Influenza A Virus Nucleoprotein Forms Viscoelastic Condensates in the Nucleus

Tomas Luka Shimkus
Division of Experimental Medicine
McGill University, Montreal

Submitted August 14, 2024
First Published August 14, 2025

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science in Experimental Medicine

Table of Contents

Abstract

Influenza A virus (IAV) is an enveloped orthomyxovirus with a segmented genome consisting of viral ribonucleoproteins (vRNPs). These segments allow reassortment between different influenza strains with the potential to create new pandemic variants. The major component of the virus' vRNPs is nucleoprotein (NP) which coats the viral RNA. In order to replicate its genome and transcribe its mRNAs IAV vRNPs must enter the nucleus. Liquid-liquid phase separation (LLPS) is a process through which proteins separate from a liquid medium via self-attraction mediated by the presence of intrinsically disordered and droplet-promoting regions (DPRs) among other factors. The vRNPs of IAV demix in the cytoplasm with this activity being important for efficient production of progeny viruses. Our group hypothesized that LLPS of NP occurs in the nucleus, resulting in the formation of biomolecular condensates that support viral polymerase activity. Experiments making use of transfection, infection, and *in vitro* conditions demonstrated that WT NP forms nuclear viscoelastic condensates, with certain deletion mutants not doing so. An R74A mutation of NP was investigated since the literature indicates that it has no effect on viral polymerase activity while preventing the formation of infectious virus and affects the RNA binding pocket which may influence LLPS. This mutation was found to affect NP condensation. Investigation of NP mutants is important for the discovery of its functional regions, allowing development of anti-influenza drugs that target NP. Future research should further examine NP, including the R74A mutant, to better understand the role of LLPS during IAV infection and potentially find new therapeutic targets.

Résumé

Le virus de la grippe A (VGA) est un orthomyxovirus enveloppé avec un génome segmenté constitué de ribonucléoprotéines virales (RNPv). Ces segments permettent un réassortiment entre différentes souches de grippe avec le potentiel de créer des nouveaux variants pandémiques. Le principal composant des RNPv du virus est la nucléoprotéine (NP) qui recouvre l'ARN viral. Afin de répliquer son génome et de transcrire ses ARNm, les RNPv IAV doivent pénétrer dans le noyau. La séparation de phase liquide-liquide (SPLL) est un processus par lequel les protéines se séparent d'un milieu liquide par auto-attraction médiée par la présence de régions intrinsèquement désordonnées et les régions favoristant de gouttelettes (RFGs), entre autres facteurs. Les ARNv de l'IAV sont démiques dans le cytoplasme et cette activité est importante pour la production efficace de virus. Notre groupe a émis l'hypothèse que le SPLL de NP se passe dans le noyau, ce qui entraîne la formation de condensats biomoléculaires qui soutiennent l'activité de la polymérase virale. Des expériences utilisant des conditions de transfection, d'infection et la reconstitution *in vitro* ont démontré que le NP forme des condensats viscoélastiques nucléaires car certains mutants de délétion ne le faisant pas. Une mutation de R74A a été étudié puisque la littérature indique qu'il n'a aucun effet sur l'activité de la polymérase virale mais prévente la formation de virus infectieux et qu'il affecte la poche de liaison de l'ARN qui peut influencer le LLPS. On a trouvé que cette mutation affecte la condensation du NP. L'investigation des mutantes de NP est importante pour la découverte des régions fonctionnelles de la protéine qui conduisent à des nouvelles thérapies qui ciblent NP. Les recherches futures devrait examiner plus en détail la NP, y compris le mutant R74A, afin de mieux comprendre le rôle du LLPS lors de l'infection par le VGA et de trouver potentiellement de nouvelles cibles thérapeutiques.

List of Abbreviations

BMC	Biomolecular condensate
DIC	Differential interference contrast
DTT	Dithiothreitol
DPR	Droplet-promoting region
CV	Column volume
FXa	Factor Xa
GFP	Green fluorescent protein
HA	Hemagglutinin
HIV	Human immunodeficiency virus
IAV	Influenza A virus
IDR	Intrinsically disordered region
IPTG	Isopropyl ß-D-1-thiogalactopyranoside
LB	Luria broth
LLPS	Liquid-liquid phase separation
MA	Matrix
MBP	Maltose binding protein
MNG	mNeonGreen
MW	Molecular weight
MWCO	Molecular weight cutoff
NA	Neuraminidase
NIH	National Institutes of Health

NP	Nucleoprotein
NS1	Non-structural 1
NS2	Non-structural 2 (also known as NEP, nuclear export protein)
OD	Optical density
PA	Polymerase acidic
PB1	Polymerase basic 1
PB2	Polymerase basic 2
PMSF	Phenylmethylsulfonyl fluoride
RNP	Ribonucleoprotein
SA	Sialic acid
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
TCEP	Tris (2-carboxyethyl) phosphine
WB	Western blot

Table of Figures

Acknowledgements

I would first and foremost like to thank my supervisor Andrew J. Mouland whose advice and expertise in biological research was invaluable and with whom the project was originally conceived. A portion of the research was conducted under the supervision of Maria João Amorim, who I would like to thank for allowing me into her laboratory and for helping me plan out my project, as well as for the essential advice, materials and protocols that she and her laboratory members were able to provide me. I would like to thank Oleg Chertkov, who provided advice and protocols and who helped me with purifying protein and preliminary *in vitro* experiments. I would also like to thank my thesis committee members Chen Liang, Rongtuan Lin, Carmen Loiselle, and David Eidelman for their advice and feedback during the thesis committee meetings. Laboratory technical training was provided by Meijuan Niu who showed me the basics of cell culture and molecular biology and assisted with archival of created materials. I would like to also thank my laboratory members Gabriel Guajardo Contreras, Ana-Luiza dos Santos and Bao-An-Chau who provided me with advice and moral support during these efforts. Influenza virus reverse genetics plasmids were a kind gift from Ervin Fodor and Maria João Amorim. HIV-1 Gag was obtained from Leslie J. Parent. The work was conducted mostly at the HIV-1 RNA Trafficking Laboratory, Lady Davis Institute, Montreal, QC, Canada. A portion of the work was conducted at the Cell Biology of Viral Infection Laboratory, Instituto Gulbenkian de Ciência, Oeiras, Portugal. Funding for this project was drawn from grants from the following organizations: McGill Department of Medicine, Division of Experimental Medicine, Canadian Institutes of Health Research, European Commission, European Research Council and the US National Institutes of Health. The funding for my laboratory stay at the Cell Biology of Viral Infection laboratory was from a McGill Graduate Mobility Award. Imaging data

was obtained using instruments and advice from the Lady Davis Institute imaging facility and the Insituto Gulbenkian de Ciência imaging facility. Image processing and analysis for this thesis was partially performed at the McGill University Advanced BioImaging Facility (ABIF), RRID:SCR_017697.

Contribution of Authors

Experiments were conceived by Tomas Shimkus, Andrew J. Mouland and Maria João Amorim. The protocol for NP expression was derived from Oleg Chertkov, who also assisted with the preliminary *in vitro* experiments. All experiments were performed by Tomas Shimkus, who also analysed the data. All thesis chapters were prepared and edited by Tomas Shimkus and reviewed by Andrew J. Mouland.

Introduction

Influenza A virus (IAV) is an enveloped orthomyxovirus with a segmented negative-sense RNA genome that is packaged into eight viral ribonucleoproteins (vRNPs). These consist mostly of nucleoprotein (NP) which bind the RNA along its length in a helical formation as well as a polymerase complex that binds to a short double-stranded portion of viral RNA at one end of the vRNP.[1] The virus' negative genome orientation means that the viral polymerase must generate both mRNA to produce proteins and a complementary positive intermediate to enable genome replication.[2] During infection, the vRNPs move to the nucleus and begin to be transcribed and replicated. The polymerase complex remains associated with the vRNP throughout this process and nascent cRNA produced immediately associates with new NP and polymerase to form additional vRNPs.[2]

Liquid-liquid phase separation (LLPS) is a phenomenon implicated in a variety of cellular and viral processes, and involves the separation of molecules from a liquid into distinct droplets.[3] In IAV, LLPS is known to be involved during the transport and assembly of its genome segments in the cytoplasm following their replication, with host Rab11a and vRNPs forming puncta with liquid properties that appear to be involved in viral reassortment of its genome segments.[4] This raises the question; does this occur elsewhere in the cell? Both Epstein-Barr virus and herpes simplex virus 1 are known to form condensates in cell nuclei to enhance viral genome replication and gene transcription respectively. Meanwhile, SARS-CoV-2 is believed to use phase-separation in the cytoplasm to enhance the replication of its genome.[3] With these factors in consideration, IAV, as an RNA virus replicating in the nucleus may make use of LLPS of vRNPs in the nucleus to concentrate cellular and viral factors to enhance transcription and genome replication.

IAV NP is likely the driving force behind any phase separation of vRNPs during infection due to its high abundance in these complexes and ability to bind RNA (which usually promotes LLPS).[2,3] In addition, the chemical nucleozin hardens IAV condensates in the cytoplasm likely via its interaction with NP, which implies a central role for this protein.[5] In this project, NP phase separation will be investigated in human cell nuclei and *in vitro*.

Review of Relevant Literature

This literature review will focus on three main areas: Influenza A virus, biomolecular condensates, and the intersection between the two.

Influenza A Virus History

It is not known when influenza virus first infected humans. There has been mention of epidemic respiratory disease with symptoms consistent with influenza in Europe since at least the 9th century CE. It was occasionally called by the Italian name "Influenza" during outbreaks in 1173 and 1387, referencing the idea that the disease was "influenced by the stars", with its source being liquid flowing from those heavenly bodies.[6,7] The first mention of pandemic influenza in the historical record comes from the early 16th century shortly after the Columbian expeditions to the Americas. It is believed that this lengthy 1510 pandemic entered Europe via Italy or North Africa, eventually spreading throughout the Old World. It was characterized as a "gasping oppression" which affected mostly the very young and very old, similar to seasonal influenza outbreaks today.[6]

By 1700 influenza was increasingly being recognized as a distinct disease with specific symptoms, mostly involving hoarseness of voice, coughing, and sweating. It was known by a multitude of different names such as “Sweate” and “Catarrh” though in 1742-43 during a European outbreak it began to be known in English as simply “Influenza”.[7] The next large pandemic in 1830-33, and the newspaper coverage it generated as it spread supported the idea of pandemics as general phenomena. The later 1847-51 pandemic was the first to be studied epidemiologically. In 1892, amidst the pandemic of 1889-94 a microbial agent dubbed “Pfeiffer’s Baccillus”, now known by the name *Haemophilus influenzae* was deemed the causative agent of influenza since it had been found in the lungs of so many diseased victims. During the especially large pandemic of 1918-19 the idea that the affliction was caused by a non-bacterial filter-passing agent became more prominent though studies at the time were inconclusive or contradictory.[8] By the time of this pandemic a number of filter-passing agents had already been identified and the term “virus” was coined to denote a hypothetical infectious organism with vaguely defined properties. The 1918 Pandemic was the most deadly event in human history with around 50 million dead and morbidity rates of between 25-40%.[9] The disease was marked by sudden onset and rapid spread and was complicated by many secondary bacterial infections. It would not be until 1933 that influenza would be confirmed to be a virus, with purified stock isolated from human patients.[10] The influenza virus strains discovered around this time, such as the WSN (1933) and PR8 (1934) strains of IAV are still used in research to this day as laboratory and vaccine strains.[11] This discovery and isolation of influenza helped increase interest in the burgeoning field of virology, leading to developments such as vaccines and anti-serums for use against the virus by the 1940s.[12] It was also during this period that human influenza was first used to infect ferrets and other mammals, supporting the concept of zoonotic transmission.[10]

Propagation of influenza in other animals and fertilized chicken eggs allowed for mass production of the virus for formalin-inactivated vaccines.[13]

During the Asian Influenza pandemic of 1957-58 the causative viral strain, otherwise similar to the human-adapted 1918 strain, had three of its genome segments replaced with those of an avian influenza virus, prompting further study into the phenomenon of zoonotic genetic shift, leading soon after to the discovery of genetic shifts caused by reassortment of influenza virus segments from different species.[13] In the 1990s efforts were made to recover samples of 1918 virus from victims in the Alaskan permafrost eventually leading to the full sequence of the 1918 virus being discovered by aligning RNA fragments recovered from a corpse.[14] Around the year 2000 reliable reverse-genetics systems were developed for influenza in order to allow the expression of virus from plasmids transfected into human cell lines, first as a 12, then an 8, and finally a single plasmid system.[11,15] The 12 plasmid system makes use of 8 plasmids containing promoters specific to RNA polymerase I which result in production of the virus' genomic RNA (vRNA) while the other four have cytomegalovirus (CMV) promoters that allow production of mRNA encoding the components of the viral polymerase complex. The polymerase complex can then go on to produce the other viral proteins from the vRNAs.[16] In 2005 this system was used to resurrect and study the 1918 strain in more detail allowing for a better understanding of the cytokine storm phenomena which resulted in high death rates among young people during the 1918 pandemic.[17] This reverse genetics system has been key to continued research on influenza, and has also enabled the production of recombinant strains for use in vaccines starting from the mid-2000s. These usually make use of PR8 with certain segments replaced with those of the seasonal circulating strain.[18] Since then, work on influenza has continued, especially in the wake

of the 2009 Swine Flu pandemic and has resulted in the creation of a number of new production methods, vaccines and antivirals to help combat this age-old menace.[19]

Influenza A Virus Biology

Influenza A virus is an enveloped orthomyxovirus with a negative-sense segmented genome. It can be either spherical or filamentous in shape. The eight segments of its genome encode ten total viral proteins, with the larger six encoding one protein each and the two smaller segments encoding two each via alternative splicing.[20,21] Additional proteins from alternative splicing or ribosomal frameshifting have also been identified in some strains.[21] Each genome segment, or viral nucleoprotein (vRNP), is made up of the viral genomic RNA (vRNA), the three polymerase components: polymerase acidic (PA), polymerase basic 1 (PB1) and polymerase basic 2 (PB2), as well as nucleoprotein (NP) which is the most abundant of the viral proteins during infection, forming most of the mass of the genome segments and coating the delicate vRNA.[22] The inner surface of the virus' outer phospholipid membrane is coated in matrix 1 (M1) protein which provides stability and structure while the outer surface is dotted with hemagglutinin (HA), involved in cell attachment and neuraminidase (NA), involved in cell release during budding. In the inner lumen space of the virus some non-structural 1 (NS1) and non-structural 2 (NS2, also known as nuclear export protein, NEP) proteins can also be found, responsible for evading host immune response and nuclear export of vRNPs respectively. The envelope is also perforated by ion channel matrix 2 (M2) proteins which enable acidification of the viral particle, mediating membrane fusion after endocytosis **(Figure 1).**[19]

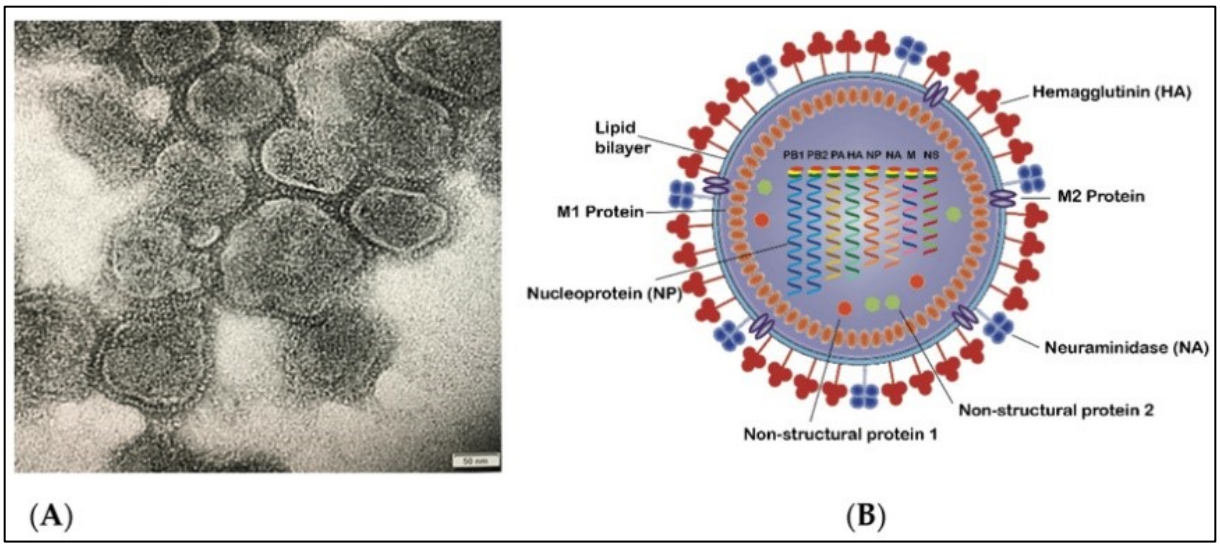


Figure 1: Appearance of Influenza A Virus

(A) A transmission EM electrograph of IAV particles. **(B)** A diagram of an IAV particle showing the locations of the viral proteins. Figure obtained from Nuwarda et al. Vaccines 2021 [https://www.mdpi.com/2076-393X/9/9/1032]. © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This Figure was published in an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

In humans, influenza A virus primarily infects airway epithelial cells with sialic acid (SA) containing an alpha-2,5 linkage. However, they can also infect alveolar macrophages and other types of immune cells, though likely without the ability to produce infectious virions.[23] The linkage type present in SA is a major restrictor of IAV infection, with the HA of avian-adapted strains of IAV unable to bind human SA.[24] The hemagglutinin attaches to the sialic acid and the viral particle scans the surface of the cell for an as-of-yet unidentified receptor (**Figure 2:1**). Once it is found, the virus undergoes clathrin-mediated endocytosis (**Figure 2:2**). Inside the

endosome, the environment begins to acidify, which results in a conformational change in HA which results in fusion between the virus membrane and the endosome. At the same time, the ion channel M2 allows the inside of the viral particle to acidify, releasing the vRNPs from M1 and allowing them to enter the cytoplasm (**Figure 2:3**). In the cytoplasm the vRNPs traffic to the nucleus via nuclear pores using importin alpha/beta complexes (**Figure 2:4**). In the nucleus the virus must produce both viral mRNAs (vmRNAs) to produce proteins (**Figure 2:5a**) as well as vRNA to encode genetic information. Transcription of vmRNAs initially prevails at earlier stages of infection, but at later stages genome replication is favored.[25] Transcription from a vRNP is primed via a 5' cap snatched and cleaved from a cellular mRNA by PB2 and PA of the polymerase complex while the base addition is performed by the PB1 component. To avoid degradation of this vmRNA poly-A tracts are added to the 3' end of the vmRNA via polymerase stuttering. A polymerase complex must be present in close proximity to an existing vRNP in order for replication to initiate, with a host protein ANP32 both bridging the gap between the polymerases and associating with NP to enable efficient encapsulation of the nascent vRNA.[26] The new polymerase is responsible for encapsulation by serving as a temple for the helical assembly of NP while the polymerase on the existing segment actually performs the replication. As the genome is negative sense, replication must occur via a positive-sense intermediate, also in the form of a vRNP, known as a complementary vRNP (cRNP) (**Figure 2:6a**). During cRNP to vRNP replication a third trans-activating polymerase is also required.[27] The polymerase subunits and NP produced in the cytoplasm from vmRNA are imported into the nucleus via their strong nuclear localization sequences to form new vRNPs **(Figure 2:5b-c**). Once produced, the vRNPs bind to M1 and NS2 and are exported to the cytoplasm using a CRM-1-dependent method (**Figure 2:7a**). In the cytoplasm the genome segments bundle together and are transported to the

plasma membrane via Rab11 on recycling endosomes or via a modified ER (also involving Rab11) with liquid biomolecular condensates **(Figure 2:7 b-c**).[28] It is not yet known how the vast majority of viral particles end up containing one of each genome segment, though correct assembly of the segments likely occurs at this stage (**Figure 2:8**).[29,30] Membrane proteins of IAV (HA, NA, and M2) are produced in the ER via co-translational translocation which strings them through translocons in the ER membrane following exposure of the SRP sequence on the proteins from a translating ribosome (**Figure 2:5d**).[28] Once produced, the membrane proteins transit through the Golgi where further HA maturation occurs, before travelling to the plasma membrane on endosomes (**Figure 2:9a-b**). Budding of the virus from the cell is believed to be initiated by M2 protein forming budding sites via membrane bending which encourages HA and NA localization (**Figure 2:10a**).[31] M1 is cytosolic and so attaches itself to the budding site on the interior side. Subsequently, vRNPs localize and bind to the M1 as budding occurs.[31] The virus then detaches itself from the cell via the action of NA cleaving the sialic acid that HA attaches to, leaving the virus free to travel through the airways to infect other cells (**Figure 2:10b**).[32,33]

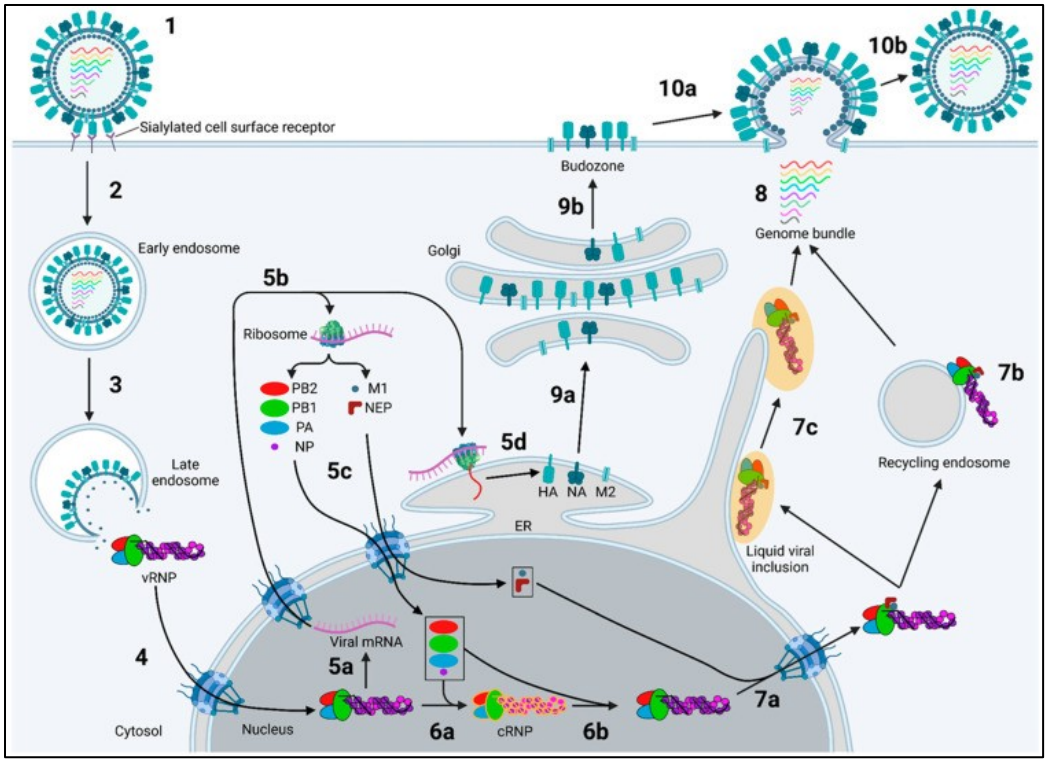


Figure 2: Diagram showing the replication cycle of IAV

(1,2) Virus enters the cell via endocytosis and **(3)** the vRNPs are released into the cytoplasm. **(4)** The vRNPs enter the nucleus where they undergo **(5a)** transcription and the resulting RNA is **(5b)** translated. Viral proteins involved in vRNP formation and export are **(5c)** re-imported to the nucleus and combine with **(6a)** newly-replicated vRNA to **(6b)** form new vRNPs. These are **(7a)** exported from the nucleus and traffic to the plasma membrane on **(7b)** recycling endosomes or **(7c)** condensates and undergo **(8)** bundling during this process. Membrane-bound viral proteins **(9a)** transit from the ER to the Golgi, **(9b)** then on to the plasma membrane where they **(10a)** induce budding. **(10b)** Viral particles are subsequently released from the cell. Figure obtained from Carter et al. Viruses 2024 [https://www.mdpi.com/1999-4915/16/2/316].

Biomolecular Condensates

Biomolecular condensates (BMCs) or liquid membraneless organelles are entities existing in biological systems that contain biomolecules such as proteins and nucleic acids and which lack an external delimiting phospholipid bilayer membrane.[34] Instead, they are formed and persist through a network of weak multivalent interactions (ability to undergo many weak interactions, stickiness) between their components which hold them in the form of a droplet separate from the environment in which they formed.[35] These interactions can involve dipole-dipole attraction, cation-anion bonding, hydrogen bonding, and hydrophobic interactions which are characterized as being weak and capable of being broken and reformed quickly.[34] Via these attractions protein-protein and solvent-solvent interactions are favored over protein-solvent interactions and if occurring under conditions that are energetically favorable result in the reaching of a lower free energy state coinciding with droplet formation.[34] The strength of these interactions determine the properties of the resulting condensate and the liquid nature of the BMC is the result of an interplay and competition between the attractive forces between the constituents of the condensate and the entropy of the system. Strong attraction drives the condensate towards a more solid aggregate state whereas weak ones maintain its liquid character or cause its dissolution completely. The liquid character referred to here is the ability of the internal components to undergo conformational, rotational, and translational changes.[36]

Liquid droplet formation from a complex solution is not limited to biological systems and generally occurs through the well-studied process of liquid-liquid phase separation (LLPS). It can be represented by a phase diagram in which certain conditions which increase interactions between dissolved components cause the separation of the liquid into two distinct phases, with

such a separation occurring in a range of physical contexts but not in others. In the case of BMCs the two separate liquids are the dense phase of the droplets themselves and the diffuse phase in which the components are still dissolved in the solvent.[34]

Under the currently favored stickers and spacers model of condensate structure particular proteins capable of undergoing multivalent interaction with several interacting regions are the drivers of BMC formation.[37] Therefore, the multivalency of a particular protein and its capacity to undergo context-dependent conformational changes is critical to its ability to condense.[38] The propensity of a protein to drive phase separation is highly dependent on the presence of these sticky intrinsically disordered regions (IDRs) within its structure. These IDRs are areas with low complexity of amino acids and no set structure, being instead flexible and able to bind in multiple different conformations.[39] IDR-rich proteins are generally known as scaffolds in the structure of a condensate. The scaffold proteins are those which drive the phase separation while the clients are proteins that have a lesser tendency to undergo phase separation and are recruited after the condensate is already formed.[40]

In the process of studying phase separation it has also been noted that regions rich in positively charged (basic) amino acids capable of binding negatively-charged DNA and RNA can also contribute to phase separation, in addition to contributions by the nucleic acids themselves.[41,42] This is likely related to the linear and flexible nature of these oligomers enhancing the multivalency of interacting proteins by connecting distant IDRs.[41] Positively charged regions of proteins have also been shown to result in localization to cellular membraneless organelles such as the nucleolus.[43] Post-transcriptional modifications to particular amino acids can also have an

effect on a protein's ability to phase separate. For example, arginine methylations can result in disruptions to cation-pi interactions while phosphorylation of serine-threonine repeats can have both positive and negative effects on LLPS depending on context.[44]

The nucleus, being a dense region with large amounts of protein and nucleic acids is a notably BMC-rich region of the cell. Within the nucleus a number of different membraneless organelles exist which contribute to processes such as replication, genome structuring, and transcription. These include the nucleolus, superenhancers, nuclear speckles, insulation loops, polycomb bodies, heterochromatin, Cajal bodies, Barr bodies, and Promyelocytic Leukemia bodies among others.[45,46] Some of these BMCs serve to regulate gene activity and are themselves regulated by it, travelling towards sites of active transcription and growing in size as the activity increases.[46]

The formation of BMCs can be achieved to carry out several different functions. It allows for the components to be highly concentrated, thus allowing the BMCs to serve as a catalytic environment for certain less energetically favorable interactions.[47,48] It can also protect the internal components from the external environment, for example in the case of virally-induced BMCs that can serve as a barrier between the virus' components and the cells' innate immune detectors such as RIG-like and toll-like receptors.[3] Finally, it has been found that condensates can also be used to detect a cell's external environment as they are highly dynamic and can rapidly shift in size, number, and density in response to changes in hypotonicity, salt concentration, and temperature.[49] The highly dynamic nature of these condensates is usually exemplified by their rapid movement, tendency to undergo fusion and fission events, and fast internal rearrangement as well as their ability to exchange components with their environment.[35]

However, some condensates may take on a more solid form with increased viscosity and elasticity or even develop irreversibly into amyloid aggregates as is the case for the condensates of prion protein implicated in human disease.[34]

The various characteristics of condensates must be proven to confirm that a particular phenomenon is indeed a BMC. Usually, a protein which is believed to drive the structure's formation is labelled fluorescently and its localization and movement within a cell or other environment is recorded. If the structure is round, moves rapidly, and undergoes fusion and fission events this is taken as evidence in favor of it being a condensate.[50] This is because it is demonstrated as being a liquid capable of mixing with like liquids and which has adopted an energetically favorable shape corresponding to phase separation from the solution. Another key characteristic of a condensate is the ability to dynamically exchange material both within itself and with its external environment. To determine this, florescence recovery after photobleaching (FRAP) experiments are conducted.[51] Broadly, FRAP involves the use of intense laser power to excite a particular region until the fluorophores in that area lose their ability to fluoresce (bleach). Following this, the intensity data for this region is recorded and plotted to determine the speed and extent of consequent florescence recovery which is indicative of the rate at which the fluorophore can diffuse into the bleached space.[52] By using FRAP on a portion of the structure being investigated internal rearrangement speed can be determined while performing FRAP on the entire structure allows for the speed of exchange with the dilute phase to be found. Certain chemicals can also be used to harden or dissolve BMCs and demonstrate the dynamism of a particular structure. For example, hexanediol is commonly used to dissolve condensates by disrupting the hydrophobic interactions that usually drive their formation.[53] Meanwhile, the

chemical nucleozin has been shown to be able to harden IAV vRNP condensates in the cytoplasm by increasing oligomerization of NP.[5] Condensates are also affected by salt concentration, which disrupts electrostatic interactions and temperature, which affects the movement of molecules and energetic favourability of certain states and varying these can provide further evidence for a structure being a condensate.[54-56]

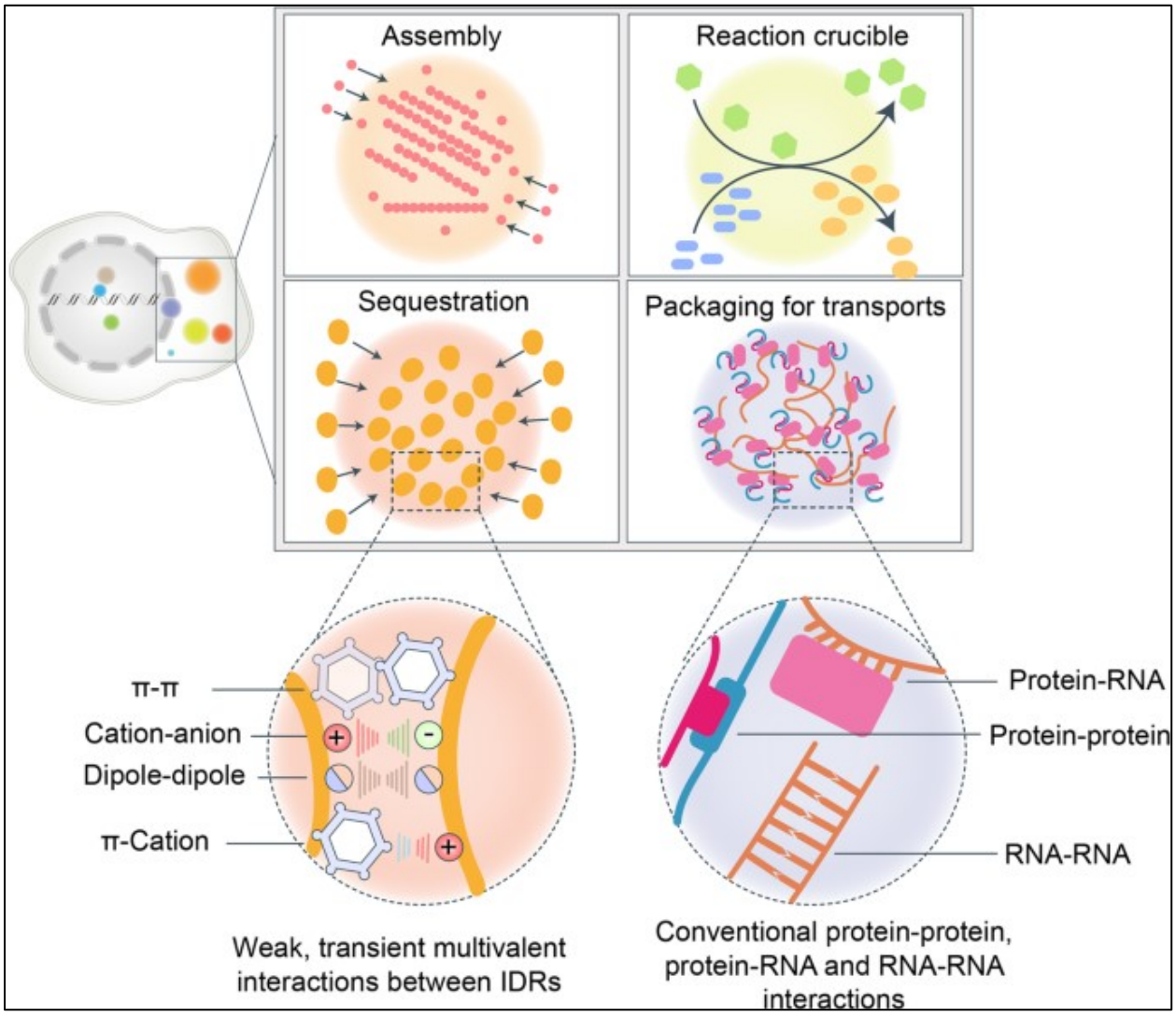


Figure 3: Functions and interactions of BMCs

Biomolecular condensates can serve purposes including assembly of complexes, creating correct conditions for certain reactions, sequestering components from the external environment, and packaging proteins for trafficking. Interactions between pi-bonds and charged sites and between protein and RNA are involved in the formation of BMCs. Figure was obtained from Wang et al. Signal Transduction and Targeted Therapy 2021 [https://www.nature.com/articles/s41392-021-00678-1].

Intersection Between Influenza A Virus and BMCs

There are several previous papers linking IAV biology to the formation of biomolecular condensates. It has been found that IAV genome segments decondense during viral uncoating and release into the cytoplasm by making use of host aggresome machinery, particularly HDAC6 as well as host import machinery such as TNPO1.[57] HDAC6 has been shown to promote LLPS under certain conditions, while TNPO1 serves as a suppressor of vRNP condensation following uncoating.[58] A further notable publication by Alenquer et al. in 2019 found that IAV vRNPs formed condensates near ER exit sites while transiting towards the plasma membrane, which likely aid assembly of the genome segments. These condensates were found to contain both vRNPs and Rab11a and to be highly dynamic, rapidly recovering in FRAP experiments and undergoing fusion and fission events. They also depend on ER-Golgi cycling and did not promote escape from the interferon response.[4] It was later found in another 2023 paper by the same group that these condensates can be hardened through the use of the anti-IAV drug

nucleozin, which encourages NP oligomerization. This paper also noted that changes in concentration of vRNPs and temperature had a lesser effect on the dynamics of these condensates than the alteration of the valency of their constituents via nucleozin.[59] The group then published a 2023 paper on methods to image the IAV condensates more efficiently, comparing different methods to obtain better estimates of the number and size of the observed condensates.[60] It has also been found by a different group that particular residues in an intrinsically disordered region of NP, including amino acids R74 and R75, are important for viral genome packaging and production of infectious progeny virus.[61] Finally, it has been noted that Kapβ2 protein involved in binding RNA-binding proteins' IDRs and transporting them to the nucleus, and which reverses condensate formation by causing disaggregation, is involved in IAV viral uncoating.[62] Little information is available on LLPS or BMCs in other viruses in the family Orthomyxoviridae to which IAV belongs such as the salmon-infecting Isovirus and the arthopod-borne Thogotovirus and Quaranjavirus.

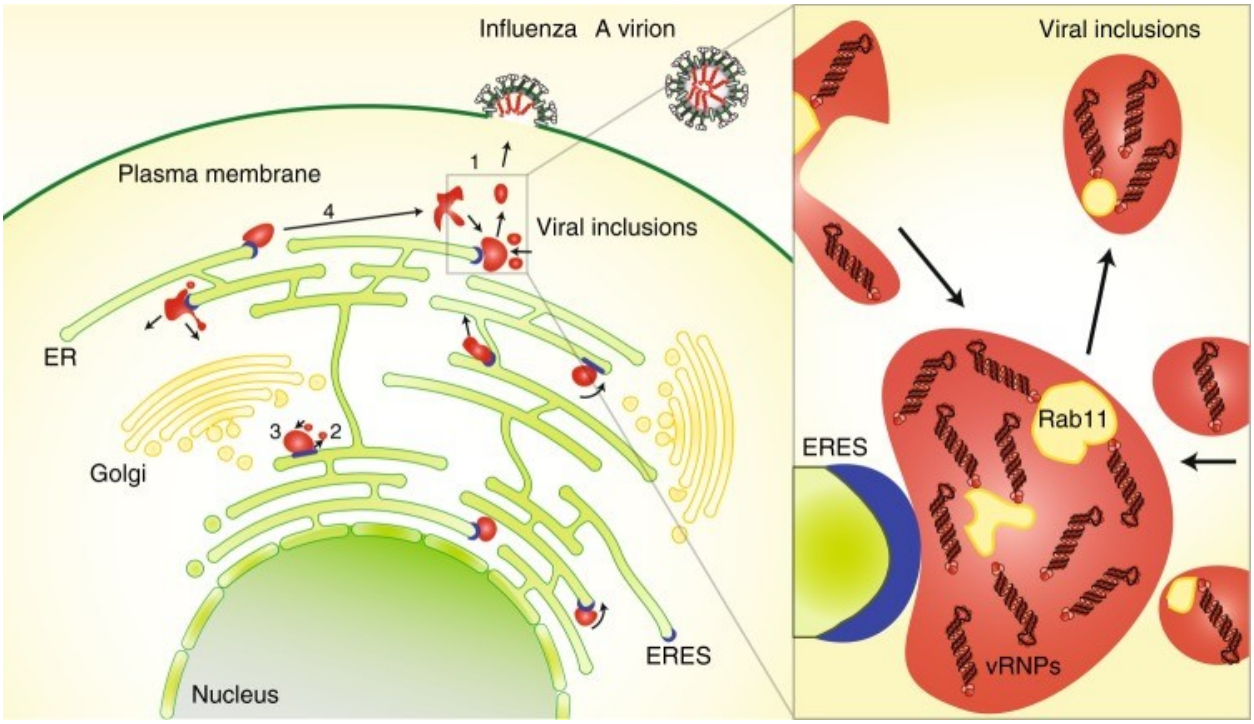


Figure 4: Diagram showing IAV BMCs in the cytoplasm

The vRNPs are transported from the nucleus to the plasma membrane via a modified ER network, with condensates containing Rab11 frequently being encountered in close proximity to ER exit sites (ERES). The Figure was obtained from Alenquer et al. Nature Communications 2019 [https://www.nature.com/articles/s41392-021-00678-1].

Methodology

There were several major techniques used in the completion of the thesis work. These included cell culture and subsequent transfection and infection as well as purification of IAV nucleoprotein and its mutants. Following purification *in vitro* observation of the protein took place. In addition, molecular biology techniques of cloning and PCR were needed to create many of the materials used in the project.

Cell culture

HeLa, 293T, and A549 cells were used over the course of this project. Culture conditions were standard, making use of Dulbecco's modified culture medium (DMEM) (Gibco, 11965-092) containing a penicillin-streptavidin mixture (10,000U penicillin, 10mg streptavidin) and fetal bovine serum (FBS) 10%. Cells were stored at 37ºC and 5% CO_2 in incubators. During transfections using lipofectamine and infection experiments cell media was changed for Opti-mem (Gibco, 01985-092) without FBS immediately prior to the start of the experiment.

Protein expression

Rosetta E.coli (NEB, C2527I) cells were transformed with the relevant pMAL plasmid (either pMAL-NP or pMAL-R74A-NP). These plasmids allow for inducible expression of NP N-terminally fused to maltose-binding protein with a linker region specific to factor Xa protease. Transformation was conducted using a standard technique. The competent bacteria were removed from storage at -80ºC and thawed on ice until liquid then 50µl of the bacteria was mixed with 10µl of the plasmid at 1ng/ul and kept on ice for 30 minutes. The bacteria were then exposed to 42.5ºC for 30 seconds before being put back on ice for 5 minutes. Outgrowth was allowed in 1ml LB or SOC (NEB, B9020S) media for 1 hour at 37ºC with shaking, after which the mix was spun down to 50µl and plated on ampicillin/chloramphenicol LB agar plates (100µg/ml each). A single colony was taken from these plates and grown in liquid LB culture (also treated with ampicillin and chloramphenicol at the same concentration) at 37 ºC in a volume of 5ml overnight. The entire volume was then added to 500ml of LB liquid (with ampicillin and chloramphenicol at the same concentration) for further growth at 37 ºC with shaking and subsequently induced with 0.5mM IPTG when the culture reached an OD600 of 0.5. The induced cells were left overnight at room temperature (measured as 23ºC) with shaking. Following this the cells were pelleted and lysed.

Lysis of bacterial cells

Base buffer was prepared containing 20µM Tris, 500mM NaCl and 5% Glycerol.
Bacterial lysis buffer was prepared by using the base buffer (20ml) and adding Lysozyme (1mg), DTT (1M stock) (20µl), RNAse A (10mg/ml stock) (20µl), DNAse 1 (500U stock) (20µl), PMSF (0.1M stock) (200µl), and complete ULTRA protease inhibitor (half regular tablet or 2x mini

tablet) (Roche, 11836170001). The bacterial lysis buffer was mixed with pelleted cells on ice, vortexed until the pellet was gone and sonicated twice for 2 minutes (pulsed 15s on, 25s off at 40% amplitude).

Protein Purification

The following buffers were prepared: Buffer A = Base buffer, Buffer B = Buffer A with 2000mM NaCl, Buffer C = Buffer A with 150mM NaCl, Buffer D = Buffer C with 10mM maltose

The following liquids were passed through an MBP-Trap Column (Cytiva, 29048641) at 3ml per minute in the order presented (CV=column volume, in this case 1ml): 5 CV H2O, 5 CV 0.5M NaOH, 5 CV H2O, 5 CV Buffer A. The following was then passed through at 0.625ml per minute: Bacterial lysate from previous step. After this the following was passed through at 3ml per minute: 30 CV Buffer A, 10 CV Buffer B, 10 CV Buffer C. The column was then eluted via syringe containing 5ml Buffer D and gathered in 0.5ml fractions. Fractions 2 and 3 contained the highest protein concentrations. After use 5 CV H2O, 5 CV 0.5M NaOH, and 5 CV 20% EtOH were passed through the column to clean and regenerate it.

Protein Processing

This part may still be streamlined by having only one digestion step of 3 days at RT with 1% Factor Xa (NEB, P8010S), which was found to result in very effective digestion during previous attempts. Western blots appear to indicate no adverse effects on NP following this procedure at room temperature. NP appears highly stable. Lower concentrations of Factor Xa may be used, but run the risk of not fully digesting one of the truncated NP products occurring at about 75kDa. Digestion reactions can be effectively stopped with 1mM PMSF. It should be noted that the size

of the Factor Xa isoform from NEB used here is 43kDa and can therefore be removed with a 50kDa MWCO ultrafilter. Not all factor Xa isoforms are this size.

Protein concentrations were determined following purification. This was done using a nanodrop device (Denovix DS-11) set to use a BSA standard as the exact MW and extinction coefficients of the MBP-NP products are not known. Once the protein concentrations were determined a portion (>1mg) was set aside to undergo digestion with Factor Xa. 0.1% (w/w) of factor Xa was added to the protein allowed it to digest for 48hrs at room temperature. This resulted in incomplete digestion of a truncated MBP-NP variant at around 75kDa (possibly due to partial factor Xa site obfuscation). 1mM PMSF was then added to stop the reaction and after a western blot which revealed this band it was ultrafiltered for 10min at 12000g with a 50kDa MWCO spin column (Amicom, 36100101) to remove smaller fragments and PMSF (though this is unlikely to have removed all PMSF the chemical loses potency as an inhibitor quickly in aqueous solution).[63] It was then digested again with 1% (w/w) Factor Xa for a further 24 hours at room temperature. The result after another round of ultracentrifugation using the 50kDa ultrafilter was a single band that appears from 48-56kDa on a membrane (depending seemingly on the extent of protein denaturation due to heat and SDS). This band can be detected using anti-NP. It should be noted that R74A is prone to not being detected by anti-NP but still appears via Ponceau on a membrane. Actions which can cause this lack of detection include repassage through the MBP-Trap and heating, though it can be overcome to an extent by simply loading more protein onto the gel. After WB the protein was diluted to 1ml/mg in PBS and was dialyzed twice for two hours against 50ml PBS using a Slide-A-Lyzer mini dialysis device with a 10kDa membrane

(Thermofisher, 88404) to prepare for labelling with AF488 and to ensure no lysis buffer remains which could affect labelling efficiency.

Protein Quality Control

Purified nucleoprotein concentration was determined using the previously mentioned nanodrop device set to use a MW of 56000 Da and an extinction coefficient (ε) of 55537 M−1·cm−1.[64] Protein was spun down using a short spin cycle of 16000g for 15 seconds prior to concentration measurement to remove aggregates. Once purified the protein underwent SDS-PAGE and Western blotting to ensure correct expression. Samples taken during each stage of the purification process were also run to monitor progress. Protein was stored at -80ºC following flash-freezing if unlabeled, or at 4ºC if labelled or in use for a maximum of 4 weeks.

SDS-PAGE and Western Blotting

About 10µg of protein at 1mg/ml was mixed with 3.33µl of 4x loading buffer then heated for 5 minutes at 95ºC to allow for denaturation and was subsequently loaded into a 10cm 10% SDS-PAGE gel and inserted into an electrophoresis unit (Biorad Mini-Protean, 1658004) and run at 120V until the dye reached the bottom of the gel. The gel was then either stained with coomassie blue or Bluesafe dye (Nzytech, MB15201), or underwent transfer to a nitrocellulose membrane using a transfer apparatus (Biorad Mini-Transblot, 1703930). The membrane was then subsequently stained with Ponceau reagent and underwent western blotting. If a western blot was performed it was blocked with 5% milk for 2 hours, then incubated with mouse primary anti-NP antibody (Abcam, ab128193) overnight at 4ºC with shaking, then incubated with secondary anti-mouse antibody linked to HRP overnight at 4ºC with shaking. Following this visualization

reagent (PerkinElmer Western Lightning ECL Pro) was added to allow visualization using the chemiluminescence setting on an imaging device (BioRad ChemiDoc) with auto-optimal settings. Images of the coomassie gel and ponceau were also taken on the Chemidoc using the visible spectrum settings. Where fluorescently tagged secondary antibodies were used (LICOR IRdye 680RD-tagged goat anti-mouse, 926-68070) imaging took place using an infrared scanner (LICOR Odessey 9120) using the 700nm channel.

Infection and Transfection Experiments (Lipofectamine)

$2.5*10^4$ A549 cells per well were seeded in 300µl DMEM on an 8-chambered imaging slide (Ibidi, µ-Slide 8 Well, 80824). Alternatively, $5*10^4$ A549 cells were seeded in 600µl DMEM (Gibco, 11965092) on a 4-chambered imaging slide (Lab-Tek, 155342). Cells were allowed to attach and replicate for 24 hours in an incubator at 37ºC and 5% CO2. Media was then replaced with 100µl or 200µl Opti-mem (Gibco, 11058021) and transfected with relevant plasmids using lipofectamine LTX (Invitrogen, 15338030) with plus reagent according to the manufacturer's protocol, leading to the addition of 100µl or 200µl more Opti-mem. Generally, transfection experiments were done with 50ng of fluorescent NP plasmid, 200ng non-fluorescent NP plasmid and optionally 250ng fibrillarin for the 8-chambered slides and twice as much for the 4-chambered slides with 1 or 2 µl Lipofectamine and 0.5 or 1 µl plus respectively. During infection experiments the plasmids encoding non-fluorescent NP were not included. If infection was not used then 100µl or 200µl more optimum was added to make up the volume to 300µl in the well. If infection was also used (only involving the smaller 8-chambered slide wells) PR8 virus (produced as described in de Wit et al. 2004.[65]) was diluted in 100µl Opti-mem to achieve an MOI of 10 and then added to the wells. Cells were observed 12, 16, or 20 hours following

transfection and/or infection. During some experiments, transfection and infection did not occur at the same time.

Transfection Experiments (Jetprime)

$2.5*10^4$ A549 per well were seeded in 300µl DMEM on an imaging plate (µ-Slide 8 Well, Ibidi). Cells were allowed to attach and replicate for 24 hours. The relevant plasmids were mixed with Jetprime and Jetprime buffer (Polyplus, 1010000 and 201000003) as appropriate for the amount of DNA used and the manufacturer's protocol. The mix was added to the wells. Cells were observed 16 hours following transfection.

Microscopy

Cells and in vitro condensates were observed using either a Zeiss LSM900 inverted confocal microscope with Airyscan 2 using a 63x oil objective, or a Zeiss LSM800 inverted confocal microscope using the 63x oil objective. A 488nm laser was used to excite mNeonGreen, GFP and AF488 fluorophores, a 561nm laser was used for mCherry and Cy3 fluorophores, and a 647nm laser was used for Cy5. Laser power and master gain settings were set to 1-2% and 500-700V respectively, as needed to ensure proper visibility for each of the fluorophores. A DIC channel was also included using the 488nm laser track. FRAP experiments were conducted using various settings. Generally, a 1uM circular area was chosen and bleached for 10-30 cycles with the laser used to excite the fluorophore of interest. Following this, the region was monitored for 30-120 seconds, recording the overall fluorescence intensity. The resulting data was normalized to the pre-bleach and immediate post-bleach intensities and plotted against time. Imaging under brightfield was also frequently conducted on both live cells and in vitro condensates using the track with the 488nm laser.

In vitro experiments

Nucleoprotein was labelled using a microscale AF488 protein labelling kit (Thermofisher, A30006) according to the manufacturer's instructions and protein was eluted in PBS. In the case of the preliminary *in vitro* experiment WT and R74A NP were instead labelled using Cy3 and Cy5 dye respectively following reduction with TCEP to reduce SH2 groups and overnight dialysis with PBS. Protein was taken from -80ºC or 4ºC (short-term) storage and centrifuged briefly (16000g for 15 seconds) to remove aggregates that formed during storage. Then, salt solution of the correct concentration using standard solutions of salt solutions of 0mM (H2O) and 1000mM [NaCl] was added to a microfuge tube followed by unlabelled and labelled protein in a 19:1 [unlabelled protein] to [dye in labelled protein] mix which was found to be approximately equivalent to 100:1 [unlabelled protein] to [labelled protein]. It should be noted that this approach based on fluorophore concentration means that the real protein concentrations are slightly lower than those written, in the following text 20µM is 19.2µM, 30µM is 28.3µM and 40µM is 38.4µM. A 50% PEG solution was subsequently added to achieve a 10% concentration in the final mix. The resulting droplet of 10µl total volume was added to a microscope slide, and a coverslip was gently deposited on top, allowed to set for 5 minutes (to prevent movement during sealing) and was sealed around the edges using clear nail polish, which was allowed to dry for a further 5 minutes. Imaging took place between 20-30 minutes after mixing depending on the number of sequential slides being imaged and how many images were taken. Calculations to achieve the correct calculations of protein and salt concentration were conducted using an Excel-based calculator which automatically gives the correct volumes to add given certain desired concentrations.

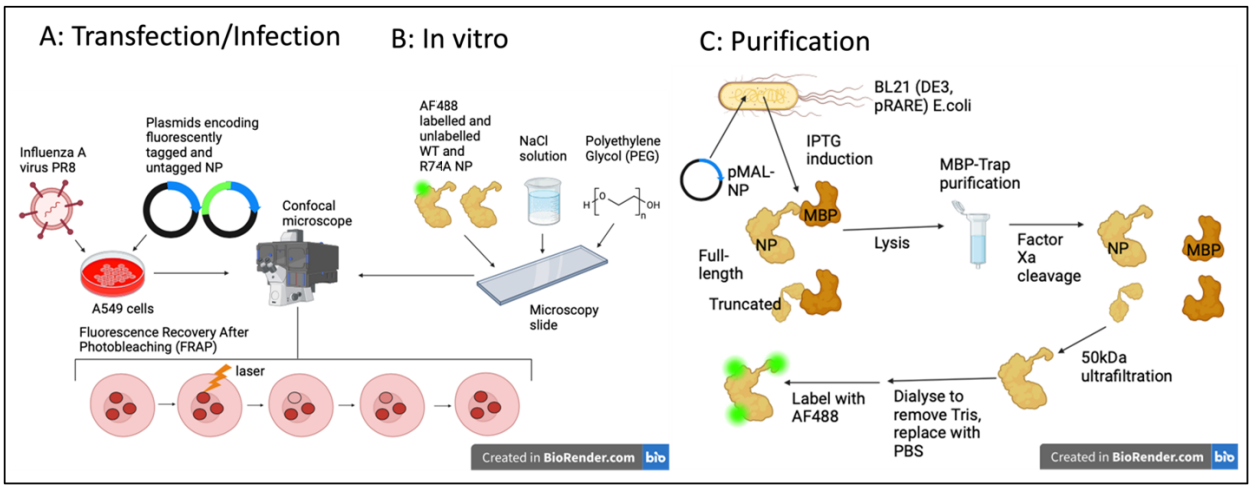


Figure 5: Diagram showing main methodologies used in this project

(A) Transfection of cells with plasmids encoding NP and infection with IAV PR8 virus followed by observation and FRAP experiments. **(B)** *In vitro* experiments using purified NP protein mixed with salt solution and PEG. The mix was added to a slide, then underwent imaging and FRAP. **(C)** Expression of NP via use of a pMAL-NP vector encoding NP N-terminally fused to maltose-binding protein (MBP) in E.coli BL21(DE3) bacteria with pRARE. This was followed by purification via MBP-Trap, tag cleavage with factor Xa, and ultracentrifugation with a 50kDa molecular weight cutoff filter to remove MBP and NP fragments. Labelling of portion of NP with AF488 also shown.

Plasmid Production

Cloning WT and DII-DIV NP into an mNeonGreen plasmid required only digestions and ligations. Briefly, pcDNA3-NP (derived from Invitrogen pcDNA3) plasmids were digested with EcoRI to remove the insert, then ligated (using Quick Ligase NEB, M2200S) into an mNeongreen-C2 (pMNG-C2) vector also digested with EcoRI following dephosphorylation of the vector with Antarctic phosphatase (NEB, M0289S). BamHI was used to confirm the directionality of the insert prior to sequencing. To clone DI-DIVNP into the Maltose binding

protein plasmid (pMAL-C2) PCR amplification of the NP with primers containing EcoRI and XbaI sites was performed using Phusion polymerase (M0530L, NEB). The resulting product was then run on a 1% agarose gel and purified using a gel extraction kit (Zymo Research, D4008). Following this, the amplicon was digested with EcoRI and XbaI and ligated into pMAL using Quick Ligase. The vector was also similarly digested, purified and dephosphorylated with Antarctic Phosphatase. Individual amino acid mutations in NP were introduced using primer mutagenesis and Phusion polymerase. MNG-NLS-DINP was produced by using a forward primer containing the NLS from SV40 and a reverse primer corresponding to the other end of the NP insert, containing EcoRI and XbaI sites respectively. The primers were used to amplify the DINP insert on the pcdna3-DINP plasmid and the resulting product was purified on a gel and digested then inserted into pMNG-C2 as before. This cloning was needed for DINP to be cloned into MNG-C2 specifically to allow the NP to be in the correct open reading frame as this deletion mutant had a single AA deletion 5' of the start site and 3' of the EcoRI site. The original version of this MNG-NLS-DINP plasmid had a frameshift mutation which had be repaired using PCR mutagenesis, with the result being the MNG-NLS-DINP (fix) plasmid. The MNG-R74A-NP plasmid was produced by cloning the R74A-NP insert from pcDNA3-R74A into MNG-NP after removing the insert from the former using EcoRI and ligating into the backbone of the latter after dephosphorylation with quick CIP (NEB, M0525) using T4 ligase (NEB, M0202).

Figure 6: Plasmids created for this project

The colors correspond to the type of backbone used for each plasmid. NP deletion shortened forms: DI = Deletion 1-31AA, DII = Deletion 63-103AA, DIII = Deletion 390-404AA, DIV = Deletion 463-481AA.

Research Findings

NP has four distinct droplet promoting regions and one intrinsically disordered region

Droplet promoting regions (DPRs) of IAV (A/PR/8/1934 H1N1) NP were predicted using Fuzdrop (University of Padova).[66-68] Four droplet promoting regions were found at amino acids 1-31, 63-103, 390-404 and 463-481. In this text these regions are referred to as DPR I-IV respectively. These regions are likely to promote phase separation of NP protein and contain many positively charged amino acids such as arginine and lysine which contribute to RNA bonding.[69] The algorithm which predicts these regions looks at short stretches of amino acids and compares them to parts of proteins known to undergo LLPS. It then assigns these amino acids a score which relates to the probability of it promoting droplet formation. If at least 5 consecutive amino acids have a high score that region is denoted as a DPR.[67] **(Figure 7)**

The intrinsically disordered regions of IAV (PR8) NP were also predicted by making use of PONDR (Molecular Kinetics, Inc) with the VLS2 algorithm. This searches for amino acid stretches of low structural complexity using a composite machine learning algorithm with known IDRs as its training set.[70] It was able to determine the presence of a large IDR at the C-terminus of NP. **(Figure 7)**

Later research was focused on the droplet promoting regions. This is because we had an interest in looking at what happens when these regions are deleted and use of the DPRs enabled the analysis of several smaller deletion mutants rather than an NP protein lacking one very large region as would be the case if we had focused on the singular IDR. It was also thought that such a large deletion in one part of the protein would significantly alter the structure and be less

informative than several smaller deletions. Following the obtaining of these predicted results we determined that the idea that NP forms condensates had at least some merit and decided to test whether any condensates would appear in cells transfected with NP plasmids.

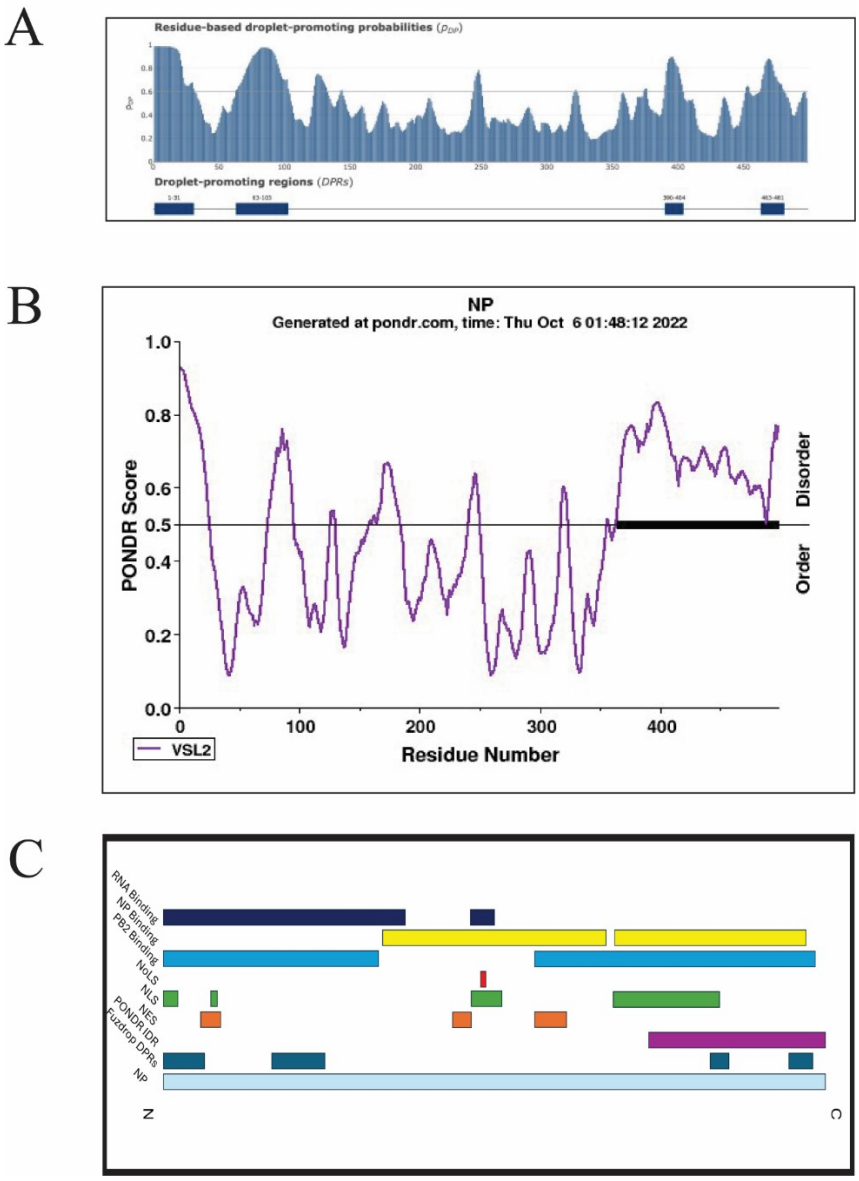


Figure 7: Regions of NP

(A) Droplet promoting regions of IAV NP according to Fuzdrop. The amino acid sequence of NP for strain A/Puerto Rico/8/1934 H1N1 was the input. **(B)** Intrinsically disordered regions of NP according to PONDR. The amino acid sequence of NP for strain A/Puerto Rico/8/1934 H1N1 was the input. **(C)** Diagram showing major functional regions of influenza A virus NP and the identified DPRs and IDR. Sources: Sobrido et al. 2018, Li et al. 2015, Ozawa et al. 2007.[71-73]

NP Forms Viscoelastic Structures in Cell Nuclei

In order to find whether NP formed biomolecular condensates it was first necessary to conduct simple observations of the behavior of the protein within cells. These experiments sought to find whether any round punctate structures appeared in nuclei when NP was introduced to cells. Hela cells were transfected with plasmids encoding GFP-NP and NP in a 1:4 ratio and observed using florescence microscopy. Cells were also transfected with plasmids encoding GFP-NP only, GFP and NP in a 1:4 ratio or GFP only as controls. These were performed in triplicate, with the same visual appearance of cells under florescence microscopy each time. It was found that the GFP-NP and NP and GFP-NP transfected cells contained nuclear puncta highly localized to the nucleus. The puncta moved and appeared to undergo fusion and fission events, though these events were difficult to confirm. The GFP and NP and GFP controls contained no puncta, and both appeared as would be expected for a cell transfected with GFP with signal present throughout the cell and no observable condensates. To show that formation of puncta was not contingent on GFP fusion to NP HeLa cells were transfected with plasmids encoding NP, fixed and stained with anti-NP primary and Alexa 488-bound anti-mouse secondary antibody. This was also done in triplicate. These immunofluorescent samples demonstrate the appearance of puncta formed of NP. In addition, plasmids encoding the various components of the vRNP (NP, PA, PB1, PB2, vRNA) were transfected into A549 cells along with plasmids encoding GFP-NP (in a 1:4 ratio with NP plasmid) in different combinations. This was done to find if there was some combination of components necessary for the formation of condensates or if certain components resulted in significant changes to the signal localization. Overall, no differences between the conditions were noted, with all showing mostly GFP-NP localization to the nucleus and nucleolus with areas of dense concentration present within the nucleus more generally. The data

overall appeared to demonstrate that NP sometimes formed small rounded structures in the nucleus that may be biomolecular condensates. (Data preliminary, not shown in Figure)

We determined that it was necessary to see what would occur given the conditions of a full viral infection in a relevant cell type and to perform more robust transfection experiments under varied conditions.. We used A549 human lung epithelial cells for more physiological relevance and mCherry-NP and newly produced MNG-NP plasmids to further exclude the possibility that the NP only formed structures due to eGFP's dimerization capability.[74]

Different time points were tested for both transfection and infection experiments. They were chosen as 4-8 hours, 16-20 hours, and 20-24 hours post-infection (early, middle and late timepoint respectively), with post-transfection time being 12 (or 16 for post-infection images) hours for the early timepoint, 16 hours for the middle timepoint and 20 hours for the late timepoint. Pre- and post- infection data for the same set of cells is available for the early timepoint for all conditions. These times were chosen because vRNPs first leave the nucleus to produce virus after around 5 hours and after 24 hours the cells generally appear dead and in very poor condition unsuitable for analysis.[75]

The transfection experiments revealed the presence of three distinct phenotypes of NP-transfected cells that were termed complex, nucleolar, and diffuse. Complex phenotype cells had NP localized to nucleoli and to large nuclear structures with irregular shapes while nucleolar phenotype cells had NP predominantly localized to the nucleoli and diffuse phenotype cells had NP localized to the nucleus generally. It is not known if these three localization types are

dependent on extent of NP expression in the cell or due to other cell-specific factors, though diffuse cells are less bright than nucleolar cells, which are in turn less bright than complex cells.

MNG-NP transfected cells tended to show a greater proportion of complex phenotypes than other fluorophores while mCherry-NP, the least bright protein, had a relatively small number of complex-type cells. During infection, at early timepoints the cells appeared very similar as during transfection, with primarily nuclear localization (diffuse and nucleolar) but at later timepoints would have a far lower (or even entirely lacking) proportion of nucleolar phenotype cells and a greater number of complex phenotype cells. These intermediate and late timepoint cells also had the previously described cytoplasmic condensates.

Complex phenotype cells were the most interesting to investigate further, as they were defined as having irregular rounded regions of dense NP concentration appear in the nucleoplasm which had the potential to be condensates. They were not reminiscent of known liquid BMCs formed though LLPS, which were generally round in shape and move rapidly. These regions were not observed to move or undergo rapid fusion or fission. To test whether the regions may be viscoelastic condensates with a less liquid nature, FRAP was performed. This revealed partial recovery (0.4 relative to initial fluorescence) over a relatively short timeframe (30s). The dynamics of NP in these regions did not vary much with timepoint for transfection or infection experiments and was similar to the recovery of NP in the nucleolus **(Figure 9).** Smaller nuclear structures with some movement capacity are sometimes observed this is only in a small minority of cells, and may just be NP localizing to various nuclear structures as has been previously found.[76] **(Figures 8,9)**

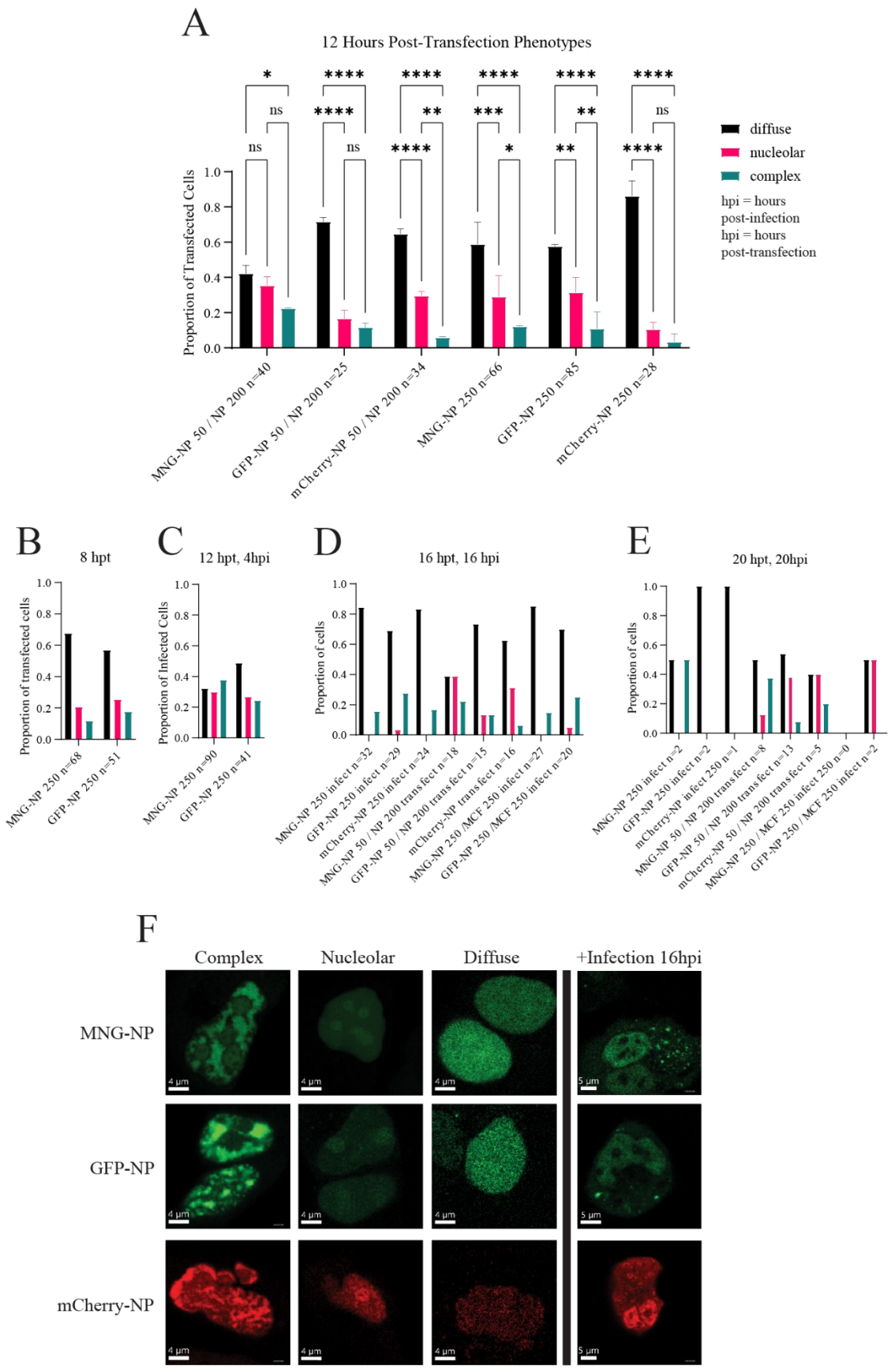
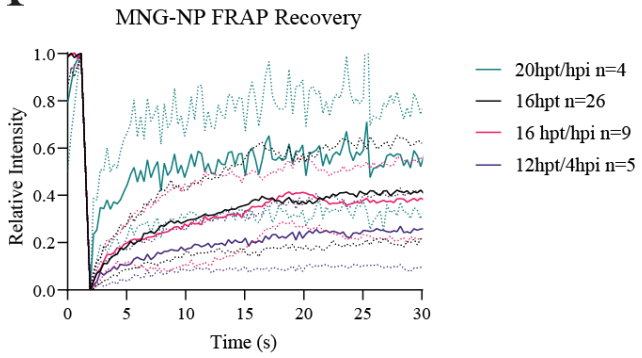

A
12 Hours Post-Transfection Phenotypes
Proportion of Transfected Cells
diffuse
nucleolar
complex
hpi = hours post-infection
hpi = hours post-transfection
MNG-NP 50 / NP 200 n=40
GFP-NP 50 / NP 200 n=25
mCherry-NP 50 / NP 200 n=34
MNG-NP 250 n=66
GFP-NP 250 n=85
mCherry-NP 250 n=28
B
8 hpt
Proportion of transfected cells
MNG-NP 250 n=68
GFP-NP 250 n=51
C
12 hpt, 4hpi
Proportion of Infected Cells
MNG-NP 250 n=90
GFP-NP 250 n=41
D
16 hpt, 16 hpi
Proportion of cells
MNG-NP 250 infect n=32
GFP-NP 250 infect n=29
mCherry-NP 250 infect n=24
MNG-NP 50 / NP 200 transfect n=18
GFP-NP 50 / NP 200 transfect n=15
mCherry-NP transfect n=16
MNG-NP 250 /MCF 250 infect n=27
GFP-NP 250 /MCF 250 infect n=20
E
20 hpt, 20hpi
Proportion of cells
MNG-NP 250 infect n=2
GFP-NP 250 infect n=2
mCherry-NP infect 250 n=1
MNG-NP 50 / NP 200 transfect n=8
GFP-NP 50 / NP 200 transfect n=13
mCherry-NP 50 / NP 200 transfect n=5
MNG-NP 250 / MCF 250 infect 250 n=0
GFP-NP 250 / MCF 250 infect n=2
F
Complex
Nucleolar
Diffuse
+Infection 16hpi
MNG-NP
GFP-NP
mCherry-NP

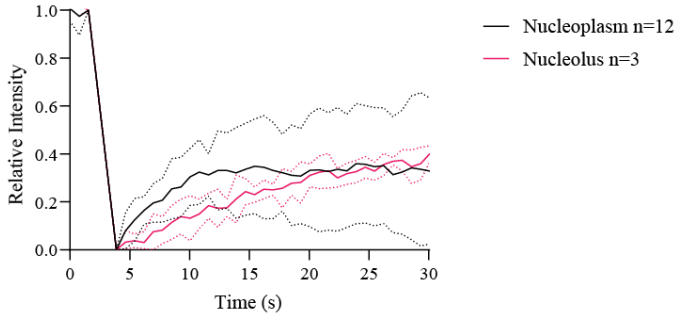
Figure 8: There is phenotypic and time-dependent variation in cells transfected with different fluorescently bound forms of NP, affected by infection.

A549 cells were transfected with the indicated plasmids and/or infected and imaged 12-20 hours later. The cells were then categorized as having the NP localized in a diffuse, nucleolar, or complex manner. Values listed on the category labels correspond to the ng of DNA used of each plasmid. The n number corresponds to the number of cells observed. Data presented in panel A from replicate experiments. Others from experiments performed once. MCF is mCherry-Fibrillarin. **(A)** Phenotypes of transfected cells at 12-16 hours post-transfection. **(B)** Phenotypes of transfected cells at 8-12 hours post-transfection. **(C)** Phenotypes of infected and transfected cells at 12-16 hours post-infection (16-20 hours post-transfection). **(D)** Phenotypes of cells at 16-20 hours post-transfection and infection. **(E)** Phenotypes of cells at 20-24 hours post-transfection and infection **(F).** Representative images showing appearance of main phenotypes observed and infected cell appearance.

A


MNG-NP FRAP Recovery
Relative Intensity
Time (s)
20hpt/hpi n=4
16hpt n=26
16 hpt/hpi n=9
12hpt/4hpi n=5

B


MNG-NP FRAP Recovery (16hpt)
Relative Intensity
Time (s)
Nucleoplasm n=12
Nucleolus n=3

C

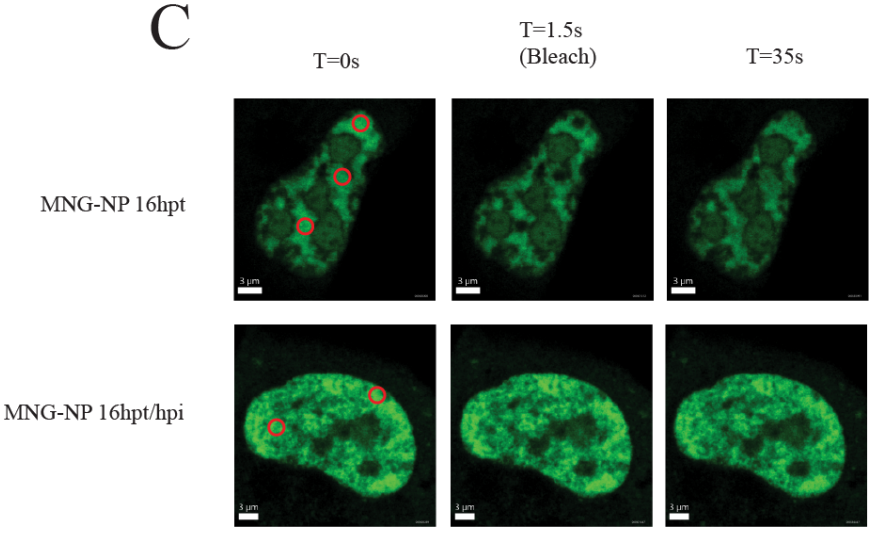

T=0s
T=1.5s
(Bleach)
T=35s
MNG-NP 16hpt
MNG-NP 16hpt/hpi

Figure 9: NP localization and dynamics change during course of infection. NP dynamics are localization-specific

A549 cells were transfected with plasmids encoding mNeongreen-NP (MNG- NP) and WT NP (in a 1:4 ratio) (50ng/200ng for A and 200ng/800ng for B) to analyze dynamics of MNG-NP. Cells were optionally transfected with MOI 10 PR8 IAV. FRAP was performed on MNG-NP regions of dense nuclear concentration or nucleolar areas 1 micron in diameter. Data was normalized so that1= immediate pre-bleaching intensity, 0=immediate post-bleaching intensity. Area bleached was monitored for 30-60s afterwards. The n number denotes the number of cells the FRAP averages are derived from. Relative fluorescence intensities from different cells or regions were averaged at each time point to obtain graphs. Experiment relating to panel B was conducted in duplicate, examining a total of 12 cells per condition. In experiment B graphs are adjusted for background and photobleaching using Nikon AX software. Dotted lines on FRAP charts are standard deviation. **(A)** Graph showing average recovery at different timepoints of infection and transfection. **(B)** Graph showing average recovery of MNG-NP transfected cells in the nucleoplasm (regions of dense concentration) and nucleolus. **(C)** Images are from representative experiments for the 16hpt condition and the 16hpt/hpi condition. The red circles denote the regions of interest that were photobleached and monitored.

NP DPR Deletion Results in Localization Differences and Abrogation of Nuclear Puncta Formation

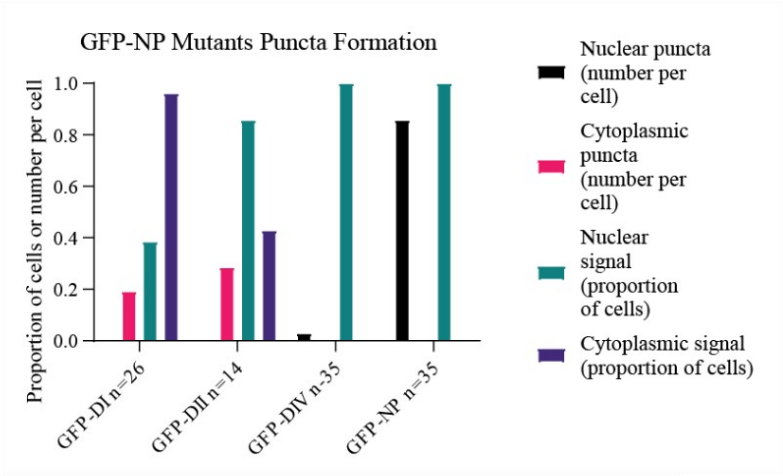
In order to investigate the effect of the identified droplet-promoting regions on the propensity of the NP to form nuclear condensates sets of plasmids were generated encoding both non-fluorescently bound forms of NP with each of the regions deleted along with plasmids encoding mNeonGreen and GFP-fusion forms of the NP deletion mutants. These NP deletion mutants are referred to as DI-IV depending on the region DPRI-IV deleted. It was not possible to generate the GFP-DIIINP mutant or investigate the MNG-NLS-DI-NP mutant.

Experiments were performed in which GFP and MNG-fused forms of NP DPR deletion mutants were transfected into cells, along with their non-fluorescent counterparts in a 1:4 ratio. These revealed that the DI mutant does not enter the nucleus, which is consistent with a lack of the N-terminal NLS. The DII mutant failed to enter the nucleoli, it is not known why as there is no known nucleolar localization sequence there, though the deletion is within the RNA binding region. The GFP-fused form of DIINP localized more to the nucleus than the cytoplasm, while the MNG-fused form tended to localize to the cytoplasm more than the nucleus. DIII and DIV of both fluorophores entered the nuclei and nucleoli but do not form the complex phenotype noted in WT, which involves certain dense localizations outside the nucleoli.

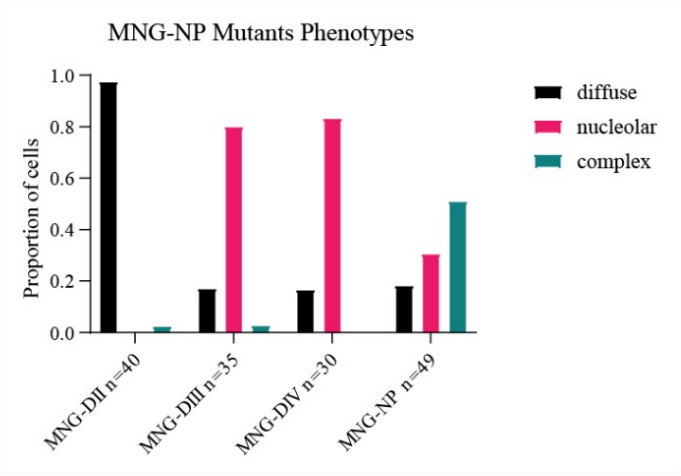
It was noted that for all deletion mutants of both GFP- and MNG-fused variants that nuclear puncta were either extremely rare or entirely absent and that phenotypes were generally far more consistent than the WT NP, with extremely few complex phenotype cells present. Instead, the mutants generally showed cell-to-cell variation relating to the amount of fluorescent signal in the nucleus versus the cytoplasm. Interestingly, as the deletions of GFP-NP go towards the C-

terminus of NP the signal tends to be increasingly localized to the nucleus, with the DPRIV mutant having all cells with primarily nuclear signal. Meanwhile, the DI mutant did not appear to localize to the nucleoplasm as frequently, with only about 40% of cells being classed as having nuclear signal. The DII mutant localized primarily to the nucleus and DIII mutants localized similarly to DIV, though with increased cytoplasmic signal. Despite the lack of nuclear puncta or dense nuclear regions cytoplasmic puncta were observed in up to 30% of observed cells depending on the region deleted. **(Figure 10)**

A


GFP-NP Mutants Puncta Formation
Proportion of cells or number per cell
1.0
0.8
0.6
0.4
0.2
0.0
GFP-DI n=26
GFP-DII n=14
GFP-DIV n-35
GFP-NP n=35
Nuclear puncta (number per cell)
Cytoplasmic puncta (number per cell)
Nuclear signal (proportion of cells)
Cytoplasmic signal (proportion of cells)

B


MNG-NP Mutants Phenotypes
Proportion of cells
1.0
0.8
0.6
0.4
0.2
0.0
MNG-DII n=40
MNG-DIII n=35
MNG-DIV n=30
MNG-NP n=49
diffuse
nucleolar
complex

C

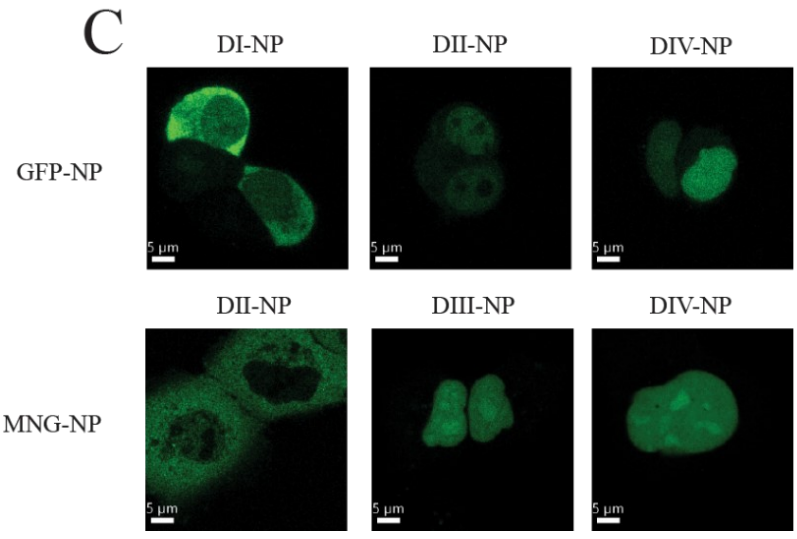

DI-NP
DII-NP
DIV-NP
GFP-NP
5 µm
DII-NP
DIII-NP
DIV-NP
MNG-NP
5 µm

Figure 10: NP DPR deletion results in localization differences

A549 cells were transfected with plasmids encoding fluorescently fused forms of NP DPR deletion mutants as well as plasmids encoding the NP deletion mutants in a 1:4 ratio and observed 16 hours later. N numbers on graphs correspond to the number of cells observed. Data is from one experiment. **(A)** Deletions of NP DPRs result in terminus proximity-dependent signal localization variation. Graph showing phenotypic proportions among the GFP-NP deletion mutants. Nuclear and cytoplasmic signal refers to whether there is signal present in the nucleus and/or cytoplasm. Note the changes as the deletions go from the N- to the C-terminus of NP. **(B)** Deletions of NP DPRs result in phenotypic changes. Graph showing phenotypic proportions among the MNG-NP deletion mutants using the same definitions as for the transfection and infection experiments (see previous section). **(C)** Appearance of NP deletion mutants.

R74A mutant nucleoprotein has altered phase separation properties

The lack of nuclear condensates in any of the cells transfected with the deletion mutants and the poor display of puncta in WT urged us to change directions. Since all DPR mutants seemed to disrupt condensate formation equally, it was decided to investigate individual amino acid mutants in the hopes of finding one that would cause a notable difference in condensate formation alone. The R74A mutation in DPRII was chosen as an amino acid of interest due to its highly conserved positively-charged nature (>99% of human H1N1 IAV sequences in the NIH flu database have this residue, the others have a lysine in its place) and literature supporting its importance to viral viability.[61] This mutation replaces a positively-charged arginine with a neutral charge alanine. A paper indicated that vRNPs containing R74A NP have similar polymerase activity to WT NP via mini-replicon assay but when present in the context of a virus do not result in subsequent production of infectious virus or efficient genome packaging.[61] It was thought that differences in phase separation and condensate formation might be responsible. Mutant R74A was produced alongside WT NP using a bacterial expression system as detailed in the methods section.

Preliminary purified NP *in vitro* experiments were performed prior to the completion of the main series of experiments. These tests involved the use of R75A and WT NP that did not have its maltose-binding protein purification tag removed. In addition, pure full length MBP-NP (ca.98kDa) was not separated from the mixture of truncated fragments acquired following expression. The bacteria tended to produce MBP-NP with C-terminal truncations, with the extent of production of truncated products varying between colonies.[61]

Purified WT and R75A NP were adjusted in salt concentration, mixed with PEG-8000 (10%) and plated on glass slides, with a small silicone gasket and glass slide added on top to enclose the resulting droplet (3ul volume). Segment 5 (NP-encoding) viral genomic RNA was also optionally added at a concentration of 33ng/µl to observe its effects on the protein's condensation. Protein was also labelled with either Cy5 (WT) or Cy3 (R75A) and added to the unlabeled NP in a 1:30 ratio. The temperature was maintained at 37°C during this experiment. Under all conditions tested WT NP appeared as small dense aggregate structures which did not readily recover from FRAP or display any dynamic activities. Meanwhile, R75A NP appeared as larger viscoelastic structures which were largest in size at a physiological salt concentration of 150mM, decreasing in size if salt is higher or lower. In the presence of vRNA the apparent condensation of R75A is decreased and the structures are less round. When FRAP experiments were performed, the condensates did not recover and no fusion or fission events were observed. Sometimes, when taking video of these condensates it was possible to see some settle on the glass slide, maintaining their form as they did so, similar to a gel sphere. It was believed that this data indicated that R75A NP was prone to phase separation while WT NP tended to aggregate, though it was necessary to perform further experiments (data not shown, preliminary).

Later, a new purification procedure for NP including the removal of the MBP tag was performed in order to obtain R74A and WT NP. Multiple attempts were needed to achieve NP of the correct size and WB results as at first the bacteria used to produce the NP did not contain the pRARE plasmid necessary to translate certain rare eukaryotic codons found in NP. After obtaining fresh stock, treating the media with chloramphenicol, and improving the protein processing steps better results were obtained.

When examining the condensates there appeared to be no difference in the shape of the WT and R74A NP puncta, with both appearing as round droplets as would be expected, this is especially evident under DIC. Images of the AF488-labelled NP condensates tend to have significant background with occasional darker regions, since these regions do not align closely with visible DIC structures they are not concluded to be evidence of a gel layer forming even at the highest concentration. NP was examined at different protein and salt concentrations centered on the physiological salt concentration of 150 mM [NaCl] found in the nucleus.[77] It was found that WT NP appeared to have no significant changes in condensate size depending on either salt or protein concentration, though sizes did generally appear larger at lower salt, there was too great a variability between experiments to denote significance. Meanwhile, the R74A NP had the greatest condensate sizes at the lowest 50mM salt concentration with decreased sizes at [NaCl] higher than 100mM. One would expect both WT and R74A NP to undergo greater phase separation at low salt since NP tends to oligomerize at high salt while maintaining itself as a monomer at low salt.[78] (**Figure 11**)

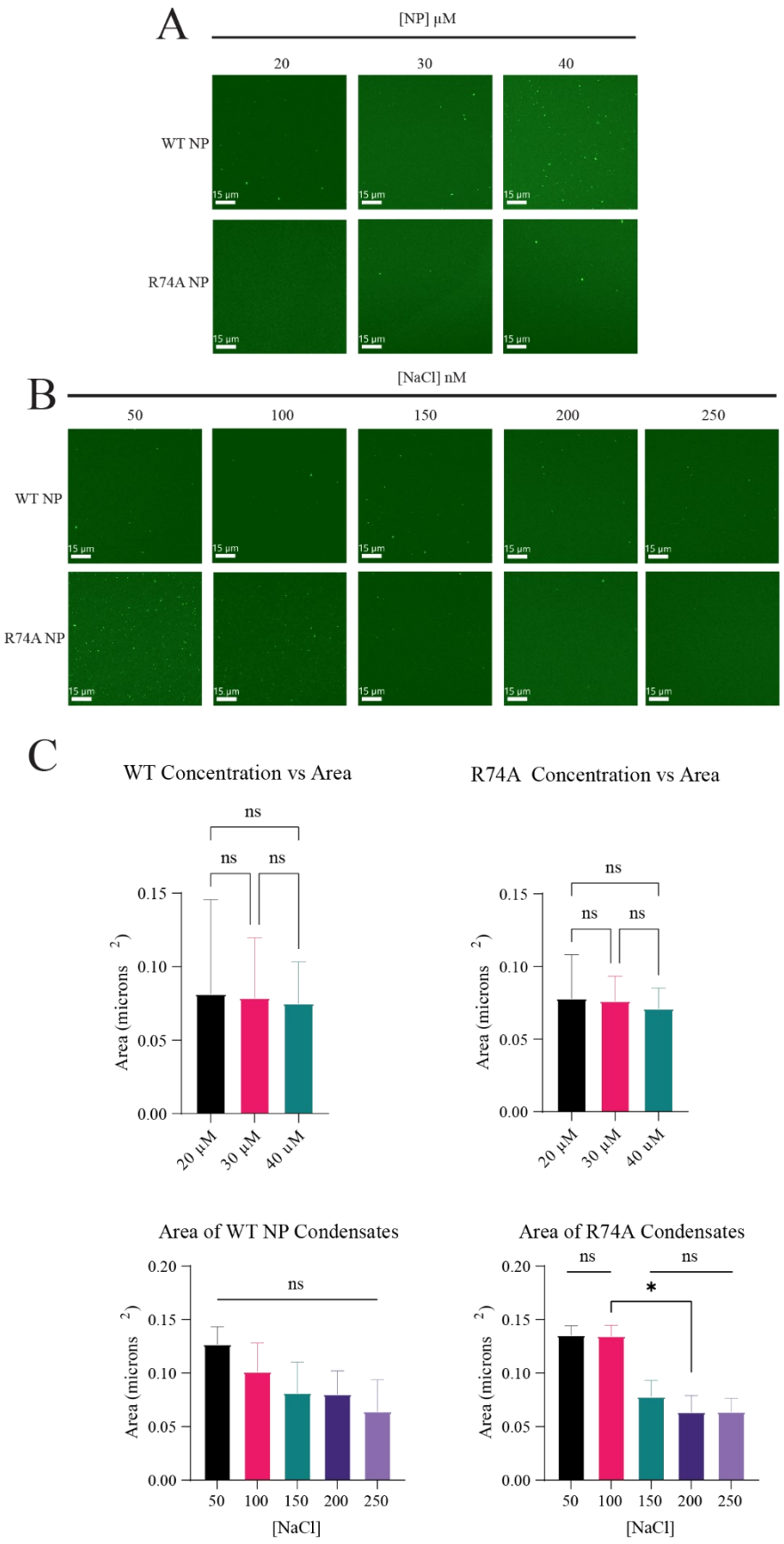

A
[NP] µM
20
30
40
WT NP
R74A NP
15 µm
B
[NaCl] nM
50
100
150
200
250
WT NP
R74A NP
15 µm
C
WT Concentration vs Area
R74A Concentration vs Area
ns
Area (microns 2)
0.00
0.05
0.10
0.15
20 µM
30 µM
40 uM
Area of WT NP Condensates
Area of R74A Condensates
0.20
*
50
100
150
200
250
[NaCl]

Figure 11: Mutant R74A NP has altered phase separation properties in vitro

Condensates in the images are small and should be inspected closely**.** Experiments were conducted in triplicate. **(A)** Purified WT and R74A NP protein (with AF488-labelled variants) was prepared at varying concentrations at 150nM NaCl **(B)** or kept at 20μM and adjusted to varying [NaCl] concentrations. Mixtures were added to a slide with 10% PEG and subsequently sealed with a coverslip and observed 20 minutes later. Three repeats were performed per condition. Images were brightened for visibility of condensates. **(C)** Graphs showing average condensate area under the conditions shown. Significance calculated with one-way ANOVA with $p<0.05$.

In addition to observing the condensates that formed in vitro, the labelling of a portion of the NP with AF488 enabled FRAP experiments to be performed. In these experiments the entire droplet was bleached with the recovery measured for up to 30s afterwards. The R74A NP did not recover as much as the WT NP, which tended to do so up to 0.4 relative to initial fluorescence compared to 0.2 for the R74A NP **(Figure S1).**

In addition to these protein and salt concentration series, the effect of 10% hexanediol on the NP condensates was tested. This was done to provide additional evidence for the structures observed being LLPS condensates, which should dissolve in the presence of hexanediol. However, the NP structures failed to dissolve, and instead increased in size (see discussion page 69). In addition, experiments were conducted where hexanediol was added prior to PEG and after PEG to see whether it functioned by penetrating existing condensates or by affecting their formation, with no significant difference in area of condensates observed between the two conditions. **(Figure S2)**

In order to ensure that the hexanediol we had in stock was still functional and could dissolve condensates it was also tested on HIV-1 Gag mixed with Gag labelled with AF647, which has been shown to form condensates under similar conditions as NP.[79] The hexanediol worked as expected and caused a decrease in the area of the condensates. **(Figure S3)**

Hexanediol experiments were performed to further study the effect of the chemical on both WT and R74A NP condensates. In order to discount the possibility that the change in condensate size observed with the WT and R74A NP was due to simply due to dilution caused by the addition of liquid to the final mix mock experiments were performed in which water was added instead. This

resulted in no change in the area of the condensates for either WT or R74A NP **(Figure 12).** In addition, two imaging timepoints were taken for the WT and R74A condensates without hexanediol in order to see whether they changed in size or number over time, with no significant difference **(Figure 13).**

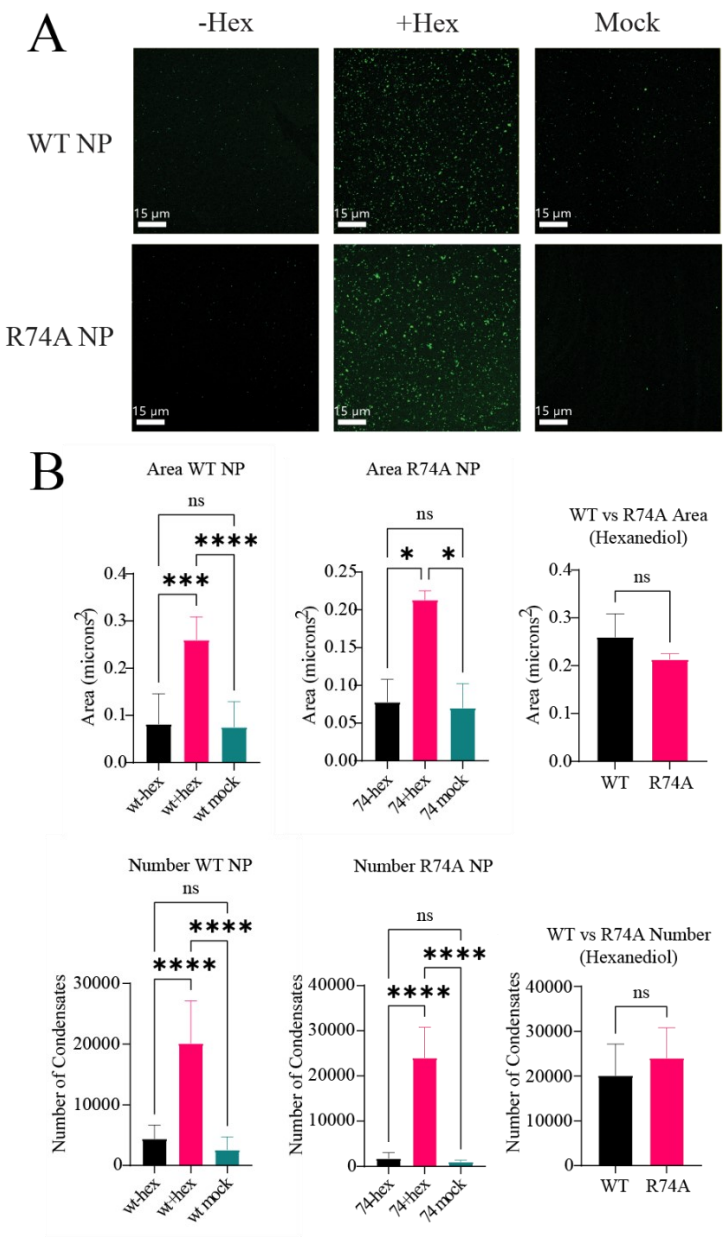


Figure 12: Hexanediol increases WT and R74A NP condensate size and number

WT NP and R74A NP (with AF488-labelled variants) were plated in concentrations of 20µM for the NPs in 150mM NaCl and 10% PEG**.** Experiments were conducted in triplicate. **(A)** In the indicated images, hexanediol was added to the mix to achieve a concentration of 10% (+Hex) or an equivalent volume of water was added (mock). **(B)** Representative graph of NP condensate areas and numbers (total of 10 images of 100x100 microns) under each condition.

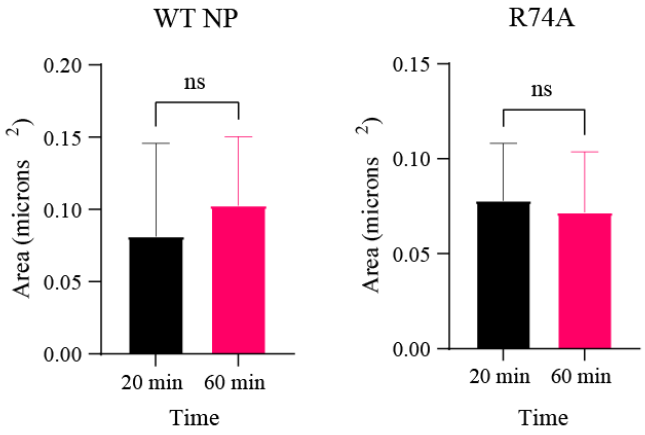


Figure 13: WT and R74A NP condensates do not change in area over time

WT NP or R74A NP (with AF488-labelled variants) were plated in concentrations of 20μM in 150mM NaCl and 10% PEG. They were imaged 20 and 60 minutes after formation (considered the time when PEG was added). Experiments were conducted in triplicate.

After this, further live cell experiments were conducted in which MNG-NP and MNG-R75A-NP along with plasmid encoding their unlabeled variants (in a 1:4 ratio) were transfected into A549 cells in order to see whether there would be alteration of localization or dynamics within cells. According to the literature, the R75A mutant is almost identical to the R74A mutant in terms of conservation, effect on polymerase activity and production of infectious viral particles.[61,80] There was no difference in regard to nuclear NP region size, MFI or FRAP dynamics between the R75A and WT NP. The complex phenotypes, defined as having areas of dense nuclear concentration were generally rare and were very bright while nucleolar and diffuse phenotypes also appeared, though the diffuse phenotype is likely simply nucleolar phenotype with low protein expression. The FRAP dynamics of the nuclear regions of complex phenotype cells was interestingly very similar to the results previously obtained at the other laboratory, with relative fluorescence intensity once again reaching about 0.4 after 25 seconds post-bleach **(Figure 14).**

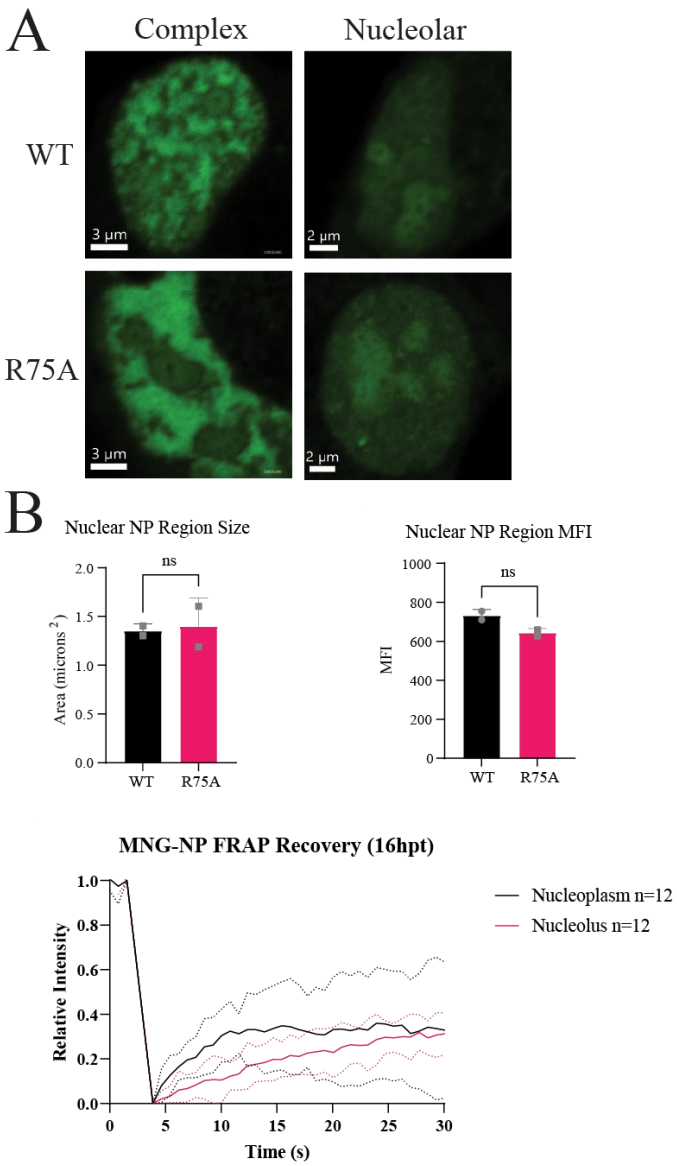


Figure 14: WT and R75A NP behave similarly in cell nuclei

A549 cells were transfected with plasmids encoding WT or mutant NP (200ng MNG-NP or MNG-R74A-NP and 800ng NP or R75A-NP). They were observed 16 hours later and underwent FRAP on regions of dense nuclear localization using regions 1uM in diameter. Experiments were conducted in duplicate, with data shown from 12 cells total. **(A)** Graphs showing average nuclear region area and MFI of complex phenotype cells. Significance calculated using unpaired T-test with $p<0.05$. **(B)** Average FRAP chart for each protein adjusted for background and photobleaching via reference regions on Nikon AX software. Dotted lines show SD.

Discussion

It was hypothesized that IAV NP can form biomolecular condensates in nuclei, potentially contributing to vRNP activity. To address this, NP was introduced to cell nuclei and observed under different conditions of transfection and infection, with an emphasis on finding condensates in the nuclei of a physiologically relevant cell line. The localization of mutant forms of NP with deletions of droplet-promoting regions were investigated in cells to find whether any of these would disrupt condensate formation. NP was also purified, along with a R74A mutant to be observed under more carefully controlled *in vitro* conditions. Its response to hexanediol, time and changes in protein and salt concentration were investigated. In addition, FRAP experiments were performed to elucidate its dynamics. These revealed an ability of NP to condensates under a variety of different conditions with strong viscoelastic properties.

Analysis of the NP sequence revealed the presence of many positively charged amino acids that can interact with nucleic acids to promote phase separation. The protein was also revealed to have droplet-promoting regions and a large C-terminal intrinsically disordered region. This would imply that NP can undergo phase separation due to its flexible and positively charged components facilitating multivalent interactions. When taken with its tendency to oligomerize we would expect NP to readily phase separate to form liquid condensates **(Figure 7).**

NP takes on three main phenotypes within the nucleus. Among these, only the complex phenotype has large regions of NP localization with some liquid properties. These make up a maximum of 37% of transfected cells, with proportions generally closer to 10% **(Figure 8A).** These large spaces on average recover from FRAP to around 0.4 of initial fluorescence, affected

by infection timepoint **(Figure 9A)** and have rounded edges and blob-like shapes that would be associated with a liquid state, but one with reduced surface tension compared to most liquid condensates, which must be round for reasons of energetic favorability **(Figure 8F, 9C)**. These large nuclear condensates, while being generally immobile, certainly have internal exchange and liquid features that we would expect from a condensate but lack the dynamism we would expect from fully liquid LLPS-derived condensates. Therefore, they are likely to be highly viscoelastic in nature. Viscoelasticity when speaking about biomolecular condensates is usually seen as relating to the finiteness of the shear relaxation rate of a structure, a merely viscous condensate would have an instantaneous shear relaxation.[81] Therefore, a viscoelastic condensate exhibits both elastic and viscous properties. Such a material would not necessarily appear round given external stressors and would have a definite shape.[82] In reality, all condensates can be viscoelastic to some extent though in most the spherical shape implies that viscosity is the main attribute (and indeed they are commonly modelled with only this feature in mind).[83]

In the case of the nucleolus, its viscoelasticity and irregular non-droplet like shape is believed to be due to rRNAs forming a gel matrix as they pass towards the outer edge from the follicular centers (FCs).[84] Indeed, it has been shown that the irregular shape of nucleoli could be predicted based only on these flux levels and the locations of FCs. It is thought that this viscoelasticity is critical to the nucleolus' function for allowing this rRNA to flow out from the nucleoli as it is processed.[84] Similar conditions and functions could be true for the NP viscoelastic condensates, with irregular shaping being due to something moving outwards from certain points combined with efflux from the viscoelastic condensates somewhere else, though unlike in nucleoli

transcriptional sources are unlikely to be responsible due to these structures appearing when NP alone is transfected.

In addition to these larger structures, there are occasionally small round mobile puncta appearing mostly during transfection with GFP-NP**.** These puncta are rarely present, appearing in at most 10% of cells when GFP-NP is used (Data not shown). Different time points of infection and transfection were attempted to try to find them in greater proportion, but this was not successful **(Figure 8).** It has been noted that NP can (though does not necessarily need to) localize to a number of different nuclear spaces, especially PML bodies and CBs.[85] These structures may be the small slow moving puncta that we observe, which would be in line with the literature.[85] In addition, eGFP has weak dimerization properties, which may be the cause of that these small round puncta.[86]

The nucleolar localization of NP is only present during transfection or in the very early stages of infection. Tests involving the cotransfection of mCherry-fibrillarin and NP showed strong colocalization **(Figure S4)** and indicated similar NP diffusion rates in both the nucleus and the nucleolus **(Figure 9B)**. Further, the apparent exclusion of NP from the fibrillar centers (believed to be the small dark regions in the nucleoli in the complex phenotype cells) but presence in the granular component is interesting and suggests that NP may bind either other nucleolar proteins or the rRNA following its transcription. This exclusion was observed with all three available NP fluorophores for primarily the complex phenotype. It is not known exactly why NP localizes to the nucleolus, but it is believed to be important for proper assembly of the vRNPs, as NP lacking a nucleolar localization sequence form irregular vRNPs which affect viral infectivity.[87] NP is also

known to form large oligomers when binding RNA of at least 24nt nonspecifically, with smaller trimers and monomers forming in the presence of shorter RNA, perhaps it is oligomerizing in contact with rRNA or snoRNA.[64]

NP localization to the nucleolus may be important for NP's interaction with nucleolin which has been found to promote localization of the vRNP to the nuclear periphery and dense chromatin regions before nuclear export.[88] Indeed, such a phenomena may also explain the appearance of the complex phenotype, as it should be noted that with the nucleolar signal phenotype the brightest region of the nucleus is the nucleolus while for the complex signal type the brightest areas are the dense nuclear regions. The cells with complex signal are also usually brighter in general, suggesting that as NP concentration increases, there is a movement from the nucleolus to the nucleoplasm, in line with the necessary passage of the NP through the nucleolus before localization to the nuclear periphery and export. During infection, the NP transitions from the nucleolus to the nucleoplasm and the virus transitions from primarily producing vmRNA to the production of vRNA and vRNPs which require NP to be present.[25] NP localization may therefore be a factor in determining the timing of this switch which could be based solely on NP quantity.

In regard to the diffuse phenotype, it is localized to the nucleoplasm generally rather than having any discernable subnuclear localization, though diffuse signal may be nucleolar signal that is simply too faint to make out as the diffuse cells are very faint in comparison to the other two types. NP at low concentrations may have no tendency to localize to nuclear structures. This is also why so many cells transfected with mCherry-NP are classed as diffuse, they were simply too faint to allow any structures to be detected.

Deletions in the DPRs alter NP localization, with C-terminal deletions favoring nuclear localization similar to WT NP and N-terminal ones favoring cytoplasmic localization. The 63-103AA deletion mutant was particularly interesting due to its nucleolar exclusion matching the phenotype observed during intermediate stages of infection where the NP is still nuclear but excluded from the nucleolus. The DIII and DIV deletion mutants (**Figure 10**) resulted in generally similar phenotypes to WT NP, though without much formation of the dense nuclear regions indicative of the complex phenotype. This lack of dense nuclear regions in response to any DPR deletion supports the idea that these regions are viscoelastic condensates.

When considering NP *in vitro* WT and R74A mutants tended to have similar properties, this is despite the fact that arginine to alanine mutations have been found to decrease fluidity, especially in the presence of RNA and should reduce phase separation.[89-91] R74A NP seems to have a tendency to undergo increased condensation and aggregation at low salt concentrations (50mM) while both WT and R74A do so in the presence of hexanediol **(Figure 11C, 12)**. At low salt concentrations NP has a tendency to form monomers instead of trimers or oligomers which implies that these latter structures help to prevent condensation and aggregation.[64]

The increase in ability to form condensates in response to hexanediol is difficult to explain. Generally, small round condensates which are dynamically arrested are associated with components that undergo phase separation but then rapidly harden (arrested phase separation).[60] This can occur when the intermolecular interactions in a condensate become too strong as the condensate grows over time or due to the presence of RNA or other binding factors.[92] The

especially small size of NP condensates given the relatively high protein concentrations tested imply this may the case. This may also explain the relatively high amount of background fluorescence observed *in vitro* as much of the NP may simply be diffuse. If we compare NP to HIV-1 Gag under similar in vitro conditions, we find a significant difference in size and visibility of the condensates with the Gag forming far larger and more numerous droplets at lower concentrations. In addition, it was found that when hexanediol was added to the Gag the size and number of condensates greatly decreased as would be expected with a liquid condensate, unlike with NP **(Figure S3).**

From this information we can conclude that the NP *in vitro* likely condenses in response to the crowding agent but then hardens before the condensates reach their largest potential size. It is known that hexanediol primarily targets hydrophobic interactions such as the weak protein-protein bonding common in phase-separated biomolecular condensates.[53,93] It is therefore likely that NP condensates are not formed due to such interactions which are based on enthalpically-favorable conditions caused by the destruction of the clathrate water molecule cage around dissolved substances.[93] This would therefore argue for a model in which viscoelastic condensates of NP are formed through non-hydrophobic effects such as electrostatic interactions. If added, the hexanediol may then disrupt the weak hydrophobic interactions of NP which further encourages ionic interaction between the proteins' positively-charged and negatively-charged regions, resulting in a collapse of the protein structure and the rapid and extensive formation of viscoelastic condensates and aggregates depending on localized effects. This is consistent with the greater condensation of R74A NP observed at lower salt concentrations as high salt tends to disrupt such ionic interactions and condensation in low-salt conditions is driven by IDRs, which

NP has.[94] An effect on condensation caused by hexanediol implies that the chemical can enter the condensates in the first place which supports the idea that the droplets seen are indeed liquid in nature. It has been demonstrated previously that hexanediol causes the dissolution of cytoplasmic NP condensates during IAV infection. Those condensates are highly liquid in nature and so are readily disrupted, though sometimes resist the treatment.[4] The lack of dissolution upon hexanediol treatment implies that the *in vitro* NP condensates are more solid than those cytoplasmic condensates.

The area of the *in vitro* condensates following 20 minutes or 60 minutes of time is similar for WT and R74A NP, further suggesting that these condensates are arrested dynamically **(Figure 11)**. It is also interesting to note that the area of NP condensates is highly inconsistent between *in vitro* experiments **(Figure 11C)**, implying that even very small changes in the system can result in far higher or lower levels of condensation and aggregation or that such aggregate formation is highly stochastic in nature.

The characteristics of WT NP and R75A NP in live cells are very similar. Since the DPRII region in which R75 is found has many arginines it may be that it requires more than one R to A mutation to have a significant effect on NP properties. The R75 residue has been found to be highly conserved (>99%) among IAV strains, and the R75A mutation is noted as resulting in reduced RNA binding. [61,80] In addition, this mutation results in non-infectious virus but no changes in polymerase activity or production of a reporter gene from an artificial vRNA.[61] However, if stretches of amino acids around the second droplet-promoting region are deleted, an effect on polymerase activity is noted.[61] It similarly appears that the disruption to RNA-binding

and the second droplet-promoting region caused by the R75A mutation is not enough to prevent formation of the structures indicative of the complex phenotype, unlike a complete deletion of DPRII **(Figure 9)**. Similar behavior between the WT and R75A in cells is overall consistent with the *in vitro* experiments where R74A and WT NP have similar areas under different conditions as well as similar reactions to hexanediol. It could be that the R75 residue only causes localization or condensation differences in the context of infection, or in the presence of viral RNA.

These results when taken together suggest that the nuclear viscoelastic condensates may be staging grounds for vRNPs, possibly involved in enhancing polymerase activity. During the process of IAV infection all viral proteins increase in cellular concentration with different strains producing different relative amounts of each protein over time.[95] In PR8 virus, NP is the most abundant protein throughout infection, levels of NP in MDCK cells have been shown to rapidly increase post-infection with it being the first protein made out of the major four viral proteins (HA, NP, NA, M1). NP concentration then plateaus at around 5-7 hours post-infection, or when NP begins to leave the nucleus and when nucleolar NP begins to disappear.[87,96] This would support a model where low NP concentration results in diffuse or nucleolar-localized NP but at higher levels it begins to build up at the nuclear periphery or in the chromatin in preparation for export. If export proteins are lacking, as they are in the transfection experiments performed NP would instead be stuck in the nucleoplasm within its staging areas forming these viscoelastic condensates. These regions also sometimes appear in infected cells, though are rare and do not appear the same as in transfected cells, generally appearing more irregular and dispersed, perhaps due to withdrawal of NP from these spaces (**Figure 8F**). As an alternative to this staging ground hypothesis, it has been observed that IAV NP colocalizes with host clusterin in the

perinuclear region at 8-12 hours post-infection. The observed viscoelastic structures may be related to this and would be consistent with the widespread pro-apoptotic effects observed in cells where NP is present.[97]

Limitations and Caveats

Phenotype frequency variations between cells transfected with different fluorescently bound forms of NP may be due to brightness of the fluorophore. It is notable that the proportion of cells displaying nucleolar signal compared to diffuse signal is higher when MNG is used, suggesting that the difference between the two phenotypes may be brightness related due to the higher quantum yield and brightness of MNG relative to the other two fluorophores (1.5-3x brighter than GFP, 5-6x brighter than mCherry).[98,99]

A limitation of transfection experiments with NP that was found in the course of this work is that there is a very limited window in which fluorescently labelled NP is visible in the nucleus before the vast majority of transfected cells begin dying due to NP's pro-apoptotic effects, especially on the A549 cells tested[97]. This means that obtaining an early timepoint for NP transfection, when very little is present or a late timepoint when we would normally observe transfected cells (24 hours) is unfortunately not ideal. While a late timepoint for transfection could be obtained, the cells are in overall very poor health in both infection and transfection conditions that it is scarcely worth observing NP dynamics within them due to the large inconsistencies in nuclear morphology. In many cases the nucleus is fragmented or the nuclear envelope is ruptured, a notable sign of apoptosis.[100]

A limitation of the *in vitro* experiments is that under different conditions non-labelled NP aggregates are sometimes visible on certain parts of the slide for certain samples, generally at high salt concentrations above 150mM or in the mock hexanediol experiments, while not appearing in the presence of hexanediol. They are sometimes present in very high numbers in the 200mM and 250mM salt conditions for both WT and R74A NP. These structures are either irregularly shaped or small and apparently round and are likely aggregates of NP that failed to incorporate labelled NP in their formation. They are frequently difficult to observe due to their small size and are not fluorescent, and so are not easily quantifiable or included in calculations of condensate number or area though their incorporation of NP would affect the amount available for the formation of condensates. This is why greater focus is given to condensate area rather than number in the data presented, as these unmarked structures are similar in size to those which are fluorescent.

Conclusion

In conclusion, IAV NP can form viscoelastic condensates in cell nuclei and *in vitro*. This is supported by data demonstrating the formation of areas of dense nuclear concentration in A549 lung epithelial cells under a variety of conditions which recovery rapidly from photobleaching and have rounded irregular forms consistent with a viscoelastic nature. In addition, the deletion of any of the identified droplet-promoting regions of NP completely abrogates their formation. The live cell data is supported by that obtained from purified NP protein, where small rounded NP condensates are observed under a variety of salt and protein concentration conditions that respond to treatment with hexanediol, with differences being noted between WT NP and R74A in low salt conditions. These viscoelastic condensates are therefore likely an example of arrested phase separation which does not rely on hydrophobic interactions to from, and are instead mostly

based on ionic interactions, explaining their more solid-like characteristics. Intriguing data from live cells and parallels between NP viscoelastic condensates and nucleoli may indicate that these nuclear viscoelastic condensates are formed as part of NP's transition from the nucleoli to the chromatin in preparation for inclusion in vRNPs and nuclear export, with the NP trafficking between the nucleoli and nucleoplasm being concentration-dependent. Alternatively, they may be the sites of other viral activities such as interaction with clusterin, causing dissociation from Bax and apoptosis[97]. Future research should examine the purpose for these observed nuclear viscoelastic condensates and whether they may serve as the site of vRNP assembly or viral polymerase activity. In addition, there is the question of how these dense NP viscoelastic condensates with slow dynamics in the nucleus turn into the highly dynamic and liquid condensates in the cytoplasm following nuclear export of vRNPs.

Future research could discover the internal organization of the condensates and their makeup by using mass spectrometry and proteomics on viscoelastic condensates hardened using nucleozin and isolated from the cells as well as by tracking the localization of host nuclear proteins known to interact with NP. To find the function of these viscoelastic condensates, one could track the production of vRNA and the localization of viral polymerase components through fluorescently-tagged proteins and nucleotides to discover whether the viscoelastic condensates are the site of polymerase activity or vRNP assembly. It would be of particular interest to determine whether the irregular shape of the viscoelastic condensates is due to outwards flow of material from certain focal points like in nucleoli. To find whether the viscoelastic condensates support polymerase activity NP mutants with known equivalent polymerase activity in the context of a mini-replicon system could be transfected into cells to observe whether there are differences in

viscoelastic condensate formation. To find how the dense NP viscoelastic condensates in the nucleus transition into the highly liquid and mobile condensates in the cytoplasm longer term observations of single cells could be performed to track the passage of tagged NP during infection.

Reference List

1. Nakano M, Sugita Y, Kodera N, et al. Ultrastructure of influenza virus ribonucleoprotein complexes during viral RNA synthesis. *Communications Biology*. 2021/07/09 2021;4(1):858. doi:10.1038/s42003-021-02388-4
2. Dou D, Revol R, Östbye H, Wang H, Daniels R. Influenza A Virus Cell Entry, Replication, Virion Assembly and Movement. *Front Immunol*. 2018;9:1581. doi:10.3389/fimmu.2018.01581
3. Wei W, Bai L, Yan B, et al. When liquid-liquid phase separation meets viral infections. *Front Immunol*. 2022;13:985622. doi:10.3389/fimmu.2022.985622
4. Alenquer M, Vale-Costa S, Etibor TA, Ferreira F, Sousa AL, Amorim MJ. Influenza A virus ribonucleoproteins form liquid organelles at endoplasmic reticulum exit sites. *Nature Communications*. 2019/04/09 2019;10(1):1629. doi:10.1038/s41467-019-09549-4
5. Etibor TA, Sridharan S, Vale-Costa S, et al. Influenza A virus liquid condensates can undergo pharmacological hardening. *bioRxiv*. 2022:2022.08.03.502602. doi:10.1101/2022.08.03.502602
6. Morens DM, North M, Taubenberger JK. Eyewitness accounts of the 1510 influenza pandemic in Europe. *Lancet*. Dec 4 2010;376(9756):1894-5. doi:10.1016/s0140-6736(10)62204-0
7. Etymologia: influenza. *Emerging Infectious Disease journal*. 2006;12(1):179. doi:10.3201/eid1201.et1201
8. Taubenberger JK, Hultin JV, Morens DM. Discovery and characterization of the 1918 pandemic influenza virus in historical context. *Antivir Ther*. 2007;12(4 Pt B):581-91.
9. Taubenberger JK, Morens DM. Influenza: the once and future pandemic. *Public Health Rep*. Apr 2010;125 Suppl 3(Suppl 3):16-26.
10. Bouvier NM, Lowen AC. Animal Models for Influenza Virus Pathogenesis and Transmission. *Viruses*. 2010;2(8):1530-1563. doi:10.3390/v20801530 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3063653/pdf/viruses-02-01530.pdf
11. Schickli JH, Flandorfer A, Nakaya T, Martinez-Sobrido L, García-Sastre A, Palese P. Plasmid-only rescue of influenza A virus vaccine candidates. *Philos Trans R Soc Lond B Biol Sci*. Dec 29 2001;356(1416):1965-73. doi:10.1098/rstb.2001.0979
12. Morens DM, Taubenberger JK, Folkers GK, Fauci AS. Pandemic influenza's 500th anniversary. *Clin Infect Dis*. Dec 15 2010;51(12):1442-4. doi:10.1086/657429
13. Barberis I, Myles P, Ault SK, Bragazzi NL, Martini M. History and evolution of influenza control through vaccination: from the first monovalent vaccine to universal vaccines. *J Prev Med Hyg*. Sep 2016;57(3):E115-e120.
14. Tumpey TM, Basler CF, Aguilar PV, et al. Characterization of the reconstructed 1918 Spanish influenza pandemic virus. *Science*. Oct 7 2005;310(5745):77-80. doi:10.1126/science.1119392
15. Nogales A, Martínez-Sobrido L. Reverse Genetics Approaches for the Development of Influenza Vaccines. *Int J Mol Sci*. Dec 22 2016;18(1)doi:10.3390/ijms18010020
16. Neumann G, Fujii K, Kino Y, Kawaoka Y. An improved reverse genetics system for influenza A virus generation and its implications for vaccine production. *Proceedings of the National Academy of Sciences*. 2005;102(46):16825-16829. doi:doi:10.1073/pnas.0505587102
17. Russell CJ, Webster RG. The Genesis of a Pandemic Influenza Virus. *Cell*. 2005;123(3):368-371. doi:10.1016/j.cell.2005.10.019

18. Hannoun C. The evolving history of influenza viruses and influenza vaccines. *Expert Rev Vaccines*. Sep 2013;12(9):1085-94. doi:10.1586/14760584.2013.824709
19. Nuwarda RF, Alharbi AA, Kayser V. An Overview of Influenza Viruses and Vaccines. *Vaccines (Basel)*. Sep 17 2021;9(9)doi:10.3390/vaccines9091032
20. Yewdell JW, Ince WL. Virology. Frameshifting to PA-X influenza. *Science*. Jul 13 2012;337(6091):164-5. doi:10.1126/science.1225539
21. Pinto RM, Lycett S, Gaunt E, Digard P. Accessory Gene Products of Influenza A Virus. *Cold Spring Harb Perspect Med*. Dec 1 2021;11(12)doi:10.1101/cshperspect.a038380
22. Liu C-L, Hung H-C, Lo S-C, et al. Using mutagenesis to explore conserved residues in the RNA-binding groove of influenza A virus nucleoprotein for antiviral drug development. *Scientific Reports*. 2016/02/26 2016;6(1):21662. doi:10.1038/srep21662
23. Reed SG, Ager A. Immune Responses to IAV Infection and the Roles of L-Selectin and ADAM17 in Lymphocyte Homing. *Pathogens*. Jan 25 2022;11(2)doi:10.3390/pathogens11020150
24. Villalón-Letelier F, Brooks AG, Saunders PM, Londrigan SL, Reading PC. Host Cell Restriction Factors that Limit Influenza A Infection. *Viruses*. Dec 7 2017;9(12)doi:10.3390/v9120376
25. Phan T, Fay Elizabeth J, Lee Z, Aron S, Hu W-S, Langlois Ryan A. Segment-Specific Kinetics of mRNA, cRNA, and vRNA Accumulation during Influenza Virus Infection. *Journal of Virology*. 2021;95(10):10.1128/jvi.02102-20. doi:10.1128/jvi.02102-20
26. Staller E, Sheppard CM, Neasham PJ, et al. ANP32 Proteins Are Essential for Influenza Virus Replication in Human Cells. *J Virol*. Sep 1 2019;93(17)doi:10.1128/jvi.00217-19
27. Te Velthuis AJ, Fodor E. Influenza virus RNA polymerase: insights into the mechanisms of viral RNA synthesis. *Nat Rev Microbiol*. Aug 2016;14(8):479-93. doi:10.1038/nrmicro.2016.87
28. Carter T, Iqbal M. The Influenza A Virus Replication Cycle: A Comprehensive Review. *Viruses*. 2024;16(2):316.
29. Chou Y-y, Vafabakhsh R, Doğanay S, Gao Q, Ha T, Palese P. One influenza virus particle packages eight unique viral RNAs as shown by FISH analysis. *Proceedings of the National Academy of Sciences*. 2012/06/05 2012;109(23):9101-9106. doi:10.1073/pnas.1206069109
30. Noda T. Selective Genome Packaging Mechanisms of Influenza A Viruses. *Cold Spring Harb Perspect Med*. Jul 1 2021;11(7)doi:10.1101/cshperspect.a038497
31. Rossman JS, Lamb RA. Influenza virus assembly and budding. *Virology*. Mar 15 2011;411(2):229-36. doi:10.1016/j.virol.2010.12.003
32. Dou D, Revol R, Östbye H, Wang H, Daniels R. Influenza A Virus Cell Entry, Replication, Virion Assembly and Movement. Review. *Frontiers in Immunology*. 2018-July-20 2018;9doi:10.3389/fimmu.2018.01581
33. Zheng W, Tao Y. Structure and assembly of the influenza A virus ribonucleoprotein complex. *FEBS letters*. 03/12 2013;587doi:10.1016/j.febslet.2013.02.048
34. Niu X, Zhang L, Wu Y, et al. Biomolecular condensates: Formation mechanisms, biological functions, and therapeutic targets. *MedComm (2020)*. Apr 2023;4(2):e223. doi:10.1002/mco2.223
35. Mitrea DM, Mittasch M, Gomes BF, Klein IA, Murcko MA. Modulating biomolecular condensates: a novel approach to drug discovery. *Nature Reviews Drug Discovery*. 2022/11/01 2022;21(11):841-862. doi:10.1038/s41573-022-00505-4

36. Schmit JD, Feric M, Dundr M. How Hierarchical Interactions Make Membraneless Organelles Tick Like Clockwork. *Trends Biochem Sci*. Jul 2021;46(7):525-534. doi:10.1016/j.tibs.2020.12.011
37. Ginell GM, Holehouse AS. An Introduction to the Stickers-and-Spacers Framework as Applied to Biomolecular Condensates. *Methods Mol Biol*. 2023;2563:95-116. doi:10.1007/978-1-0716-2663-4_4
38. Sagan SM, Weber SC. Let’s phase it: viruses are master architects of biomolecular condensates. *Trends in Biochemical Sciences*. 2023;48(3):229-243. doi:10.1016/j.tibs.2022.09.008
39. Martin EW, Holehouse AS. Intrinsically disordered protein regions and phase separation: sequence determinants of assembly or lack thereof. *Emerg Top Life Sci*. Dec 11 2020;4(3):307-329. doi:10.1042/etls20190164
40. Chu X, Sun T, Li Q, et al. Prediction of liquid–liquid phase separating proteins using machine learning. *BMC Bioinformatics*. 2022/02/15 2022;23(1):72. doi:10.1186/s12859-022-04599-w
41. Guo Q, Shi X, Wang X. RNA and liquid-liquid phase separation. *Noncoding RNA Res*. Jun 2021;6(2):92-99. doi:10.1016/j.ncrna.2021.04.003
42. Dai Z, Yang X. The regulation of liquid-liquid phase separated condensates containing nucleic acids. *The FEBS Journal*. 2023/09/21 2023;n/a(n/a)doi:https://doi.org/10.1111/febs.16959
43. Miyagi T, Yamazaki R, Ueda K, et al. The Patterning and Proportion of Charged Residues in the Arginine-Rich Mixed-Charge Domain Determine the Membrane-Less Organelle Targeted by the Protein. *Int J Mol Sci*. Jul 11 2022;23(14)doi:10.3390/ijms23147658
44. Hofweber M, Dormann D. Friend or foe-Post-translational modifications as regulators of phase separation and RNP granule dynamics. *J Biol Chem*. May 3 2019;294(18):7137-7150. doi:10.1074/jbc.TM118.001189
45. Rubio K, Dobersch S, Barreto G. Functional interactions between scaffold proteins, noncoding RNAs, and genome loci induce liquid-liquid phase separation as organizing principle for 3-dimensional nuclear architecture: implications in cancer. *The FASEB Journal*. 2019/05/01 2019;33(5):5814-5822. doi:https://doi.org/10.1096/fj.201802715R
46. Schede HH, Natarajan P, Chakraborty AK, Shrinivas K. A model for organization and regulation of nuclear condensates by gene activity. *Nature Communications*. 2023/07/12 2023;14(1):4152. doi:10.1038/s41467-023-39878-4
47. Banani SF, Lee HO, Hyman AA, Rosen MK. Biomolecular condensates: organizers of cellular biochemistry. *Nat Rev Mol Cell Biol*. May 2017;18(5):285-298. doi:10.1038/nrm.2017.7
48. Wang B, Zhang L, Dai T, et al. Liquid–liquid phase separation in human health and diseases. *Signal Transduction and Targeted Therapy*. 2021/08/02 2021;6(1):290. doi:10.1038/s41392-021-00678-1
49. Field S, Jang G-J, Dean C, Strader LC, Rhee SY. Plants use molecular mechanisms mediated by biomolecular condensates to integrate environmental cues with development. *The Plant Cell*. 2023;35(9):3173-3186. doi:10.1093/plcell/koad062
50. Garabedian MV, Su Z, Dabdoub J, et al. Protein Condensate Formation via Controlled Multimerization of Intrinsically Disordered Sequences. *Biochemistry*. Nov 15 2022;61(22):2470-2481. doi:10.1021/acs.biochem.2c00250

51. Taylor NO, Wei MT, Stone HA, Brangwynne CP. Quantifying Dynamics in Phase-Separated Condensates Using Fluorescence Recovery after Photobleaching. *Biophys J*. Oct 1 2019;117(7):1285-1300. doi:10.1016/j.bpj.2019.08.030
52. Lippincott-Schwartz J, Snapp EL, Phair RD. The Development and Enhancement of FRAP as a Key Tool for Investigating Protein Dynamics. *Biophys J*. Oct 2 2018;115(7):1146-1155. doi:10.1016/j.bpj.2018.08.007
53. Liu X, Jiang S, Ma L, et al. Time-dependent effect of 1,6-hexanediol on biomolecular condensates and 3D chromatin organization. *Genome Biology*. 2021/08/17 2021;22(1):230. doi:10.1186/s13059-021-02455-3
54. Garaizar A, Espinosa JR. Salt dependent phase behavior of intrinsically disordered proteins from a coarse-grained model with explicit water and ions. *The Journal of Chemical Physics*. 2021;155(12)doi:10.1063/5.0062687
55. Hazra MK, Levy Y. Cross-Talk of Cation-π Interactions with Electrostatic and Aromatic Interactions: A Salt-Dependent Trade-off in Biomolecular Condensates. *J Phys Chem Lett*. Sep 28 2023;14(38):8460-8469. doi:10.1021/acs.jpclett.3c01642
56. Cinar H, Fetahaj Z, Cinar S, Vernon RM, Chan HS, Winter RHA. Temperature, Hydrostatic Pressure, and Osmolyte Effects on Liquid–Liquid Phase Separation in Protein Condensates: Physical Chemistry and Biological Implications. *Chemistry – A European Journal*. 2019/10/11 2019;25(57):13049-13069. doi:https://doi.org/10.1002/chem.201902210
57. Husain M, Cheung CY. Histone deacetylase 6 inhibits influenza A virus release by downregulating the trafficking of viral components to the plasma membrane via its substrate, acetylated microtubules. *J Virol*. Oct 2014;88(19):11229-39. doi:10.1128/jvi.00727-14
58. Etibor TA, Yamauchi Y, Amorim MJ. Liquid Biomolecular Condensates and Viral Lifecycles: Review and Perspectives. *Viruses*. 2021;13(3). doi:10.3390/v13030366
59. Etibor TA, Vale-Costa S, Sridharan S, et al. Defining basic rules for hardening influenza A virus liquid condensates. *Elife*. Apr 4 2023;12doi:10.7554/eLife.85182
60. Alberti S, Gladfelter A, Mittag T. Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell*. Jan 24 2019;176(3):419-434. doi:10.1016/j.cell.2018.12.035
61. Kakisaka M, Yamada K, Yamaji-Hasegawa A, Kobayashi T, Aida Y. Intrinsically disordered region of influenza A NP regulates viral genome packaging via interactions with viral RNA and host PI(4,5)P2. *Virology*. 2016/09/01/ 2016;496:116-126. doi:https://doi.org/10.1016/j.virol.2016.05.018
62. Miyake Y, Keusch JJ, Decamps L, et al. Influenza virus uses transportin 1 for vRNP debundling during cell entry. *Nature Microbiology*. 2019/04/01 2019;4(4):578-586. doi:10.1038/s41564-018-0332-2
63. James GT. Inactivation of the protease inhibitor phenylmethylsulfonyl fluoride in buffers. *Analytical Biochemistry*. 1978/06/01/ 1978;86(2):574-579. doi:https://doi.org/10.1016/0003-2697(78)90784-4
64. Labaronne A, Swale C, Monod A, Schoehn G, Crépin T, Ruigrok RW. Binding of RNA by the Nucleoproteins of Influenza Viruses A and B. *Viruses*. Sep 13 2016;8(9)doi:10.3390/v8090247
65. Wit Ed, Spronken MIJ, Bestebroer TM, Rimmelzwaan GF, Osterhaus ADME, Fouchier RAM. Efficient generation and growth of influenza virus A/PR/8/34 from eight cDNA fragments. *Virus Research*. 2004/07/01/ 2004;103(1):155-161. doi:https://doi.org/10.1016/j.virusres.2004.02.028

66. Hardenberg M, Horvath A, Ambrus V, Fuxreiter M, Vendruscolo M. Widespread occurrence of the droplet state of proteins in the human proteome. *Proceedings of the National Academy of Sciences*. 2020;117(52):33254-33262. doi:doi:10.1073/pnas.2007670117
67. Vendruscolo M, Fuxreiter M. Sequence Determinants of the Aggregation of Proteins Within Condensates Generated by Liquid-liquid Phase Separation. *Journal of Molecular Biology*. 2022/01/15/ 2022;434(1):167201. doi:https://doi.org/10.1016/j.jmb.2021.167201
68. Hatos A, Tosatto SCE, Vendruscolo M, Fuxreiter M. FuzDrop on AlphaFold: visualizing the sequence-dependent propensity of liquid–liquid phase separation and aggregation of proteins. *Nucleic Acids Research*. 2022;50(W1):W337-W344. doi:10.1093/nar/gkac386
69. Corley M, Burns MC, Yeo GW. How RNA-Binding Proteins Interact with RNA: Molecules and Mechanisms. *Mol Cell*. Apr 2 2020;78(1):9-29. doi:10.1016/j.molcel.2020.03.011
70. Garner E, Romero P, Dunker AK, Brown C, Obradovic Z. Predicting Binding Regions within Disordered Proteins. *Genome Inform Ser Workshop Genome Inform*. 1999;10:41-50.
71. Martínez-Sobrido L, Peersen O, Nogales A. Temperature Sensitive Mutations in Influenza A Viral Ribonucleoprotein Complex Responsible for the Attenuation of the Live Attenuated Influenza Vaccine. *Viruses*. Oct 15 2018;10(10)doi:10.3390/v10100560
72. Ozawa M, Fujii K, Muramoto Y, et al. Contributions of two nuclear localization signals of influenza A virus nucleoprotein to viral replication. *J Virol*. Jan 2007;81(1):30-41. doi:10.1128/jvi.01434-06
73. Li J, Yu M, Zheng W, Liu W. Nucleocytoplasmic Shuttling of Influenza A Virus Proteins. *Viruses*. 2015;7(5):2668-2682. doi:10.3390/v7052668
74. Myatt DP, Hatter L, Rogers SE, Terry AE, Clifton LA. Monomeric green fluorescent protein as a protein standard for small angle scattering. *Biomedical Spectroscopy and Imaging*. 2017;6:123-134. doi:10.3233/BSI-170167
75. Abdoli A, Soleimanjahi H, Tavassoti Kheiri M, Jamali A, Jamaati A. Determining influenza virus shedding at different time points in madin-darby canine kidney cell line. *Cell J*. Summer 2013;15(2):130-5.
76. Höfer CT, Jolmes F, Haralampiev I, Veit M, Herrmann A. Influenza A virus nucleoprotein targets subnuclear structures. *Cellular Microbiology*. 2017/04/01 2017;19(4):e12679. doi:https://doi.org/10.1111/cmi.12679
77. Terry CA, Fernández M-J, Gude L, Lorente A, Grant KB. Physiologically Relevant Concentrations of NaCl and KCl Increase DNA Photocleavage by an N-Substituted 9-Aminomethylanthracene Dye. *Biochemistry*. 2011/11/29 2011;50(47):10375-10389. doi:10.1021/bi200972c
78. Tarus B, Bakowiez O, Chenavas S, et al. Oligomerization paths of the nucleoprotein of influenza A virus. *Biochimie*. 2012/03/01/ 2012;94(3):776-785. doi:https://doi.org/10.1016/j.biochi.2011.11.009
79. Monette A, Niu M, Maldonado RK, et al. Influence of HIV-1 Genomic RNA on the Formation of Gag Biomolecular Condensates. *J Mol Biol*. Aug 15 2023;435(16):168190. doi:10.1016/j.jmb.2023.168190
80. Kukol A, Hughes DJ. Large-scale analysis of influenza A virus nucleoprotein sequence conservation reveals potential drug-target sites. *Virology*. 2014/04/01/ 2014;454-455:40-47. doi:https://doi.org/10.1016/j.virol.2014.01.023
81. Ghosh A, Kota D, Zhou H-X. Shear relaxation governs fusion dynamics of biomolecular condensates. *Nature Communications*. 2021/10/13 2021;12(1):5995. doi:10.1038/s41467-021-26274-z

82. Deshmukh K, Kovářík T, Muzaffar A, Basheer Ahamed M, Khadheer Pasha SK. Chapter 4 - Mechanical analysis of polymers. In: AlMaadeed MAA, Ponnamma D, Carignano MA, eds. *Polymer Science and Innovative Applications*. Elsevier; 2020:117-152.
83. Michieletto D, Marenda M. Rheology and Viscoelasticity of Proteins and Nucleic Acids Condensates. *JACS Au*. Jul 25 2022;2(7):1506-1521. doi:10.1021/jacsau.2c00055
84. Riback JA, Eeftens JM, Lee DSW, et al. Viscoelasticity and advective flow of RNA underlies nucleolar form and function. *Molecular Cell*. 2023/09/07/ 2023;83(17):3095-3107.e9. doi:https://doi.org/10.1016/j.molcel.2023.08.006
85. Takizawa N, Watanabe K, Nouno K, Kobayashi N, Nagata K. Association of functional influenza viral proteins and RNAs with nuclear chromatin and sub-chromatin structure. *Microbes and Infection*. 2006/03/01/ 2006;8(3):823-833. doi:https://doi.org/10.1016/j.micinf.2005.10.005
86. Costantini LM, Fossati M, Francolini M, Snapp EL. Assessing the tendency of fluorescent proteins to oligomerize under physiologic conditions. *Traffic*. May 2012;13(5):643-9. doi:10.1111/j.1600-0854.2012.01336.x
87. Miyamoto S, Nakano M, Morikawa T, et al. Migration of Influenza Virus Nucleoprotein into the Nucleolus Is Essential for Ribonucleoprotein Complex Formation. *mBio*. Jan-Feb 2022;13(1)
88. Terrier O, Carron C, De Chassey B, et al. Nucleolin interacts with influenza A nucleoprotein and contributes to viral ribonucleoprotein complexes nuclear trafficking and efficient influenza viral replication. *Sci Rep*. Jul 4 2016;6:29006. doi:10.1038/srep29006
89. Fuxreiter M. Spot in a drop: mutations in aberrant condensates. *Nature Reviews Molecular Cell Biology*. 2021/03/01 2021;22(3):162-163. doi:10.1038/s41580-021-00338-w
90. Fisher RS, Elbaum-Garfinkle S. Tunable multiphase dynamics of arginine and lysine liquid condensates. *Nature Communications*. 2020/09/15 2020;11(1):4628. doi:10.1038/s41467-020-18224-y
91. Hong Y, Najafi S, Casey T, Shea JE, Han SI, Hwang DS. Hydrophobicity of arginine leads to reentrant liquid-liquid phase separation behaviors of arginine-rich proteins. *Nat Commun*. Nov 28 2022;13(1):7326. doi:10.1038/s41467-022-35001-1
92. Linsenmeier M, Hondele M, Grigolato F, Secchi E, Weis K, Arosio P. Dynamic arrest and aging of biomolecular condensates are modulated by low-complexity domains, RNA and biochemical activity. *Nature Communications*. 2022/05/31 2022;13(1):3030. doi:10.1038/s41467-022-30521-2
93. Sun Q. The Hydrophobic Effects: Our Current Understanding. *Molecules*. Oct 18 2022;27(20)doi:10.3390/molecules27207009
94. Lin Y, Protter DS, Rosen MK, Parker R. Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Mol Cell*. Oct 15 2015;60(2):208-19. doi:10.1016/j.molcel.2015.08.018
95. Bogdanow B, Wang X, Eichelbaum K, et al. The dynamic proteome of influenza A virus infection identifies M segment splicing as a host range determinant. *Nat Commun*. Dec 4 2019;10(1):5518. doi:10.1038/s41467-019-13520-8
96. Küchler J, Püttker S, Lahmann P, et al. Absolute quantification of viral proteins during single-round replication of MDCK suspension cells. *Journal of Proteomics*. 2022/05/15/ 2022;259:104544. doi:https://doi.org/10.1016/j.jprot.2022.104544
97. Tripathi S, Batra J, Cao W, et al. Influenza A virus nucleoprotein induces apoptosis in human airway epithelial cells: implications of a novel interaction between nucleoprotein and host

protein Clusterin. *Cell Death & Disease*. 2013/03/01 2013;4(3):e562-e562. doi:10.1038/cddis.2013.89
98. Shaner NC, Lambert GG, Chammas A, et al. A bright monomeric green fluorescent protein derived from Branchiostoma lanceolatum. *Nat Methods*. May 2013;10(5):407-9. doi:10.1038/nmeth.2413
99. Shaner NC, Campbell RE, Steinbach PA, Giepmans BNG, Palmer AE, Tsien RY. Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nature Biotechnology*. 2004/12/01 2004;22(12):1567-1572. doi:10.1038/nbt1037
100. Suzuki J, Nagata S. Chapter Fifteen - Phospholipid Scrambling on the Plasma Membrane. In: Ashkenazi A, Yuan J, Wells JA, eds. *Methods in Enzymology*. Academic Press; 2014:381-393.

Supplemental Figures

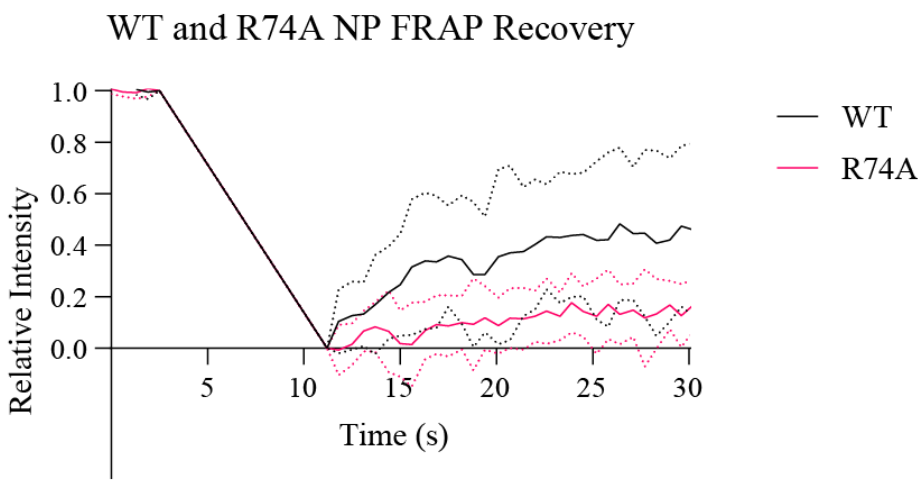


***Figure S 1 WT NP has faster dynamics than R74A NP* in vitro**

WT NP and R74A NP at 80µM (with AF488-labelled fluorescent variants) were mixed with salt solution to achieve 150mM [NaCl] and 10% PEG as a crowding agent. The resulting solution was added to a slide and imaged. Entire punctate condensates underwent FRAP with 1 minute follow-up, data to 30 seconds shown as after bleaching is apparent. Data was normalized so 1=pre-FRAP fluorescence intensity and 0=post-FRAP fluorescence intensity. Data shown is from 12 puncta bleached for each protein. Dotted lines represent SD.

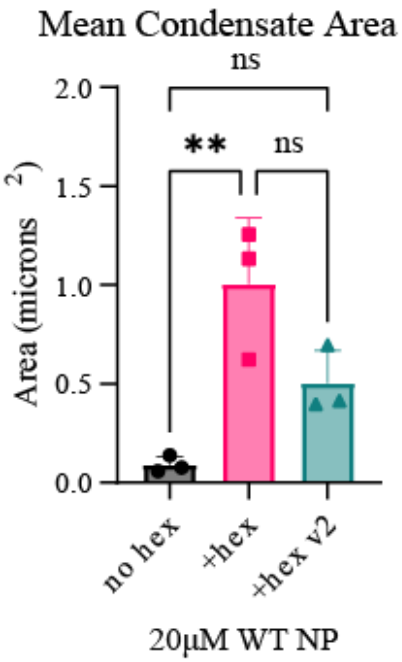


Figure S 2: Hexanediol effect on WT NP condensate area depending on addition order relative to PEG.

10µM WT NP at 150mM NaCl was exposed to 10% hexanediol with adjusted volume (+hex) or without adjusted volume (+Hex v2). Adjusted volume refers to whether the amount of salt-less solution added was lowered so that concentration of salt and protein would reach the intended amounts after addition of hexanediol. The second variant has no adjusted volume so the concentrations are the same as the -Hex before hexanediol addition. Representative graph showing area of condensates observed. Experiment was performed in triplicate. Significance calculated with one-way ANOVA with $p<0.05$.

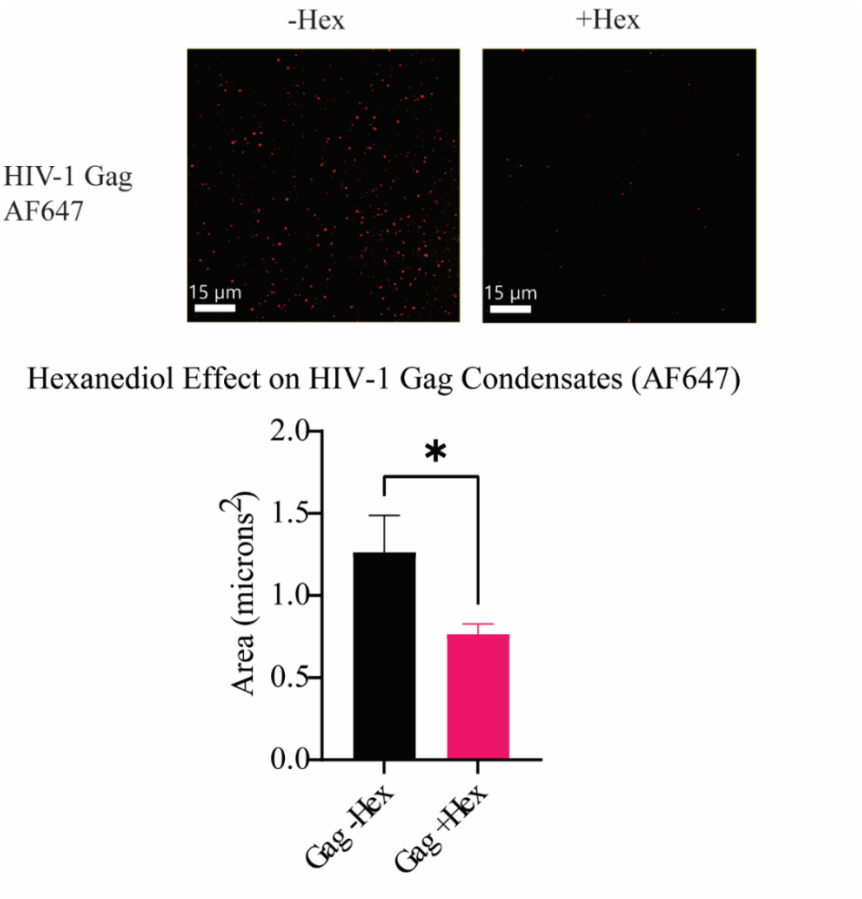


Figure S 3: Hexanediol reduces the area of AF-647 labelled HIV-1 Gag condensates.

Recombinant HIV-1 Gag was mixed 19:1 with AF647-labelled Gag and 10µM was plated in a solution containing 150mM NaCl and 10% PEG-8000. (A) Representative images showing AF647-labelled Gag signal with and without the addition of 10% hexanediol. (B) Representative graph showing the change in area of the condensates in this test. Experiment performed in triplicate. Significance calculated with unpaired t-test with $p<0.05$. See methods section (page 32) for labelling procedure. HIV-1 Gag was obtained from the laboratory of Leslie J. Parent at the Pennsylvania State University College of Medicine and is described in Monette et al. 2023.

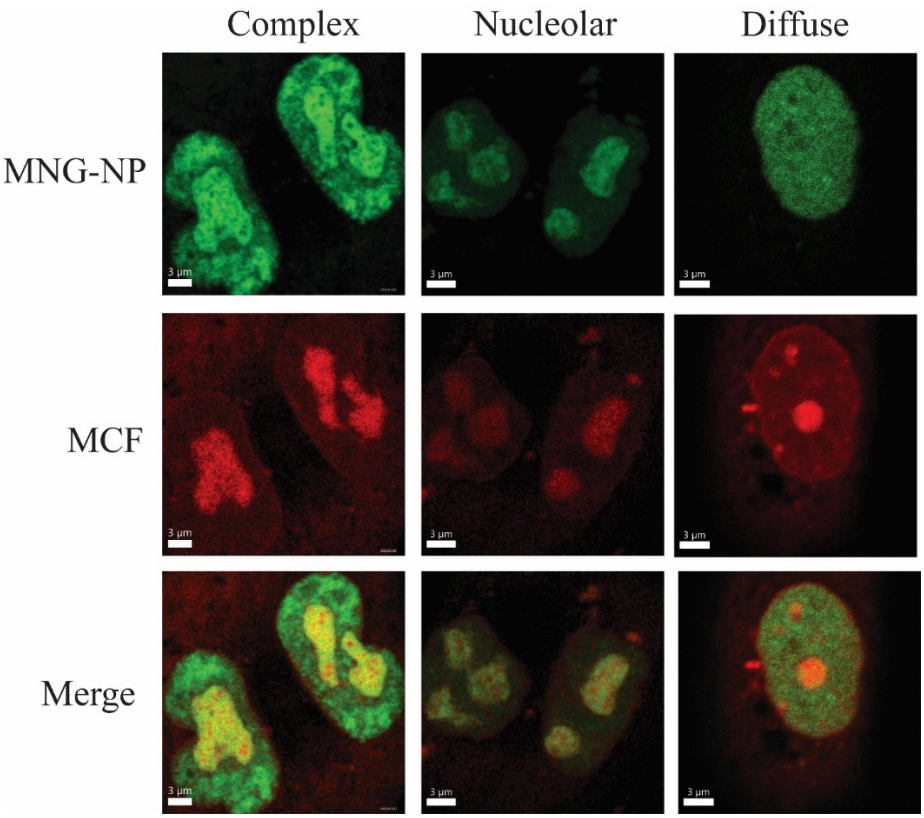


Figure S 4: NP and fibrillarin colocalize, indicating nucleoli where NP is present

A549 cells were transfected with varying amounts of MNG-NP and mCherry-fibrillarin plasmid, depending on experiment and well size, with clear entry of the NP to the nucleoplasm as can be seen here. The three main phenotypes previously identified are shown. These experiments were able to prove that certain spaces in other images and in experiments where mCherry-fibrillarin is not used are indeed nucleoli.