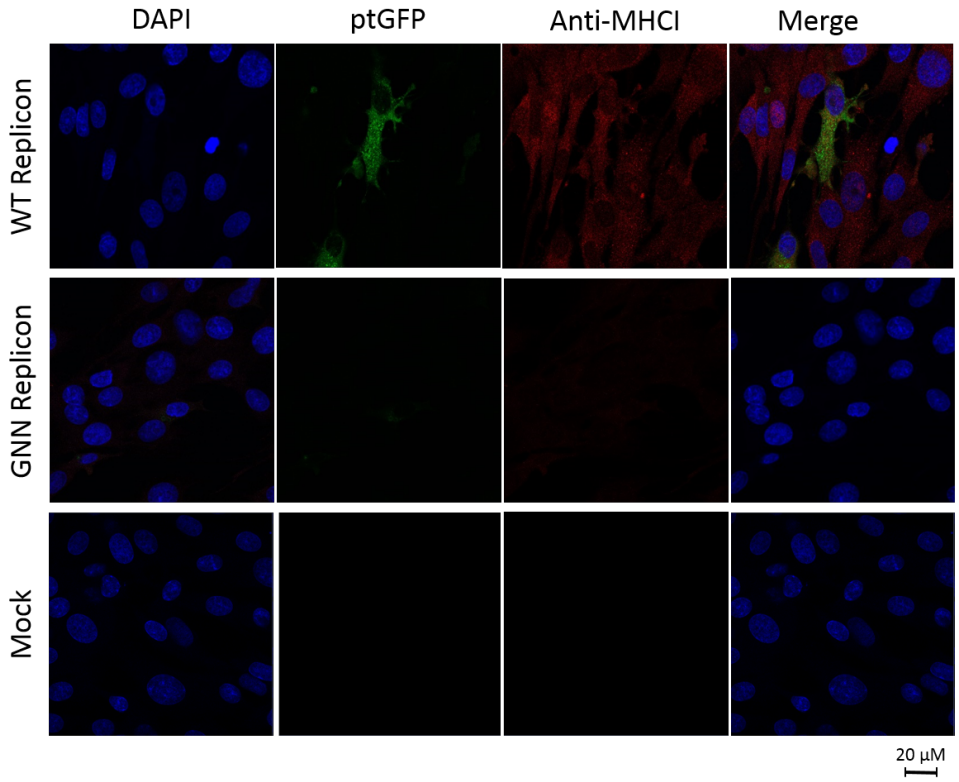

Figure 4.35: Visualisation of HLA in FMDV transfected BHK-21 cells. BHK-21 cells were transfected with wither WT, 3D-GNN or ΔS GFP expressing replicons. Cells were fixed using 4 % PFA four hours post-transfection and permeabilised overnight in saponin buffer. HLA was visualised using an anti-HLA antibody. Replicon transfected cells were visualised by GFP expression and nuclei stained using DAPI. Images are representative of 2 independent experiments.

4.2.9 Potential IRES binding proteins

Due to the number of controls used in the identification of potential S-fragment binding proteins, it was possible to re-analyse the data to identify proteins that could be binding to the IRES. Although IRES interactions are better characterised, it was interesting to see if we observed any novel interactions using the high specification mass spectrometry equipment available at the University of Dundee.

Potential binders were identified by identifying proteins that were enriched in the full UTR RNA pull down and UTR ΔS, but not seen in the UTR ΔIRES or AR RNA control. This produced a list of potential IRES binding proteins, several of which

were known interactors including the eIF4B and PTBP1 (Introduction Fig1.5)(Table 4.5). One protein was considered of interest, HNRNP D, otherwise known as AUF1. Communication with Professor Bert Semler (University of California) revealed that AUF1 is a negative regulator of PV replication, this observation was later published by his research group. Additionally, Prof. Semler also reported that AUF performs a similar role in coxsackievirus B3 infection and could possibly play a role in the lifecycle of other enteroviruses. Professor Semler kindly provided HEK 293-T cell lines expressing either inducible siRNA to AUF1 or a scrambled siRNA as a control.

Protein	Protein ID	Function
hnRNPR	A3KMV6	mRNA processing
CSDA protein	A5D7Q4	Polysome binding
MORF4L2	A6H7C1	Histone deacetylation/transcription
eIF2A	F1MBF0	Translation
PTBP1	A8YXY5	mRNA processing
Unchar	E1B8C6	RNA binding
DAZ associated	E1BAK6	RNA stem-loop binding
ATP binding cassette subfamily F member 1	E1BGH0	ATP binding / ATPase
40s Ribosomal protein S30	Q56JV7	Translation
Unchar	G3N1N7	Unknown
40s Ribosomal protein S26	F1MB60	Translation
U5 SNRNP40	F1MLP7	mRNA splicing
H1 histone family member x	F1MMU4	Nucleosome assembly
DNA topoisomerase I	F1MN93	DNA binding / Chromatin remodeling
Acyl-coA dehydrogenase family member 11	Q0P5G8	Oxidoreductase activity
Fragile X mental retardation 1	F1MXQ7	Negative regulation of translation
eIF4B	Q3MHP6	Translation
Unchar	F1N4K3	mRNA export/ processing
U5 snRNP component	F1N6D5	mRNA splicing
eIF5B	F1N6Y7	Translation
Protein quaking	F1N7N1	mRNA binding
hnRNP A0	F6RSR1	mRNA processing
Histone cluster 1 H1 family member e	G3MWV5	Nucleosome assembly
Glyceraldehyde 3-phosphate dehydrogenase	O77679	Oxidoreductase activity
CD3e molecule associated protein	Q0VD10	RNA polymerase activity
Metadherin	Q24JZ4	dsRNA binding, positive regulation of NF-KB, negative regulation of transcription by RNA pol II
RNA polymerase I	Q3SZ29	RNA polymerase
ErbB3-binding protein	Q3ZBH5	Negative regulation of transcription/ regulation of translation
Aminoacyl tRNA synthetase complex	Q3ZBX5	Translation
GAPDH	Q712W6	Glycolysis
Similar to glyceraldehyde 3-phosphate dehydrogenase	Q862P5	Glycolysis
hnRNPD	A6H6Y0	mRNA processing

Table 4.5: Identification of IRES specific binding proteins.
Identified proteins, that were found bound to the full 5′ UTR but not in the UTR ΔIRES or AR RNA control were isolated using the Perseus software. Protein IDs were entered into uniProt and protein identification listed alongside a general function of that protein.

4.2.10 AUF1 in FMDV replication

Upon transfection of HEK 293-T AUF1 knockdown or control cells with either WT or 3D-GNN PV replicons, replication, in the form of GFP expression, was observed. As the comparison was looking at two different cells lines (knock down and control), replication is shown here as the average mean GFP intensity (GCU) to provide a more suitable comparison than counting GFP positive cells. Consistent with the data observed by Semler et al, an increase in PV replication could be seen upon knock down of AUF1 (Fig 4.36A).

To determine whether a similar enhancement of replication could be observed in FMDV replicon replication, the experiment was repeated by transfection of WT or 3D-GNN replicon RNA into HEK 293-T AUF1 knock down or control cells. A slight increase in WT replicon replication upon AUF1 knock down was observed, (Fig 4.36B) with no difference seen when transfecting 3D-GNN replicons (Fig 4.36C). Although, FMDV replication was seen to be poor in HEK 293-T cells, and as such the difference between knock down and scrambled controls was small.

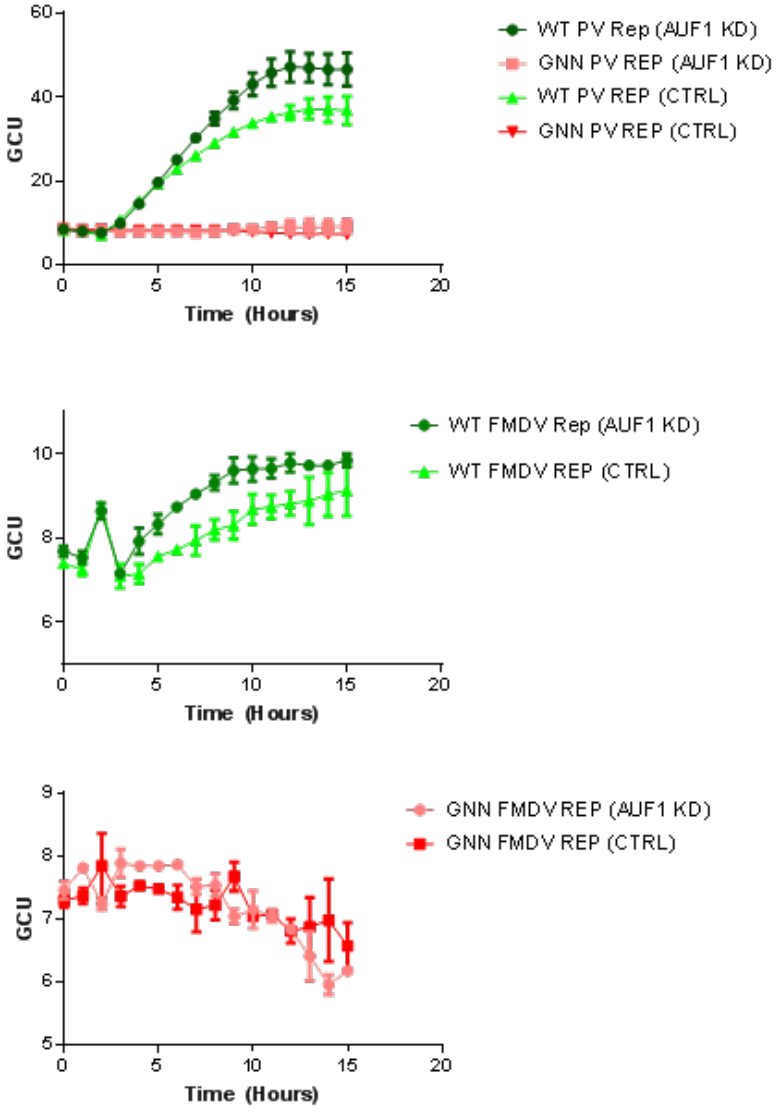

Figure 4.36: Replication of FMDV and PV replicons in HEK 293T AUF1 knock down, or control siRNA cells.
(A) PV WT and GNN replicons were transfected into HEK 293T cells expressing either an siRNA to AUF1 or a control scrambled siRNA. Comparison of GFP expression was monitored over 16 hours and shown here as the mean GFP intensity (GCU). FMDV WT (B) and GNN (C) replicons were transfected into similarly prepared HEK 293T siRNA expressing cells and replication monitored by GFP expression over 16 hours. n=2, error bars represent SEM.

4.3 Discussion

4.3.1 Improved S-fragment structure prediction

Previously, knowledge of S-fragment secondary structures has been largely provided by computational mFold approaches. While these can offer useful insights, they are not entirely reliable as in silico methods cannot account for physiological factors that may affect RNA structure. Here, SHAPE analysis was used to further enhance and confirm previous predicted structures. Overall, the SHAPE data shown here mostly corroborates with previous predictions. However, there appears to be some differences in the large stem-loop with bulges of differing sizes appearing in different locations. It is important to note that the ΔG of the S-fragment is low, suggesting high stability, therefore adding more confidence to our predicted structures. While this is true for the structure as a whole, it is possible that the bulges found throughout the length of the stem may be more flexible and indeed some SHAPE reactivity suggests a certain amount of breathing at the ends of bulges, represented by reactivity at base-paired nucleotides found proximal to bulges. Although the use of SHAPE has increased our knowledge of the S-fragment structure, further probing performed in a physiologically relevant environment, such as with-in a cell, would improve confidence in its true structure as potential influencing factors such as RNA-RNA interactions and RNA-protein interactions would be maintained.

4.3.2 S-fragment chimeras

The generation of the Aerie replicon has allowed for easy manipulation of the FMDV S-fragment to generate mutants. Using this system, we have confirmed previous findings that the S-fragment is essential for replication. Furthermore, we showed that the reverse complement of the sequence, was not capable of providing the same function.

We also demonstrate that whilst sequence similarity may differ between S-fragments of different isolates, the function it provides to the lifecycle remains the same and is interchangeable between different FMDV viruses. This is evident by the lack of change in replication when the S-fragment of the O1K replicon was replaced by the SAT2 virus S-fragment, suggesting that unless there are hidden motifs in the small areas of conservation, the S-fragment sequence is not critical for its functions.

Notably, although S-fragments from different FMDV isolates can be exchanged, it is not possible to replace this with RNA structures found in similar positions from other picornaviruses. Whilst these structures are predicted to have similar roles, (e.g in genome circularisation) replacement of the S-fragment with these structures always resulted in non-replicative constructs. It is likely that these structures were not compatible due to differences in viral or host cell protein interactions with the structures, thus preventing their roles being effectively carried out.

4.3.3 Truncations

Following the identification of FMDV isolates with truncations to their S-fragments, deletion mutants that increased in size were generated targeting both the top and bottom of the stem-loop structure. The data showed that removing over half of the S-fragment from either the top or bottom still permitted replication, whereas replication is prevented in the replicons containing larger deletions. SHAPE probing confirmed that these truncations maintained the stem loop structure of the RNA. These findings are interesting as the overlapping deletions in T-148 and B-195 suggest that there is no requirement for nucleotide specificity and that RNA structure alone may be the sole determinant for function. However, this conclusion is somewhat contradicted by the fact that B-195 is the only mutant that behaves differently when introduced into the virus system, where no virus could be recovered. This suggests that B-195 contains a sequence essential for virus lifecycle stages not represented in the replicon system, such as aspects virus assembly. Alternatively, the levels of GFP

expression observed for this construct in the replicon system could be produced as a result of enhanced translation rather than replication. Further investigation of this hypothesis could exploit translation reporter systems to observe changes in IRES translation with different sized S-fragments.

It is surprising that we observed disparities in S-fragment mutant replication between MDBK and BHKI-21 cells, which could be explained by the difference in immunocompetence between the two cell lines. Indeed in 2017, Reider et al, published that S-fragment truncations targeting the top of the structure caused an increase in immune activation in PK(15) cells, a porcine kidney cell line. This suggests that the length of the S-fragment is important for evasion of host immune responses, with longer S-fragments capable of avoiding immune detection whilst smaller ones induce higher levels of IFNβ and interferon stimulated genes[95]. However, information in this area is still unclear as different size truncations seemed to have different effects on immune activation in a non-linear fashion. It is possible that the observed effect where a smaller S-fragment causes increased host immune response may not be directly due to the role of the S-fragment in immune evasion, but rather having an altered version of the S-fragment could have downstream effects on other aspects of the virus lifecycle, possibly eliciting an immune response. To better understand this aspect of FMDV virology, experiments using would need to be performed including investigation into dominant negatives – where co-transfection of one replicon with a shortened S-fragment might lead to decrease in replication of a WT.

It is evident that more work in this area needs to be carried out, but it is promising that a potential function may have been uncovered.

4.3.4 Virus recovery of S-fragment truncated mutants

Several differences were noted when attempting to recover infectious virus harbouring S-fragment truncations when compared to the replicon experiments. Firstly,

T-148, which replicated well in BHK-21 cells in replicon, only produced infectious virus once out of 6 attempts. Despite this inconsistency, it could be argued that since infectious virus was recovered with no change to the input sequence, that this virus is capable of replication. It is still however unknown why it took so many attempts to produce infectious virus. Similarly, the T-246 virus recovery was successful in four of the 6 attempts, in contrast to the replicon data which showed no replication in either BHK-21 or MDBK cell lines. This disparity between the virus and replicon data pointed to the possibility of compensatory mutations elsewhere in the genome of the T-246 recovered virus, which permitted replication despite such a large S-fragment truncation. However, upon sequencing, no change in input sequence was reported.

It is possible that the T-246 replicon was replicating but at a very low level, undetectable by the measure of GFP fluorescence. It is possible that by introduction of this mutation into infectious clone replication was visible due to viral amplification by the ability for virus to cause cell-cell spread.

Whilst T-246 showed no replication in replicon, but produced infectious virus, B-195 showed the opposite phenotype. Although reduced, the B-195 replicon showed GFP expression approximately 10 fold higher than that of the 3D-GNN control when transfected into BHK-21 cells. However, when introduced into the infectious clone, this mutation was unable to produce infectious virus. It should be noted that recovery of this mutation was repeated 3 times with no success, whilst T-148 was repeated 6 times to produce one successful recovery of infectious virus. It could be that upon repetition B-195 might produce infectious virus.

Most of the other truncated S-fragment viruses were recovered with no change in sequence, however mutant B-97 was observed to display CPE at a faster rate than the smaller truncation B-48, during virus recovery. All three replicates recovering the B-97 virus showed an increase in CPE speed by passage 5, exceeding that of the B-48 virus. However, only two of the replicates showed mutations elsewhere in the genome,

whilst the third showed the same sequence as the input RNA. How or why the speed of CPE increased here remains unknown. The other two replicates showed two different sets of mutations, one revealed the I189L mutation in the polymerase, whilst the other revealed a collection of synonymous and non-synonymous 2C mutations. The transfection which led to the observation of mutations within the polymerase resulted in higher rates of CPE when compared to B-48 as soon as CPE could be detected (passage 3), enrichment of the I189L mutation was also observed as the virus was passaged suggesting a preference to keep this mutation within the population. Interestingly, upon investigation into this residue, comparisons of 103 FMDV field isolates revealed that the isoleucine found here is 100 % conserved (personal communication, Lidia Laseka-Dyes, the Pirbright Institute). Previously published structural studies investigating the interaction between $3D^{pol}$ and RNA has revealed that this residue is thought to interact with the phosphate backbone of RNA[57]. When mapped onto a crystal structure of the $3D^{pol}$ complexed with RNA, the mutant residue shows a subtle change compared to the WT structure, however, proximity to the RNA backbone supports a potential interaction. Together this suggests that this mutation might alter RNA binding in a way that allows for improved replication when combined with truncations to the bottom of the S-fragment.

The mutations discovered in the 2C protein of replicate 1 (B-97 (1)) were reported in the fourth passage. At this time only virus from this passage has been sequenced, meaning no data exists on when and how these mutations appeared. Although all three were present at 70 % in the population, it is possible they arose independently, further sequencing is ongoing to investigate this further. It is interesting that this attempt at virus recovery showed the slowest production of CPE, being slower than B-97 (2) and (3), as well as B-48. Rate of CPE remained slow until passage 5 when the speed increased to faster than that of the B-48 virus, akin to B-97 (1) and (2) recovered viruses.

It is possible that the introduction of 2C mutations slowed the replicative rate of

B-97 (1), and once all three mutations were stabilised within the population, replication increased as seen by the increased rate of CPE. Little is known about the FMDV 2C protein and as such, prescribing mechanisms to these mutations is difficult. It is potentially not surprising that the mutation of the first amino acid is synonymous as alteration of the amino acid sequence could affect $3C^{pro}$ cleavage of the polyprotein, leading to profound effects on viral replication. The location of the other two mutations does not reveal much with regards to potential function, when comparing the locations to PV 2C, the S318R mutation does map to the RNA binding region. Although the structure of FMDV 2C is unknown, comparison with the PV 2C annotation would support the idea that an alteration in RNA binding permits fast replication with the B-97 truncation to the S-fragment.

Neither 3D or 2C mutations have yet been re-introduced into replicon, doing so could provide insights into their mechanism and will be targeted as future experiments.

4.3.5 Dissecting the I189L $3D^{pol}$ mutation

Sym/sub experiments were used to explore the functionality of $3D^{pol}$ when containing the I189L mutation. Experiments looking at the addition of +1 and +4 nucleotides revealed little difference in the kinetics of polymerase elongation when comparing WT and I189L polymerases. However, the production of a 'super product' was unexpectedly observed in both WT and I189L conditions, interestingly more so when the I189L polymerase was used.

Further investigation into the differences in 'super product' formation is currently ongoing, with the hypothesis that this product is produced as a result of increased strand switch recombination, allowing the polymerase to switch template and perform further elongation. This hypothesis relies on the misincorporation of a GTP nucleotide at the end of the initial sym/sub template. Experiments were nucleotides were added in sequential fashion revealed the possibility that when A, C and G nu-

cleotides where added to the reaction that a +4 band could be seen, supporting the idea of misincorporation at this residue.

Validity is added to this hypothesis as the only time a larger product has been noted in sym/sub experiments is in the investigation of poliovirus 3D polymerase recombination[213]. Here the super product was seen, but only upon addition of another RNA which could then be used as the product for further elongation. It is possible that the error rate of FMDV 3D polymerases is higher, allowing the misincorporation to occur and thereby creating the possibility of recombination by strand switching. To confirm this hypothesis, it would be important to uncover the sequence of the super product, as well as performing additional experiments investigating recombination rates of the I189L containing viruses in a cellular environment. Further controls to the sym/sub experiment could also be included in the form of a high-fidelity recombinant 3D polymerases which would be less likely to cause GTP misincorporation to create a new template.

4.3.6 Exploration of 2C mutations

Exploration into the role of the S-fragment mutations discovered in the 2C protein is ongoing. Little is known about the specific activity of FMDV 2C, as such attempting to predict a function of the mutant is difficult. Using experimental evidence on the roles of poliovirus 2C as a template for potential FMDV 2C function, it is possible to suggest that one of the mutations, S318R, could lie in an RNA binding domain of 2C[170]. This would corroborate with the I189L mutation which is thought to interact with the RNA phosphate backbone, suggesting these mutations arose to compensate for the B-97 mutation by allowing alterations to RNA binding. RNA binding studies using both the I189L mutant 3D and a recombinantly produced 2C (carrying the mutations observed) would be an interesting way to investigate whether RNA binding is altered.

4.3.7 Transcomplementation

With the recent investigation into elements of the FMDV genome that could be complemented by another construct, it was decided to see if the S-fragment, like many other viral functions, could be provided in trans to a deficient replicon. Initial experiments co-transfecting a ΔS replicon with either WT, cre A1G or 3D-GNN replicon showed that under certain conditions, recovery of the ΔS replicon could be observed. However, subsequent passaging experiments suggested that this recovery was as a result of recombination rather than complementation. Surprisingly, recombination could occur even when cells were co-transfected with non-replicon RNA such as RNA representing only the 5′UTR. Furthermore, an increase in replication could be seen even when using the S-fragment followed by lengths of scrambled sequence.

It would be interesting to determine where recombination took place, as there are no conserved sequences between the 918 and 783 scrambled UTRs and the ΔS replicon. This could reveal an area susceptible to insertions and a site that presumably does not disrupt other essential structures in the UTR such as the IRES and the cre. This site would likely lie 5′ to the IRES and cre structures, lying in the areas shown to be non-essential for replicon replication such as the pseudoknots and poly-C-tract (discussed in Chapter 5.1).

Interestingly, RNA length was also shown to be important in the recovery of the ΔS replicon. S-fragments followed by either 918 or 783 nucleotides of scrambled sequence showed that longer UTRs provide a more efficient recovery, it would be interesting to continue this line of experimentation to determine the minimal size of the RNA required for function as well as whether recurrent recombination sites occur.

4.3.8 Potential S-fragment binding proteins

RNA-protein pull down experiments revealed several potential S-fragment protein interactors. Interestingly the previously reported RHA was not isolated as a potential binder in our assay. Upon close examination of the raw data this is due to RHA binding with high efficiency to both IRES and the AR 3′ UTR control and it was therefore removed from the list of potential S-fragment interacting proteins. However, this does not rule out RHA as a potential S-fragment interacting protein and is best explained by the ability of this protein to bind to other structured RNA elements, which as an RNA binding protein might not be surprising. Using the AR 3′ UTR RNA as a non-specific RNA control may therefore not be the best choice, partially due to the RNA-binding properties of these proteins which means they may bind non-specifically to RNA, but also due to the fact that the AR RNA is a 3′ UTR. As an untranslated region, likely containing structured RNA elements, it is possible that several of these proteins may bind to this region organically and therefore get excluded from the potential interacting proteins.

Our results highlight the possibility that multiple other S-fragment binders are "lost" in the analysis due to possible binding elsewhere. For example, during the protein analysis, any proteins that also showed overlapping binding with the ΔS IRES were excluded from the list of potential binders. Whilst this approach would be appropriate for identifying specific interactive proteins for the S-fragment, it eliminates any proteins that could be binding to both the S-fragment and IRES. Binding to multiple RNA structures in the 5′ UTR is a characteristic seen with many proteins during PV infection, for example, PCBP2 is shown to bind to both IRES and the 5′ CL. Further analysis of the mass spectrometry data could reveal proteins that bind both the IRES and S-fragment, providing insights into the switch between translation and genome replication.

Changes in TIA-1 expression over time when BHK-21 cells are transfected with WT replicon, but not with ΔS or 3D-GNN replicons reveals a potential interac-

tion with stress granules as previously highlighted with PV[207]. Furthermore, novel data presented by the von Kupperfeld research group at the European picornavirus meeting (EuroPic) also revealed a role of G3BP3, another stress granule marker, in FMDV replication, further supporting a distinct role for stress granules in FMDV replication. Due to the observed dispersal after initial formation, it is likely that stress-granules form in response to FMDV transfection or infection which is then dispersed by viral proteins during replication. Prevention of stress-granule formation inhibits sequestration of translational machinery away from viral replication sites. This was not observed upon transfection of replication incompetent replicons, potentially due to limitations of input translation in producing enough viral proteins to adequately disperse stress granules. Co-localisation studies of viral proteins with TIA-1 and other stress granule markers would be interesting to investigate if any viral proteins interact with stress granules. The use of a mechanism to detect replicon RNA, such as fluorescent in situ hybridisation (FISH), could also provide evidence supporting the RNA-Protein interaction between the FMDV genome (specifically the S-fragment) and TIA-1. How the S-fragment plays a role in the phenotype of TIA-1 is unknown, but it is possible that it acts as another potential binding partner for TIA-1 to prevent downstream interactions for stress granule formation.

The observation of increased MHC class I antigen associating with the S-fragment was intriguing, especially due to recent publications suggesting the S-fragment might play a role in immune modulation of the host cell. A similar observation has been reported recently for murine norovirus (MNV), where infection upregulated MHC class I expression in a cell-dependent manner[221]. Immunofluorescent visualisation of MHC class I, however, was not as convincing as hoped. While it is interesting that an increase in fluorescent signal was observed upon transfection of only WT replicons, the localisation to the cytoplasm is inconsistent to normal MHC class I localisation at the plasma membrane. A possible explanation is that replication of FMDV causes an upregulation of MHC I expression but subsequent localisation to the plasma membrane is inhibited as a mechanism of immune escape. Alternatively,

the natural localisation of this protein in BHK-21 cells could be defective, as is true for other aspects of the BHK-21 immune response.

Interestingly, the increase in MHC I signal was observed in all cells, not only in those showing signs of FMDV replication. One explanation of this is that transfected cells are providing immune signalling in an endocrine fashion, causing an upregulation in neighbouring cells in preparation of an infection. However, it cannot be ignored that the signal seen here could be non-specific and additional experimentation is required for confirmation.

Other proteins identified by mass spectrometry would also be interesting to pursue, the presence of both HSP70 and HSP70 binding protein provides increased confidence that these may be legitimate binding partners, with HSP70 binding to the S-fragment, and HSP70 binding protein binding to HSP70 in a "daisy-chain like" fashion. The role of HSP70 in stress response also correlates with the presence of TIA-1, suggesting the S-fragment might play a role in modulating stress responses within the cell. The identification of several cytoskeletal and cell adhesion proteins suggests a potential role of the S-fragment in reorganisation of the cellular environment. Electron microscopy experiments show that following transfection, the cell undergoes massive membrane rearrangement (Section 3.2.9) and the S-fragment could be involved in this as seen in many other viruses such as poxviruses, HIV and herpesviruses, which all manipulate the cytoskeleton to permit viral replication.

4.3.9 Potential IRES binding proteins

Re-analysis of the mass spectrometry data allowed for the isolation of potential IRES-specific binding proteins. This list comprises many expected proteins such as elongation initiation factors. However, there were proteins notably absent from this list such as eIF4A, despite this being used as a positive control by western blot to confirm successful a successful binding and pull down, it did not appear in the list of

identified proteins at all. This could be due to incomplete annotation in the MDBK (bos Taurus) library, or misidentification of fragments of this protein.

Despite stringent selection criteria, several novel proteins were identified, including a number of hnRNPs which are novel for FMDV, but have been previously reported for multiple other picornaviruses. The presence of HNRNP D (otherwise known as AUF1) was exciting, especially due to the recent work by Semler et al showing an increase in PV replication upon AUF1 knock down in HEK 293-T cells[198]. HEK 239-T AUF1 knockdown cells were provided by Professor Bert Semler and while these showed an increase in PV replicon replication as expected, the poor replication of FMDV in this cell line made it difficult to determine if there was any effect on the knock down of AUF1 on FMDV replication. Work is currently on going to knock down AUF1 in FMDV permissive cell lines such as BHK-21 and MDBKs. If a similar phenotype to PV is observed, these knock down cell lines could be an exciting advancement in FMDV tissue culture methodology and could even be used to replace traditional BHK-21 cells as the vaccine producer cell line to obtain higher virus yields.

Chapter 5

Untangling the roles of the FMDV pseudoknots

5.1 Introduction

As discussed in previous chapters, the 5′UTR of FMDV contains many highly ordered structural elements. In this chapter, the role of one of the lesser studied elements of the 5′ UTR, the FMDV pseudoknots is investigated.

The FMDV pseudoknots are each made of around 48 nucleotides which are tandemly repeated to make up a total of two to four[38]. They lie directly 3′ to the poly-C-tract, a stretch of mostly cytosine residues which can vary in length from 50 to 200 depending on virus type and isolate, although actual quantification of this is incredibly difficult due to difficulties in sequencing large repeats of the same nucleotide (Fig 5.1)[127,76].

The first pseudoknot model for FMDV was predicted by computational sequence analysis by Clarke et al in 1987, apart from this initial observation, very little investigation at the molecular level has been carried out on this region. The 1987 prediction suggests that within each pseudoknot section, a small stem-loop forms with a four-nucleotide high stem and a five to six nucleotide containing bulge. Predictions suggest that three or four nucleotides present in the bulge may interact with nucleotides downstream creating an H-type pseudoknot interaction (Fig 5.2)[38].

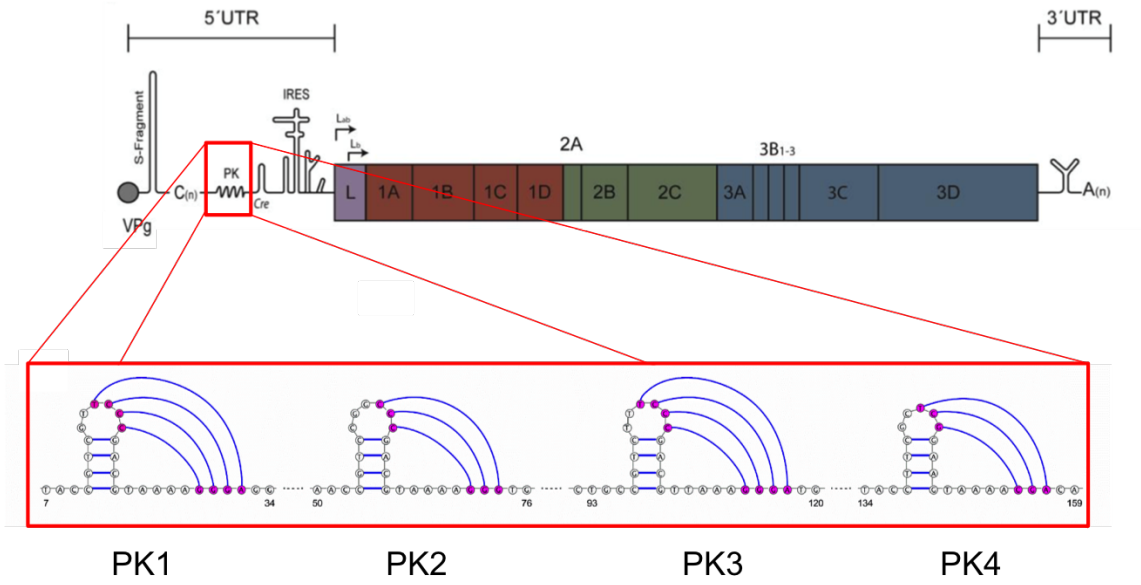

Figure 5.1: $5'$ UTR and pseudoknot schematic.
A cartoon representation of the FMDV O1K genome. The pseudoknot region, is shown enlarged with the predicted pseudoknot interactions highlighted in hot-pink. Nucleic acid number is relative to the nucleotide position after the poly-C-tract.

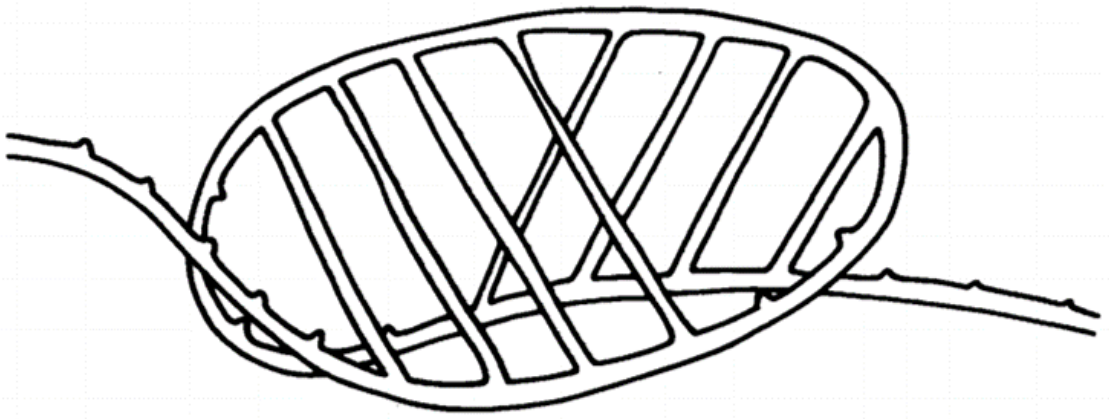

Figure 5.2: Artists rendition of an FMDV pseudoknot.
A 3D representation of how the pseudoknot structures may fold. Figure adapted with permission from Clarke et al 1987.

No isolate of FMDV has yet to be recorded to have less than two pseudoknots, however, variations between two and four have been described. Analysis of over 103 novel sequenced genomes by our collaborator, Lidia Laseka-Dykes, at the Pirbight institute also reveals no serotype bias for a distinct number of pseudoknots. Due to the sequence similarity between the pseudoknots, it is hard to distinguish them apart, as such it cannot be concluded that there is a preference for specific pseudoknots in particular sero or sub types. However, due to the sequence similarity, it is likely that a preference is not seen. Analysis of the novel sequences also revealed the interesting observation that deletions of individual pseudoknots are almost solely seen in the context of removal of the entire 48 nucleotide region with sightings of partial deletions not reported (Personal communication Lidia Laseka-Dykes, *the Pirbright Institute*).

For the experiments described in this thesis, a replicon derived from the FMDV O1K strain, which has four pseudoknots, was used. Comparison of the nucleotide sequence of the pseudoknots from the O1K genome shows a high level of sequence similarity, with PK 4 being the most diverse (Fig 5.3). This could suggest that the repetition of the pseudoknots may have arisen during duplicative events over the evolution of FMDV, potentially through mechanisms such as recombination, followed by error-prone replication.

```
PK1  AAGTTTTACCGTCGTTCCCGACGTAAAAGGGAGGTAACCACAAG---
PK2  -CTTGAAACCGTCCGGCCCGACGTAAAAGGGTGGTAACCACAAG---
PK3  -CTTACTGCCGTCTTTCCCGACGTTAAAGGGATGAAACCACAAG---
PK4  ---ACTTACC-TTCGCTCGGAAGTAAAACGACAAACACACACAGTTT
             ** *     * ** ** *** *      **    **
```

Figure 5.3: Sequence comparison of pseudoknots in the O1K $5'$ UTR. Sequence alignment of the four pseudoknots found in the O1K $5'$ UTR. Stars represent invariant nucleotides and nucleotides predicted to interact for the formation of the pseudoknot are highlighted in hot pink, as in Fig 5.1.

Pseudoknots have been known to play a variety of roles throughout the field of virology. In particular, they have been reported to have a role in protection against digestion by RNase in flaviviruses such as dengue virus[33]. In dengue virus, a pseudoknot structure found in the 3′ UTR has shown resistance to XRN-1 degradation, allowing the production of a small non-coding piece of RNA (sfRNA). Production of this sfRNA is related to an increase in pathogenicity, although the mechanisms behind this are not yet elucidated[92]. A role for pseudoknots in splicing is seen in influenza, whereby a conformational switch from a hairpin to a pseudoknot makes nucleotides necessary for splicing less accessible[72]. The formation of pseudoknots therefore helps modulate the splicing and release of mRNA essential to produce the M2 ion-channel protein[136,159]. Genome analysis of a murine coronavirus revealed a pseudoknot used for ribosomal frameshifting, allowing the switch between translation of ORF1a and ORF1b. FMDV is not the only picornavirus to have pseudoknots described in the genome. Presence of PKs have been reported in the 5′ UTR of EMCV and ERAV, however, in both cases the pseudoknots are located at the 5′ end of the poly-C-tract, making their position in the FMDV genome unique. As for FMDV, a potential function of the pseudoknots found in EMCV and ERAV has not yet been described.

Pseudoknots have also been reported to have multiple functions in cellular RNAs and are present in a number of different organisms, most of these act as riboswitches in regulatory systems, controlling protein expression depending on the environment of the cell[184]. For example, pseudoknots are used by some bacterial species to prevent fluoride toxicity by acting as a fluoride ligand binding receptor. When fluoride levels rise within the cell they can bind to the pseudoknot and encourage the expression of fluoride transporters, thereby preventing the build-up of toxic products within the cell[25].

Cellular pseudoknots have also been reported to function as ribozymes, catalysing RNA scission and ligation in several processes such as RNA splicing, tRNA biosynthesis and mRNA cleavage. An example of this includes the glmS ribozyme, which

acts in a feedback loop system to prevent the over-production of GlcN6P sugar[56]. This pseudoknot-ribozyme is located in the 5′UTR of the glmS mRNA, encoding an enzyme necessary for the synthesis of the sugar, GlcN6P. When the cellular concentration of GlcN6P is high, the ribozyme can bind GlcN6P molecules, causing the cleavage of the ribozyme off the mRNA leaving a 5′ hydroxyl on the end of the newly cleaved mRNA[124]. This 5′ hydroxyl can be recognised by RNase J and target the mRNA for downstream degradation and reduce GlcN6P production. Other cellular ribozymes have been reported which resemble that of the double-pseudoknotted hepatitis D ribozyme. Upon folding, the HDV ribozyme can self-cleave away from the rest of the RNA, indeed the HDV ribozyme is so characterised that it is utilised in research during the creation of expression constructs where the precise sequence at 5′ end of the RNA is essential[12]. The HDV ribozyme is used in the FMD replicon system described in this thesis, as it has been shown that picornaviruses with the authentic 5′ end perform better than those with additional nucleotides that can occur during T7 polymerase transcription[80].

5.2 Results

5.2.1 SHAPE analysis of the pseudoknot region

It is important to try and confirm the predicted structure of the FMDV pseudoknots. Computational structural prediction of pseudoknots remains challenging, with most prediction occurring via manual sequence analysis. Here we used chemical mapping overlaid onto the 1987 predicted pseudoknot structures to evaluate if the existence of these structures could be confirmed or disproved.

Selective-2'hydroxyl acetylation analysed by primer extension (SHAPE), uses chemical modification of available nucleotides, to form 2' adducts. As described previously (Section 4.2.1), modified RNA is used in reverse transcription reactions to look for

early termination products caused by reverse transcriptase termination at sites of modification.

SHAPE analysis was performed as described in materials and methods 2.7, with replacement of the S-fragment specific primer with a fluorescent primer located further downstream to allow the analysis of the PK region.

Visualisation of the "raw" data, comparing NMIA and DMSO early termination events reveal that, unlike, in the S-fragment reactions, few DMSO terminations are seen. Suggesting that, when unmodified, this region is not highly structured, allowing the reverse transcriptase to read through easily. In contrast to the S-fragment reactions, the whole PK region was more reactive when compared to that of the S-fragment, suggesting a less defined or more flexible structure. This might be expected from the previous pseudoknot predictions as the hairpins themselves are relatively small, with large regions of predicted single stranded nucleotides (Fig 5.4).

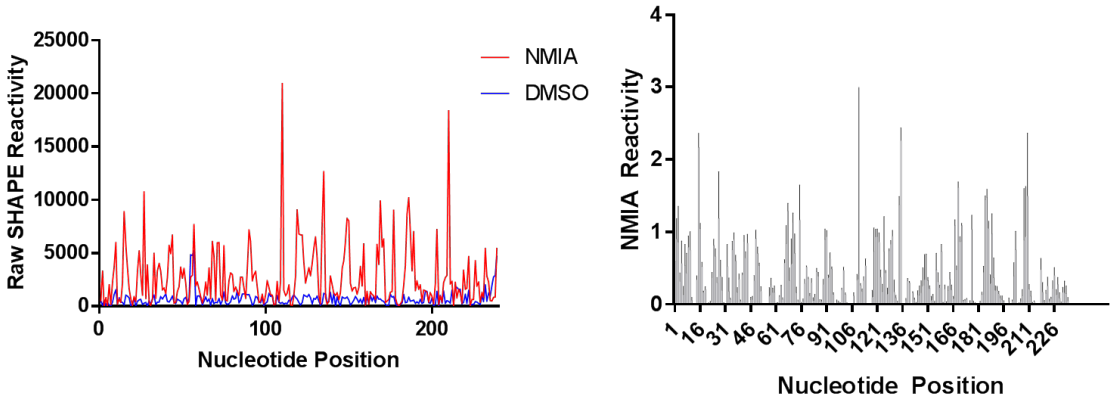

Figure 5.4: SHAPE reactivity of the pseudoknot region.
A. Raw SHAPE reactivity data representing early terminations caused by treatment with NMIA (red) or DMSO (blue). B. Read out of individual nucleotide NMIA reactivity as analysed by SHAPE reactions and capillary electrophoresis. High reactivity indicates high probability of non-base paired nucleotides. Data was analysed using QuSHAPE where negative reactivates were set to 0 and 'super reactive' nucleotides set to a maximum of 3, error bars represent SEM (n=4).

The SHAPE data largely agree with the predicted structures, with the stems of PK 1, 2 and 3 being generally unreactive, suggesting they are base paired. Formation of the stem of PK 4 was, however, less convincing, with the stem showing low levels of reactivity throughout as indicated by the white colouring of the nucleotides, this suggests a low likely hood of these nucleotides being involved in base pairing (Fig 5.5). This could be reflective of the fact that PK 4 has the highest sequence diversity amongst all the PKs, suggesting that it might not always form a stable structure (Fig 5.5).

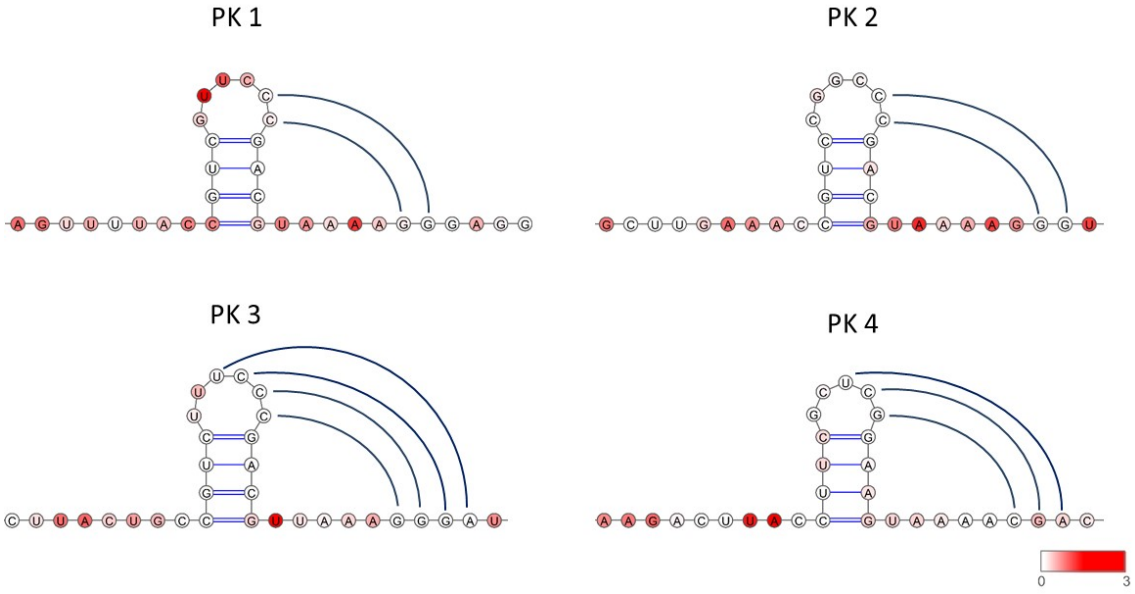

Figure 5.5: SHAPE NMIA reactivity of the pseudoknot region.
NMIA reactivity of each pseudoknot overlaid onto the predicted pseudoknot structure using VARNA. Loop and downstream interactions represent those supported by SHAPE data. NMIA reactivity represented on a colour scale from low (white) to high (red) (n=4).

Looking at the NMIA reactivity of the bulges and down-stream nucleotides also supports the predicted structure, although the SHAPE data shown here suggests that there might be fewer interactions than previously predicted. PK 1 was predicted to have four nucleotides in the loop which interact with four nucleotides downstream. Conversely, the SHAPE data suggests that only two nucleotides might be involved in this reaction, with two of the four showing low reactivity whilst the other two

were more reactive. This is mirrored with the downstream interactive nucleotides where we can see two potential interaction partners which are unreactive. The remaining nucleotides in the bulge are strongly reactive, further supporting the stem-loop structure.

PKs 2 and 3 show a similar pattern, with the three-interacting loop cytosines being unreactive as well as two down-stream guanosines having no or low reactivity. The SHAPE analysis for PK 3 shows NMIA reactivity most like that of the predicted structure with three unreactive cytosines in the loop as well as three unreactive downstream guanosines, suggesting that three interactions could take place in this structure. Interestingly for these two pseudoknots, as well as PK 4, the nucleotides in the loop, not predicted to be interacting with downstream elements, also show very low reactivity.

PK 4 is less typical, with the stem showing more reactivity as well as the predicted downstream interaction site 'CGA'. This suggests along with the larger sequence diversity seen in PK 4 that it may be more flexible and form a less stable PK.

Whilst the difficulties in predicting pseudoknot structures remain, the use of SHAPE data overlaid onto previously predicted structures helps increase our confidence in these structures.

5.2.2 Deletion of replicon encoded pseudoknots

The O1K replicon used in these studies possess four PKs (Fig 5.1). To investigate if all four are needed, the pseudoknots were deleted in a sequential fashion from the $3'$ end (i.e PK 4 – PK 1) and investigated to see if any effect on replication was observed. Since wild strains have been documented with only two PKs, it is predicted that no change in replication when reducing the number of pseudoknots in the O1K replicon down to two[38]. However, when reducing the number of pseudoknots further, a phenotype might emerge.

Deletion was performed by complete removal of each individual 48 nucleotide pseudoknot fragment from the pRep-ptGFP replicon plasmid via PCR mutagenesis. This better reflected the sequences found in recorded isolates than just deleting the pseudoknot structure itself. Deletion was performed sequentially starting with removal of pseudoknots 3 and 4 (ΔPK 34 to produce a replicon containing only 2 pseudoknots. Pseudoknot 2 was then also deleted to produced ΔPK 234, leaving PK 1 as the sole remaining pseudoknot. Due to the location of PK 1 being immediately $3'$ to the poly-C-tract removal proved difficult. To allow for deletion to occur an AflII site was introduced at the end of the poly-C-tract and at the beginning of PK 1. During the mutagenesis to introduce this site the poly-C-tract was truncated, probably due to the polymerase skipping over the repetitive region. The resultant poly-C-tract was reduced to 11 cytosines, creating a replicon termed C11. Previous investigation into the effect of poly-C-tract truncation on viral replication showed reducing the poly-C-tract to two cytosines (C2) still permitted viral replication [163]. To check the same was true with our mutant, the replication of the C11 replicon was investigated alongside WT and GNN controls to ensure there was no negative effect on replication. For comprehensive analysis, the entire length of the poly-C-tract was also removed creating the C0 replicon. Replication was observed by transcription of the mutant replicon clones alongside wild-type and 3D-GNN controls. Transcribed replicon RNAs were then transfected into BHK-21 cells, the bovine cell line (MDBKs), and a porcine cell line (SK-RST). Replication was observed by monitoring GFP-expression using the Incucyte Zoom as initially described in section 3.2.1. Consistent with previous investigation into the poly-C-tract, we saw no significant drop in replication in either the C11 or C0 replicons when compared to the WT (Fig 5.6) [163].

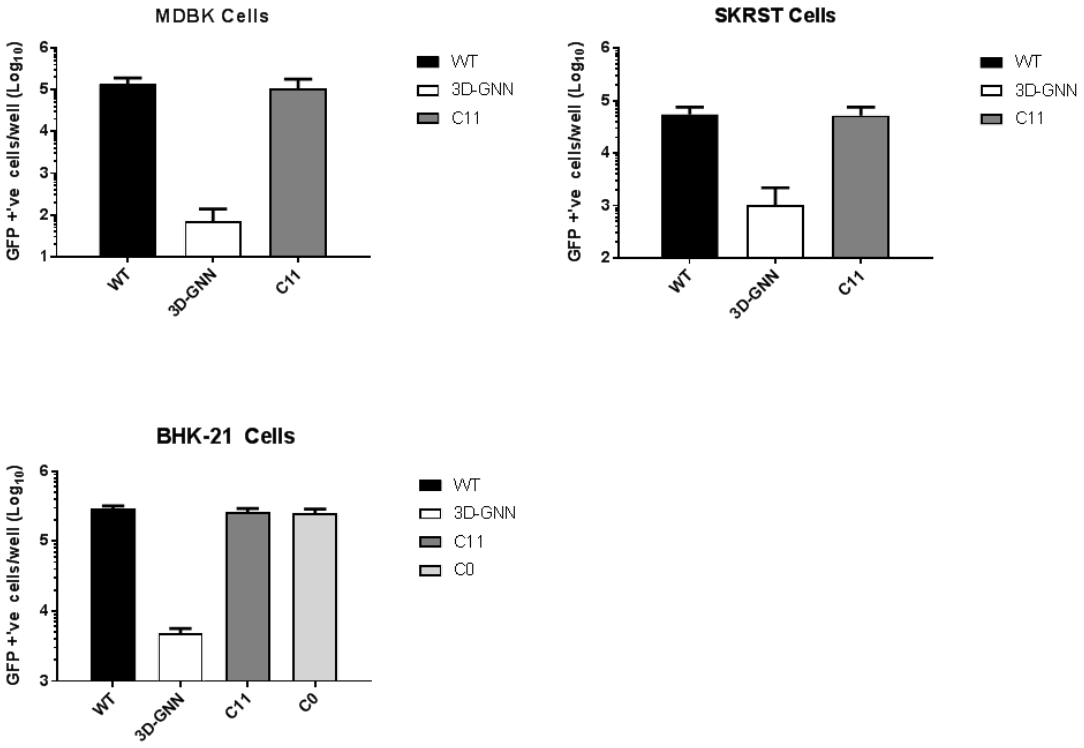

Figure 5.6: Replicative ability of replicons containing poly-C-tract truncations.
MDBK, SKRST and BHK-21 cells where transfected with a replicon containing a poly-C-tract of 11 cytosines in length (C11). BHK-21 cells were also transfected with a replicon where the poly-C-tract was removed entirely (C0). Mutant replicons were transfected alongside WT and 3D-GNN controls, GFP was measured by IncuCyte, data shown is representative of 8 hours post transfection. Error bars shown calculated by SEM, n=2.

After confirming that the C11 replicon could replicate at a similar level to the WT, it was used as the template for deleting all four pseudoknots (C11 ΔPK 1234). Replicons containing the various pseudoknot deletions (Fig 5.7) were transcribed and transfected into the three cell lines BHK-21, MDBK and SK-RST. Replication was then monitored using an Incucyte Zoom. Credit for some of this work should be given to an MBiol student under my supervision, Emma Warner, who helped in the creation of deleted pseudoknot mutants and transfection experiments.

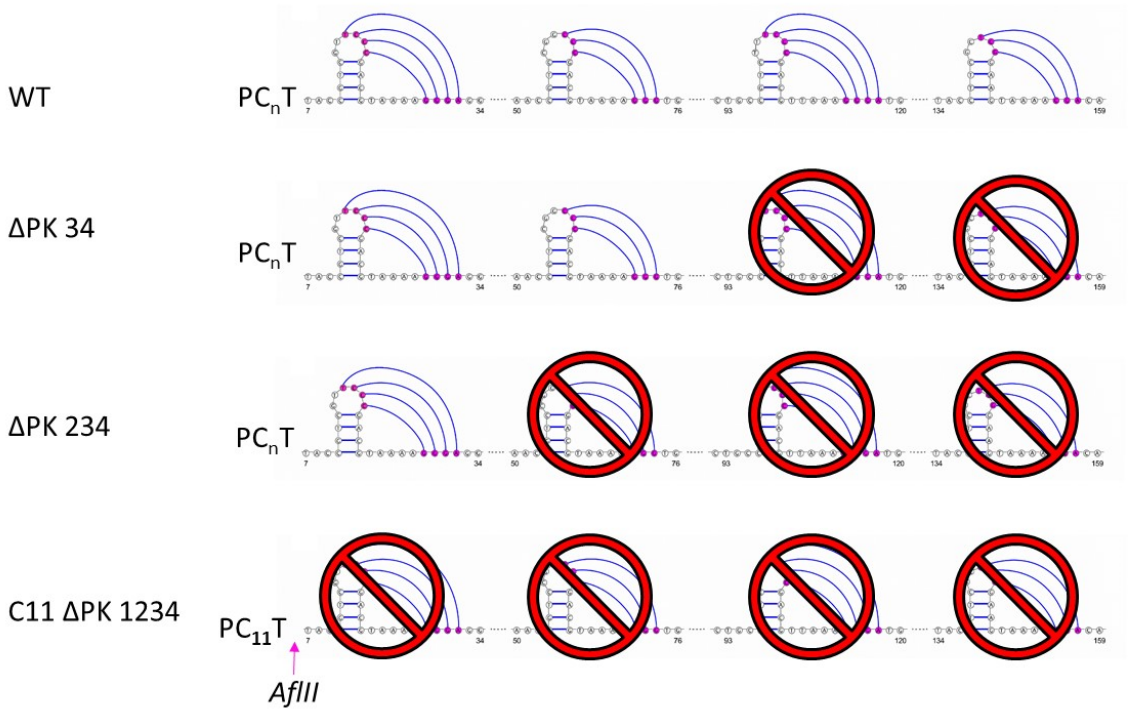

Figure 5.7: Schematic of pseudoknot deletions
Cartoon displaying pseudoknot deleted replicons

It was observed that replicons lacking PK 34 and PK 234 (ΔPK 34 and ΔPK 234 respectively) displayed no significant difference in replication when compared to the wild-type replicon (Fig 5.8). The replicon containing no PKs (C11 ΔPK 1234) did, however, show a significant decrease in replication when compared to the wild-type replicon in all three cell lines (Fig 5.8). This suggests that the pseudoknots must play a role in genome replication. However, as replication was still over 10-fold higher than that of our negative control, 3D-GNN, replication must be still occurring and therefore the role the pseudoknots play in genome replication may not be essential. It also shows that there may be some redundancy in the number of pseudoknots as only one PK was required to conserve wild-type replication in any of the cell lines used. The requirement of only a single pseudoknot is striking, especially since no isolate to date has been recovered with fewer than two pseudoknots. The decrease in replication observed upon removal of all four pseudoknots was more obvious when

transfected into SK-RST and MDBK cells, although this was still approximately 10-fold higher than the 3D-GNN. The greater difference observed in these cell lines might reflect the lower levels of replication seen in the SK-RST and MDBK cell lines, making any differences more obvious. This difference is not seen here due to the metric of GFP cells per well, however when looking at overall GFP production from constructs transfected into the different cell types, it is observed that both MDBK and SK-RSTs show lower levels of GFP expression (Fig 5.9).

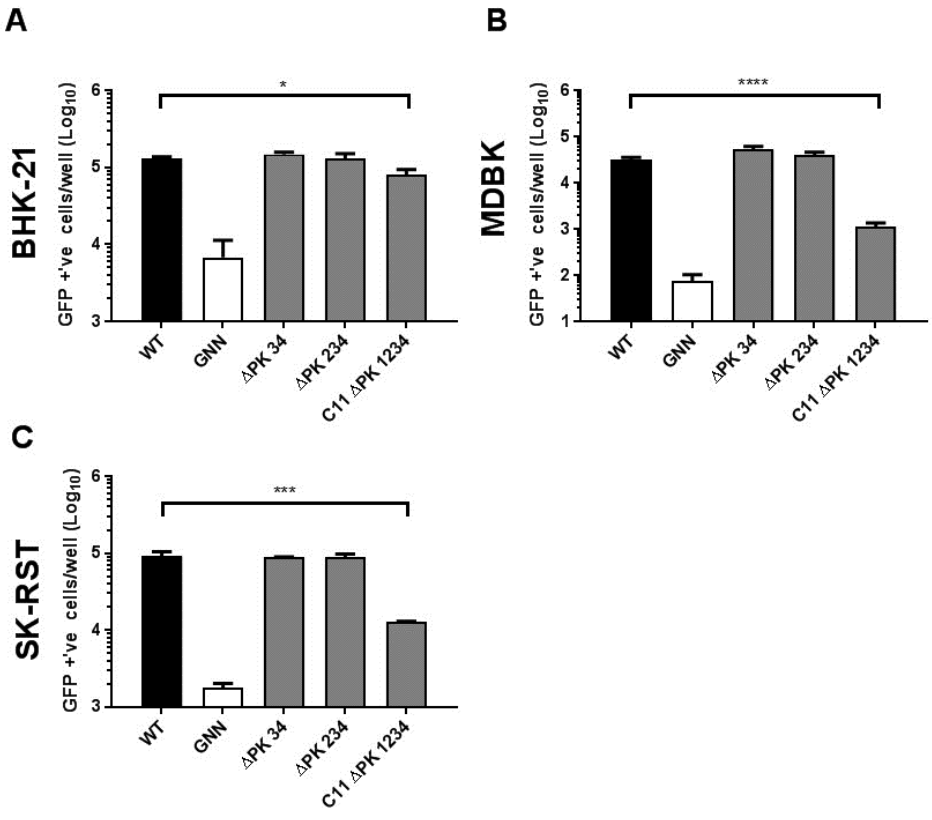

Figure 5.8: Replicon replication can occur with all four pseudoknots deleted.

A, B, C. Replicons with sequentially deleted pseudoknots were assayed for replication in 3 different cell lines (labelled). Replication was measured by counting the number of GFP positive cells per well using the Incucyte Zoom. Error bars shown are calculated by SEM, n=3 * $p < 0.05$, *** $p < 0.001$, **** $p < 0.0001$. Statistical analysis performed by paired T-test.

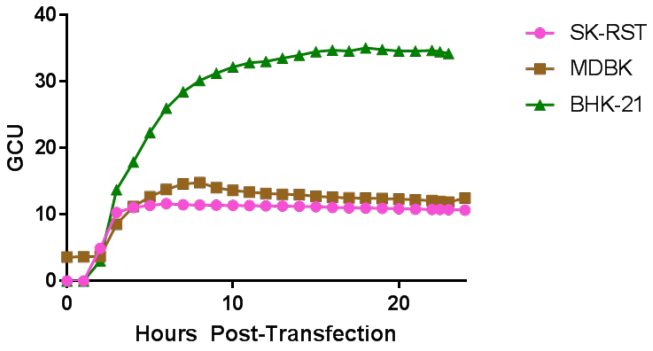

Figure 5.9: Differences in GFP intensity in different FMDV replicon transfected cell lines.
BHK-21 (green), SK-RST (pink) and MDBK (brown) cell lines were transfected with WT FMDV replicon RNA. Replication is shown here over 24 hours and expressed as the average green fluorescent intensity (GCU).

5.2.3 New genome synthesis in pseudoknot mutants

To investigate genome replication in further detail, the number of new genomes made was measured utilising incorporation of tritiated uridine (^{3}H) into newly synthesised RNA. This provided both a more direct and sensitive way to observe any differences between the WT and C11 ΔPK 1234 replicons.

Preliminary experiments were performed using replicon transfected BHK-21 cells, with and without the addition of actinomycin D. The role of actinomycin D was to inhibit the activity of the host cell RNA polymerases, thereby preventing the incorporation of any ^{3}H into cellular mRNA [160]. At one hour post-transfection the cells were treated with 10 μg/ml of actinomycin D, or an equal volume of DMSO as a control. At two hours post-transfection 18 μCi of ^{3}H uridine was added per well of a 6 well plate, to the BHK-21 cells and incubated for 2 hours. The media was removed and total RNA harvested using trizol. Equal amounts of purified total RNA was then applied to a 5-25 % sucrose gradient containing 0.1 % SDS and 100 mM sodium acetate. The gradient was centrifuged in a high-speed centrifuge at 150,000

RCF for 5 hours at room temperature. Gradients could then be fractionated and added to a tube containing scintillation fluid. Presence of ^{3}H in the fractions was analysed using a scintillation counter.

When comparing the amount of new strand synthesis upon transfecting WT, 3D-GNN and C11 ΔPK 1234 replicons, a clear peak could be seen in the WT and C11 ΔPK 1234 samples at around fractions 11-12, although the latter was at a much lower level (Fig 5.10AB). This replicon peak correlated with a control experiment where purified transcribed replicon RNA was separated on the sucrose gradient and fractions analysed by absorbance at 260 nm (Fig 5.10E). The observed peak was also absent in the 3D-GNN and un-transfected controls (Fig 5.10CD). A second peak could be seen higher up the gradient, and was seen in all four samples at similar levels, suggesting this was a peak corresponding to some cellular RNA. The peak corresponding to newly generated replicon genomes created by the C11 ΔPK 1234 replicon was only at 3 % of the WT, representing a much larger difference than that seen measuring replication using the incucyte suggesting that the pseudoknots are playing a much larger role in genome replication.

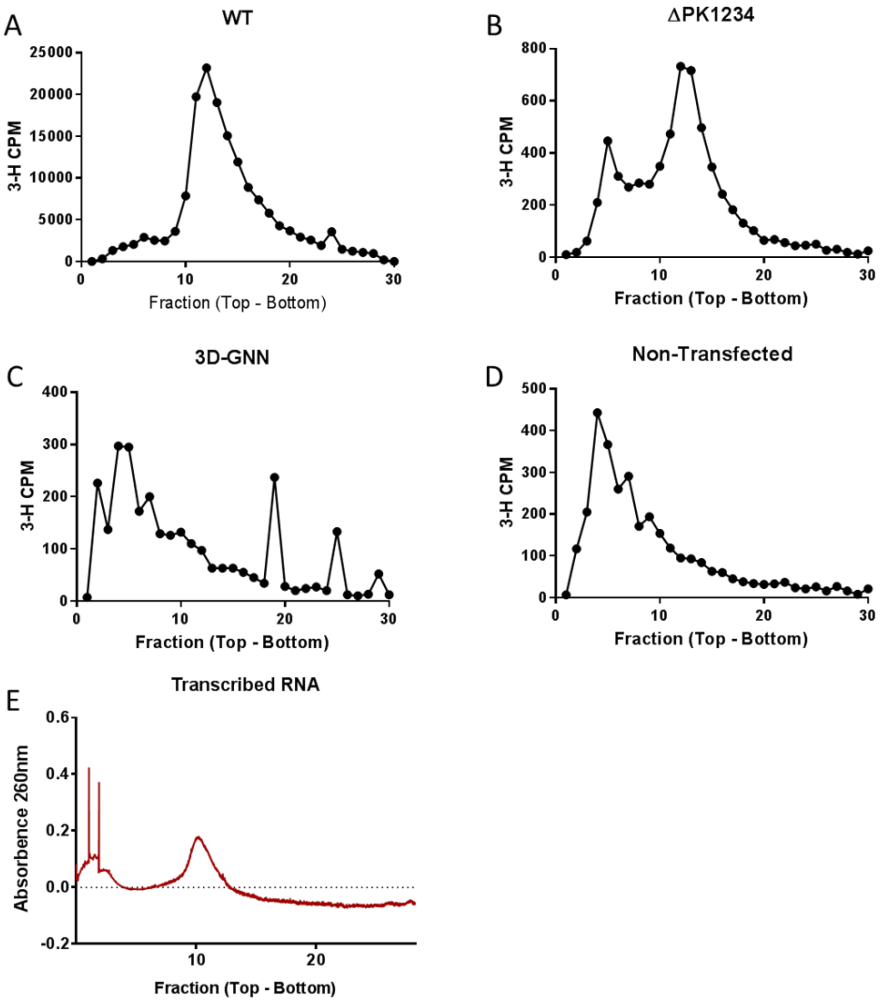

Figure 5.10: Comparison of new replicon genomes made from different FMDV replicons.

Acintomycin D treated BHK-21 cells, transfected with either WT (A), C11 ΔPK 1234 (B), 3D-GNN (C), alongside non-transfected BHK-21 cells (D) were harvested for total RNA at 4 hours post transfection. Total RNA was subjected to ultracentrifugation on 5-25 % sucrose gradients. Gradients were fractionated and analysed by scintillation counting to detect sedimentation of RNA species. Purified transcribed RNA run on parallel 5-25 % sucrose gradients was fractionated and analysed by monitoring absorbance at 260 nm (E), n=1.

5.2.4 Pseudoknot deletions in infectious virus

Since genome replication could be detected in all the replicons containing pseudoknot deletions, pseudoknot deletions were introduced into the infectious clone to observe if virus could be recovered. As described in section 4.2.5, this was done in collaboration with the Pirbright Institute with the help of the Toby Tuthill research group.

The pseudoknot deleted replicons were converted into infectious clones by re-introduction of the capsid proteins, infectious clones were transcribed and resultant RNA transfected into BHK-21 cells. Virus recovery was performed as described previously (Section 4.2.5), briefly, infectious clone RNA was transfected into 25 cm^2 flasks containing BHK-21 cells. At 24 hours pos-transfection or at appearance of CPE, cell lysates were freeze thawed and clarified by centrifugation. Clarified lysates were then reintroduced onto naïve BHK-21 cells. This blind passaging occurred 5 times and at passage 4 any recovered virus was sequenced to oserved changes from the input sequence.

CPE was observed in the WT, ΔPK 34 and ΔPK 234 transfected cells and when sequenced after the fourth passage no change in sequence was noted. A similar result was seen for the C11 mutant, however after passaging, sequencing revealed an expansion of the poly-C-tract, although the total length could not be reliably measured. Surprisingly despite genome replication being observed in the replicon system, C11 ΔPK 1234 produced no CPE and infectious virus could not be recovered (Fig 5.11). This suggests that the pseudoknots may play another role in the viral life cycle which is separate to that of genome replication. This could include roles such as packaging of the viral genome or capsid assembly. Further work is currently ongoing to investigating the potential role in genome packaging. Since this involves the use of capsid sequences, biosecurity mandates insist this must be performed at the Pirbright Institute.

	Appearance of CPE				
Recovered Virus	Passage 1	Passage 2	Passage 3	Passage 4	Sequence of Rescued virus
WT	✓	✓	✓	✓	No Change
C11	✓	✓	✓	✓	No Change
ΔPK34	✓	✓	✓	✓	No Change
ΔPK234	✓	✓	✓	✓	No Change
C11 ΔPK1234	✗	✗	✗	✗	N/A

Figure 5.11: Viral recovery of pseudoknot deleted mutants.
Infectious clones containing pseudoknot deleting mutants (ΔPK34, ΔPK234, ΔPK1234), poly-C-tract truncation (C11) or WT sequence were transcribed and transfected into BHK-21 cells. Recovered virus was blind passaged five times and sequenced to identify additional mutations or reversions. Transfections were all performed in duplicate. *Data provided by Lidia Laseka-Dykes, the Pirbright Institute.*

5.2.5 Investigation into the effects of pseudoknot removal on genome replication

The studies above suggested that the pseudoknots are involved in two stages of the viral lifecycle, genome replication and production of mature infectious virus particles. Since the role in generation of infectious virus particles is limited to study at the Pirbright Institute, investigation into the role of the pseudoknots in genome replication was continued at Leeds. We sought to better understand the requirements of the pseudoknots to maintain the wild-type levels of replication, with the hope of better understanding the mechanism of action. Some of the work described below was performed with the assistance of a second MBiol student under my supervision, Niall McLean.

Within the poliovirus 5′ UTR is a stretch of approximately 34 nucleotides which provide a role as a spacer between the cloverleaf and IRES. As the pseudoknots lie in a similar location, it was thought they may be providing a similar role. To

investigate this, the pseudoknots were removed and replaced with a stretch of equal length randomised sequence. The resultant construct produced was named 'C11 PK Replace'. The replication of this construct was tested as before using WT, 3D-GNN and C11 ΔPK 1234 as positive and negative controls. Interestingly, the replication of C11 PK Replace was lower than that of constructs having no pseudoknots at all, suggesting that having the random sequence in its place was detrimental. This suggest that the pseudoknot function was not as a simple spacer between other structural elements of the 5′ UTR (Fig 5.12).

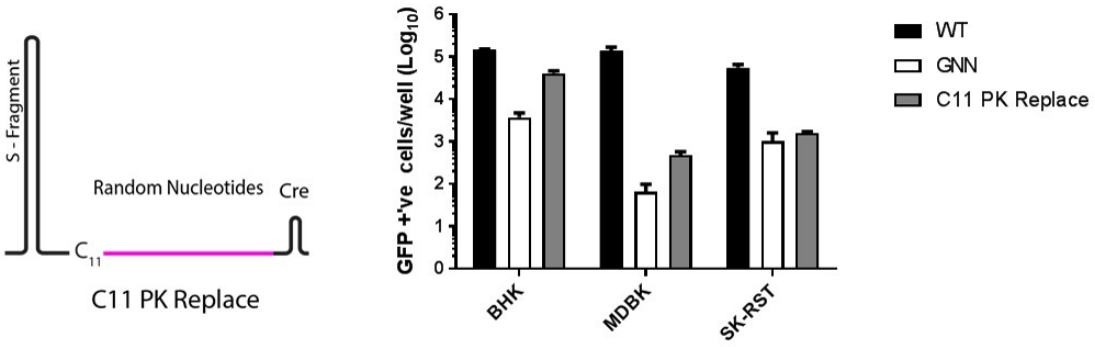

Figure 5.12: Replacing the pseudoknots with a random spacer decreases replicon replication.
A. Cartoon showing the deletion of pseudoknots and replacement with random nucleotides. B. Replicative ability of replicons with replacement of the PK region with a series of random nucleotides of equal length, along side WT and 3D-GNN controls. Replication is shown as GFP positive cells/well at 8 hours post transfection of BHK-21, MDBK and SK-RST cells. Error bars shown calculated by SEM, n=2.

As the experiment above revealed that the sequence of the pseudoknot region was important for its function, we probed the requirements of this sequence to maintain pseudoknot structure and function. For simplicity, the C11 construct with PKs 2, 3 and 4 removed was used, leaving one functional pseudoknot (PK 1) in the replicon. Mutations were made in the loop of PK 1 and the corresponding downstream nucleotides designed to interrupt base pairing and abrogate formation of the PK structure. Nucleotide changes (shown in red), created a GAGA motif in both the

bulge and downstream region (Fig 5.13). The created construct named C11 PK Disrupt, was transcribed and transfected into BHK-21 cells and compared to the replication of WT, 3D-GNN and C11 ΔPK 1234 replicons. Replication of the PK disrupt replicon showed equivalent replication to the C11 ΔPK 1234 replicon, suggesting this interaction was essential for function (Fig 5.13). These mutations also helped support the predicted structure proposed by Clarke et al.

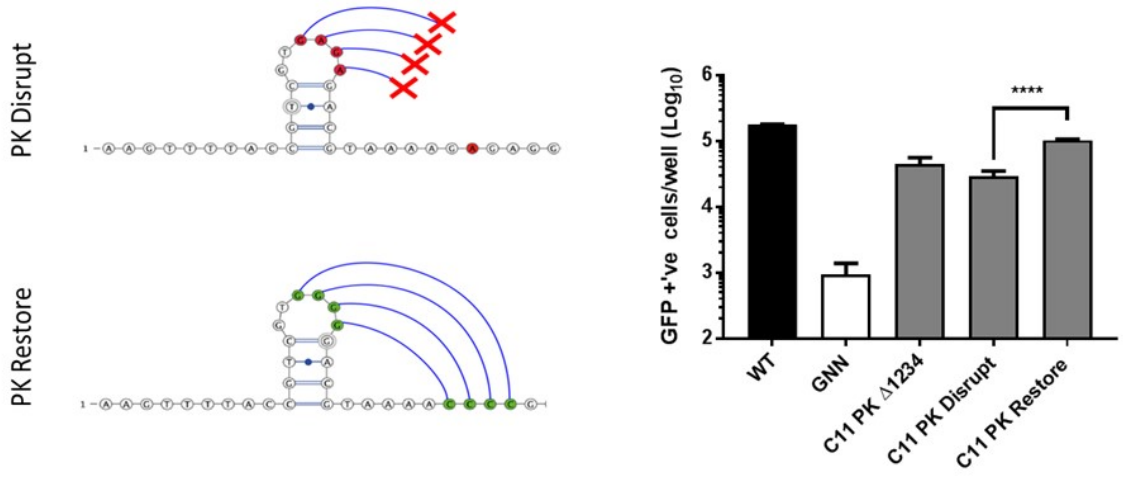

Figure 5.13: Disrupting the pseudoknot structure reduces replication. Cartoon representations of mutations made to the pseudoknots, where nucleotides in the bulge of the stem loop and interacting region downstream are mutated to disrupt structure formation 'PK disrupt' and mutated to maintain bulge and downstream interaction but with different nucleotides 'PK restore'. Replicative ability of mutant replicons shown as number of fluorescent cells/well of replicons 8 hours post transfection of BHK-21 cells. Error bars shown calculated by SEM, n=3, statistics shown represent comparing C11 PK disrupt and C11 PK restore, **** $p < 0.0001$.

The pseudoknot structure was then restored by mutation of the relevant nucleotides in the loop and downstream region to GGGG and CCCC respectively, producing the C11 PK Restore replicon. Restoring this interaction produced a significant increase in replication compared to the disrupted PK replicon, although this was still below that of the WT. The slight reduction in replication still observed in the C11 PK Restore replicon could be due to the alternate nucleotides used to make the interaction. As these are the non-native nucleotides, artificially increasing the GC

content this could influence the strength of the nucleotide interaction and stability of the structure, therefore having a knock-on effect on pseudoknot function.

The orientation of PK 1 was also reversed by "flipping" the nucleotide sequence. This reversal of the sequence should make the pseudoknot fold towards the 5′ end of the genome as opposed to the 3′ end. This replicon, 'C11 PK Rvs', was assayed for replication alongside WT, 3D-GNN and C11 ΔPK 1234 in BHK-21 cells. Interestingly, it was observed that reversing the orientation of PK 1 gave a similar replication phenotype to having no pseudoknots at all (Fig 5.14). This suggests that the role of the pseudoknots in genome replication is dependent not only on their structure but also on their orientation.

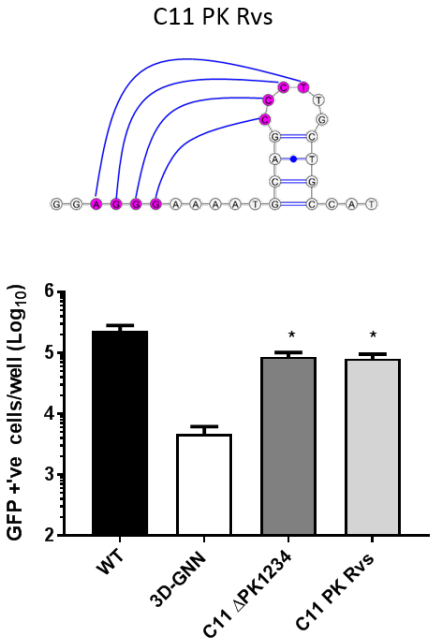

Figure 5.14: Reversing the pseudoknot structure impairs function.
Cartoon representations of mutation made to the pseudoknots where the nucleotide sequence is flipped to reverse the structure of the pseudoknot 'C11 PK Rvs'. B. Replicative ability of mutant replicons shown as number of fluorescent cells/well of replicons 8 hours post transfection of BHK-21 cells. Error bars shown calculated by SEM, n=3, statistics shown represent comparing C11 ΔPK1234 and C11 PK Rvs to the WT replicon and calculated using a paired T-test, * $p < 0.05$.

As previously stated the number of pseudoknots in the FMDV genome can vary from strain to strain, with a minimum of two and a maximum of four detected in field isolates. We wondered if adding additional pseudoknots might have an effect on genome replication, and if so, what insights into potential functions in genome replication this might provide.

A construct was created where PK 1-4 was duplicated and inserted immediately after the original four pseudoknots. This created a replicon containing the following PKs, PK 1, 2, 3, 4 , 1, 2, 3, 4. This replicon was named PK 8 and any effects on replication were investigated by transfection of replicon RNA into BHK-21 cells.

Introduction of four more pseudoknots did not seem to have any effect on genome replication, and GFP production akin to the WT replicon was observed (Fig 5.15). This observation supports previous data and the hypothesis that there is potential redundancy in the number of pseudoknots within the FMDV genome .

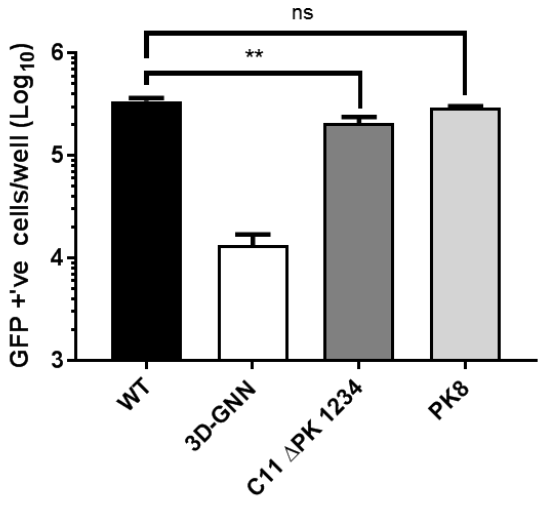

Figure 5.15: Adding additional pseudoknots does not effect FMDV replication.
A construct with twice the number of pseudoknots (PK8), was transfected alongside, WT, 3D-GNN and C11 ΔPK 1234 controls. Replicative ability of mutant replicons shown as number of fluorescent cells/well 8 hours post transfection of BHK-21 cells. Error bars shown calculated by SEM, n=3, statistics shown represent comparing C11 ΔPK 1234 and PK8 to the WT replicon using a paired T-test, ** $p < 0.01$.

All mutational study performed so far focussed on the deletion of pseudoknots 2, 3 and 4, leaving PK 1 as the sole remaining pseudoknot. To check that all pseudoknots were providing the same role, PK 1 was replaced with the most diverse pseudoknot (in terms of sequence variation), PK 4. The resultant replicon 'C11 PK 4' was transcribed and replication measured against that of the WT and 3D-GNN replicons in BHK-21, SKRST and MDBK cell lines. Having PK 4 as the only remaining pseudoknot did show a slight drop in replication compared to the WT, more noticeable in SK-RSTs than BHK-21 or MDBK cells (Fig 5.16). This drop was not significant and still considerably less than that of the drop in replication seen from the C11 ΔPK 1234 replicon in these cell lines. If we take the sequence diversity and SHAPE data into account, it might not be surprising that we observe a small decrease in replication, as all evidence so far suggests that PK 4 may form weaker interactions

to maintain its structure.

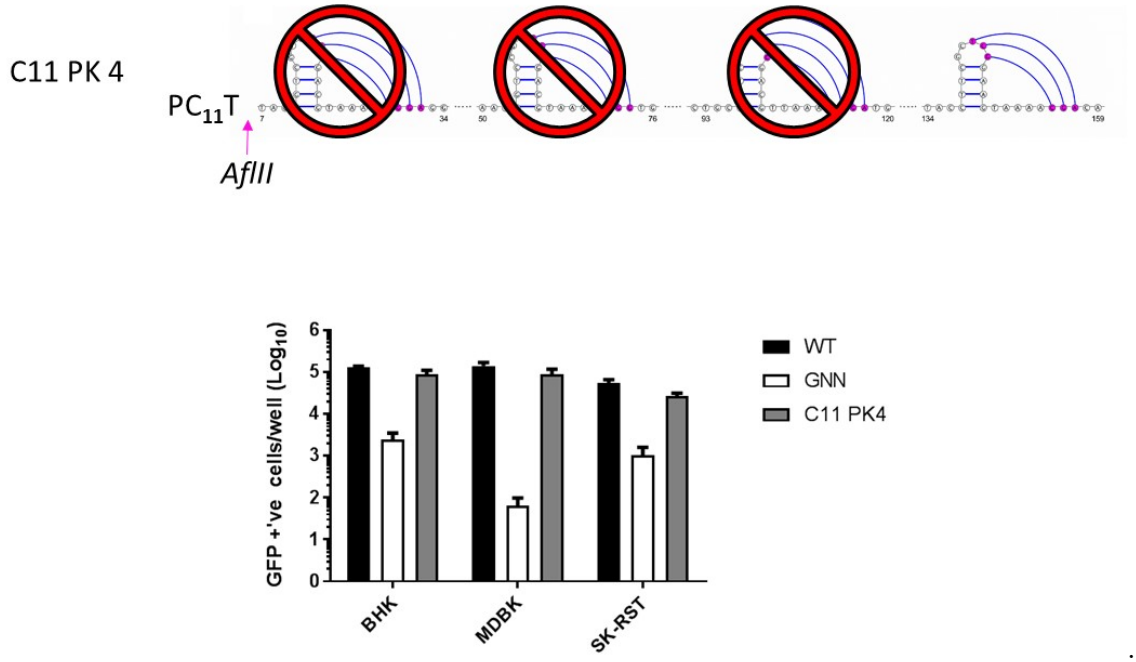

Figure 5.16: There is no preference to which pseudoknot is maintained. Top. Cartoon showing the deletion of pseudoknots 1, 2 ad 3 and maintenance of PK 4. Bottom. Replicative ability of replicons when the PK 4 is the sole remaining pseudoknot compared to WT and 3D-GNN replicons. Replication is shown as GFP positive cells/well at 8 hours post transfection of BHK-21, MDBK and SK-RST cells. Error bars shown calculated by SEM, n=2.

5.2.6 Removal of pseudoknots could have a negative impact on replicon translation

The replicons containing deleted pseudoknots were combined with the lethal 3D-GNN mutation, creating C11 ΔPK 1234 GNN, ΔPK 234 GNN and ΔPK 34 GNN. These were transfected alongside WT and 3D-GNN control replicons to see if any effect on input translation (the translation seen from transfected RNA) could be seen with the removal of pseudoknots. It was observed that the ΔPK 234 and ΔPK 34 GNN replicons showed a similar amount of input translation to that of the 3D-GNN replicon. However, the C11 ΔPK 1234 GNN replicon showed a significant 6-fold

reduction in GFP expression, suggesting the pseudoknots could play a role in the control of IRES driven translation (Fig 5.17).

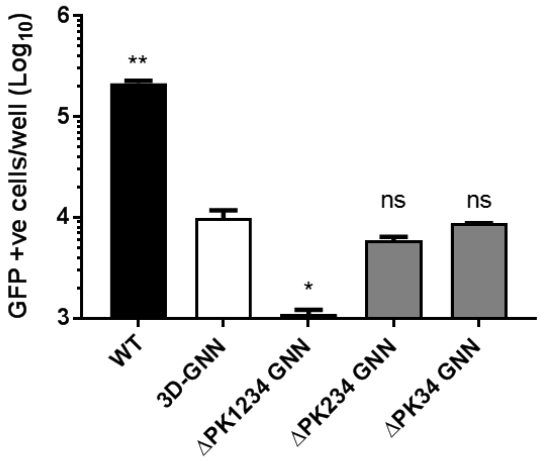

Figure 5.17: Removing all pseudoknots reduces input translation from replicative defective replicons.
Pseudoknot deleted replicons were combined with the 3D active site mutation 'GNN' to look at the effect of the pseudoknots on translation. Replicons where transfected into BHK-21 cells and where monitored for replication by reporter gene expression via IncuCyte Zoom. Data shown represents number of GFP positive cells/well 8 hours post transfection. Error bars shown calculated by SEM, n=2. Significance shown when compared to the ptGFP GNN control. ns $p > 0.05$, * $p < 0.05$, ** $p < 0.01$

5.2.7 There is a competitive advantage to possessing more pseudoknots

The question remained, why FMDV can possess up to 4 PKs when evidentially only one appears sufficient for wild type levels of genome replication and in fact genome replication could still occur with no pseudoknots at all.

Competition experiments where performed with the assistance of a post-doctoral scientist in the FMDV team at Leeds, Oluwapelumi Adeyemi, to investigate if having

more pseudoknots provided a competitive advantage in replicon replication. Replicons comprising differential fluorescent reporters (either GFP of mCherry) were used to easily distinguish replication of the different competitors. GFP expression replicons WT, ΔPK 34, ΔPK 234 and C11 ΔPK 1234 were co-transfected with either a WT mCherry replicon or yeast tRNA as a control (Fig 5.18).

The mCherry and GFP replicons were co-transfected into BHK-21 cells and replication was monitored by observing the levels of GFP and mCherry expression over three sequential passages. At the end of the three passages replicons were harvested, reverse transcribed and sequenced to ensure no recombination had taken place. For each passage at 8 hours post-transfection, the total RNA was harvested in trizol-reagent and purified, before re-transfection into new naïve BHK-21 cells.

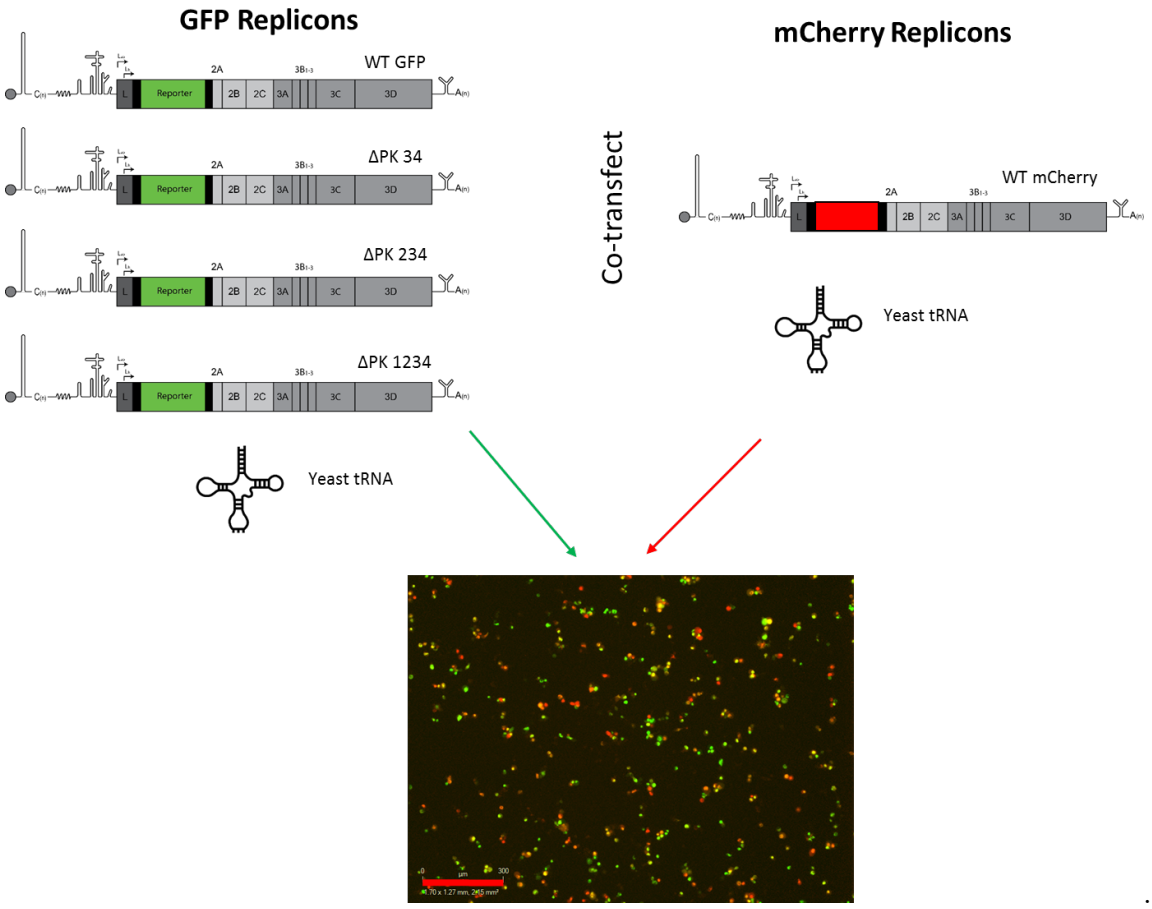

Figure 5.18: Diagram representation of competition experiment.
Co-transfection of GFP replicons, or yeast tRNA was performed with either WT mCherry replicon or yeast tRNA. 500ng of each construct was mixed before transfection in the previously described manner. Scale bar (red) represents 300 μM.

Co-transfection of any of the GFP replicons had no obvious effect on the mCherry WT replicon as expected, demonstrating there no dominant negative effects were occurring (Fig 5.19A). Furthermore, when co-transfected with the WT mCherry replicon, the replication of the WT and ΔPK 34 GFP replicons were comparable. Both see an initial reduction in replication after the first passage, but recover to near that of the original transfection level by the third passage. The ΔPK 234 replicon, however, showed a similar initial drop in replication to the WT and ΔPK 34 constructs, but showed no sequential recovery with each passage. Instead it is observed that replication steadily decreases until at passage three there is a 2.5

fold decrease in GFP positive cells compared to that at passage 0 (Fig 5.19B). This suggests that there might be a disadvantage to having only a single PK when competing with a wild-type replicon, supporting the observation that virus with less than two PKs has not been isolated to date.

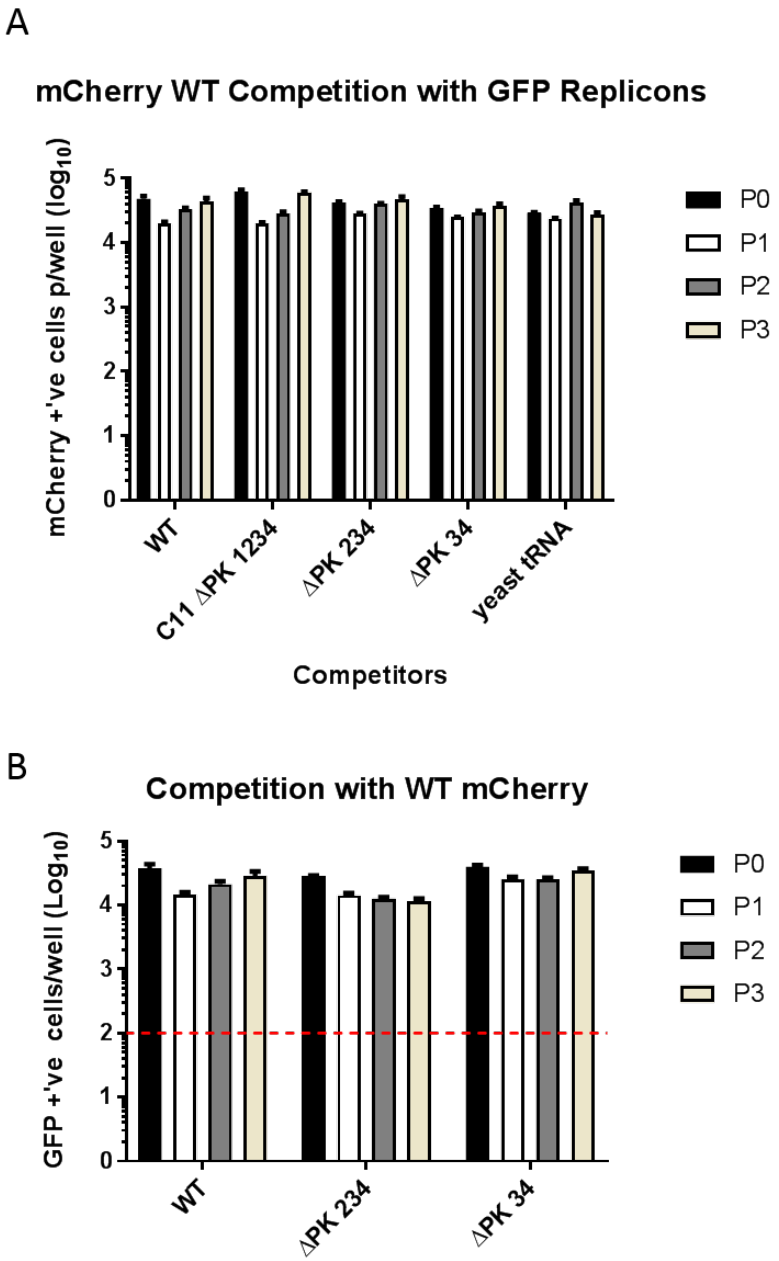

Figure 5.19: Replication of GFP and mCherry replicons in competition experiment over 3 passages.

A. Replication of WT mCherry replicon when co-transfected with all of the GFP containing replicons. B. Replication of WT, ΔPK 234, ΔPK 34 when co-transfected with a WT mCherry replicon over 3 passages. Replication shown as GFP or mCherry positive cells/well at 8 hours post transfection. Red dotted line marks the point of detection over background. Error bars shown calculated by SEM, n=2.

When the C11 ΔPK 1234 replicon was co-transfected with the WT mCherry replicon, a large decrease in replication can be seen at passage 0 when comparing it to the WT mCherry replicon. By passage two, the GFP signal of the C11 ΔPK 1234 replicon was no longer detectable, suggesting that this replicon has been out competed efficiently by the WT mCherry replicon. Interestingly, when C11 ΔPK 1234 was co-transfected with yeast tRNA (as a non-competitive control RNA), the GFP signal was higher than when co-transfected with the mCherry WT replicon at passage 0. Suggesting the competitive advantage of possessing pseudoknots was evident immediately. Surprisingly, when co-transfected with yeast tRNA, the GFP signal of the C11 ΔPK 1234 replicon is reduced at passage two and then further reduces at passage three to a point of no detection (Fig 5.20). This suggests that even without a competitive replicon present, the C11 ΔPK 1234 replicon struggles to persist after multiple passages. This proposes that removal of the pseudoknots could be having a larger impact on replication than first thought. This data supports the findings of the ^{3}H experiment which showed a large reduction in genome replication for this construct, together these observations suggest that genome replication could be severely reduced inhibiting the ability of the replicon to survive multiple passages.

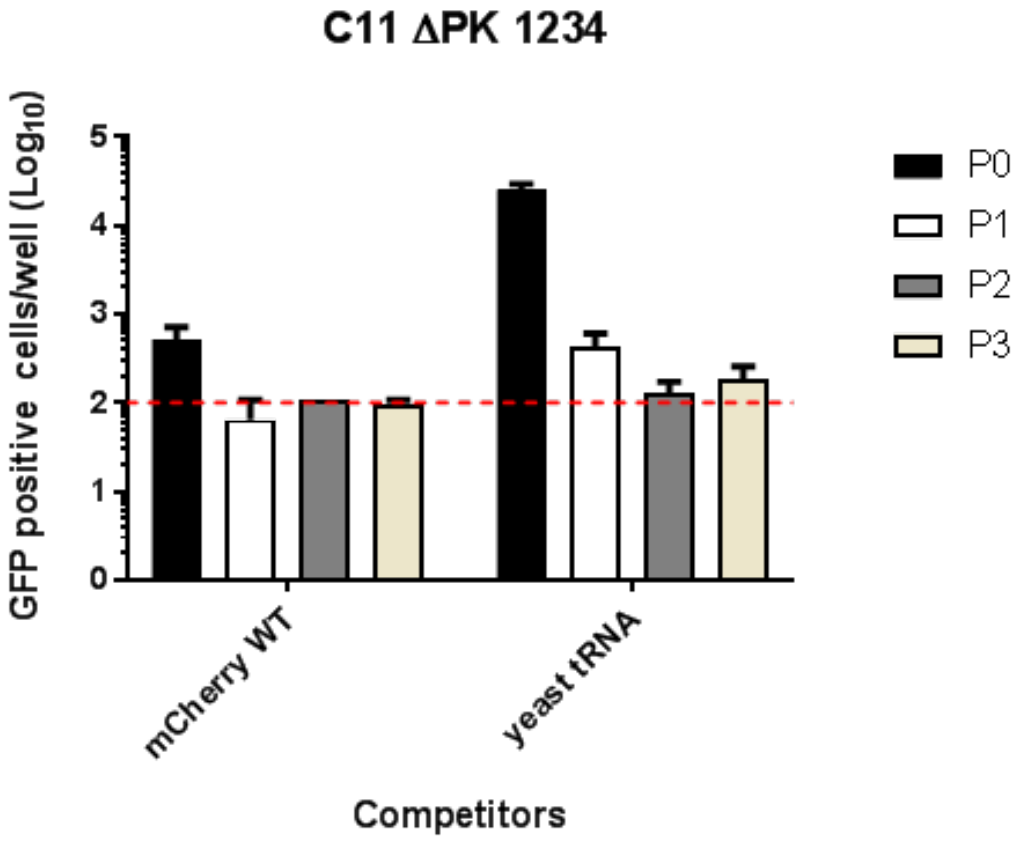

Figure 5.20: Replication of GFP and mCherry replicons in competition experiment over 3 passages.
Replication of C11 ΔPK 1234 replicon, when co-transfected with either a WT mCherry replicon or yeast tRNA over three sequential passages. Replication shown as GFP or mCherry positive cells/well at 8 hours post transfection. Red dotted line marks the point of detection over background. Error bars shown calculated by SEM, n=2.

5.3 Discussion

5.3.1 Improved prediction of pseudoknot structure

Although approaches in predicting pseudoknot structures are improving, confirmation of these structures remains difficult. Using SHAPE approaches we hoped to

gain further insight into the potential structure of the pseudoknot region. Whilst not fitting perfectly with the previously proposed model, the SHAPE data does support the prediction of the pseudoknot structure, although the number of interacting nucleotides required to form the downstream interaction may be less than originally thought. Mutagenesis of the pseudoknot loop region and downstream nucleotides also helped support this structural prediction, with mutation showing a decrease in replicon replication akin to removing the pseudoknots altogether. It is possible that in the conditions the SHAPE reactions were performed, fewer nucleotides were needed or were available to make interactions. It was noted that some nucleotides in the PK 2 and 3 loop were not SHAPE reactive, even though these were not predicted to interact with other nucleotides. This could be due to steric hindrance, where folding of the pseudoknots might block the ability of NMIA to interact with them, thereby providing a false negative reactivity.

Performing these reactions in a cellular environment or in more appropriate conditions could see subtle changes in the structure. In-cell SHAPE would be difficult to perform using a replicon due to the large amounts of RNA required, coupled with the lack of replicon ability to have cell-cell spread. A more realistic alternative would be to use a cell free system, a system that has previously shown to allow for viral replication and infectious particle formation in a micro centrifuge tube. This was performed by taking the S10 extracts from a cell line, commonly HeLa cells, and combining them with enriched initiation factors as well as a reaction buffer containing amongst others creatine phosphate and kinase. Using this methodology would produce an environment closer to that of inside a cell and might provide more reliable information. Alternatively, use of infectious virus and a variety of relevant cell lines, would allow easier amplification of the viral genome and therefore permit harvest of sufficient material for SHAPE reactions. However, due to containment restrictions, such experiments would be limited to the facilities such as the Pirbright Institute.

The observation that a virus lacking pseudoknots might be deficient in infectious

virus production, could be due, to defects in capsid assembly or genome packaging. It would be interesting to perform *in-virion* SHAPE, a protocol which takes purified full virions and performs the SHAPE reactions on these, in a non-cell environment. Differences in SHAPE reactivity compared to that of in-cell SHAPE could highlight interactions between RNA structures and the protein capsid of the virus, further supporting the packaging/assembly hypothesis.

5.3.2 Pseudoknots and the FMDV replicon

Deletion of two pseudoknots to reproduce the minimum seen in field isolates showed similar replication to that of the WT replicon, suggesting that there did not seem to be a benefit to possessing more than two pseudoknots. Indeed, replication persisted at wild type levels with as few as one, and could still occur with no pseudoknots at all, although a slight, but significant, decrease in the level of GFP expression was observed under this condition.

While Incucyte experiments suggested that replication with no pseudoknots was only slightly reduced, experiments trying to visualise the synthesis of new strands suggested a larger effect on genome replication. In experiments using ^{3}H to visualise newly synthesised replicon RNA, the level of RNA produced from the ΔPK 1234 replicon was reduced to only 3 % of the WT levels (Fig 5.10). It is puzzleing why such a larger difference in new strand synthesis was observed while the Incucyte experiment showed only a small reduction in GFP expression (Fig 5.8). One explanation would be that with pseudoknot deletion, translation was upregulated, making it look like replication was taking place. However, efforts to observe input translation by incorporation of 3D-GNN mutations to pseudoknot deleted replicons suggested the opposite, with lower amounts of GFP observed when all four pseudoknots were removed. While this reduction was repeatable, the level of GFP signal was very low, being just above the level of background fluorescence. As such this hypothesis should be further tested using more sensitive methodologies, such as the use of dual

luciferase translation reporter systems[118].

The major peak seen in the RNA sucrose gradient is thought to correspond to replicon RNA, a second peak, higher up the gradient, was also seen in all controls (Fig 5.10). This could be an aspect of incomplete RNA polymerase inactivation by the Actinomycin D treatment, or could represent degraded replicon RNA, although its presence within the 3D-GNN control makes this unlikely.

Replicon experiments suggested that pseudoknots do play a role in genome replication, although to what degree is as yet uncertain, as well as a potential regulating effect on translation. Despite this, it does appear possible for replication to occur without any pseudoknots at all and at WT levels with only one, begging the questions why sometimes four are seen in field isolates.

Replacement of the pseudoknot region with random nucleotides showed that it did not function as a spacer in the 5′UTR. This was as expected as the spacer seen in other viruses tend to not have a specific RNA structure. However, the lack of controls for this experiment makes it hard to determine if this had an overall negative effect on replication, more so than simple removal of the pseudoknots (Fig 5.12). Inclusion of C11 and C11 ΔPK 1234 replicons as experimental controls in this experiment would have provided a clearer insight into any effect on replication. Other mutagenic experimentation to this region revealed that adding extra pseudoknots had no effect on replication, and that the function the replicon was having in genome replication was dependent on its orientation, with reversing the polarity showing similar levels of replication to removing it entirely.

The observation that the direction of pseudoknot folding was important for its function suggested an interaction with replication that's that may function directionaly (i.e effects a polymerase moving in the 3′ - 5′ orientation). This could include the viral polymerase or ribosome, both of which have specific orientations which they work in. It would be interesting to design experiments looking at ribosome or polymerase stalling at the pseudoknots, to see if they have a role in slowing down these

factors, potentially describing a role in both translation and genome replication.

5.3.3 Pseudoknots are essential for virus recovery

With the discovery that replication could occur in the replicon system with no pseudoknots present, deleted mutants were introduced into the FMDV infectious clone to see if any effect on the replication life cycle was observed.

Similar to replicon expeirments, infectious virus was recovered in constructs containing deletions with PK 34 and PK 234 removed. Intersetingly however, deletion of all four pseudoknots led to an absence in the recovery of infectious virus. The large disparity between replicon and viral experiments suggested a major role of the pseudoknots in parts of the life cycle not seen in the replicon system. One obvious target for this is interaction with the capsid proteins in a function such as genome packaging into the capsid, capsid assembly, or viral stability. Deep sequence analysis of many field isolates of FMDV revealed several putative packaging signals, one of which has mapped to the pseudoknot region[114]. While this does not map perfectly to our predicted pseudoknots, it does suggest that this region may be important in the packaging of viral genomes into capsids[114]. This is a role seen by viral RNA secondary structure in viruses such as in HIV-1. While this putative packaging was not investigated further by Logan et al it is the logical next step in the efforts to assign a function to these structures. Efforts into the production of a trans-encapsidation system for FMDV is currently under way to better investigate this potential function along with experimentation to ensure genome replication is occurring in the d.pk1234 virus.

5.3.4 Two is better than one: Replicon competition studies

Co-transfection and passaging studies revealed that the reason for possessing at least two pseudoknots could be due to competition. Those replicons with two or more

pseudoknots had no reduction in replication over 3 passages, whilst the replicon containing only one showed a steady decrease over time when co-transfected with a WT competitor (Fig 5.20). One explanation for this could be that possessing a backup pseudoknot that performs the same function would act as a safeguard for replication. Picornaviruses replicate their genomes using an error-prone RdRp, with the small nature of the pseudoknots, it is likely the structure could be disrupted with very few error driven changes. The position of the pseudoknots, adjacent to the poly-C-tract makes this error seem more likely as due to the large number of repetitive nucleotides the polymerase may struggle to read through this region properly. Therefore, keeping a spare pseudoknot could allow for continued replication should one be disrupted by RdRp error. The mystery remains why some isolates would go as far as having further additional pseudoknots which do not seem to perform any additional function. It is interesting that this somewhat contradicts the dogma of virus biology, where genetic economy is thought to be key and no part of the genome is wasted or superfluous to requirement.

5.3.5 The poly-C-tract and genome replication

In a similar vein of genetic economy, we also confirm the findings of Reider and Mason, and demonstrate that truncation and full deletion of the poly-C-tract seemed to have no effect on replicon replication[163]. Whilst this confirms Reider and Masons findings, who used an A12 type virus, it contradicts that of Zibert and Beck, who reported an O1 type virus which could not produce infectious virus with a poly-C-tract of less than 32 C residues[223].

We show here an advancement on the Reider and Mason story by removing the poly-C-tract in its entirety, although, semantics could be argued that the 2 cytosines left by Reider and Mason does not justify the name of a cytosine tract. We also show that, (contradictory to Zibert and Becks findings) that an O1 virus can replicate with a shortened, (C11), poly-C-tract and that replicon replication can continue

with no poly-C-tract at all.

It is interesting that whilst there is a growing body of evidence for the poly-C-tract having minimal effect on virus replication in FMDV, it has been observed by Reider and Mason, Zibert and Beck as well as ourselves that extension of a shortened poly-C-tract was observed upon sequential viral passage, although absolute numbers of the length of the extensions is difficult to quantify. The growth of the poly-C-tract reported by Reider and Mason was only observed on tracts with an initial length of six cytosines. This growth is likely to be due to polymerase slippage.

Deleting of all four pseudoknots, in combination, led to the truncation of the poly-C-tract. Reider and Mason had the reverse problem, with accidental deletion of PK 1 when they attempted to truncate the poly-C-tract to C2[164]. They found that removal of PK 1 in the context of their C2 virus showed reduced replication when compared to the wild-type, but this replication was like that of the C2 containing all pseudoknots. Suggesting the reduction in replication was not due to the removal of the pseudoknot, but the shortening of the poly-C-tract. The serendipitous complementation of these two studies helps support the finding of both pieces of work, suggesting that there is redundancy in the number of pseudoknots present in the viral genome. It also helps support our finding that the potential multiple roles the pseudoknots are playing in the viral life cycle is not specific to a single pseudoknot.

Whilst the role of the pseudoknots is still unclear, the work described above takes a positive step forward in refining the requirements of the pseudoknots for wild-type levels of replication. Their role in infectious virus production is currently ongoing in collaboration with the Pirbright institute. It is hoped that the use of *in-vitro* packaging assays will help the confirmation of this as a potential role of the pseudoknots. If found to be true, this add to the growing amount of data supporting the role of RNA packaging signals in picornavirus particle assembly.

Chapter 6

Discussion

The body of work described in this thesis betters our understanding of the mechanisms of FMDV replication, with the aim of developing novel ways of disease control.

My primary focus was to investigate the unusual RNA structure found at the $5'$ end of the FMDV $5'$ UTR. This "S-fragment" has not been well characterised, although many mechanisms and functions have been suggested due to comparison with the PV cloverleaf found in the similar location. As the project progressed, I also started investigating other unique aspects of the FMDV genome including the $5'$ UTR pseudoknots and the multiplicity of the 3B gene. Throughout this project I have introduced over 800 different mutations to the FMDV genome, allowing the study of many aspects of viral replication. So far, this work has contributed towards two publications, with another three in the final stages of experimentation or in preparation.

6.0.1 The importance of the 3B proteins in FMDV replication

Previous research has suggested that although conserved amongst FMDV isolates, infectious virus can be produced with only one functional copy of the 3B. This was confirmed in this thesis using the versatile replicon system. Our publication (Herod et al 2017) also delved into observed requirements that the presence of 3B3 was essential for replication. By removing the C-terminus of 3B3 and replacing it with

that of 3B2, and vice-versa, we discovered that the C-terminus of 3B3 is essential for correct polyprotein processing and release of mature $3D^{pol}$. We believe that replacing this region with the C-terminus of 3B2 enhances processing at the 3B3/3C junction creating an alternate dead end pathway of protein processing. This observation suggests an alternate processing pathway to the one proposed by from the study of PV. The processing of this region in FMDV is not well characterised, and as PV only contains one copy of 3B, there lies much room for variation away from the PV dogma. I suggest an alternate pathway with changes to the 'major' route outlined in (Fig 1.9), here I suggest that the C-terminus of 3B2 promotes quick cleavage to produce 3B12 and 3B3CD, which are then processed to mature proteins (Fig 6.1). Experimentation probing polyprotein processing could be carried out to test this hypothesis further. The difficulty lies in the small size of the different molecules of 3B and determining how many 3B proteins are attached within precursors. For example, it would be incredibly difficult to determine if a protein band on an SDS-PAGE gel contains 3B23CD or just 3B3CD.

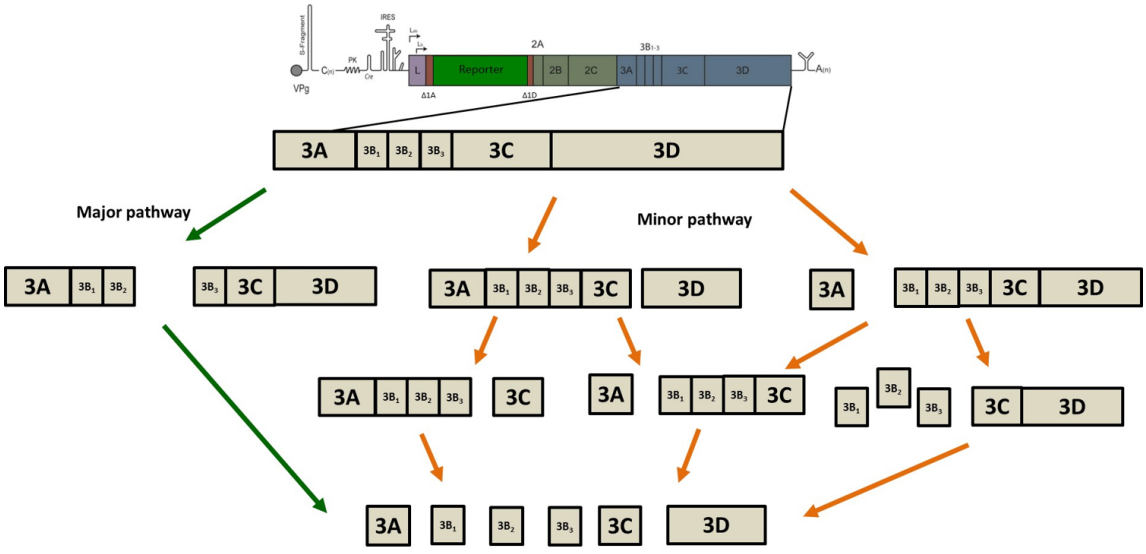

Figure 6.1: Suggested FMDV P3 processing pathway.
An alteration on the classical P3 processing pathway suggested by the study of PV. Changes include alteration in the major pathway to produce 3B12 and 3B3CD precursors, as well as expanding the previous processing pathways described for PV to include the multiple copies of 3B present in FMDV. The two pathways are indicated by the major (green arrows) and minor (orange arrows) routes.

By altering the cleavage boundaries between the proteins in the P3 region, we discovered that the N-terminus of FMDV 3B was essential for its function. Uncleavable fusions to the N-terminus would result in inactivity of the 3B molecule, although fusions to the C-terminus were tolerated, even with large additions such as the entire 3C protein. Although N-terminal fusions inactivated the function of 3B, it remains unknown whether this is due to a block in its function as a primer, where N-terminal fusions could limit $3D^{pol}$ extension by steric hindrance, or whether uridylylation of 3B is inhibited. Previous research studying the uridylylation of 3B molecules has shown that in *in vitro* experiments only 3B or 3BC precursors were functional as uridylylation targets, with no uridylylation of 3AB seen[110]. This would suggest the latter hypothesis, that fusion on the 3B N-terminus inhibits 3B uridylylation. However, these experiments are performed in heavily controlled environments, not reflective of the replication conditions within a cell. Furthermore, evidence to the contrary shows that 3A is essential when attempting to provide functional 3B protein in cis to a 3B inactivated PV with 3BC proteins showing no rescue[110]. This suggests that 3AB could be the uridylylated precursor with the action of 3A anchoring 3B into the membrane through its transmembrane domain, allowing uridylylation before cleavage to a mature 3BpUpU. It is possible that both scenarios are correct but are only providing parts of the story. Combining this information it could be suggested that 3A is essential to provide 3B as it anchors the small protein within a membrane, presumably within the replication complex. Taking the information from the *in vitro* uridylylation experiments, suggesting 3AB is not uridylylateable, the next step could be interaction of proteins such as 3CD or 3C (known to be important as proteins involved in the uridylylation process), which cleave away 3B from the 3AB precursor and traffic this towards the cre for uridylylation (Fig 6.2). While plausible, this theory is hard to demonstrate, and would involve extensive tracking of protein movement and methods of identifying uridylylated 3B proteins within a cellular environment. It is possible that the use of tolerant fusions and epitope tag insertions within 3B molecules described here could be used as tools to help this investigation.

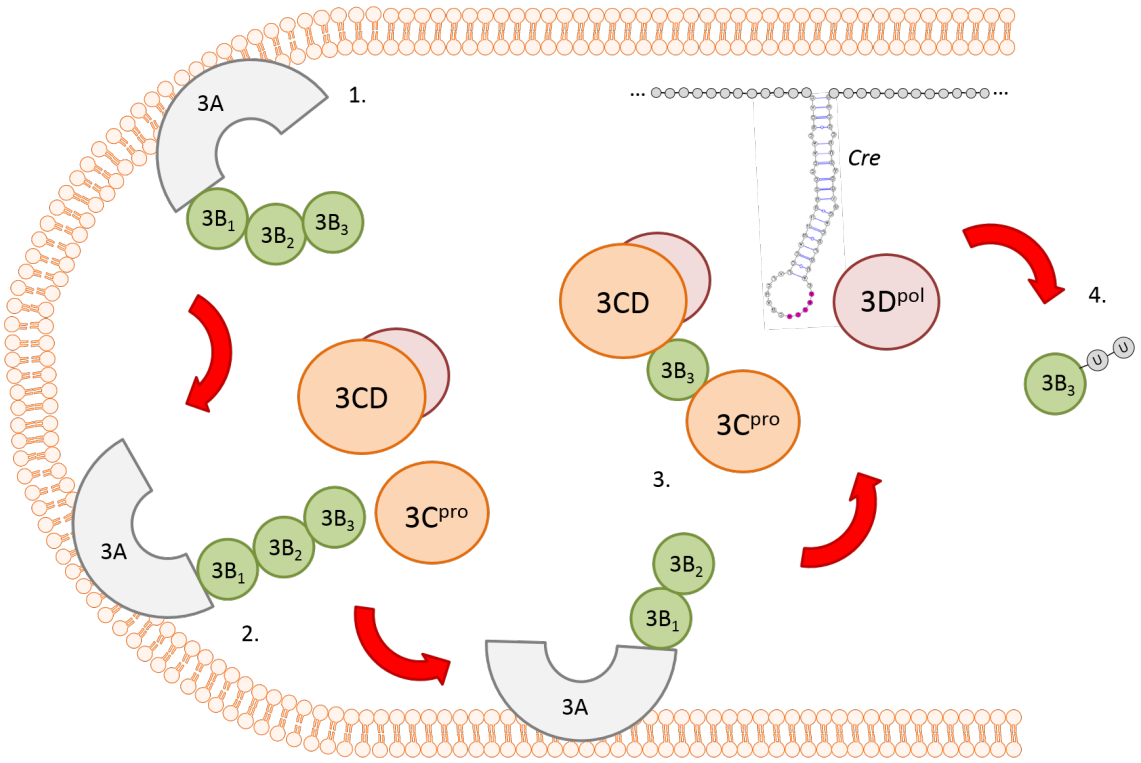

Figure 6.2: Suggested FMDV 3B uridylylation pathway.
1. 3A transmembrane domain anchors 3AB123 precursor into cellular membranes. 2. A 3B molecule is removed from 3AB123 by the action of $3C^{pro}$. 3. $3CD/3C^{pro}$ traffic mature 3B molecule to the cre from uridylylation by $3D^{pol}$. 4. Mature 3BpUpU can be used as a primer for replication.

With the disruption of the cleavage boundaries between the P3 proteins, it was discovered that individual 3B proteins could be fused together without negative impacts to replication. This was exploited to use the C-terminal 3B molecule as a carrier for epitope tag insertion. It is possible that these epitope tags could be tolerated without the necessity for the fusion of 3Bs, however this line of investigation has not been explored. Insertion of V5 and HA epitope tags, allows the detection of 3B molecules from different replicons when co-transfected within the same cells, which could be used as tools to study the requirements of 3B complementation. We also suggest the use of these tags to help the isolation of newly synthesised RNA strands as, when transcribed, the 3B-tag protein would be covalently linked to the ends of both positive and negative strands. The tags could then be used in immunopre-

cipitation experiments to purify new strands away from the total RNA of the cell. This would solve a problem experienced for many years with the replicon system, where direct qPCR of transfected whole-cell RNA has proved troublesome, mainly due to the large amounts of input RNA present within the cell. The inclusion of a large MT tag has led towards a line of investigation with the end goal of visualising replicon protein within a transfected cell by electron microscopy. While these experiments are still in their infancy and were not feasible in the time frame of this PhD the project, they are planned to be carried forward in collaboration with Dr Juan Fontana (University of Leeds).

Experiments investigating the ability of replicons containing different amounts of active 3B genes to provide 3Bs in trans revealed surprising differences between FMDV and the single 3B containing PV. While PV could provide its single 3B in trans, when reduced to two copies FMDV could no longer do so (Figs 3.19 and 3.17). This suggests a potential reason behind why FMDV may contain three copies of the 3B gene. It is beneficial for a viral population to be able to share proteins within an infected cell. If, during error prone RdRp replication an inactivating mutation arises, rather than that RNA not being replicated, complementation of the inactivated protein may allow for replication to continue until reversion or compensatory mutations arise, further boosting viral evolution. It could be this complementation of proteins combined with error prone replication that helps make RNA viruses evolve so rapidly. For reasons so far unknown, it has been demonstrated here that to be able to share 3B proteins in trans, FMDV must possess all three copies. This could be due to the rate of FMDV replication, meaning that if less than three active 3Bs are present there are not enough mature 3B molecules present to spare some to an inactivated replicon. If so, experiments looking at replication kinetics should reveal differences in replication upon deletion or inactivation of 3B genes, with replication slowing as the number of copies of 3B reduce. These experiments could be performed using the replicon and Incucyte system, with an increase in resolution of replication kinetics observed by reducing the time points between scans from hourly to every

10 minutes (or similar). Experimentation using infectious virus could also be performed, for example by observing the rate of CPE using cell-killing assay similar to those described in Section 4.2.5.4, would provide further detail in the effects of 3B mutation on viral replication kinetics.

As described in the introduction (Section 1.2.6), determination between virus infected and vaccinated animals can be difficult. It is possible that the discoveries reported here could be used as DIVA factors within a vaccine. With both 3B deletions, not found in wild type isolates, or the inclusion of epitope tags being used as determinants for the identification of vaccinated animals. The epitope tags could, theoretically, be used to identify vaccinated animals through simple lateral flow devices. Where animal samples are applied to the flow device preloaded with antibodies against the epitope tags, vaccination could be indicated in a manner similar to that of a pregnancy test, with vaccinated animals generating a positive antibody-antigen complex. Likewise, the absence or 3Bs or presence of epitope tag sequences could be used in PCR detection to identify vaccinated animals.

6.0.2 The roles of the S-fragment during FMDV replication.

The S-fragment was one of the first RNA elements discovered when initial investigation into the FMDV genome revealed that this could be easily digested into two segments; the short, S-fragment and the long, L-fragment. Despite the presence of the S-fragment being known for a long time, little is published about its function in replication. Here the structure of the S-fragment was confirmed using SHAPE analysis to indeed match the predicted single large stem loop. However, SHAPE analysis has so far only been performed in a minimalistic environment. Therefore it would be interesting to perform further analysis to look for structural differences at different points in replication, such as within the infected/transfected cell or within the virus capsid. It would also be interesting to observe the structure of the same region in the negative sense strand, as the structure is predicted to be very different.

Truncation of the top of the S-fragment recapitulated isolates found in the SAT serotype with an approximate 70 nucleotide truncation. These showed replication in both virus and replicon, suggesting the entire length of the S-fragment was not needed to provide its function. Further truncations from the top revealed limitations in the size of the S-fragment that would still permit replication. Interestingly, disparity existed here between replicon and virus experiments, with the T-246 truncation showing no replication in replicon, but producing infectious virus (Fig 4.1). It could be that replication of this mutant was so low that GFP expression could not be detected, and only when converted into the infectious clone could replication be detected, amplified by the ability for virus to cause cell-cell spread.

Truncations from the bottom of the S-fragment revealed a similar phenotype, with replication decreasing with the size of the S-fragment. Interestingly, despite the B-195 mutation showing replication in replicon, it did not allow the recovery of virus. However, as discussed in section 4.2.5.3 additional attempts could prove successful.

Only one truncation revealed additional mutations elsewhere in the genome upon sequential passaging. Truncating 97 nucleotides from the bottom of the S-fragment established two populations of mutations, one within the 3D polymerase and the second producing three mutations within the 2C protein. Interestingly these mutations evolved separately, suggesting there are multiple ways to compensate for the reduction of S-fragment size. It is thought that both of these mutations could effect RNA binding, with the 2C mutation thought to lie within the RNA binding domain of 2C (predicted from PV), and the $3D^{pol}$ mutation thought to be in a residue that interacts with the phosphate backbone of RNA.

Investigation into the 3D mutation has revealed a potential increase in strand-switch recombination, further supporting a change in RNA binding potential, as weaker binding to the template might permit an increase in strand-dissociation and re-binding to a new template (Fig 4.22). Further experimentation into this is ongoing, with introduction of both the $3D^{pol}$ and 2C mutations into replicon, as well as

investigation into the potential RNA binding capabilities of the 2C protein.

Recent publications have highlighted that the size of the S-fragment could have effects on immune recognition and activation of the innate immune response within cells. Studies by Reider et al, suggest that shortening the S-fragment causes an increase in interferon stimulated genes[95]. I report an observation which could be explained by an interaction with the immune system of cells, where some truncations, such as the T-148 and B-195 mutants, could replicate in the IFN deficient BHK-21 cells, but not within the immunocompetent MDBK cells. Further experimentation is planned, involving the addition of exogenous IFN to stimulate the innate immune system of BHK-21 cells as well as using specific inhibitors to impair immune activation of MDBK cells.

Better characterisation of recovered virus containing truncated S-fragments is currently on going, including viral growth curves and experimentation investigating the rate of CPE induction. The reported work by Reider et al suggests that with truncation attenuation could also appear, providing a potentially good candidate for introduction into an attenuated virus vaccine. Trans-complementation of S-fragment deletions however, has revealed the ready recombination of this region, so if carried forward, additional mutations should be incorporated such as high fidelity polymerase mutations to limit the possibility of recombination and reversion to virulence[158].

My RNA-Protein pull down studies revealed potential S-fragment interacting proteins. Several of these related to cell-stress responses, such as HSP70 and stress granule protein TIA-1, as well as several cytoskeletal proteins (Fig 4.30). It is possible that the S-fragment acts to sequester cellular stress related proteins, preventing their function within the cell. If this were the case, it might be expected that observation of these proteins by immunofluorescent microscopy would reveal distinct localisation to sites of RNA replication. However, over time TIA-1 is seen to do the opposite, and instead disperses throughout the cell (Fig 4.34). To date, no ev-

idence of distinct FMDV replication complexes has been reported in BHK-21 cells. As such, this could indicate that replication is occurring in many locations throughout the cytoplasm, and not localised to a few key locations as has been reported for other viruses. It would be interesting to perform co-localisation experiments between replicon RNA and TIA-1 to support these hypotheses.

Several potential S-fragment binders are listed in this thesis, but so far, only one (TIA-1) has been investigated in any detail. Investigation into the other proteins using immunofluorescence or siRNA knock down could reveal key insights into the function of the S-fragment, as well as potentially idetenifying targets for the production of new cell lines.

The mass spectrometry data also allowed the identification of proteins that could be binding to the IRES. While there were many expected proteins listed here, including translation factors and ribosome subunits, there were several unreported potential binding factors. These include hnRNPD, otherwise known as AUF1. Professor Bert Semler (University of California), demonstrated that AUF1 is a negative regulator of PV replication, and upon siRNA knock down, PV replication has been seen to increase. Here, initial investigation of this with FMDV showed a small increase in FMDV replication. However, it is important to mention that this was not performed in cell lines optimal for FMDV replication and therefore the signal was very low. Production of BHK-21 AUF1 knock out cell lines could better investigate this interaction. If successful, and an increase in virus replication is seen, this cell line could be used to replace current BHK-21 cells used in the bulking of virus for use in inactivated virus vaccines.

6.0.3 Untangling the roles of the FMDV pseudoknots in viral replication

The pseudoknots are another example of poorly understood RNA structures in the $5'$ UTR of FMDV. FMDV pseudoknot structure is only predicted with no biochemical or functional data to suggest a function, or even their existence in the FMDV genome. SHAPE analysis was used to better refine these structures, which largely agreed with previous predictions, and was complemented using mutagenesis to disrupt the pseudoknot structures. As with the S-fragment, it would be interesting to perform these SHAPE reactions in a more authentic environment to see if the structures are maintained or switch between different conformities.

As with the 3B gene, the reason behind the multiplicity of the FMDV pseudoknots is unknown. Deletion experiments revealed that although replication could occur at wild-type levels with as few as one pseudoknot, competition experiments suggested that at least two were required to permit maintenance of WT replication. Replication was also detected upon the complete removal of the pseudoknots in the replicon system, although this was significantly reduced. Surprisingly, when introduced into the infectious clone, no virus could be recovered upon the deletion of all four pseudoknots.

The combination of replicon and virus data suggests a multiplicity in the roles of the pseudoknots in FMDV replication. The significant reduction of replicon replication suggests a role in enhancing genome replication, whilst the inability to recover infectious virus suggests another role in infectious virus production. Further investigation into the potential role in genome replication is ongoing, with initial evidence that new strand synthesis may be reduced more than first thought. While replicon experiments show a small, but significant, decrease in GFP production, ^{3}H incorporation assays showed a large difference between WT and replicons lacking all pseudoknots. The experiment shown in this thesis has so far only been done once, and as such

continued investigation is critical to explore this further. The ^{3}H signal of ΔPK 1234 replicon was at approximately 3 % of the WT so it could be hypothesised that this corresponds to the production of negative strands, with no positive strands actually produced (Fig 5.10). While this would complement exsisting virus studies, an explanation for the large amounts of GFP produced from this replicon is still lacking.

A recent publication by Logan et al, reported the potential presence of a packaging signal in the pseudoknot region[114]. While this does not map perfectly to any of the pseudoknots, it does suggest that this region might play an important role in the packaging of the viral genome into the capsid. Trans-encapsidation experiments are currently being designed by our collaborators at the Pirbright Institute to further investigate the potential role of the pseudoknots as a packaging signal.

6.0.4 Concluding Remarks

The work presented in this thesis provides a foundation for further studies into the replication of FMDV. Dissecting understudied RNA structures within the 5′ UTR has revealed novel insights into their potential function in the FMDV lifecycle. Further work is ongoing to elucidate their functional relevance. The overall aim is to develop cell-lines for efficient and safe production of infectious virus for vaccine manufacture.

Investigation into the multiplicity of the 3B genes has revealed the requirements of these proteins in the processing of the P3 region of FMDV. Future investigations will explore the mechanisms of 3B uridylylation, aiming to show any preference for the individual 3B proteins during viral replication as well as the precursors required throughout replication.

References

[1] Acevedo, A. and Andino, R. (2014). Library preparation for highly accurate population sequencing of RNA viruses. Nature Protocols *9*, 1760–1769.

[2] Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990). Basic local alignment search tool. Journal of Molecular Biology *215*, 403–410.

[3] Ao, D., Guo, H.-C., Sun, S.-Q., Sun, D.-H., Fung, T. S., Wei, Y.-Q., Han, S.-C., Yao, X.-P., Cao, S.-Z., Liu, D. X. and Liu, X.-T. (2015). Viroporin Activity of the Foot-and-Mouth Disease Virus Non-Structural 2B Protein. PLOS ONE *10*, e0125828.

[4] Arias, A., Arnold, J. J., Sierra, M., Smidansky, E. D., Domingo, E. and Cameron, C. E. (2008). Determinants of RNA-dependent RNA polymerase (in)fidelity revealed by kinetic analysis of the polymerase encoded by a foot-and-mouth disease virus mutant with reduced sensitivity to ribavirin. Journal of virology *82*, 12346–55.

[5] Arias, A., Perales, C., Escarmís, C. and Domingo, E. (2010a). Deletion Mutants of VPg Reveal New Cytopathology Determinants in a Picornavirus. PLoS ONE *5*, e10735.

[6] Arias, A., Perales, C., Escarmís, C. and Domingo, E. (2010b). Deletion Mutants of VPg Reveal New Cytopathology Determinants in a Picornavirus. PLoS ONE *5*, e10735.

[7] Arnold, J. J. and Cameron, C. E. (2000). Poliovirus RNA-dependent RNA polymerase (3D(pol)). Assembly of stable, elongation-competent complexes by using a symmetrical primer-template substrate (sym/sub). The Journal of biological chemistry *275*, 5329–36.

[8] Banerjee, R. and Dasgupta, A. (2001). Interaction of picornavirus 2C polypeptide with the viral negative-strand RNA. J. Gen. Virol. *82*, 2621–2627.

[9] Barton, D. J., O'Donnell, B. J. and Flanegan, J. B. (2001). 5' cloverleaf in poliovirus RNA is a cis-acting replication element required for negative-strand synthesis. The EMBO journal *20*, 1439–48.

[10] Bashirullah, A., Cooperstock, R. L. and Lipshitz, H. D. (2001). Spatial and temporal control of RNA stability. Proceedings of the National Academy of Sciences *98*, 7025–7028.

[11] Bazin, J., Baerenfaller, K., Gosai, S. J., Gregory, B. D., Crespi, M. and Bailey-Serres, J. (2017). Global analysis of ribosome-associated noncoding RNAs unveils new modes of translational regulation. Proceedings of the National Academy of Sciences of the United States of America *114*, E10018–E10027.

[12] Been, M. D. and Wickham, G. S. (1997). Self-cleaving ribozymes of hepatitis delta virus RNA. European journal of biochemistry *247*, 741–53.

[13] Belsham, G. J. and Brangwyn, J. K. (1990). A region of the 5' noncoding region of foot-and-mouth disease virus RNA directs efficient internal initiation of protein synthesis within cells: involvement with the role of L protease in translational control. Journal of virology *64*, 5389–5395.

[14] Belsham, G. J., McInerney, G. M. and Ross-Smith, N. (2000). Foot-and-mouth disease virus 3C protease induces cleavage of translation initiation factors eIF4A and eIF4G within infected cells. Journal of virology *74*, 272–80.

[15] Berger, H. G., Straub, O. C., Ahl, R., Tesar, M. and Marquardt, O. (1990). Identification of foot-and-mouth disease virus replication in vaccinated cattle by antibodies to non-structural virus proteins. Vaccine *8*, 213–6.

[16] Bevilacqua, P. C. and Blose, J. M. (2008). Structures, Kinetics, Thermodynamics, and Biological Functions of RNA Hairpins. The Annual Review of Physical Chemistry is online at *59*, 79–103.

[17] Bienz, K., Egger, D., Troxler, M. and Pasamontes, L. (1990). Structural organization of poliovirus RNA replication is mediated by viral proteins of the P2 genomic region. J. Virol. *64*, 1156–1163.

[18] Birtley, J. R., Knox, S. R., Jaulent, A. M., Brick, P., Leatherbarrow, R. J. and Curry, S. (2005a). Crystal Structure of Foot-and-Mouth Disease Virus 3C Protease: NEW INSIGHTS INTO CATALYTIC MECHANISM AND CLEAVAGE SPECIFICITY. Journal of Biological Chemistry *280*, 11520–11527.

[19] Birtley, J. R., Knox, S. R., Jaulent, A. M., Brick, P., Leatherbarrow, R. J. and Curry, S. (2005b). Crystal Structure of Foot-and-Mouth Disease Virus 3C Protease: NEW INSIGHTS INTO CATALYTIC MECHANISM AND CLEAVAGE SPECIFICITY. Journal of Biological Chemistry *280*, 11520–11527.

[20] Biswal, J. K., Subramaniam, S., Ranjan, R. and Pattnaik, B. (2016). Partial deletion of stem-loop 2 in the 3 untranslated region of foot-and-mouth disease virus identifies a region that is dispensable for virus replication. Archives of Virology *161*, 2285–2290.

[21] Bittle, J. L., Houghten, R. A., Alexander, H., Shinnick, T. M., Sutcliffe, J. G., Lerner, R. A., Rowlands, D. J. and Brown, F. (1982). Protection against foot-and-mouth disease by immunization with a chemically synthesized peptide predicted from the viral nucleotide sequence. Nature *298*, 30–3.

[22] Blyn, L. B., Swiderek, K. M., Richards, O., Stahl, D. C., Semler, B. L. and Ehrenfeld, E. (1996). Poly(rC) binding protein 2 binds to stem-loop IV of the poliovirus RNA 5' noncoding region: identification by automated liquid chromatography-tandem mass spectrometry. Proceedings of the National Academy of Sciences of the United States of America *93*, 11115–20.

[23] Bradrick, S. S., Dobrikova, E. Y., Kaiser, C., Shveygert, M. and Gromeier, M. (2007). Poly(A)-binding protein is differentially required for translation mediated by viral internal ribosome entry sites. RNA (New York, N.Y.) *13*, 1582–93.

[24] Bravo de Rueda, C., de Jong, M. C. M., Eblé, P. L. and Dekker, A. (2014). Estimation of the transmission of foot-and-mouth disease virus from infected sheep to cattle. Veterinary research *45*, 58.

[25] Breaker, R. R. (2012). New insight on the response of bacteria to fluoride. Caries research *46*, 78–81.

[26] Brierley, I., Pennell, S. and Gilbert, R. J. C. (2007). Viral RNA pseudoknots: versatile motifs in gene expression and replication. Nature Reviews Microbiology *5*, 598–610.

[27] Brown, C. C., Piccone, M. E., Mason, P. W., McKenna, T. S. and Grubman, M. J. (1996). Pathogenesis of wild-type and leaderless foot-and-mouth disease virus in cattle. Journal of virology *70*, 5638–41.

[28] Cammas, A. and Millevoi, S. (2016). RNA G-quadruplexes: emerging mechanisms in disease. Nucleic Acids Research *45*, gkw1280.

[29] Cao, Y., Lu, Z. and Liu, Z. (2016). Foot-and-mouth disease vaccines: progress and problems. Expert Review of Vaccines *15*, 783–789.

[30] Carocci, M. and Bakkali-Kassimi, L. (2012). The encephalomyocarditis virus. Virulence *3*, 351–67.

[31] Carrillo, C., Tulman, E. R., Delhon, G., Lu, Z., Carreno, A., Vagnozzi, A., Kutish, G. F. and Rock, D. L. (2005). Comparative genomics of foot-and-mouth disease virus. Journal of virology *79*, 6487–504.

[32] Castelló, A., Alvarez, E. and Carrasco, L. (2011). The multifaceted poliovirus 2A protease: regulation of gene expression by picornavirus proteases. Journal of biomedicine & biotechnology *2011*, 369648.

[33] Chapman, E. G., Moon, S. L., Wilusz, J. and Kieft, J. S. (2014). RNA structures that resist degradation by Xrn1 produce a pathogenic Dengue virus RNA. eLife *3*.

[34] Chase, A. J., Daijogo, S. and Semler, B. L. (2014). Inhibition of poliovirus-induced cleavage of cellular protein PCBP2 reduces the levels of viral RNA replication. Journal of virology *88*, 3192–201.

[35] Chawla, M., Oliva, R., Bujnicki, J. M. and Cavallo, L. (2015). An atlas of RNA base pairs involving modified nucleobases with optimal geometries and accurate energies. Nucleic Acids Research *43*, 6714–6729.

[36] Chen, S.-J. (2008). RNA folding: conformational statistics, folding kinetics, and ion electrostatics. Annual review of biophysics *37*, 197–214.

[37] Chong, P., Guo, M.-S., Lin, F. H.-Y., Hsiao, K.-N., Weng, S.-Y., Chou, A.-H., Wang, J.-R., Hsieh, S.-Y., Su, I.-J. and Liu, C.-C. (2012). Immunological and biochemical characterization of coxsackie virus A16 viral particles. PloS one *7*, e49973.

[38] Clarke, B., Brown, A., Currey, K., Newton, S., Rowlands, D. and Carroll, A. (1987). Potential secondary and tertiary structure in the genomic RNA of foot and mouth disease virus. Nucleic Acids Research *15*, 7067–7079.

[39] Clarke, B. E., Sangar, D. V., Burroughs, J. N., Newton, S. E., Carroll, A. R. and Rowlands, D. J. (1985). Two Initiation Sites for Foot-and-Mouth Disease Virus Polyprotein in vivo. Journal of General Virology *66*, 2615–2626.

[40] Cowan, K. M. and Graves, J. H. (1966). A third antigenic component associated with foot-and-mouth disease infection. Virology *30*, 528–40.

[41] Crawford, N. M. and Baltimore, D. (1983). Genome-linked protein VPg of poliovirus is present as free VPg and VPg-pUpU in poliovirus-infected cells. Proceedings of the National Academy of Sciences of the United States of America *80*, 7452–5.

[42] Curry, S., Fry, E., Blakemore, W., Abu-Ghazaleh, R., Jackson, T., King, A., Lea, S., Newman, J. and Stuart, D. (1997). Dissecting the roles of VP0 cleavage and RNA packaging in picornavirus capsid stabilization: the structure of empty capsids of foot-and-mouth disease virus. Journal of virology *71*, 9743–52.

[43] Dedepsidis, E., Kyriakopoulou, Z., Pliaka, V. and Markoulatos, P. (2010). Correlation between recombination junctions and RNA secondary structure elements in poliovirus Sabin strains. Virus Genes *41*, 181–191.

[44] Delli Ponti, R., Marti, S., Armaos, A. and Tartaglia, G. G. (2017). A high-throughput approach to profile RNA structure. Nucleic acids research *45*, e35.

[45] den Boon, J. A., Diaz, A. and Ahlquist, P. (2010). Cytoplasmic viral replication complexes. Cell host & microbe *8*, 77–85.

[46] Devaney, M. A., Vakharia, V. N., Lloyd, R. E., Ehrenfeld, E. and Grubman, M. J. (1988). Leader protein of foot-and-mouth disease virus is required for cleavage of the p220 component of the cap-binding protein complex. Journal of virology *62*, 4407–9.

[47] Domingo, E. and Holland, J. J. (1997). RNA VIRUS MUTATIONS AND FITNESS FOR SURVIVAL. Annual Review of Microbiology *51*, 151–178.

[48] Donnelly, M. L., Luke, G., Mehrotra, A., Li, X., Hughes, L. E., Gani, D. and Ryan, M. D. (2001). Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'. The Journal of general virology *82*, 1013–25.

[49] Duke, G. M., Hoffman, M. A. and Palmenberg, A. C. (1992). Sequence and structural elements that contribute to efficient encephalomyocarditis virus RNA translation. Journal of virology *66*, 1602–9.

[50] Duke, G. M., Osorio, J. E. and Palmenberg, A. C. (1990). Attenuation of Mengo virus through genetic engineering of the 5' noncoding poly(C) tract. Nature *343*, 474–6.

[51] Echeverri, A., Banerjee, R. and Dasgupta, A. (1998). Amino-terminal region of poliovirus 2C protein is sufficient for membrane binding. Virus research *54*, 217–23.

[52] Emerson, S. U., Nguyen, H. T., Torian, U., Mather, K. and Firth, A. E. (2013). An essential RNA element resides in a central region of hepatitis E virus ORF2. The Journal of general virology *94*, 1468–76.

[53] Falk, M. M., Grigera, P. R., Bergmann, I. E., Zibert, A., Multhaup, G. and Beck, E. (1990). Foot-and-mouth disease virus protease 3C induces specific proteolytic cleavage of host cell histone H3. Journal of virology *64*, 748–56.

[54] Falk, M. M., Sobrino, F. and Beck, E. (1992). VPg gene amplification correlates with infective particle formation in foot-and-mouth disease virus. Journal of virology *66*, 2251–60.

[55] Feng, S., Sekine, S., Pessino, V., Li, H., Leonetti, M. D. and Huang, B. (2017). Improved split fluorescent proteins for endogenous protein labeling. Nature Communications *8*, 370.

[56] Ferré-D'Amaré, A. R. (2010). The glmS ribozyme: use of a small molecule coenzyme by a gene-regulatory RNA. Quarterly Reviews of Biophysics *43*, 423–447.

[57] Ferrer-Orta, C., Arias, A., Perez-Luque, R., Escarmís, C., Domingo, E. and Verdaguer, N. (2004). Structure of Foot-and-Mouth Disease Virus RNA-dependent RNA Polymerase and Its Complex with a Template-Primer RNA. Journal of Biological Chemistry *279*, 47212–47221.

[58] Firth, A. E. (2014). Mapping overlapping functional elements embedded within the protein-coding regions of RNA viruses. Nucleic Acids Research *42*, 12425–12439.

[59] Forrest, S., Lear, Z., Herod, M. R., Ryan, M., Rowlands, D. J. and Stonehouse, N. J. (2014). Inhibition of the foot-and-mouth disease virus subgenomic replicon by RNA aptamers. Journal of General Virology *95*, 2649–2657.

[60] Forss, S. and Schaller, H. (1982). A tandem repeat gene in a picornavirus. Nucleic acids research *10*, 6441–50.

[61] Forss, S., Strebel, K., Beck, E. and Schaller, H. (1984). Nucleotide sequence and genome organization of foot-and-mouth disease virus. Nucleic acids research *12*, 6587–601.

[62] Fujita, H., Ohshima, T., Oishi, T., Aratani, S., Fujii, R., Fukamizu, A. and Nakajima, T. (2005). Relevance of nuclear localization and functions of RNA helicase A. International journal of molecular medicine *15*, 555–60.

[63] Gao, Y., Sun, S.-Q. and Guo, H.-C. (2016). Biological function of Foot-and-mouth disease virus non-structural proteins and non-coding elements. Virology journal *13*, 107.

[64] García-Nuñez, S., Gismondi, M. I., König, G., Berinstein, A., Taboga, O., Rieder, E., Martínez-Salas, E. and Carrillo, E. (2014). Enhanced IRES activity by the 3'UTR element determines the virulence of FMDV isolates. Virology *448*, 303–13.

[65] Gladue, D. P., O'Donnell, V., Baker-Bransetter, R., Pacheco, J. M., Holinka, L. G., Arzt, J., Pauszek, S., Fernandez-Sainz, I., Fletcher, P., Brocchi, E., Lu, Z., Rodriguez, L. L. and Borca, M. V. (2014). Interaction of foot-and-mouth disease virus nonstructural protein 3A with host protein DCTN3 is important for viral virulence in cattle. Journal of virology *88*, 2737–47.

[66] Gladue, D. P., O'Donnell, V., Baker-Branstetter, R., Holinka, L. G., Pacheco, J. M., Fernandez-Sainz, I., Lu, Z., Brocchi, E., Baxt, B., Piccone, M. E., Rodriguez, L. and Borca, M. V. (2012). Foot-and-mouth disease virus nonstructural protein 2C interacts with Beclin1, modulating virus replication. Journal of virology *86*, 12080–90.

[67] Gmyl, A. P., Belousov, E. V., Maslova, S. V., Khitrina, E. V., Chetverin, A. B. and Agol, V. I. (1999). Nonreplicative RNA recombination in poliovirus. Journal of virology *73*, 8958–65.

[68] Goodfellow, I., Chaudhry, Y., Richardson, A., Meredith, J., Almond, J. W., Barclay, W. and Evans, D. J. (2000). Identification of a cis-acting replication element within the poliovirus coding region. Journal of virology *74*, 4590–600.

[69] Gorbalenya, A. E., Koonin, E. V. and Wolf, Y. I. (1990). A new superfamily of putative NTP-binding domains encoded by genomes of small DNA and RNA viruses. FEBS letters *262*, 145–8.

[70] Grubman, M. J. and Baxt, B. (2004). Foot-and-mouth disease. Clinical microbiology reviews *17*, 465–93.

[71] Guarné, A., Hampoelz, B., Glaser, W., Carpena, X., Tormo, J., Fita, I. and Skern, T. (2000). Structural and biochemical features distinguish the foot-and-mouth disease virus leader proteinase from other papain-like enzymes 1 1Edited by R. Huber. Journal of Molecular Biology *302*, 1227–1240.

[72] Gultyaev, A. P. and Olsthoorn, R. C. L. (2010). A family of non-classical pseudoknots in influenza A and B viruses. RNA biology *7*, 125–9.

[73] Haimov, O., Sinvani, H. and Dikstein, R. (2015). Cap-dependent, scanning-free translation initiation mechanisms. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms *1849*, 1313–1318.

[74] Harrich, D., Mavankal, G., Mette-Snider, A. and Gaynor, R. B. (1995). Human immunodeficiency virus type 1 TAR element revertant viruses define RNA structures required for efficient viral gene expression and replication. Journal of virology *69*, 4906–13.

[75] Harris, K. S., Reddigari, S. R., Nicklin, M. J., Hämmerle, T. and Wimmer, E. (1992). Purification and characterization of poliovirus polypeptide 3CD, a proteinase and a precursor for RNA polymerase. Journal of virology *66*, 7481–9.

[76] Harris, T. J. R. and Brown, F. (1976). The Location of the Poly(C) Tract in the RNA of Foot-and-Mouth Disease Virus. Journal of General Virology *33*, 493–501.

[77] Hentze, M. W., Caughman, S. W., Rouault, T. A., Barriocanal, J. G., Dancis, A., Harford, J. B. and Klausner, R. D. (1987). Identification of the iron-responsive element for the translational regulation of human ferritin mRNA. Science (New York, N.Y.) *238*, 1570–3.

[78] Herod, M. R., Ferrer-Orta, C., Loundras, E.-A., Ward, J. C., Verdaguer, N., Rowlands, D. J. and Stonehouse, N. J. (2016). Both cis and trans Activities of Foot-and-Mouth Disease Virus 3D Polymerase Are Essential for Viral RNA Replication. Journal of Virology *90*, 6864–6883.

[79] Herod, M. R., Gold, S., LaseckaDykes, L., Wright, C., Ward, J. C., McLean, T. C., Forrest, S., Jackson, T., Tuthill, T. J., Rowlands, D. J. and Stonehouse, N. J. (2017). Genetic economy in picornaviruses: Foot-and-mouth disease virus replication exploits alternative precursor cleavage pathway. PLOS Pathogens *13*, e1006666.

[80] Herold, J. and Andino, R. (2000). Poliovirus requires a precise 5' end for efficient positive-strand RNA synthesis. Journal of virology *74*, 6394–400.

[81] Herold, J. and Andino, R. (2001). Poliovirus RNA Replication Requires Genome Circularization through a Protein–Protein Bridge. Molecular Cell *7*, 581–591.

[82] Hill, M. K., Hooker, C. W., Harrich, D., Crowe, S. M. and Mak, J. (2001). Gag-Pol Supplied in trans Is Efficiently Packaged and Supports Viral Function in Human Immunodeficiency Virus Type 1. Journal of Virology *75*, 6835–6840.

[83] Holbrook, S. R. and Kim, S.-H. (1997). RNA crystallography. Biopolymers *44*, 3–21.

[84] Homan, P. J., Favorov, O. V., Lavender, C. A., Kursun, O., Ge, X., Busan, S., Dokholyan, N. V. and Weeks, K. M. (2014). Single-molecule correlated chemical probing of RNA. Proceedings of the National Academy of Sciences of the United States of America *111*, 13858–63.

[85] Huo, X., Han, S., Wu, G., Latchoumanin, O., Zhou, G., Hebbard, L., George, J. and Qiao, L. (2017). Dysregulated long noncoding RNAs (lncRNAs) in hepatocellular carcinoma: implications for tumorigenesis, disease progression, and liver cancer stem cells. Molecular cancer *16*, 165.

[86] Hüttenhofer, A. and Vogel, J. (2006). Experimental approaches to identify non-coding RNAs. Nucleic acids research *34*, 635–46.

[87] Jablonski, S. A. and Morrow, C. D. (1995). Mutation of the aspartic acid residues of the GDD sequence motif of poliovirus RNA-dependent RNA polymerase results in enzymes with altered metal ion requirements for activity. Journal of virology *69*, 1532–9.

[88] Jansen, R.-P. (2001). mRNA localization: message on the move. Nature Reviews Molecular Cell Biology *2*, 247–256.

[89] Kanda, T., Ozawa, M. and Tsukiyama-Kohara, K. (2016). IRES-mediated translation of foot-and-mouth disease virus (FMDV) in cultured cells derived from FMDV-susceptible and -insusceptible animals. BMC veterinary research *12*, 66.

[90] Karabiber, F., McGinnis, J. L., Favorov, O. V. and Weeks, K. M. (2013). QuShape: Rapid, accurate, and best-practices quantification of nucleic acid probing information, resolved by capillary electrophoresis. RNA *19*, 63–73.

[91] Kertesz, M., Wan, Y., Mazor, E., Rinn, J. L., Nutter, R. C., Chang, H. Y. and Segal, E. (2010). Genome-wide measurement of RNA secondary structure in yeast. Nature *467*, 103–107.

[92] Kieft, J. S., Rabe, J. L. and Chapman, E. G. (2015). New hypotheses derived from the structure of a flaviviral Xrn1-resistant RNA: Conservation, folding, and host adaptation. RNA biology *12*, 1169–77.

[93] King, A. M., Sangar, D. V., Harris, T. J. and Brown, F. (1980). Heterogeneity of the genome-linked protein of foot-and-mouth disease virus. Journal of virology *34*, 627–34.

[94] King, A. M. Q., Blakemore, W. E., Ellard, F. M., Drew, J. and Stuart, D. I. (1999). Evidence for the role of His-142 of protein 1C in the acid-induced disassembly of foot-and-mouth disease virus capsids. Journal of General Virology *80*, 1911–1918.

[95] Kloc, A., Diaz-San Segundo, F., Schafer, E. A., Rai, D. K., Kenney, M., de los Santos, T. and Rieder, E. (2017). Foot-and-mouth disease virus 5'-terminal S fragment is required for replication and modulation of the innate immune response in host cells. Virology *512*, 132–143.

[96] Knight-Jones, T. and Rushton, J. (2013). The economic impacts of foot and mouth disease – What are they, how big are they and where do they occur? Preventive Veterinary Medicine *112*, 161–173.

[97] Knowles, N. J., Samuel, A. R., Davies, P. R., Kitching, R. P. and Donaldson, A. I. (2001). Outbreak of foot-and-mouth disease virus serotype O in the UK caused by a pandemic strain. The Veterinary record *148*, 258–9.

[98] Komar, A. A. and Hatzoglou, M. (2011). Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states. Cell cycle (Georgetown, Tex.) *10*, 229–40.

[99] Kurisu, T., Tanaka, T., Ishii, J., Matsumura, K., Sugimura, K., Nakatani, T. and Kawashima, H. (2006). Expression and function of human steroid receptor RNA activator in prostate cancer cells: role of endogenous hSRA protein in androgen receptor-mediated transcription. Prostate Cancer and Prostatic Diseases *9*, 173–178.

[100] Lanz, R. B., McKenna, N. J., Onate, S. A., Albrecht, U., Wong, J., Tsai, S. Y., Tsai, M. J. and O'Malley, B. W. (1999). A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex. Cell *97*, 17–27.

[101] Lawrence, P. and Rieder, E. (2009a). Identification of RNA helicase A as a new host factor in the replication cycle of foot-and-mouth disease virus. Journal of virology *83*, 11356–66.

[102] Lawrence, P. and Rieder, E. (2009b). Identification of RNA helicase A as a new host factor in the replication cycle of foot-and-mouth disease virus. Journal of virology *83*, 11356–66.

[103] Lawson, M. A. and Semler, B. L. (1992). Alternate poliovirus nonstructural protein processing cascades generated by primary sites of 3C proteinase cleavage. Virology *191*, 309–320.

[104] Lee, C.-G. and Hurwitzs, J. (1992). A New RNA Helicase Isolated from HeLa Cells That Catalytically Translocates in the 3' to 5' Direction*. The Journal of Biological Chemistry *267*, 4398–4407.

[105] Leontis, N. B., Stombaugh, J. and Westhof, E. (2002). The non-Watson-Crick base pairs and their associated isostericity matrices. Nucleic acids research *30*, 3497–531.

[106] Leppek, K., Das, R. and Barna, M. (2018). Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. Nature reviews. Molecular cell biology *19*, 158–174.

[107] Lescar, J., Soh, S., Lee, L. T., Vasudevan, S. G., Kang, C. and Lim, S. P. (2018). The Dengue Virus Replication Complex: From RNA Replication to Protein-Protein Interactions to Evasion of Innate Immunity. Advances in experimental medicine and biology *1062*, 115–129.

[108] Li, D., Lei, C., Xu, Z., Yang, F., Liu, H., Zhu, Z., Li, S., Liu, X., Shu, H. and Zheng, H. (2016). Foot-and-mouth disease virus non-structural protein 3A inhibits the interferon-β signaling pathway. Scientific Reports *6*, 21888.

[109] Li, H. and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics *26*, 589–595.

[110] Liu, Y., Franco, D., Paul, A. V. and Wimmer, E. (2007). Tyrosine 3 of Poliovirus Terminal Peptide VPg(3B) Has an Essential Function in RNA Replication in the Context of Its Precursor Protein, 3AB. Journal of Virology *81*, 5669–5684.

[111] Liu, Y., Zhu, Z., Zhang, M. and Zheng, H. (2015). Multifunctional roles of leader protein of foot-and-mouth disease viruses in suppressing host antiviral responses. Veterinary research *46*, 127.

[112] Lobert, P. E., Escriou, N., Ruelle, J. and Michiels, T. (1999). A coding RNA sequence acts as a replication signal in cardioviruses. Proceedings of the National Academy of Sciences of the United States of America *96*, 11560–5.

[113] Logan, G., Freimanis, G. L., King, D. J., Valdazo-González, B., Bachanek-Bankowska, K., Sanderson, N. D., Knowles, N. J., King, D. P. and Cottam, E. M. (2014). A universal protocol to generate consensus level genome sequences for foot-and-mouth disease virus and other positive-sense polyadenylated RNA viruses using the Illumina MiSeq. BMC Genomics *15*, 828.

[114] Logan, G., Newman, J., Wright, C. F., Lasecka-Dykes, L., Haydon, D. T., Cottam, E. M. and Tuthill, T. J. (2017). Deep sequencing of foot-and-mouth disease virus reveals RNA sequences involved in genome packaging. Journal of virology *92*.

[115] Lopez de Quinto, S. (2002). IRES-driven translation is stimulated separately by the FMDV 3'-NCR and poly(A) sequences. Nucleic Acids Research *30*, 4398–4405.

[116] Lopez de Quinto, S., Sáiz, M., de la Morena, D., Sobrino, F. and Martínez-Salas, E. (2002). IRES-driven translation is stimulated separately by the FMDV 3'-NCR and poly(A) sequences. Nucleic Acids Research *30*, 4398–4405.

[117] López-Manríquez, E., Vashist, S., Ureña, L., Goodfellow, I., Chavez, P., Mora-Heredia, J. E., Cancio-Lonches, C., Garrido, E. and Gutiérrez-Escolano, A. L. (2013). Norovirus genome circularization and efficient replication are facilitated by binding of PCBP2 and hnRNP A1. Journal of virology *87*, 11371–87.

[118] Loundras, E.-A., Herod, M. R., Harris, M. and Stonehouse, N. J. (2016). Foot-and-mouth disease virus genome replication is unaffected by inhibition of type III phosphatidylinositol-4-kinases. The Journal of general virology *97*, 2221–30.

[119] Lowry, K., Woodman, A., Cook, J. and Evans, D. J. (2014). Recombination in enteroviruses is a biphasic replicative process involving the generation of greater-than genome length 'imprecise' intermediates. PLoS pathogens *10*, e1004191.

[120] Martínez-Salas, E., Francisco-Velilla, R., Fernandez-Chamorro, J., Lozano, G. and Diaz-Toledano, R. (2015a). Picornavirus IRES elements: RNA structure and host protein interactions. Virus Research *206*, 62–73.

[121] Martínez-Salas, E., Francisco-Velilla, R., Fernandez-Chamorro, J., Lozano, G. and Diaz-Toledano, R. (2015b). Picornavirus IRES elements: RNA structure and host protein interactions. Virus research *206*, 62–73.

[122] Mathews, D. H., Moss, W. N. and Turner, D. H. (2010). Folding and finding RNA secondary structure. Cold Spring Harbor perspectives in biology *2*, a003665.

[123] Matzke, M. A. and Birchler, J. A. (2005). RNAi-mediated pathways in the nucleus. Nature Reviews Genetics *6*, 24–35.

[124] McCown, P. J., Winkler, W. C. and Breaker, R. R. (2012). Mechanism and distribution of glmS ribozymes. Methods in molecular biology (Clifton, N.J.) *848*, 113–29.

[125] McDonald, S., Block, A., Beaucourt, S., Moratorio, G., Vignuzzi, M. and Peersen, O. B. (2016). Design of a Genetically Stable High Fidelity Coxsackievirus B3 Polymerase That Attenuates Virus Growth in Vivo. The Journal of biological chemistry *291*, 13999–4011.

[126] Mellor, E. J. C., Brown, F. and Harris, T. J. R. (1985a). Analysis of the Secondary Structure of the Poly(C) Tract in Foot-and-Mouth Disease Virus RNAs. Journal of General Virology *66*, 1919–1929.

[127] Mellor, E. J. C., Brown, F. and Harris, T. J. R. (1985b). Analysis of the Secondary Structure of the Poly(C) Tract in Foot-and-Mouth Disease Virus RNAs. Journal of General Virology *66*, 1919–1929.

[128] Mercer, T. R., Gerhardt, D. J., Dinger, M. E., Crawford, J., Trapnell, C., Jeddeloh, J. A., Mattick, J. S. and Rinn, J. L. (2012). Targeted RNA sequencing reveals the deep complexity of the human transcriptome. Nature Biotechnology *30*, 99–104.

[129] Merrick, W. C. (2004). Cap-dependent and cap-independent translation in eukaryotic systems. Gene *332*, 1–11.

[130] Mignone, F., Gissi, C., Liuni, S. and Pesole, G. (2002). Untranslated regions of mRNAs. Genome biology *3*, REVIEWS0004.

[131] Minskaia, E. and Ryan, M. D. (2013). Protein coexpression using FMDV 2A: effect of linker residues. BioMed research international *2013*, 291730.

[132] Mirsafian, H., Manda, S. S., Mitchell, C. J., Sreenivasamurthy, S., Ripen, A. M., Mohamad, S. B., Merican, A. F. and Pandey, A. (2016). Long non-coding RNA expression in primary human monocytes. Genomics *108*, 37–45.

[133] Mirzayan, C. and Wimmer, E. (1992). Genetic analysis of an NTP-binding motif in poliovirus polypeptide 2C. Virology *189*, 547–55.

[134] Moffat, K., Howell, G., Knox, C., Belsham, G. J., Monaghan, P., Ryan, M. D. and Wileman, T. (2005). Effects of Foot-and-Mouth Disease Virus Nonstructural Proteins on the Structure and Function of the Early Secretory Pathway: 2BC but Not 3A Blocks Endoplasmic Reticulum-to-Golgi Transport. Journal of Virology *79*, 4382–4395.

[135] Morris, D. R. and Geballe, A. P. (2000). Upstream open reading frames as regulators of mRNA translation. Molecular and cellular biology *20*, 8635–42.

[136] Moss, W. N., Dela-Moss, L. I., Priore, S. F. and Turner, D. H. (2012). The influenza A segment 7 mRNA 3' splice site pseudoknot/hairpin family. RNA biology *9*, 1305–10.

[137] Moye, A. L., Porter, K. C., Cohen, S. B., Phan, T., Zyner, K. G., Sasaki, N., Lovrecz, G. O., Beck, J. L. and Bryan, T. M. (2015). Telomeric G-quadruplexes are a substrate and site of localization for human telomerase. Nature Communications *6*, 7643.

[138] Muckenthaler, M. U., Rivella, S., Hentze, M. W. and Galy, B. (2017). A Red Carpet for Iron Metabolism. Cell *168*, 344–361.

[139] Murray, K. E. and Barton, D. J. (2003). Poliovirus CRE-dependent VPg uridylylation is required for positive-strand RNA synthesis but not for negative-strand RNA synthesis. Journal of virology *77*, 4739–50.

[140] Nayak, A., Goodfellow, I. G. and Belsham, G. J. (2005a). Factors Required for the Uridylylation of the Foot-and-Mouth Disease Virus 3B1, 3B2, and 3B3 Peptides by the RNA-Dependent RNA Polymerase (3Dpol) In Vitro. Journal of Virology *79*, 7698–7706.

[141] Nayak, A., Goodfellow, I. G. and Belsham, G. J. (2005b). Factors required for the Uridylylation of the foot-and-mouth disease virus 3B1, 3B2, and 3B3 peptides by the RNA-dependent RNA polymerase (3Dpol) in vitro. Journal of virology *79*, 7698–706.

[142] Nayak, A., Goodfellow, I. G., Woolaway, K. E., Birtley, J., Curry, S. and Belsham, G. J. (2006). Role of RNA structure and RNA binding activity of foot-and-mouth disease virus 3C protein in VPg uridylylation and virus replication. Journal of virology *80*, 9865–75.

[143] Newton, S. E., Carroll, A. R., Campbell, R. O., Clarke, B. E. and Rowlands, D. J. (1985a). The sequence of foot-and-mouth disease virus RNA to the 5 side of the poly(C) tract. Gene *40*, 331–336.

[144] Newton, S. E., Carroll, A. R., Campbell, R. O., Clarke, B. E. and Rowlands, D. J. (1985b). The sequence of foot-and-mouth disease virus RNA to the 5' side of the poly(C) tract. Gene *40*, 331–6.

[145] Novikova, I. V., Hennelly, S. P. and Sanbonmatsu, K. Y. (2012). Structural architecture of the human long non-coding RNA, steroid receptor RNA activator. Nucleic acids research *40*, 5034–51.

[146] O'Donnell, V. K., Pacheco, J. M., Henry, T. M. and Mason, P. W. (2001). Subcellular Distribution of the Foot-and-Mouth Disease Virus 3A Protein in Cells Infected with Viruses Encoding Wild-Type and Bovine-Attenuated Forms of 3A. Virology *287*, 151–162.

[147] Ogawa, T., Watanabe, Y., Meshi, T. and Okada, Y. (1991). Trans complementation of virus-encoded replicase components of tobacco mosaic virus. Virology *185*, 580–4.

[148] Pacheco, J. M., Henry, T. M., O'Donnell, V. K., Gregory, J. B. and Mason, P. W. (2003). Role of nonstructural proteins 3A and 3B in host range and pathogenicity of foot-and-mouth disease virus. Journal of virology *77*, 13017–27.

[149] Palazzo, A. F. and Lee, E. S. (2015). Non-coding RNA: what is functional and what is junk? Frontiers in genetics *6*, 2.

[150] Paolini, N. A., Moore, K. S., di Summa, F. M., Fokkema, I. F. A. C., 't Hoen, P. A. C. and von Lindern, M. (2018). Ribosome profiling uncovers selective mRNA translation associated with eIF2 phosphorylation in erythroid progenitors. PLOS ONE *13*, e0193790.

[151] Park, J.-H., Tark, D., Lee, K.-N., Lee, S.-Y., Ko, M.-K., Lee, H.-S., Kim, S.-M., Ko, Y.-J., Seo, M.-G., Chun, J.-E., Lee, M.-H. and Kim, B. (2016). Novel foot-and-mouth disease virus in Korea, July-August 2014. Clinical and experimental vaccine research *5*, 83–7.

[152] Pathak, H. B., Oh, H. S., Goodfellow, I. G., Arnold, J. J. and Cameron, C. E. (2008). Picornavirus Genome Replication. Journal of Biological Chemistry *283*, 30677–30688.

[153] Paul, A. V., Peters, J., Mugavero, J., Yin, J., van Boom, J. H. and Wimmer, E. (2003). Biochemical and genetic studies of the VPg uridylylation reaction catalyzed by the RNA polymerase of poliovirus. Journal of virology *77*, 891–904.

[154] Peixeiro, I., Silva, A. L. and Romão, L. (2011). Control of human beta-globin mRNA stability and its impact on beta-thalassemia phenotype. Haematologica *96*, 905–13.

[155] Pelletier, J. and Sonenberg, N. (1988). Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. Nature *334*, 320–325.

[156] Peng, Y., Leung, H. C. M., Yiu, S. M. and Chin, F. Y. L. (2012). IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics *28*, 1420–1428.

[157] Perera, R., Daijogo, S., Walter, B. L., Nguyen, J. H. C. and Semler, B. L. (2007). Cellular Protein Modification by Poliovirus: the Two Faces of Poly(rC)-Binding Protein. Journal of Virology *81*, 8919–8932.

[158] Pfeiffer, J. K. and Kirkegaard, K. (2003). A single mutation in poliovirus RNA-dependent RNA polymerase confers resistance to mutagenic nucleotide analogs via increased fidelity. Proceedings of the National Academy of Sciences *100*, 7289–7294.

[159] Plant, E. P. and Dinman, J. D. (2008). The role of programmed-1 ribosomal frameshifting in coronavirus propagation. Frontiers in bioscience : a journal and virtual library *13*, 4873–81.

[160] Polatnick, J. and Arlinghaus, R. B. (1967). Effect of actinomycin D on virus-induced ribonucleic acid polymerase formation in foot-and-mouth disease virus-infected baby hamster kidney cells. Journal of virology *1*, 1130–4.

[161] Ponjavic, J., Ponting, C. P. and Lunter, G. (2007). Functionality or transcriptional noise? Evidence for selection within long noncoding RNAs. Genome research *17*, 556–65.

[162] Reuter, J. S. and Mathews, D. H. (2010). RNAstructure: software for RNA secondary structure prediction and analysis. BMC Bioinformatics *11*, 129.

[163] Rieder, E., Bunch, T., Brown, F. and Mason, P. W. (1993a). Genetically engineered foot-and-mouth disease viruses with poly(C) tracts of two nucleotides are virulent in mice. Journal of virology *67*, 5139–45.

[164] Rieder, E., Bunch, T., Brown, F. and Mason, P. W. (1993b). Genetically engineered foot-and-mouth disease viruses with poly(C) tracts of two nucleotides are virulent in mice. Journal of virology *67*, 5139–45.

[165] Rieder, E., Paul, A. V., Kim, D. W., van Boom, J. H. and Wimmer, E. (2000). Genetic and biochemical studies of poliovirus cis-acting replication element cre in relation to VPg uridylylation. Journal of virology *74*, 10371–80.

[166] Rieder, E., Xiang, W., Paul, A. and Wimmer, E. (2003). Analysis of the cloverleaf element in a human rhinovirus type 14/poliovirus chimera: correlation of subdomain D structure, ternary protein complex formation and virus replication. The Journal of general virology *84*, 2203–16.

[167] Risco, C., Sanmartín-Conesa, E., Tzeng, W.-P., Frey, T., Seybold, V. and de Groot, R. (2012). Specific, Sensitive, High-Resolution Detection of Protein Molecules in Eukaryotic Cells Using Metal-Tagging Transmission Electron Microscopy. Structure *20*, 759–766.

[168] Roberts, P. J. and Belsham, G. J. (1995). Identification of Critical Amino Acids within the Foot-and-Mouth Disease Virus Leader Protein, a Cysteine Protease. Virology *213*, 140–146.

[169] Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G. and Mesirov, J. P. (2011). Integrative genomics viewer. Nature Biotechnology *29*, 24–26.

[170] Rodríguez, P. L. and Carrasco, L. (1995). Poliovirus protein 2C contains two regions involved in RNA binding activity. The Journal of biological chemistry *270*, 10105–12.

[171] Rodríguez Pulido, M., Sobrino, F., Borrego, B. and Sáiz, M. (2009). Attenuated foot-and-mouth disease virus RNA carrying a deletion in the 3' noncoding region can elicit immunity in swine. Journal of virology *83*, 3475–85.

[172] Romero-Barrios, N., Legascue, M. F., Benhamed, M., Ariel, F. and Crespi, M. (2018). Splicing regulation by long noncoding RNAs. Nucleic acids research *46*, 2169–2184.

[173] Rowlands, D. J., Harris, T. J. and Brown, F. (1978). More precise location of the polycytidylic acid tract in foot and mouth disease virus RNA. Journal of virology *26*, 335–43.

[174] Rust, R. C., Landmann, L., Gosert, R., Tang, B. L., Hong, W., Hauri, H.-P., Egger, D. and Bienz, K. (2001). Cellular COPII Proteins Are Involved in Production of the Vesicles That Form the Poliovirus Replication Complex. Journal of Virology *75*, 9808–9818.

[175] Ryan, M. D. and Drew, J. (1994). Foot-and-mouth disease virus 2A oligopeptide mediated cleavage of an artificial polyprotein. The EMBO journal *13*, 928–33.

[176] Ryan, M. D., King, A. M. Q. and Thomas, G. P. (1991). Cleavage of foot-and-mouth disease virus polyprotein is mediated by residues located within a 19 amino acid sequence. Journal of General Virology *72*, 2727–2732.

[177] Sáiz, M., Gómez, S., Martínez-Salas, E. and Sobrino, F. (2001). Deletion or substitution of the aphthovirus 3 NCR abrogates infectivity and virus replication. Journal of General Virology *82*, 93–101.

[178] Scott, L. G. and Hennig, M. (2008). RNA Structure Determination by NMR. Methods in molecular biology (Clifton, N.J.) *452*, 29–61.

[179] Serrano, P., Pulido, M. R., Sáiz, M. and Martínez-Salas, E. (2006). The 3' end of the foot-and-mouth disease virus genome establishes two distinct long-range RNA-RNA interactions with the 5' end region. Journal of General Virology *87*, 3013–3022.

[180] Shabman, R. S., Hoenen, T., Groseth, A., Jabado, O., Binning, J. M., Amarasinghe, G. K., Feldmann, H. and Basler, C. F. (2013). An upstream open reading frame modulates ebola virus polymerase translation and virus replication. PLoS pathogens *9*, e1003147.

[181] Shu, B. and Gong, P. (2016). Structural basis of viral RNA-dependent RNA polymerase catalysis and translocation. Proceedings of the National Academy of Sciences of the United States of America *113*, 4005–14.

[182] Simmonds, P. (2006). Recombination and selection in the evolution of picornaviruses and other Mammalian positive-stranded RNA viruses. Journal of virology *80*, 11124–40.

[183] Spealman, P., Naik, A. W., May, G. E., Kuersten, S., Freeberg, L., Murphy, R. F. and McManus, J. (2018). Conserved non-AUG uORFs revealed by a novel regression analysis of ribosome profiling data. Genome research *28*, 214–222.

[184] Staple, D. W. and Butcher, S. E. (2005). Pseudoknots: RNA Structures with Diverse Functions. PLoS Biology *3*, e213.

[185] Steil, B. P. and Barton, D. J. (2009). Cis-active RNA elements (CREs) and picornavirus RNA replication. Virus Research *139*, 240–252.

[186] Suhy, D. A., Giddings, T. H. and Kirkegaard, K. (2000). Remodeling the endoplasmic reticulum by poliovirus infection and by individual viral proteins: an autophagy-like origin for virus-induced vesicles. Journal of virology *74*, 8953–65.

[187] Svitkin, Y. V., Costa-Mattioli, M., Herdy, B., Perreault, S. and Sonenberg, N. (2007). Stimulation of picornavirus replication by the poly(A) tail in a cell-free extract is largely independent of the poly(A) binding protein (PABP). RNA (New York, N.Y.) *13*, 2330–40.

[188] Sweeney, T. R., Cisnetto, V., Bose, D., Bailey, M., Wilson, J. R., Zhang, X., Belsham, G. J. and Curry, S. (2010). Foot-and-mouth disease virus 2C is a hexameric AAA+ protein with a coordinated ATP hydrolysis mechanism. The Journal of biological chemistry *285*, 24347–59.

[189] Sweeney, T. R., Roqué-Rosell, N., Birtley, J. R., Leatherbarrow, R. J. and Curry, S. (2007). Structural and Mutagenic Analysis of Foot-and-Mouth Disease Virus 3C Protease Reveals the Role of the β-Ribbon in Proteolysis. Journal of Virology *81*, 115–124.

[190] Terhuja, M., Saravanan, P. and Tamilselvan, R. P. (2015). Comparative efficacy of virus like particle (VLP) vaccine of foot-and-mouth-disease virus (FMDV) type O adjuvanted with poly I:C or CpG in guinea pigs. Biologicals *43*, 437–443.

[191] Teterina, N. L., Gorbalenya, A. E., Egger, D., Bienz, K., Rinaudo, M. S. and Ehrenfeld, E. (2006). Testing the modularity of the N-terminal amphipathic helix conserved in picornavirus 2C proteins and hepatitis C NS5A protein. Virology *344*, 453–467.

[192] Tiley, L., King, A. M. Q. and Belsham, G. J. (2003). The foot-and-mouth disease virus cis-acting replication element (cre) can be complemented in trans within infected cells. Journal of virology *77*, 2243–6.

[193] Tulloch, F., Pathania, U., Luke, G. A., Nicholson, J., Stonehouse, N. J., Rowlands, D. J., Jackson, T., Tuthill, T., Haas, J., Lamond, A. I. and Ryan, M. D. (2014a). FMDV replicons encoding green fluorescent protein are replication competent. Journal of Virological Methods *209*, 35–40.

[194] Tulloch, F., Pathania, U., Luke, G. A., Nicholson, J., Stonehouse, N. J., Rowlands, D. J., Jackson, T., Tuthill, T., Haas, J., Lamond, A. I. and Ryan, M. D. (2014b). FMDV replicons encoding green fluorescent protein are replication competent. Journal of virological methods *209C*, 35–40.

[195] Tulloch, F., Stonehouse, N. J., Ward, J. C., Loundras, E.-A., Herod, M. R. and Rowlands, D. J. (2015a). Employing transposon mutagenesis to investigate foot-and-mouth disease virus replication. Journal of General Virology *96*, 3507–3518.

[196] Tulloch, F., Stonehouse, N. J., Ward, J. C., Loundras, E.-A., Herod, M. R. and Rowlands, D. J. (2015b). Employing transposon mutagenesis to investigate foot-and-mouth disease virus replication. Journal of General Virology *96*, 3507–3518.

[197] Tuthill, T. J., Groppelli, E., Hogle, J. M. and Rowlands, D. J. (2010). Picornaviruses. Current topics in microbiology and immunology *343*, 43–89.

[198] Ullmer, W. and Semler, B. L. (2018). Direct and Indirect Effects on Viral Translation and RNA Replication Are Required for AUF1 Restriction of Enterovirus Infections in Human Cells. mBio *9*, 01669–18.

[199] Valdazo-González, B., Timina, A., Scherbakov, A., Abdul-Hamid, N. F., Knowles, N. J. and King, D. P. (2013). Multiple introductions of serotype O foot-and-mouth disease viruses into East Asia in 2010-2011. Veterinary Research *44*, 1–12.

[200] van der Velden, A. W. and Thomas, A. A. (1999). The role of the 5' untranslated region of an mRNA in translation regulation during development. The international journal of biochemistry & cell biology *31*, 87–106.

[201] Vangaveti, S., Ranganathan, S. V. and Chen, A. A. (2017). Advances in RNA molecular dynamics: a simulator's guide to RNA force fields. Wiley Interdisciplinary Reviews: RNA *8*, e1396.

[202] Varani, G. (1995). Exceptionally Stable Nucleic Acid Hairpins. Annual Review of Biophysics and Biomolecular Structure *24*, 379–404.

[203] Vázquez, A. L., Alonso, J. M. and Parra, F. (2000). Mutation analysis of the GDD sequence motif of a calicivirus RNA-dependent RNA polymerase. Journal of virology *74*, 3888–91.

[204] Vincent ter Beek (2018). PigProgress - FMD returns to South Korea's pig herd.

[205] Wang, D., Fang, L., Li, K., Zhong, H., Fan, J., Ouyang, C., Zhang, H., Duan, E., Luo, R., Zhang, Z., Liu, X., Chen, H. and Xiao, S. (2012). Foot-and-mouth disease virus 3C protease cleaves NEMO to impair innate immune signaling. Journal of virology *86*, 9311–22.

[206] Westhof, E. and Auffinger, P. (2000). RNA Tertiary Structure. In Encyclopedia of Analytical Chemistry. John Wiley & Sons, Ltd Chichester, UK.

[207] White, J. P. and Lloyd, R. E. (2011). Poliovirus unlinks TIA1 aggregation and mRNA stress granule formation. Journal of virology *85*, 12442–54.

[208] White, K. A., Enjuanes, L. and Berkhout, B. (2011). RNA virus replication, transcription and recombination. RNA biology *8*, 182–3.

[209] Wilce, J. A., Leedman, P. J. and Wilce, M. C. J. (2002). RNA-Binding Proteins That Target the Androgen Receptor mRNA. IUBMB Life (International Union of Biochemistry and Molecular Biology: Life) *54*, 345–349.

[210] Wilkinson, K. A., Merino, E. J. and Weeks, K. M. (2006). Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. Nature protocols *1*, 1610–6.

[211] Woese, C. R., Winker, S. and Gutell, R. R. (1990). Architecture of ribosomal RNA: constraints on the sequence of "tetra-loops". Proceedings of the National Academy of Sciences of the United States of America *87*, 8467–71.

[212] Wölk, B., Büchele, B., Moradpour, D. and Rice, C. M. (2008). A dynamic view of hepatitis C virus replication complexes. Journal of virology *82*, 10519–31.

[213] Woodman, A., Arnold, J. J., Cameron, C. E. and Evans, D. J. (2016). Biochemical and genetic analysis of the role of the viral polymerase in enterovirus recombination. Nucleic acids research *44*, 6883–95.

[214] Xia, S., Ji, R. and Zhan, W. (2017). Long noncoding RNA papillary thyroid carcinoma susceptibility candidate 3 (PTCSC3) inhibits proliferation and invasion of glioma cells by suppressing the Wnt/β-catenin signaling pathway. BMC neurology *17*, 30.

[215] Xiao, Y., Chen, H.-Y., Wang, Y., Yin, B., Lv, C., Mo, X., Yan, H., Xuan, Y., Huang, Y., Pang, W., Li, X., Yuan, Y. A. and Tian, K. (2016). Large-scale production of foot-and-mouth disease virus (serotype Asia1) VLP vaccine in Escherichia coli and protection potency evaluation in cattle. BMC Biotechnology *16*, 56.

[216] Yan, Y., Rout, S. N., Kim, S.-H. and Samal, S. K. (2009). Role of Untranslated Regions of the Hemagglutinin-Neuraminidase Gene in Replication and Pathogenicity of Newcastle Disease Virus. Journal of Virology *83*, 5943–5946.

[217] Young, S. K. and Wek, R. C. (2016). Upstream Open Reading Frames Differentially Regulate Gene-specific Translation in the Integrated Stress Response. The Journal of biological chemistry *291*, 16927–35.

[218] Zacharias, M. and Hagerman, P. J. (1995). Bulge-Induced Bends in RNA: Quantification by Transient Electric Birefringence RNA splicing; human immunodeficiency virus (HIV). J. Mol. Biol *247*, 486–500.

[219] Zhao, X., Patton, J. R., Ghosh, S. K., Fischel-Ghodsian, N., Shen, L. and Spanjaard, R. A. (2007). Pus3p- and Pus1p-Dependent Pseudouridylation of Steroid Receptor RNA Activator Controls a Functional Switch that Regulates Nuclear Receptor Signaling. Molecular Endocrinology *21*, 686–699.

[220] Zhou, Z., Mogensen, M. M., Powell, P. P., Curry, S. and Wileman, T. (2013). Foot-and-mouth disease virus 3C protease induces fragmentation of the Golgi compartment and blocks intra-Golgi transport. Journal of virology *87*, 11721–9.

[221] Zhu, S., Jones, M. K., Hickman, D., Han, S., Reeves, W. and Karst, S. M. (2016a). Norovirus antagonism of B-cell antigen presentation results in impaired control of acute infection. Mucosal immunology *9*, 1559–1570.

[222] Zhu, Z., Wang, G., Yang, F., Cao, W., Mao, R., Du, X., Zhang, X., Li, C., Li, D., Zhang, K., Shu, H., Liu, X. and Zheng, H. (2016b). Foot-and-Mouth Disease Virus Viroporin 2B Antagonizes RIG-I-Mediated Antiviral Effects by Inhibition of Its Protein Expression. Journal of Virology *90*, 11106–11121.

[223] Zibert, A., Maass, G., Strebel, K., Falk, M. M. and Beck, E. (1990). Infectious foot-and-mouth disease virus derived from a cloned full-length cDNA. Journal of virology *64*, 2467–73.

[224] Zuker, M. (1989). On finding all suboptimal foldings of an RNA molecule. Science (New York, N.Y.) *244*, 48–52.