Fluorescence imaging and modelling of virus structure and infection

Andrew McMahon

St Cross College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Hilary 2023

Abstract

Diseases caused by viruses represent a serious global health concern. In this thesis, I present novel methods for labelling, imaging and analysing virus particles. I also describe mathematical and computational models that can be used to describe biological structures, such as viruses.

In studies of viral structure, pleomorphism, and particularly filamentous morphologies, are frequently overlooked.To characterise and better understand viral filaments, I developed and optimised an imaging protocol for individual virus particles that captured the heterogeneity of viral shape. I also produced analysis pipelines that took this image data to study the structure of viruses at the individual virion level in order to find an overall picture of the populations of virions that exist. These pipelines investigated virus size, patterning of surface proteins and the location of viral genome. I also demonstrated a novel virus labelling strategy that specifically labels virus that will aid in the future study of viral assembly and infection.

I also used mathematical modelling to further understand the role of viral morphologies. I modelled virus particles as a pressure vessel to study engineering stress in relation to virus, bacteria, and cellular morphology.

Together these advancements will further our ability to study viruses and our knowledge of virus structure.

Fluorescence imaging and modelling of virus structure and infection



Andrew McMahon

St Cross College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hilary 2023

Acknowledgements

The first and largest thank you must go to my supervisors, Nicole Robb and Achillefs Kapanidis who put up with me for 4 years, supervised me, and harnessed the wanderings of my mind to produce this thesis. Secondly, those I had the pleasure to work closest with: Becky, Ben, Dani, Hafez, Nic, and all members of the Robb and Kapanidis labs. To my family and friends who have kept me sane through the DPhil through talks, walks, meals, trialling, TV watching, and rowing- especially those who constantly put up with me through lockdown (Andrew, John, Maggie, Peter, and Steph) and those who frequently took the time to hang out (Chris, Jimmy, Kat, Liv, Marcus, and Oli). A shout-out to George Clooney and the boys in the boat for my brief sojourn into the world of film. Finally, to Rand: thank you for everything.

Abstract

Diseases caused by viruses represent a serious global health concern. In this thesis, I present novel methods for labelling, imaging and analysing virus particles. I also describe mathematical and computational models that can be used to describe biological structures, such as viruses.

In studies of viral structure, pleomorphism, and particularly filamentous morphologies, are frequently overlooked.To characterise and better understand viral filaments, I developed and optimised an imaging protocol for individual virus particles that captured the heterogeneity of viral shape. I also produced analysis pipelines that took this image data to study the structure of viruses at the individual virion level in order to find an overall picture of the populations of virions that exist. These pipelines investigated virus size, patterning of surface proteins and the location of viral genome. I also demonstrated a novel virus labelling strategy that specifically labels virus that will aid in the future study of viral assembly and infection.

I also used mathematical modelling to further understand the role of viral morphologies. I modelled virus particles as a pressure vessel to study engineering stress in relation to virus, bacteria, and cellular morphology.

Together these advancements will further our ability to study viruses and our knowledge of virus structure.

Contents

1 Introduction **1**

1.1 Introduction 1
1.2 Virus structure 3
1.3 Influenza infection cycle 7
1.4 Fluorescence microscopy 9
1.5 Models used for viral studies 15
1.6 Chapter organisation 19

2 Imaging of fixed virus particles **21**

2.1 Introduction 22
2.2 The development of the imaging protocol 23
2.3 High-throughput analysis of influenza virus pleomorphism 31
2.4 Methods and materials 38
2.5 Discussion and conclusion 44

3 Viral organisation and virus-cell interactions **49**

3.1 Introduction 50
3.2 The spatial organisation of viruses 51
3.3 Fixed cell imaging of virus-cell interactions 58
3.4 Labelling of virus particles through the infection of labelled cells . . 61
3.5 Methods and materials 65
3.6 Discussion and conclusion 68

4 Surface stress as a selective pressure for filamentous virus formation **71**

4.1 Introduction 71
4.2 Theory 74
4.3 Results 76
4.4 Methods and materials 84
4.5 Discussion and conclusion 86

5 Conclusions **91**

5.1 Conclusions 91
5.2 Future work 92

Appendices

A Publications **97**

B Numerical modelling of SARS-CoV-2 reinfection **99**
B.1 Introduction 99
B.2 Methods 101
B.3 Results 108
B.4 Discussion and conclusion 112

C Computational simulations showing an improvement to dense fitting using intensity differences **117**
C.1 Introduction 117
C.2 1D simulations 118
C.3 2D simulations 118
C.4 Conclusion 120

D The collection of aerosolised virus particles **121**
D.1 Introduction 121
D.2 Results 122
D.3 Conclusions 123

References **125**

It is a strange fate that we should suffer so much fear and doubt over so small a thing. Such a little thing.

— Borimir in J.R.R. Tolkein's *The Lord of the Rings: The Fellowship of the Ring*

1 Introduction

Contents

1.1 Introduction 1
1.2 Virus structure 3
Influenza 3
SARS-CoV-2 5
Respiratory syncytial virus (RSV) 6
1.3 Influenza infection cycle 7
1.4 Fluorescence microscopy 9
Fluorescence 9
Immunofluorescence 10
TIRF (Total Internal Reflection Fluorescence) microscopy . . . 11
Super-resolution microscopy 12
1.5 Models used for viral studies 15
Monte Carlo simulations 15
Pressure vessel models 16
SIR (Susceptible, Infected, Recovered) Model 17
1.6 Chapter organisation 19

1.1 Introduction

Viruses are microscopic infectious particles (typically less than 300 nm in diameter) that replicate within cells of host organisms and cause viral diseases. These diseases result in widespread human and animal morbidity and mortality each year, as well as economic and social impacts. Viruses are composed of an outer protein coat that

surrounds the DNA or RNA core. In some viruses, this is all contained within a lipid envelope derived from the membrane of the cells of the host.

The use of microscopy to directly image viral particles can be split into two fields: the study of the virus structure and mechanisms, and viral diagnostics. Mechanisms of particular interest include virus/host interactions such as viral entry, propagation, and viral egress. Through these studies, novel viral targets can be identified. Viral diagnostics is essential as a way of safeguarding a population as viral spread from known, infected individuals can be reduced through distancing/isolation measures and targeted antivirals for specific viral infections. By reducing false bacterial diagnoses, viral diagnostics also reduces prescribed antibiotics and so has a place in the combat against antibiotic-resistance in bacteria.

The imaging of viruses by conventional light microscopy is limited as virus particles are generally less than the $\sim$250 nm resolution limit given by the Rayleigh criterion for visible light. This means that it is not possible to image the internal structure and organisation of viruses and so, historically, the imaging of viruses has been done with electron microscopy (EM) [1]. Since its invention in the 1930s, EM rapidly became the gold standard for fixed virus imaging, providing a resolution at the length scales of $\sim$0.1 nm [2]. The utility of EM has only increased with the invention of cryo EM, which has a resolution greater than any other imaging modality. However, this method has limited utility in the study of the dynamic processes of viruses and in the study of complex biological environments as the preparation of samples for electron microscopy is difficult and high energy electron beams damage live cells. Furthermore, specific structures can only be identified if they have a dissimilar size or shape to other structures and specific labelling with gold-conjugated antibodies is inefficient [3].

In order to compliment EM microscopy, new techniques have been developed, including super-resolution fluorescence microscopy (SRFM) [4]. SRFM is a general term given to optical-based microscopy methods that bypass the resolution limit imposed by the diffraction of light. These techniques enable the imaging of single molecules which enables the heterogeneity within a population to be studied. There

are numerous SRFM methods; many of these are capable of live cell imaging, and have a large field of view (FOV), enabling large-number statistics to be collected. Also, labelling samples with fluorescent molecules (fluorophores) for imaging enables SRFM to image specific structures in multiple colours giving functional information. The trade off with these techniques is a lower spatial resolution than EM.

In this thesis, I aim to use SRFM, image analysis and a mathematical and computational models to study filamentous morphologies of influenza in order to better understand their structure and organisation. Their structure is of interest as filaments may represent the more infectious and pathogenic forms of influenza, as well as offering insights into virus assembly and mechanisms of infection [5]. I also show an extension to the susceptible, infected, recovered (SIR) model which allows the study of viral reinfection which is important in the understanding of novel, epidemic virus infections early in their discovery.

1.2 Virus structure

Due to viruses evolving in different hosts, models of virus structure do not generalise over every virus (though families of viruses do exist). As such, novel features of each virus/viral family need to be studied individually. Experiments in this thesis have focused on influenza A, SARS-CoV-2 and respiratory syncytial virus (RSV).

Influenza

Of all of the major, potentially deadly, uncontrolled diseases in the world, influenza, caused by the influenza virus, is among the most common and deadly [6] [7]. Annual epidemics result in 290,000 to 650,000 respiratory deaths per year [6]), and rapid mutation can result in additional circulating strains, such as the 'swine flu' pandemic strain ((H1N1)pdm09) of 2009. Despite this, many aspects of the structure and cellular interactions of the virus are poorly understood. Their small size makes influenza viruses difficult to study - a typical spherical virion is of the order of 100 nm in diameter [8]. However, they are important to study as they result in

significant morbidity and mortality in both human and animal populations, resulting in a large healthcare and economic burden to countries around the world.

Influenza A virus (IAV) is an RNA virus- a virus with RNA as its genetic material, and the only species in the Alphainfluenzavirus genus of the Orthomyxoviridae viral family. IAV forms many different pleiomorphic virions [9] - spherical virions that are approximately 100nm in diameter to filaments of similar widths but reaching many micrometers in length. Filamentous strains have frequently been overlooked in virus research, both because laboratory-passaged viruses tend to be spherical (with the filaments more typical of clinical isolates) and also because ultra-centrifugation and other purification and storage procedures tend to damage filaments [10].

Influenza virus particles are surrounded by a lipid bilayer, in which two varieties of surface glycoproteins are embedded, the hemagglutinin (HA) and neuraminidase (NA) (Fig. 1.1) [11]. HA is a trimeric glycoprotein that has a role as both an attachment factor and a membrane fusion protein that conformationally changes to an active state at a low pH [12] [13], while NA cleaves sialic acid and is required for influenza to be released from its host cell [14]. A membrane channel (M2) is also embedded in the lipid bilayer. It is a homotetrameric protein responsible for acidifying the interior of the virus, allowing virions to escape endosomes and infect cells [15]. Electron micrograph images of spherical viruses have suggested that there are 375 surface protein spikes, of which approximately one seventh are NA and the rest are HA[11] [16]. The distributions of the HA and NA glycoproteins over the virion surface are not entirely random; with studies showing that small clusters of NA are formed within the more abundant HA [11] which is uniformly distributed across the surface of the virus [17]; that NA has a tendency to cluster at one of the poles of filaments, [18] [11] [19] [20] [21] and that HA and NA proteins may alternate on the filament surface [18]. Beneath the lipid bilayer lies the structural matrix protein M1 - in filaments, this has been observed to have a helical structure [21] and has been implicated as a major determinant of the pleomorphic structure of influenza [22]. Inside the M1 core lies the viral genome, which is comprised of 8 viral RNA segments, each of which is bound by multiple nucleoproteins (NP)

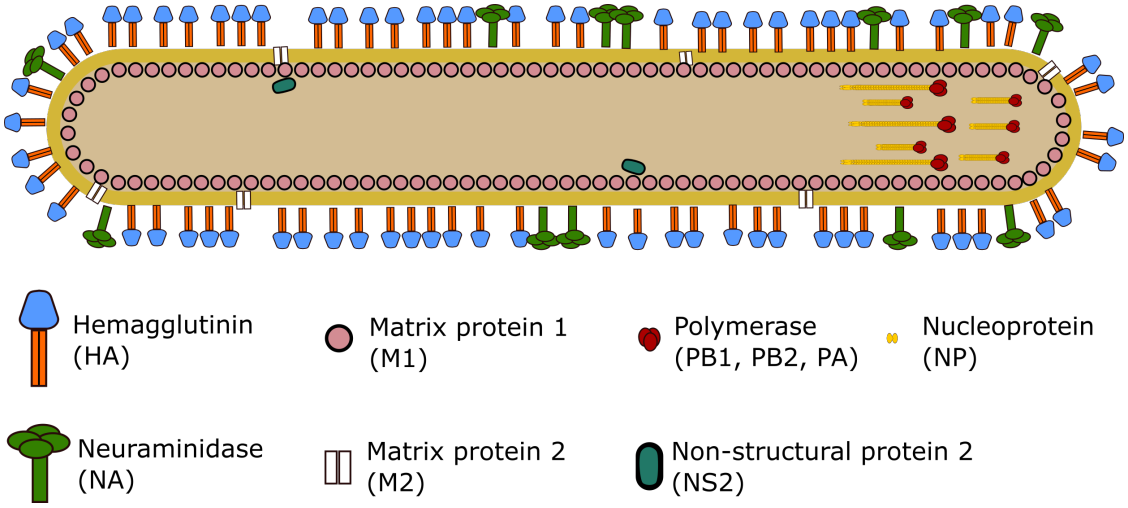


Figure 1.1: Schematic of the protein structure of filamentous influenza.

and the RNA polymerase (formed of the subunits PB1, PB2 and PA) to make ribonucleoprotein (RNP) complexes (Fig. 1.1). There is some suggestion that RNP complexes cluster at one end of filamentous virus particles [18] [19] [20] [21] [23]. Some filamentous virions have been observed to have large bulges at one end known as Archetti bodies [24]; previous electron microscopy images have indicated that these bulges do not appear to contain RNPs [21] and their exact purpose in viral replication remains unknown. NS2 is a non-structural protein and is responsible for mediating the nuclear export of viral RNA; it is not clear whether it is present in filamentous virus particles [25].

SARS-CoV-2

Covid-19, caused by the SARS-CoV-2 virus of the genus Coronavirus in the family of Coronaviridae, is a major virus currently in circulation within the population. The virus emerged in 2019 and by 2020, 4% of scientific research involved the virus [26] due to the number of deaths caused by the Covid-19 pandemic and the international effects of the pandemic. Virions have only been seen in spherical morphologies and their size is in the range of 50 to 140nm. [27] [28] [29] [30] It is important to study SARS-CoV-2 due to the severity of the pandemic that the virus has caused, and also to understand why SARS-CoV-2 and other coronaviruses

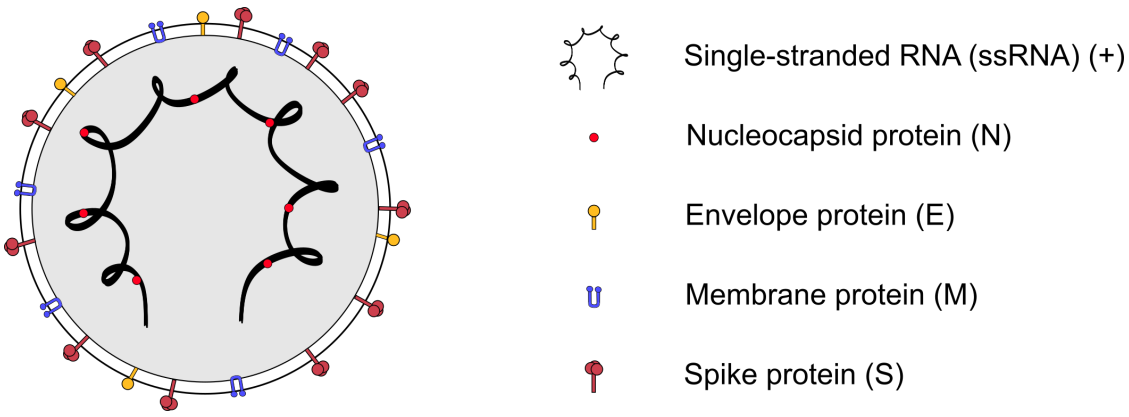


Figure 1.2: Schematic of the protein structure of SARS-CoV-2.

in the past two decades (MERS and SARS) commonly result in mortality while previous common coronaviruses (e.g. HKU, 229E and hCoV-OC43) do not.

SARS-CoV-2 viral particles are formed of a lipid envelope into which a spike (S) protein is embedded. The S protein is responsible for membrane fusion and allows endocytosis into cells and the infection of an organism. Also embedded in the envelope are the envelope (E) protein and the membrane (M) protein. They are responsible for entry, replication and viral assembly with the M protein being the more prevalent and important for the budding of new viral particles [31]. The genome is contained in the viral membrane in positive-sense single-stranded RNA which is bundled by the nucleocapsid (N) protein [32] (Fig. 1.2).

Respiratory syncytial virus (RSV)

RSV of the genus Orthopneumovirus within the family Pneumoviridae, is a major cause of lower respiratory tract infections in the young, elderly and immunocompromised. Infections are the number one cause of pneumonia and bronchiolitis in children under one year of age, and it has been estimated that there are approximately 60 million cases of RSV per year worldwide resulting in over 250 000 deaths [33][34].

The virus is formed from 11 structural and non-structural proteins [35]. The membrane contains 3 membrane proteins: the receptor attachment glycoprotein (G) which attaches the virus to the host cell; a short hydrophobic (SH) protein

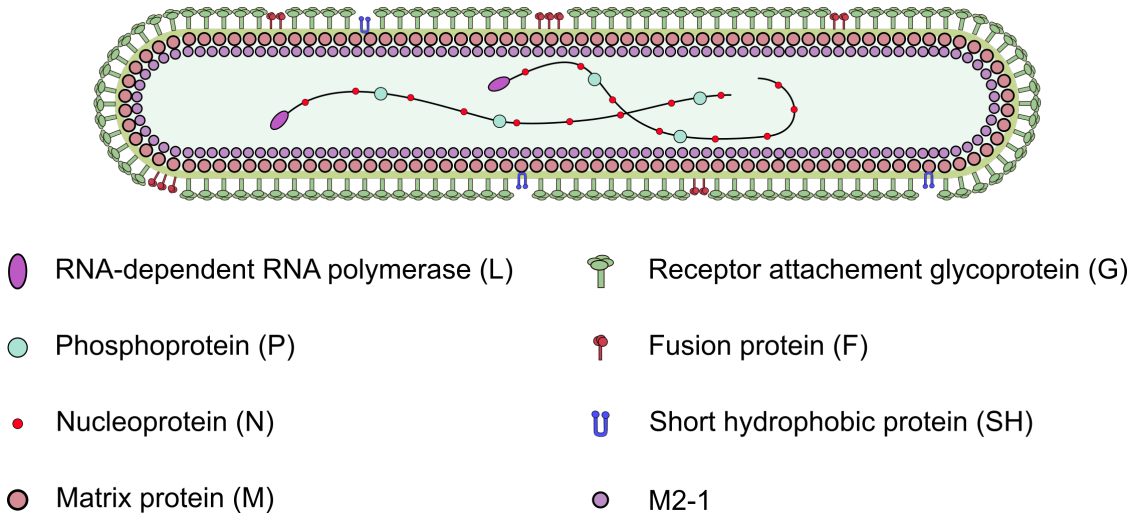


Figure 1.3: Schematic of the protein structure of RSV.

which forms an ion channel, and the fusion protein (F) which fuses the virus with the host. The F protein expressed on the surface of infected cells is able to fuse with other virions, forming multinucleated syncytia- the major characteristic of RSV virus infection [36]. The RNA is protected by and replicated by the ribonucleoprotein complex (RNP) formed of the nucleoprotein (N), phosphoprotein (P), and RNA-dependent RNA polymerase (L) (Fig. 1.3).

1.3 Influenza infection cycle

The influenza viral infection cycle describes the way influenza particles enter host cells and hijack the host machinery to create new virions (Fig. 1.4) [37].

The infection cycle of influenza particles begins with the virion viral entry. The HA of the influenza particle binds to the sialic acid of the host cell [38]. At this point, spherical virions usually undergo receptor-mediated endocytosis and enter in an endosome [37] while filamentous virions undergo macropinocytosis and enter as an intact filament in macropinasome [39].

Low pH in the late-endosomal compartments lead to fusion of the endosome and membrane; the opening of M2 ion channels, and the acidification of the viral core, allowing RNA to enter the cell nucleus [40]. In filaments, the low pH appears to fragment the filaments, releasing the RNA.

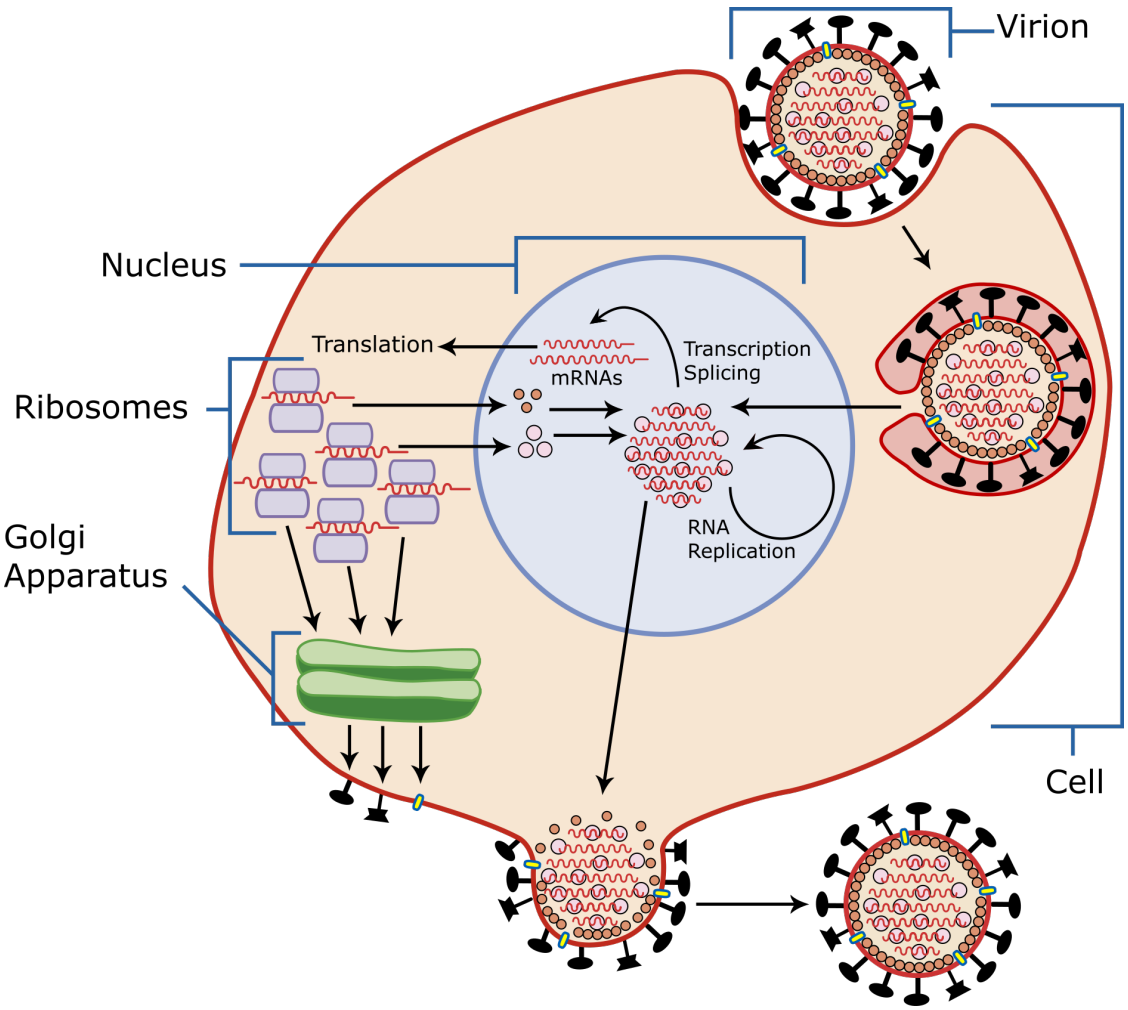


Figure 1.4: The virus cycle for influenza particles. In order to replicate, virions enter the host cells; the vRNPs are replicated and transcribed in the nucleus, and the virus assembles and buds from the host plasma membrane. Figure edited from Microbiology [46].

Inside the nucleus, negative sense viral RNA (vRNA) is transcribed to positive sense vRNA using the CAP structure from cellular mRNA [41]. This is translated to form the viral proteins and negative sense RNA is replicated to form new RNPs which are transported to the cell membrane.

Cells require HA, NA and M2 to form a viable viral particle [42]. These proteins are transported to the host cell membrane at the budding site. The absence of the M2 tail appears to produce filamentous virion growth [43]. M1 closes off the budding particle [44], forming a helical structure in filamentous particles [45]. Finally, NA cleaves the sialic acid residues, allowing the virus particle to bud.

1.4 Fluorescence microscopy

Fluorescence microscopes use light collected from fluorescence to study materials. They are often used in biological studies due to their ability to give both structural and functional information and the possibility of carrying out fluorescence live cell imaging.

Fluorescence

Fluorescence is a phenomenon whereby light is emitted from a molecule when an electron relaxes from an exited singlet state (a state in which all electrons are spin paired) to the singlet ground state (see Fig. 1.5 [47]). It typically occurs in the nanoseconds timescale. After 'absorbtion' of a high energy photon, electrons can relax to a lower excited state before returning to the ground state. The lower energy difference results in a longer wavelength photon and a Stokes shift between the incident and emitted radiation. Also shown in Fig. 1.5 is the process of phosphorescence: a photon is emitted from the excited triplet (three possible values of spin along a given axis) state to the ground state after an intersystem crossing. As an excited triplet state has electrons with the same total spin, the crossing would result in particles of the same spin state and is a forbidden transition- it is slow as it requires further vibrational relaxation. Phosphorescence lifetimes are typically of the order of milliseconds.

Fluorescence is useful as a method of indirectly visualising structures, molecules and reactions of interest using fluorescent labels. This includes immunofluorescence (see below); click chemistry: the use of a simple chemical reaction to join two molecules; quantum dots: nanometer sized semiconductors that luminesce due to the promotion of an electron from the valence band to the conduction band; fluorescent dyes and fluorescence in situ hybridisation (FISH) probes: short DNAs that are complimentary to DNA sites of interest.

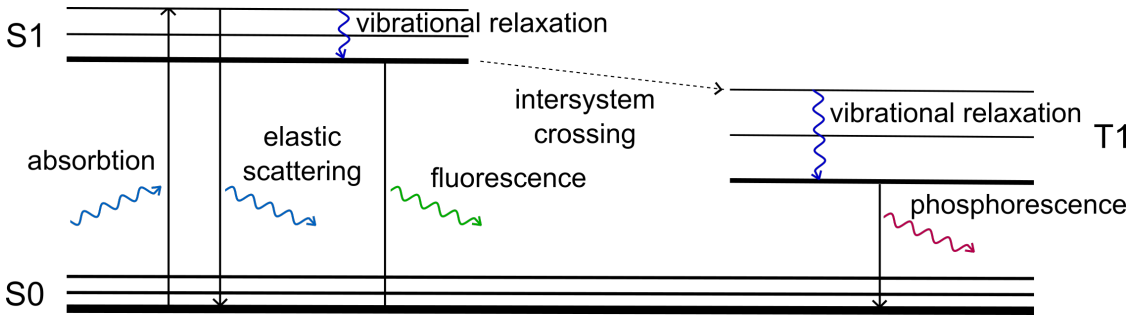


Figure 1.5: The Jablonski diagram for fluorescence emission. An absorbed photon may be absorbed by a fluorophore, causing the transition from the S_0 state to the S_1 state. By emitting a photon (fluorescence), the molecule may return to the S_0 state; this happens on the order of nanoseconds. If instead, it undergoes the forbidden transition from S_1 to the triplet state, T_1, and then decays; the molecule has undergone phosphorescence which occurs on the order of milliseconds.

Immunofluorescence

Antigens, including proteins (e.g. the SARS-CoV-2 spike protein), peptides, polysaccharides, lipids and nucleic acids among other biological molecules, cause immune systems to produce antibodies. By fluorescently labelling these antibodies, any antigens can be specifically labelled for fluorescence imaging, this is called immunofluorescence. Direct labelling occurs when there is a single antibody that is attached to a fluorophore; while indirect has 2 antibodies, one specific against the antigen of interest and the other specific against the animal the primary antibody was grown in (see Fig. 1.6).

The affinity of an antibody describes the strength of the binding to the target of interest and is expressed through the dissociation equilibrium constant- the ratio of the dissociation rate to the association rate. The affinity affects the labelling efficiency in immunofluorescence and so, impacts the uses of an antibody. They can be used in live cell studies though antibodies can affect the function of labelled structures [48] [49].

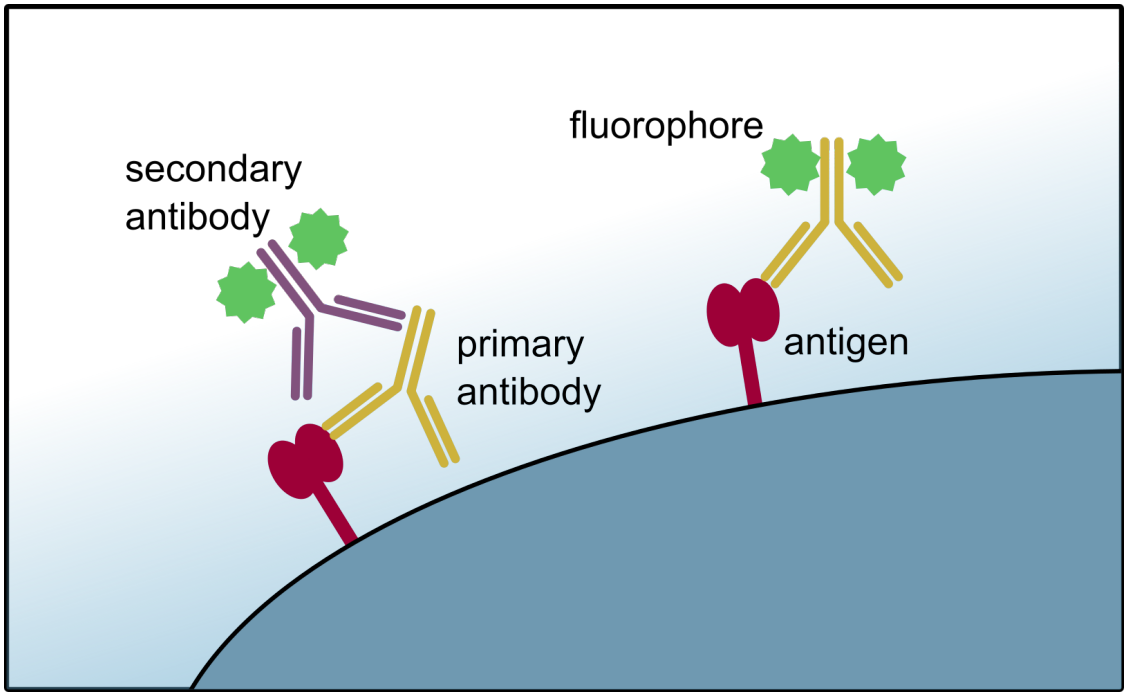


Figure 1.6: Schematic of the workings of immunofluorescence. The antigen of interest is fluorescently labelled using antibodies with fluorophores attached. The labelling can either be direct; whereby there is only one antibody which is fluorescently labelled, or indirect; when there are two antibodies, one specific against the target antigen and the other with an antigenic target that includes the primary antibody.

TIRF (Total Internal Reflection Fluorescence) microscopy

Total internal reflection fluorescence (TIRF) microscopy [50] utilises the process of total internal reflection to only illuminate a very thin section of a sample close to the surface interface. When light is incident upon an interface at greater than a critical angle, the light is reflected and none of the light is refracted into the sample. The critical angle is given by:

$$sin(\theta_{critical}) = n_r/n_i$$

Where n_r is the refractive index of the 2nd medium and and n_i is the refractive index of the incident medium. When this happens, although no light is refracted into the 2nd medium, an evanescent wave is produced which penetrates into the second medium (Fig. 1.7). The amplitude of an evanescent wave decays exponentially and so it only illuminates a very small volume (approximately 100 nm deep) and excites a thin section of the sample of interest. It has been widely used to image

exocytosis and endocytosis using fluorescently labelled vescicles that appear when exocytosis has occured, or disappear when vescicles are endocytosed [51].

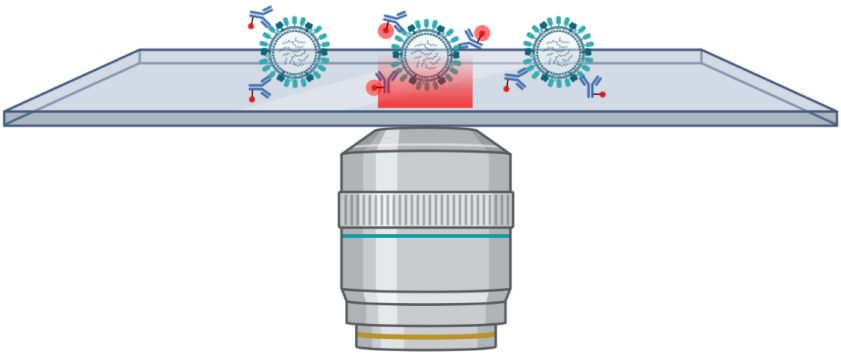


Figure 1.7: A schematic of the imaging of the influenza virus through TIRF microscopy. The illumination of a sample is beyond the critical angle for the interface of materials. This sets up an exponentially decaying evanescent wave into the sample so that only a small distance into the sample is illuminated, decreasing noise in the image. [52]

Super-resolution microscopy

Due to their use of visible light, the resolution of fluorescent microscopes is ordinarily limited to the diffraction limit of light [53]. However, various techniques have overcome this limit. Structured illumination microscopy (SIM) does this using sampling in the frequency domain. Some techniques use the nonlinear response of fluorophores (fluorescent chemical compounds) to excitations; examples of this include stimulated emission depletion (STED) and reversible saturable optical fluorescence transitions (RESOLFT). Other methods make use of the stochastic photoswitchable or photoactivatable nature of fluorophores to provide temporal resolution; examples of this include direct stochastic optical reconstruction microscopy (dSTORM) and photoactivated localisation microscopy (PALM). The techniques, how they overcome the resolution limit of light and their advantages and disadvantages are given in table 1.1 [54][55][56][57][58][59][60][61][62][63].

Technique	Method of resolution increase	Pros	Cons
STED (Stimulated emission depletion)	Minimisation of the illumination area at the focal point by selectively deactivating nearby fluorophores. This area is then scanned across a sample.	• Greater sectioning in thick samples.	• High laser power leads to toxicity to biological materials • High complexity of microscope required • Limited usable fluorescent dyes
SOFI (Super-resolution optical fluctuation imaging)	Correlation of light from a single fluctuating photon across the field of view to pinpoint the location of a photon.	• Usable in conjunction with other super resolution techniques • Rejects out of focus light	• Lowers temporal resolution as fluctuations in time are needed.
PALM/STORM/GSD (Photo-activated localization microscopy/Stochastic optical reconstruction microscopy/Ground state depletion)	Uses the blinking of fluorescent particles to fit functions to the intensity profile that can give sub pixel accuracy. The image is obtained from the sum over all the frames of the acquisition.	• Highest increase in the resolution of the image • Can be performed on most setups	• Long acquisition time per image
SIM (Structured illumination microscopy)	A spatially patterned light is used to excite the sample which produces interference patterns which can be deconvolved to produce the image.	• Lower laser power is less toxic to biological materials • Almost continuous acquisition is possible	• Only a modest increase in resolution • Specialised microscope required
PAINT (Points accumulation for imaging in nanoscale topography)	Fluorophores are non-fluorescent before binding to the target at which point they fluoresce. Multiple binding events are accumulated to form the image.	• Multiplexing DNA allows many structures to be labelled in the same image	• Requires many binding and unbinding events so has a long acquisition time

		• Highest increase in the resolution of the image • Can be performed on most setups	
Localization by stepwise photobleaching	Localise fluorophores based on bleaching between frames.	• High resolution increase	• Requires optimisation based on dyes, sample and lasers • Not good with dense labelling • High laser powers used • Long acquisition time per image
Stepwise optical saturation (SOS)	Eliminates the lowest powers of the emission of a fluorophore by linear combinations of images with differing intensities	• No additional setup required	• Additional images at different intensities required per improvement • Modest resolution increase

Table 1.1: Super-resolution methods. A table of super-resolution techniques; how they break the resolution limit and their advantages and disadvantages.

dSTORM (direct Stochastic Optical Reconstruction Microscopy)

dSTORM is a form of microscopy that relies on the stochastic activation of only a small subset of the fluorophores that are used to label a sample (Fig. 1.8) [64]. By forcing the majority of fluorophores into a dark radical or a triplet state (see 1.5), only those not in the dark state can be activated. dSTORM makes use of the photophysics of certain fluorescent dyes, in conjunction with an oxidising and reducing system to reduce photobleaching and increase the population of molecules in the dark state. The fluorescent signal of each of the subset of fluorophores that are 'on' in each frame of the acquisition can be fit using a two-dimensional Gaussian function, localising them with nanometer accuracy. The localisations from all frames can then be summed to create a super-resolved image (Fig. 1.8). This technique greatly increases the resolution of an optical image without increasing the complexity of the imaging system used. The greatest resolution limitation comes from the precision to which individual fluorophores can be localised, which is compounded by other errors such as motion blur, drift, defocus effects and the size of fluorescent tags. Using dSTORM, sub 20 nm precision can often be achieved with the trade off coming from the drastically reduced temporal resolution (acquisitions take on the order of 10 minutes) which usually requires samples to be fixed.

1.5 Models used for viral studies

Models are useful as a way of simplifying real life and seeing how an idealised scenario would behave. By comparing models with experimental results, we can see if the models are valid and able to give predictions for real world behaviour. Or, if they produce incorrect results, we can see that there are more important effects or if the model needs refinement.

Monte Carlo simulations

Monte Carlo simulations use stochastics and random sampling to obtain relevant information about a complex physical system. By using the numerical mean of a

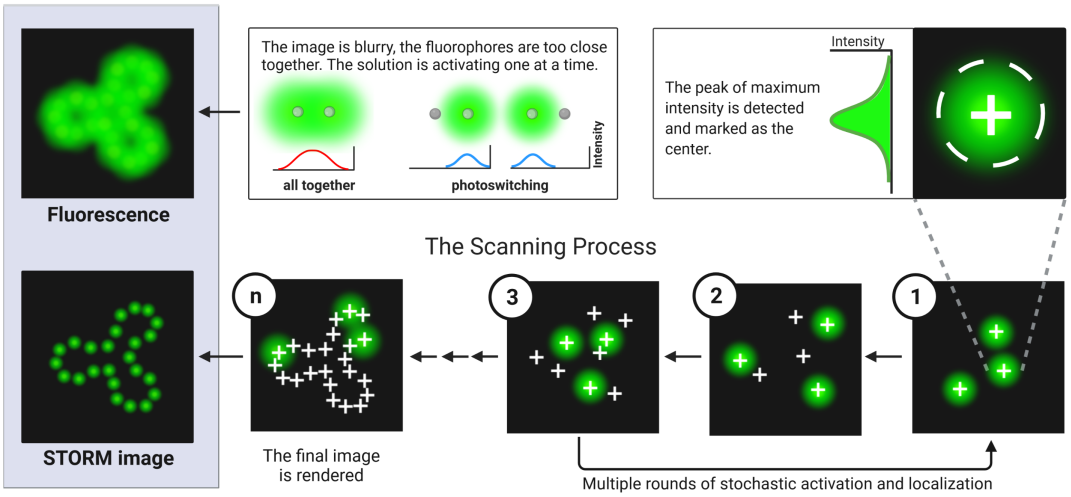


Figure 1.8: A flow diagram illustrating the acquisition of a STORM image. By performing multiple rounds of activating a subset of fluorophores in a sample and localising the centre of the fluorescence, a higher-resolution image can be built up over time, trading temporal resolution for spatial resolution. Edited from biorender.com [65]

number of simulations, the expected value of a problem is found via the law of large numbers. The larger the number of iterations of the simulation, the greater the accuracy to which the expectation value is found. An example using the probability of placing a point within a quadrant of radius 1 compared to a square of side length 1 to find the value of pi is shown in Fig. 1.9. The random sampling shown is capable of determining information about a real model (the value of pi from the area of a circle), with the precision increasing for a greater number of simulations completed.

Monte Carlo simulations were developed and used systematically during World War II while investigating neutron travel through shielding [66]. They are often used to investigate risk, stocks and sales [67] which have a large amount of uncertainty and can incorporate sensitivity analysis due to the probability space sampled.

Pressure vessel models

This is a simplistic model that finds the relationship between the stress in the outer surface of a container and the internal pressure at a cross section of a 3D hollow shape. As the body is at rest, we can equate the forces between the internal pressure and the shear stress which gives the relation between the 2 quantities:

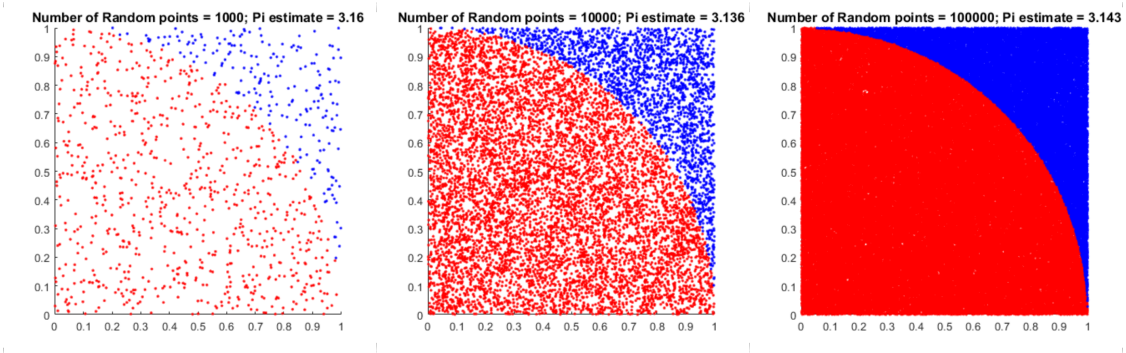


Figure 1.9: Monte Carlo simulations placing a changing number of points in a quadrant of radius one within a square of side length one to estimate pi. Three Monte Carlo simulations placing 1000, 10000 and 100000 points within a square of side length 1. The probability of being placed within the quadrant is $\pi/4$ and so the relative number of points in the quadrant gives an estimate for π. The estimate improves the more points are used in the simulation (3.16 to 3.136 to 3.143).

$$PA_1 = \sigma A_2$$

Where the internal area of the cross section is given by A_1, the area of the walls of the shape are given by A_2 and the internal pressure and shear stress are given by P and σ respectively as shown in Fig. 1.10.

With the assumption that the internal pressure of a body has equilibrated and is constant across the body, this gives us a relation between the dimensions of a hollow body and the shear stress in the walls.

Pressure vessel models have not been used in the context of viral studies previously and are of interest to study viral shape and the viral envelope.

SIR (Susceptible, Infected, Recovered) Model

SIR (Susceptible, Infected, Recovered) models look at the number of susceptible, infected and recovered members of a population change over time by following a set of differential equations. It was first used in 1927 by Kermack and McKendrick [68] and has since found use in modelling epidemics from Acquired Immune Deficiency Syndrome (AIDS) [69] to SARS [70].

The SIR model measures disease spread over time by tracking 3 populations of people and is an example of a compartmental model [68]. The number of people

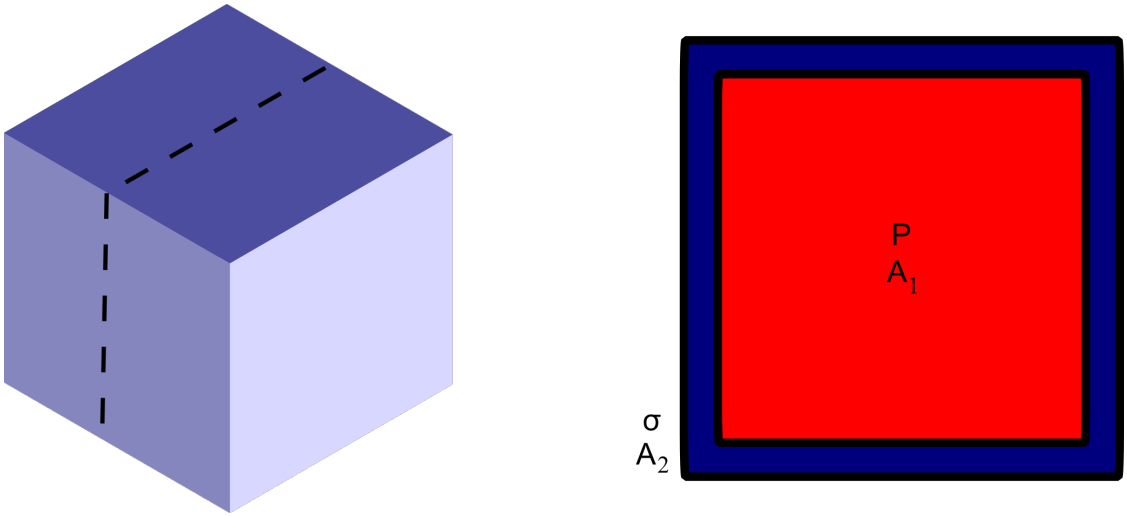


Figure 1.10: An example of the pressure vessel model that is generally applicable to any arbitrary 3D shape. By taking a cross-section of a 3D shape (in this example, a cube) and equating the force due to the internal pressure normal to the cross section and the force due to the shear stress normal to the cross section, the relationship between this shear stress at this point on the 3D body and the internal pressure can be found.

susceptible to the disease, S(t), those infected, I(t), and those recovered from the disease R(t), are the values measured over time with their sum, N, constant over time (i.e. S(t) + I(t) + R(t) = N) within a simulation. The model is controlled by a coupled set of differential equations with 2 constants controlling the rate of infection, β, and the rate of recovery, γ:

$$\frac{dS(t)}{dt} = -\beta SI$$
$$\frac{dI(t)}{dt} = \beta SI - \gamma I$$
$$\frac{dR(t)}{dt} = \gamma I$$

The model is initiated with the number of people in each population and the rate constants (S(0), I(0), R(0), β and γ) and then allowed to evolve over time. Assumptions involved in this model are that births and deaths happen too slowly to alter the model; there is an equal probability of any person catching the virus or recovering from the virus and all of the infections and recoveries from the model occur simultaneously.

This is a relatively simple model but it continues to be studied and different adaptations to the model are used to study different infection dynamics of a viral infection.

1.6 Chapter organisation

In this work I focus on improving our knowledge of viruses and methods to study and detect them.

In chapter 2, an imaging protocol is optimised and single virus particles are analysed en masse to study their size and shape. In chapter 3, influenza particle protein distribution is studied including the alternation of their surface proteins and the localisation of RNPs; the budding of filamentous influenza is studied through fixed cell fluorescence imaging, and the production of labelled virus particles using labelled cells is described. In chapter 4, I use a thin-wall pressure model to investigate shear stress in microbiology and why filamentous morphologies exist and, finally, in chapter 5, I extend the SIR (Susceptible, Infected, Recovered) simulation framework to investigate the reinfection dynamics of a novel, pandemic-causing virus.

The principal hypothesis of this thesis is that the shape, size and organisation of virus particles can be studied through high throughput fluorescence imaging. Filamentous morphologies of influenza were investigated by studying the shape and protein distribution of fixed filamentous particles, modelling them as pressure vessels to understand why filaments are formed, and imaging their cellular interactions and budding. The labelling of viruses by infecting cellular hosts which have had their cell membrane fluorescently stained is a novel method described in this thesis.

The specific contributions of other persons are identified at the beginning of any chapter in which they contributed. Appendix A contains a list of pre-prints and published works to which I contributed as a D. Phil. student; appendix B details work using a Susceptible, Infectious, Recovered (SIR) model to investigate reinfection of a novel, pandemic-causing virus (SARS-CoV-2) which may be important for understanding future viral outbreaks and rapidly collecting data on the infection dynamics of future novel viruses, and further appendices describe short-term projects

which yielded interesting results but weren't substantive enough to warrant full thesis chapters.

2

Imaging of fixed virus particles

Contents

2.1 Introduction 22
2.2 The development of the imaging protocol 23
Growing the viral samples 24
Immobilisation of virus to microscopy slides 24
Labelling the viral particles 25
Testing different blocking strategies 29
High-throughput microscopy imaging of virus particles 29
2.3 High-throughput analysis of influenza virus pleomorphism 31
Length analysis of filamentous particles 31
Small virion shape analysis 33
2.4 Methods and materials 38
Virus strains 38
Virus immobilization 40
Sample preparation 41
Imaging 41
Channel mapping 42
Drift correction 42
Filament length analysis 42
Spherical size analysis 43
Aggregating STORM images 43
2.5 Discussion and conclusion 44

2.1 Introduction

Viral disease results in significant illness and deaths in humans each year [71], and thus represents a large healthcare and economic burden to countries around the world. The COVID-19 pandemic has resulted in the deaths of millions of people [72], while annual influenza epidemics can result in up to 650,000 respiratory deaths per year [6], with significantly more during intermittent influenza pandemics. Despite the high mortality and morbidity associated with viruses, many aspects of their structure and morphology are poorly understood, in part due to their pleomorphic nature and small size, which represent a challenge when it comes to imaging them via diffraction limited microscopy.

Influenza virus particles are highly pleomorphic [9], ranging in size from spherical virions ∼100 nm in diameter [8] to filaments of a similar width but reaching many micrometers in length. Similar pleomorphism has been seen for many other viruses, such as Respiratory Syncytial Virus (RSV) [73], Ebola [74], measles [75], human parainfluenza virus type 2 (HPIV2) [76] and Newcastle Disease Virus [77]. Filamentous strains of influenza are not well studied, both because laboratory-passaged viruses tend to be spherical (with a filamentous morphology more typical of clinical isolates) and also because ultracentrifugation and other purification and storage procedures tend to damage filaments [10]. Filamentous virus structure is of interest as it has been linked to increased pathogenicity [78][79][80], resistance to neutralizing antibodies[81] and penetration through host mucus barriers [18], as well as offering a tool to study virus assembly and novel mechanisms of infection [5] [18].

To study the pleomorphism of viruses, detailed information on the shape, size and protein organisation of individual virions is required. Diffraction limited fluorescence microscopy is of limited utility to study viral shape and organisation as virus particles are generally smaller than the ∼250 nm resolution limit of optical microscopy for visible light. Electron microscopy (EM), and more recently cryo-EM, are excellent methods for high-resolution imaging of virus structure, but can be comparatively low throughput, time consuming, and offer limited molecular identification. The pleomorphy of many viruses also precludes visualization by

procedures that rely on averaging many identical particles, and as a result, it remains challenging to measure protein distributions at the single-virus level across an entire population. To overcome these limitations of throughput and sample averaging, I used direct stochastic optical reconstruction microscopy (dSTORM)[57][64], a technique that allows the location of fluorescent molecules to be determined with a resolution of less than 20 nm and offers an alternative for virus imaging [82][83]. This method is compatible with immunostaining, allowing us to specifically label multiple viral proteins within virions.

In this study, I combined dSTORM imaging methodology with rapid automated analysis software to carry out a high throughput and high-resolution analysis of thousands of virions at a time. I used these tools to investigate the size and shape of influenza H3N2 A/Udorn/72; a common laboratory strain of influenza which, unlike other laboratory-adapted strains, still exhibits filamentous morphology as well as spherical morphologies[21]. This analysis allowed the how many populations are involved in the range of pleomorphism seen in influenza. The analysis pipeline is versatile and can be adapted for use on multiple other pathogens, as demonstrated by its application to SARS-CoV-2. The ability to study the populations that make up viral pleomorphism is important to study viral assembly.

In section 2.2 I discuss the development of the imaging protocol, and in section 2.3 I show that A/Udorn/72 particle sizes falls into 2 distinct populations, In section 2.4 I detail the imaging protocol and methods used in my analyses and, finally, in section 2.5 these results are discussed in the context of the current literature.

Danielle Groves completed the imaging of SARS-CoV-2. All other work was carried out independently.

The work presented in this chapter is in preprint as Ref. [84].

2.2 The development of the imaging protocol

Before being able to image and automatically analyse thousands of virus particles from various experiments, a reproducible method of immobilising, labelling and imaging viruses was required.

Growing the viral samples

To produce filament-containing virus stock, I used the H3N2 A/72/Udorn virus, an influenza strain with a well-documented spherical and filamentous phenotypes[21]. Viruses were grown as described in the methods section. Briefly, confluent Madin-Darby Canine Kidney (MDCK) cells were infected at a multiplicity of infection of 0.1 and incubated for 24 hours before the supernatant was harvested and clarified without any further purification. These particles were then immobilised on the surface of a pegylated glass slide using biotin, and imaged using a non-specific fluorescent staining method that uses calcium ions to bind fluorescent DNA to the virus surface [85]. This offered a rapid, non-specific method to label virus particles to quickly ensure filamentous morphologies were present. Diffraction-limited images of both spherical and filamentous viruses were observed on the glass slide 2.1, confirming that filamentous influenza particles could be produced. Stocks were grown freshly for experiments (a new stock grown monthly) as the freezing process, and the freeze-thaw process in particular, damages filamentous morphologies [10].

Immobilisation of virus to microscopy slides

Initially, virus particles were immobilised for fluorescence imaging by non-specifically biotinylating the virus surface and incubating particles on the surface of a pegylated glass slide. However, this method did not result in many immobilized particles (Fig. 2.2A), was not consistent, and required time-consuming slide preparation and so, other immobilisation techniques were tested.

Testing of polymer coatings

Tested immobilisation methods included coating the glass slides with the positively-charged linear polymers poly-L-lysine (Fig. 2.2B) or chitosan (Fig. 2.2C), which electrostatically attracted the negatively charged virus particles. Both chitosan and poly-L-lysine immobilised virus particles well and with low background. However, poly-L-lysine was chosen for subsequent experiments due to its long-term stability, ease of preparation and the large structures that could be immobilised.

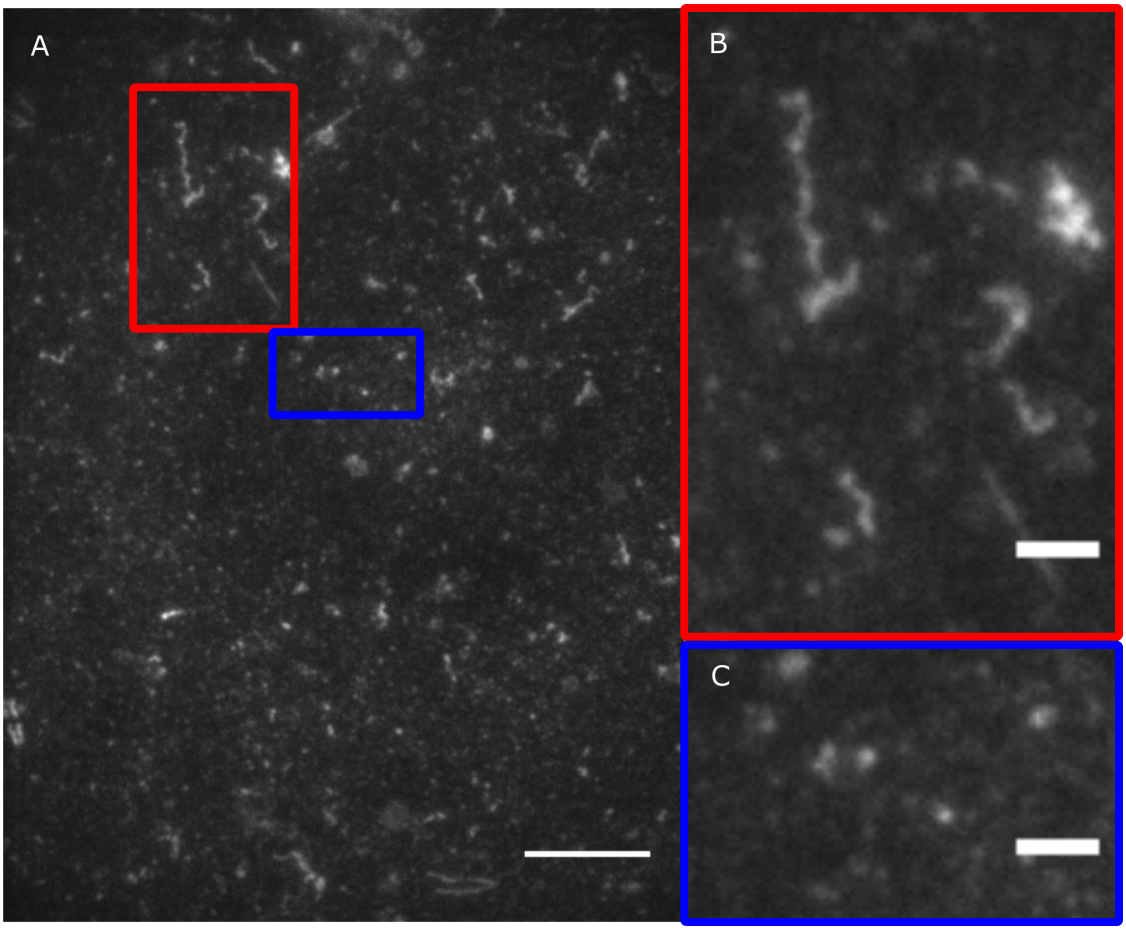


Figure 2.1: Diffraction-limited images of filamentous A/Udorn/1972 virus. Virus particles were biotinylated and immobilised on a pegylated glass slide, before being fluorescently labelled using a mixture of calcium chloride and fluorescent DNA[85] A) A field of view of labelled Udorn. Scale bar 10μm. B) Zoomed in image of filamentous virions. Scales bar 2.5μm. C) Zoomed in image of spherical virions. Scales bar 2.5μm.

Testing drying viral samples onto slides

Virus samples were dried directly onto glass coverslips with and without a pre-coating of poly-L-lysine, a process which took approximately 10 minutes. I showed that a combination of drying the sample and using poly-L-lysine increased the number of immobilized virus filaments on the slide surface (Fig. 2.3) and that the filament appearance was visibly unaffected by the drying protocol.

Labelling the viral particles

I conducted a thorough search in order to identify the best antibodies for A/Udorn/72 that specifically labelled proteins of potential interest. They brightly labelled both

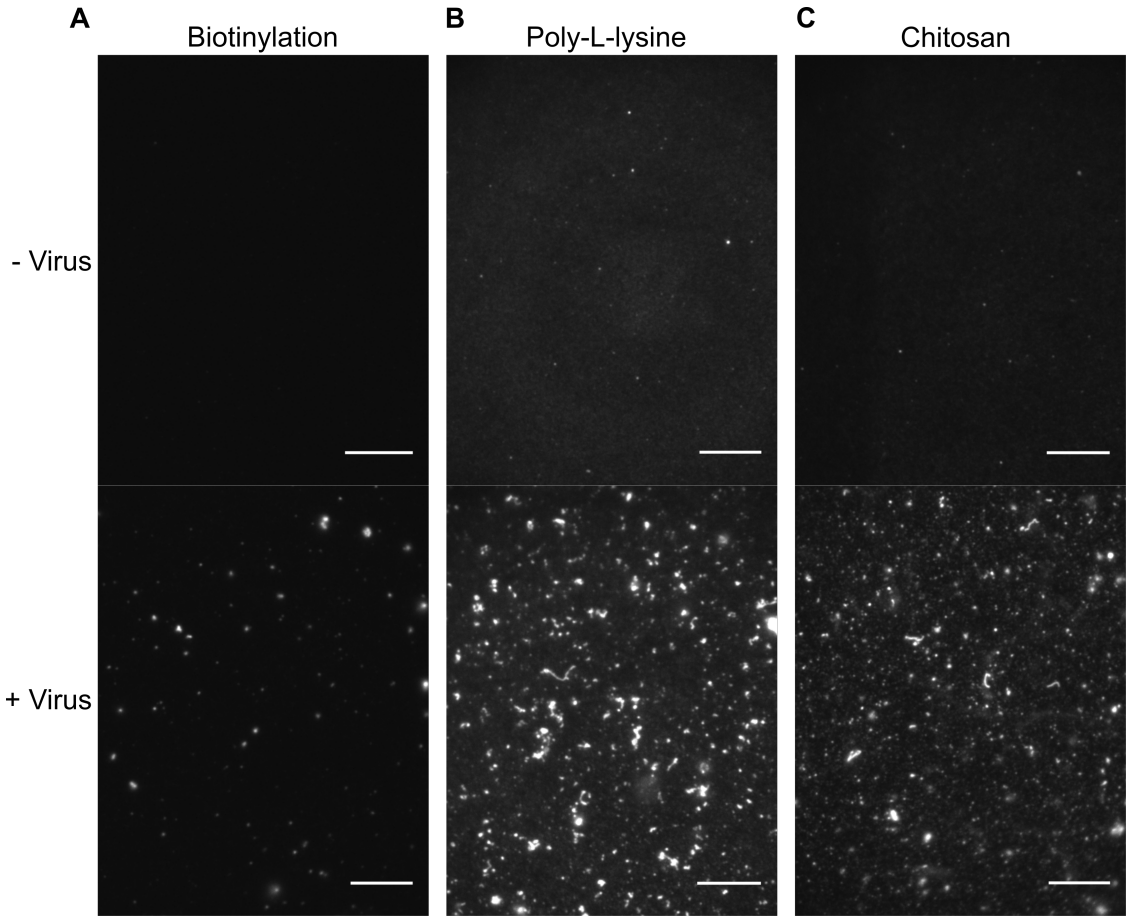


Figure 2.2: Immobilisation methods for filamentous viruses. A) A virus negative control (top) or an A/Udorn/72 virus sample (bottom), were immobilized via a specific biotin/PEG linkage. Virus particles were biotinylated by incubation with 1 mg/mL Sulfo-NHS-LC-Biotin (ThermoFisher) for 3 hours at 37°C before being immobilised on a pegylated slide. The virus was labelled with an anti-Udorn primary antibody and an Alexa647 secondary antibody and imaged on a widefield TIRF microscope. Scale bars 10 μm. B) A virus sample was incubated on a slide pre-treated with 0.01% poly-L-lysine. C) A virus sample was incubated on a slide pre-treated with 0.015 mg/mL chitosan in 0.1 M acetic acid.

filamentous and spherical virions and had a sufficiently high labelling efficiency to study the protein organisation with a large number of particles.

The best antibodies against HA, NA and M1 are shown in Fig. 2.4 A, B and C respectively, labelled with a secondary antibody labelled with Alexa Fluor 647. The comparison of widefield microscopy images of a viral sample and a negative control show the specificity of the antibodies, and super resolution images of the viral samples using the antibodies show they are suitable for a high throughput, super resolution study.

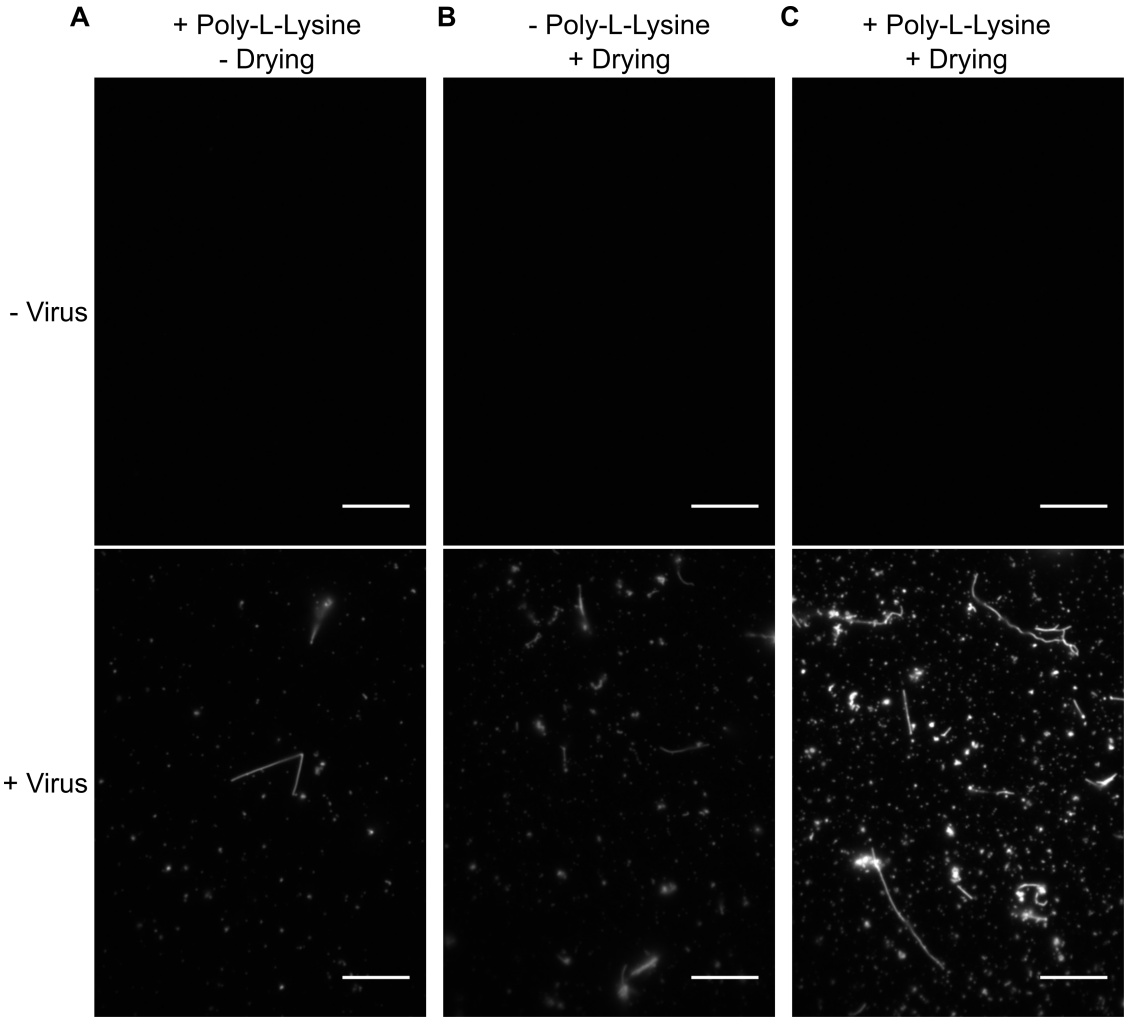


Figure 2.3: A combination of drying the sample and using poly-L-lysine increases the number of immobilized viruses on the slide surface. A) A virus negative control (top) or an A/Udorn/72 virus sample (bottom), were incubated on a slide pre-treated with 0.01% poly-L-lysine at 4°C for 10 minutes. The excess sample was removed from the well and the immobilized virus was fixed and labelled with an anti-Udorn primary antibody and an Alexa Fluor 647 secondary antibody before being imaged on a widefield TIRF microscope. Scale bars 10 μm. B) As in A) but the slide was not pre-treated with poly-L-lysine and the samples were dried directly onto glass coverslips by heating at 45°C for 10 minutes. C) As in A) but the samples were dried directly onto glass coverslips by heating at 45°C for 10 minutes.

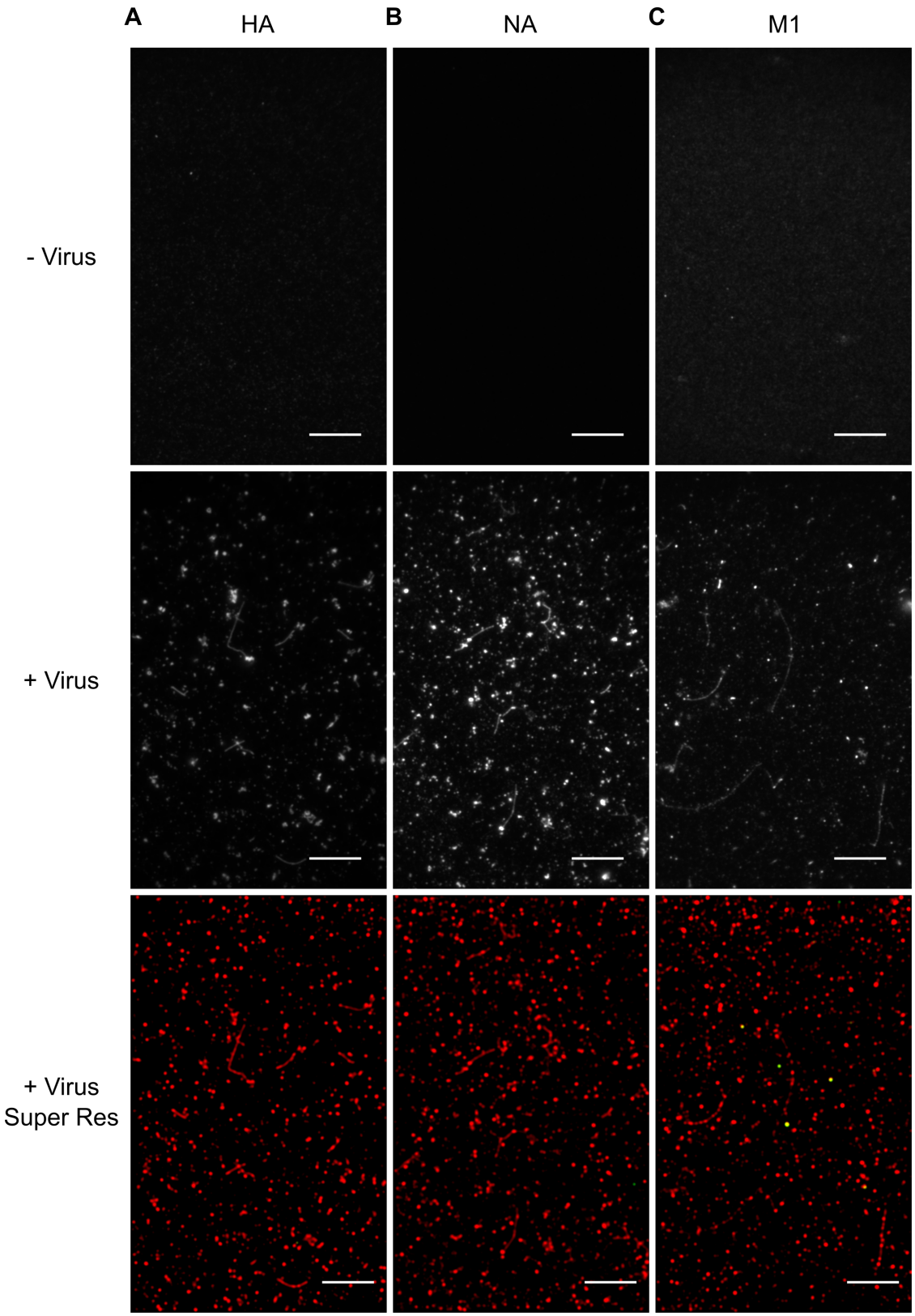


Figure 2.4: The testing of antibodies found 3 specific antibodies for the HA, NA and M1 proteins. A virus negative control (top), an A/Udorn/72 virus sample (middle) and a super resolution image of the virus sample (bottom) for the antibodies that were specific against A/Udorn/72, all labelled with an Alexa Fluor 647 secondary antibody. Scale bars 10 μm. A) HA antibody. B) NA antibody. C) M1 antibody.

Testing different blocking strategies

With a protocol that could produce filamentous virions, stick them to a slide for imaging, and label the proteins of interest, I wanted to investigate which blocking agent worked best to remove the background signal from images. I tested various proteins (bovine serum albumin (BSA), donkey serum, gelatin, and milk powder) and also included a control with no blocking agent. The images produced using these different blocking agents are shown in Fig. 2.5. As assessed visually, the lowest background (and also signal to noise ratio) was produced by the sample using normal donkey serum and so this was changed in the imaging protocol.

High-throughput microscopy imaging of virus particles

Immobilised viruses were fixed, permeabilised and the HA and NA was labelled with antibodies using the protocol developed above (see Methods and Materials) (Fig. 2.6A) to investigate the throughput of our imaging and the suitability of the images to study viral shape and organisation. Next, the immobilized virus particles were imaged using a widefield total internal reflection fluorescence (TIRF) microscope. A/Udorn/72 virions were dual-stained using anti-HA and anti-NA primary antibodies and secondary antibodies labelled with Alexa546 (green) and Alexa647 (red) respectively (Fig. 2.6B). A single field-of-view (FOV; measuring 50 x 80 μm) showed examples of multiple viruses, including large numbers of spherical particles and filaments of a wide range of different lengths (Fig. 2.6B). A negative control consisting of media from non-infected cells showed that the labelling was specific (Fig. 2.6C). A long acquisition (10,000 frames) of the FOV was taken to generate a super resolution image, which showed the virus particles at high resolution, revealing the spatial organization of HA and NA on the surface of both spherical particles and filaments (Fig. 2.6 D-F). Filaments were seen to be relatively uniformly covered in HA with patches of NA as previously seen [18] and expected from mass spectrometry [16]. Multiple FOVs of the same sample can be imaged in this way, providing high-resolution information on thousands of spherical influenza virions and hundreds of filaments in each run. Taken together,

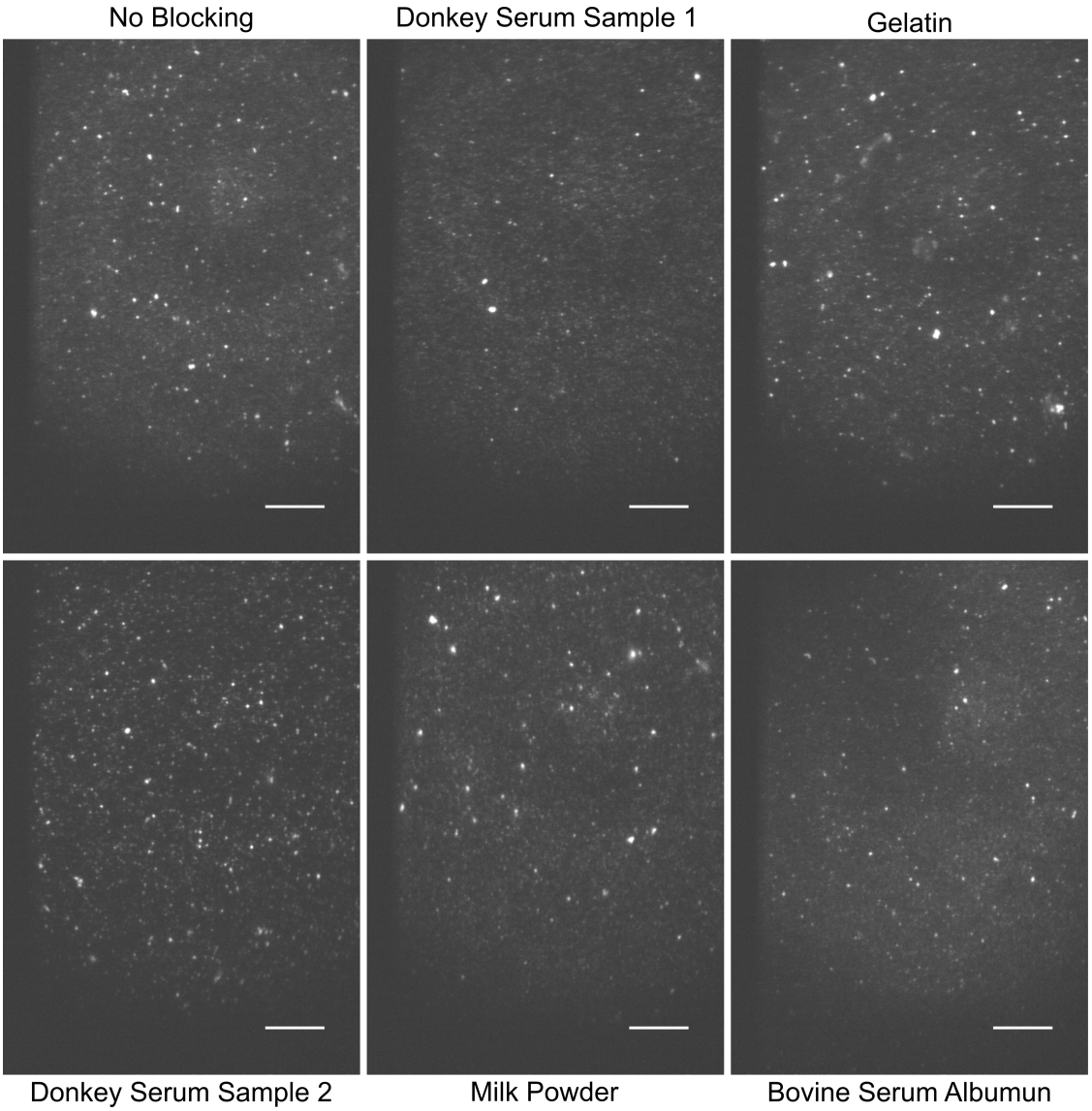


Figure 2.5: The testing of different blocking agents found normal donkey serum to be the best. 4 different proteins were tested for their blocking ability on a sample of A/Udorn/72 with our found HA antibody with an Alexa Fluor 546 secondary antibody. The samples were of donkey serum (brands unknown, obtained from Ervin Fodor, University of Oxford), gelatin, milk powder and bovine serum albumin. The best chemical was normal donkey serum which had the lowest background which meant, on a positive sample, it had the highest signal to noise ratio.

this approach allows us to efficiently, rapidly and easily immobilize and image large numbers of virus particles at high resolution, in order to gain structural information on large virus populations.

2.3 High-throughput analysis of influenza virus pleomorphism

Length analysis of filamentous particles

Our initial images showed a population of elongated virus particles with large variability in length, ranging from approximately 250 nm to several microns. Due to the large size of the viral filaments (larger than the diffraction limit), I initially performed a high-throughput length analysis of multiple filaments using widefield images containing diffraction-limited signals. Multiple images of A/Udorn/72 virus labelled with an anti-HA antibody were taken and a rapid automated analysis pipeline was used to measure filament length (Fig.2.7A). Each image was adjusted and binarized to pick out the virus filaments, using a lower threshold of 234 nm (a length of 2 pixels) to exclude any particles that would be below the diffraction limit. Morphological closing operations were used to fully connect the virus particles if there were regions without labelled protein and then a skeletonization operation[86] which reduced the particles to the simplest shape to give the shape of the filaments was applied. The lengths of the resulting skeletons were taken as the lengths of the virus filaments [87][88].

This process was completed on 486 individual FOVs, allowing the lengths of 46,872 filamentous particles to be measured. The resulting histogram was fitted with a single exponential function (Fig. 2.7 B) showing virions longer than 234 nm to belong to a single population that are produced stochastically. Repeating the process on 243 FOVs negative for virus showed that any background signal due to spurious noise or background labelling was negligible (Fig. 2.7C&D) and could be subtracted from the histogram of viral lengths. The resulting length distribution revealed that there were significantly more filaments of shorter lengths (<1000 nm), with the frequency of filaments decreasing at longer length scales. This is

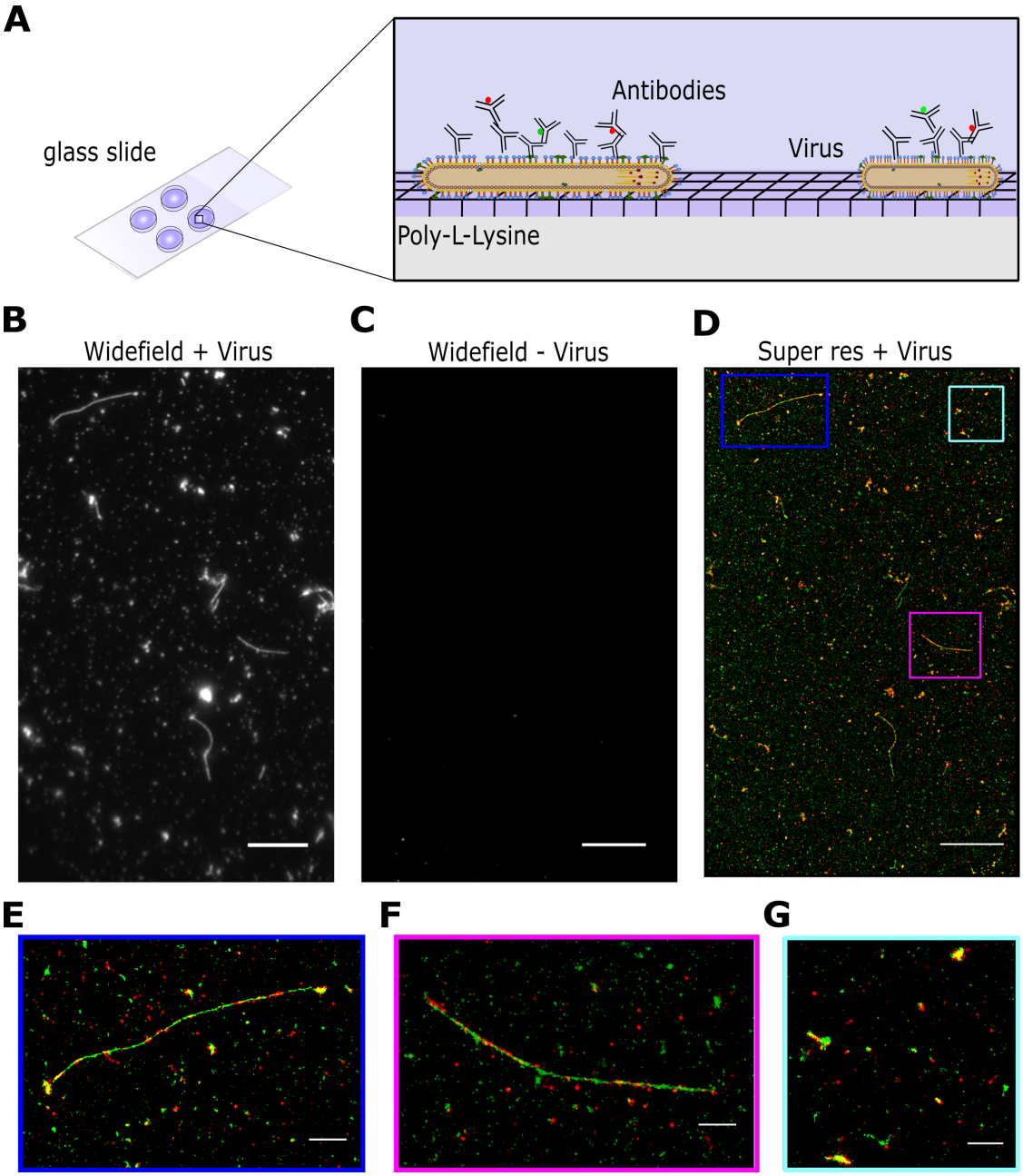


Figure 2.6: High throughput imaging of influenza using super-resolution microscopy. A) Schematic of the labelling protocol. Virus samples were dried directly onto glass coverslips pre-coated with poly-L-lysine before being fixed, permeabilised and stained with antibodies using a standard immunofluorescence protocol. B) A representative field of view (FOV) of a widefield image of labelled A/Udorn/72 influenza, imaged in the green channel. Scale bar 10 μm. C) A representative FOV of a widefield image of a virus negative sample, imaged in the green channel. Scale bar 10 μm. D) The corresponding dSTORM image of the FOV in B), where HA is labelled in green and NA is labelled in red. Scale bar 10 μm. E-G) Zoomed in images from D) showing individual filaments and spherical particles. Scale bars 5 μm.

consistent with previous observations that filaments are fragile[10] and can break or fracture at longer lengths; alternatively, length may be limited by the greater amount of membrane and viral proteins required to form each filament compared to a spherical particle. Due to the fragility of filaments, it is thought longer filaments are more likely to break, potentially biasing the data further towards the smaller population seen. The large variability in filament lengths and lack of distinctive sub-populations of filamentous virions of a particular size also suggests a model of virus particle assembly where membrane scission and the release of progeny filaments from cells occurs in a stochastic way.

Small virion shape analysis

Having completed a comprehensive length analysis of long virions we next considered how to do this for smaller virus particles (specifically, those that fall into the spherical, ellipsoidal/bacilliform and short filament (<600 nm) categories). The small size of these viruses precluded the use of widefield images to get accurate length data, and so I turned to super-resolution images to obtain their size distribution. A/Udorn/72 virus was immobilized, stained and imaged using dSTORM as described above; the DBSCAN clustering algorithm [89] was used to cluster the resulting localisations in each FOV, and an ellipse was fitted to each cluster (Fig. 2.8A). The best-fit ellipses were used to give the major and the minor axis lengths of the clusters, which were taken as a measure of virus particle size and the results were plotted as histograms.

This process was completed on 46 individual FOVs, allowing the size of 41,754 particles to be measured. By repeating the process on virus-negative FOVs, the data was corrected for noise or background (Sup. Fig. 2.9). The resulting histogram of the major axis length was fitted using a triple log-normal function, chosen to take into account the two distinctive virion populations as well as a minor third distribution representing the small amount of background signal still remaining (Fig. 2.8B). This analysis revealed two distinct populations of virions, centred at ∼75 nm (54.6% of the total particles) and ∼160 nm (44.2% of the total particles), likely corresponding to spherical and elongated virions respectively. The third distribution, representing

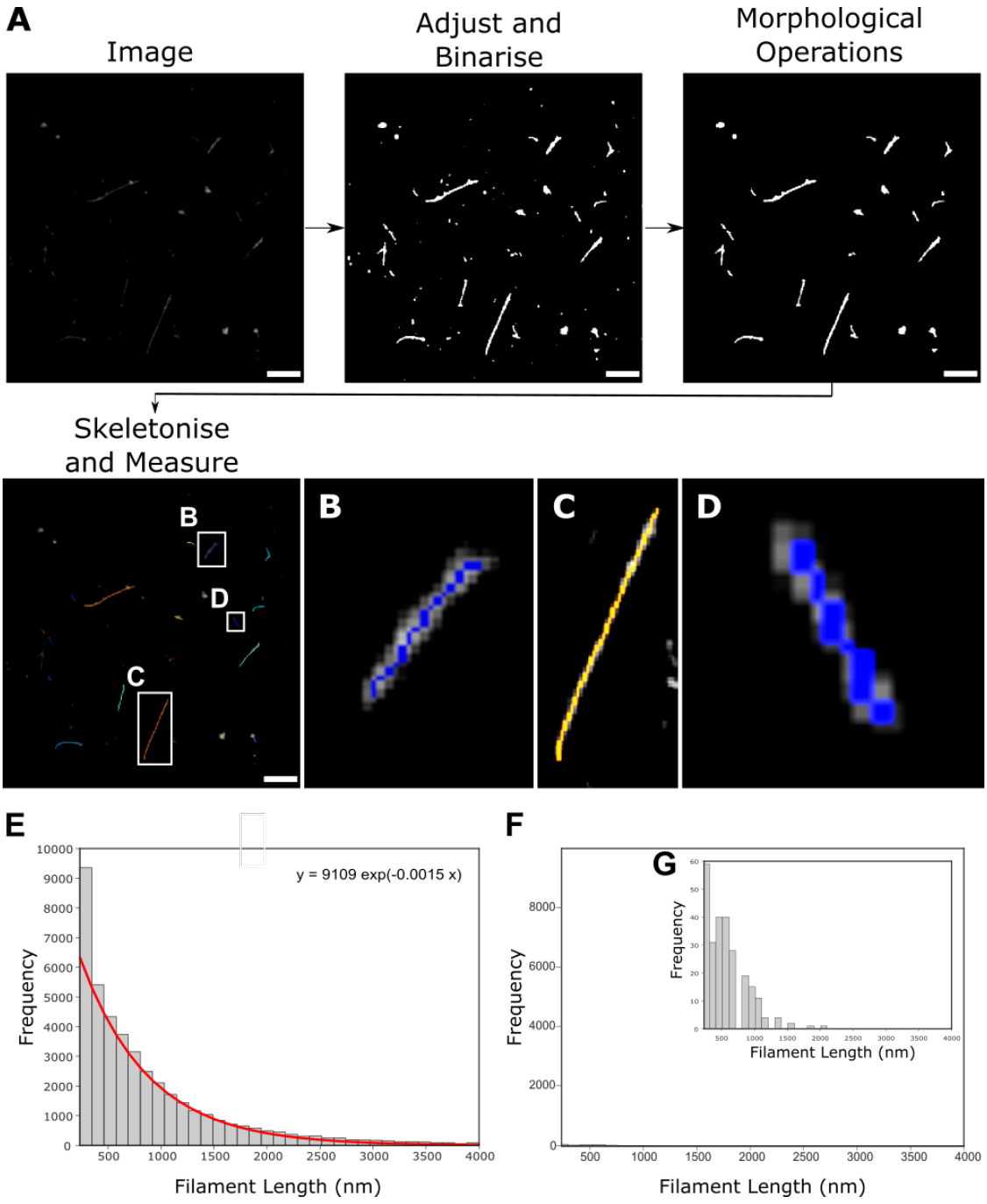


Figure 2.7: Widefield imaging and automated length analysis of A/Udorn/72 reveals that filaments fit a single distribution. A) A pipeline for analyzing filamentous Udorn particles uses multiple widefield images, which are adjusted and binarised before being skeletonized to get a measure of filament length. Scale bar 5 μm. B,C&D) Zoomed in skeletons from A). E) Widefield images of A/Udorn/72 virus stained with an antibody against the HA protein were analysed using the pipeline in A). The resulting frequency distribution of the lengths of 46,872 filamentous particles, fit with a single exponential (red line; equation above plot). F) Histogram from the analysis of 243 virus-negative FOVs demonstrates that the background was negligible. G) Zoomed in histogram of F).

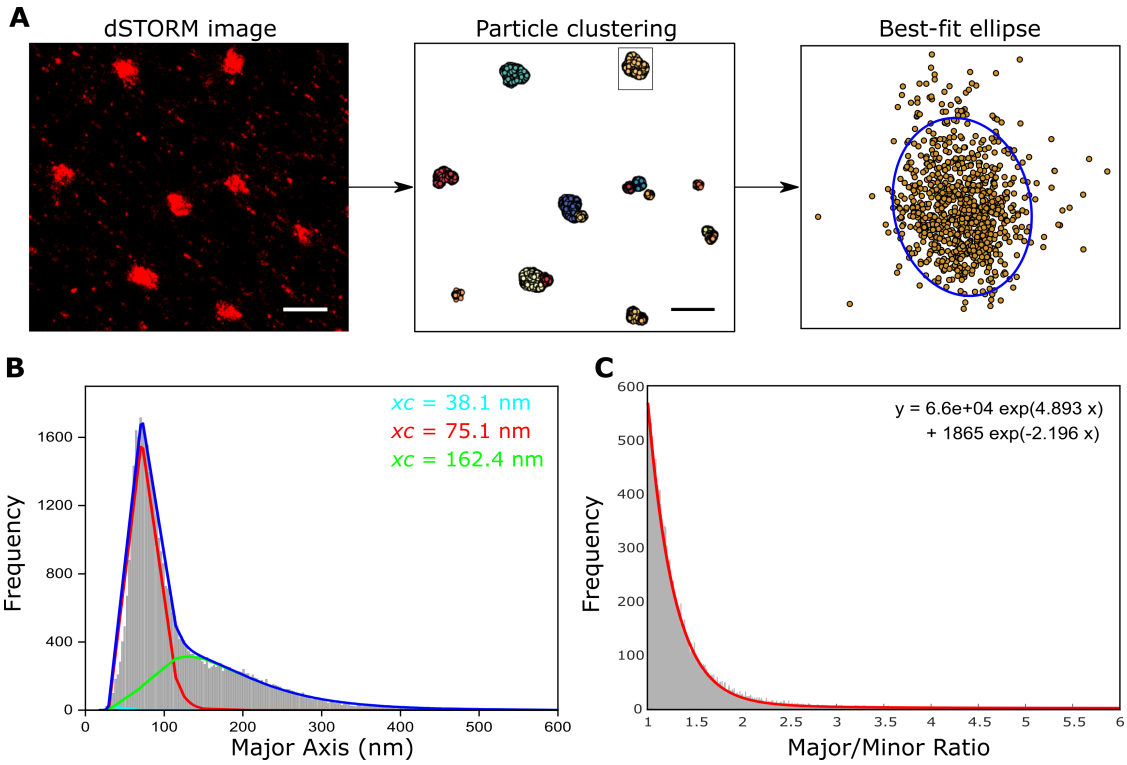


Figure 2.8: Super-resolution imaging and size analysis of spherical virus particles. A) A pipeline for analyzing spherical and bacilliform particles using multiple super-resolution images. Super-resolution localisations are clustered using DBSCAN and each cluster is fitted with an ellipse. Particle dimensions are extracted using the major and minor diameters of each ellipse. Scale bars 1 μm. B) A histogram of the major axis lengths shows that influenza virions stained for the HA protein fall into two distinct populations, centred at 75.1 nm and 162.4 nm. C) Histogram of the major/minor axis ratio of influenza particles shows a single distribution.

1.2% of all particles, was centred at ~40 nm; the extremely small size of this distribution suggests that our clustering methods and subtraction of data points from virus-negative experiments were sufficient to remove most non-specific localisations.

The fitting error in individual localisations was plotted for the super-resolution images used in this analysis, which gave a median error of 7.4 nm in the x and 7.2 nm in the y direction. This is only part of the localisation error which includes fitting error, label length and various imaging artefacts. Given a large number of localisations, these errors are too small to make any meaningful difference to our results (Sup. Fig. 2.10).

EM images of virus particles have previously been categorized based on their axial ratio (< 1.2 for spherical virions, > 1.2 for bacilliform virions and filaments), as

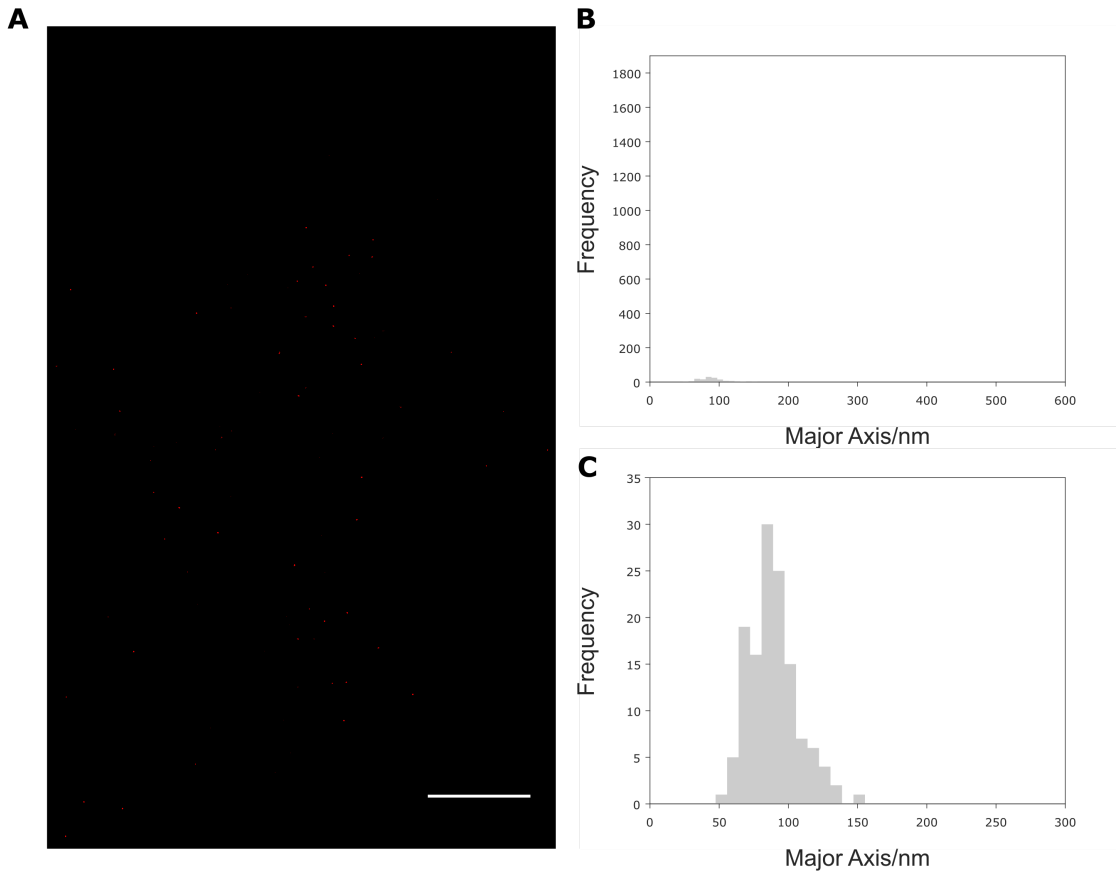


Figure 2.9: Virus-negative FOV and histograms for the size analysis of spherical and bacilliform influenza particles. A) A representative super-resolution image of a virus-negative sample virus stained with an antibody against the HA protein. Scale bar 10 μm. B) Super-resolution localisations were clustered and each cluster fitted with an ellipse to extract particle dimensions. A histogram of the major axis lengths shows that background signal is negligible. C) Zoomed in histogram of B).

well as length. I therefore plotted the major/minor axis histogram from the measured particles, and fitted the resulting plot with a single exponential function (Fig. 2.8 C). Rather than being able to distinguish two distinct axial ratio populations (of greater than and less than 1.2) [23], a single population was observed. This suggests that when analysing large numbers of virions (many thousands compared to a few hundred via EM) a much larger variability in particle shape may be captured with no distinct populations determined by axial ratio; thus suggesting that a greater heterogeneity in virus shape than previously thought may occur.

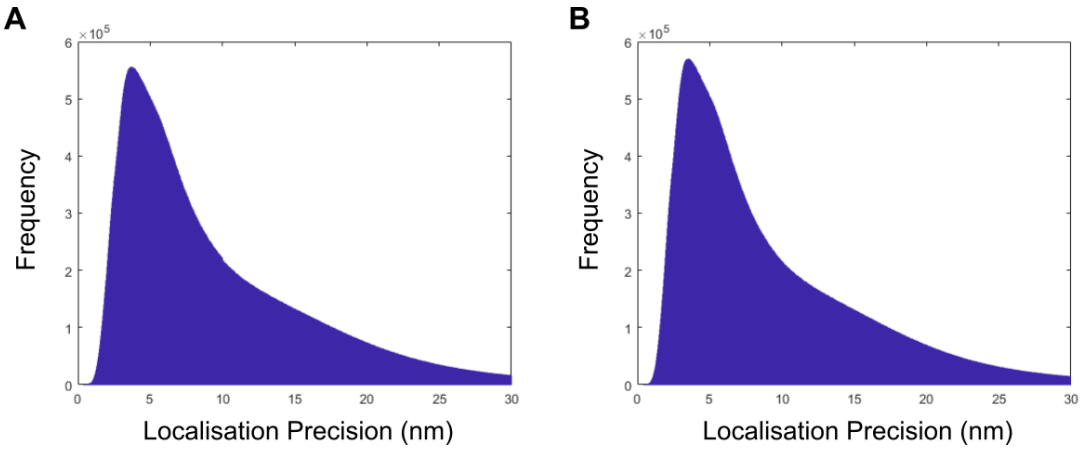


Figure 2.10: Histograms of the localisation precisions of the data from Figure 2.8. A) Each FOV was drift corrected, and a Gaussian function was fitted to each detected localization in every frame of the acquisition, using the inbuilt Nanoimager software. The fitting error (or localization precision) in the x direction of each localization was exported and plotted as a histogram, providing a median error of 7.4 nm. B) Plot of the localisation precision in the y direction, providing a median error of 7.2 nm.

In order to demonstrate the versatility of our imaging and analysis pipelines, we carried out similar experiments on the SARS-CoV-2 virus, the causative agent of the COVID-19 pandemic. Fixed SARS-CoV-2 virions were dual-labelled for the spike (red) and nucleocapsid (green) proteins (Fig. 2.11A), or the envelope (red) and nucleocapsid (green) (Fig. 2.11B). Multiple co-localised particles were observed, corresponding to double-labelled virions. Super-resolution localisations in each channel were clustered and each cluster fitted with an ellipse to extract particle dimensions. Analysis of the spike protein localisations showed that virions fell into a single population centred at ∼94 nm (Fig. 2.11C), in keeping with an expected size for SARS-CoV-2 virus particles[27][28][29][30], while the major/minor axis histogram from the labelled SARS-CoV-2 particles showed a tighter distribution of virion axial ratio than that found for influenza, in keeping with observations that SARS-CoV-2 forms largely spherical particles (Fig. 2.11D). Repeating the analysis for the nucleocapsid protein revealed a distribution centred around ∼82 nm, as expected from an internally labelled structure within the virion (Fig. 2.11E). We also found that the signal from the viral envelope proteins produced a distribution centred around ∼83 nm, which is slightly larger than the nucleocapsid (Fig. 2.11F). Comparing the distributions found using an unpaired t test, the difference between

the spike protein and both internal proteins is statistically significant while the difference between the envelope protein and nucleocapsid protein is not.

From the SARS-CoV-2 data, we were also able to build up an average structure from many virions from multiple FOVs. This allows the average internal structure of a virion to be seen. Fig. 2.12A shows there to be a consistent channel mapping offset or a consistent drift between imaging different channels that is not adjusted for though there is no drift through an acquisition seen. The average structure can still be found by centring proteins on their own centroids (Fig. 2.12B) or by centring every displayed protein off another protein.

All together, these results are in keeping with previously observations on the size and distribution of SARS-CoV-2 virus particles and demonstrate the general applicability of our analysis pipeline to multiple other viruses.

2.4 Methods and materials

Virus strains

The influenza strain H3N2 A/Udorn/72 (Udorn) was grown in Madin-Darby Canine Kidney (MDCK) cells as described previously [85]. To produce filament-containing stocks for analysis, confluent MDCK cells were infected at a multiplicity of infection of 0.1 from a stock of known titre and incubated at 37°C in serum-free media (Dulbecco's Modified Eagle Medium, Gibco) supplemented with 2 μg/mL TPCK-treated trypsin (Sigma) for 24 hours. Supernatants were harvested and clarified at 4000 rpm for 5 minutes at room temperature to remove cell debris before being used without any further purification or concentration steps. 0.2% formaldehyde was added to inactivate the virus and the supernatant was stored at 4°C to prevent filament degradation. SARS-CoV-2 was grown in Vero E6 cells and collected as above. The virus was inactivated by addition of 4% formaldehyde before use and stored at -80°C.

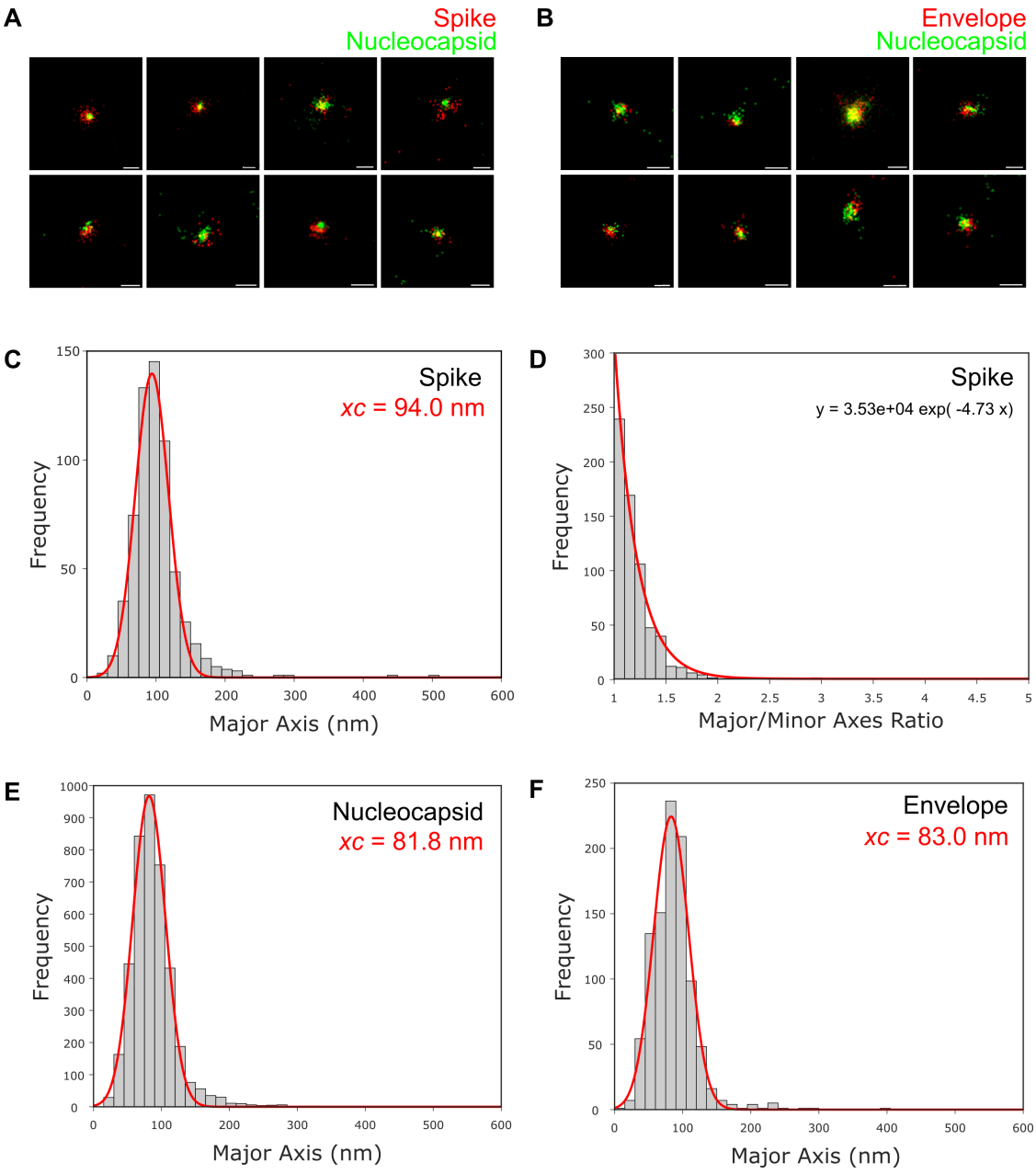


Figure 2.11: Super-resolution imaging and size analysis of SARS-CoV-2 virus particles. A) Representative super-resolution images of SARS-CoV-2 virions dual-labelled with anti-spike and anti-nucleocapsid primary antibodies and secondary antibodies labelled with Alexa647 (red) and Alexa546 (green) respectively. Scale bar 10 μm. B) Representative super-resolution images of SARS-CoV-2 virions dual-labelled with anti-envelope and anti-nucleocapsid primary antibodies and secondary antibodies labelled with Alexa647 (red) and Alexa546 (green) respectively. Scale bar 10 μm. C) A histogram of the major axis lengths fitted with a Gaussian function shows that SARS-CoV-2 virions fall into a single population, centred at 94.0 nm. D) Histogram of the major/minor axis ratio of spike localisations shows a single distribution. E&F) Analysis of the nucleocapsid protein and envelope protein also fall into single populations centred at 81.8 nm and 83.0 nm respectively.

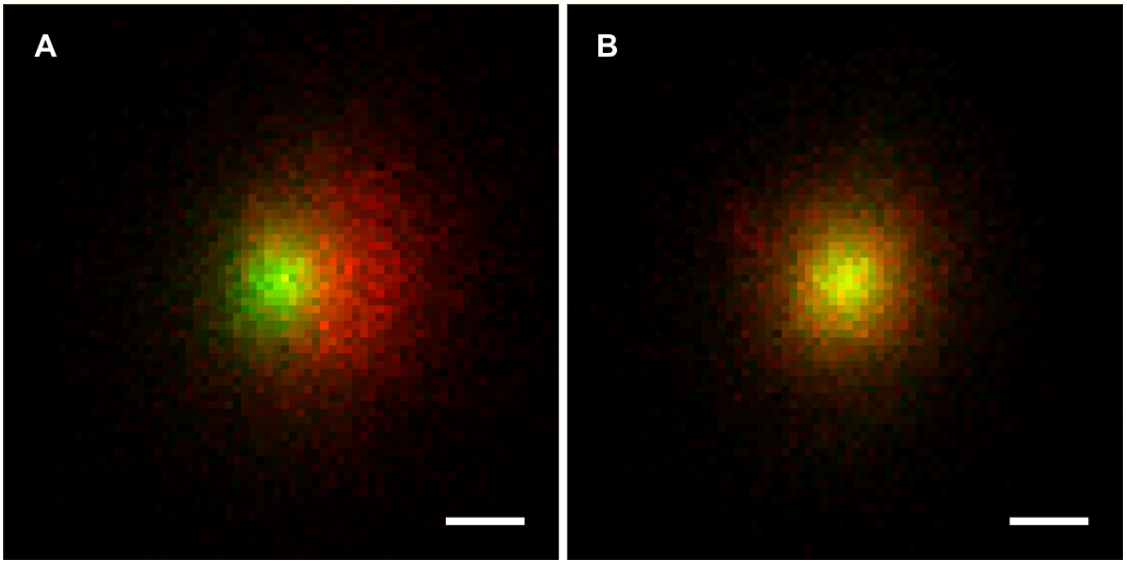


Figure 2.12: The average structure of the SARS-CoV-2 virion. A) The average structure when centring spike protein localisations on the centroid of nucleocapsid protein. B) The average structure when centring all proteins on their own centroid and aligning these centres.

Virus immobilization

Unless otherwise specified, viruses were immobilized using poly-L-lysine. Coverslips (25x65mm, thickness number 1, VWR) were heated in a furnace at 500°C for 1 hour to burn off any dust. A silicone gasket (Grace Bio-Labs, USA) was added to the glass slide and 0.01% poly-L-Lysine (Sigma) solution was added to the wells for 30 minutes. The excess was removed and the chambers were washed three times with MilliQ water and the slide allowed to dry. As an alternative chitosan was used in place of poly-L-lysine (0.015% chitosan powder (Sigma) in 0.1M ethanoic acid). For specific immobilisation of viruses via biotinylation, passivated microscope slides were prepared by washing in acetone and Vectabond solution (Vector Laboratories) before being incubated with NHS- PEG:Biotin-NHS-PEG in an 80:1 ratio. 0.5 mg/mL neutravidin was incubated for 10 minutes at room temperature on the slide shortly before virus was added. Viruses were biotinylated by incubation in a 1 mg/mL Sulfo-NHS-LC-Biotin (ThermoFisher) for 3 hours at 37°C before being fixed and immunolabelled as described below.

Sample preparation

Virus was incubated in the wells at 45°C until the well was dry. Immobilised virus was fixed with 4% formaldehyde (Thermo Scientific) in phosphate buffered saline (PBS) for 10 minutes at room temperature, before being permeabilised in 0.5% Triton-X-100 (MP Biomedicals) for 15 minutes. The sample was blocked to prevent non-specific binding with 10% donkey serum (Sigma) and 0.1% Tween-20 in PBS for 1 hour at room temperature. Primary antibodies were diluted in blocking buffer and incubated with the sample for 1 hour, followed by washing in PBS and labelling with secondary antibodies (Invitrogen). A final washing step was carried out and samples were stored in PBS at 4°C until imaging. For super resolution imaging, the PBS was replaced with a STORM imaging buffer comprising an enzymatic oxygen scavenging system consisting of 1 mg/mL glucose oxidase and 40 μg/mL catalase, 10% glucose and 0.05 M mercaptoethylamine (MEA) in 1x PBS. Multiple independent repeats, with different virus preparations, were taken over different days. A/Udorn/72 virions were labelled with the mouse anti-HA primary antibody Hc83x (a kind gift from Stephen Wharton, Francis Crick Institute) and goat anti-NA and anti-M1 (a kind gift from Jeremy Rossman, University of Kent). SARS-CoV-2 virions were labelled with the human anti-spike antibody EY6A37 (a kind gift from Tiong Tan and Alain Townsend, University of Oxford) and a SARS-CoV-2 nucleocapsid antibody (GTX632269) from Genetex. Secondary antibodies labelled with either Alexa647 or Alexa 546 were purchased from Invitrogen.

Imaging

All experiments were performed on a commercially available Nanoimager fluorescence microscope (Oxford Nanoimaging). The Nanoimager used a 1.4 NA 100x aperture with an effective magnification for the microscope of 110x. Images were acquired on an sCMOS Orca flash4 V3 camera which overall gave a pixel size of 117 nm/pixel. The sample was imaged using total internal reflection fluorescence (TIRF) microscopy. The laser illumination was focused at an angle of 53° with respect to the default position. Images of a field of view (FOV) measuring 80 x 50 μm

were taken with an exposure time of 30 ms. For super resolution imaging the laser intensities were gradually increased up to a maximum of 780 kW/cm^2 for both the green (532 nm laser filtered using a 530.5 – 533.5 nm dichroic) and red (647 nm laser filtered using a 632.8 – 647.1 nm dichroic) lasers to control the number fluorescent molecules per frame, and movies of between 5,000 and 15,000 frames were taken. The red channel was imaged first, followed by the green channel. DSTORM sub-pixel localisation was computed using Gpufit [90] to fit Gaussian profiles to signals.

Channel mapping

Channel mapping is computed using the inbuilt Nanoimager software. This software uses a mapping computed from a sample which fluoresces in both channels- usually a Tetraspek bead slide. The channel offset is found from the phase correlation shift between the two channels and then, using a maximum distance, it finds colocalised localisations in the channels that do not colocalise with multiple localisations in the other channel. Finally, the mapping matrix is found using least squares minimisation for up to order 2 (i.e. 1, x, y, xy, x^2, y^2) in the x and y coordinates of channel 0 for all colocalised localisations.

Drift correction

Drift correction is computed using the inbuilt Nanoimager software which uses phase correlation between consecutive frames to find a translative offset between the two images.

Filament length analysis

Images were saved as .tif files, opened in ImageJ and each FOV was cropped so that just the single channel in which the HA protein was labelled was used. The images were adjusted using the MATLAB imadjust function and binarized using the MATLAB imbinarize function with a threshold of 0.001. The background of the resulting black and white image was removed using the bwpropfilt MATLAB function and the resulting image was closed, skeletonised and labelled using the

MATLAB bwmorph, bwskel and bwlabel functions respectively. Any skeletons of less than 234 nm (2 pixels) long were excluded from the analysis, as were overlapping skeletons (due to overlapping filaments dried onto the slide. Results were plotted as a histogram and fitted with a bi-exponential function, detailed on the graph.

Spherical size analysis

Each FOV was drift corrected, and super-resolution localisations were extracted in each frame using the inbuilt Nanoimager software. The localisations were exported and analysed further in Matlab. Localisations were clustered using the sklearn library implementation of the DBScan clustering algorithm with a minimum cluster size of 200 and an epsilon of 30nm. The clusters were fit using the confidence_ ellipse library method of fitting a confidence ellipse to a set of points with 2.0 as the standard deviation of the ellipse. This algorithm uses the covariance of a set of points, the Pearson coefficient and the mean of the points to fit an ellipse to a set of points. The major and minor axes of the ellipse were taken as the size of the viral particles. Protein distribution and spatial frequency analysis. FOVs with multiple filaments ($>$ 615 nm) labelled with the protein of interest (HA or NA) were skeletonized as in the filament length analysis. For each pixel in the skeleton of filaments, a fit line was produced (using the MATLAB polyfit function) using the next 5 points in each direction from the pixel of interest. From this quadratic, the normal was found and the intensity of the pixels in the normal were summed. After iterating this over the skeleton, the calculated line profile of each filament were Fourier transformed using the MATLAB fft function and the resulting distributions across all filaments were summed to produce the average frequency spectrum for that surface protein.

Aggregating STORM images

Colocalised clusters (virus particles) between channels were found by clustering and fitting best fit ellipses to data (see 2.4 and finding clusters whose centroids were within 50 nm of each other. The position of localisations in these clusters

were taken with reference to either the centroid of one cluster or the centroid of the cluster that a localisation belonged to. These were plotted on heat maps with a pixel size of 3 nm and overlaid in ImageJ.

2.5 Discussion and conclusion

The small size and pleomorphic nature of viruses precluded their visualization by single-particle averaging methods or conventional fluorescence imaging, making it difficult to measure protein distributions at the single-virus level across an entire population. The high throughput required to investigate populations precluded electron microscopy. In this work, we have shown that a combination of super-resolution imaging and high-throughput analysis of virus particles allows for the size and shape of multiple virions to be examined at a time. Using the highly pleomorphic influenza strain A/Udorn/72 as a proof of principle, we imaged thousands of filamentous and spherical virions, allowing the features of a large population of virus particles to be ascertained at high-resolution.

Large viral filaments that measured over 250 nm in length were analysed using widefield microscopy; analysis of the size distribution of over 40,000 filamentous virus in this way showed a broad range of filament sizes- fitting previous observations that filaments are microns in length, with some exceeding 30 000 nm [10]. The frequency of filaments decreased with length, suggesting that long filaments are easily broken or that production of longer filaments may be restricted by the need for the large amounts of viral protein required to form each filament compared to a spherical particle. Virion size at short length scales was investigated using super resolution microscopy, where the analysis revealed two distinct size distributions, in keeping with previous observations that influenza forms both spherical and bacilliform particles[77]. In contrast to previous analyses[23], axial ratio analysis did not reveal distinctive populations, instead observing just a single population of viruses, perhaps because our analysis of such a large number of virions captured a greater heterogeneity in particle shape than previously observed. We concluded

from this that there exists a mechanism to produce large or small virions but that selection does not occur due to axial ratio.

The super resolution method of choice in this study was dSTORM, a single-molecule localization microscopy technique that exploits the ability of photoswitchable fluorophores, attached to antibodies that bind to specific viral protein targets, to stochastically blink. As the number of fluorophores that are in the 'on' state is carefully controlled, only a sparse subset is imaged at any one time; hence signals from individual fluorophores can be spatially discriminated. dSTORM achieves super-resolution via the precise localization of the signals emitted from individual molecules by fitting a Gaussian profile, which allows the reconstruction of an image with a localization precision below the diffraction limit. We have identified a number of possible sources of error in this methodology that may contribute to small overestimates or underestimates in our analyses: i) the use of antibodies, which can add an additional linkage error between reporter and epitope which has been estimated to be 10–15 nm[91], ultimately limiting measurement accuracy; ii) the localization error from the Gaussian fit to each localisation during super-resolution image reconstruction, which we calculated to be around 7 nm in our experiments (Fig. 2.10); and iii) the fitting error from the ellipse fitted to localization clusters to estimate virion size, which may be an underestimate due to the labelling density being greater in the center of a cluster than around the exterior. In spite of these potential sources of error the average sizes that we obtained from our analysis of many thousands of influenza and SARS-CoV-2 particles were in agreement with data obtained through other imaging techniques, confirming that fluorescence super resolution microscopy offers an excellent choice for the rapid and specific study of large numbers of pleomorphic structures without particle averaging.

My analysis demonstrated large variability in filament length, as well as a lack of distinctly sized sub-populations within a filamentous virus population. In influenza, the initiation of virus budding is thought to be initiated by clustering of HA and NA in lipid raft domains on the cell surface, followed by recruitment of M1 which serves as a docking site for the RNPs just below the plasma membrane[92].

Elongation of the budding virion is caused by polymerization of the M1 protein, and the viral M2 protein is thought to localize at the periphery of the budding virus through interactions with M137. Finally, membrane scission, mediated by M2, leads to release of the budding virus[93]. The regulation of filament formation is not fully understood, but appears to be a complex process driven by at least M1 and M2, and possibly other viral or cellular proteins. Our analysis suggests a model of virus particle assembly where the release of progeny filaments from cells occurs in a stochastic way, leading to filaments of a wide range of lengths. This model is supported by a previous observation that influenza produces virions with broad variations in size and protein composition, which led to the suggestion that phenotypic variability may contribute to virus survival under stress conditions such as in the presence of antiviral NA inhibitor drugs[94].

The analysis on the average structure of a virion was a useful technique that worked with STORM or PALM data. It expanded on the work of Laine et. al. who studied the positions of different protein layers in fixed Herpes-Simplex Virus type-1 (HSV-1) using STORM [95]. From the averaged structure, there was clearly no drift during an acquisition. There was a channel mapping error though this did not affect particle size calculations. The fluorophore attached to the secondary antibodies affected the point spread function size and photons collected- affecting the localisation error. By swapping the secondary antibodies, this could be controlled in the future. This method is broadly applicable and could be used to analyse images of many structures including other viruses and bacteria.

STORM microscopy has been used to study virus particles sizes previously. In HIV, it has been used to compare free virions to those that had entered host cells to understand the change in capsid dimensions during viral entry [96]. SARS-CoV-2 was shown to be transported in clathrin-coated vesicles from the cell surface when STORM images of the lower membrane were taken, this also allowed particle measurements to be taken [97].

A limitation of this work is the lack of functional information on the virus populations. It would be of interest to ensure the filaments seen were infectious and to link the pleomorphisms seen to the pathogenicity and infectivity of the virions.

In this work, I have shown that high-resolution microscopy and rapid automated analysis software can be used to gain structural information at the single-virus level on thousands of virions at a time. This type of analysis can easily be extended to the study of other pleomorphic viruses or structures, and so our methods may also be useful for future studies of virus assembly and budding, cell-to-cell transmission and virus pathogenesis, or the development of viral diagnostic strategies that rely on virus morphology for identification.

3

Viral organisation and virus-cell interactions

Contents

3.1 Introduction 50
3.2 The spatial organisation of viruses 51
The spatial frequency of filamentous surface proteins 51
The spatial organisation of filamentous virus RNA 54
3.3 Fixed cell imaging of virus-cell interactions 58
Confocal imaging of virus-cell interactions 58
The study of NA polarisation in budding filaments 59
Imaging of filamentous virus budding 60
3.4 Labelling of virus particles through the infection of labelled cells 61
3.5 Methods and materials 65
Fluorescence in situ hybridisation 65
Virion simulation 66
FISH image analysis 66
Infected cell labelling 67
Confocal microscopy of infected cells 67
Labelling of virus particles through the infection of labelled cells 68
3.6 Discussion and conclusion 68

3.1 Introduction

I hypothesised that the plomorphism of viruses and the patterning of surface proteins was broadly variable and could be studied using my imaging protocol, and so have expanded the pipelines developed in the previous section to investigate the organization of influenza H3N2 A/Udorn/72 filaments. These analysis pipelines are versatile and can be adapted for use on multiple other pathogens. The ability to gain nanoscale structural information from many thousands of viruses in just a single experiment is valuable for the study of virus assembly mechanisms, host cell interactions and viral immunology, and should be able to contribute to the development of viral vaccines, anti-viral strategies and diagnostics.

Also, I have imaged infected fixed cells, specifically investigating the protein organisation and large-scale structure during budding. The study of filaments in infected cells is important to understand their role in viral infection.

In order to study anything using SRFM, they must first be fluorescently labelled. Many labelling strategies, including immunofluorescence previously described, produce a background within the sample studied. I present a novel viral labelling strategy that labels virus membranes during viral budding from host cells, producing fluorescently labelled virus particles that need no future labelling, these will be of use for future SRFM viral studies.

In section 3.2 I show that there is no spatial patterning of the filament surface proteins and that the filament RNP is polarised and found in a bulbous structures called an Archetti body when they are present in the virion. In section 3.3, I image infected cells, showing NA enrichment in their distal ends and novel structures along filament bridges and finally, in 3.4 I show the novel labelling strategy with multiple dyes.

Rebecca Andrews carried out the FISH imaging and carried out the FISH analysis routine and Sohail Ghani modified the appearance of the simulated data.

The work presented in this chapter is in preprint as Ref. [84].

3.2 The spatial organisation of viruses

The spatial frequency of filamentous surface proteins

The distributions of the HA and NA glycoproteins over the virion surface are not entirely uniform; indeed, when I stained filaments for either HA (Fig. 3.1A) or NA (Fig. 3.1B) the resulting super-resolution images showed a distribution of clustered protein signals which I hypothesised were produced stochastically. I therefore used intensity analysis and Fourier transforms to investigate the patterning of surface proteins on a large number of filaments. A/Udorn/72 was immobilised, labelled and imaged using dSTORM; the filaments were fit with skeletons as above and the intensity profiles of filaments were found by fitting the skeleton with a polynomial, finding the normal, summing the intensity over the normal and plotting this over the length of the filament (Fig. 3.1C). Finally, the overall patterning was shown by summing the normalised Fourier transforms of all of the filaments, in order to form an average distribution of the patterning.

We started by analyzing the frequency distribution of the HA protein. The resulting distribution showed no peaks apart from the high signal at close to 0 nm^{-1} which is due to the spatial frequency of the filament as a whole (Fig. 3.1D). This suggests that there is no set spatial frequency at which HA proteins are located along influenza virions, which in turn suggests that the surface virion proteins are not patterned deterministically. The broad distribution at low frequencies indicates that individual filaments can have patterning at a certain spatial frequency, but it is not common among all filaments. Similar analysis was conducted for the NA protein (Fig. 3.1E), with the same conclusion that no set frequency of NA spatial patterning occurs on influenza filaments.

Monte Carlo simulations of virus particles

To confirm our observation that no spatial frequency arrangement of the surface proteins occurs, representative filamentous virions with a length of 500 nm and a diameter of 80 nm [5] were simulated through Monte Carlo methods (Fig. 3.2).

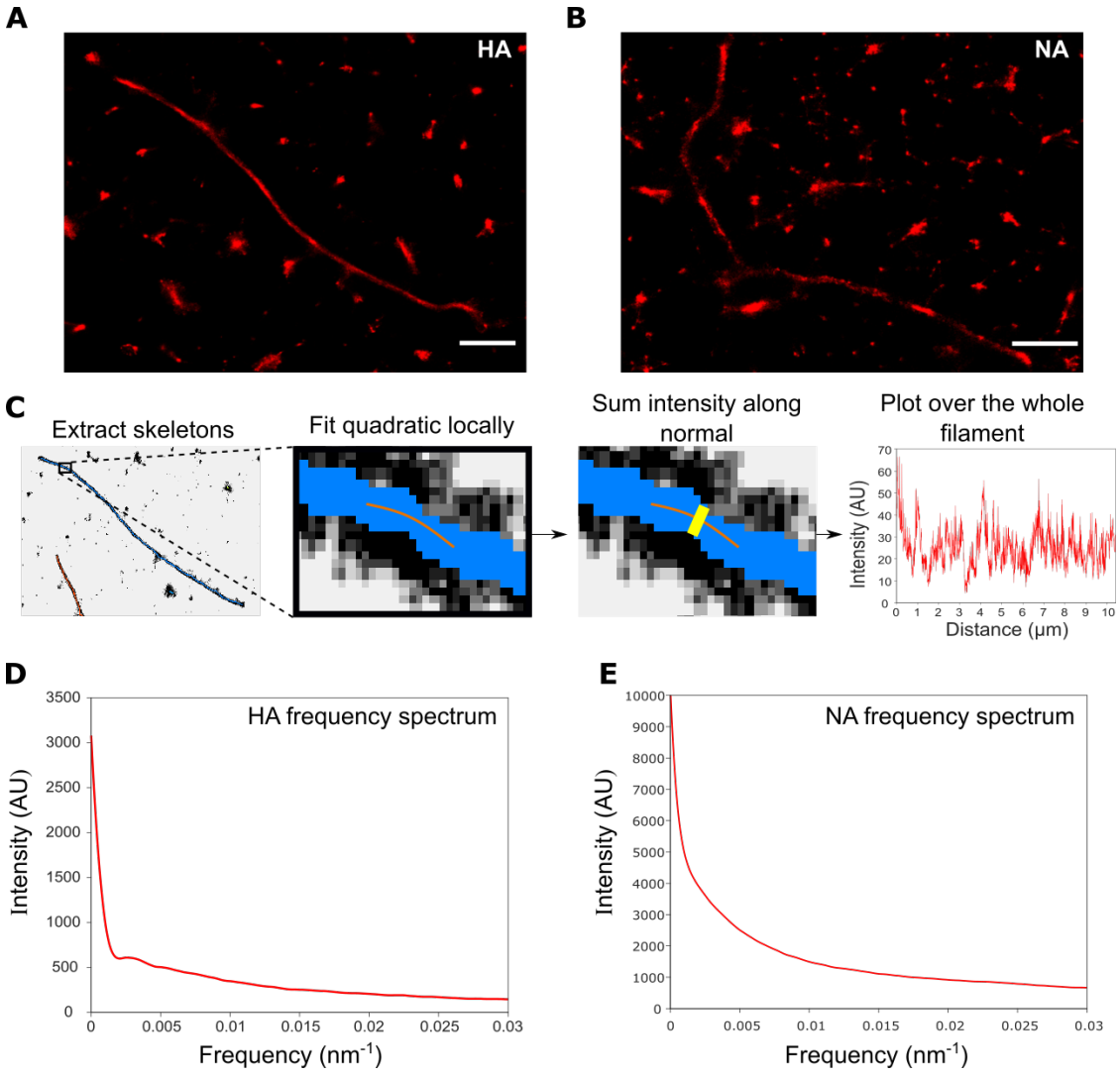


Figure 3.1: Analysis of the frequency distribution of influenza surface proteins. A&B) Representative super-resolution images of filaments labelled with A) HA antibody and B) NA antibody. Scale bars 1 μm. C) A pipeline for analyzing the influenza filament protein distribution using multiple super-resolution images. Super-resolution images are skeletonized, fit with quadratics at a point along the skeletons; the normal is found and the intensity across this normal is summed and the Fourier transform is taken and summed across filaments. D) The frequency distribution of the HA proteins from 1,067 filaments from 8 FOVs. The frequency spectrum for HA has no distinctive peaks, suggesting that there is no common spatial patterning of HA across filaments. E) The frequency spectrum for NA also has no distinctive peaks, suggesting that there is no common spatial patterning of NA across filaments.

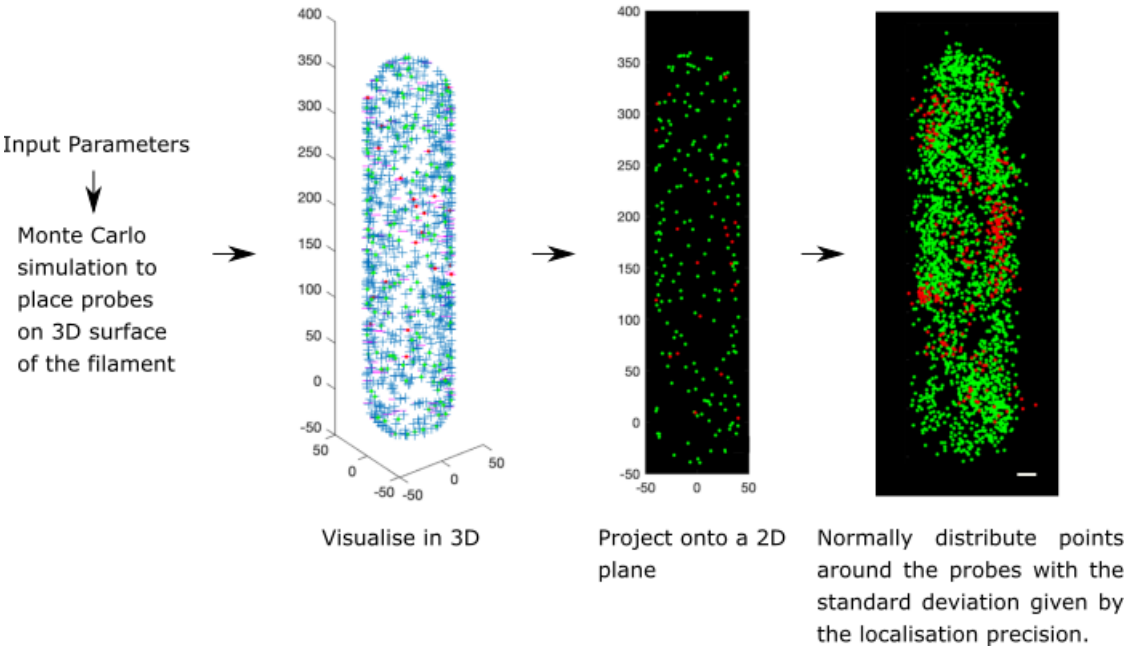


Figure 3.2: Summary of the dSTORM simulation method. Input parameters of virion size, number of localisations and localisation precision were used for Monte Carlo simulations to create simulated dSTORM images of filamentous virions. Filaments were modelled as cylinders with hemispherical caps. After projecting the simulated filaments into 2D, localisations were randomly placed with a normal distribution about the protein location with a standard deviation of 7.4nm and the image was coloured as a STORM image. Scale bar 20nm.

Initially, points corresponding to the HA and NA proteins were positioned randomly on a cylinder with hemispherical caps; by projecting these localisations into two dimensions and randomly fitting points with a normal distribution around the protein positions, simulated dSTORM images could be built up (Fig. 3.3A). For comparison, we simulated dSTORM images of filaments with a forced frequency of protein alternation, representing filaments with a deterministic protein placement (Fig. 3.3B). The relative positioning of the surface proteins in our simulations was confirmed by plotting the HA intensity profiles of representative simulated filaments with either randomly distributed HA and NA, or with alternating HA and NA (Fig. 3.3 C&D). Next, we investigated the spatial patterning of these simulated filaments by computing the sum of the Fourier transforms of the histograms of protein location along 1000 randomly distributed, and 1000 alternating, simulated filaments. The frequency distribution for filaments with completely randomly chosen

points (Fig. 3.3E) was similar to the distributions obtained from experimental data for the HA and NA proteins (Fig. 3.1D&E), with no peaks and a broad spread of frequencies at low numbers. The plot for simulations with a forced alternation of 0.01 nm^{-1} however, showed a peak at 0.01 nm^{-1} and a further small peak of the harmonic at $0.02nm^{-1}$ (Fig. 3.3F).

To ensure the simulations could handle higher and lower frequencies and simultaneous frequencies, the simulations were repeated with alternations at 0.024 nm^{-1} and mixed frequencies at 0.007 nm^{-1} and 0.033 nm^{-1}. These frequencies were chosen so that there was a range that could not be mistaken for harmonics and could be seen with virions of $\sim$1 μm with our localisation precision. The analysis of these simulations is shown in figure 3.4 with the peak at 0.024 nm^{-1} in figure 3.4A clearly showing the patterning simulated. Figure 3.4B contains a peak at 0.007 nm^{-1} and a harmonic at 0.014 nm^{-1} which seem to be picked out for these low frequencies with a small peak at 0.033 nm^{-1} also clearly picked up, showing the ability to pick out multiple frequencies.

Together, this data suggests that no deterministic spatial patterning of the HA and NA glycoproteins occurs over a large population of filaments. This HA and NA are recruited to the cell membrane before budding, but no patterning or organisation occurs during filamentous budding, with both spherical and filamentous virions lacking patterning apart from some clustering of NA [5].

The spatial organisation of filamentous virus RNA

Electron microscopy images of bacilliform and short filaments of influenza have suggested that the RNPs that contain the viral RNA cluster at one end of the virus interior, however clear images of the genome have only been obtained in a minority of longer filaments (reviewed in [10]). In order to clarify the location of RNPs within filaments we used single-molecule fluorescence in situ hybridization (smFISH) to specifically label the influenza NA segment RNA [98] combined with antibody labelling against the HA protein to outline the filament shape. smFISH uses an array of fluorescently labelled DNA probes that bind to several sites on

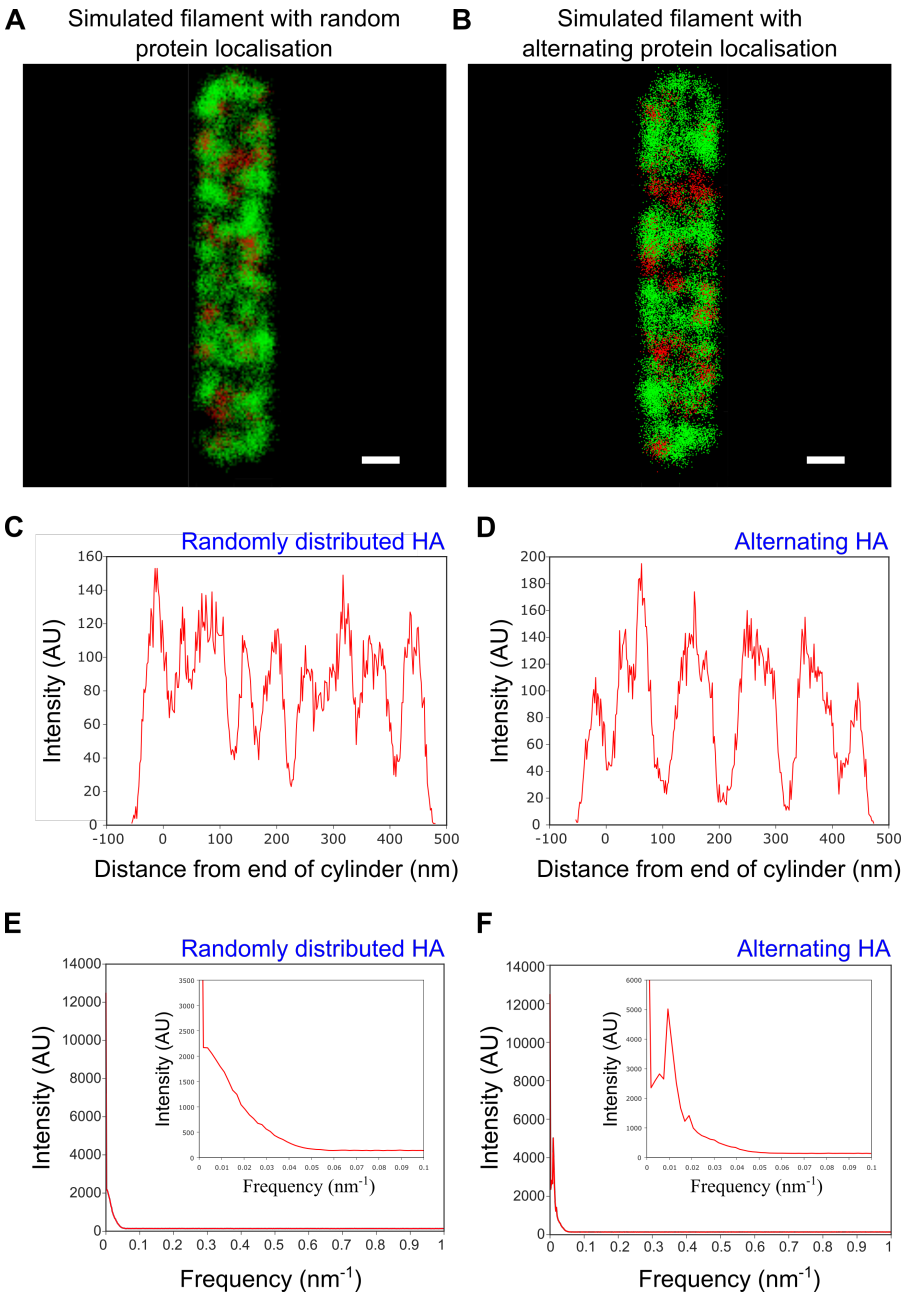


Figure 3.3: Filamentous virion simulations and analysis of alternation on simulated dSTORM images. A&B) Representative simulated dSTORM images of filamentous influenza particles with A) random protein placement and B) 0.01 nm^{-1} frequency alternating protein placement. HA is shown as green and NA as red, on a filament of 500 nm in length and 80 nm in width. Scale bars 50 nm. C) Intensity profile for a representative simulated filament with randomly distributed HA and NA. D) Same as C) but for filaments with alternating HA/NA. E) The sum of the Fourier transforms of the intensity plots of 1000 simulated filaments with randomly distributed HA and NA, suggesting no common patterning exists. F) Same as E) but for filaments with alternating HA/NA. The peak at 0.01 nm^{-1} and harmonic at 0.02 nm^{-1} confirm that protein patterning can be detected using this method.

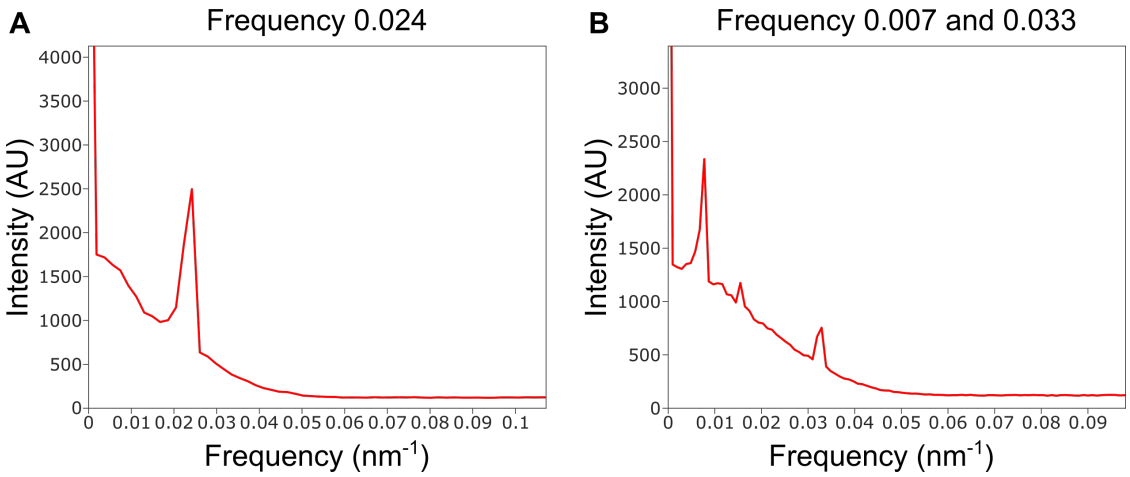


Figure 3.4: Protein patterning analysis of simulations can pick up alternations smaller than the filament length and also patterning of multiple frequencies. A) Frequency distribution from 500 nm simulated filaments with protein frequency patterned at 0.024 nm^{-1}, the limit of visible frequencies occur when the length scales are on the order of half of filament lengths. B) Frequency distribution from 1000 nm simulated filaments with protein frequency patterned at 0.007 nm^{-1} and 0.033 nm^{-1} show patterning at multiple frequencies can be seen. A small peak from a harmonic of the 0.007 nm^{-1} frequency is also seen.

the viral RNA, thereby accumulating several fluorescent dye molecules on a viral RNA, making it visible as a bright spot. A/Udorn/72 virus was immobilized and stained for HA (green) before being incubated overnight with Quasar 670 (red)-labelled FISH probes. Widefield images of the sample showed multiple long filaments of up to several microns in length (green), which were then overlaid with super-resolution localisations from the FISH probes against the NA gene segment (red) (Fig. 3.5A&B). 83 filaments were imaged in total, of which 15 were manually discarded as they overlapped with other filaments. Out of the remaining 68 filaments, 14 (20.1%) of these showed a distinctive bulbous shape at one end in the green channel, corresponding to Archetti bodies (Fig. 3.5A; white arrows), while the remaining filaments had no visible bulge at the end (Fig. 3.5B).

In order to determine the position of the RNA in each filament, a line was drawn along the axis of the filaments (using the green HA signal to approximate the filament outline) and the intensity profiles of the line in both the green and red channels was measured. The averaged intensity of the HA signal from all 68 filaments analysed was uniformly distributed over the length of the filaments

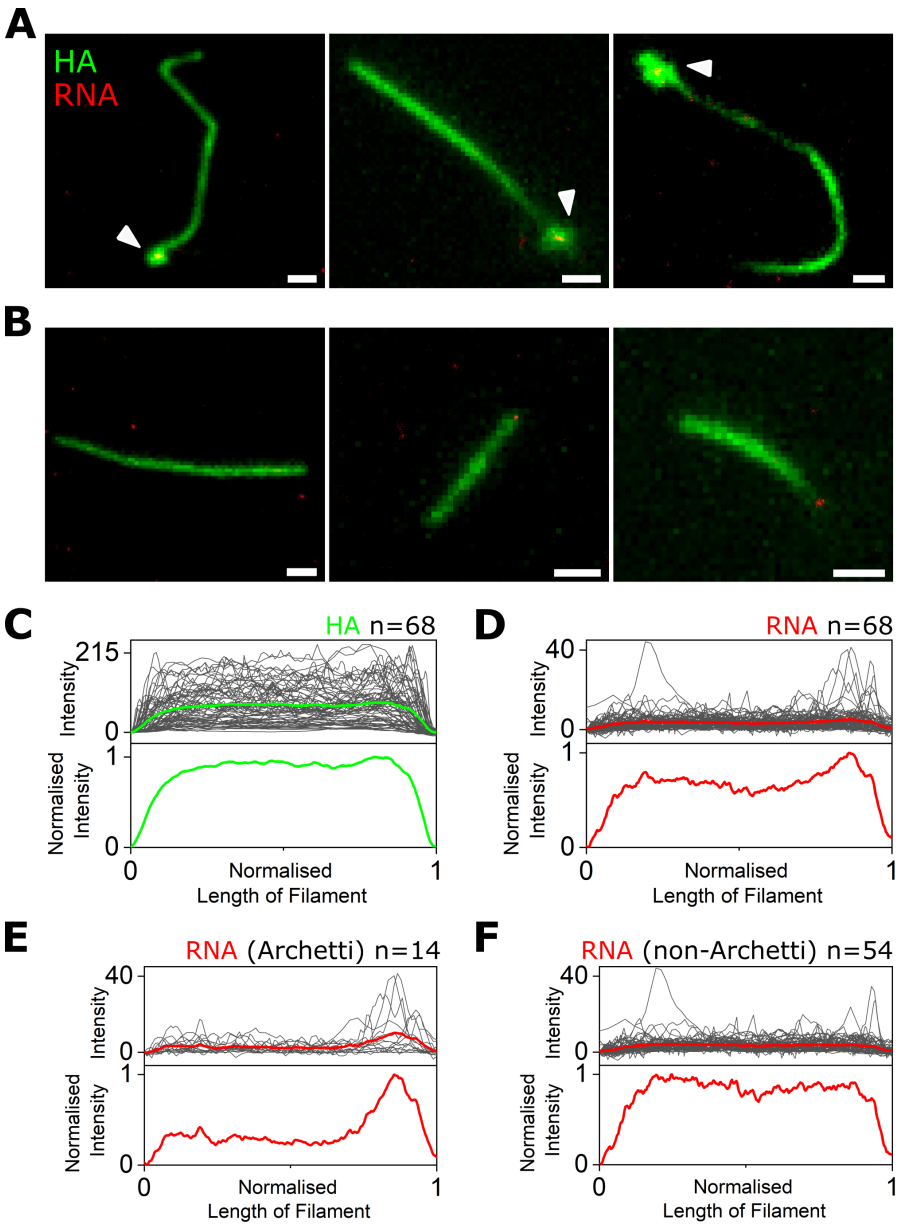


Figure 3.5: RNA is located at one end of influenza filaments. A&B) Diffraction limited images of filamentous A/Udorn/72 stained with an anti-HA antibody (green) overlaid with super-resolution localisations from an array of FISH probes against the NA gene segment (red). Scale bars 1 μm. White arrows denote Archetti bodies. C) Top: Intensity traces (grey) of the HA signal from 68 filaments, with the average intensity profile shown as a green line. Bottom: Normalised average HA signal from 68 filaments. D) Same as C) but for the red super-resolution signal corresponding to NA RNA from the 68 filaments. E) Top: Intensity traces (grey) of the RNA signal from 14 filaments with a visible Archetti body, with the average intensity profile shown as a green line. Bottom: Normalised average RNA signal from the 14 filaments. F) Same as E) but for the RNA signal from 54 filaments with no visible Archetti body.

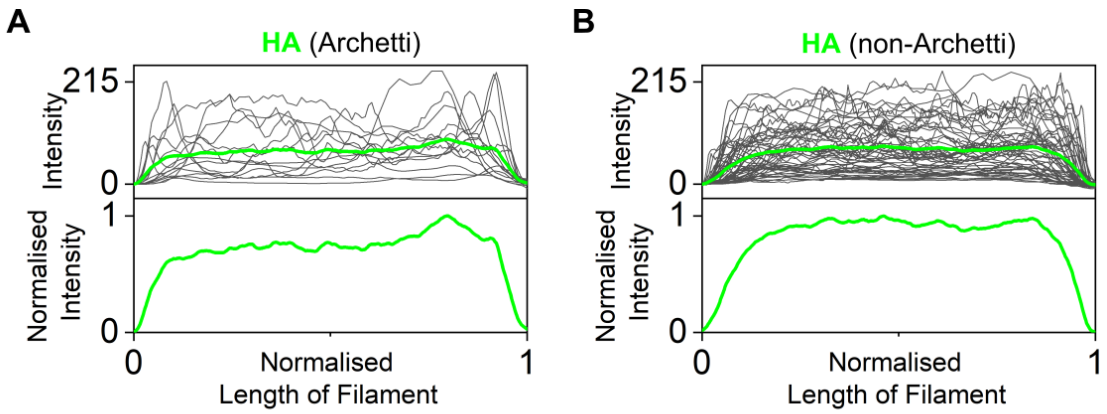


Figure 3.6: HA distribution of filaments shows when Archetti bodies are present. A) Top: Intensity traces (grey) of the HA signal from 14 filaments with a visible Archetti body, with the average intensity profile shown as a green line. Bottom: Normalised average RNA signal from the 14 filaments. B) Same as A) but for the HA signal from 54 filaments with no visible Archetti body.

(Fig. 3.5C), however, the averaged intensity traces of the super-resolution signal corresponding to NA RNA from the 68 filaments showed distinct peaks with a tendency to be located towards the filament ends (Fig. 3.5D). This trend was only obvious for filaments that had Archetti bodies (Fig. 3.5E and Fig. 3.6A), compared to those without (Fig. 3.5F and Fig. 3.6B); and all filaments with Archetti bodies had a FISH signal corresponding to viral RNA within the structure. Together, these results suggest that Archetti bodies may play a role in housing RNPs and potentially in virus transmission.

3.3 Fixed cell imaging of virus-cell interactions

After studying individual filamentous virus particles, I studied their host cell interactions in order to study their role in infection.

Confocal imaging of virus-cell interactions

Firstly, we tested whether an adaptation of the viral labelling strategy was suitable in the study of virus-infected cells. We infected MDCK cells and labelled the cell nucleus with DAPI (diamidino-2-phenylindole hydrochloride) and the viral HA

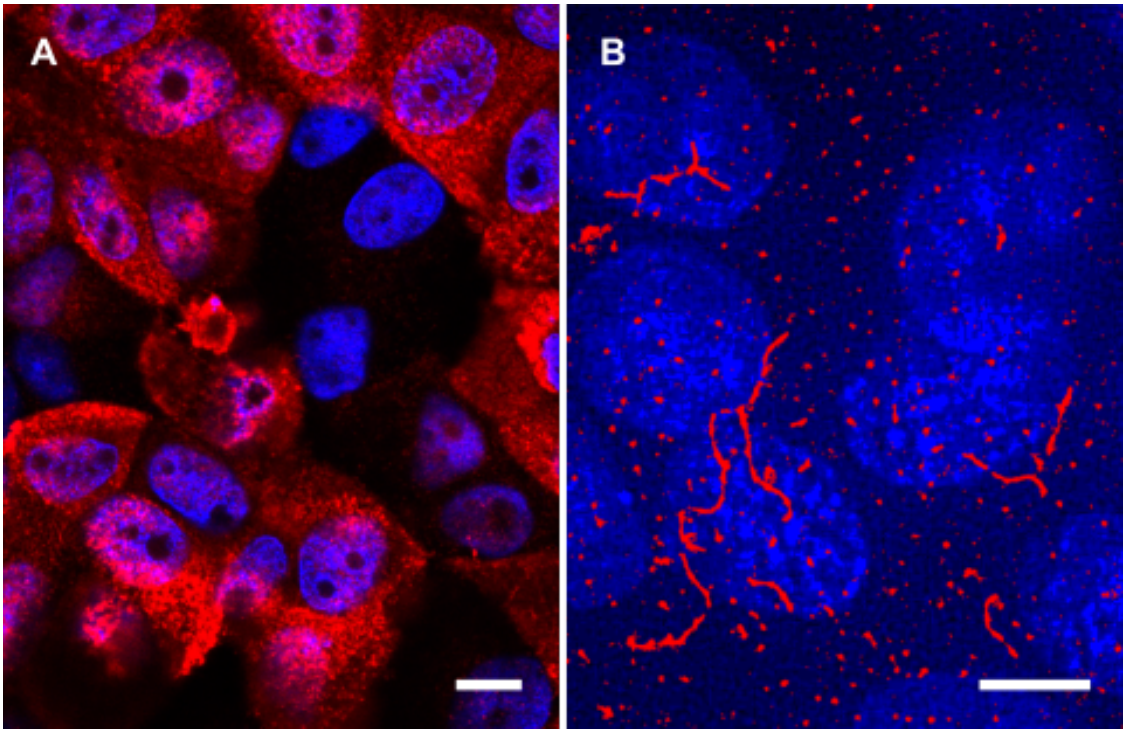


Figure 3.7: HA labelled influenza in infected cells. A) Confocal imaging of a slice through virus infected cells. Scale bar 20 μm B) Super-resolution image of filamentous influenza on the surface of infected cells. Red: viral HA. Blue: Nucleus labelling. Scale bar 5 μm.

proteins were labelled with antibodies. Infected cells were imaged with a confocal microscope (Fig. 3.7A) which clearly shows infected and non-infected cells and super-resolution TIRF images were overlain upon widefield images of cell nuclei (Fig. 3.7B) which showed filaments around the infected cells.

The study of NA polarisation in budding filaments

Previous work has shown that NA has a tendency to cluster at one of the poles of influenza filaments; however, while tomograms of Udorn virions suggested that NA clusters at the end of the virion opposite to the end where the RNPs are located[21], a recent analysis using high throughput fluorescence microscopy to visualize endogenously labelled NA within filaments revealed a tendency for NA to colocalize with NP, in turn suggesting that NA localisation is linked to the location of the viral genome[18]. We confirmed previous observations that NA was enriched at one of the filament poles by measuring the intensity profiles of fluorescence signals

from immunostained HA (green) and NA (red) in filaments budding out of infected MDCK cells (Fig. 3.8A). Budding filaments were clearly defined in the green channel (corresponding to labelled HA), whilst an overlay of the red and green signals showed filaments with largely continuous HA signal punctuated with infrequent clusters of NA (Fig. 3.8B). Using data from both channels as a guide, intensity profiles were measured by drawing a line from the tip of each filament to just before the signal broadened at the cell membrane. The averaged intensity of the HA signal from 34 filaments was uniformly distributed over the length of the filaments (Fig. 3.8C), however the averaged intensity traces of the NA signal showed a roughly 2-fold enrichment at the filament ends furthest from the cell membrane (Fig. 3.8D).

Two of the filaments had large bulges at the tip which we identified as Archetti bodies (Fig. 3.9A). Intensity profiles of these individual filaments showed a strong enrichment of NA signal at the Archetti body (Fig. 3.9B); however, we confirmed that exclusion of these two filaments from our analysis still resulted in stronger NA signal at the filament tip (Fig. 3.9C).

Imaging of filamentous virus budding

To study the budding of filamentous influenza, an appropriate time point to image infected MDCK cells was required. To do this I infected confluent cells grown in an 8 well plate at an MOI of 0.1 (see section 2.4) and imaged them at 3 hour, 6 hour, 9 hour and 24 hour time points (Fig. 3.10). Spherical virions, but no filaments were seen before 6 hours and very rarely, small filaments were seen at 9 hours. By 24 hours, many cells had a number of filamentous protrusions and so this was chosen as the time for images to be taken.

When imaging cells at the 24 hour time point, we saw multiple cells with filaments bridging cells (Fig. 3.11) that were of particular interest. Some of these filaments contained previously unidentified structures along their lengths (Fig. 3.11). By the small-scale diffusive motion, these bridges were only attached at their ends (at the cells) and had some amount of tension along their lengths, keeping them taut.

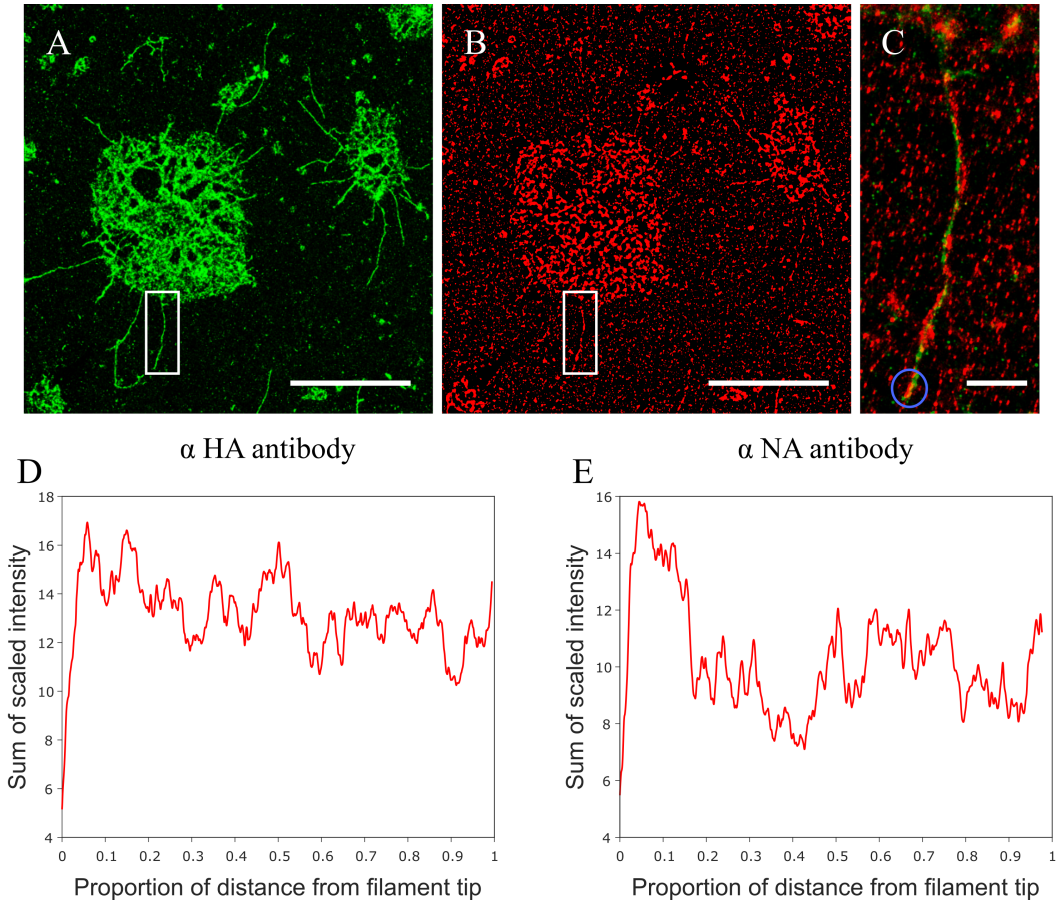


Figure 3.8: NA is enriched at the tip of budding influenza filaments. A&B) Super resolution images of filamentous A/Udorn/72 stained with for HA (green; A) and NA (red; B) budding out of infected cells. Scale bars 10 μm. C) Zoomed in merged image of filament highlighted in white box in A&B). Scale bar 1 μm. D) Normalised averaged intensity traces of the HA signal from 34 filaments showing a uniformly distributed distribution along the filament length. E) Normalised averaged intensity traces of the NA signal showing a ∼2-fold enrichment at the filament ends furthest from the cell membrane.

3.4 Labelling of virus particles through the infection of labelled cells

Virus particles can be labelled with membrane dyes. This non-specific labelling has uses such as live cell microscopy, though it has the disadvantage of also labelling other structures within a sample with a cell membrane e.g. cells, exosomes or cell debris. It was hypothesised that, as influenza virus particles bud from the membrane of their cellular hosts, a labelled cell membrane would lead to labelled virus particles that could be purified and used for other studies (e.g. infections). To test this, cells were labelled using dyes which labelled different cellular structures, infected, and then the

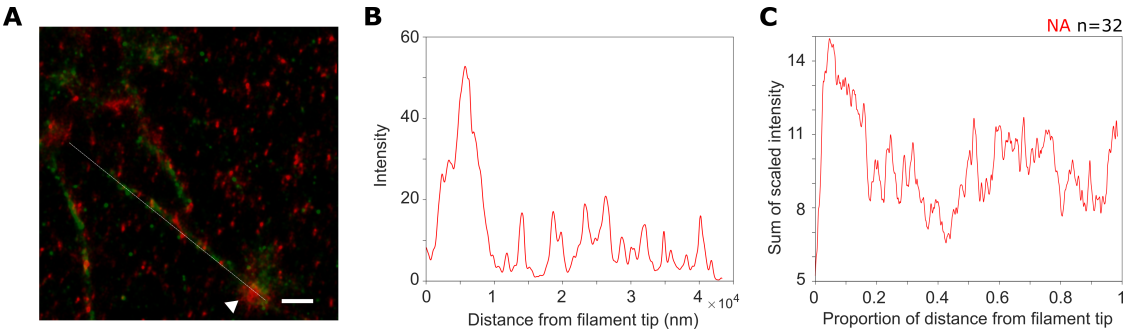


Figure 3.9: NA is enriched in Archetti bodies as well as at the tip of non-Archetti budding influenza filaments. A) Diffraction limited image of A/Udorn/72 filaments stained with for HA (green) and NA (red) budding out of an infected cell. White line denotes a filament with an Archetti body (arrow) at the tip. Scale bar 1 μm. B) Intensity trace of the NA signal from the filament highlighted in A), showing higher NA signal at the Archetti body at the tip. C) Normalised averaged intensity traces of the NA signal from filaments without obvious Archetti body structures at the tip, showing an enrichment at the filament ends furthest from the cell membrane.

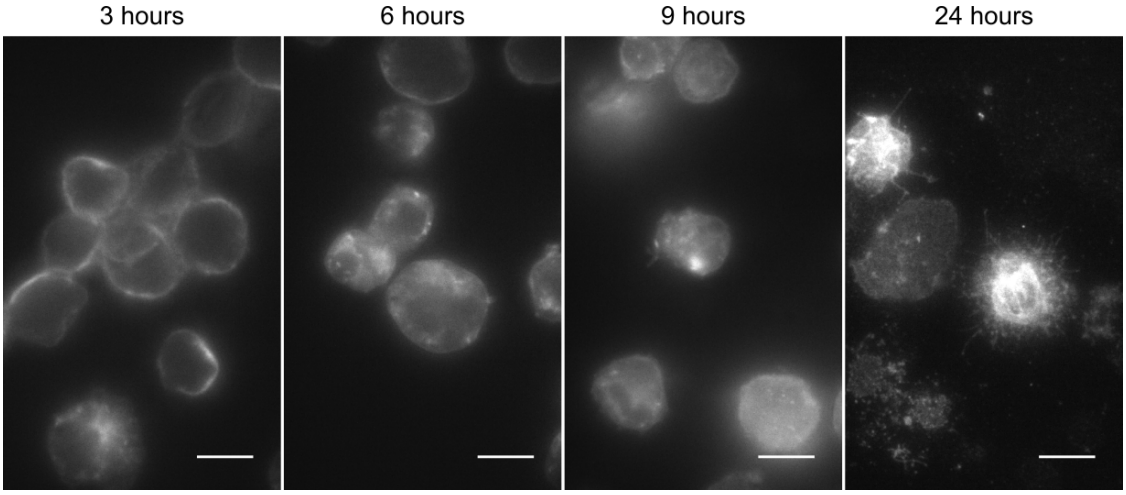


Figure 3.10: Time course of A/Udorn/72 infected MDCK cells. A time course of infected MDCK cells with time points at 3, 6, 9, and 24 hours. Many filaments were present at the 24 hour time point. Scale bars 10 μm.

supernatant was collected (section 3.5). The collected images are shown in Fig. 3.12.

As can be seen, the method clearly works effectively at producing labelled virus particles as signal is only seen when cells are infected and dye is used at the same time. Wheat germ agglutinin (wga) and FM1-43 both label cell membrane, bodipy ceramide labels the Golgi apparatus and DiOC6(3) labels the endoplasmic reticulum. As can be seen, different dyes appear to label with different efficiency. This suggests that these structures are particularly relevant in the budding of new virus particles.

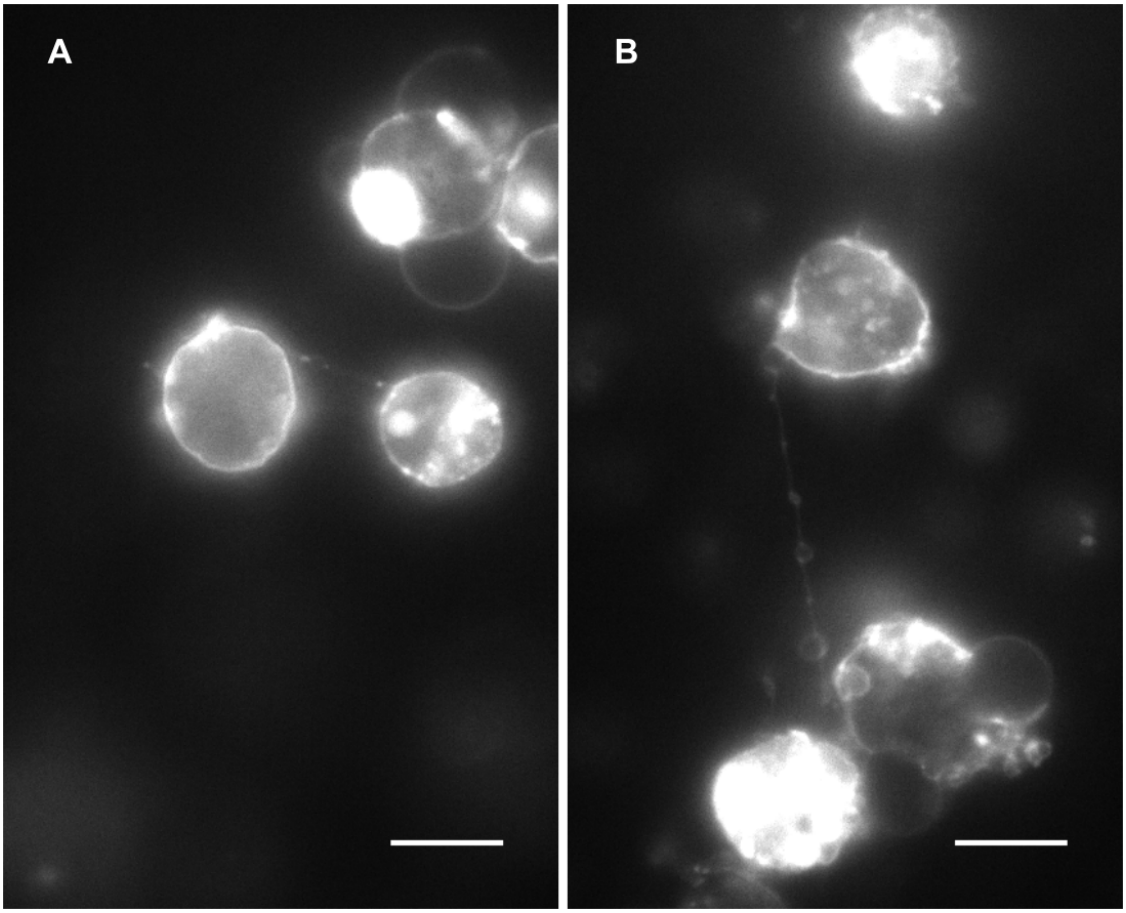


Figure 3.11: Fixed cell imaging of infected cells shows viral bridging and new structures along the viral bridges. A) Viral bridging between two MDCK cells. B) Viral bridging with unrecognised structures between two MDCK cells. Scale bars 10 μm.

Many applications of labelled virus particles require dual labelling and colocalisation to ensure a structure is a virus. To test if this was possible, cells were labelled with 2 dyes that emit at different wavelengths (wga 555 and wga 647) before being infected with A/Udorn/72. FOVs were taken of both channels and then merged in order to see colocalised particles (Fig. 3.13). As can be seen by comparing the merged image from the virus positive sample to the merged virus negative sample, colocalisations are only seen in the virus positive sample and no signal is seen in the virus positive sample that does not colocalise with the other channel. This demonstrates both the ability of the technique to label in 2 colours simultaneously but also the high labelling efficiency of the technique.

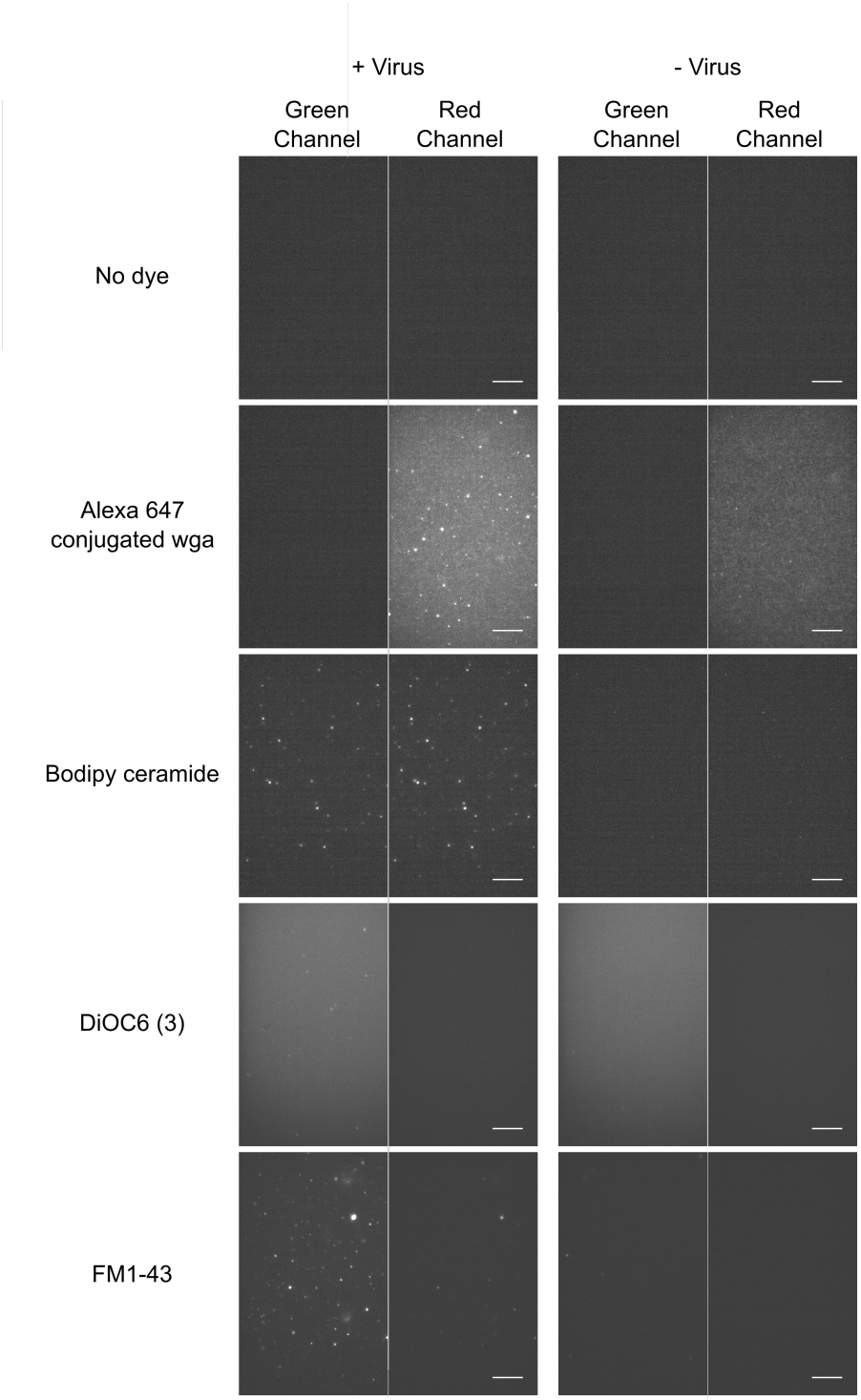


Figure 3.12: Virus positive and negative fields of view from cells labelled with different dyes. Cells were labelled with dyes and then infected with A/Udorn/72 before the supernatent was collected and imaged. Cells labelled on their membrane or their golgi apparatus clearly show positive signal that only appears when cells were infected. Scale bars 10 μm.

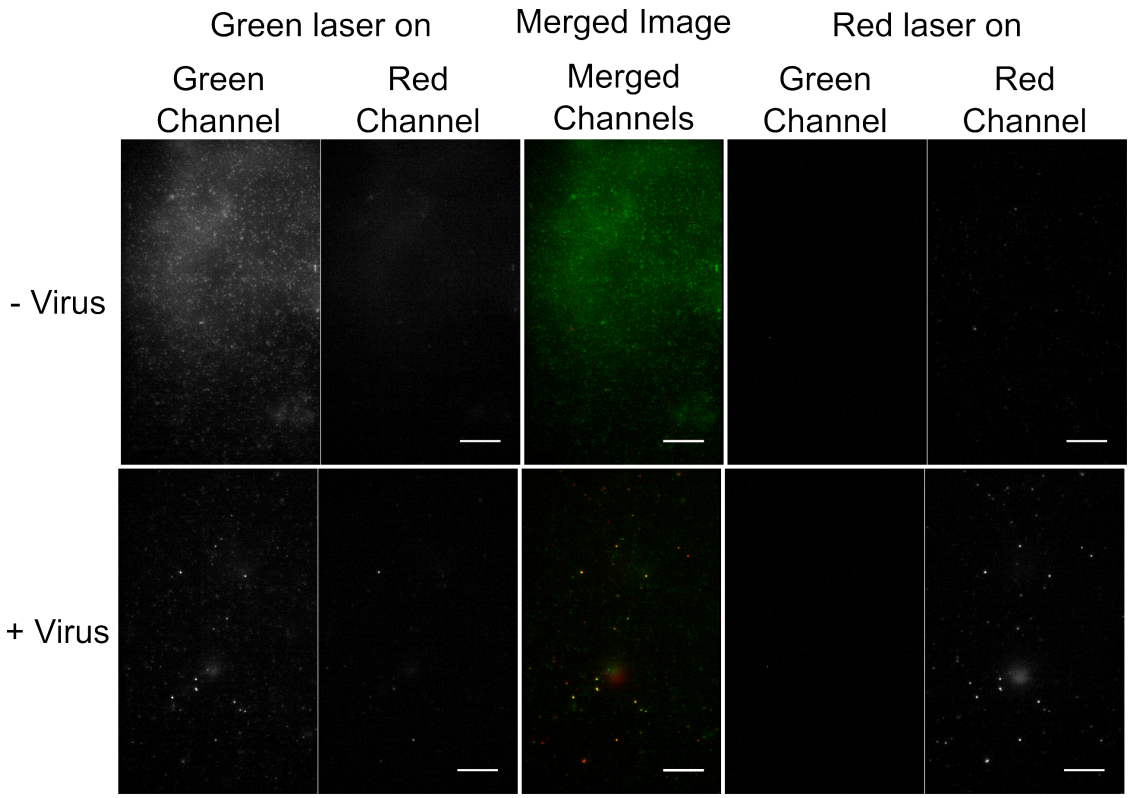


Figure 3.13: Virus particles can be dual labelled specifically using this technique. Cells labelled with wheat germ agglutinin (wga) conjugated with Alexa 555 and Alexa 647 were infected and the supernatent collected. By imaging the samples through exciting with 532 nm and 647 nm lasers separately, and then merging the collected images, we can see colocalised signal. This was only seen in the virus positive sample. Scale bars 10 μm.

3.5 Methods and materials

Viral samples were prepared and imaged as in section 2.4. Quasar 670 dye was reactivated with increasing intensity of 405 nm light (370 nm - 410 nm filtered).

Fluorescence in situ hybridisation

For fluorescence in situ hybridisation, the samples were immunostained for HA as described above, then fixed again with 4% formaldehyde in PBS for 10 minutes, washed with 2x saline sodium citrate (SSC) and subsequently permeabilized with 0.5% Triton X-100 in PBS for 10 min. After washing with 2x SSC, the surface was blocked with blocking buffer (2x SSC, 10% formamide, 10% dextran sulfate, 0.02% RNase free-BSA, 0.2 mg/ml E. coli tRNA, 1% RNasin Plus) for 30 minutes

at 37°C. The sample was then incubated with 1 μM FISH probes (Stellaris LGC; 48 probes designed against the NA gene segment and labelled with Quasar 670) in blocking buffer for 1-3h at 37°C. After hybridization, the sample was washed thrice with 2x SSC, 10% formamide, 1 % RNasin Plus and then washed three times with 2x SSC before being imaged.

Virion simulation

dSTORM images of viruses were simulated using a Monte Carlo method with the probability of a protein being placed equally across the virion and an exclusion distance of 20 nm. Filaments were modelled as cylinders with hemispherical caps. After projecting the simulated filaments into 2D, localisations were randomly placed with a normal distribution about the protein location with a standard deviation of 7.4 nm (derived from the localisation precision of our data) and the image was coloured as a STORM image. Alternating proteins were placed by choosing the random number for the protein location along the filament body from probability distributions derived from the functions |sin(kx/2)| and |cos(kx/2)|.

FISH image analysis

The inbuilt Nanoimager software was used to create dual-colour images of diffraction-limited green fluorescence and super-resolution localisations of red fluorescence of the filaments. To create the intensity along the filament figures, movies recorded at 30 ms exposure with 1000 frames were projected in FIJI to make a single summation image. In FIJI, a line was taken, with no directionality, along the axis of the filaments measuring the intensity in the green and red emission channels of the microscope. The background intensity (the first value in the intensity line) was subtracted to produce the raw intensity of the fluorescence, the length in pixels was normalised (0-1), and graphs plotted for all lines.

Infected cell labelling

Monolayers of MDCK cells in 8-well plates were infected with A/Udorn/72 at an MOI of 0.1 diluted in serum free, phenol red free DMEM supplemented with 2μg/mL trypsin, The virus was incubated on the cells for 40 minutes at 37°C with shaking every 10 minutes. The wells were then topped up to 0.5mL with DMEM and incubated for a further 24 hours at 37°C with 5% CO_2. The wells were washed with 200μl sterile PBS and fixed with 4% FA (Thermo Scientific) in PBS for 15 minutes at room temperature before another wash. The virus was permeabilised with 0.5% Triton-X-100 (MP Biomedicals) for 15 minutes and blocked, to prevent non-specific binding, with PBS+ for 1 hour at room temperature. For immunostaining: 0.1mL of the primary antibodies (provided by Dr. Jeremy Rossman, University of Kent) (1:200 in PBS+) was incubated with the cells for 1 hour at room temperature, after which they were washed three times with sterile PBS with rocking with each wash. 0.1mL of secondary antibodies (Invitrogen) (1:5000 in PBS+) was incubated with the cells for 1 hour at room temperature. For membrane staining: cells were labelled with 1 μg/mL Alexa 647 conjugated wheat germ agglutinin (wga) for one hour. Finally, they were washed five times with sterile PBS with shaking and left covered with sterile PBS for storage before imaging.

Confocal microscopy of infected cells

Confluent MDCK cells were infected at a multiplicity of infection of 0.1 and incubated at 37°C in serum-free media (Dulbecco's Modified Eagle Medium, Gibco) supplemented with 2 μg/mL TPCK- treated trypsin (Sigma) for 36 hours. Infected cells were fixed with 4% formaldehyde (Thermo Scientific) in phosphate buffered saline (PBS) for 10 minutes at room temperature, before being permeabilised in 0.5% Triton-X-100 (MP Biomedicals) for 15 minutes. The sample was blocked to prevent non-specific binding with 10% donkey serum (Sigma) and 0.1% Tween-20 in PBS for 1 hour at room temperature. Primary antibodies were diluted in blocking buffer and incubated with the sample for 1 hour, followed by washing in PBS and labelling with secondary antibodies (Invitrogen). Finally, cells were

washed five times in PBS, rinsed with water and mounted onto glass microscope slides with Mowiol containing 0.001% 4',6'-diamidino-2-phenylindole hydrochloride (DAPI) (Sigma). Cells were viewed using a Zeiss microscope with Zeiss X63/1.25 oli immersion objective and images processed using ImageJ.

Labelling of virus particles through the infection of labelled cells

MDCK were grown to a low confluency (∼20%). After 24 hours, they were labelled in serum-free, phenol red-free media (Dulbecco's Modified Eagle Medium, Gibco) with Alexa 555/647 conjugated wheat germ agglutinin (wga) (ThermoFisher) (1 μg/mL), bodipy ceramide (1 μg/mL), DiOC6(3) (1 μM) or FM1-43 (1 μg/mL). After 1 hour they were washed three times with PBS and infected, harvested and clarified as in 2.4. Particles were filtered (0.2 μm pore size (thermo)) and imaged as in 2.4.

3.6 Discussion and conclusion

In this chapter, I have used our analysis pipelines to investigate the arrangement of viral proteins to test the hypothesis that apparent patterning appeared stochastically. The analysis demonstrated that no generalized spatial frequency patterning of HA or NA on the virion surface occurs as suggested previously [18], and observed that RNPs are preferentially located at filament ends as previously suggested in cryo EM studies by Virjayakrishnan et al. [23].

Electron microscopy images of influenza virions allow NA and HA spikes on the viral surface to be distinguished by length and density, leading to the observation that HA and NA separate into distinct clusters. Our large-scale analysis showed however that there isn't a common spatial frequency patterning of HA and NA amongst filaments, suggesting that HA/NA patterning isn't driven deterministically and further supporting our hypothesis that filament phenotype is broadly variable. Limitations of this study include the lack of functionality assays on the filaments with no spatial patterning, and the small number of glycoproteins (HA and NA) studied using our analysis pipeline.

NA has been shown previously to preferentially cluster at the filament pole, either at the same[18] or the opposite[21] end to the RNPs, perhaps to facilitate virus scission and release or directional virus movement. Our analysis confirmed that there was an enrichment of NA signal at the filament tip of budding viral filaments, which has also been suggested to be the location of the RNPs[18],[19][20][21][23]. Interestingly, we also observed a strong NA enrichment in the small number of Archetti bodies found at budding filament tips. In addition, our smFISH analysis showed that RNP complexes were located preferentially at the filament pole within Archetti bodies in filaments with obvious Archetti bodies at the tip. Together, these results suggest that NA-rich Archetti bodies may play a role in housing RNPs, however, further work is needed to elucidate their potential role in virus transmission and infection. Also, to show the applicability to human infection, the experiments should be repeated in a human epithelial cell line.

Structures called tunneling nanotubes (TNTs) have previously been shown to form connections between cells [99]. This machinery may be connected to the formation of viral filaments, with evidence for the movement of viral proteins and RNA between cells as a means for viruses to escape the immune system [100][101]. However, the structures I saw have not been documented previously, the function of which need further research.

The viral labelling strategy shown has great potential to compliment other viral labelling strategies for fluorescence microscopy. It has two use cases: the labelled particles can be used for single virus tracking (SVT), and the mechanism of labelling is dependent on the membrane of the cell labelled, this gives information on the structures of importance in the budding of a viral species. For this to give cellular membrane data, intensities of the differing labels would first need to be quantified. The infectivity of the virus particles and quantification of the labelling efficiency of this technique will also increase its utility further. Future studies that can apply this method include studying the formation of lipid rafts within cells and studying viral fusion and entry.

4

Surface stress as a selective pressure for filamentous virus formation

Contents

4.1 Introduction 71
4.2 Theory 74
Model assumptions 74
The derivation of a relation between the major and minor axis 74
4.3 Results 76
The fit of the data to viruses 76
The fit of the data to bacteria 77
Filamentous virion radii are half the radii of spherical virions . 79
Growth in one, two or three dimensions and non-uniform geometries 81
The application of the model to red blood cells 83
4.4 Methods and materials 84
Super resolution fluorescence imaging and analysis 84
E.coli imaging and analysis 85
Cryo EM imaging and analysis 85
4.5 Discussion and conclusion 86

4.1 Introduction

A pressure vessel is designed to hold liquids or gases at substantially higher or lower pressures than ambient pressure. A high pressure difference requires a correct design in order to avoid catastrophic failures, and so pressure vessel models [102] are used

to assess (engineering) stress development and whether vessels of a known geometry will be able to sustain the applied internal pressure. Stress is the force applied to a material divided by the area over which the force acts before deformation occurs. By equating the force due to pressure and the equal and opposite force from the stress, a relation between stress, pressure, and the dimensions of a given cross section of the container can be derived. In chemical engineering, pressure vessel models using theoretical reaction pressures and the critical stress of a material are used to determine the geometries of reaction vessels (size and wall thickness) so that their operation falls within allowed safety margins.

A filamentous, or elongated, morphology is a common feature of many biological microorganisms, and can be observed in multiple species of fungi, bacteria and viruses. Filamentous morphologies have many diverse functions, for example, some fungi form filamentous structures known as hyphae which play a role in reproduction, while certain species of bacteria form filaments with a higher surface-to-volume ratio than that of their non-filamentous counterparts, which helps them survive under low oxygen conditions [103]. We set out to investigate whether pressure vessel models could be applied to filamentous biological structures and provide insight into pressure-stress relations, using the filamentous influenza and respiratory syncytial viruses as model organisms.

The role of filaments in influenza is particularly poorly understood. Current suggestions for a filamentous morphology includes the increased ability of filaments to penetrate mucus layers [104]; their increased directionality of travel on mucus layers compared to spherical forms, which enable filaments to infect cells further from the original host cell in a shorter time [18]; their increased resilience to UV radiation [105] or their increased ability to evade antibodies due to their increased number of surface proteins per virion [81]. In contrast, the larger surface area of a filament in comparison to a spherical virion means that significantly more proteins are required to form each progeny virion, and so overall far fewer filamentous virions than spherical can be produced by each infected cell; which may explain why spherical virions are also formed during new virion production. In the case of RSV,

released virions have been observed to be mostly filamentous, with evidence that filamentous virions represent the infectious form of the virus associated with the prefusion form of the surface F protein and are required for fusion of the virus with the host cell [73]. For IAV, filamentous virus represent an alternative, non-clathrin, infection pathway via pinocytosis [39], however other particles are infectious.

In this study, we modeled filaments as cylindrical particles with hemispherical caps, and used known results from thin-walled pressure vessels [102] to form a theory of how the major axis of a filament would change in relation to its minor axis given a critical stress [106]. This derived relationship was tested using microscopy measurements of the length and width of influenza and RSV. Overall, our modelling provides insights into how pressure and stress may influence the size and shape of biological organisms and suggests filamentation as a method of increasing surface area and volume of a single virion. This theory agrees with the model of filamentous formation as a method of evading the immune system [81].

In section 4.2 I discuss the theory behind the model, the assumptions used and the experimental predictions made by the model, and in section 4.3 I show that A/Udorn/72 and E.coli particle sizes follow the distribution predicted by the model, and Udorn and RSV filaments fit the prediction that they are approximately half the diameter of spherical virions. The model explains the shape of red blood cells; suggests the functionality of filamentous production is to increase the surface area or volume, and explains why some viruses grow as filaments and spheres, but not other shapes. In section 4.4 I detail the imaging protocol and methods used in my analyses and, finally, in section 4.5 these results are discussed in the context of the current literature.

Hafez El-Sayyed carried out preparation and imaging of E.coli and Swetha Virjayakrishnan carried out all cryo electron tomography preparation and imaging. All other work was carried out independently.

A manuscript for the work presented in this chapter is in preparation.

4.2 Theory

In order to see if a pressure vessel model was useful for biological studies, we had to derive a relation that could be tested experimentally, and make a number of assumptions so that we could use a simplified pressure vessel model:

Model assumptions

In the use of a (thin-walled) pressure model to derive the three normal stresses, a number of assumptions have been made. They are as follows:

1. The internal pressure was uniform across different virus particles of the same species as the internal pressure is equilibrated between budding virions and their host cell.

2. The surface environment across the surface of virus particles was similar, and the material was isotropic, and so the maximum stress was constant.

3. The bending stress in the virus membrane was not considered in using the thin wall approximation.

4. When cutting shapes to find a cross section, perpendicular lines on the neutral axis remain perpendicular after cutting.

The derivation of a relation between the major and minor axis

In order to investigate any physical motivation for filamentous morphologies, a pressure vessel model was formed. We did this by modelling the virus as a cylinder of length L and radius r with hemispherical caps of radius, r (Table 4.1). Taking a cross section of this and equating the force due to pressure with the oppositely directed force due to the stress in the walls of the filament we could derive the known relation for a pressure vessel (Eq. 4.1). This pressure stress relation was valid for viruses of any size, from a sphere to a filament. The relation showed that the stress increased as the radius increased for spherical morphologies. However, for a filamentous morphology, the stress in the membrane did not continue to increase as the length increased. Instead, as the filament length tends to an infinite length, the stress tends to a value twice that of the spherical particle with the same radius (Table 4.1).

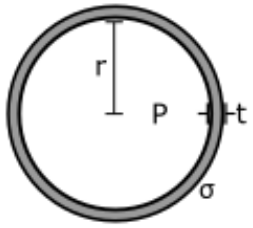
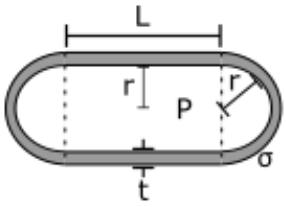
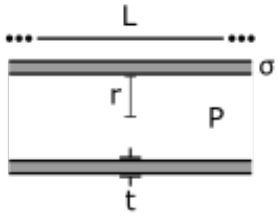
Length	Shape	Pressure/stress relation
L=0		$P(\pi r^2) = \sigma(2\pi rt)$ $\sigma = P\frac{r}{2t}$
L>0		$P(\pi r^2+2rL) = \sigma(2\pi rt+2Lt)$ $\sigma = P\frac{\pi r^2+2rL}{2\pi rt+2Lt}$
L→∞		$P(2rL) = \sigma(2Lt)$ $\sigma = P\frac{r}{t}$

Table 4.1: Diagrams of the cross sections of interest and the pressure/stress relations derived from them for differing lengths of virion models. Viruses are modelled as cylinders with hemispherical caps. For different lengths of cylinders: L=0, L>0 and L→∞, the pressure/stress relation is given along with the cross section of interest. In the limits of L→0 and L→∞, the model used for filamentous viruses tends to the known results for spheres and pipes of an infinite length.

This gave principle stress/pressure relations for the longitudinal stress, σ_1, hoop stress, σ_2, and normal/external stress, σ_3:

$$\sigma_1 = \frac{P(\pi r^2 + 2Lr)}{t(2\pi r + 2L)} \tag{4.1}$$

$$\sigma_2 = \frac{Pr}{2t} \tag{4.2}$$

$$\sigma_3 = -P \tag{4.3}$$

Where P is the difference in pressure over the surface ($P_{internal} - P_{external}$), t is the thickness of the lipid layer, r is the radius of the cylinder and hemispheres

and L is the length of the cylinder.

Assuming the breaking stress occurs when $\sigma_1 = \sigma_{max}$, $\sigma_2 = \sigma_{max}$ or $\sigma_3 = \sigma_{max}$, given $\sigma_1 > \sigma_2 > \sigma_3$, and using $r = r_\infty$ (the radius of an infinitely long filament) when L$\rightarrow\infty$, this gives us:

$$r_\infty = \frac{\sigma_{max} t}{P} \tag{4.4}$$

This gives the relationship between L (the length of the cylindrical segment) and the radius, r, in terms of radius of an infinitely long filament, r_∞, of:

$$L = \frac{2\pi r r_\infty - \pi r^2}{2(r - r_\infty)} \tag{4.5}$$

Which, using the major axis length, m_1 = L+2r, and the minor axis length, m_2 = 2r = m_1 – L, gives us a relation between the major and the minor axis:

$$m_1 = m_2 + \frac{\pi m_2 (r_\infty - \frac{m_2^2}{4})}{m_2 - 2r_\infty} \tag{4.6}$$

This relation shows a quadratic relation between major and minor axis for a value of r_∞, and that as the major axis increases, the minor axis decreases, tending to half the value when $m_2 = m_1$. This relation can be applied to any pressure vessel with this approximate geometry, though we will use it for the study of filamentous morphologies of viruses.

This gives us an relation to compare experimentally by looking at the sizes of virus particles. We can also test the model at the extremes given by table 4.1, which suggests spherical particles to reach a maximum diameter twice that of the longest filaments.

4.3 Results

The fit of the data to viruses

To experimentally test the relation between major and minor axis lengths of spherical and bacilliform virus particles, we performed super-resolution imaging of the influenza virus A/Udorn/72, a strain with a well-characterized spherical

and filamentous phenotype [21]. Virus particles were immobilised, fixed and immunolabelled, and direct stochastic optical reconstruction microscopy (dSTORM) was carried out to obtain high resolution reconstructed images (Fig. 4.1A). Both spherical (Fig. 4.1B) and filamentous (Fig. 4.1C) particles were observed in the resulting images. Minimal signal was observed in a negative control consisting of cell culture media lacking virus particles, confirming that the virus particles were specifically labelled (Fig. 4.1D).

Measurements of the major and minor axes of spherical and bacciliform particles were then taken from the super-resolution images by clustering using the DBScan algorithm and then fitting a confidence ellipse to clustered data. 41,754 viruses were measured from 46 fields of view. The final bivariate histogram (Fig. 4.1E) was produced by subtracting the distribution from negative FOVs (Fig. 4.1F) from the distribution from positive FOVs.

The final histogram showed two populations of large virions: those with a major axis equal to their minor axis, and those with a larger major axis and a minor axis approaching 140 nm (Fig. 4.1E). The boundary between these two lines followed the predicted equation (Eq. 4.6). A scatter plot of the major axis against minor axis for the 41,754 particles is shown as Fig. 4.1G with the lines $m_1 = m_2$ (red) and $m_2 = 2r_\infty$ (green) plotted and the derived relation between minor axis and major axis (Eq. 4.6) with $r_\infty = 70$ nm plotted in yellow. The predicted limit to the size of virus particles for a fixed shear stress contains the majority of the data points and follows the contour formed by the extremes of the data.

The fit of the data to bacteria

To compare our findings on viruses to another biological system, we imaged Escherichia coli (E. coli) bacteria. Bacteria can be classified on the basis of their shape, and tend to fall within five main groups: spherical (cocci), rod (bacilli), spiral (spirilla), comma (vibrios) or corkscrew (spirochaetes). E. coli is a straight, rod-shaped bacterium with a regular length of approximately 1.5 μm [107]. In a natural, unstressed population of E. coli, approximately 10% of the cells form

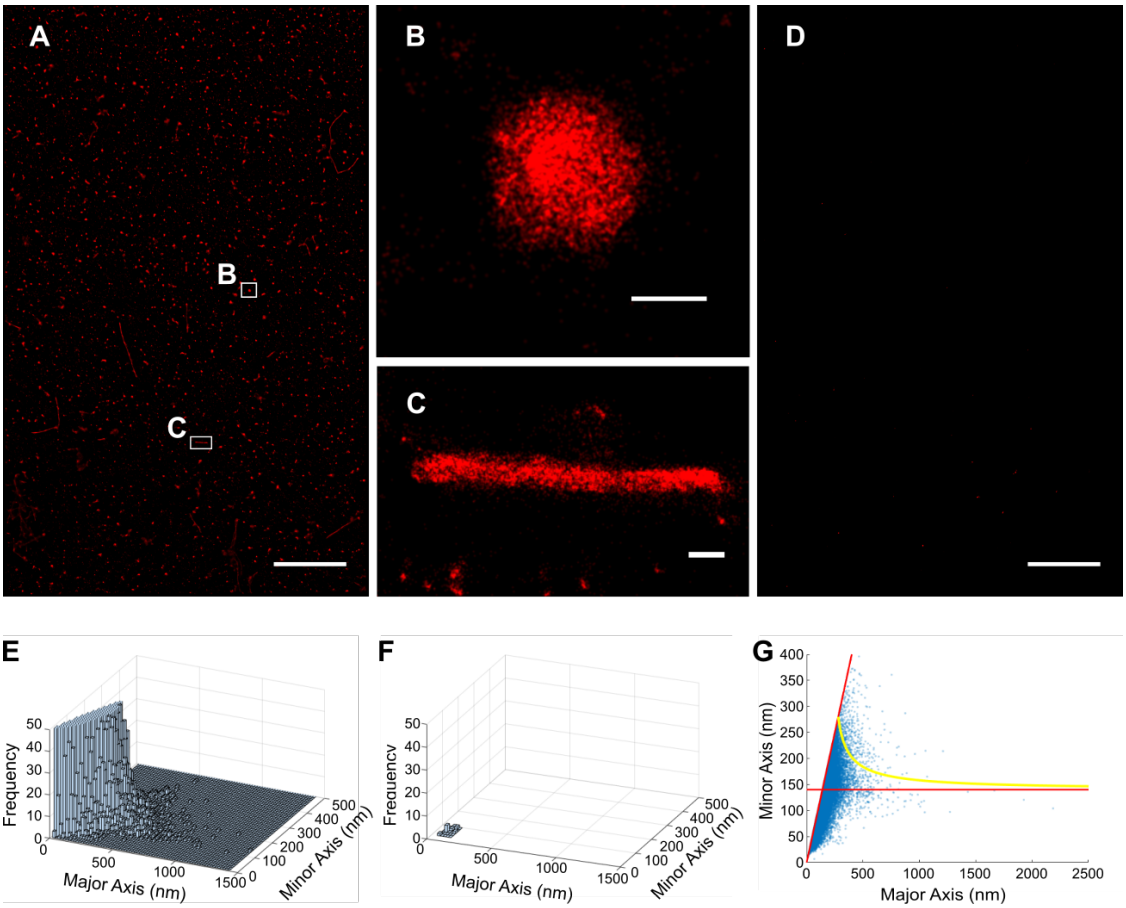


Figure 4.1: Super-resolution imaging of spherical and bacilliform influenza particles fit against the predicted theory from pressure vessel analysis. A) A representative dSTORM field of view (FOV) of labelled A/Udorn/72 influenza where HA is imaged in the red channel. Scale bar 10 μm. B&C) Zoomed in images from A) showing individual filaments and spherical particles. Scale bar 200 nm. D) A representative dSTORM FOV of a virus negative sample, imaged in the red channel. Scale bar 10 μm. E) A bivariate histogram of the major/minor axis lengths. The frequency decreases exponentially as the major axis increases tending towards a specific minor axis value. It also falls away as the major and minor axis both increase. F) A 2D histogram of the major/minor axis lengths of the single negative FOV D). All clusters are far from the boundary of interest and there are very few of them. G) The major/minor axis scatter plot with lines at i) minor axis = 140nm (red) ii) major axis = minor axis (red) and with the line as given in equation 4.6 describing the derived allowable relation between major and minor axes (Methods) with r_∞ = 70nm (yellow).

filamentous structures [108]. When stressed, the proportion of filamentous structures increases, which suggests that filaments may play a role in protecting the bacteria from predation, phagocytosis and other stressors [108][109][110]. Here, we imaged untreated E. coli (Fig. 4.2A) and in the presence of antibiotic rifampicin (Fig. 4.2B) and segmented the FOVs to measure the bacteria (Fig. 4.2C&D). By eye, we observed that rifampicin resulted in elongation of the bacteria and a higher proportion of filamentous morphologies (Fig. 4.2B).

To look at the sizes of the bacteria, cells from both treated and untreated populations were segmented with a bacterial neural network model (Fig. 4.2C&D) (described in section 4.4) [111] from which their lengths and widths were measured. A bivariate histogram of this data shows the trend for a decrease in the maximum width of the cells as their lengths increase (Fig. 4.2E) and a scatter plot with the modelled maximum size (yellow) with an r_∞ of 1.1 μm (red) shows the main density of cells fitting the model with very good agreement (Fig. 4.2F). The process of segmentation, and bacteria at the edges of FOVs led to cut filaments, leading to a decreased population of longer filaments, and an increase in the number of small bacteria and unphysical bacterial < 0.8 μm in length. As this was equally likely to cut wide and narrow cells, this did not affect the finding that longer filaments were narrower.

Taken together, our results from images of both filamentous influenza virus and rifampicin-treated E. coli suggest that elongation of viruses and bacteria (non-induced or induced through chemical stress) requires the narrowing of the particles along the minor axis.

Filamentous virion radii are half the radii of spherical virions

To further investigate the application of our derived pressure stress model to filamentous viruses, we looked at the extrema of the model when the filaments were very long (and the pressure stress relation can be modelled as that of a pipe), against which we could compare literature values for spherical virions. Our model predicted that the radius of filaments should be half the radius of the largest allowed spherical particles.

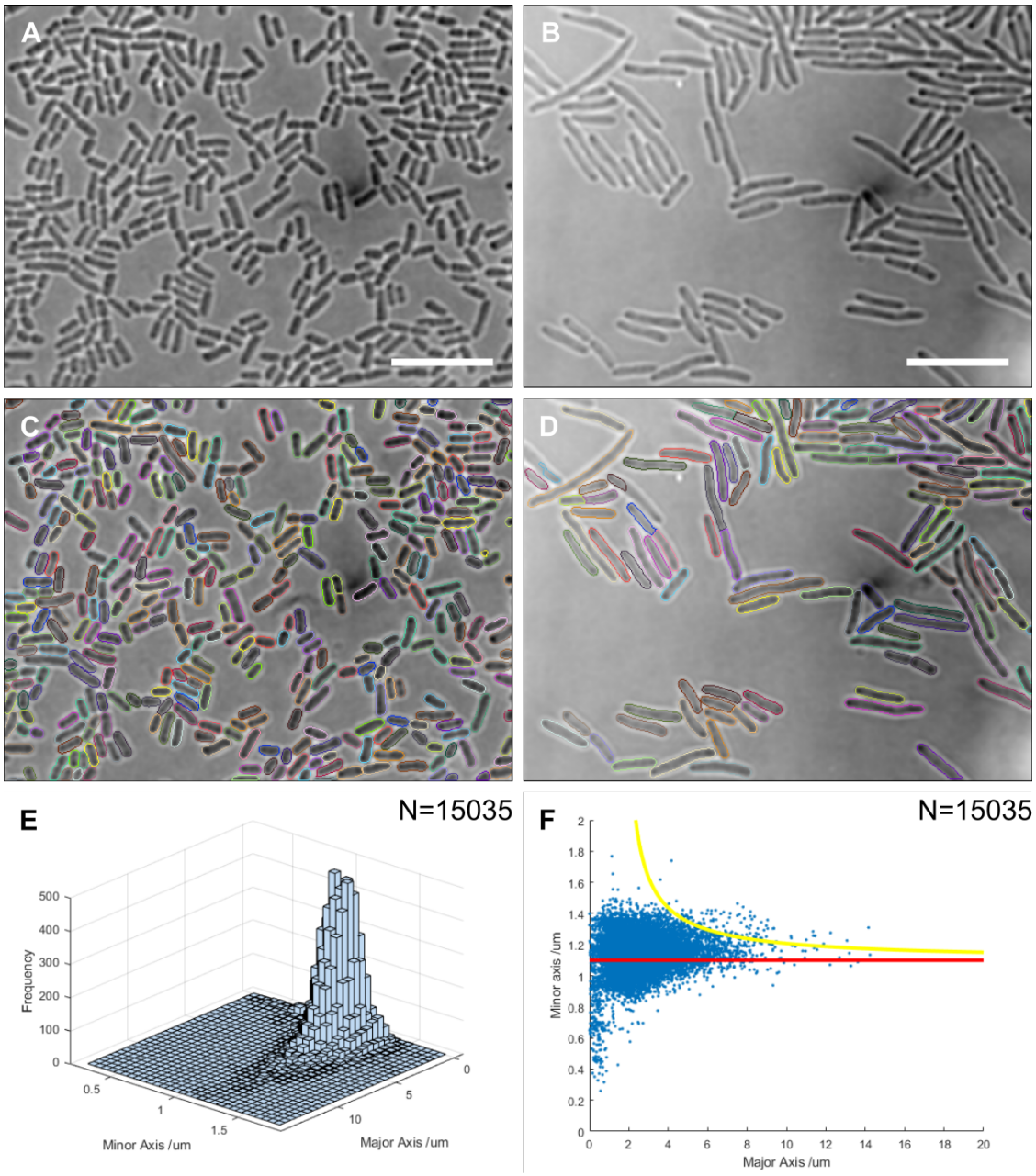


Figure 4.2: Phase contrast imaging of E.coli cells fit against the predicted theory from pressure vessel analysis. A&B) Representative FOVs of phase contrast imaging of E.coli cells when untreated (A) or treated with 50μg/ml rifampicin. C&D) Segmentations of the FOVs in A&B) from which sizes of cells were taken. E) A bivariate histogram of the major and minor -of the major and minor axis lengths with the fit of the model showing the majority of points to be contained within the expected theoretical region. F) The major/minor axis scatter plot with lines at i) minor axis = 1.1 μm (red) and with the line as given in equation 4.6 describing the derived allowable relation between major and minor axes (Methods) with r_∞ = 550 nm (yellow).

In order to get experimental data to support our model, the influenza virus A/Udorn/72 and RSV (strain A2-RSV) were imaged using cryo-electron tomography to produce tomograms of filaments with nanometer resolution (Fig. 4.3A&B). Measuring the widths of individual filaments in the resulting images five times and averaging these measurements, we found the averaged external diameter, internal diameter and thickness of the membrane for both viruses over 8 filaments of Udorn and 5 filaments of RSV (Fig. 4.3C&D). This analysis provided measurements of 55.8±0.5 nm, 37.8±0.4 nm and 9.0±0.6 nm for the external diameter, internal diameter and thickness of the influenza filament membrane respectively (Fig. 4.3C). This is in agreement with previously published values for influenza filaments e.g. 74.7 nm measured to the tips of the glycoproteins [23]. In contrast, SEM images of RSV filaments revealed that they were wider than influenza filaments, with average measurements for the external diameter, internal diameter and membrane thickness of 94.9±1.0 nm, 81.6±1.2 nm and 6.7±0.8 nm respectively. Similarly to the influenza measurements, these figures are in broad agreement with previously published values of between 71 and 166 nm [112].

The wall thickness of RSV was 7.6% of its diameter at its neutral axis (taken as the centre of the viral membrane); for Udorn it was 19.6%. This puts RSV filaments comfortably within the thin wall approximation and Udorn at the extreme of the model. Comparing these values to maximum values in the literature for spherical particle diameter for Udorn, 140 nm (including surface glycoproteins) [23], and for RSV, 250 nm (including surface glycoproteins) [113][114] we find good agreement with our model that predicts filament membrane diameter to be half of the size of spherical membrane diameter.

Growth in one, two or three dimensions and non-uniform geometries

As smooth, symmetrical geometries minimize stress concentrations, we then considered the different geometries that could be formed as spheres grew symmetrically in one, two or three dimensions to investigate the properties of filaments that

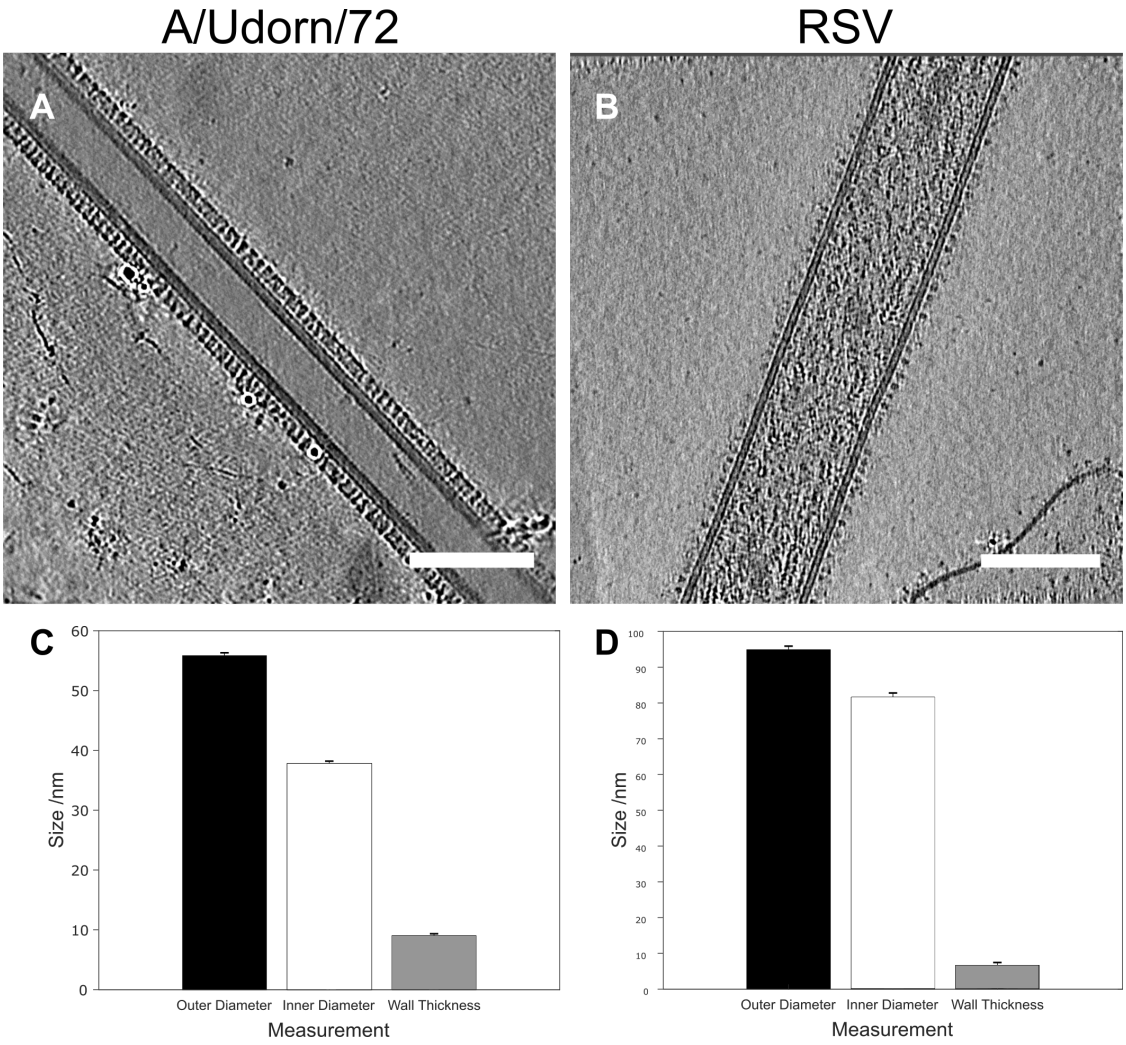


Figure 4.3: Cryo electron tomography of filamentous virus particles and measurements of their internal and external diameters and membrane thickness. A&B) Representative tomograms of A/Udorn/72 influenza and respiratory syncytial virus respectively. Scale bars 100 nm. C&D) Bar graphs of the average outer and inner diameters and the wall thickness of Udorn and RSV show RSV to be wider but with narrower membrane walls (6.4 nm compared to 9.8 nm for Udorn). Error bars: standard deviation from measuring 8 Udorn filaments and 5 RSV filaments.

may be important. In 1D, a filament is formed, in 2D a circular sheet and in 3D, a sphere (Fig. 4.4).

Increasing stress as the virus particles grow to an infinite size explained the absence of large spherical viruses, but not why large disks and non-uniform geometries were not formed. Due to the deformable nature of viral membranes, 2D growth would not be a stable geometry and particles would "balloon" to the 3D, spherical, geometry upon budding from cells. Above a certain size, as shown

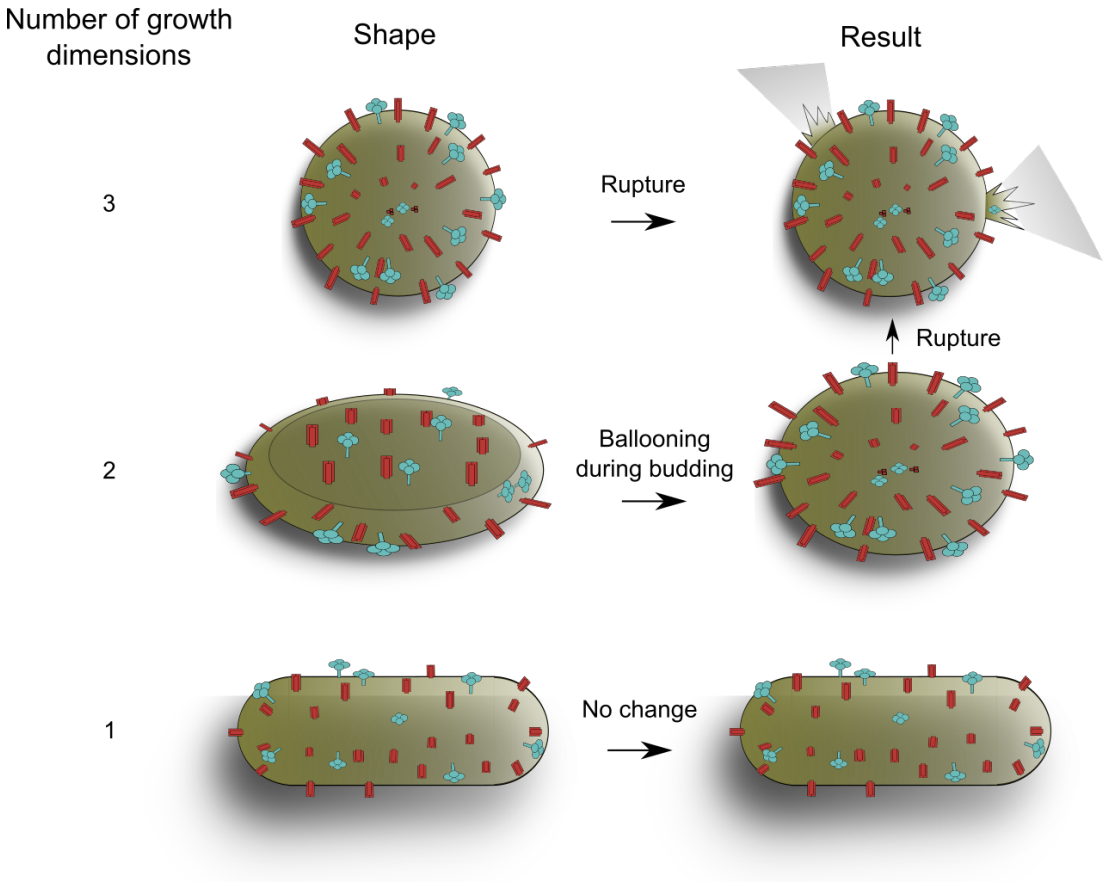


Figure 4.4: Schematic of the shapes formed from symmetric growth in 3D, 2D and 1D and the outcome for virions with those geometries. The three possible shapes of virus when they grow symmetrically in three, two or one dimension by adding material are shown. 2D growth results in ballooning during budding that results in the 3D case. The 3D case ruptures at a size when the maximum stress is reached and so a filamentous shape is the only case with no size limit.

above, these particles would burst and so 2D growth is also not viable. Non-uniform geometries produce local maxima in the stress of the membrane due to their local cross section and so they are not formed due to the limitation of the rupture stress of the viral membrane.

The application of the model to red blood cells

Red blood cells require a high cross sectional surface area and total surface area so that they move with blood pumped around the body and also rapidly transfer gases across their membrane. Though their biconcave shape is ideal for this, we wanted to see what effect this had on the membrane stress. We compared the

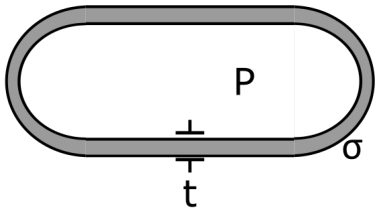
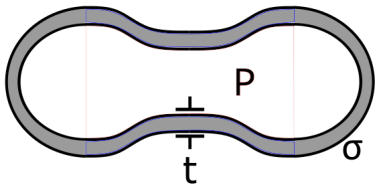
Shape	Pressure/stress relation
Perimeter = L Area = A  	$\sigma = \frac{PA}{Lt}$
Perimeter = L + δL Area = A - δA  Cross-section of a red blood cell	$\sigma = \frac{P(A - \delta A)}{(L + \delta L)t}$

Table 4.2: Diagrams of the cross sections for a red blood cell compared to a disk shape. The stress relation for the biconcave geometry of a red blood cell is lower than that of the disk geometry, allowing red blood cells to be deformed more because of their shape.

biconcave geometry with the closest symmetrical geometry, in this case, 2D growth in a circular disk (Table 4.2).

The comparative increase in the cross sectional perimeter along with the decrease in the cross sectional area results in a decrease in the stress for the same pressure and wall thickness. This allows more deformation of the cell for the same critical stress which means this geometry is advantageous for a structure likely to undergo deformation.

4.4 Methods and materials

Super resolution fluorescence imaging and analysis

Imaging and analysis of A/Udorn/72 was as described in section 2.4.

E.coli imaging and analysis

Images of Escherichia coli (E. coli) bacteria were acquired as described previously [115]. Briefly, bacteria were grown to their exponential phase when OD600 was measured at 0.2, treated with 50 μg/ml rifampicin for 30 mins when required and imaged using brightfield microscopy. Brightfield microscopy images were illuminated using a white LED light source (CoolLED pE-100). Collected light was passed back through the multi-notch filter with transmission windows at 439±15 nm, 21±17 nm, and 605±25 nm, with an achromatic doublet lens (AC508-300-A, ThorLabs, Newton, New Jersey, USA) forming an image on an EMCCD camera (iXon 897 Ultra, Andor Technology Ltd, Belfast, UK). Image acquisition was performed using the software package Andor SOLIS (Andor Technology, Belfast, UK).

Bacterial image analysis was carried out on the bacterial segmentation software AKSeg [111]. Cells were segmented using a convolutional neural network trained on manually labelled data; dividing and incomplete cells were excluded from the segmentation, and then imperfect segmentations were cleaned up manually. From the segmentation mask, the lengths and widths of cells were extracted.

Cryo EM imaging and analysis

Scanning electron microscopy (SEM) preparation and imaging was carried out as described previously [23]. Briefly, Vero, U2-OS or MDCK cells were seeded on to cryo-EM grids as follows. Freshly glow-discharged finder gold quantifoil grids (200 mesh R2/2—Quantifoil MicroTools GmbH, Germany) were sterilised in 70% EtOH and placed in a glass-bottomed MatTek dish (MatTek Corporation, MA, USA). 200 μl of laminin (50 $\mu g ml^{-1}$) was added and incubated overnight. Grids were then washed in DMEM. 105 cells were added in 2 ml of DMEM supplemented with 10% FBS and incubated at 37°C in an atmosphere containing 5% CO2 overnight before infection. Cells were infected with human RSV (strain A2-RSV) or A/Udorn/72 at an MOI of 1 and incubated for a further 72h before being prepared for cryo-EM.

Grids were frozen for cryo-EM by plunge freezing. Briefly, 3 μl of a suspension of 5 nm colloidal gold beads was pipetted onto each grid (BBI Solutions, United

Kingdom). Grids were then transferred to a Vitrobot Mk IV (Thermo Fisher Scientific), blotted for four seconds and then immediately plunged into a bath of liquid ethane [116].

Tilt-series imaging was performed on a JEOL 2200FS energy-filtering transmission electron microscope equipped with a Gatan Ultrascan 4 k×4 k CCD camera and a Gatan 914 high-tilt cryo-stage. The microscope was operated at 200 kV and zero-loss energy filtered imaging with a slit-width of 30 eV was used to enhance image contrast. Tilt series were recorded using the SerialEM software package[117]. Images were acquired at two-degree increments from 70° to +70° between 10,000× and 20,000× magnification for cells on grids and at 20,000× or 40,000× magnification for purified virus. Images were recorded with two-times binning, corresponding to a pixel sizes ranging from 21.2 to 5.4 Å/pixel in the specimen. The target defocus was set to between 4 and 6μm under-focus and the electron dose ranged from 83 e/Å^2 to 100 e/Å^2 per tilt-series. Tomograms were calculated and visualized using the IMOD software package [118]. Reconstruction was performed using weighted back projection followed by denoising using non-linear anisotropic diffusion. Figures were prepared by averaging 10 tomogram sections using IMOD's 3dmod slicer routine. Segmentation was performed manually using Amira (Visual Sciences Group).

The contrast of each tomogram taken was adjusted to make the boundaries of the filaments clear. The outer and inner diameters of each filament in the acquired tomograms were measured in 5 different locations with the line tool of ImageJ. These were used to calculate the average width of each filament and the error in these measurements. From the diameter information, the wall thickness was calculated with the error in wall thickness calculated through error propagation.

4.5 Discussion and conclusion

In this work, we have presented a model of filamentous morphologies of viruses to explain their formation. A pressure vessel model was formed that related the stress, internal pressure and the length scales of a virus particle. It showed that the stress increased with the radius of a spherical virion and that a filament was

a method of avoiding the size constraint that this imposed upon virion size due to the bursting stress of the membrane. Bursting stress also explains the absence of irregular geometries and other shapes of viral growth.

The bending stress in the walls has been ignored using the thin wall approximation. This is a very good approximation for RSV as diameter/wall thickness > 20. We are at the extrema of the model for filamentous Udorn as $20 >$ diameter/thickness > 10. However, it allows a relation without knowing $P_{internal}$ or $P_{external}$ though it slightly underestimates the (maximum) spherical radius for a given filamentous radius. As enveloped viral particles bud from cells, the internal pressure is equalized between cell and virus leading to a greater internal pressure than external which allows the approximation that the pressure is uniform across virions of the same species.

The 2D histogram of A/Udorn/72 major and minor axis sizes (Fig. 4.1E), showed virions tend to a decreased minor axis at large major axis length scales. The yellow theoretical line of Fig. 4.1E passed through the population of the largest spherical virions and then tended to the minor axis size at large length scales and showed that beyond a certain size, virions need to become filamentous to increase their size.

Other studies [23][119][84] have also found an average diameter of approximately 80 nm for filamentous influenza with the discrepancy to our study arising from the measurements including the surface proteins of the virus. Our model's prediction that filamentous virions should have half the diameter of spherical virions appears to fit experimental data for Udorn with a filamentous diameter of 56 nm and a spherical diameter of 140 nm, and RSV with a filamentous diameter of 95 nm and spherical diameter of 250 nm; especially when considerations for the measurement of surface proteins for spherical virions and the neglection of bending stress are considered. This gives a strong indication that our model is capturing real behaviour of viral particles.

The model is applicable to many cells and biological particles as shown by our data on E.coli and suggests stress to be an important factor in the geometries of many particles including viruses and bacteria. The above analysis of filaments applies for bacteria and explains why they can grow to be tens of microns without

stresses due to their internal pressure becoming too large- with the diameter of the elongated filamentous E.coli being reduced when compared to the unstressed bacilliform E.coli. The bacteria do not form spherical particles which may derive from their reproduction through cell division. The analysis is also applicable to other biological structures, in particular, they may be most relevant for structures prone to external stresses.

Li et al. show the lipid bilayer to be an easily deformable, non-rigid container [120]. This means that geometries with large, flat surfaces would "balloon". When the 2D growth scenario (Fig. 4.4) "balloons", a more spherical geometry is formed and would exceed the maximum stress and, so, viruses do not grow in disk shapes along 2 principle directions.

As shown in Fig. 4.2, the red blood cell geometry allows a greater stress and so, a greater deformation than a disk-shaped geometry. This is an important feature of its shape as, to provide oxygen to the whole body, they must travel down capillaries and are often deformed in this process. [121] This geometry minimises the risk of rupture when they are carrying out their role in circulating oxygen around the body.

Our model has a number of limitations including: the pressure throughout filaments and particles may not be fully uniform; the structure of the M1 protein in filaments is helical [45] which may change the maximum stress of filaments (assumed to be a constant); the walls have a finite thickness and so the assumption that the radius of the lipid layer is equal to the radius of the inside of the virus is not perfect; bending stress is not included in the model; the surface of virions, with surface proteins and matrix proteins, is not perfectly uniform which leads to a slight deviation in the maximum stress between virions and, finally, the model does nothing to explain the mechanism for filamentous formation or why some strains of viruses form filaments, but others do not. In spite of this, this model suggests filamentous formation to be the method through which enveloped viruses can become larger than the limit placed upon them from the stress of a lipid bilayer. As larger viruses require more material from the cell and decreases the number of virions produced; this suggests the mechanism is to either increase their volume

to contain multiple copies of genomic segments [105][84] or to increase the surface area, making them less susceptible to the immune system [81].

5 Conclusions

5.1 Conclusions

The use of fluorescence microscopy has revolutionised the study of single molecules and the micro-scale biological world, with its ability to image with a high temporal resolution or image specific structures at a nanometer resolution. Notable examples include: the arrangement nuclear pore complex proteins through the use of dSTORM [122], and 3D STORM imaging was used in the discovery of the actin ring building blocks that make up actin within a neuronal axon [123].

Mathematical and computational models simplify individual mechanisms and phenomena in the biological world to a state where they can be explained or make predictions about the world around us. Turing patterns [124] explain the patterning of a tiger and the differentiation of cells early in the development of a embryo, while computational modelling is at the forefront of modern drugs research and the identification of novel drug targets [125][126].

In this thesis, I presented a pipeline that enables the imaging and analysis of thousands of individual virus particles in order to study viral populations; investigate the average structure of a virus particle, and understand the protein organisation of viral filaments. Using widefield microscopy and dSTORM, the size distribution of virus particles was shown. Udorn, a pleomorphic virus, was split into two

populations by size, suggesting a stochastic mechanism for filament formation, while SARS-CoV-2 had a single population. By computing the Fourier transform of the intensity profile of HA and NA along filaments, I found that there was no discernible spatial patterning and, by summing the intensity profiles of filaments with the RNA labelled with FISH probes and budding filaments labelled with an NA antibody, I found RNPs are primarily located at the ends of filaments, located in Archetti bodies when present, and NA is enriched at the distal end of a filament.

By infecting and fixing cells at various time points, we saw that filamentous budding could be seen by the 24 hour time point. At this time point, we also saw filaments bridging between cells, with structures along the length of the bridges. These bridges and structures have not been explained. I also demonstrated the labelling of virus particles through the infection of labelled cells. The labelling was achieved using multiple fluorescent dyes, and depending on the dye used (and the labelled structure) the intensity of fluorescent virus particles differed.

In order to explain the existence of filamentous morphologies of viruses, I modelled viruses as pressure vessels to derive a relation between their major and minor axis given a constant critical stress. This found relation was tested against the major-minor axis distribution for Udorn and E.coli which fit our predictions with good agreement. This model also predicted that filaments should have half the diameter of spherical particles which fit measurements from cryo EM data collected for both Udorn and RSV which further confirmed the model. The model also explained the absence of non-uniform virus geometries, the shape of red blood cells and suggests filament formation as a method of increasing virus volume without exceeding the membrane critical stress.

5.2 Future work

The protocols developed in this thesis to image virus particles have since been used to image virus particles by other lab groups (e.g University of York) and the computational tools and methods are generally applicable, and would be of interest for other filamentous viruses and the study of clinical strains of influenza.

A functional assay on filtered virus particles is required to elucidate the relationship between pleomorphism, pathogenesis, and infection to show how the studied viral particles are important in the infection cycle of virions, and the fixed cell studies should be repeated with human epithelial cells to make them more relevant to human infection. The filamentous particles bridging between cells and the structures seen along their lengths have not been explained and their further study to understand their role in the life-cycle of influenza is required.

The novel labelling strategy is already being used by the Robb lab and collaborators. It is being used to create specifically labelled virus particles, in particular to study the infection of cells, and, by changing the labelled structure in the cell, lipid rafts and the assembly of virus particles is being studied. It will be important to investigate the infectivity of the virus particles produced to ensure the dye is not hampering viral function, and investigate the specificity of the dyes to the membranes of choice.

The pressure vessel model and the SIR model for virus reinfection are generally applicable to other biological structures and diseases respectively. Their application and modification may lead to future developments in the understanding of bacteria, viruses, funghi, vesicles etc. and the reinfection dynamics of future epidemics/pandemics.

Appendices

A
Publications

McMahon A, Robb NC, *Reinfection with SARS-CoV-2: Discrete SIR (Susceptible, Infected, Recovered) Modeling Using Empirical Infection Data.* JMIR Public Health Surveill 2020; 6(4):e21168. DOI: 10.2196/21168

Shiaelis N, Tometzki A, Peto L, **McMahon A**, Hepp C, Bickerton E, Favard C, Muriaux D, Andersson M, Oakley S, Vaughan A, Matthews PC, Stoesser N, Crook D, Kapanidis AN, Robb NC. *Virus detection and identification in minutes using single-particle imaging and deep learning.* ACS Nano 2023 17 (1), 697-710. DOI: 10.1021/acsnano.2c10159

McMahon A, Andrews R, Ghani SV, Cordes T, Kapanidis AN, Robb NC. *High-throughput super-resolution analysis of influenza virus pleomorphism reveals insights into viral spatial organization.* bioRxiv; 2021. DOI: 10.1101/2021.09.23.461536.

McMahon A, Vijayakrishnan S, el Sayyed H, Robb NC. *Engineering stress as a motivation for filamentous virus morphology* In preparation.

B

Numerical modelling of SARS-CoV-2 reinfection

Contents

B.1 Introduction 99
B.2 Methods 101
Choice of geographical regions and data sources 101
Assumptions 101
Framework 102
Assigning probabilities 105
B.3 Results 108
United Kingdom data 108
Simulations of infection data in other regions 110
Comparison of infection and hospitalisation data 110
Comparison of simulated and real recovery data 111
Changing the recovery lag time 111
B.4 Discussion and conclusion 112
The results in relation to what was known about Covid-19 at the time of the study (17th May 2020) 113
The results in relation to what was known about Covid-19 at the time of writing (13th January 2023) 116

B.1 Introduction

The outbreak of the SARS-CoV-2 coronavirus, causing the disease COVID-19, is thought to have originated in late 2019 and spread to cause the a global pandemic.

Up to 12th January 2023, it has caused a confirmed 6.71 million deaths from 665.52 million confirmed cases[127]. This work was completed with data up to 17th May 2020. By that time, it had caused over 820,000 deaths from over 24 million infections and had caused major social and economic disruption.

During the COVID-19 pandemic, variations on the susceptible, infected, recovered (SIR) model have been used to investigate the varying burden on healthcare systems based on public health intervention [128], the absence of a stable disease-free equilibrium [129] and infection rate [130], and the eventual size of the pandemic [131]. An Susceptible-Exposed-Infectious-Removed-Undetectable-Susceptible (SEIRUS) extension to the model has also found that an equilibrium point will be reached where there are no further reinfections [132].

At the time this study was completed, little was known about the epidemiology of SARS-CoV-2, including the reinfection potential of the virus or biologically what happened to a person's immunity upon infection with SARS-CoV-2. The aim of this study was to build upon the simple SIR model to extend the framework to distinguish between infected (for the first time) and reinfected individuals to model empirical data, in order to investigate the effectiveness of primary SARS-CoV-2 infection at preventing reinfection in the short term.

In Section B.2, I define the 6 state SIR model that was used, give the assumptions required to use the model, explain the choice of geographical regions chosen to base simulations on and show how the probabilities of infection, recovery and death are assigned in a simulation. In Section B.3, I show an example of a simulation, give the number of reinfections for the geographical regions chosen and investigate how reinfection data would look if only hospitalisations were modelled, how the use of modelled recovery data vs empirical recovery data changes the number of reinfections, and determine how the recovery time changes the number of reinfections,

The work presented in this chapter has been published as Ref. [133].

B.2 Methods

Choice of geographical regions and data sources

The simulations were initially completed for the United Kingdom (UK) where, at the time the data was accessed, there were a high number of confirmed cases. Australia was selected as an example of a region with low numbers of recorded cases in order to investigate the limit of expected reinfections; Germany was selected as it was one of few countries with recorded recovery data; Italy was studied as the number of infections and deaths had peaked by the 17th May 2020; Singapore was unique as a city-state so population density for the nation was very high; Switzerland was selected as hospitalisation data was available at the time the data was accessed, and the United States of America (USA) as a whole was compared with New York City, which was the worst affected state in the USA at the time.

We used national infection and mortality data from a variety of sources to investigate the reinfection dynamics of SARS-CoV-2. Unless specified, the national data on infections and deaths from SARS-CoV-2 was acquired from the 'Our World in Data' database compiled by the Oxford Martin School at the University of Oxford [72]; the hospitalisation cases in Switzerland were obtained from the Federal Office of Public Health in Switzerland [134]; the data for the city of New York was obtained from the New York City Health website [135]; the population of New York City was obtained from the 2019 New York census [136] and the recovery data for Germany was sourced from Trading Economics [137] which gets data from WHO (World Health Organisation). For each geographical region, the data was taken from the date of the first recorded infection worldwide up until the 17th May 2020 when the data was accessed.

Assumptions

A number of assumptions were made. Where possible, they have been made so that the number of reinfections is underestimated. These assumptions are:

1. There is a large lag time for recovery to take place (28 days) [138][139].

2. The incubation period was modelled as 6 days [140].

3. The model does not consider social distancing or shielding and so assigns an equal probability of an infection to all individuals.

4. Not all infections have been recorded due to lack of testing, misdiagnosis or asymptomatic infection. [141]

5. Infections and recoveries are not necessarily recorded on the date that they first occurred.

6. There is no emigration out of, or immigration into, a population of interest.

7. The model assumes a homogeneous population density, with no societal structure (e.g. equal residents per household).

Framework

We based our model on the compartmental SIR framework, but differentiated between initial and subsequent infections, resulting in a six-state model (Susceptible, Infected, Recovered, Infected (two or more times), Recovered (two or more times) and Deceased) (Figure B.1 & Table B.1). The number of infections and deaths each day were taken from national statistics (as described above). Where available, recovery data was used, and otherwise recoveries were modelled with a 28 day lag time (with the number of recoveries representing those individuals that did not die during the 28 day recovery time). 'Recovered' individuals were selected stochastically from the populations of the states 28 days prior [138][139]; 'infected' individuals from the populations of the states 6 days prior (due to the incubation period [140]), and 'deceased' individuals from the populations of the states 1 day prior.

When using the model, the rate of infection, recovery and fatality for each state were assumed to be independent of how many previous infections a host had previously had. The number of susceptible persons at the beginning of the simulation, N, was taken to be the population of the region of interest [72][135]. After all infections, recoveries and deaths for a day, the number of days into the

x_i	Random number
t	The number of days into the simulation. t=1 for the day of the first infection
$t_{recovery}$	The average number of days for recovery. Here, we have used $t_{recovery}$=28
$t_{incubation}$	The average number of days for an infection to be seen. Here, we have used $t_{incubation}$ = 6
t_{max}	The number of days over which the simulation is run
S_t	The number of susceptible on day t
I_t	The number of Infected (once) on day t
R_t	The number of Recovered (once) on day t
I'_t	The number of Infected (two or more times) on day t
R'_t	The number of Recovered (two or more times) on day t
D_t	The number of deceased from SARS-CoV-2 on day t
B_t	The infection rate on day t
γ	The recovery rate on day t
m_t	The death rate on day t
N	The total population in the model $N = S_t + I_t + R_t + I'_t + R'_t + Dt \ \forall\ t$
$N_t^{infected}$	The number of infected individuals on day t $N_t^{infected} = I_t + I'_t$
$N_t^{uninfected}$	The number of uninfected individuals on day t $N_t^{uninfected} = S_t + R_t + R'_t$
$n_t^{infected}$	The number of new infections on day t $n_t^{infected} = \beta_t S_{t-t_{incubation}}(I_{t-t_{incubation}}+I'_{t-t_{incubation}}) + \beta_t R_{t-t_{incubation}}(I_{t-t_{incubation}}+I'_{t-t_{incubation}}) + \beta_t R'_{t-t_{incubation}}(I_{t-t_{incubation}}+I'_{t-t_{incubation}})$
$n_t^{recovered}$	The number of recovering individuals on day t $n_t^{recovered} = \gamma_t I_{t-t_{recovery}} + \gamma_t I'_{t-t_{recovery}}$
n_t^{deaths}	The number of deaths on day t $n_t^{deaths} = m_t I_{t-1} + m_t I'_{t-1}$
$i_t^{first\ time}$	The number of first time infections on day t. i.e. the number of $S_t \rightarrow I_t$ transitions on day t

Table B.1: Definition of the model parameters.

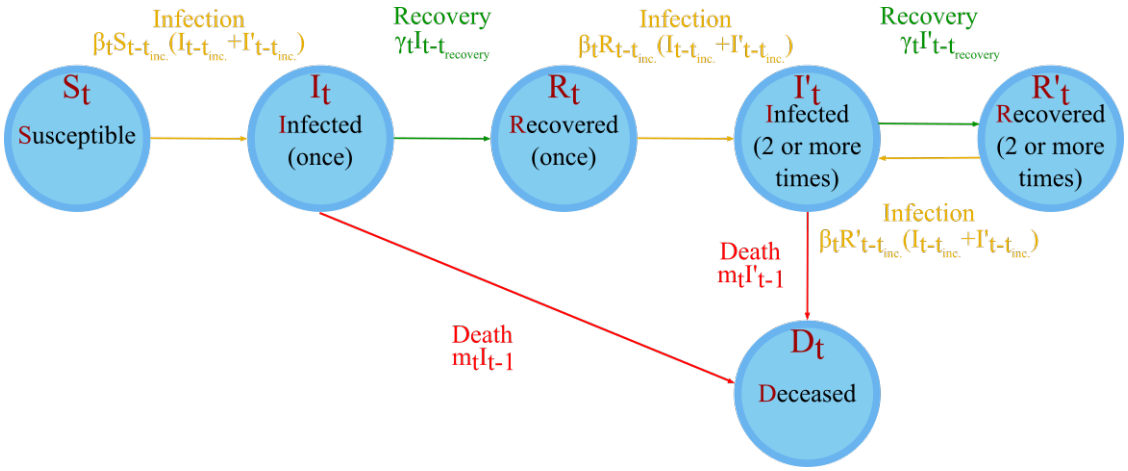


Figure B.1: A simple representation of the model. S_t represents the number of persons susceptible to infection who have had no prior infections on day t; It is the number of people currently infected for the first time on day t; R_t is the number of people who have recovered once on day t; I'_t is the number of people who have been infected two or more times and are infected on day t; R'_t is the number of people who have recovered two or more times and are not infected on day t of the model and D_t is the number of deceased persons on day t of the model. Further symbols are defined in Table B.1.

simulation was increased by one, t→t+1 up to tmax. The simulation was repeated 10,000 times to produce expectation values and standard deviations for the number of individuals classified as reinfections.

By pooling the number of cases in the Infected (two or more times), Recovered (two or more times) and those deceased from the Infected (two or more times) states at the end of the simulation, we calculated an estimate of the number of reinfections that would be expected to occur. This number represents the total population that had passed through the infected (two or more times) state by the end of the simulation.

Unless otherwise stated, the average recovery time used in the simulations was set as 28 days, as this is greater than the median recovery time suggested in the report of the WHO-China joint mission on coronavirus disease 2019 [138].

The model equations

The equations relating how each of the states (Susceptible, Infected (once), Recovered (once), Infected (two or more times), Recovered (two or more times) and Deceased respectively) changed for each iteration of the simulation were:

$$S_t = S_{t-1} - \beta_t S_{t-t_{incubation}}(I_{t-t_{incubation}} + I'_{t-t_{incubation}}) \tag{B.1}$$

$$I_t = I_{t-1} + \beta_t S_{t-t_{incubation}}(I_{t-t_{incubation}} + I'_{t-t_{incubation}}) - \gamma_t I_{t-t_{recovery}} - m_t I_{t-1} \tag{B.2}$$

$$R_t = R_{t-1} + \gamma_t I_{t-t_{recovery}} - \beta_t R_{t-t_{incubation}}(I_{t-t_{incubation}} + I'_{t-t_{incubation}}) \tag{B.3}$$

$$\begin{aligned} I'_t = I'_{t-1} + \beta_t R_{t-t_{incubation}}(I_{t-t_{incubation}} + I'_{t-t_{incubation}}) - \gamma_t I'_{t-t_{recovery}} \\ +\beta_t R'_{t-t_{incubation}}(I_{t-t_{incubation}} + I'_{t-t_{incubation}}) - m_t I'_{t-1} \end{aligned} \tag{B.4}$$

$$R'_t = R'_{t-1} + \gamma_t I'_{t-t_{recovery}} - \beta_t R'_{t-t_{incubation}}(I_{t-t_{incubation}} + I'_{t-t_{incubation}}) \tag{B.5}$$

$$D_t = D_{t-1} + m_t I_{t-1} + m_t I'_{t-1} \tag{B.6}$$

where the processes are shown in Fig. B.1 and the symbols are defined in Table B.1.

Assigning probabilities

Assignment of probabilities is given by the size of the state and the total number of infected individuals

The probability of an infection is determined by the ratio of the population of the state being infected to the total number of uninfected individuals, $N_t^{uninfected}$. The probability of any recovery or death is given by the ratio of the population recovering/dying to the total number of infected states, $N_t^{infected}$. For example, for the probability of an infection from the Susceptible state S_t to the infected (first time) state I_t is given by:

$$\begin{aligned} p \quad & (S_{t-1} \to I_t | Infection) \\ = \quad & \frac{\beta_t (I_{t-t_{incubation}} + I'_{t-t_{incubation}}) S_{t-t_{incubation}}}{\beta_t (I_{t-t_{incubation}} + I'_{t-t_{incubation}})(S_{t-t_{incubation}} + R_{t-t_{incubation}} + R'_{t-t_{incubation}})} \\ = \quad & \frac{S_{t-t_{incubation}}}{S_{t-t_{incubation}} + R_{t-t_{incubation}} + R'_{t-t_{incubation}}} \\ = \quad & \frac{S_{t-t_{incubation}}}{N^{uninfected}_{t-t_{incubation}}} \end{aligned} \tag{B.7}$$

Attributing infections/recoveries/deaths

We have used the uniform distribution with range 1 to R, where R is a real integer, to determine states involved in infections, recoveries and deaths. The probability for every number in the range is equal and the density function is given by

$$p(x, R) = \frac{1}{R-1}, 1 \leq x \leq R. \tag{B.8}$$

For sampling, we have used the cumulative distribution function:

$$P(x, R) = \int_1^x p(x, R)dx = \frac{x-1}{R-1}, 1 \leq x \leq R \tag{B.9}$$

Assigning infections

The patients that were infected each day were chosen at random from the 3 states that are not infected with the virus (weighted by the number of persons in each state), $n_t^{infected}$ random numbers, x, are taken from the distribution $p(x, N_{t-1}^{uninfected})$

The infections are assigned to states going through all the random numbers x_i. A person in the susceptible state is infected if:

$$P(x_i, N_{t-t_{incubation}}^{uninfected}) \leq P(S_{t_t_{incubation}}, N_{t-t_{incubation}}^{uninfected}) \tag{B.10}$$

A person in the recovered state is infected if:

$$P(S_{t-t_{incubation}}, N_{t-t_{incubation}}^{uninfected}) \leq P(x_i, N_{t-t_{incubation}}^{uninfected}) \leq P(S_{t_t_{incubation}} + R_{t-t_{incubation}}, N_{t-t_{incubation}}^{uninfected}) \tag{B.11}$$

A person in the recovered (2 or more times) state is infected if:

$$P(x_i, N_{t-t_{incubation}}^{uninfected}) > P(S_{t_t_{incubation}} + R_{t-t_{incubation}}, N_{t-t_{incubation}}^{uninfected}) \tag{B.12}$$

The number of infections each day, $n_t^{infected}$, is taken from the real number of infections that occurred in the region of interest on each day.

Assigning recoveries

The patients that recover each day were chosen at random from the two states that were infected with the virus (weighted by the number of persons infected in each state a number of days earlier, $t_{recovery}$). $n_t^{recovered}$ random numbers, x, are taken from the distribution $p(x, n_{t-t_{recovery}}^{infected})$.

The recoveries are assigned to states going through all the random numbers x_i. A person from the infected state recovers if:

$$P(x_i, n_{t-t_{recovery}}^{infected}) \leq P(i_{t-t_{recovery}}^{firsttime}, n_{t-t_{recovery}}^{infected}) \tag{B.13}$$

A person from the infected (two or more times) state recovers if:

$$P(x_i, n_{t-t_{recovery}}^{infected}) \ > \ P(i_{t-t_{recovery}}^{firsttime}, n_{t-t_{recovery}}^{infected}) \tag{B.14}$$

$n_{t-t_{recovery}}^{infected}$ is the number of people on day t-1 who are infected and $i_{t-t_{recovery}}^{firsttime}$ is the number of people infected for the first time on day $t - t_{recovery}$.

Where recovery data was available, the number of infections each day, $n_t^{recovered}$, is taken from the real number of recoveries that occurred in the region of interest on each day.

Where recovery data was not available, the number of infections each day, $n_t^{recovered}$, is given by the number of persons infected a number of days earlier, $n_{t-t_{recovery}}^{infected}$ who would not be expected to have died

$$n_t^{recovered} = (1 - \frac{\Sigma_t n_t^{deaths}}{\Sigma_t n_t^{infected}}) n_{t-t_{recovery}}^{infected} \tag{B.15}$$

Where n_t^{deaths} is the number of deaths on day t and $n_t^{infected}$ is the number of people infected on day t.

The cycle between Infected (two or more times) and Recovered (two or more times) is permitted as we are only interested in reinfection rates.

Assigning deaths

The patients that die each day were chosen at random from the two states that were infected with the virus (weighted by the number of persons in each state). n_t^{deaths} random numbers, x, are taken from the distribution $p(x, N_{t-1}^{infected})$.

The recoveries are assigned to states going through all the random numbers x_i. A person from the infected state dies if:

$$P(x_i, N_{t-1}^{infected}) \leq P(I_{t-1}, N_{t-1}^{infected}) \tag{B.16}$$

A person from the infected (two or more) state dies if:

$$P(x_i, N_{t-1}^{infected}) \ > \ P(I_{t-1}, N_{t-1}^{infected}) \tag{B.17}$$

where $N_{t-1}^{infected}$ is the total number of people on day t-1 who are in one of the infected states and the number of people dying from the Infected (two or more times) population are counted so the total number of individuals passing through the Infected (two or more times) state can be counted.

The number of deaths each day, n_t^{deaths}, is taken from the real number of deaths that occurred in the region of interest on each day.

B.3 Results

United Kingdom data

We initially ran the simulation for data in the United Kingdom over the course of 106 days (from the first recorded case on the 1st of February until the 17th of May 2020 when the data was accessed). Fig. B.2 shows how the population of each state in the model changed over the course of a typical simulation. The number of susceptible individuals initially remained steady, until day 55, when there was a sharp decline due to the increase in primary infections (Fig. B.2A). The number of individuals infected just once started to increase steadily after day 40, and continued to do so throughout the simulation until day 92. After the 28-day lag time, the individuals infected once started to recover, resulting in an increase

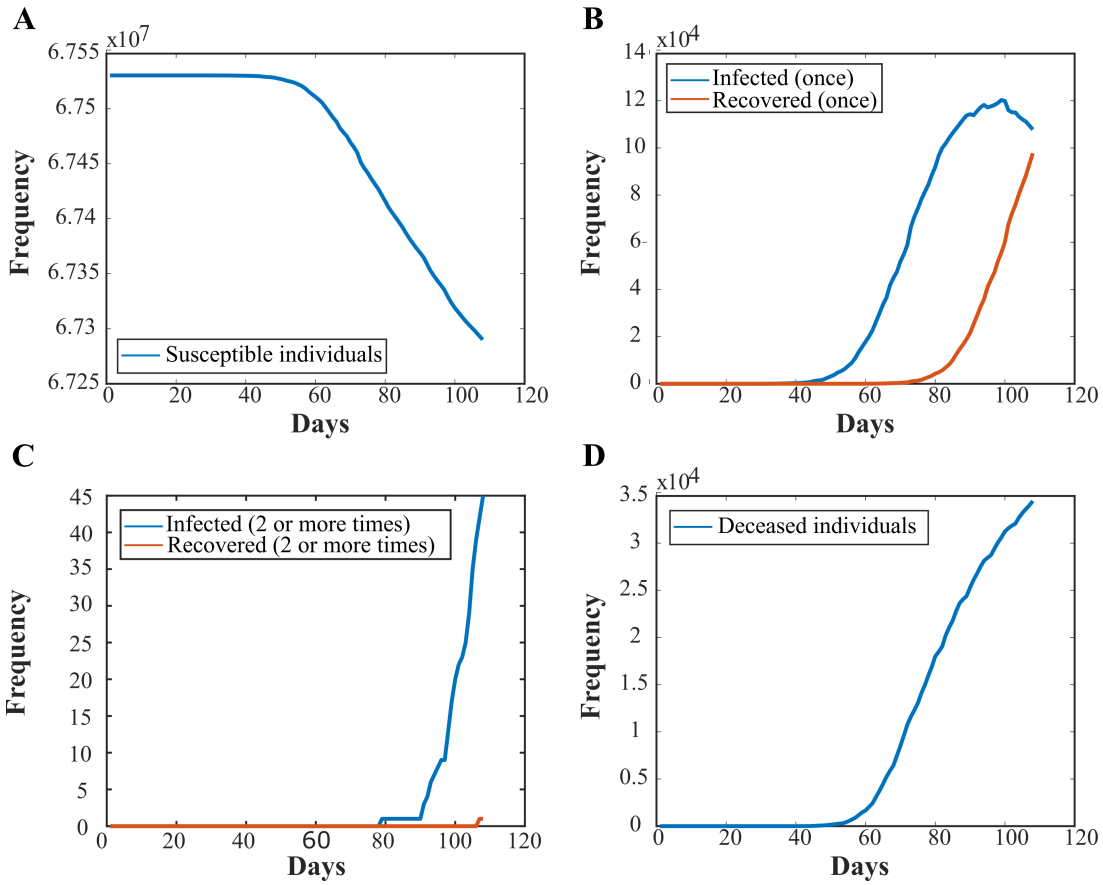


Figure B.2: Plots of the populations of each state in the model over the course of a typical simulation, using infection data from the United Kingdom. A) An example plot of the susceptible population in the model over the course of 106 days (from the first recorded case on the 1st of February until the 17th of May 2020 when the data was accessed). B) An example plot of the populations that are infected for the first time or recovered from a single infection. C) An example plot of the populations of the simulation that have been reinfected and recovered from infection twice. D) An example plot of the number of deceased individuals through the course of the simulation.

in the Recovered (once) state through to the end of the simulation (Fig. B.2B). As the number of recovered individuals started to increase, so did the number of people infected for a second time. The number of people recovered for the second time started to increase after the 28 day recovery lag time (Fig. B.2C). The number of deaths started to rise from day 55 onwards, and fatalities continued to increase through to the end of the simulation (Fig. B.2D). In the United Kingdom, the number of expected reinfections was calculated to be 43±7, which makes up 0.018% of the total infections (Table B.2). The first reinfection for the UK was on day 82±5, corresponding to the 22nd of April 2020.

Region	Mean number of reinfections	Standard deviation of number of reinfections	Total number of infections	Reinfections as a % of total infections
Australia	0.1	0.3	7036	0.0014
Germany	20	5	174355	0.011
Germany (with recovery data)	79	9	174355	0.05
Italy	63	8.0	224760	0.028
New York City	209	15	189027	0.11
New York City (hospitalisations)	7	3	48462	0.004
Singapore	4	2	27356	0.014
Switzerland	4	2	30587	0.013
United Kingdom	43	7	240161	0.018
United States of America	402	20	1467884	0.027

Table B.2: The number of predicted reinfections and their standard deviation in different locations worldwide as predicted from the model. Unless otherwise stated, these figures represent simulations using the total number of infections of each region and are modelled without the data on number of recoveries. S.D. standard deviation.

Simulations of infection data in other regions

The simulations were repeated with data from Australia, Italy, New York City, Singapore, Switzerland and the United States of America (USA). The mean number of expected reinfections in each region or country for the 10,000 simulations that were run are shown in Table B.2. In all cases, with the exception of Australia, our model predicts that reinfection cases should occur.

Comparison of infection and hospitalisation data

Next, we repeated our simulation for New York City, with the total number of infections replaced by the number of hospitalisations. When we ran the simulation with an input of the total number of infections, the number of secondary infections continued to increase throughout the simulation, when the numbers appear to start to peak (Fig. B.3A). This was followed by an increase in the number of secondary recoveries after the 28-day recovery lag time. In comparison, the hospitalisation

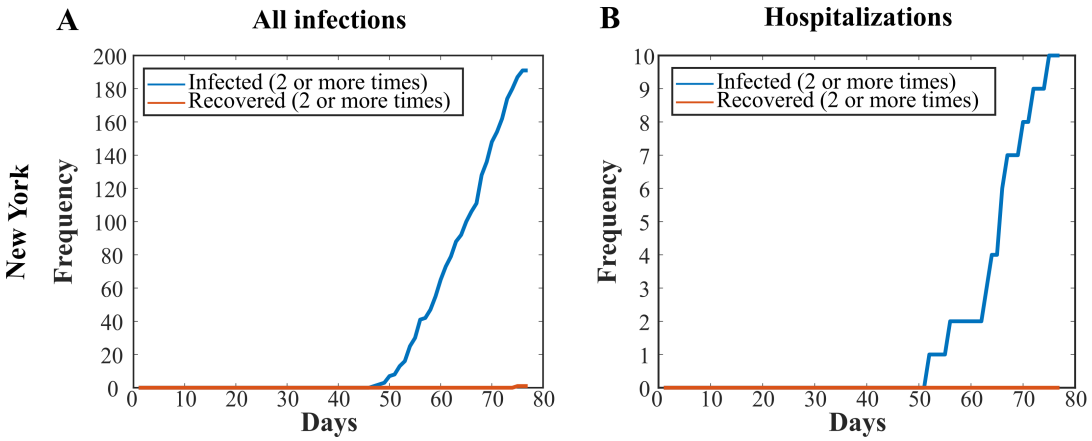


Figure B.3: Comparison of total infections vs. hospitalisation data in New York City. Plots of the Infected (two or more times) and Recovered (two or more times) states for A) New York using all infection data B) New York using only the hospitalisations data.

data for New York showed no secondary recoveries as the reinfections occurred later into the simulation (Fig. B.3B). The total number of predicted reinfections from the New York hospitalised data was 12±4 (Table B.2).

Comparison of simulated and real recovery data

Recovery data was sparse or unavailable for most regions, likely due to lack of follow-up testing. There was recovery available from Germany and we therefore compared the results of our simulation for Germany with and without the recovery data as an input. The models used a 28-day lag before recoveries started, meaning very few secondary recoveries took place (Fig. B.4A and Fig. B.4B). There were 73 more reinfections with the reinfection data than with the modelled data (Table B.2).

Changing the recovery lag time

The 28-day lag time used for the modelled recovery data ensured that we underestimated the recovery rate, and so also the rate of reinfection. To investigate a more life-like recovery rate, the United Kingdom simulations were repeated again using the modelled recovery data, while shortening the lag time for recovery. As expected, we found that the rate of reinfection increased as the lag time was decreased from 28 days through to 7 days, as there was a larger population that recovered from

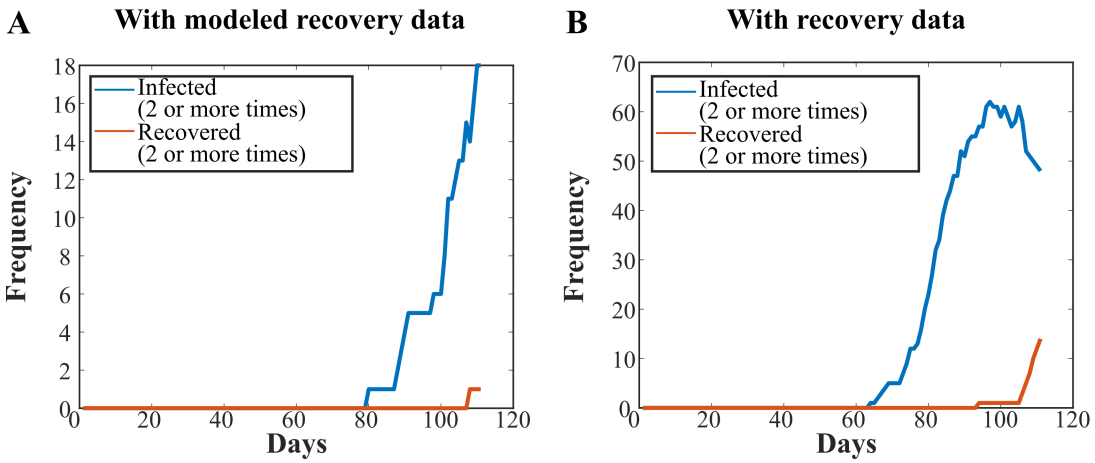


Figure B.4: Use of actual recovery data from Germany suggests that the number of recovered individuals, and hence reinfections, are underestimated in our model. Plots of Infected (two or more times) and Recovered (two or more times) states with A) modelled recovery data and B) actual recovery data.

a primary infection. With a 7-day lag time, the number of people in the infected (2 or more times) state peaked at day 101 of the simulation (Fig. B.5). The total number of people reinfected throughout the simulation increased with a decreasing lag time, with 43±7, 83±9, 139±12 and 209±14 reinfections for 28-day, 21-day, 14-day and 7-day recovery lag times respectively.

B.4 Discussion and conclusion

In this work we have presented a modelling strategy used to determine whether SARS-CoV-2 reinfections can occur. We modelled actual infection and fatality data from different regions around the world and found that all regions investigated, with the exception of Australia, should have recorded cases of reinfections if primary infection with SARS-CoV-2 did not provide some level of immunity.

Our model has a number of limitations, including the lack of modelling of any social structure, the fact that individuals who have been infected may change their shielding behaviours, differing recovery times from person to person, and missing information regarding immigration into and out of regions of interest.

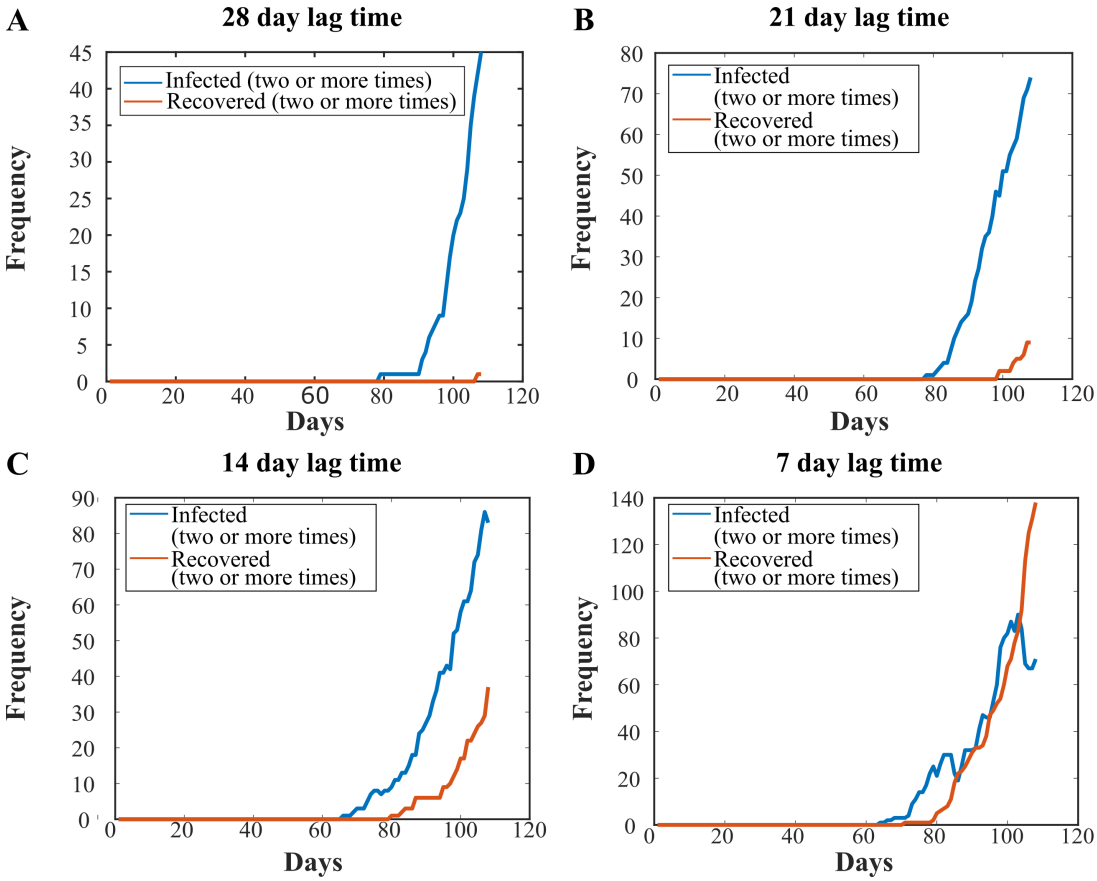


Figure B.5: Comparison of the reinfection dynamics when changing the recovery lag time. Plots of Infected (two or more times) and Recovered (two or more times) states with A) 28 day lag time, B) 21 day lag time, C) 14 day lag time, and D) 7 day lag time shows there to be reinfections expected even with an overestimate of the recovery lag time.

The results in relation to what was known about Covid-19 at the time of the study (17th May 2020)

In Australia, the number of confirmed SARS-CoV-2 infections at the time the data was accessed was relatively low [72], possibly due to early social distancing measures, the closing of international borders and mass testing and tracing measures. The number of modelled reinfections (0.1±0.3; Table B.2) reflects this, and so even without immunity from infection no reinfections would be expected to occur. Similarly, in Switzerland and Singapore, very low numbers of reinfections were predicted by the model (6.2±2.5 and 6±2 respectively; Table B.2). It is possible that these very low numbers of reinfection cases could have been missed due to misdiagnosis or lack of follow up testing. We therefore applied our model to data

from Germany [72], Italy [72], New York City [135][136] and the USA as a whole [72], which have recorded far higher numbers of SARS-CoV-2 infections (174355, 224760, 189027 and 1467884 respectively, when the data was accessed). The number of reinfection cases predicted for these countries was 30±6, 89±9, 335±18 and 635±25 for Germany, Italy, New York and the USA respectively (Table B.2). We conclude that it is therefore very unlikely that all of these predicted cases, if true, were missed due to misdiagnosis or lack of testing.

We also found that rehospitalisation cases should have been seen amongst hospitalised cases in New York City – it is unlikely that these cases would be missed as people are processed and tested on admission into hospital. To date, however, no reinfections have verifiably been recorded anywhere in the world. A report from South Korea suggested that 116 patients recovered from COVID-19 had tested positive by RT-PCR for the virus again, however, this has since been explained as the 'false-positive' detection of remnants of viral RNA rather than reactivation or reinfection [142]. The lack of documented reinfections suggests that short-term immunity to the virus is produced by an initial infection, however, our model cannot predict whether this immunity will last over longer time scales.

At the time of the study, our results were supported by a number of animal challenge studies which also showed immunity to SARS-CoV-2 can be conferred. A study in rhesus macaques showed that, following initial viral clearance, the monkeys showed a reduction in their median viral load in comparison with primary infection when rechallenged with SARS-CoV-2 [143]. Similarly, Ryan et al. demonstrated that re-challenged ferrets were fully protected from acute lung pathology [144]. An adenovirus-vector vaccine tested on rhesus macaques elicited a humoral and cellular response that, on challenge with the virus, proved to significantly reduce the viral load in bronchoalveolar lavage fluid and respiratory tract tissue [145]. However, a longitudinal study by Seow et al. [146] showed that the immunity conferred against SARS-CoV-2 may only be short-term. Our model proposes that reinfection cases should have already started to appear by April 2020, suggesting a possible lower limit for immunity duration.

A report from the WHO-China joint mission on Coronavirus disease estimated the recovery time for SARS-CoV-2 infection as 2 weeks for mild cases and 3-6 weeks for severe or critical disease [138]; based on this we used a long (28 day) recovery lag time in the modelled data. Comparison with real-world recovery data from Germany suggested that the actual recovery time may be significantly shorter, giving rise to an underestimation of the reinfection rate in our modelled data. This was supported by an increase in the number of predicted reinfections in the United Kingdom simulations when we used a shorter recovery lag time of 7, 14 or 21 days. In addition, there were no allowances in our model for transmission being localised to regions smaller than a nation or city; the daily infection data was likely to be only a fraction of the total number of infections due to asymptomatic or mild infections not being recorded, and infections were recorded on the date of testing, not the actual date of infection. We also note that significant differences in testing, reporting and shielding of the vulnerable exist between the different regions in this study and that a large number of SARS-CoV-2 cases were missed in every region of interest (for example, in Geneva, unreported cases were estimated to be 11.6 infections per reported infection from April 6 to May 9 [141]). In every region, we expect that the impact on our simulation would be to underestimate the number of reinfections. Taken together, this suggests that the actual reinfection rate would be significantly higher than that predicted by our model if there was no immunity conferred by prior infection.

The actual number of cases of reinfection that had been reported in any of the studied regions or countries to the study date was zero, and so the results documented here provide strong evidence, based on real data, to suggest that that there is at least short-term immunity conferred by an initial infection of SARS-CoV-2. This has implications for serological testing strategies, lockdown easing timescales and vaccine development. Our modelling strategy can also be extended to understand the reinfection dynamics of future pandemics.

The results in relation to what was known about Covid-19 at the time of writing (13th January 2023)

Up to 3 September 2022, the estimated reinfection rate by the UK was 44.4 per 100 000 participant days at risk with the probablility of reinfection decreasing for those who had a symptomatic first infection [147]. The risk of a subsequent infection within 6 months of a previous infection was classed as low by the CDC [148] which led to many countries adopting Covid strategies that included recent PCR tests as evidence for being Covid negative. This strongly supports the findings of this study that short term immunity is conferred upon infection with SARS-CoV-2.

Vaccination makes it difficult to look at long-term immunity and compare infection rates as over 50 million people in the UK alone have now had 2 or more vaccines which decreases the chance of Covid infection [149]. Offered vaccines include adenoviral vector vaccines (e.g. ChAdOx1) and mRNA-based vaccines. The equal footing of vaccination and a recent (cleared) infection with Covid 19 strongly supports the conclusion of this study that there is strong short term immunity after an infection with SARS-CoV-2.

Reinfection has massively increased since the mutation of the virus and the circulation of multiple strains that are capable of antigenic escape. The greatest mutations responsible for antigenic escape have come in Omicron variants with 91.6% of confirmed reinfections coming during the period that Omicron was the dominant strain [147]. Of all reinfections from the five strains of note so far (alpha, delta, Omicron BA.1, Omicron BA.2 and Omicron BA.4/5), 1% have been alpha reinfections from an initial alpha infection; 1% have been delta reinfections from an initial delta infection and 1% have been an Omicron BA.4/5 reinfections from an initial Omicron BA.4/5 infection [150]. The other 97% have been reinfections with a mutated strain of the virus compared to the initial infection. This adds the caveat to the findings of the study that the immunity is less powerful against different strains of the virus, though, as only the alpha variant was in circulation at the time of the study, the conclusion at the time of the study was correct.

C

Computational simulations showing an improvement to dense fitting using intensity differences

Contents

C.1 Introduction 117
C.2 1D simulations 118
C.3 2D simulations 118
C.4 Conclusion 120

C.1 Introduction

Although the accuracy of super-resolution dSTORM is a huge improvement upon traditional fluorescence imaging, it is still limited to 10-20nm [57]. In addition, localisation-based super-resolution methods, which require low labelling density of the sample, tend to have slow imaging rates (typically <0.01 images/s), limiting their utility in live-cell studies [151]. Increased labelling density can result in increased imaging rates, however these experiments are limited by the fitting of multiple overlapping PSFs (and require a dense fitting algorithm like the HAWK algorithm [152] or DAOSTORM [151]). The aim of the performed simulations was

to investigate the improvement in fitting the localisations from densely labelled structures when labelled with fluorophores of differing intensities.

C.2 1D simulations

The first experiment undertaken was to simulate PSFs in one dimension with a Gaussian-like intensity profile. The error in the localisation of Gaussian fitting against the separation of simulated peaks is shown in Figure C.1. The experiment was undertaken multiple times, changing the relative intensity of the moved peak in relation to the peak at the origin each time. As can be seen from by comparing Figure C.1A with C.1B, C.1C and C.1D, the fitting of the peaks appears to be more stable when the peaks are of differing intensities up to a certain amplitude ratio. The 20% amplitude simulation (Figure C.1E), is fit worse. Ignoring the points that are fit poorly due to the instability in the fitting procedure, the fitting is improved for peaks of differing amplitudes. An error of consistently less than 0.01 in the fit of both peaks is found for separations down to 0.63 for the 100% secondary peak compared to 0.50 for the 80% secondary peak, 0.43 for the 60% secondary peak and 0.56 for the 40% secondary peak. Also, there are fewer points where the error changes by multiple orders of magnitude for the fitting of mixed-amplitude peaks (compared to the oscillatory behaviour of the fitting of equally-sized peaks) suggesting the fitting is more stable under these conditions. The error was only less than 0.01 for three separations for the 20% secondary peak - probably due to the small peak size hindering accurate fitting, meaning this relative peak amplitude worsens the fitting.

C.3 2D simulations

The fitting of two 2D Gaussians was completed (see Figure C.2) to investigate whether the affects seen in the 1D simulations were also valid for 2D Gaussians. Similar separation values and peak heights were to be investigated in multiple directions from the origin before the project was discontinued.

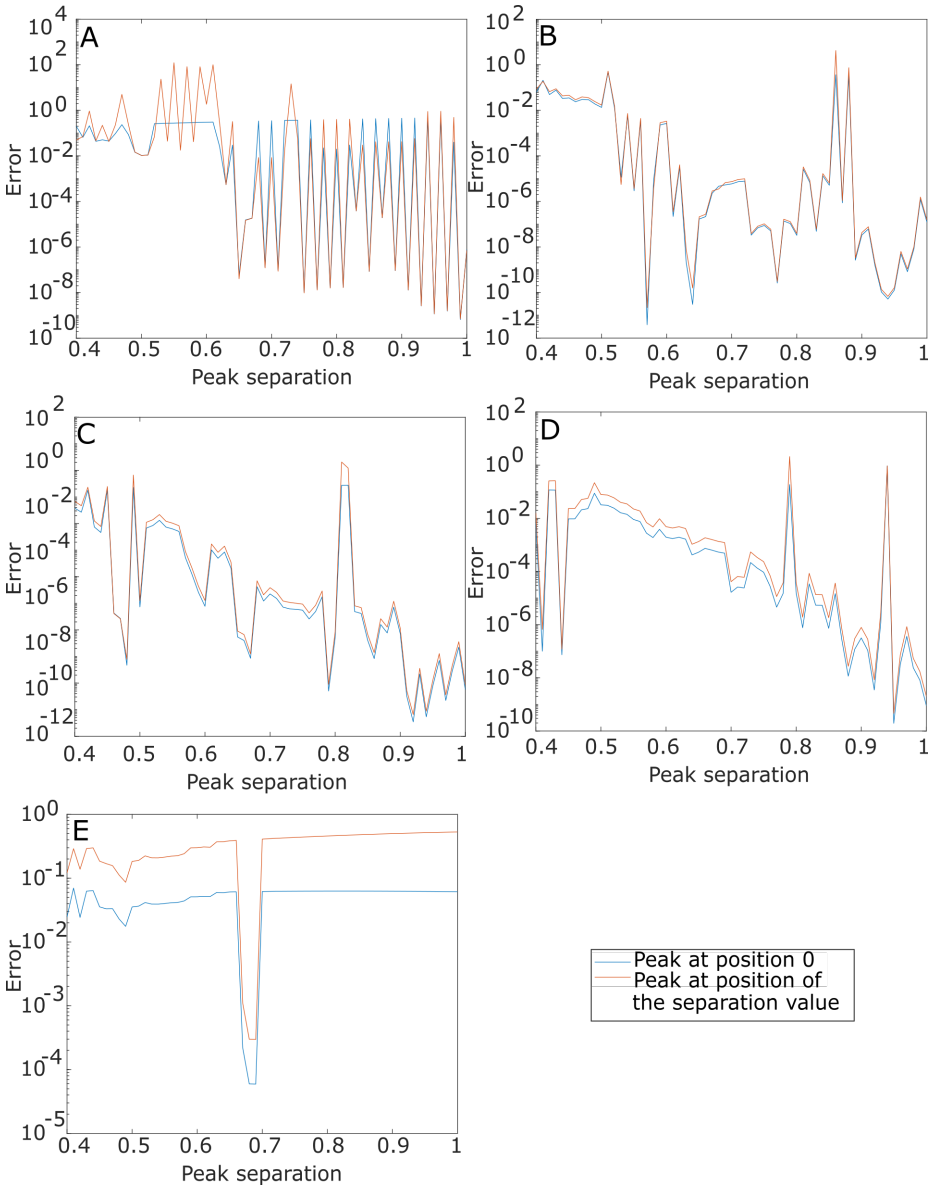


Figure C.1: Log plots of the error in the mean value of two fitted Gaussians when the separation between the Gaussians is changed. The error is the the difference between the fit and actual values for the mean of the Gaussians. The amplitude of the separated peak is: (A) 100% of the centre peak, (B) 80% of the centre peak, (C) 60% of the centre peak (D) 40% of the centre peak (E) 20% of the centre peak. Each line on a plot shows variation in the error of one of the Gaussians.

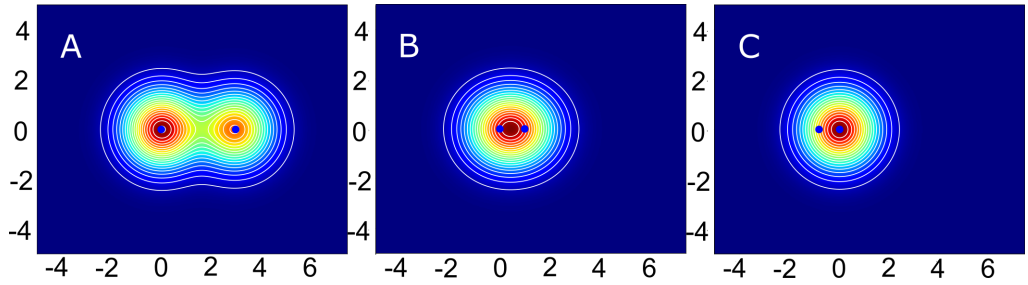


Figure C.2: The fitting of peaks in 2D. Colours represent the intensities of the added Gaussians, isobars represent the fit of the Gaussians and the blue dots represent the centres of the two Gaussian peaks. The peak at x≠0 is 80% of the amplitude of the peak at x=0. A) The second peak was placed at x=3, the peaks are easily distinguishable by eye and are accurately fit. B) The second peak was placed at x=1, the peaks are not easily distinguishable by eye but are accurately fit. C) The second peak was placed at x=0.1, the peaks are not easily distinguishable and are not accurately fit.

C.4 Conclusion

In conclusion, this experiment suggests that the minimum distance that can be resolved between two peaks is increased when the peaks are of different amplitudes and the fit is more stable with respect to the separation distance. This may be utilised to improve the speed at which dSTORM acquisitions can be completed by using a dense-emitter fitter in conjunction with multiply labelling structures with different fluorescent labels of differing intensity in order that PSFs can be closer together and more localisations found in a unit time. Different relative intensities of fluorophores can also be achieved by using a second fluorophore of a lower intensity, or by using a fluorophore designed to be used at a different wavelength to its optimal wavelength.

However, due to the requirement for multiple fluorophores of differing intensities and the limit in the improvements achievable, this project was halted at the point of simulating 2D gaussians. This is because these improvements would be possible with multiple fluorophores of different wavelengths which wouldn't need to be separated at all using multiple frequency filters and optical paths- the only advantage would be the ability to perform this improvement in conjunction with multiple colours and the fact it can be performed without altering any optical elements.

D

The collection of aerosolised virus particles

Contents

D.1 Introduction 121
D.2 Results 122
D.3 Conclusions 123

D.1 Introduction

By January 2022, 400,000,000 viral tests had been performed due to the SARS-CoV-2 pandemic in the UK alone [153]. Polymerase chain reaction (PCR) and lateral flow based tests both require a swab sample of either or both of the mouth and or nose. The swabbing of oronasal areas, while yielding a good titre of virus, is uncomfortable for the patient which leads to difficulty collecting viable samples from children and those with learning difficulties, and does not investigate the viral load being shed- this is the virus that could cause infections in other people. Also, each test performed produces large amounts of non-recycled waste in the form of swabs. To avoid these problems we investigated the collection of virus samples through the condensation of breath samples in a reuseable, printable device.

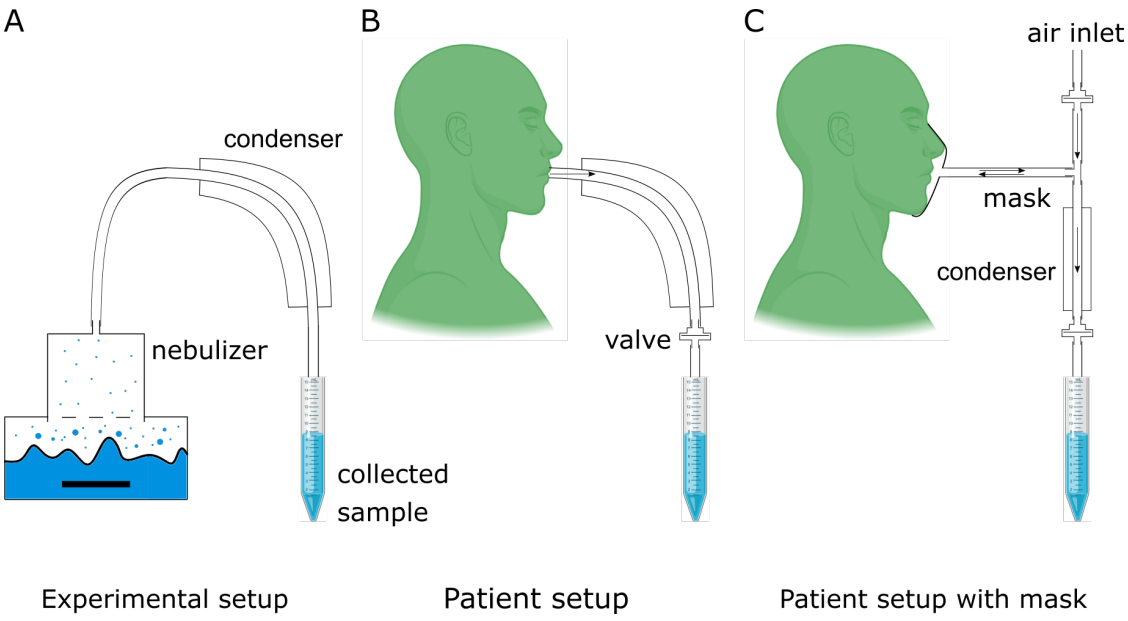


Figure D.1: A schematic of the apparatus used in tests with and without a patient A virus sample could be collected in multiple modalities: A) a nebulised virus sample; B) a patient or C) a masked patient. [52]

D.2 Results

During the Covid-19 pandemic, I investigated the trapping of aerosolised virus particles using a reusable, portable, 3D printed condenser. The functionality could be examined with the setup in Fig. D.1A using the nebulisation of a virus sample as the source of virus, and the testing of patients could be completed with the setups in Fig. D.1B and D.1C dependent on the condition of the person being tested- these setups with valves allow a person to breath normally and prevent the inhalation of the condensed oral sample..

The device was modelled using 3D CAD software (Fig. D.2A) [154]. The CAD rendering of the device was converted to a 3D printable file with supports and drainage holes (Fig. D.2B) [155]. Finally, the device was printed and cured on an ANYCUBIC Photon X [156] resin 3D printer (Fig D.2C).

An internal chamber is part-filled with water and frozen before a sample was taken by nebulising a sample of A/Udorn/72. By labelling diffusing virus particles with CaCl2 and fluorescently labelled DNA [85], diffusing virus particles were imaged. The amount of virus condensed from the nebulised sample was not large enough to

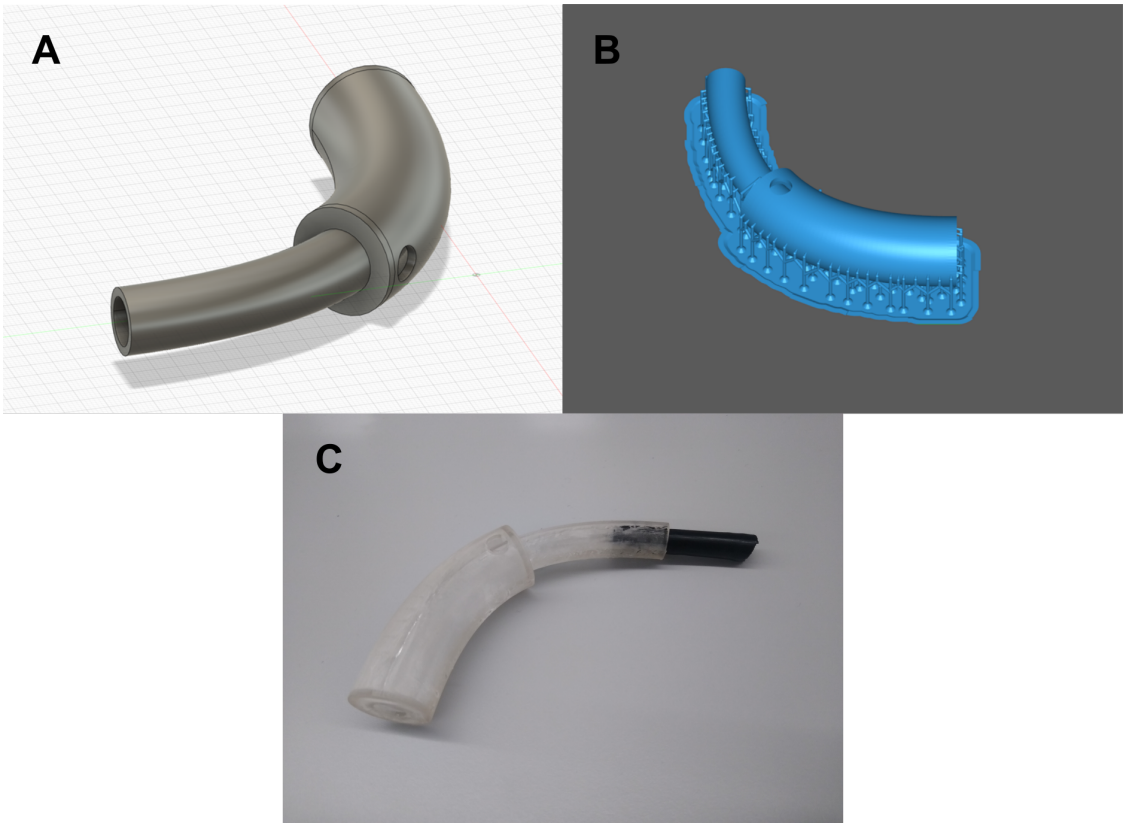


Figure D.2: The designs and final 3D printout of the device A) The 3D CAD model of the device. B) The 3D printable design with supports and holes for the support of the device and the draining of the resin during printing. C) The final, printed device with an inserted PET tube for use on multiple patients.

be detected using single molecule fluorescent microscopy (Fig. D.3A). A sample from a swab from the inside of the device found 2 viral particles in a 30 s acquisition (Fig D.3B), this was deemed to be too low in sensitivity as thousands of viral particles were detected in the original viral sample that was nebulised (Fig. D.3C).

D.3 Conclusions

Although the device meets the demands of being cleaned and reused, eliminating the majority of waste from other virus sampling methods, and was also cheap and mass producible through 3D printing, the fact that the collected sample did not contain sufficient virus to be detected meant that the development of the device was stopped. The plans for the device are available publicly at https://www.thingiverse.com/thing:5121572 [157].

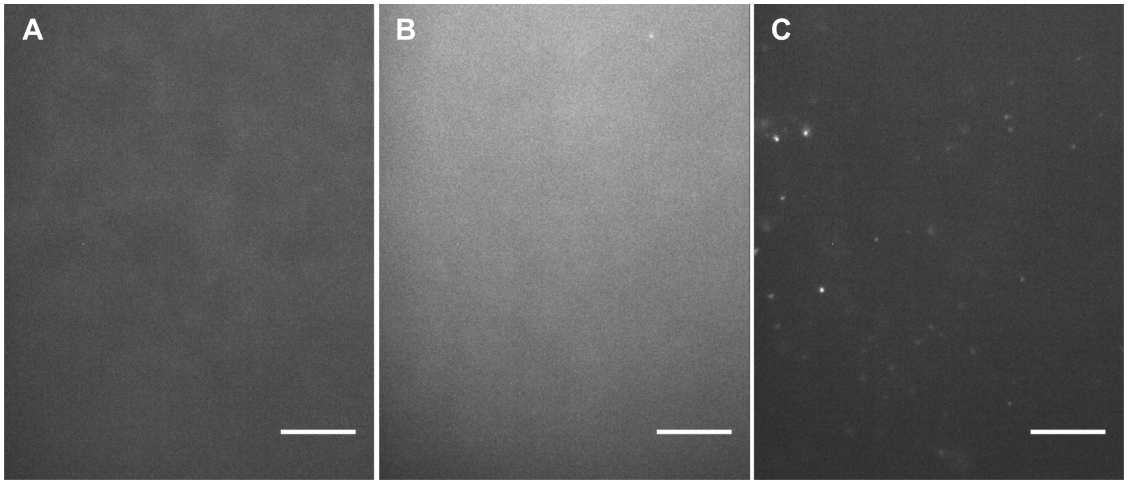


Figure D.3: Single molecule fluorescent imaging of diffusing viral particles captured and in the original sample. A) An image taken from a 30 s acquisition of the condensation sample. No virus particles were seen. B) An image taken from a 30 s acquisition of the sample swabbed from the inside of the device. 2 virus particles were seen. C) An image taken from a 30 s acquisition of the original sample. Thousands of virus particles were seen. Scale bars 10 μm.

References

[1] Helmut Ruska, Bodo von Borries, and Ernst Ruska. “Die Bedeutung der Übermikroskopie für die Virusforschung”. In: *Archiv für die gesamte Virusforschung* 1 (2005), pp. 155–169.

[2] Maximilian Haider et al. “Towards 0.1 nm resolution with the first spherically corrected transmission electron microscope”. In: *Journal of Electron Microscopy* 47.5 (Jan. 1998), pp. 395–405.

[3] O.P. Choudhary and Priyanka Choudhary. “Scanning Electron Microscope: Advantages and Disadvantages in Imaging Components”. In: *International Journal of Current Microbiology and Applied Sciences* 6 (May 2017), pp. 1877–1882.

[4] E. Betzig et al. “Breaking the Diffraction Barrier: Optical Microscopy on a Nanometric Scale”. In: *Science* 251.5000 (1991), pp. 1468–1470.

[5] B. et al. Dadonaite. “Filamentous influenza viruses”. In: *J Gen Virol* 97 (2016), pp. 1755–1764.

[6] WHO. *Influenza (Seasonal)*. Online Multimedia. Accessed: 20/06/2023. 2019. URL: `https://www.who.int/news-room/fact-sheets/detail/influenza-(seasonal)`.

[7] WHO. *HIV and AIDS*. Online Multimedia. Accessed: 21/06/2023. 2023. URL: `https://www.who.int/news-room/fact-sheets/detail/hiv-aids`.

[8] N. M. Bouvier and Palese P. “The biology of influenza viruses”. In: *Vaccine* 26(Suppl 4).D48-53 (2008).

[9] T. Noda. “Native Morphology of Influenza Virions”. In: *Front. Microbiol* 2 (2012).

[10] Jack C. Hirst and Edward C. Hutchinson. “Single-particle measurements of filamentous influenza virions reveal damage induced by freezing”. In: 100.12 (2019), pp. 1631–1640.

[11] A. Harris, G. Winkler Cardone, et al. “Influenza virus pleiomorphy characterized by cryoelectron tomography”. In: *Proc. Natl. Acad. Sci. USA* (2006).

[12] I. A. Wilson, J. J. Skehel, and D. C. Wiley. “Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å Resolution”. In: *Nature* 289.5796 (1981), pp. 366–373.

[13] T Stegmann, F P Booy, and J Wilschut. “Effects of low ph on influenza virus. activation and inactivation of the membrane fusion capacity of the hemagglutinin.” In: *Journal of Biological Chemistry* 262.36 (1987), pp. 17744–17749.

[14] J. N. Varghese, W. G. Laver, and P. M. Colman. “Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 Å Resolution”. In: *Nature* 303.5912 (1983), pp. 35–40.

[15] C. Wang, R.A. Lamb, and L.H. Pinto. “Activation of the M2 ion channel of influenza virus: A role for the transmembrane domain histidine residue”. In: *Biophysical Journal* 69.4 (1995), pp. 1363–1371.

[16] E.C. Hutchinson, P.D. Charles, S.S. Hester, et al. “Conserved and host specific features of influenza virion architecture”. In: *Nat. Communications* 5.4816 (2014).

[17] D.M. McCraw, J.R. Gallagher, U. Torian, et al. “Structural analysis of influenza vaccine virus-like particles reveals a multicomponent organization”. In: *Scientific Reports* 8.1 (2018).

[18] M.D. Vahey and D.A. Fletcher. “Influenza A virus surface proteins are organized to help penetrate host mucus”. In: *eLife* (2019).

[19] Petr Chlanda et al. “Structural analysis of the roles of influenza A virus membrane-associated proteins in assembly and morphology”. In: *Journal of Virology* 89.17 (2015), pp. 8957–8966.

[20] Sebastian Wasilewski et al. “Distribution of surface glycoproteins on influenza A virus determined by electron cryotomography”. In: *Vaccine* 30.51 (2012), pp. 7368–7373.

[21] L.J. Calder, S. Wasilewaki, J.A. Berriman, et al. “Strucural organization of filamentous influenza A virus”. In: *Proc. Natl. Acad. Sci. USA* 107.23 (2010), pp. 10685–10690.

[22] S.V. Bourmakina and A. Garcí-Sastre. “Reverse genetics studies on the filamentous morphology of influenza A virus”. In: *J Gen Virol* 84 (2003), pp. 517–527.

[23] Swetha Vijayakrishnan et al. “Cryotomography of Budding Influenza A virus reveals filaments with diverse morphologies that mostly do not bear a genome at their distal end”. In: *PLoS Pathogens* 9.6 (2013).

[24] I. Archetti. “Appearances associated with filamentous forms of influenza viruses”. In: *Arch Gesamte Virusforschung* 6.1 (1955), pp. 29–35.

[25] R.E. O’Neill, J. Talon, and P. Palese. “The influenza virus NEP (NS2 protein) mediates the nuclear export of viral ribonucleoproteins”. In: *The EMBO Journal* 17 (1998), pp. 288–296.

[26] Holly Else. “How a torrent of Covid Science Changed Research Publishing — in Seven charts”. In: *Nature* 588.7839 (2020), pp. 553–553.

[27] Zunlong Ke et al. “Structures and distributions of SARS-COV-2 spike proteins on intact virions”. In: *Nature* 588.7838 (2020), pp. 498–502.

[28] Chuang Liu et al. “The architecture of inactivated SARS-COV-2 with Postfusion spikes revealed by cryo-EM and Cryo-et”. In: *Structure* 28.11 (2020).

[29] Sébastien Lyonnais et al. “Atomic Force Microscopy analysis of native infectious and inactivated SARS-COV-2 virions”. In: *Scientific Reports* 11.1 (2021).

[30] Hangping Yao et al. “Molecular architecture of the SARS-COV-2 virus”. In: *Cell* 183.3 (2020).

[31] Bertrand Boson et al. “The SARS-COV-2 envelope and membrane proteins modulate maturation and retention of the spike protein, allowing assembly of virus-like particles”. In: *Journal of Biological Chemistry* 296 (2021), p. 100111.

[32] Paul S. Masters. “Coronavirus genomic RNA packaging”. In: *Virology* 537 (2019), pp. 198–207.

[33] Harish Nair et al. In: *BMC Public Health* 11.S3 (2011).

[34] In: *The Lancet* 385.9963 (2015), pp. 117–171.

[35] Yung T. Huang, Peter L. Collins, and Gail W. Wertz. “Characterization of the 10 proteins of human respiratory syncytial virus: Identification of a fourth envelope-associated protein”. In: *Virus Research* 2.2 (1985), pp. 157–173.

[36] J. A. Morris, R. E. Blount, and R. E. Savage. “Recovery of cytopathogenic agent from chimpanzees with goryza”. In: *Experimental Biology and Medicine* 92.3 (1956), pp. 544–549.

[37] T. Samji. “Influenza A: understanding the viral life cycle.” In: *Yale J Biol Med.* 82(4).40 (2009), pp. 153–159.

[38] John J. Skehel and Don C. Wiley. “Receptor binding and membrane fusion in virus entry: The influenza hemagglutinin”. In: *Annual Review of Biochemistry* 69.1 (2000), pp. 531–569.

[39] Jeremy S. Rossman, George P. Leser, and Robert A. Lamb. “Filamentous influenza virus enters cells via macropinocytosis”. In: *Journal of Virology* 86.20 (2012), pp. 10950–10960.

[40] Lawrence H. Pinto and Robert A. Lamb. “The M2 proton channels of influenza A and B viruses”. In: *Journal of Biological Chemistry* 281.14 (2006), pp. 8997–9000.

[41] Stephen J. Plotch, Michele Bouloy, and Robert M. Krug. “Transfer of 5-terminal cap of globin mrna to influenza viral complementary RNA during transcriptionnbsp; nbsp; nbsp; nbsp; nbsp; nbsp; in vitro”. In: *Proceedings of the National Academy of Sciences* 76.4 (1979), pp. 1618–1622.

[42] Debi P. Nayak et al. “Influenza virus morphogenesis and budding”. In: *Virus Research* 143.2 (2009), pp. 147–161.

[43] Kiyoko Iwatsuki-Horimoto et al. “The cytoplasmic tail of the influenza A virus M2 protein plays a role in Viral Assembly”. In: *Journal of Virology* 80.11 (2006), pp. 5233–5240.

[44] Debi P. Nayak, Eric Ka-Wai Hui, and Subrata Barman. “Assembly and budding of influenza virus”. In: *Virus Research* 106.2 (2004), pp. 147–165.

[45] Julia Peukes et al. “The native structure of the assembled matrix protein 1 of influenza A virus”. In: *Nature* 587.7834 (2020), pp. 495–498.

[46] Boundless. *Microbiology*. LibreTexts, 2023.

[47] A. Jablonski. “Efficiency of anti-stokes fluorescence in dyes”. In: *Nature* 131.3319 (1933), pp. 839–840.

[48] J.M. Bator and C.L. Reading. “Measurement of antibody affinity for cell surface antigens using an enzyme-linked immunosorbent assay”. In: *Journal of Immunological Methods* 125.1–2 (1989), pp. 167–176.

[49] Bjoern Traenkle and Ulrich Rothbauer. “Under the microscope: Single-domain antibodies for live-cell imaging and super-resolution microscopy”. In: *Frontiers in Immunology* 8 (2017).

[50] D Axelrod. “Cell-substrate contacts illuminated by total internal reflection fluorescence.” In: *Journal of Cell Biology* 89.1 (1981), pp. 141–145.

[51] William J Betz, Fei Mao, and Corey B Smith. “Imaging exocytosis and endocytosis”. In: *Current Opinion in Neurobiology* 6.3 (1996), pp. 365–371.

[52] Created with Biorender.com.

[53] E. Abbe. “Beiträge zur Theorie des Mikroskops und der Mikroskopischen Wahrnehmung”. In: *Archiv für Mikroskopische Anatomie* 9.1 (1873), pp. 413–468.

[54] Thomas A. Klar and Stefan W. Hell. “Subdiffraction resolution in far-field fluorescence microscopy”. In: *Opt. Lett.* 24.14 (July 1999), pp. 954–956.

[55] Eric Betzig et al. “Imaging Intracellular Fluorescent Proteins at Nanometer Resolution”. In: *Science* 313.5793 (2006), pp. 1642–1645.

[56] Samuel T. Hess, Thanu P.K. Girirajan, and Michael D. Mason. “Ultra-High Resolution Imaging by Fluorescence Photoactivation Localization Microscopy”. In: *Biophysical Journal* 91.11 (2006), pp. 4258–4272.

[57] Michael J. Rust, Mark Bates, and Xiaowei Zhuang. “Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)”. In: *Nature Methods* 3 (2006), pp. 793–796.

[58] Stefan Bretschneider, Christian Eggeling, and Stefan W. Hell. “Breaking the Diffraction Barrier in Fluorescence Microscopy by Optical Shelving”. In: *Phys. Rev. Lett.* 98 (21 May 2007), p. 218103.

[59] John M. Guerra. “Super-resolution through illumination by diffraction-born evanescent waves”. In: *Applied Physics Letters* 66.26 (1995), pp. 3555–3557.

[60] Alexey Sharonov and Robin M. Hochstrasser. “Wide-field subdiffraction imaging by accumulated binding of diffusing probes”. In: *Proceedings of the National Academy of Sciences* 103.50 (2006), pp. 18911–18916.

[61] Matthew P. Gordon, Taekjip Ha, and Paul R. Selvin. “Single-molecule high-resolution imaging with photobleaching”. In: *Proceedings of the National Academy of Sciences* 101.17 (2004), pp. 6462–6465.

[62] Xiaohui Qu et al. “Nanometer-localized multiple single-molecule fluorescence microscopy”. In: *Proceedings of the National Academy of Sciences* 101.31 (2004), pp. 11298–11303.

[63] Yide Zhang et al. “Generalized stepwise optical saturation enables super-resolution fluorescence lifetime imaging microscopy”. In: *Biomedical Optics Express* 9.9 (2018), p. 4077.

[64] Mike Heilemann et al. “Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes”. In: *Angewandte Chemie International Edition* 47.33 (2008), pp. 6172–6176.

[65] Adapted from “How To Break The Diffraction Limit With STORM” by BioRender.com. 2021. URL: `https://app.biorender.com/biorender-templates`.

[66] J Neumann and R. D. Richtmyer. “2. statistical methods in neutron diffusion”. In: *Analogies Between Analogies* (1990), pp. 17–36.

[67] Victor Platon and Andreea Constantinescu. "Monte Carlo method in risk analysis for investment projects". In: *Procedia Economics and Finance* 15 (2014), pp. 393–400.

[68] William Ogilvy Kermack, A. G. McKendrick, and Gilbert Thomas Walker. "A contribution to the mathematical theory of epidemics". In: *Proc. R. Soc. Lond. A* 115 (1997), pp. 700–721.

[69] A. O. Victor and H. K. Oduwole. "Evaluating the deterministic SERIUS model for disease control in an age-structured population". In: *Global Sci J* 8.3 (2020).

[70] Tuen Wai Ng, Gabriel Turinici, and Antoine Danchin. "A double epidemic model for the SARS propagation". In: *BMC Infect DIs.* 3.19 (2003).

[71] Catherine Armitage. "The high burden of infectious disease". In: *Nature* 598.7882 (2021).

[72] Edouard Mathieu et al. "Coronavirus Pandemic (COVID-19)". In: *Our World in Data* (2020). https://ourworldindata.org/coronavirus.

[73] Zunlong Ke et al. "The morphology and assembly of respiratory syncytial virus revealed by cryo-electron tomography". In: *Viruses* 10.8 (2018), p. 446.

[74] Tanmay A. Bharat et al. "Structural dissection of ebola virus and its assembly determinants using cryo-electron tomography". In: *Proceedings of the National Academy of Sciences* 109.11 (2012), pp. 4275–4280.

[75] Masuyo Nakai and David T. Imagawa. "Electron microscopy of measles virus replication". In: *Journal of Virology* 3.2 (1969), pp. 187–197.

[76] Qizhi Yao and Richard W. Compans. "Filamentous particle formation by human parainfluenza virus type 2". In: *J Gen Virol* 81.5 (2000), pp. 1305–1312.

[77] Romain F Laine et al. "Structured illumination microscopy combined with machine learning enables the high throughput analysis and classification of virus structure". In: *eLife* 7 (2018).

[78] Jill Seladi-Schulman, John Steel, and Anice C. Lowen. "Spherical influenza viruses have a fitness advantage in embryonated eggs, while filament-producing strains are selected in vivo". In: *Journal of Virology* 87.24 (2013), pp. 13343–13353.

[79] Seema S. Lakdawala et al. "Eurasian-origin gene segments contribute to the transmissibility, aerosol release, and morphology of the 2009 pandemic H1N1 influenza virus". In: *PLoS Pathogens* 7.12 (2011).

[80] Patricia J. Campbell et al. "The M segment of the 2009 pandemic influenza virus confers increased neuraminidase activity, filamentous morphology, and efficient contact transmissibility to a/puerto rico/8/1934-based reassortant viruses". In: *Journal of Virology* 88.7 (2014), pp. 3802–3814.

[81] Tian Li et al. "The shape of pleomorphic virions determines resistance to cell-entry pressure". In: *Nature Microbiology* 6.5 (2021), pp. 617–629.

[82] Emma Touizer et al. "Application of super-resolution and advanced quantitative microscopy to the spatio-temporal analysis of influenza virus replication". In: *Viruses* 13.2 (2021), p. 233.

[83] Maria Arista-Romero, Silvia Pujals, and Lorenzo Albertazzi. "Towards a quantitative single particle characterization by Super Resolution Microscopy: From virus structures to antivirals design". In: *Frontiers in Bioengineering and Biotechnology* 9 (2021).

[84] Andrew McMahon et al. "High-throughput super-resolution analysis of influenza virus pleomorphism reveals insights into viral spatial organization". In: (2021).

[85] N. C. Robb et al. "Rapid functionalisation and detection of viruses via a novel Ca(2+)-mediated virus-DNA interaction". In: *Sci Rep* 9.1 (2019), p. 16219.

[86] Sonka, Hlavac, and Boyle. *Image Processing Analysis and machine vision*. Cengage, 2015.

[87] S.L. Wearne et al. "New techniques for imaging, digitization and analysis of three-dimensional neural morphology on multiple scales". In: *Neuroscience* 136.3 (2005), pp. 661–680.

[88] Yuan Xiong et al. "Automated characterization of cell shape changes during amoeboid motility by skeletonization". In: *BMC Systems Biology* 4.1 (2010).

[89] Martin Ester et al. "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise". In: KDD'96 (1996), pp. 226–231.

[90] Adrian Przybylski et al. "Gpufit: An open-source toolkit for GPU-accelerated curve fitting". In: *Scientific Reports* 7.1 (2017).

[91] Susanna M. Früh et al. "Site-specifically-labeled antibodies for super-resolution microscopy reveal in situ linkage errors". In: *ACS Nano* 15.7 (2021), pp. 12161–12170.

[92] Jeremy S. Rossman and Robert A. Lamb. "Influenza virus assembly and budding". In: *Virology* 411.2 (2011), pp. 229–236.

[93] Jeremy S. Rossman et al. "Influenza virus M2 protein mediates ESCRT-independent membrane scission". In: *Cell* 142.6 (2010), pp. 902–913.

[94] Michael D. Vahey and Daniel A. Fletcher. "Low-fidelity assembly of influenza A virus promotes escape from host cells". In: *Cell* 176.1-2 (2019).

[95] Romain Laine et al. "Structural analysis of herpes simplex virus by optical super-resolution imaging". In: *Nature Communications* 6.1 (2015).

[96] Cândida F Pereira et al. "HIV taken by storm: Super-resolution fluorescence microscopy of a viral infection". In: *Virology Journal* 9.1 (2012).

[97] Barbara Storti et al. "A spatial multi-scale Fluorescence Microscopy Toolbox discloses entry checkpoints of SARS-COV-2 variants in vero E6 cells". In: *Computational and Structural Biotechnology Journal* 19 (2021), pp. 6140–6156.

[98] M. BUONGIORNO-NARDELLI and F. AMALDI. "Autoradiographic detection of molecular hybrids between rrna and DNA in tissue sections". In: *Nature* 225.5236 (1970), pp. 946–948.

[99] Maeva Dupont et al. "Tunneling nanotubes: Intimate communication between myeloid cells". In: *Frontiers in Immunology* 9 (2018).

[100] Kari L. Roberts, Balaji Manicassamy, and Robert A. Lamb. "Influenza A virus uses intercellular connections to spread to neighboring cells". In: *Journal of Virology* 89.3 (2015), pp. 1537–1549.

[101] Amrita Kumar et al. “Influenza virus exploits tunneling nanotubes for cell-to-cell spread”. In: *Scientific Reports* 7.1 (2017).

[102] Timothy Philpot and Jeffery S. Thomas. *Mechanics of Materials: An Integrated Learning System*. Wiley, 2020.

[103] Gilberte Gaval and Jean-Jacques Pernelle. “Impact of the repetition of oxygen deficiencies on the filamentous bacteria proliferation in activated sludge”. In: *Water Research* 37.9 (2003), pp. 1991–2000.

[104] Paul C. Roberts and Richard W. Compans. “Host cell dependence of viral morphology”. In: *Proceedings of the National Academy of Sciences* 95.10 (1998), pp. 5746–5751.

[105] Yu. A. Smirnov, M. A. Kuznetsova, and N. V. Kaverin. “The genetic aspects of influenza virus filamentous particle formation”. In: *Archives of Virology* 118.3-4 (1991), pp. 279–284.

[106] T.H.G. Megson. “Structural and stress analysis.” In: (1996).

[107] Daisuke Shiomi, Hideo Mori, and Hironori Niki. “Genetic mechanism regulating bacterial cell shape and metabolism”. In: *Communicative amp; Integrative Biology* 2.3 (2009), pp. 219–220.

[108] T. P. Cushnie, Noëlle H. O’Driscoll, and Andrew J. Lamb. “Morphological and ultrastructural changes in bacterial cells as an indicator of antibacterial mechanism of action”. In: *Cellular and Molecular Life Sciences* 73.23 (2016), pp. 4471–4492.

[109] Yuly A. Jaimes-Lizcano, Dayton D. Hunn, and Kyriakos D. Papadopoulos. “Filamentous Escherichia coli cells swimming in tapered microcapillaries”. In: *Research in Microbiology* 165.3 (2014), pp. 166–174.

[110] Martin W. Hahn and Manfred G. Höfle. “Grazing pressure by a bacterivorous flagellate reverses the relative abundance of comamonas acidovorans PX54 and vibrio strain CB5 in chemostat Cocultures”. In: *Applied and Environmental Microbiology* 64.5 (1998), pp. 1910–1918.

[111] *AKSeg*. Online Multimedia. 2023. URL: `https://github.com/KapanidisLab/Deep-Learning-and-Single-Cell-Phenotyping-for-Rapid-Antimicrobial-Susceptibility-Testing`.

[112] Gabriella Kiss et al. “Structural analysis of respiratory syncytial virus reveals the position of M2-1 between the matrix protein and the ribonucleoprotein complex”. In: *Journal of Virology* 88.13 (2014), pp. 7602–7617.

[113] Thomas Bächi and Calderon Howe. “Morphogenesis and ultrastructure of respiratory syncytial virus”. In: *Journal of Virology* 12.5 (1973), pp. 1173–1180.

[114] Thomas J. Utley et al. “Respiratory syncytial virus uses a VPS4-independent budding mechanism controlled by RAB11-FIP2”. In: *Proceedings of the National Academy of Sciences* 105.29 (2008), pp. 10209–10214.

[115] Hafez el Sayyed et al. “Single-molecule tracking reveals the functional allocation,in vivointeractions and spatial organization of universal transcription factor NUSG”. In: (2022).

[116] Marc Adrian et al. “Cryo-electron microscopy of viruses”. In: *Nature* 308.5954 (1984), pp. 32–36.

[117] David N. Mastronarde. “Automated electron microscope tomography using robust prediction of specimen movements”. In: *Journal of Structural Biology* 152.1 (2005), pp. 36–51.

[118] James R. Kremer, David N. Mastronarde, and J.Richard McIntosh. “Computer visualization of three-dimensional image data using IMOD”. In: *Journal of Structural Biology* 116.1 (1996), pp. 71–76.

[119] Councilman Morgan, Harry M. Rose, and Dan H. Moore. “Structure and development of viruses observed in the electron microscope”. In: *Journal of Experimental Medicine* 104.2 (1956), pp. 171–182.

[120] Sai Li et al. “Bending and puncturing the influenza lipid envelope”. In: *Biophysical Journal* 100.3 (2011), pp. 637–645.

[121] R. Skalak and P I Brånemark. “Deformation of red blood cells in capillaries”. In: *Science* 164.3880 (1969), pp. 717–719.

[122] Anna Szymborska et al. “Nuclear pore scaffold structure analyzed by super-resolution microscopy and particle averaging”. In: *Science* 341.6146 (2013), pp. 655–658.

[123] Ke Xu, Guisheng Zhong, and Xiaowei Zhuang. “Actin, spectrin, and associated proteins form a periodic cytoskeletal structure in axons”. In: *Science* 339.6118 (2013), pp. 452–456.

[124] Alan Turing. “The chemical basis of morphogenesis”. In: *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 237.641 (1952), pp. 37–72.

[125] E. Lee. “Corrections: The roles of APC and axin derived from experimental and theoretical analysis of The wnt pathway”. In: *PLoS Biology* 2.3 (2004).

[126] B.S. Hendriks et al. “Computational modelling of erbb family phosphorylation dynamics in response to transforming growth factor alpha and heregulin indicates spatial compartmentation of phosphatase activity”. In: *IEE Proceedings - Systems Biology* 153.1 (2006), p. 22.

[127] Online Multimedia. URL: `https://ourworldindata.org/covid-vaccinations?country=OWID_WRL` (visited on 01/12/2023).

[128] W. K. Ming and C. J.P. Zhang. “Breaking down of healthcare system: mathematical modelling for controlling the novel coronavirus(2019-nCoV) outbreak in Wuhan, China”. In: *bioRxiv* (2020).

[129] A. Victor. “Mathematical predictions for COVID-19 as a global pandemic”. In: *SSRN Journal* (2020).

[130] I. Nesteruk. “Statistics based predictions of coronavirus 2019-nCoV spreading in mainland China”. In: *medRxiv* (2020).

[131] M. Batista. “Estimation of the final size of the coronavirus epidemic by the SIR model”. In: *ResearchGate* (2020).

[132] Alexander Victor Okheuse. “Estimation of the Probability of Reinfection With COVID-19 by the Susceptible-Exposed-Infectious-Removed-Undetectable-Susceptible Model”. In: *JMIR Public Health and Surveillance* (2020).

[133] A. McMahon and N. C. Robb. "Reinfection with SARS-CoV-2: Discrete SIR (Susceptible, Infected, Recovered) Modeling Using Empirical Infection Data". In: *JMIR Public Health Surveill* 6.4 (2020), e21168.

[134] Federal Office of Public Health FOPH. *Coronavirus*. URL: `https://www.bag.admin.ch/bag/en/home/krankheiten/ausbrueche-epidemien-pandemien/aktuelle-ausbrueche-epidemien/novel-cov/situation-schweiz-und-international.html` (visited on 05/17/2020).

[135] NYC Health. *Covid-19: Data*. URL: `https://www1.nyc.gov/site/doh/covid/covid-19-data.page` (visited on 05/17/2020).

[136] Census.gov. *United States Census Bureau. QuickFacts: Richmond County (Staten Island Borough), New York; New York County (Manhattan Borough), New York; Bronx County (Bronx Borough), New York; Kings County (Brooklyn Borough), New York; Queens County (Queens Borough), New York*. 2019. URL: `tinyurl.com/ybmwj89r` (visited on 05/17/2020).

[137] Trading Economics. *Germany Coronavirus Recovered*. URL: `https://tradingeconomics.com/germany/coronavirus-recovered` (visited on 05/17/2020).

[138] World Health Organization. *Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19)*. URL: `https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf` (visited on 05/17/2020).

[139] Qifang Bi et al. "Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study". In: *The Lancet Infectious Diseases* 20.8 (2020), pp. 911–919.

[140] Jantien A Backer, Don Klinkenberg, and Jacco Wallinga. "Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020". In: *Euro Surveill.* 25.5: 2000062 (2020).

[141] Silvia Stringhini et al. "Seroprevalence of anti-SARS-CoV-2 IgG antibodies in Geneva, Switzerland (SEROCoV-POP): a population-based study". In: *The Lancet* 396.10247 (2020), P313–319.

[142] Sangmi Cha and Josh Smith. *Explainer: South Korean findings suggest 'reinfected' coronavirus cases are false positives*. Report. Reuters, 2020.

[143] Abishek Chandrashekar, Jinyan Liu, and Amanda J. et al. Martinot. "SARS-CoV-2 infection protects against rechallenge in rhesus macaques". In: *Science* (2020).

[144] Kathryn A. Ryan et al. "Dose-dependent response to infection with SARS-CoV-2 in the ferret model: evidence of protection to re-challenge". In: *bioRxiv* (2020).

[145] Neeltje van Doremalen et al. "ChAdOx1 nCoV-19 vaccination prevents SARS-CoV-2 pneumonia in rhesus macaques". In: *bioRxiv* (2020).

[146] Jeffrey Seow et al. "Longitudinal evaluation and decline of antibody responses in SARS-CoV-2 infection". In: *MedRXiv* (2020).

[147] ONS. *Coronavirus (COVID-19) Infection Survey, characteristics of people testing positive for COVID-19, UK: 21 September 2022*. Online Multimedia. URL: https://tinyurl.com/46v6cw8d (visited on 01/13/2023).

[148] CDC. *Covid-19: Data*. Online Multimedia. URL: https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/vaccine-induced-immunity.html (visited on 01/13/2022).

[149] UK government. *Covid-19: Data*. Online Multimedia. URL: https://coronavirus.data.gov.uk/details/vaccinations (visited on 01/13/2023).

[150] Online Multimedia. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19latestinsights/infections#reinfections (visited on 01/13/2023).

[151] Seamus J. Holden, Stephan Uphoff, and Achillefs N. Kapanidis. "DAOSTORM: an algorithm for high- density super-resolution microscopy". In: *Nature Methods* 8 (2011), pp. 279–280.

[152] R. J. Marsh et al. "Artifact-free high-density localization microscopy analysis". In: *Nature Methods* 15 (2018), pp. 689–692.

[153] Online Multimedia. 2022. URL: https://coronavirus.data.gov.uk/details/testing.

[154] Siemens. 2022. URL: https://structures.aero/software/solid-edge/.

[155] CHITUBOX. 2022. URL: https://www.chitubox.com/en/index.

[156] ANYCUBIC. 2022. URL: https://www.anycubic.com.

[157] A. McMahon. 2021. URL: https://www.thingiverse.com/thing:5121572.