Structure and Interactions of Endonuclease in the Influenza A Polymerase Complex

Prepared by Isabella Rhyu

Supervised by Professor Jason Schnell

A thesis submitted in partial fulfilment of the requirements for the degree of Master of Science by Research in Biochemistry at the University of Oxford

January 2025

Abstract

The influenza viral genome is contained in eight discrete segments which are transcribed and replicated separately by influenza viral polymerases (FluPol) before being assembled into nascent virions and trafficked out of the host cell (Zhu et al., 2023). FluPols undergo dramatic conformational changes and can perform either transcription or replication depending on cellular conditions and interaction with different host factors (Carrique et al., 2020). The endonuclease at the N-terminal end of the Polymerase Acidic (PA) subunit of FluPol is necessary for transcription; endonuclease performs "cap-snatching," where host mRNA caps are cleaved and used to initiate transcription of viral mRNA (Dias et al., 2009). Interestingly, endonuclease remains structured in the cap-snatching incompetent conformations of FluPol and appears to interact with different subunits or external proteins depending on the FluPol conformation (Staller et al., 2024, Thierry et al., 2016; Vidic et al., 2016). Whether the interactions with different FluPol subunits regulate endonuclease activity is unknown. Based upon the hypothesis that regulation of promiscuous endonuclease activity could mediate conformational changes in the FluPol, I tested the interactions of endonuclease and three potential binding partners: viral nucleoprotein, a peptide from the Polymerase Basic 1 (PB1) subunit, and a peptide from the Polymerase Basic 2 (PB2) subunit. The same unstructured region of the endonuclease appeared to interact with viral nucleoprotein and the PB1 peptide. I introduced point mutations to knock out the interaction, confirming the residues implicated in binding, and performed activity assays to assess the enzymatic activity of wild type and mutant endonuclease with each potential binding partner, where the

mutants displayed substantial loss of function. These results increase understanding of endonuclease interactions in the influenza polymerase and identify a single residue mutation which confers almost complete loss of function, offering insight into potential targets for novel antiviral therapies.

Acknowledgments

I will first express my gratitude to Professor Jason Schnell for his support in the completion of this thesis and especially for his encouragement of the warm, collaborative atmosphere that makes it a pleasure to be in the lab.

Most sincere thanks to Daniel Molland for providing insight, ingenuity, and aggravation. Without his supervision on the bench I would have been truly lost. And to Maddie Eaton – "'I wonder what Piglet is doing,' thought Pooh. 'Wish I were there to be doing it too.'" Thank you to Lara Nikel for sharing in the ups and downs and for providing great emotional warmth when our heaters were broken.

I wish to acknowledge Ervin Fodor, who is incredibly generous with his time and his knowledge, and Christina Redfield, it seems there is no question for which she does not have an answer. Further thanks are due to Geidre Ratkeviciute and Finaritra Raoelijaona for their support in the lab and their willingness to answer my endless questions about cloning.

Of course the greatest thanks is due to my parents, who taught me everything I know and then some.

Table of Contents

Chapter 1: Introduction

Framework

Influenza is considered one of the greatest modern public health threats. The devastating pandemic potential of influenza virus infection was demonstrated to catastrophic effect during the 1918 “Spanish flu,” when as many as 100 million people died in sixteen weeks (Barry et al., 2004). Subsequent epidemics of influenza infection caused by influenza type A including H1N1 have cause hundreds of thousands of excess deaths, and the ongoing epizootic of avian influenza (H5N1) is considered by some microbiologists and epidemiologists to be a “pandemic-in-waiting” (Taubenberger and Morens, 2010). The need for effective treatment and intervention is pressing.

As with many RNA viruses, influenza mutates rapidly, leading to immune escape, where the virus mutates to evade recognition by the host immune system, and rapid development of resistance against vaccines and antivirals. Effective therapies will consider rapid viral evolution and remain potent. Treatment development can be aided by increasing understanding of viral replication to provide insight into influenza viruses’ mechanism of infection and to outline potential drug targets.

The influenza polymerase is ubiquitous in flu replicative processes, conducting both transcription and translation, and therefore provides an attractive platform for pharmaceutical intervention. The polymerase is composed of three distinct subunits and interacts with several host factors which mediate dramatic conformational changes and help determine polymerase function. This thesis focuses on endonuclease, an enzyme contained in one of the polymerase subunits which is crucial for normal influenza

replication, with the aim of better understanding its role in different conformations of the polymerase.

Endonuclease activity is not required in all forms of the influenza polymerase, and the specifics of its regulation and interactions are not entirely clear. I hypothesized that endonuclease binding to other subunits driven by the need to regulate endonuclease activity contributes to stabilization of alternative polymerase conformations. By increasing understanding of endonuclease binding partners and activity, I hope to contribute to a more complete picture of influenza replication and to elucidate potential treatment avenues.

Influenza Viruses

Influenza is a virus in the family Orthomyxoviridae which causes acute respiratory infection in humans and many other animals, notably swine and fowl. There are four strains of influenza viruses, types A, B, C, and D, though only types A and B, which are responsible for circulating seasonal epidemics, pose a substantial public health threat. Influenza type C causes a mild, self-limited infection in humans and type D primarily infects cattle. Influenza type A is associated with zoonotic outbreaks and has substantial pandemic potential; influenza A viruses were responsible for the 1918 Spanish flu pandemic and for the 2009 H1N1 “swine flu” and are considered one of the great uncontrolled illnesses of the 21st century.

Influenza viruses are enveloped RNA viruses; the virus is composed of a lipid bilayer studded with hemagglutinin (HA) and neuraminidase (NA) surface proteins which envelops the viral genetic material, viral polymerase (FluPol), and viral

nucleoprotein (NP) (Neverov et al., 2015; Ivanova et al., 2015). Influenza A virus subtypes are identified by their surface proteins (for example, the Spanish flu is subtype H5N1 and the so-called swine flu, type H1N1), which mutate rapidly, especially under selection pressure from antiviral therapies (Garjani et al., 2023; Neverov et al., 2015). The high mutation rate observed in RNA viruses in general makes vaccination and treatment more difficult because the virus quickly acquires resistance (Sanjuán et al., 2010). In the case of influenza A, the consequence of rapid acquisition of resistance and immense diversity of viral subtypes can be seen in the need for annual flu vaccines, which are designed every year from predictions of the viral strains likely to be circulating.

The World Health organization estimates 1 billion human cases of influenza annually, with the most substantial burden of mortality in the elderly (>65 years old) and in the very young (<5 years old), especially in developing countries. Though vaccination is an extremely impactful public health intervention and substantially decreases odds of hospitalization and death from influenza infection, the existing trivalent vaccine does not confer long-lasting immunity, especially in aging populations (Liao 2022). Further, the rapid mutation rate of RNA viruses has made development of successful antivirals difficult, and recent meta-analysis of antivirals in treatment of seasonal influenza found very low certainty in the efficacy of the drugs assessed (Gao 2024). The combination of relatively short-term vaccine protection, minimal antiviral treatment options, and pandemic potential makes influenza an extremely pressing public health concern.

Increasing understanding of the viral life cycle could offer insight into potential treatment avenues which target aspects of the virus with lower propensity for mutation.

The central aim of this thesis was to study the structure and potential interactions of the PA endonuclease, which is a region of the influenza polymerase (FluPol) and a key enzyme in the influenza life cycle. The influenza endonuclease has previously been considered an interesting target for antiviral intervention and several endonuclease inhibitors have been discovered and even approved as therapies for influenza infection. However, rapid resistance tends to evolve in influenza viruses and several of these compounds, including Baloxavir which was used as a control endonuclease inhibitor in this thesis, have little evidence for meaningful reduction of morbidity and mortality. I sought to improve the picture of the PA endonuclease's role in influenza polymerase (FluPol) function as a whole, with the hope that it would provide some insight into the mechanism of FluPol conformational changes.

FluPol is a heterotrimeric polymerase responsible for both transcription – the conversion of viral RNA (vRNA) to messenger RNA (mRNA), and replication – the conversion of mRNA to complementary (cRNA) and then vRNA, of viral genetic material. The intrinsic flexibility and regions of disorder of different FluPol subunits contributes to its ability to act as both transcriptase and replicase. However, these same areas of flexibility often lack resolution when FluPol is visualised by structural techniques such as Cryo-EM and x-ray crystallography. Nuclear magnetic resonance spectroscopy (NMR) allows for high resolution investigation of flexible regions and for *in situ* titration of potential binding partners and analysis of their effect on the target protein. In this thesis I used NMR and functional methods to assess the structure and activity of PA endonuclease and its potential binding partners.

Viral Life Cycle

FluPol is responsible for both transcription and replication of viral RNA and is able to perform different functions depending on cellular conditions and interactions with host factors (Krischuns et al., 2024). Increasing understanding of the conformational changes that allow FluPol to serve such diverse and crucial roles in replication of the influenza viral genome offers insight into viral machinery and potential treatment targets.

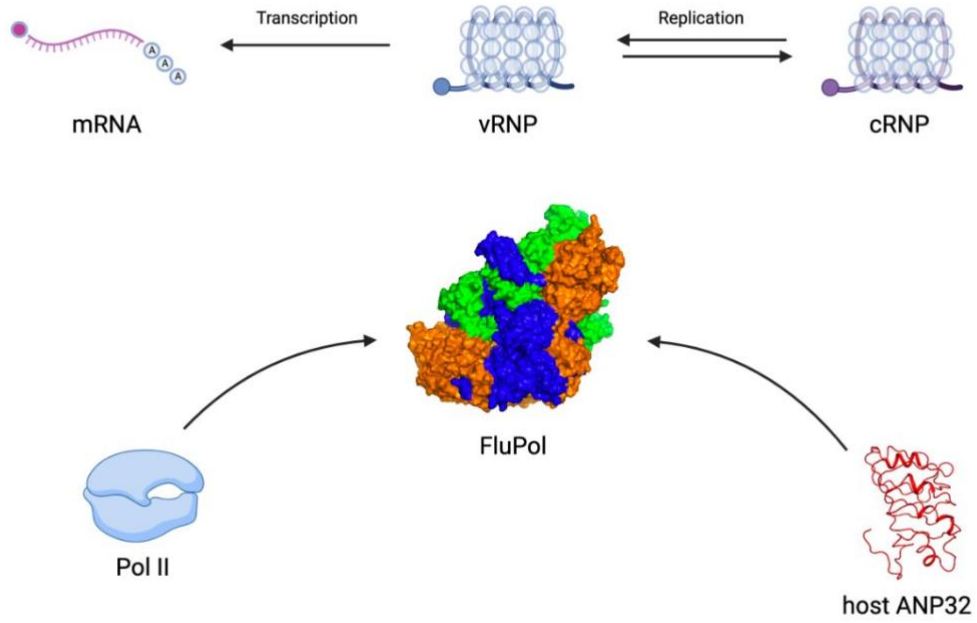
The influenza viral genome is encoded into 8 segments of single-stranded, negative-sense RNA, each of which is individually transcribed and replicated during the viral life cycle (Carrique et al., 2020). During influenza infection, the virus first enters the host cell by binding to sialic acid on the cell surface and entering through the host plasma membrane, most commonly through clathrin-mediated endocytosis (Ni et al., 2024; Sieczkarski & Whittaker, 2002). After entry, the virus is trafficked to the nucleus where the C-terminus of the FluPol PA subunit binds the C-terminus of host PolII (Staller et al., 2024). The PolII/FluPol association supports the transcription-competent conformation of FluPol and facilitates viral usage of host polymerase products and capabilities, including production of nascent mRNAs (Engelhardt et al., 2005; Krischuns et al., 2022). After sufficient mRNA transcription by the FluPol – PolII complex, FluPol dissociates from PolII and adopts a slightly different conformation, including repacking of the endonuclease against the PB1 subunit to limit access to RNA, which allows for replication.

During replication, FluPol first produces complementary RNA (cRNA) from the template mRNAs produced during transcription, then progeny vRNA from the cRNA template. In both cases, the nascent replicated RNA is pseudo-circularized through the binding of a single FluPol to the 5’ and 3’ end of the RNA segment (Figure 1a) (Zhu et

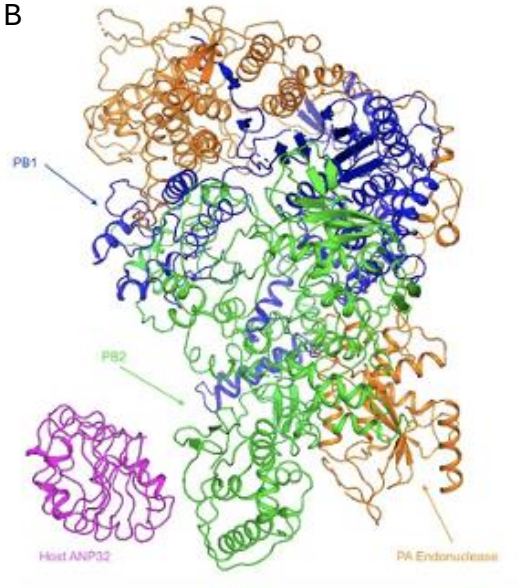
al., 2023). The remaining RNA in the segment is bundled around influenza protein ‘NP’ at a rate of approximately one NP per 25 bases (Knight et al., 2021); NP then oligomerizes to form the backbone of ribonucleoprotein complexes (RNP), which prevent degradation (Ye et al., 2006). After replication of all eight segments of the viral genome and assembly of viral RNP (vRNP) complexes, the viral genetic material is packaged with NP and FluPols into progeny virions which are then trafficked out of the cell to continue the cycle of infection.

FluPol was initially observed as a monomeric heterotrimer characterized by a stable polymerase core with surrounding peripheral domains (Figure 1b). More recently, it was confirmed that FluPol dimerizes upon interaction with host acidic nuclear phosphoprotein (ANP32) in order to perform replication (Figure 2). This is a sophisticated replicative platform where nascent cRNAs or vRNAs are suggested to move directly from one polymerase in the dimer (called the replicating polymerase, or $FluPol_R$) towards the second polymerase, called the encapsidating polymerase ($FluPol_E$). $FluPol_E$ is able to shield newly produced viral genetic material from detection by the host immune system and facilitates the rapid formation of ribonucleoprotein complexes. Though $FluPol_R$ in the dimer is structurally very similar to the replicase and monomeric transcriptase previously solved, $FluPol_E$ exhibits an entirely novel conformation, where the structure, function, and interactions especially of flexible peripheral regions may be quite different from previously solved structures (Figure 2).

A



B



C

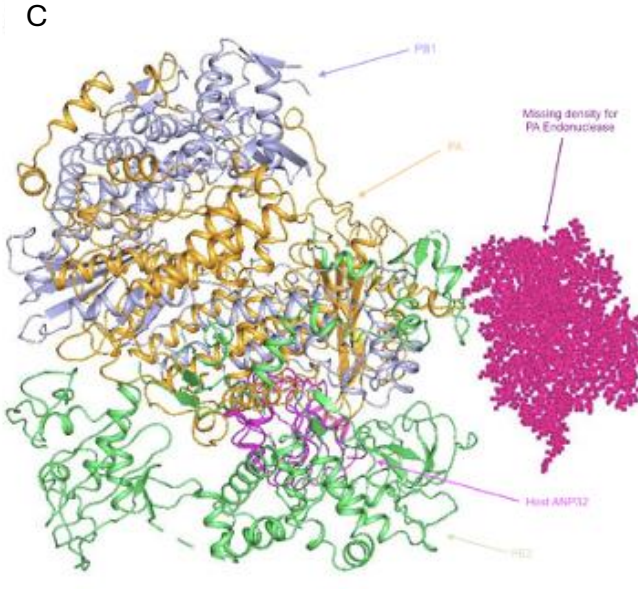


Figure 1. Influenza virus life cycle and structure of the influenza virus RNA polymerase. *A. Characterization of major processes in the influenza life cycle involving FluPol. Interaction between*

host Pol II and FluPol facilitates mRNA synthesis while dimerization of FluPols mediated by host ANP32 allows for formation of the replicative platform and subsequent replication of the viral genome. Created with BioRender.com B. PyMol figure showing a structure of the influenza A FluPol$_R$, with the PA subunit and PA endonuclease indicated in orange, PB1 subunit in blue, PB2 subunit in green, and host ANP32 used in crystallization and crucial for dimerization shown in pink. C. PyMol figure showing a structure of influenza type A FluPol$_E$, with the PA subunit and PA endonuclease indicated in light orange, PB1 subunit in light blue, PB2 subunit in light green, and host ANP32 shown in pink. Missing density for PA endonuclease indicated in magenta. All structural credit: Staller et al., 2024, PDB 8R1L.

Structure and Function of the Influenza A Polymerase

The overall polymerase structure for influenza types A, B, and C have been solved via Cryo-EM and have significant similarities. This thesis focuses on influenza type A because of its endemic and pandemic potential in humans and will therefore focus on the structure and sequence of the influenza A polymerase, though the body of research into influenza types B and types C polymerases is helpful and often referenced. FluPol is a heterotrimer composed of three subunits of relatively similar sizes, called Polymerase Basic 1 (PB1), Polymerase Basic 2 (PB2), and Polymerase Acidic (PA). In the monomeric transcriptase and replicase forms of the polymerase, the PB1 subunit, PA C-terminal domain, and PB2 N-terminal domain have substantial stabilizing surface contacts which form the structured core of the FluPol heterotrimer (Figure 2). The polymerase catalytic site is formed by interaction between PB1 and PB2-N within this structured core, and conformational changes in the remaining, flexible polymerase dictate function (Figure 1b) (Zhu et al., 2023).

The interaction between the PB2 N-terminus and PB1 forms a canonical cavity in the influenza polymerase which contains the FluPol active site and allows for entry of template genetic material and exit of nascent product. This conformation is observed in both the transcription and replication conformations of FluPol. The N-terminal PA

endonuclease and C-terminal region of the PB2 subunit, including the nuclear localization signal (NLS) and pathogenicity-determining 627 domain, are more flexible. Structural changes in these flexible peripheral domains inform changes in polymerase structure and function and facilitate interaction with host factors. It has been hypothesized that conformational changes within FluPol mediated by an internal or external molecular switch determines the functional state of the polymerase (Zhu et al., 2020).

Polymerase Dimer Formation

Interaction with host factors is a crucial component of FluPol function and differences in host factor identity between host species plays a key role in virulence of influenza as a zoonotic pathogen. Oligomerization of FluPols had long been suggested (Thierry et al., 2016) as a potential conformational change necessary for replication; the formation of an influenza polymerase dimer was confirmed in influenza type C by Carrique et al. in 2020. Dimer formation is facilitated by interaction with host ANP32.

ANP32 proteins belong to a family of highly evolutionary conserved proteins with diverse functions in the host species (Yu et al., 2022). ANP32 proteins are composed of a structured, N-terminal, leucine rich region (LRR) and a long C-terminal tail called the low-complexity acidic region (LCAR) which appears flexible in Cryo-EM maps. Though the LCAR is not directly involved with the FluPol dimer, it appears to enrich the surrounding area with NP, potentially facilitating RNP assembly (Wang et al., 2022). The LRR appears to interact directly with both the encapsidating and replicating polymerases in the dimer, albeit with different regions, and allows FluPols to dimerize.

FluPol dimer formation is mediated by interaction with host ANP32, and the success of influenza viruses in interspecies infection depends predominantly on the virus' ability bind ANP32 from different hosts. Avian influenza A is generally not able to dimerize with mammalian ANP32 due to differences in the host factors, but a point mutation at residue 627 of the PB2 subunit allows influenza to replicate in both humans and fowl (Zhang et al., 2019). ANP32-mediated dimer formation therefore has substantial implications for the human pandemic potential of influenza A viruses, as it allows for assembly of the replication platform. This thesis sought to investigate other interactions contributing to dimer formation to increase understanding of the flu replicative process and of potential intervention points.

FluPol Dimer Structure

Dimerization of FluPols to form a replicative platform has been confirmed in influenza types A, B, and C (Staller et al., 2024; Arragain et al., 2024; Carrique et al.). Staller et al.'s 2024 results are, to the best of my knowledge, the first published structure of a polymerase dimer for influenza A with successfully resolved density for the endonuclease (Figure 2). They found that in influenza A, the $FluPol_R$ of the polymerase dimer adopts a conformation similar to that of the cap-snatching incompetent replicase, where endonuclease is packed against the PB1 C-terminal/PB2 terminal helical bundle to prevent enzymatic activity.

The structure of the encapsidating FluPol, as observed in the ANP32-mediated FluPol dimer which forms the replicative platform, involves radical repacking of the three FluPol subunits. Cryo-EM structures suggest that formation of FluPolE is stabilized

primarily by rearrangement of PB2 subdomains. The PB2 midlink and cap-binding domains (CBD) shift away from packing against the PA C-terminal domain in the FluPol globular core, with the PB2 CBD shifting to interact with the PA endonuclease, which rotates 70 degrees to secure the PB2 lid domain between the endonuclease and the PB2 CBD, stabilising the overall FluPolE conformation (Fifugre NEW). The PA C terminal region of FluPolE is then free to interact with the PB2 subunit of FluPolR, forming extensive surface contacts between the two polymerases of the dimer. The PA C-terminal region of $FluPol_R$ then packs against the LRR of host ANP32 (Figure 2a). The endonuclease-PB2 interaction in FluPolE is suspected to be necessary for formation and stabilization of the encapsidase conformation (Staller et al., 2024).

Structures very similar to the dimer $FluPol_R$ have been observed monomerically, both alone and in complex with ANP32, but the encapsidating conformation of FluPol has not been observed outside of the replicative platform. ANP32 interacts with the 627 domain of $FluPol_E$ but encapsidase formation is primarily dictated by the interaction between the PB2 subunit and the PA endonuclease, with the PB2 lid subdomain flipping over entirely from its replicase positioning to hold the endonuclease against the PB2 cap-binding domain (Figure 2) (Thierry et al., 2016). Stabilization of the encapsidase is mediated by the endonuclease-PB2 interaction, and the PA subunit of $FluPol_E$ then interacts with the PB2 subunit of $FluPol_R$ through a substantial surface interface to complete assembly of the replicative platform (Figure 2).

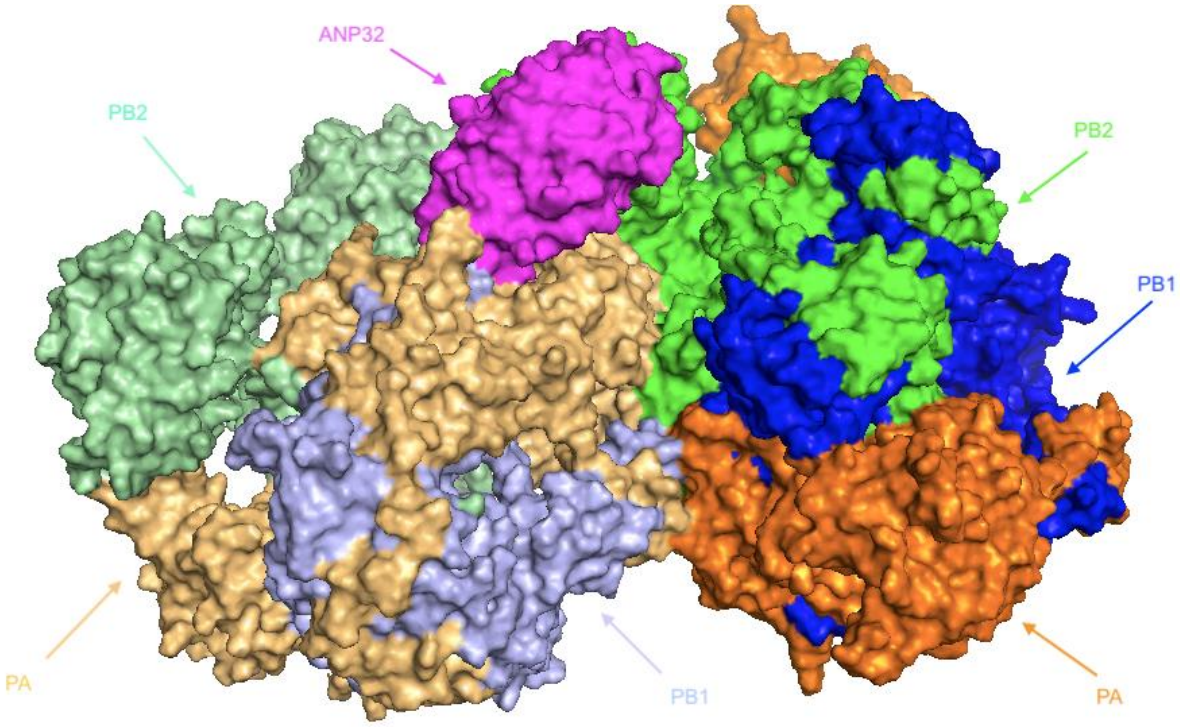
The dimerization of FluPols effectively extends the protective exit channel observed in monomeric $FluPol_R$. In the dimer, nascent vRNAs or cRNAs need to be trafficked directly from $FluPol_R$ towards $FluPol_E$ for efficient assembly into RNPs. Newly

synthesized RNA therefore avoids exposure to the host immune machinery and is at reduced risk of degradation. Functional studies have assessed that formation of the dimeric replication platform is necessary for replication (Carrique et al., 2020).

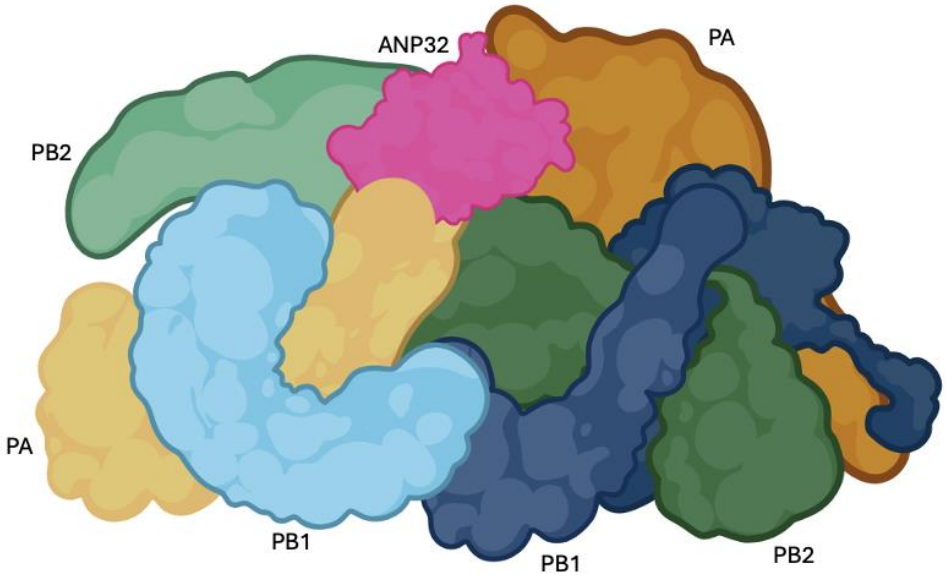
Though there is strong evidence for the stabilization of $FluPol_E$ through interactions between the PA endonuclease and the PB2 subunit, the $FluPol_E$ endonuclease is difficult to resolve via Cryo-EM because of its flexibility (Figure 1c) (Staller et al., 2024). The importance of the endonuclease-PB2 interaction was highlighted by Thierry et al. in 2016, when the authors observed an entirely novel PB2 conformation, now associated with the encapsidase, upon co-crystallization with endonuclease. These data suggest an interaction between PB2 and the PA endonuclease which does not depend on stabilization from other elements of the FluPol dimer.

Given the importance of endonuclease to the influenza life cycle, I sought to better understand its structural role in the dimer and in assembly of the replicative complex.

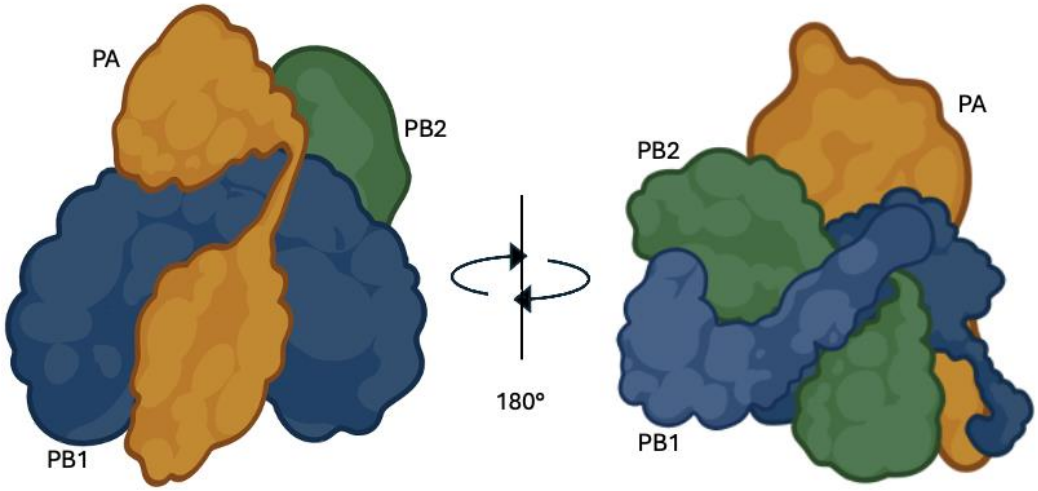
A



B



C



D

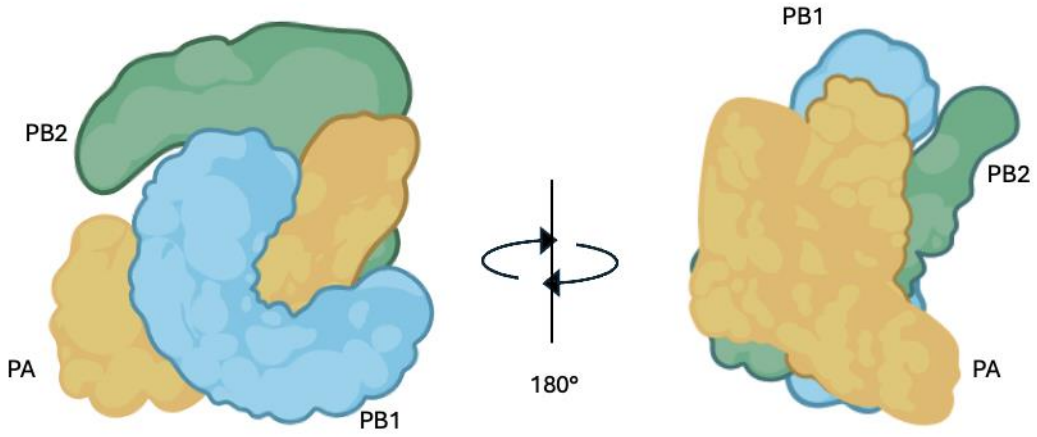


Figure 2. Dimerization of influenza type A polymerases to form the replicative platform and cartoon schematics of interactions between the polymerase subunits. *A. Surface representation of the FluPol dimer with ANP32 (bright pink), the encapsidating FluPol at left (PB1 light blue, PB2 light green, PA light orange), and replicating FluPol at right (PB1 blue, PB2 green, PA orange) (PDB 8R1L) B. Cartoon representation of the FluPol dimer, ANP32 (bright pink), the encapsidating FluPol at left (PB1 light blue, PB2 light green, PA light orange), and replicating FluPol at right (PB1 blue, PB2 green, PA orange). C. Cartoon representation of $FluPol_R$ (PB1 blue, PB2 green, PA orange). D. Cartoon representation of $FluPol_R$ (PB1 light blue, PB2 light green, PA light orange).*

Cap-snatching

During the influenza life cycle the polymerase first facilitates production of template mRNA (transcription); once sufficient levels of mRNA are produced in the host cell, the polymerase generates cRNA, which is finally replicated into product vRNA to form novel virions (Carrique et al., 2020). Replication is initiated without the need for an external primer (Shapiro & Krug, 1988). Instead, replication is initiated *de novo* on the terminal 3' residues of the template strand with stabilization by a priming loop present in the polymerase active site (te Velthius et al, 2021).

Unlike replication, transcription cannot proceed without an mRNA cap. Influenza viruses lack the ability to generate their own mRNA caps, instead relying upon a process called cap-snatching. The 5' mRNA cap is composed of a 7-methylguanosine residue linked via a triphosphate bridge to the 5' nucleotide of the RNA. The virus cleaves the first 10-13 residues of host mRNAs which are then used by the influenza virus as a primer. The 5' mRNA cap prevents the rapid identification and degradation of nascent viral mRNA by the host immune system (Reich et al., 2014). During cap-snatching, the mRNA binding region of the PB2 subunit (DuBois et al., 2012) binds nascent host mRNAs and positions them towards the PA endonuclease. Endonuclease then cleaves, or "snatches," the host-generated mRNA cap. The PB2 subunit rotates back by 70 degrees to position the "snatched" cap in the active site of the polymerase, at which point transcription can begin (te Velthuis et al., 2021). Figure 3 is included as a demonstration of the relative size and positionings of endonuclease and the PB2 subunit where PB2 is bound to RNA. However, it should be noted that the RNA-bound structure shown in Figure 3 is solved for influenza type C in the cap-snatching incompetent replicase form of the polymerase, and should therefore be considered only

for insight into the approximate size and structure of the endonuclease relative to the PB2 subunit.

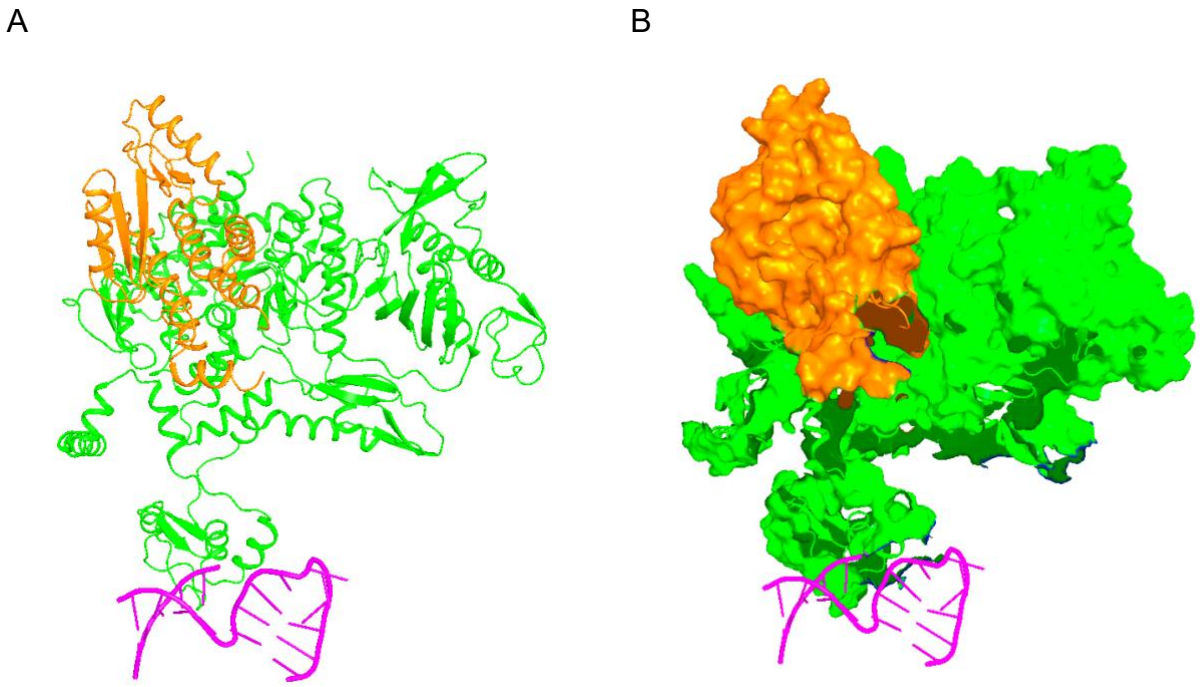


Figure 3. Structure of PA endonuclease and RNA-bound PB2 in FluPol$_R$ of influenza type C. *A. Cartoon representation with endonuclease (orange), PB2 (green), and RNA (pink). B. Surface representation with endonuclease (orange), PB2 (green), and RNA (pink). PDB 6XZR*

Transcription would not be able to proceed without the catalytic activity of the PA endonuclease, but, as noted above, cap-snatching is not required in the replicative platform. Endonuclease is a promiscuous enzyme and existing studies of endonuclease activity exploring changes in catalytic activity upon interaction with potential binding partners, namely the PB2 subunit of FluPol (Thierry et al., 2016), have not observed substantial regulation of enzymatic activity, even when crystal structures strongly indicate interaction. Instead, it is suggested that enzymatic activity is regulated in the replicase by burying the endonuclease to restrict substrate access. The endonuclease in the cap-snatching incompetent replicase has substantial interaction with a helical

bundle formed at the PB1 C-terminus and PB2 N-terminus, which likely limits endonuclease activity (Thierry et al., 2016). It should be noted that, while there are significant regions of conserved structure between influenza types A, B, and C, there are notable differences in sequence and structure in some regions, including the endonuclease, and these distinctions are sufficient to warrant type-specific investigation.

Interestingly, though the replicative platform generated by FluPol dimerization should not require cap-snatching competency in the replicating or the encapsidating polymerases, the PA endonuclease is not buried against the PB1 subunit in $FluPol_E$. Rather, in the substantial rearrangement of FluPol domains first reported in influenza type B by Thierry et al. in 2016, endonuclease interacted with a smaller interface of the PB1/PB2 helical bundle and much more extensively with the PB2 C-terminus. These data were corroborated by subsequent structural investigations of the dimer structure in influenza type A (Staller et al., 2024). Though endonuclease could not be resolved in the monomeric FluPol-ANP32 complex, $FluPol_E$ in the polymerase dimer displayed a substantial conformational change with the PA endonuclease sandwiched between the PB2 lid and PB2 cap binding domains. In particular, the novel positioning of endonuclease in the encapsidase begs the question of whether endonuclease could be acting as a conformational switch – the dynamics of enzymatic regulation could stabilize different conformations of the polymerase.

We hypothesized that if endonuclease were acting as a conformational switch, introduction of peptides from the PB1 of PB2 subunit which interact closely with endonuclease in one conformation of the polymerase could induce substantial

rearrangement of endonuclease from the apo state to the replicase or encapsidase conformation. After initial experiments introducing peptides into solution with endonuclease and seeing small shifts in the endonuclease structure, I concluded that titration of relatively small (18-22 residue) peptides was not enough to trigger the substantial conformational changes observed in endonuclease and FluPol as a whole between the replicating and encapsidating states. Instead, I identified a region of interest in the endonuclease from my early experiments and continued to probe the interactions between endonuclease and potential binding partners in the dimerized replicative platform.

Potential Binding Partners

Evidence for binding between endonuclease and the PB1 and PB2 subunits within FluPol is strongly supported by a large body of research detailed above. Externally, it has been suggested that NP may interact with endonuclease in $FluPol_R$ (Vidic et al., 2016). Increased density around the encapsidating endonuclease has been observed in cryo-EM trials, suspected to be NP. Since NP is a key component of vRNP and cellular RNP (cRNP) construction post-replication (Turrell et al., 2014), it has been suggested that interaction between NP and FluPol may recruit NP to the polymerase dimer for use in vRNP and cRNP complexes (Szeto et al., 2020). Alternatively, it has been posited that NP may stabilize endonuclease for interaction with PB2, which could mediate repacking of PB2 against endonuclease in the encapsidating polymerase conformation (E. Fodor, personal communication).

Introduction to NMR

Existing structures of the PA endonuclease have been solved primarily by x-ray crystallography and Cryo-EM (Thierry et al., 2016; Zhu et al., 2023; Staller et al., 2024). However, the existence of flexible, unstructured regions likely associated with functional sites in the endonuclease necessitated truncation for successful crystallization, and in some cases the endonuclease either could not be resolved or was resolved to a lower resolution than canonically structured regions (Zhu et al., 2023; Staller et al., 2024). The existing structures for the endonuclease are therefore limited by the techniques which have been used thus far in FluPol analysis, and NMR was selected as a visualization technique to facilitate analysis of the unstructured regions of the endonuclease and their potential interactions.

NMR has been used for detailed characterization of small molecules with a range of real world applications (Brinson, 2021). Use of NMR for protein analysis allows real time, solution state visualization at the atomic level (Puthenveetil & Vinogradova, 2019), including providing detailed insight into unstructured regions which often appear as low-density or cannot be resolved at all using crystallization-based visualization techniques.

NMR measures and facilitates differentiation between the specific properties of individual atoms in the target protein by relying upon the fundamental biophysical property of atomic spin. When placed in a strong magnetic field and excited with radio-frequency pulses, atoms respond to the perturbation and, over time, ‘relax’ back to equilibrium (Bruker, 2025). Minute differences in atomic response are affected by the biochemical properties of the individual atom and by its immediate environment, including close neighboring atoms, allowing detailed analysis of the two and three-dimensional structure of the selected protein.

However, ideal protein conditions for NMR analysis are severely limited, with spectral quality often dependent on small size as well as long-term stability in low-pH and at relatively high temperatures. Size was a particularly notable limiting factor in NMR analysis for this thesis since the PA endonuclease (27kDa including the His-tag) falls at the upper range of NMR-suitable protein sizes. As protein size increases, line broadening and overlapping signals make it difficult to differentiate atoms or reliably identify next residues during assignment (Fernandez & Wider, 2003). This challenge can be managed in some part by introducing isotopic sample labeling with 15N and 13C, which allows for multiple reference dimensions for comparison. However, spectral quality is still significantly influenced by size, and even with double labeling, solution of protein structures beyond ~25kDa is challenging (Puthenveetil & Vinogradova, 2019).

Transverse relaxation-optimized spectroscopy (TROSY) experiments were introduced to improve the quality of spectra recorded for larger proteins by using constructive interference to improve signal quality (Salzmann et al., 1998). BEST-TROSY (BTROSY) experiments were used throughout this thesis in collection of two-dimensional spectra, comparing 1H and 15N signals, for improved sensitivity and increased time efficiency (Solyom et al., 2013).

Thesis Aims

Interactions and regulation of the influenza PA endonuclease present a viable pathway for the regulation of FluPol function and viral replication (Omoto et al., 2018). In the transcriptase, endonuclease is enzymatically active and facilitates the generation of nascent viral mRNA by performing cap-snatching. In the replicase form of the

polymerase, where cap-snatching is not necessary because replication initiation proceeds in a *de novo* manner, endonuclease activity is regulated by interaction with the PB1 C-terminal/PB2 N-terminal helical bundle. This interaction suggests that regulation may be necessary to prevent promiscuous enzymatic activity. The substantial interface between the PB1 C-terminal/PB2 N-terminal bundle and endonuclease does not exist in $FluPol_E$, perhaps indicating that another interaction contributes to the regulation of enzymatic activity. From these observations and existing research, I hypothesized that endonuclease would undergo structural changes to bind PB1, PB2, and NP individually in interactions important to the structure of the broader endonuclease; conformation of these interfaces may offer targets for antiviral intervention (Thierry et al., 2016; Staller et al., 2024; Vidic et al., 2016). I further posited that interactions with external binding partners would alter endonuclease activity. Based upon these hypotheses, the specific aims of this thesis were threefold:

1. Visualize the influenza type A PA endonuclease in its solution state using NMR, providing a foundation for the subsequent research undertaken in this thesis.
2. Confirm the interaction between endonuclease and likely binding partners in the hopes of specifically identifying binding regions which could be targeted to interfere with successful assembly of FluPol.
3. Assess whether structural changes as a result of ligand binding regulate enzymatic activity, elucidating viral methods of endonuclease regulation.

Chapter 2: Purification, Validation, and NMR Characterization of PA endonuclease

The PA Endonuclease

The PA endonuclease domain comprises the N-terminal ~195 residues of the PA subunit of FluPol and facilitates the unique cap-snatching capabilities of the influenza polymerase (Doan et al., 1999). Initially, it was thought that the endonuclease domain was in either of the basic subunits, PB1 or PB2 (Shi et al., 1995), but the enzyme was later found to be a domain of the PA subunit by Dias et al. (2009). Endonuclease belongs to a family of type II metal-binding restriction enzymes, which require divalent cations for activity (Dias et al., 2009). Some debate exists about the metal most likely bound to the endonuclease active site *in vitro*, with studies reporting the highest levels of endonuclease activity in the presence of manganese and cobalt, but magnesium considered potentially more biologically relevant due to its significantly higher concentration in the cell (DuBois et al., 2012). Manganese ions are paramagnetic and result in significant line broadening of the NMR signals arising from nuclei within ~10Å (Schnell, personal communication). The presence of manganese is therefore prohibitive for NMR studies of the endonuclease and magnesium was used instead for stabilization and elucidation of the metal-bound form of the enzyme.

An N-terminally GST-tagged construct of the PA endonuclease comprising the first 198 residues of the influenza A PA subunit was generously provided by Ervin Fodor. The construct was cloned into an N-terminally His-tagged pET15b vector and purified via gravity flow immobilized metal ion affinity chromatography before being isotopically (^{13}C, ^{15}N) labeled for characterization by NMR spectroscopy. NMR was

useful in visualizing endonuclease because it is not restricted by regions of disorder and was particularly relevant to this thesis because it facilitates observation of what were expected to be relatively weak interactions between endonuclease and potential binding partners.

The backbone amide resonances were assigned to >85% and validated using TALOS-N (Shen & Bax, 2015). While structures of the influenza endonuclease, alone and in the polymerase trimer, were previously determined by X-ray crystallography and cryo-EM (Dias et al., 2009; Carrique et al., 2020; Staller et al., 2024), however studies of the solution state of endonuclease by NMR have not previously been reported. NMR enables high resolution studies of the native sequence, including any flexible loops, which may also provide information on the overall flexibility of the endonuclease within the larger FluPol complex. Previous structural studies visualizing the PA-endonuclease by cryo-EM and s-ray crystallography noted that these techniques might favor one endonuclease structure over another, since endonuclease interactions in the polymerase are quite varied (Thierry et al., 2016). It was therefore important to identify the apo (RNA-free) state of endonuclease by NMR to compare these spectra to those of endonuclease with potential binding partners.

Purification

Initial purification of endonuclease was conducted by gravity flow over Glutathione Sepharose 4B resin using an existing protocol provided by the Fodor lab in the Dunn School of Pathology at the University of Oxford. While the purified protein appeared to be free from contaminants based upon gel electrophoresis analysis (Figure 4b), the yield from this purification protocol was relatively low (<1mg/L of culture). In an

effort to increase yield, I cloned the endonuclease into another vector resulting in an N-terminally His-tagged construct (pET15b).

After attempts to purify the His-tagged construct with immobilized metal ion affinity chromatography and size-exclusion chromatography using the Äkta HisTrap and Superdex 75 Increase columns (Figure 4c), I determined that gravity flow gave the highest yield with sufficient protein purity (Figure 4d). Successful purification of the endonuclease was confirmed by gel electrophoresis and yield was consistently greater than 10mg/liter of labeled culture.

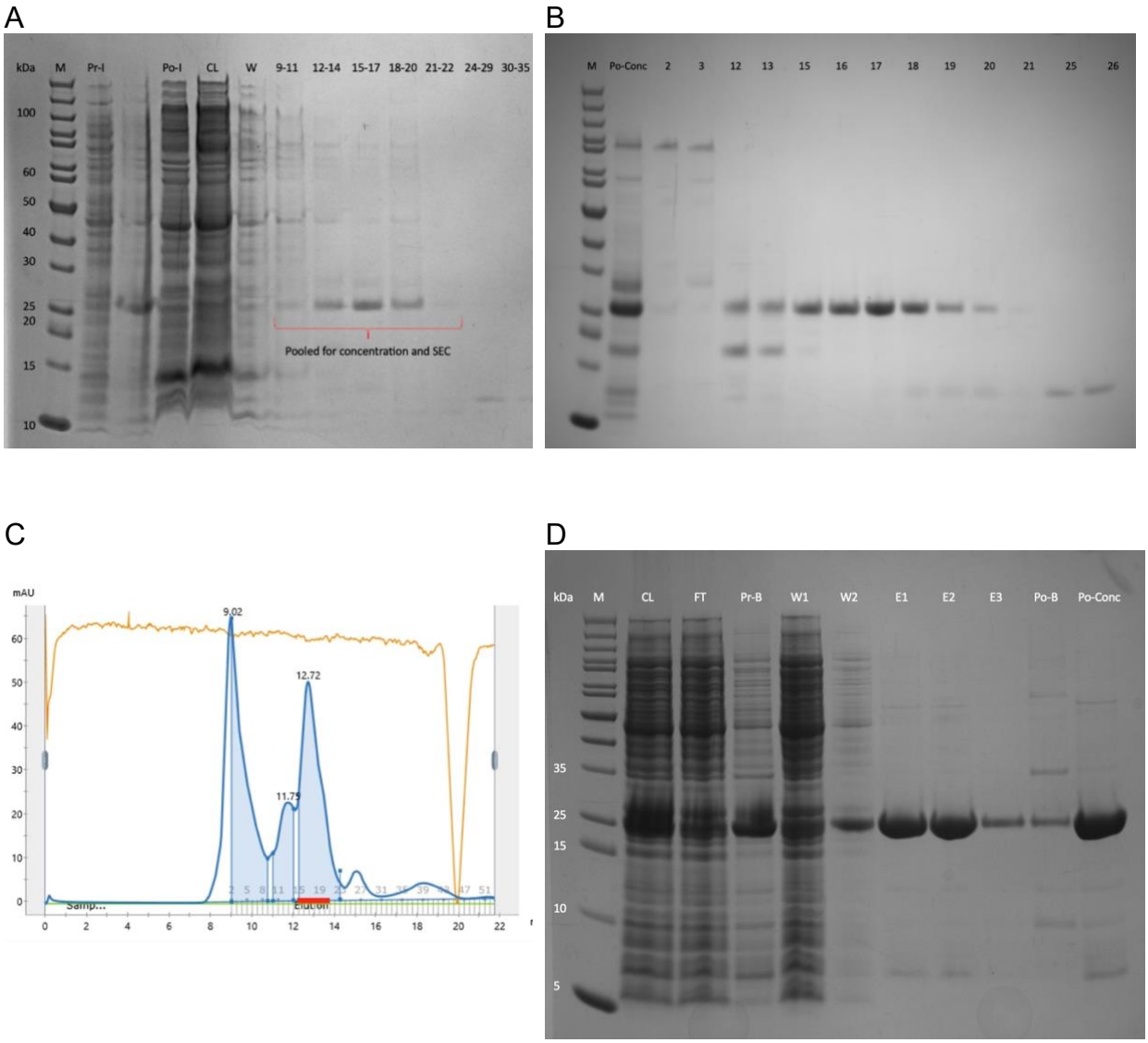


Figure 4. Purification of His-tagged endonuclease. *A. HisTrap purification of endonuclease with lanes pooled for Size Exclusion Chromatography (SEC) marked. B. Post-SEC purified endonuclease. C. Superdex 75 Increase Äkta trace corresponding to the gel shown in B. Fractions considered pure and concentrated are marked in red along the x-axis. D. Purification by gravity flow without subsequent SEC, yielding a significantly higher final concentration of protein (~1mM, 17mg per liter of cell culture in unlabeled purifications).*

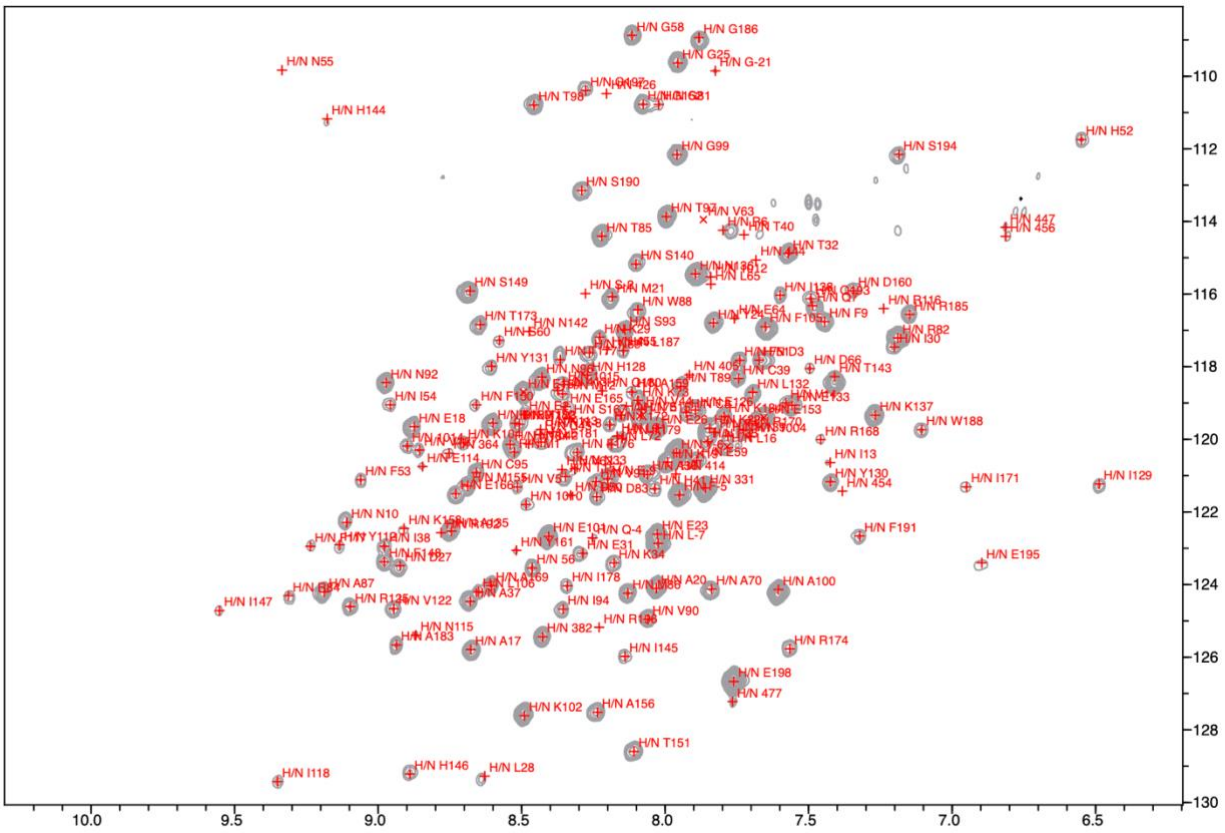
Characterization via NMR

The NMR signal intensities of protein amide groups typically increase under increasingly acidic solution conditions due to slower rates of ^{1}H exchange between the protein amides and water. Samples of ^{15}N labelled endonuclease at three pH values

were tested to optimize spectral quality and the protein aggregated and became insoluble at pH 6.5. The poor solubility at low pH is likely related to endonuclease's theoretical pI, which is 5.26. To increase protein stability and longevity, I conducted all NMR experiments at the purification pH 7.5, unless otherwise noted. I compensated for the decreased signal to noise resulting from high sample pH by using samples at a concentrations up to 600 μM. Further experiments tested the stability of endonuclease with or without metal bound and found qualitatively that the inclusion of magnesium salts in the sample buffer resulted in more stable samples.

2D backbone amide NMR experiments (^{1}H,^{15}N BEST-TROSY) conducted at pH 7.5 and 25°C were well-dispersed, consistent with a well-folded protein with stable secondary and tertiary structure. After confirmation of successful endonuclease purification via gel electrophoresis and 2D NMR check experiments, I isotopically labeled the protein with ^{15}N and ^{13}C to conduct 3D triple-resonance NMR experiments and assign the backbone amide chemical shift resonances. The experiments used for assignment were: HNCA, HNCO, HN(CA)CO, HNCACB, CBCA(CO)NH, and NOESY. After assigning backbone amides to 85% (Figure 5a), TALOS-N was used to assess the likelihood of the predicted assignments, with the prediction summary indicated high confidence in the assignments provided (Figure 6a). TALOS-N secondary structure predictions were then compared to an existing structure of the endonuclease in the influenza type A polymerase heterotrimer (PDB 8R1L) and found to be similar in the assigned residues, supporting the validity of the assignments.

A



B

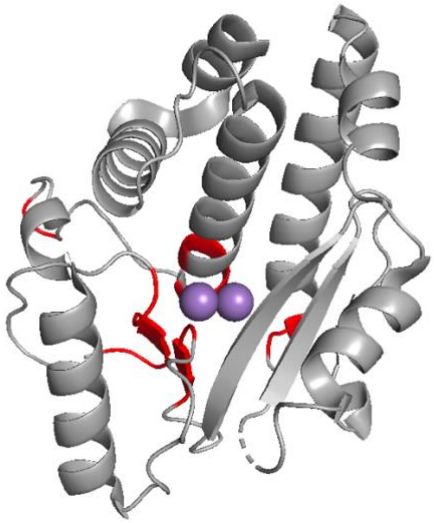
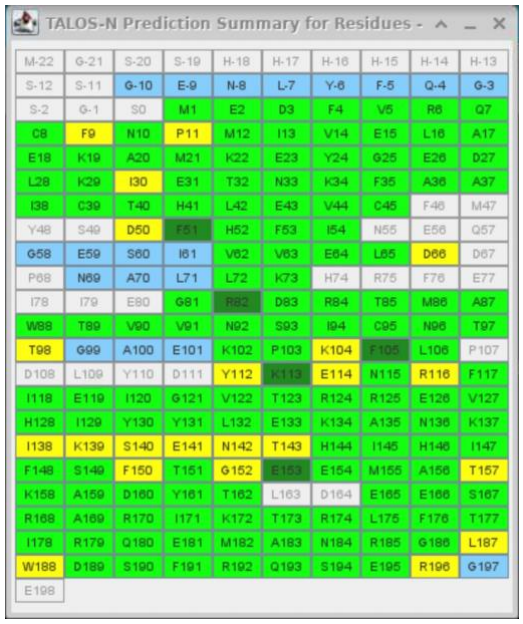
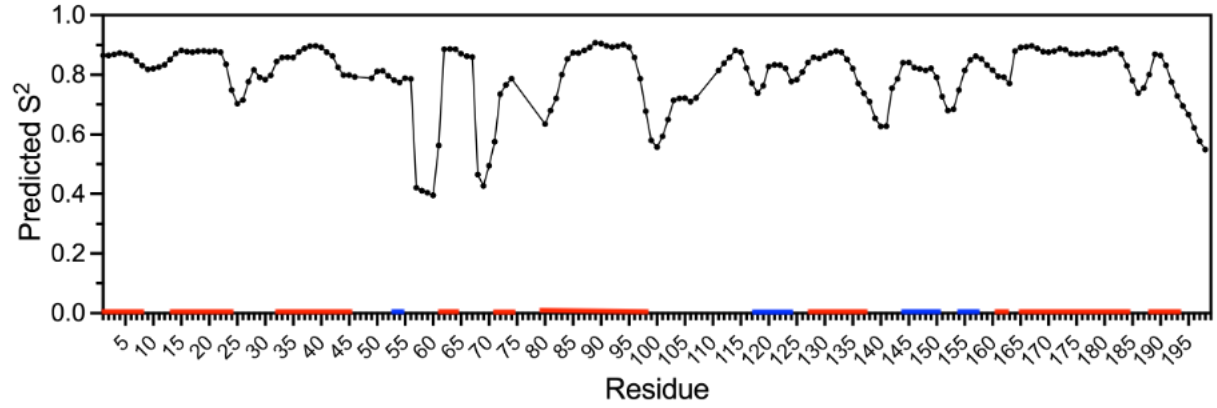


Figure 5. PA endonuclease assignment. *A. Projection of the 1H,15N plane of the 3D HNCA spectrum of ^{13}C, ^{15}N-labelled PA endonuclease with assignments indicated, x-axis (^{1}H, ppm), y-axis (^{15}N, ppm),. B. X-ray crystallography solved structure of endonuclease and Mn^{2+} (purple) with unassigned residues noted in red (PDB 2W69). Unassigned residues (colored red) were primarily confined to the core of the protein, rather than the peripheral flexible regions of interest.*

A



B



C

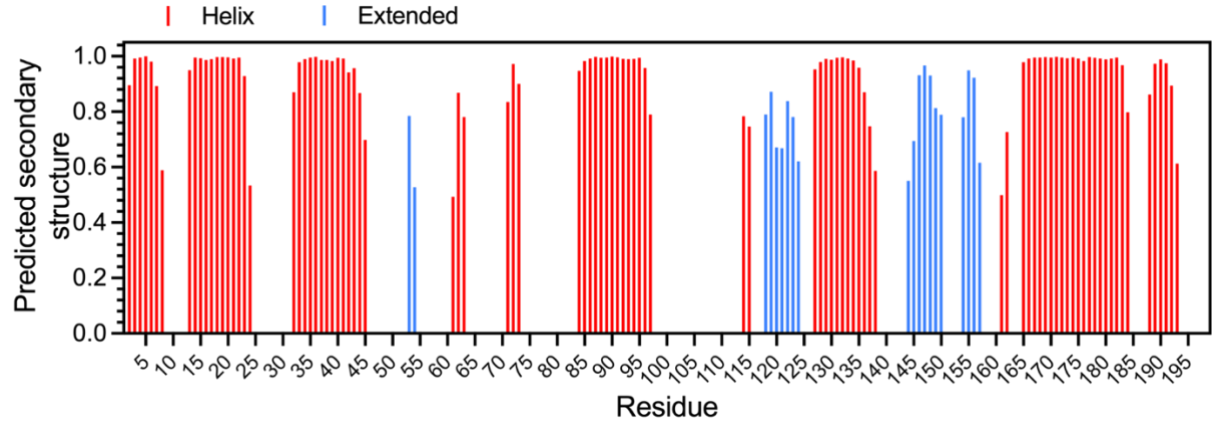
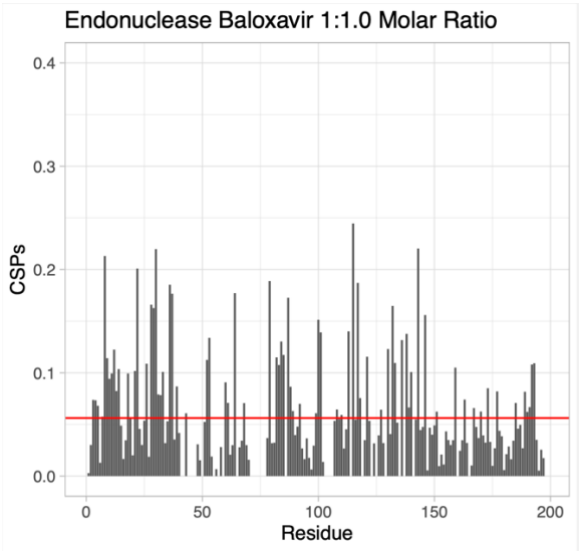


Figure 6. TALOS-N predicted flexibility and secondary structure from assigned ^{15}N and ^{13}C chemical shifts. *A. Prediction summary based upon provided chemical shift information, with gray indicating missing information (residue unassigned), blue indicating highly dynamic residues, yellow*

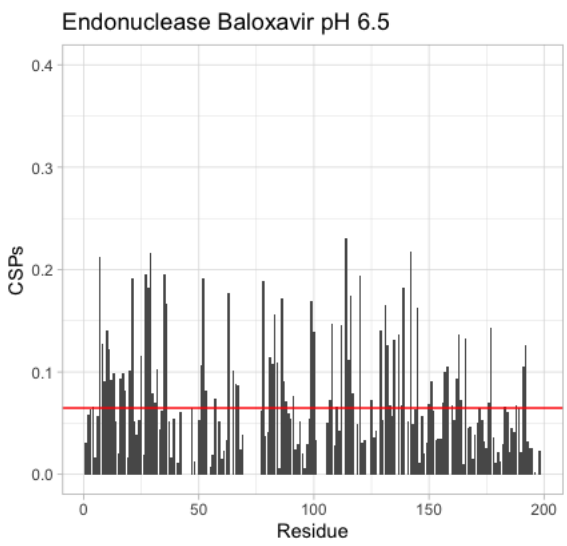
suggesting dynamic residues, and green indicating high a high confidence prediction. B. Predicted RCI S^2 values, with predicted helical regions indicated on the x-axis in red, and predicted beta-sheet regions indicated on the x-axis in blue. C. Predicted secondary structure. Predicted alpha-helices shown in red and predicted beta-sheets shown in blue, with the y-axis indicating the prediction confidence.

Having assigned the protein, a Baloxavir screen was conducted to stabilize the endonuclease at a lower pH to enable further functional screens and to confirm protein identity. Baloxavir is an antiviral which binds tightly to the endonuclease to inhibit cap-snatching capabilities (Todd et al., 2021). Significant chemical shift perturbations (CSP) were observed upon titration of Baloxavir into endonuclease (Figure 7), corroborating the identity and native conformation of the purified protein. Since Baloxavir binds tightly to endonuclease, the peaks in the NMR spectra could not be tracked reliably from one titration point to the next, so a “minimal CSP,” which measured the minimum movement of each peak after progressive addition of the ligand was calculated (Figure 7a, b). I attempted to use Baloxavir-bound, ^{13}C-^{15}N labeled endonuclease sample to increase the quality of my data for assignment but due to the slow ligand exchange, I was unable to transfer peak shifts back to the unbound state with sufficient confidence.

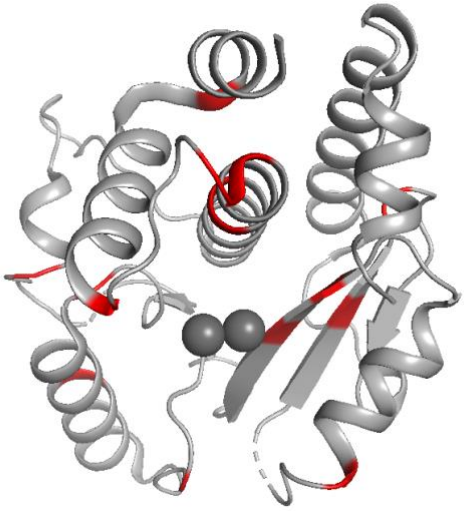
A


Endonuclease Baloxavir 1:1.0 Molar Ratio
CSPs
0.0
0.1
0.2
0.3
0.4
0
50
100
150
200
Residue

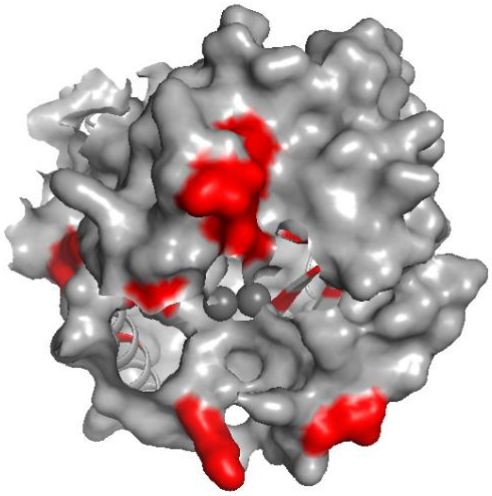
B


Endonuclease Baloxavir pH 6.5
CSPs
0.0
0.1
0.2
0.3
0.4
0
50
100
150
200
Residue

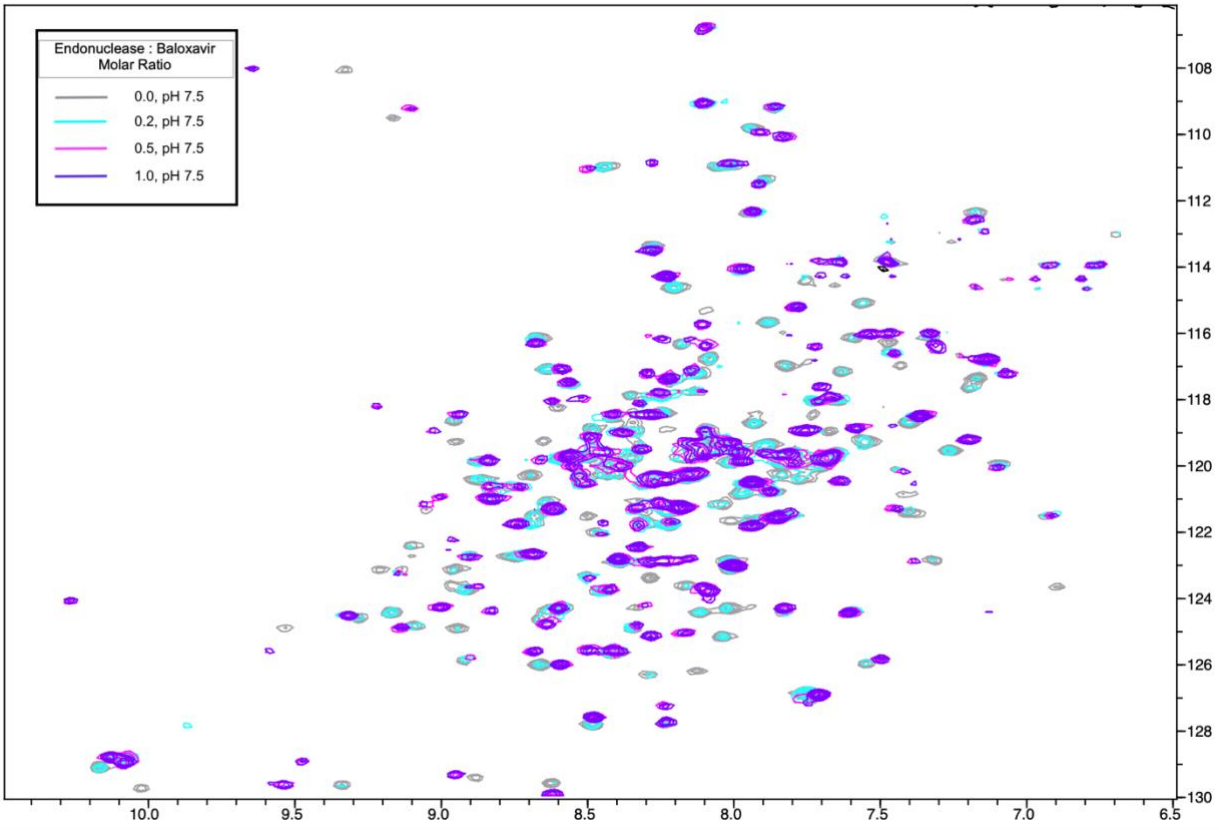
C



D



E


Endonuclease : Baloxavir
Molar Ratio
0.0, pH 7.5
0.2, pH 7.5
0.5, pH 7.5
1.0, pH 7.5
108
110
112
114
116
118
120
122
124
126
128
130
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5

F

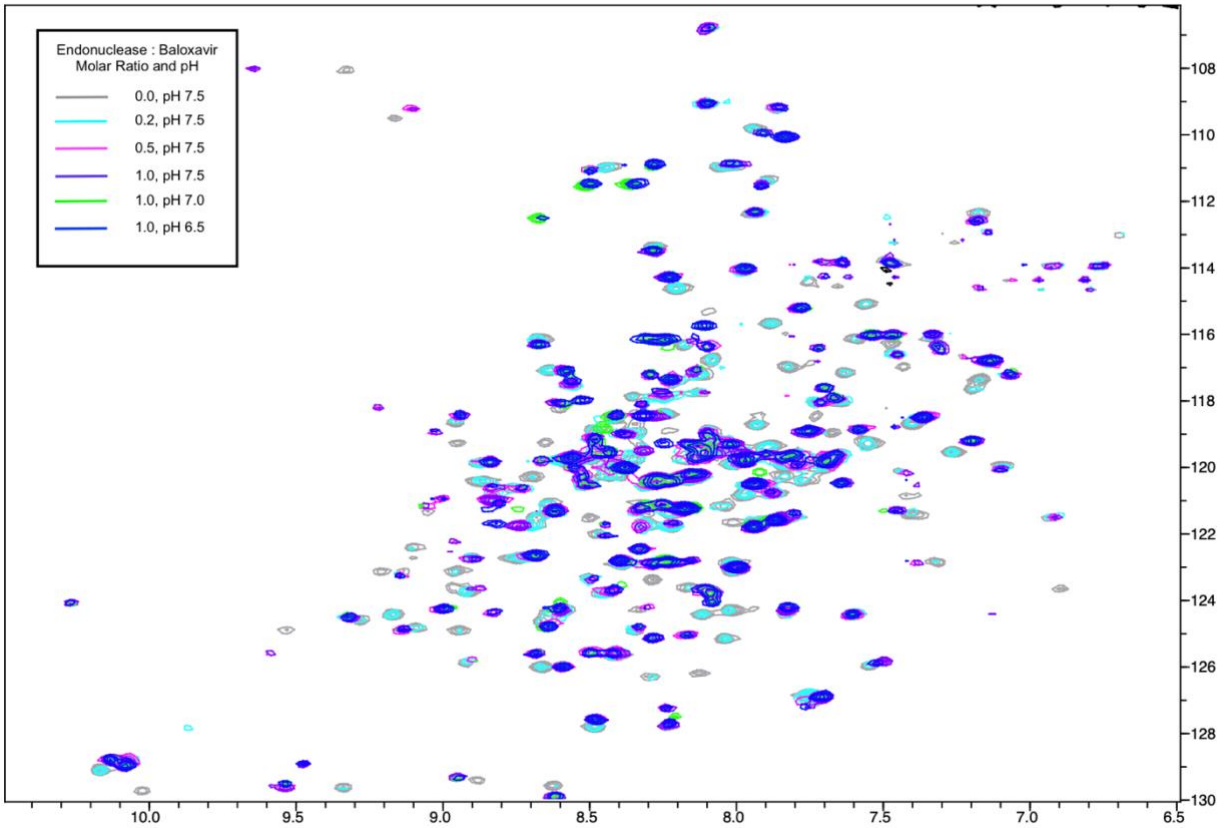


Figure 7. Endonuclease-Baloxavir titrations. *A. Chemical shift perturbations (ppm) of endonuclease with a 1.0:1.0 molar ratio of target to ligand at pH 7.5 as compared to the Baloxavir-free spectra, with the mean chemical shift perturbation indicated in red. B. Chemical shift perturbations of endonuclease with a 1.0:1.0 molar ratio of target to ligand at pH 6.5. The mean chemical shift perturbation is indicated in red. C. PyMol cartoon structure of endonuclease with minimal CSPs from equimolar Baloxavir mapped onto the structure. All CSPs over 0.15 indicated in red and all other residues shown in grey. D. PyMol surface representation of endonuclease with minimal CSPs from equimolar Baloxavir mapped onto the surface. All CSPs over 0.15 indicated in red and all other residues shown in grey. E. Overlay of endonuclease ^{1}H, ^{15}N BEST-TROSY spectra, x-axis (^{1}H, ppm), y-axis (^{15}N, ppm), with increasing concentrations of Baloxavir or decreasing pH. Control spectrum (gray); 1.0:0.2 molar ratio of endonuclease to Baloxavir, pH 7.5 (turquoise); 1.0:0.5 molar ratio, pH 7.5 (pink); 1.0:1.0 molar ratio, pH 7.5 (purple). F. Overlay of endonuclease ^{1}H, ^{15}N BEST-TROSY spectra, x-axis (^{1}H, ppm), y-axis (^{15}N, ppm), with increasing concentrations of Baloxavir or decreasing pH. Control spectrum (gray); 1.0:0.2 molar ratio of endonuclease to Baloxavir, pH 7.5 (turquoise); 1.0:0.5 molar ratio, pH 7.5 (pink); 1.0:1.0 molar ratio, pH 7.5 (purple) 1.0:1.0 molar ratio, pH 7.0 (green), 1.0:1.0 molar ratio, pH 6.5 (blue).*

Discussion and Next Steps

These data can be used to support successful purification of endonuclease with structural components similar to those of previously solved structures. More importantly, the results from purification and visualization of endonuclease in its solution state via NMR

provided the framework for further experimentation. Ideally, I would have been able to confidently assign endonuclease to >95%, which have been achieved by replacing non-exchangeable ^{1}H spins with deuterium (^{2}H) in order to remove the strong ^{1}H contribution to dipole relaxation, in order to slow relaxation and sharpen the resonances. However, this project was limited in time and expense, and the unassigned regions of the protein, which were primarily confined to the core, structured region (Figure 5b), were assessed as likely not interacting with potential binding partners.

It should be noted that the residues which were not assigned also correspond closely to the divalent cation binding site, and could therefore have been the result of obfuscation of signal due to the paramagnetic effects on manganese binding. No chelating step was used in the endonuclease purification protocol, so manganese binding during expression would not have been replaced. As has been established, endonuclease has a high affinity for Mn^{2+} and therefore the molar excess of magnesium added to the dialysis buffer, which may have displaced other, less-tight binders, may not have replaced magnesium from the active site. Further experimentation with the aim of more complete assignment of the protein could add a chelating step to determine whether signal in the core of the protein is better with higher confidence of magnesium binding.

Successful purification and visualization of the PA endonuclease by NMR accomplished the first goal of this thesis: to give what is, to my knowledge, the first solution state characterization of predicted secondary structure based on chemical shifts given of the enzyme. Further, NMR spectra of the free, wild-type endonuclease established

a baseline to which spectra of ligand-bound and mutant endonucleases could be compared.

Chapter 3: NP and Peptide Interactions

Potential Endonuclease Binding Partners

Published structures of the influenza A polymerase dimer reveal substantial changes in conformation and surface interactions between the three subunits (Staller et al., 2024). Notably, there is a radical repacking of the PB2 domain against the PA subunit evident in the encapsidating form of the endonuclease that has not been observed outside of the polymerase dimer (Thierry et al., 2016; Zhu et al. 2023). Since the polymerase dimer seems to function for replication rather than transcription, it is assumed that neither the replicase nor the encapsidase form of the dimer is cap-snatching competent (Staller et al., 2024). Endonuclease's role, especially in the encapsidase, is not clear. I sought to investigate interactions between endonuclease and potential binding partners, as identified by analysis of published structures, in an effort to elucidate changes in endonuclease structure and function and interactions between different subunits of the influenza A polymerase.

We used existing structures obtained through Cryo-EM and x-ray crystallography to select peptides from other influenza subunits that seemed to have relevant surface contacts with endonuclease (Staller et al., 2024; Thierry et al., 2016). AlphaFold 3 models of these peptides with endonuclease were generated to see whether high-confidence interactions were predicted, and unlabeled peptides were titrated into ^{15}N-labelled wild-type endonuclease to observe any chemical shift perturbations (CSPs) that might be attributable to peptide binding.

Putative endonuclease interactions with NP

The assembly of the FluPol dimer is mediated by ANP32, a host factor composed of a LRR, which interacts with the polymerases in the dimer, and a C-terminal LCAR. The LCAR of ANP32 is proposed to play a role in recruiting NP to the polymerase complex, presumably to facilitate assembly of vRNPs and cRNPs from the nascent RNA produced by the FluPols (Wang et al., 2022). Increased density in cryo-EM maps around the endonuclease of $FluPol_R$ have been proposed to be NP, perhaps acting to stabilize the interaction between endonuclease and PB2 or perhaps as a transient interaction before assembly of the RNP complex (E. Fodor, personal communication). Interaction between the C-terminal domain of the PA subunit in influenza A is supported by surface plasmon resonance and cryo-EM experiments (Moeller et al., 2021; Vidic et al., 2016). There is strong evidence from pull-downs for interaction between NP and endonuclease in influenza type C, but pull-downs between influenza type A endonuclease and NP are less indicative of an interaction (E. Fodor, personal communication).

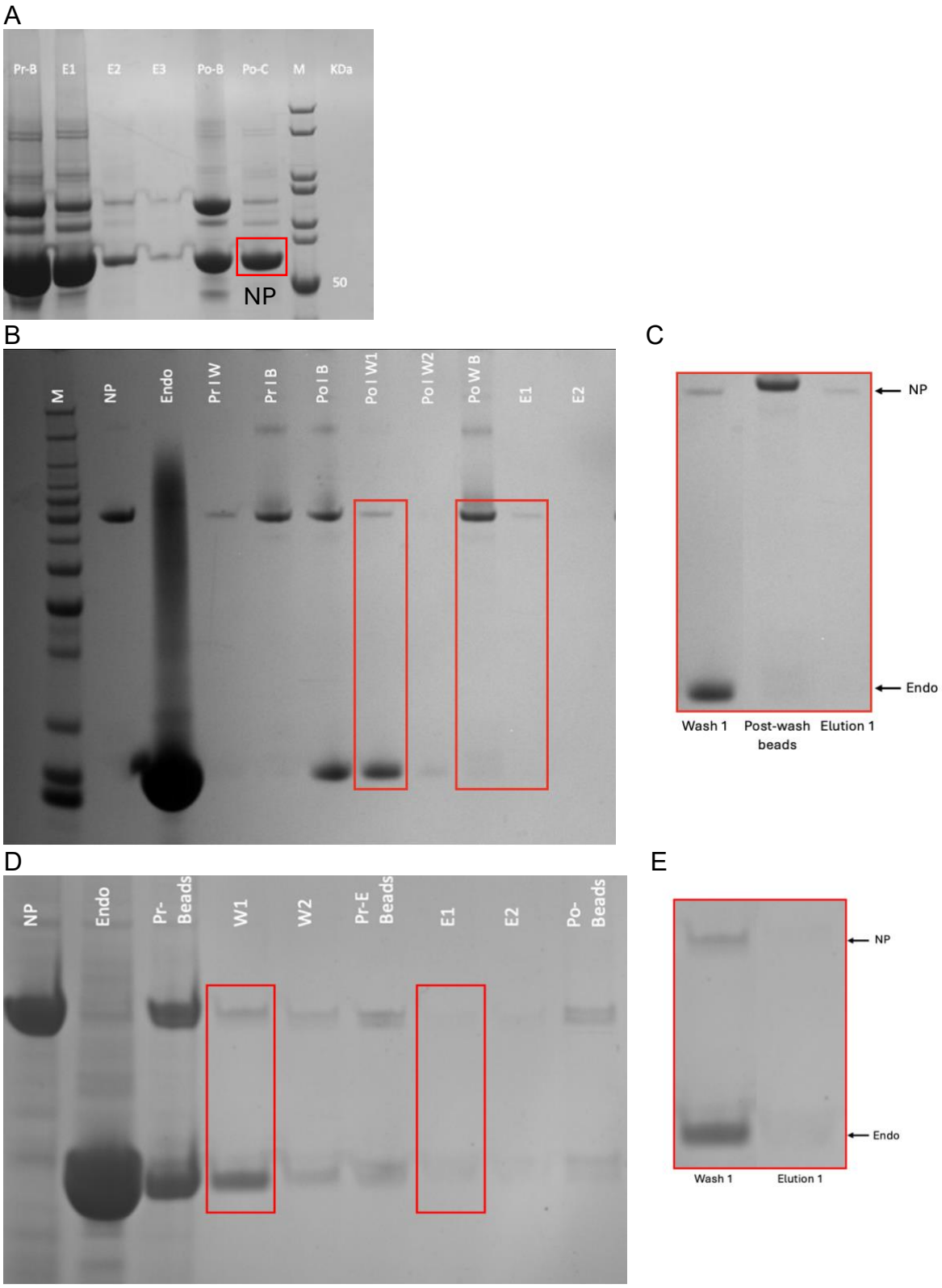


Figure 8. NP R416A Purification and pull-downs with endonuclease. *A. Gravity flow affinity purification of GST-tagged R416A NP. Clear lysate was incubated with equilibrated Glutathione Sepharose 4B Beads (Cytiva). Beads were washed, then incubated with 40uL HRV 3C overnight. Flow through containing purified protein was collected and concentrated for further applications (Samples*

E1, E2, E3, Po-Conc). Samples of the resin after washing (Sample Pr-B) and after elution (Sample Po-B) indicated incomplete elution of protein off of the resin. B. 0.3 mg purified, GST-tagged NP(Sample NP) was added to equilibrated Glutathione Sepharose 4B beads (Cytiva) and washed. The wash and pre-incubation beads were collected (Samples Pr I W, Pr I B) and allowed to incubate for 2 hours at *4˚C on the column before 0.2 mg purified WT endonuclease (Sample Endo) was applied to the column. The beads were sampled again (Sample Po I B). The column was incubated a further two hours, then washed (Samples Po I W1, Po I W2), then eluted with 25 mM reduced glutathione. Eluate and post-elution beads were sampled (Samples E1, E2, Po-B) C. NP and Endonuclease pull-down. 3.3 mg purified GST-cleaved NP (Sample NP) were combined with 4.6 mg PA endonuclease (Sample Endo) and incubated overnight at 4˚C. This sample was applied to 500uL bead volume of equilibrated Nickel beads and the beads were washed and sampled (Samples W1, W2, Pr-Beads) before being incubated for 2 hours at 4˚C. A pre-elution sample of the washed beads was collected (Sample Pr-E Beads) before the protein was eluted using 300 mM imidazole (Samples E1, E2). A final sample of the post-elution beads was collected (Sample Po-Beads).*

NP R416A, used throughout this thesis because of its low propensity for oligomerization, was purified and used in a pull-down with endonuclease to assess potential binding (Figure 8). In Figure 8b, control samples of NP and undiluted endonuclease alone were run to confirm the position they migrate to on the gel. The control endonuclease gel sample was run at a much higher concentration (600 μM, or 13.8 mg/mL) than the pull-down gel samples, resulting in substantial swelling in the corresponding band on the gel. Some NP washed off the resin before introduction of the endonuclease sample. The affinity of NP for endonuclease was then assessed based upon comparison to the post-incubation beads, labeled “Po I B” in Figure 8b. Most of the added endonuclease appeared to wash off the beads before incubation with the elution buffer, though a faint band remains on the post-wash beads ("Po W B") and post-elution beads ("Po B").

NP came off of the resin in greater concentration in the wash than the elution, making it difficult to discern whether interaction was occurring, and the post-elution beads sample (Po-B) show a significant amount of NP remaining on the beads. This result suggests that GST-tagged NP bound strongly to the beads and informed the later

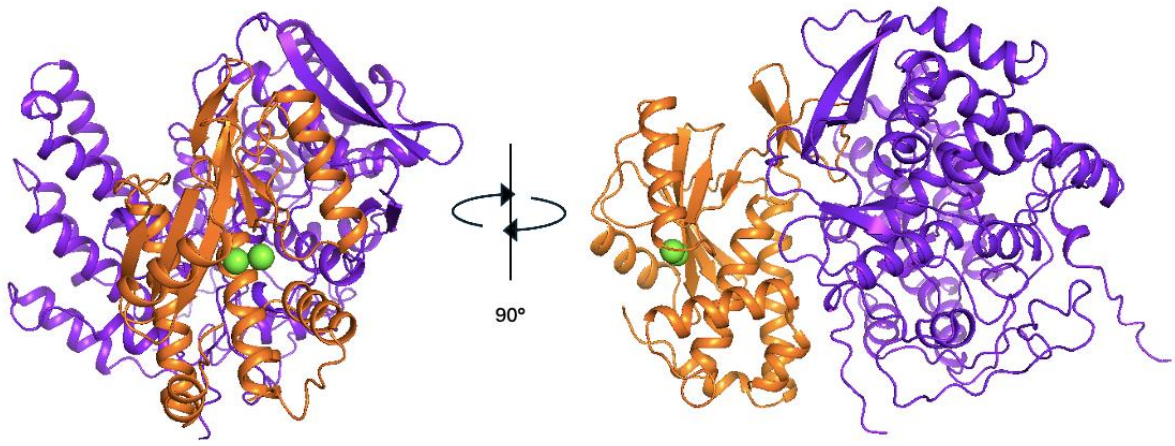
change in NP-purification protocol to cleave the protein off the beads before eluting GST-cleaved protein for further investigation. When taken with the high volume of GST-cleaved protein remaining on the glutathione resin as shown in Figure 8a, this is compelling evidence that NP binds strongly and non-specifically to GST beads. In an effort to mediate the effects of this binding on the pull-down result, a repeat pull-down using His-endonuclease as the bait was conducted (Figure 8c), again with inconclusive results.

Had there been a strong interaction between endonuclease and NP, one would have expected to see very little to no bait (resin-bound) protein washing out in the wash buffer, with increased concentrations of both endonuclease and NP eluting off of the beads after the introduction of reduced glutathione or imidazole in the elution buffer. In contrast, complete washing out of the target (unbound) protein during the wash stages followed by clean elution of the bait protein would have been evidence against interaction. Instead, both NP and endonuclease were observed in the wash buffer and the elution buffer in both pull-down experiments, albeit faintly. Since there also appeared to be both NP and endonuclease left on the resin in both cases, especially in Figure 8c, where endonuclease was used as the bait protein, these results were considered very weakly suggestive of interaction between endonuclease and NP. However, the pulldown method was judged to be overall ineffectual in this case because no conclusive result could be determined, highlighting the need for further experimentation using NMR.

Before NMR experiments to test the interaction between NP and endonuclease, AlphaFold3 was used to predict potential interaction points (Figure 9). The predicted

error, provided in AlphaFold 3 in angstroms, was used to assess confidence in the provided prediction, where a lower error in expected position equated to higher likelihood that the predicted positioning was correct (Figure 9). Unsurprisingly, given the suspected transient nature of the interaction between endonuclease and NP *in vivo,* the confidence of the predicted structure was very low.

A



B

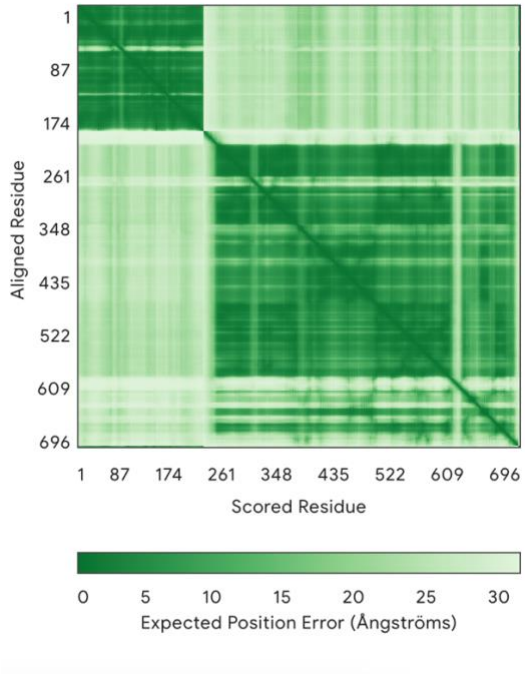


Figure 9. AlphaFold3 predictions of NP R416A interaction with endonuclease. *A. AlphaFold 3-generated structures showing proposed interaction between magnesium-bound endonuclease and R416A NP, with endonuclease (orange), NP (purple), and manganese (grey). B. AlphaFold 3-generated graph of position and expected error, used to assess the confidence of the prediction.*

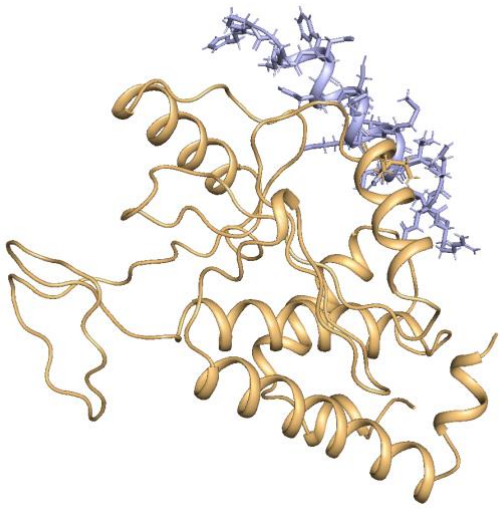
Design of PB1 Peptide

Published cryo-EM structures of the influenza A polymerase complex in the replicase and transcriptase forms show extensive surface contacts between the PA endonuclease and PB1 (Reich et al., 2014; Fan et al, 2019; Carrique et al., 2020). In the transcription-competent form of the polymerase, where cap-snatching is necessary for the synthesis of mRNA, the active site of the endonuclease is exposed and available for substrate access. In the replicase, however, packing of the endonuclease against the PB1 C-terminal/PB2 N-terminal helix restricts RNA access for the endonuclease and therefore apparently prevents promiscuous enzymatic activity (Figure 1b). Since the interactions of endonuclease in both of these conformations are relatively well investigated, I sought to better understand endonuclease interactions in $FluPol_E$.

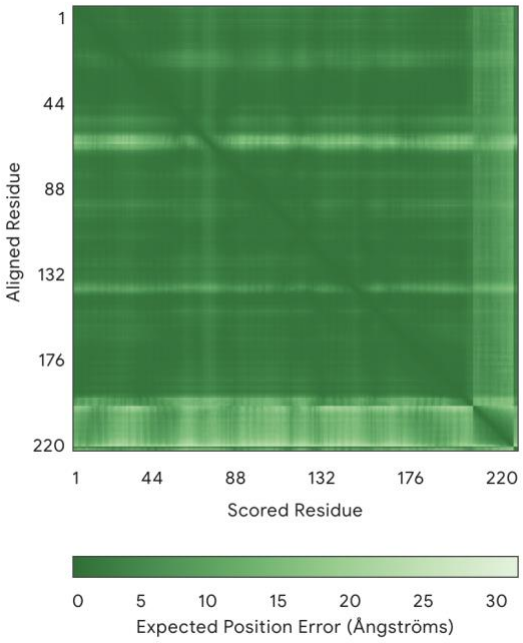
Where endonuclease has been successfully resolved in the $FluPol_E$ polymerase of the dimer, the encapsidase conformation appears to be anchored by the interaction between endonuclease and the PB2 subunit, where the PB2 lid subdomain and cap-binding domain sandwich the endonuclease to stabilize the encapsidase (Staller et al., 2024). However, the peripheral interactions of endonuclease with the PB1 subunit are also altered; potential binding of a PB1 peptide was therefore tested with the hypothesis that PB1/endonuclease interactions could also help stabilize the encapsidase. The S100-T115 region of PB1 was selected for investigation based upon a published influenza A FluPol dimer structure (8R1L; Staller et al., 2024), where the encapsidating endonuclease appears to have surface contacts with the region between residues 100-115 in PB1 (Figure 1c, Figure 10). In addition to selection based upon the solved structures, AlphaFold3 was used to assess the likely binding pattern of the selected

peptides (Figure 10). The predicted interactions were not high confidence, though there was relatively high confidence in the predicted structure of the selected peptide (Figure 10).

A



B



C

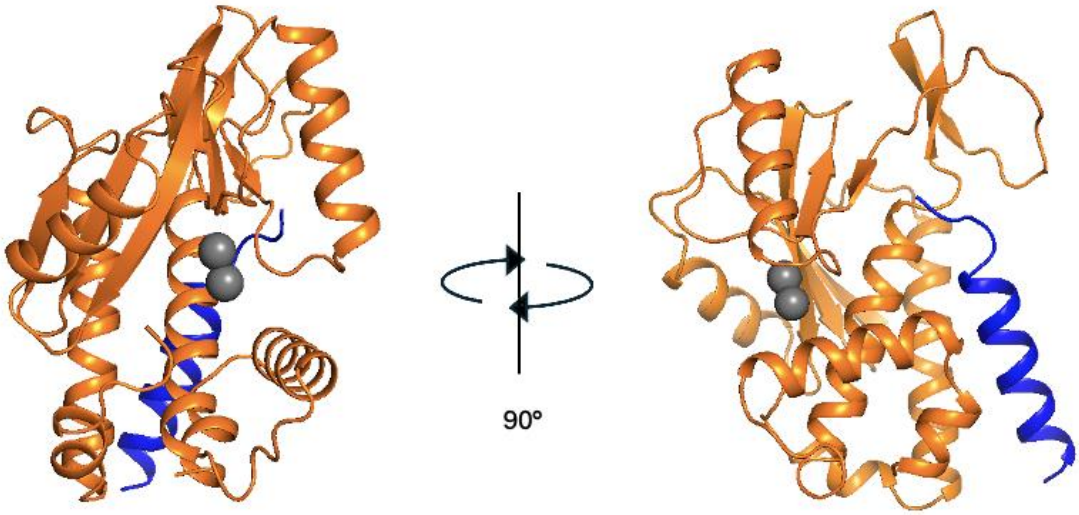


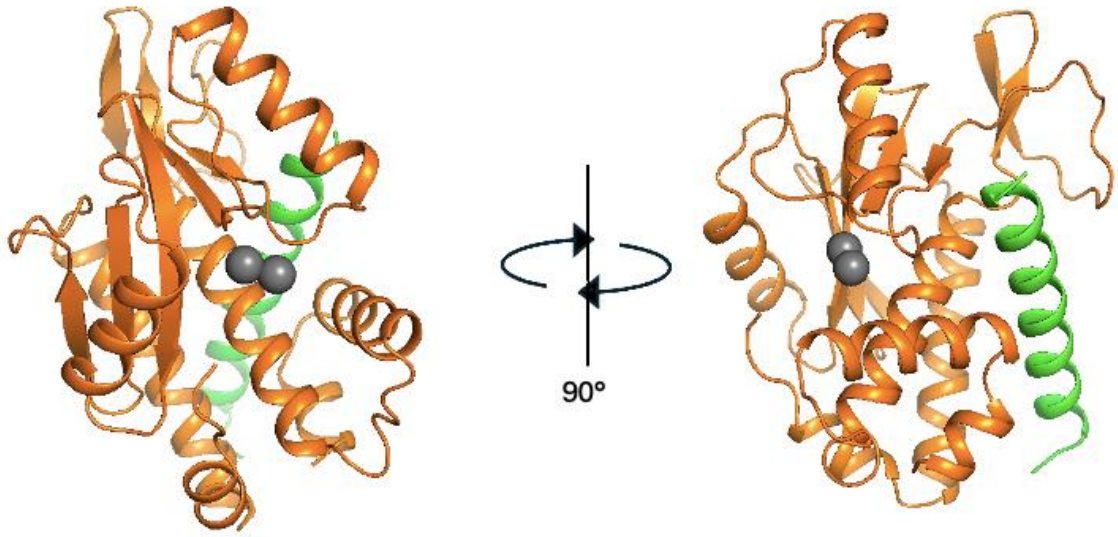
Figure 10. Selection and AlphaFold3 predictions of PB1 peptide interaction with endonuclease. *A. Cryo-EM solved structure of the influenza type A endonuclease in the encapsidating confirmation (orange), with stick-representation of the region selected as the PB1 peptide (blue), PDB 8R1L. B. AlphaFold 3-generated graph of position and expected error, used to assess the confidence of the prediction. C. AlphaFold 3-generated structures showing proposed interaction between magnesium-bound endonuclease and the PB1 peptide, with endonuclease (orange), PB1 peptide (blue), and manganese (grey).*

Design of PB2 Peptide

The positioning and structure of endonuclease in the encapsidating form of the influenza polymerase has not always been resolved with high resolution in published structures, likely due to the flexibility and regions of intrinsic disorder in the endonuclease (Fan et al., 2019; Zhu et al., 2023). However, evidence from Thierry et al. in 2016 and Staller et al. in 2024 suggest interaction between endonuclease and the C-terminal domain of the PB2 subunit. The 745-770 region of PB2 was identified to interact with endonuclease in Flu B by Thierry et al. in 2016, where they identified a novel repacking of the PB2 subunit characteristic of the encapsidating form of the endonuclease. Based on these data, I selected a C-terminal peptide composed of

residues M740-N758. A structural model of the predicted complex between endonuclease and this peptide was generated by AlphaFold 3 though confidence was low for both the selected peptide structure and the regions of interaction. Given the magnitude of the surface contacts between endonuclease and the PB2 subunit in the encapsidase structure, it seemed likely that binding with the single selected peptide would not be strong enough to result in a high-confidence AlphaFold3 prediction (Figure 11).

A



B

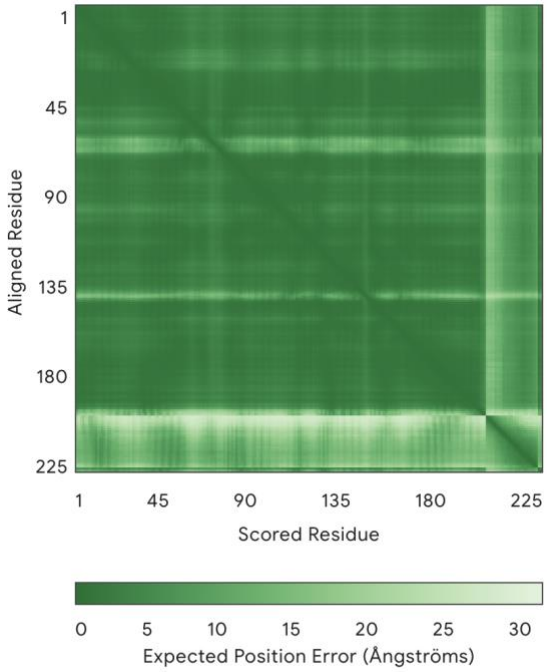


Figure 11. AlphaFold3 predictions of PB2 peptide interaction with endonuclease. *A. AlphaFold 3-generated structures showing proposed interaction between magnesium-bound endonuclease and the PB2 peptide, with endonuclease (orange), PB2 peptide (green), and manganese (grey). B. AlphaFold 3-generated graph of position and expected error, used to assess the confidence of the prediction*

Endonuclease-Peptide Titrations

After potential binding partners were selected, I proceeded by titrating PB1 peptide, PB2 peptide, and purified R416A NP into purified, ^{15}N labeled endonuclease and observing shifts via NMR.

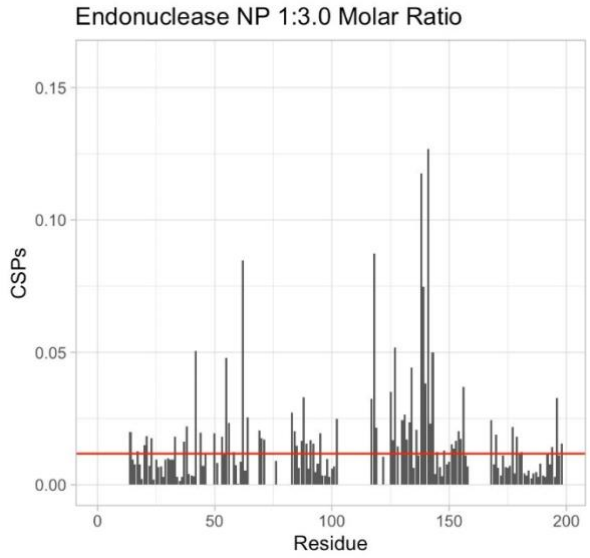
Endonuclease-NP Titrations

The titration of R416A NP into endonuclease indicated an interaction in the region between residues 134-142. CSP data for all spectra across the titration series corroborated these results (Appendix Figure 1), with CSP magnitude directly correlated to NP concentration of the titration point. The mean CSP for the experiment comparing

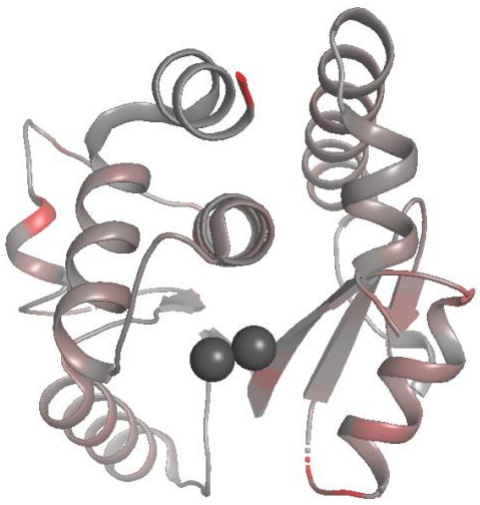
endonuclease in buffer to endonuclease with a threefold molar excess of NP was 0.012, with a maximum CSP = 0.126 at residue 141. Throughout titration experiments, CSPs greater than five times the mean were considered suggestive of interaction and those greater than eight times the mean were considered likely interaction sites.

Mapping of the measured CSPs to the endonuclease structure revealed that the primary areas of interaction are in the unstructured region from residues 137-142 at the C-terminal end of α-helix 5. Residues 63 and 119 also demonstrated CSPs significantly greater than the mean, with calculated shifts of 0.085 and 0.087, respectively. However, in both cases only individual residues showed a dramatic increase in CSP, which lowers confidence in those reflecting a binding interface. An overlay of spectra from the titration points (Figure 12c) showed small peak shifts between titration points, indicating that the binding event was fast on the NMR timescale and therefore low affinity.

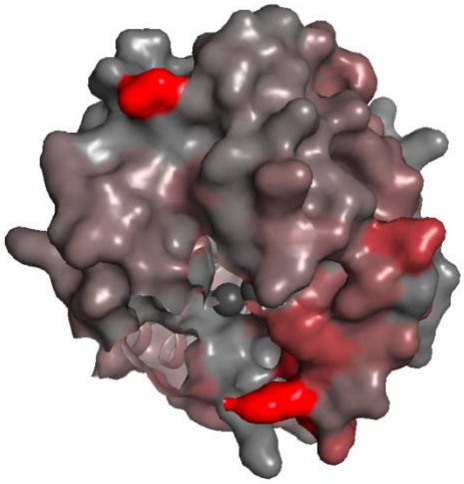
A



B



C



D

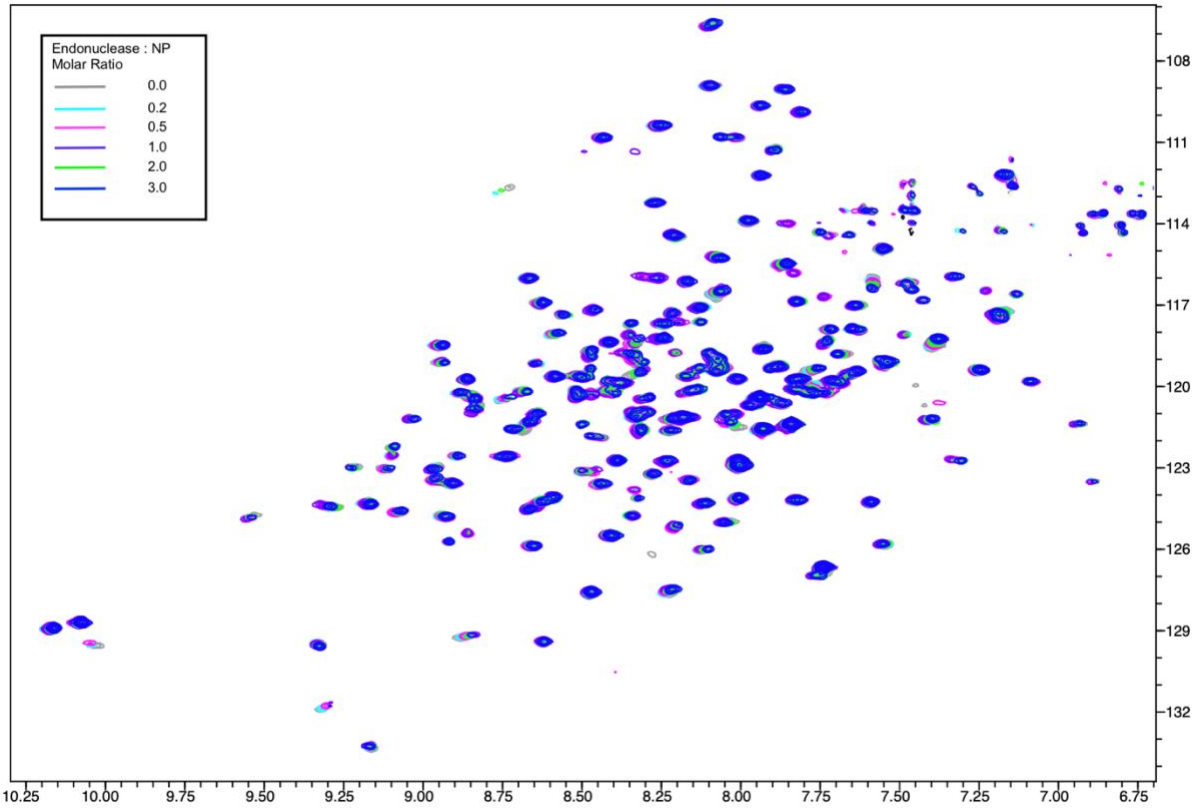


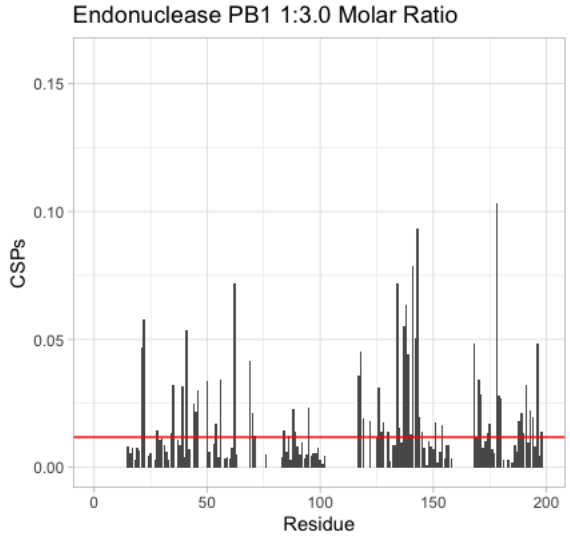
Figure 12. Titration of NP into solution-state endonuclease indicates interaction*. A. CSPs of endonuclease with three-fold molar excess of R416A NP. All CSPs given in ppm. Mean CSP value = 0.012 indicated in red. B. Metal-bound structure of WT endonuclease (PDB 2W69) with residues color-coded on a scale from gray (CSP = 0.0) to red (CSP = 0.12). C. Surface representation of WT endonuclease with three-fold molar excess of R416A NP, residues color-coded on a scale from gray (CSP = 0.0) to red (CSP = 0.12). D. Overlay of NMR spectra showing NP (unlabeled) titration of purified ^{15}N-labelled endonuclease sample, x-axis (^{1}H, ppm), y-axis (^{15}N, ppm).*

Endonuclease-PB1 Peptide Titrations

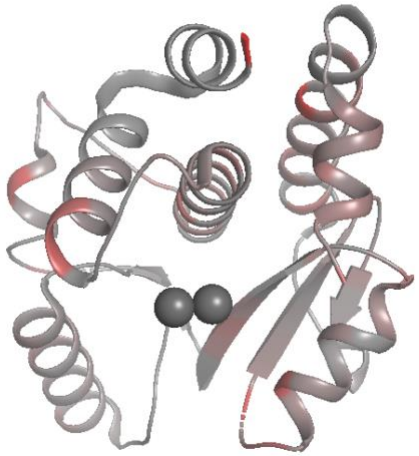
Titration of the PB1 peptide into endonuclease gave some evidence of binding at the highest titration point, a three-fold molar excess of PB1 peptide to endonuclease, though at no point was saturation achieved (Figure 13). In this experiment, shifts appeared in the endonuclease residues 134-143, where I observed consistently elevated CSPs, similar to those observed in the NP experiments. The maximum CSP in this region was 0.09 for residue 143. A maximum CSP of 0.103 occurred at residue 178.

When CSPs were mapped onto the endonuclease structure, it appeared that the shift at residue 178 occurred in a short α-helix in the otherwise disordered region towards the endonuclease C-terminus. Shifts in individual residues are not usually considered indicative of binding since a protein-protein interaction surface is unlikely to consist of a single residue. Though residue 178 had a CSP greater than 0.08 in the presence of one, two, and three-fold molar excess of PB1 peptide (Appendix Figure 4), it is an arginine, a charged amino acid which would likely be sensitive to changes in salt or pH. Both salt concentrations and pH were kept consistent across the samples, but local charge differences could have contributed to the apparent shifts.

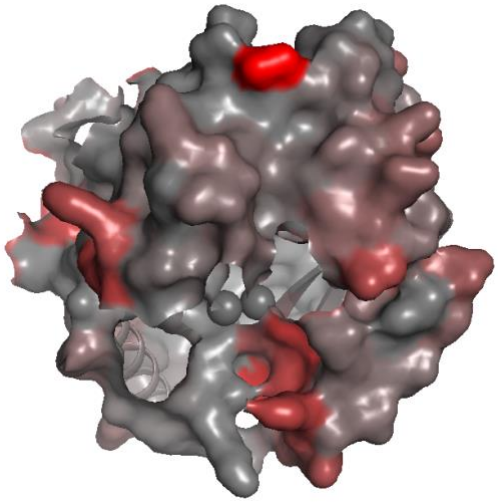
A


Endonuclease PB1 1:3.0 Molar Ratio
CSPs
0.15
0.10
0.05
0.00
0
50
100
150
200
Residue

B



C



D

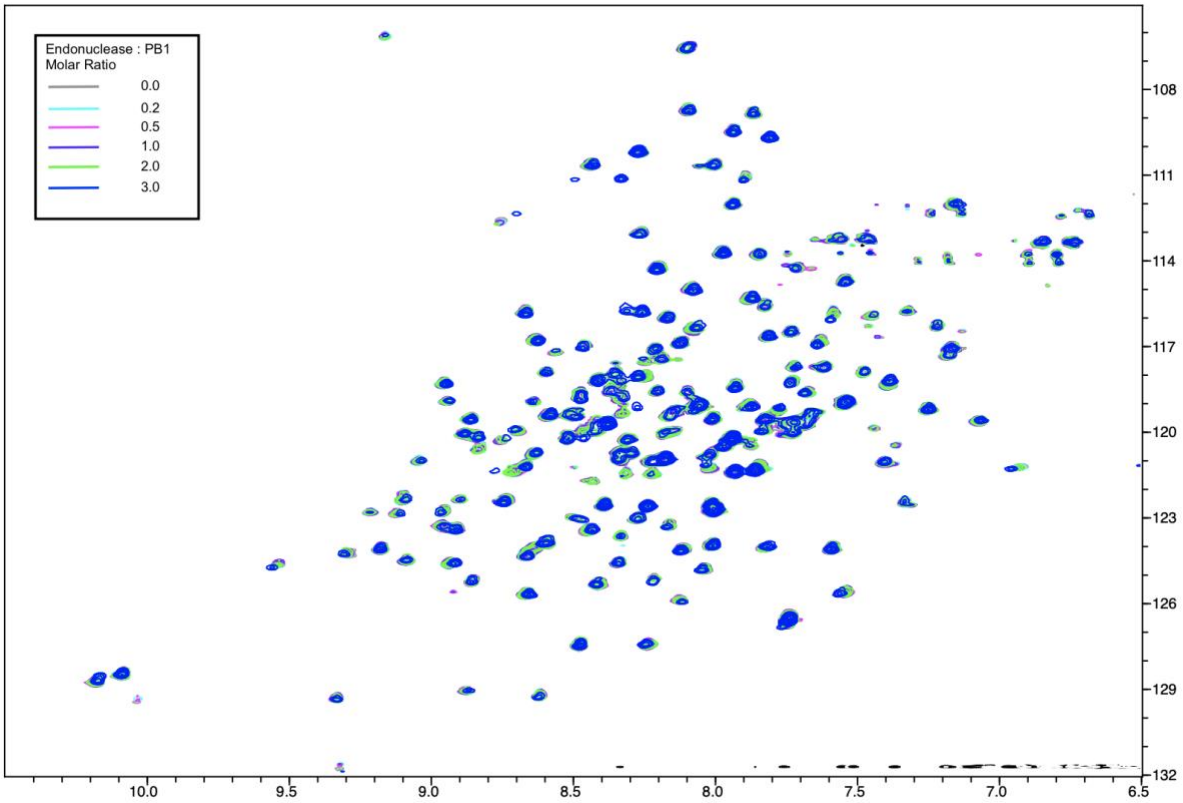
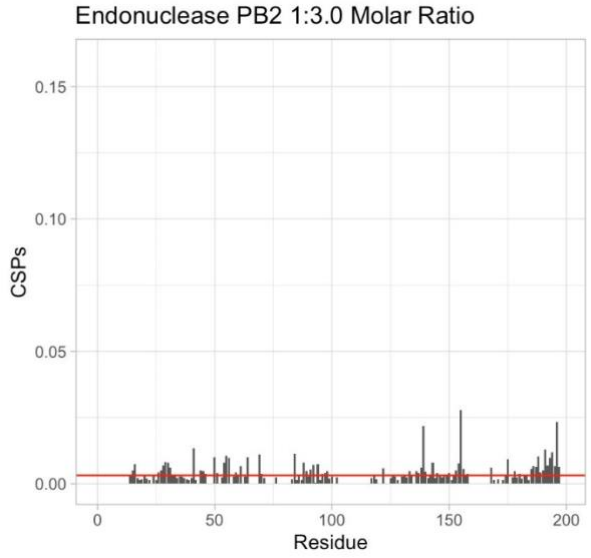


Figure 13. Titration of PB1 peptide into solution-state endonuclease indicates interaction. *A. CSPs of endonuclease with three-fold molar excess of PB1 peptide. All CSPs given in ppm. Mean CSP value = 0.12 indicated in red. B. Metal-bound structure of WT endonuclease (PDB 2W69) with residues color-coded on a scale from gray (CSP = 0.0) to red (CSP = 0.12). C. Surface representation of WT endonuclease with three-fold molar excess of PB1 peptide, residues color-coded on a scale from gray (CSP = 0.0) to red (CSP = 0.12). D. Overlay of ^{1}H-^{15}N BEST-TROSY spectra showing titration of PB1 peptide into purified endonuclease sample, x-axis (^{1}H, ppm), y-axis (^{15}N, ppm).*

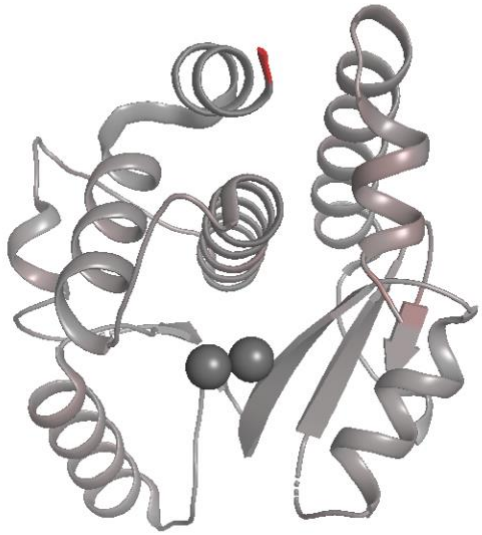
Endonuclease-PB2 Peptide Interaction

The maximum CSP observed in endonuclease upon addition of a three-fold molar excess of PB2 peptide was 0.028 at residue 155 (Figure 14). The mean CSP for this experiment was 0.003 and is lower than the mean CSPs observed in the NP and PB1 experiments, indicating minimal shifts throughout the spectrum and suggesting that endonuclease did not interact with the PB2 peptide.

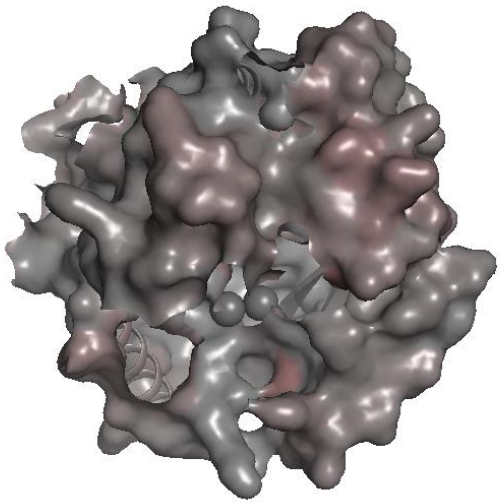
A


Endonuclease PB2 1:3.0 Molar Ratio
CSPs
0.15
0.10
0.05
0.00
0
50
100
150
200
Residue

B



C



D

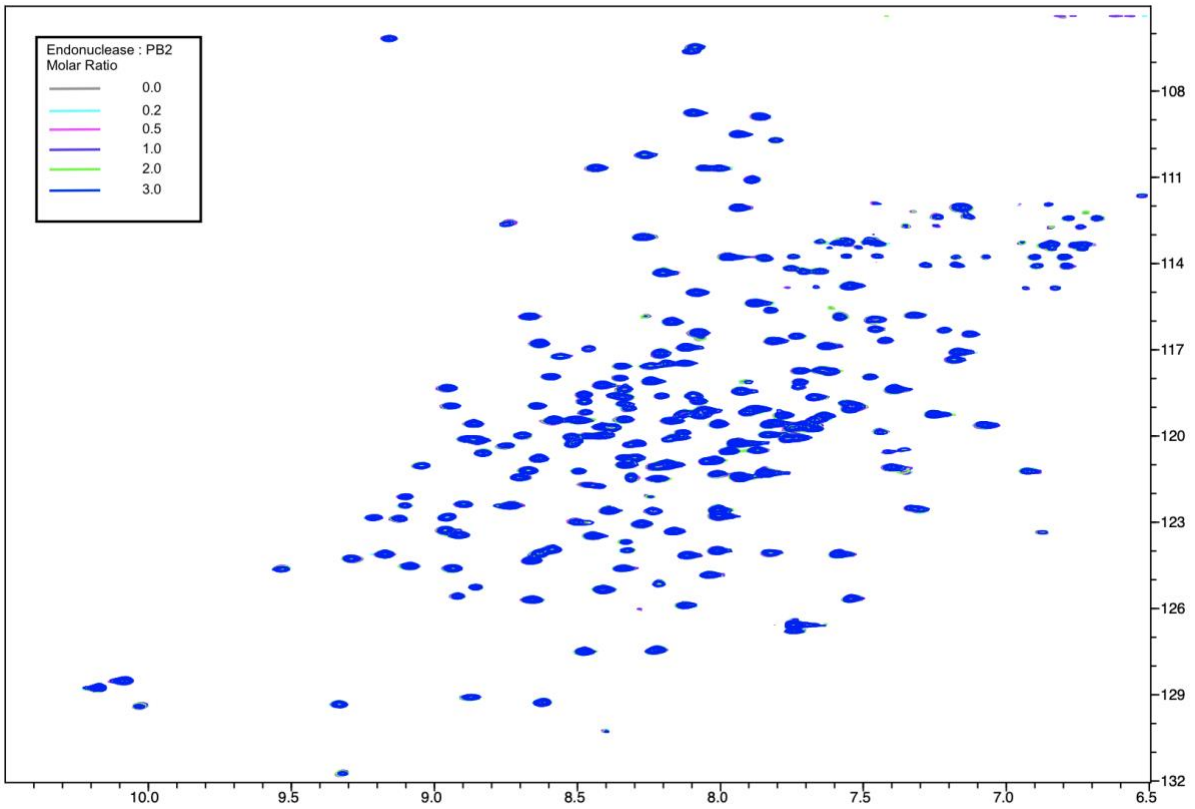


Figure 14. Titration of PB2 into solution-state endonuclease does not indicate interaction. *A. CSPs of endonuclease with three-fold molar excess of PB2 peptide. All CSPs given in ppm. Mean CSP value = 0.003 indicated in red. B. Metal-bound structure of WT endonuclease (PDB 2W69) color coded to indicate shifts with CSP = 0.0 in gray and CSP = 0.12 in red. C. Surface representation of WT endonuclease with three-fold molar excess of PB2 peptide, color coded from gray (CSP = 0.0) to red (CSP = 0.12). D. CARA spectral overlay showing titration of PB2 peptide into purified endonuclease sample, x-axis (*^{1}H, ppm*), y-axis (*^{15}*N, ppm).*

Endonuclease-PB2 Peptide Titrations in the Presence of NP

It has also been proposed that the increased density around endonuclease in cryo-EM maps presumed to be NP is a transient, stabilizing interaction which allows for endonuclease to bind PB2 (E. Fodor, personal communication). Having observed some chemical shifts in endonuclease with NP, I therefore decided to repeat the PB2 peptide titration with NP present to see whether NP might facilitate interaction between endonuclease and the PB2 peptide.

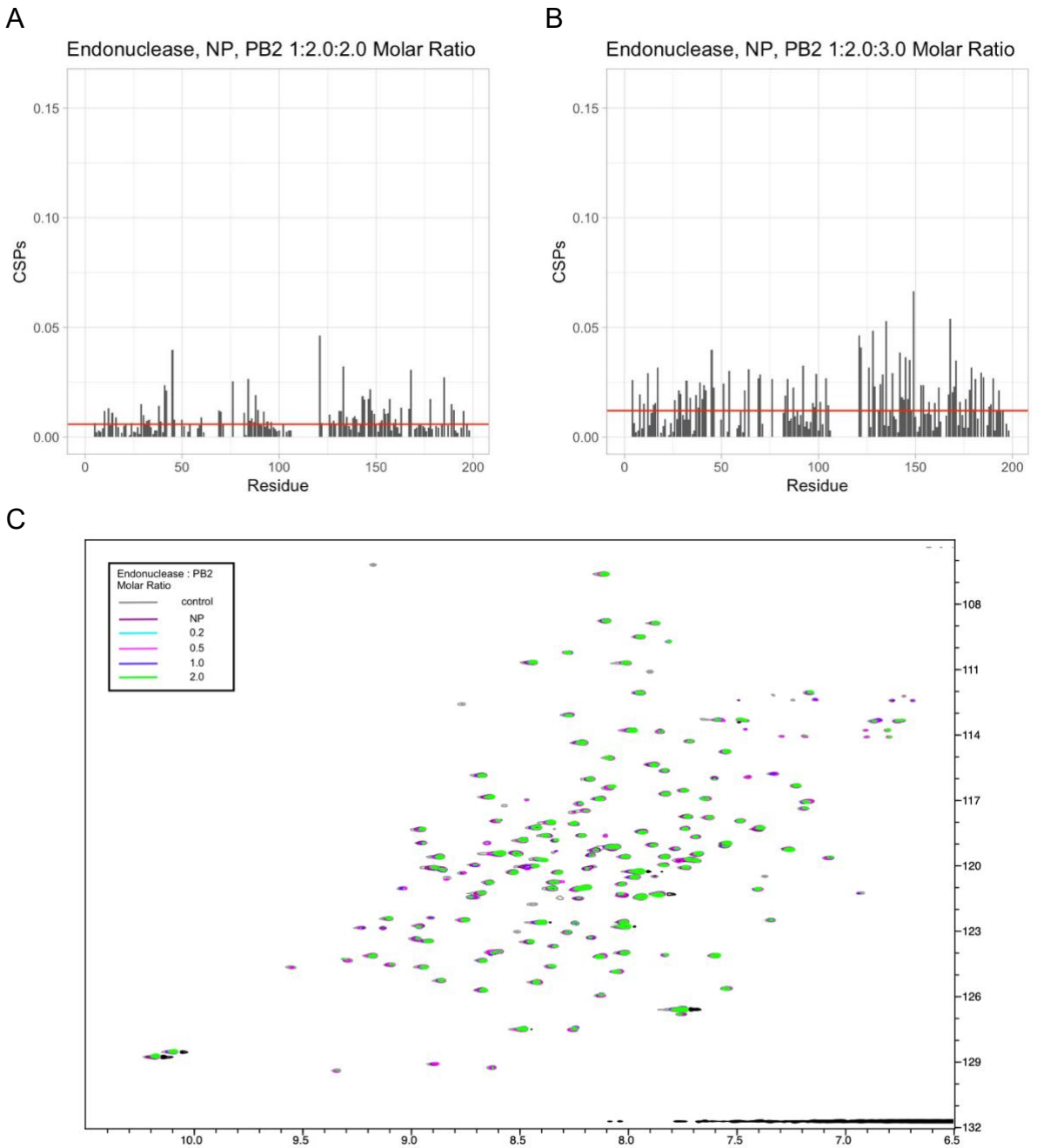


Figure 15. The presence of NP does not appear to stabilize an endonuclease-PB2 peptide interaction*. A. CSPs of endonuclease with a 1.0:2.0 molar ratio of endonuclease to NP and a 1.0:2.0 molar ratio of endonuclease to PB2 peptide, as compared to the control spectrum of free endonuclease. The mean CSP is indicated by a horizontal red line. B. CSPs of endonuclease with a 1.0:2.0 molar ratio of endonuclease to NP and a 1.0:3.0 molar ratio of endonuclease to PB2 peptide. Mean CSP indicated in red. C. NMR spectral overlay showing NP and PB2 peptide titrated into an endonuclease sample, with a maximal molar ratio of endonuclease to PB2 peptide of 1.0:2.0. All spectra except the control contained a 1.0:2.0 molar ratio of endonuclease to NP with varying concentrations of PB2 peptide, x-axis (*^{1}H, ppm*), y-axis (*^{15}N, ppm*), . The spectrum with a three-fold molar excess of PB2 peptide to endonuclease had very low signal to noise and, once contoured with the same parameters as the other spectra, had no remaining visible peaks. All spectra except the*

control contained a 1.0:2.0 molar ratio of endonuclease to NP with varying concentrations of PB2 peptide

The titration experiment conducted with PB2 peptide was repeated with the addition of a two-fold molar excess of NP in each sample (Figure 15). The NP/PB2 spectra were compared to control spectrum of endonuclease alone for calculation of CSPs. The signal to noise for these spectra was less than those of the previously recorded spectra, making analysis of CSPs more difficult, especially in the case of the three-fold molar excess of PB2 peptide titration point (Appendix Figure 5), where the signal intensity had decreased to the point that analyzing shifts with confidence was not possible because peaks were not sufficiently distinct to track their movement. The reason for lower signal to noise in this case is not clear – titrations with both NP and PB2 peptide independently were high quality, including in the higher-concentration NP titration points, where the total protein in solution was much higher than in the NP/PB2 peptide experiments. It is possible that the addition of NP and PB2 to the endonuclease sample formed a large, stable complex, reducing the signal to noise at each titration point, but it seems more likely that the endonuclease used in these samples was otherwise degraded. Specifically, the endonuclease sample used in this experiment had been frozen and then thawed before visualization rather than being made fresh due to errors in sample preparation. Though the sample concentrations were confirmed via NanoDrop, there was notable visual evidence of precipitation suggesting possible instability of the protein in solution.

The addition of NP did not appear to facilitate interaction between endonuclease and PB2 peptide in these samples, though this could have been affected by the quality

of the endonuclease sample (Appendix Figure 4). Residues A135, S149, R168 showed CSPs greater than 0.05 in the experiment with a two-fold molar excess of NP and three-fold molar excess of PB2, but these were compared to a mean perturbation of 0.012, or 0.016 excluding CSPs that equal 0 ppm. It is possible that the lower quality spectra in this experiment (Figure 15d) made it more difficult to identify shifts faithfully and therefore skewed the results towards non-interaction.

None of the ligands introduced showed compelling evidence of stabilization of peak positions across the spectra which would have suggested saturation. There was evidence of saturation at residues K139 and E141 at the one-to-one molar ratio of endonuclease to NP, giving evidence of a specific interaction between the residue and the ligand (Figure 16b, c).

Neither K139 nor E141 appeared to saturate with the introduction of PB1 peptide, and residues K137 and T143 did not show evidence of saturation with NP or with PB1 peptide, perhaps giving evidence of some sort of non-specific interaction (Figure 16). Given the physiological context of endonuclease interactions with likely binding partner, that is, with substantial external stabilization from other subunits of the polymerase, it is difficult to judge whether interacting endonuclease residues might saturate *in vivo*. However, the lack of saturation, coupled with the fact that each of the residues with notable CSPs carries a charge or is polar, suggests that the rearrangement observed especially upon the introduction of PB1 peptide could have been explained by charge interaction shifts.

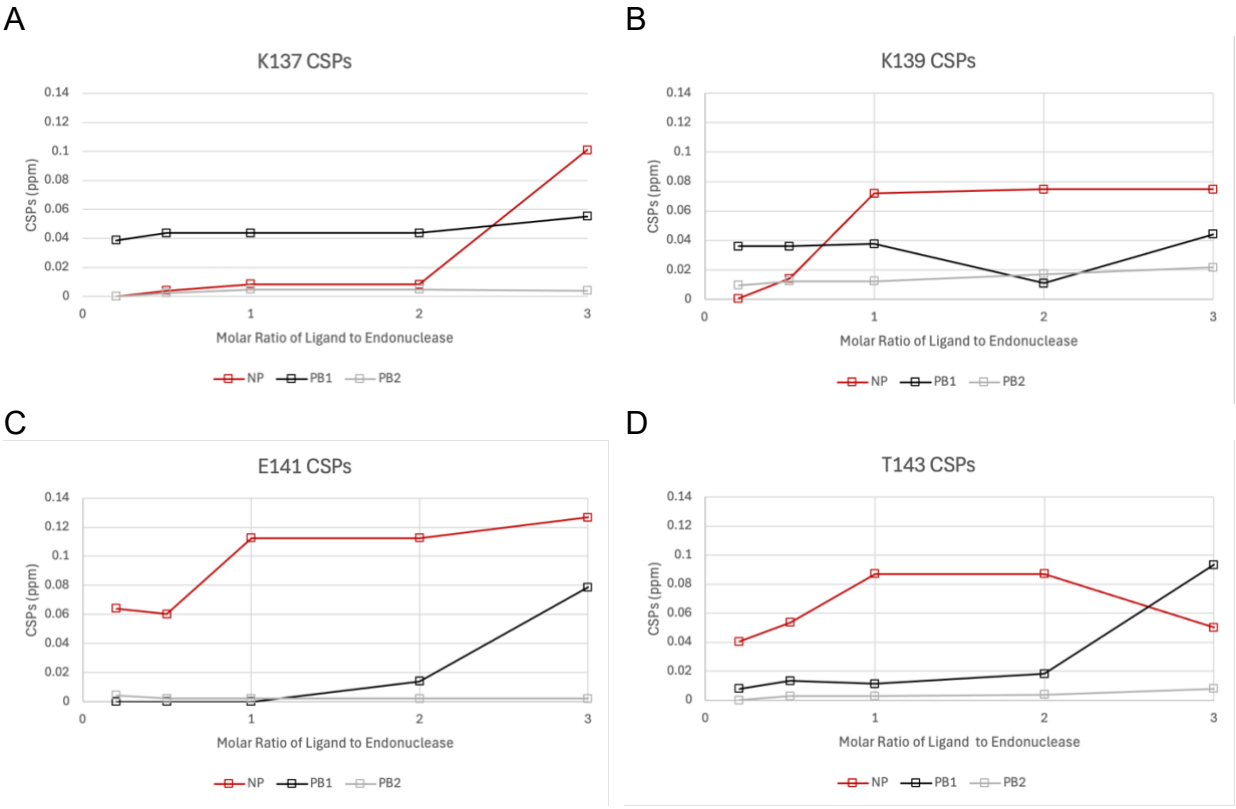


Figure 16. CSP of individual residues as a function of molar ratio of titrated ligands NP (red), PB1 (black), and PB2 (gray). *A. CSPs for residue K137. B. CSPs for residue K139. C. CSPs for residue E141. D. CSPs for residue T143.*

Discussion

AlphaFold 3 Predictions of NP and Peptide Interactions

AlphaFold 3 predictions for endonuclease interactions with all three of the selected potential binding partners were low confidence. It is possible that the flexible regions of endonuclease, which have reduced density of Cryo-EM maps, are the interaction sites for the ligands of interest and are not easily modelled by AlphaFold 3. The AlphaFold 3 models for all three ligands differed significantly in endonuclease-ligand binding sites from predictions based on the solved structures. After completing the titration experiments, I was able to compare my *in vitro* results to the *in silico* AlphaFold 3 predictions and found that AlphaFold 3 incorrectly predicted interaction

sites in all three models. This was unsurprising since the pLDDTs were very low (<50) at the predicted sites for endonuclease-ligand interactions in every case.

Endonuclease-NP Interactions

Previous work solving Cryo-EM structures of the FluPol dimer noted some increased density around endonuclease $FluPol_R$ which was postulated to be NP, but endonuclease and NP could not be fully resolved in the context of the larger FluPol structure (Carrique et al., 2020; E. Fodor, personal communication). The potential endonuclease-NP interaction is particularly interesting given the interaction suspected between NP and ANP32, where the ANP32 LCAR is hypothesized to interact with NP, increasing the relative concentration of NP around the FluPol dimer to facilitate RNP formation. In this proposed model, NP could then be further recruited towards nascent cRNAs and vRNAs via transient interaction with endonuclease.

I was particularly interested to observe that both NP and PB1 appeared to interact with the endonuclease in residues 137-142. These data suggest that the PB1-endonuclease interaction may be replaced by transient interaction with NP in $FluPol_R$, when the region is not buried. Given the packing of endonuclease against PB1 in the cap-snatching incompetent form of FluPol, where burying of endonuclease against the PB1 subunit is suggested to prevent unwanted enzymatic activity, the interface overlap between PB1 and NP points to a possible mechanism for regulating endonuclease activity.

Endonuclease-PB1 Peptide Interactions

The shifts observed upon titration of PB1 peptide into endonuclease provide some evidence of an interaction. Given the lack of saturation, it is possible that the observed shifts were due to charge interactions, though there were relatively consistently elevated CSPs in the region between residues 137-142 which would suggest transient binding. Notably, there was substantial overlap in the endonuclease residues which appeared to be interacting both NP and PB1. The solved crystal structure of manganese bound endonuclease (PDB 2W69) shows an α-helix from residues 126 to 136 followed by a disordered region with some missing density, which overlaps with the apparent endonuclease interaction site with both NP and PB1 peptide, where the last two residues of the α-helix and following disordered region show evidence of perturbation.

FluPol dimer structures indicate that endonuclease interacts with different parts of the PB1 subunit in $FluPol_R$ and $FluPol_E$, which helped inform selection of the PB1 peptide. In representations of the replicating form of the endonuclease published by Staller et al. in 2024 (PDB 8R1L), the replicating form of the endonuclease does not appear to have any surface contacts with the relevant region of PB1. Instead, the region from residues 134-142, including the disordered region beginning at lysine 137, are relatively exposed on the surface of $FluPol_R$. In contrast, endonuclease appears to have substantial surface contacts with the PB1 peptide in $FluPol_E$, an interaction which would be supported by the above results, which suggest that endonuclease and PB1 interact directly in $FluPol_E$ (Staller et al., 2024). In this model, the interaction between PB1 and the 137-142 disordered region in $FluPol_E$ could play a role in stabilizing the novel

conformation observed in the encapsidating form of the polymerase or in regulating endonuclease activity in the replicase. In this proposed model, PB1 would interact with and potentially regulate endonuclease activity in $FluPol_E$ and NP would interact with the same region in $FluPol_R$.

Endonuclease-PB2 Peptide Interactions

In contrast to results reported in influenza B, I did not observe any meaningful interactions between endonuclease and the PB2 peptide. I had anticipated some movement in residues 164-179 of the enzyme since previous publications have confirmed an interaction between the influenza B endonuclease and the PB2 NLS (Thierry et al., 2016). It should be noted that the solved structure of $FluPol_E$ in influenza A does exhibit extensive contacts between endonuclease and the PB2 subunit. In particular, the interaction between the PB2 cap-binding domain, endonuclease, and the PB2 "lid" subdomain appears to be decisive in the formation of the encapsidase structure. These interactions are stabilized by significant surface contacts and hydrogen bonding between the endonuclease and relevant PB2 domains, so it is very plausible that binding was not observed because the relatively short peptide lacked the stabilizing interactions which would occur in the full heterotrimeric complex.

Further experiments to investigate conformational changes in endonuclease upon interaction with PB2 could be conducted using a construct of the full-length PB2 subunit, though the size of this ligand (MW = 85k Da) would make detection of NMR signals challenging since the slow tumbling times of large proteins lead to peak broadening and loss of signal in NMR experiments (Xing et al., 2017). I did not observe

any evidence of interaction between endonuclease and the PB2 peptide of choice, contradictory to published data. However, the lack of peak shifts observed were likely due to the inability of this experiment to mimic *in vivo* conditions. Previous experiments providing evidence of the interaction between endonuclease and the PB2 subunit were able to investigate the interaction between endonuclease and the full length of the subunit. I therefore suspected that the interaction between PB2 and endonuclease requires stabilization by hydrogen bonding from various subunits of the full PB2 complex and therefore could not be replicated with a single, twenty-five amino acid long peptide.

Summary and Next Steps

These data imply that both NP and the 100-115 region of PB1 interact with the PA endonuclease of FluPol at the N-terminal end of the α-helix and subsequent flexible region from residues 137-142. Evidence for interaction with NP is particularly strong given the saturation of two residues – K139 and E141 – at an equimolar ratio of endonuclease to NP. Though significant interactions (CSPs > 5 times the mean) were observed upon the introduction of PB1 peptide, evidence for binding was weaker because of the lack of saturation. Though the solution conditions provided had consistent salt concentrations, pH, and other buffer conditions, making transient interaction likely, sample conditions could not be made identical, so it is possible that the shifts observed upon the introduction of PB1 peptide were due to charge interactions.

In this chapter I show evidence of interaction between endonuclease and NP and endonuclease and PB1 peptide, provide evidence against binding between endonuclease and the PB2 peptide selected, and indicate regions where interaction likely occurred, accomplishing the second aim of this thesis: to establish and identify regions of interaction between endonuclease and binding partners of interest.

Chapter 4: Ligand Interactions with Endonuclease Mutants

Use of Mutations to Validate Ligand Binding

The titrations conducted in chapter three indicated residues with CSPs in the PA endonuclease as a result of PB1 or NP binding. I sought to validate the apparent interactions of endonuclease with PB1 and NP by testing point mutations in endonuclease designed to disrupt the putative interface.

Selection of Mutants

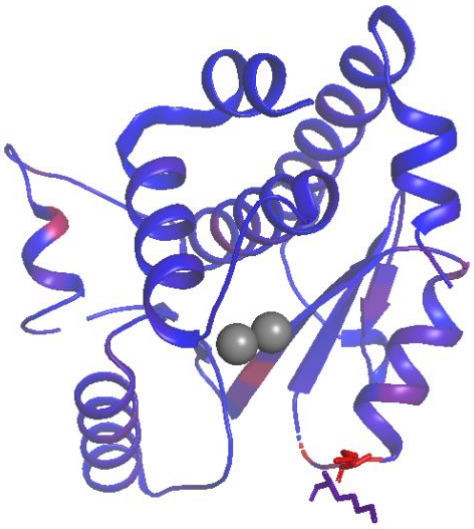
The CSPs measured in Chapter 3 after titrations of unlabeled NP and PB1 peptide into ^{15}N-labelled endonuclease indicated overlapping interaction sites at residues 136-144 and residues 134-146, respectively. As noted in the previous chapter, there was little evidence of an interaction between endonuclease and the PB2 peptide. The charged residues K137 and K139 were chosen as targets for mutagenesis because they are on the protein surface, are conserved, and may stabilize a protein-protein interaction through salt bridges (Figure 17a, b).

Three substitutions at two different sites in the WT endonuclease (K137 and K139), were selected with the aim of knocking out interactions while maintaining the

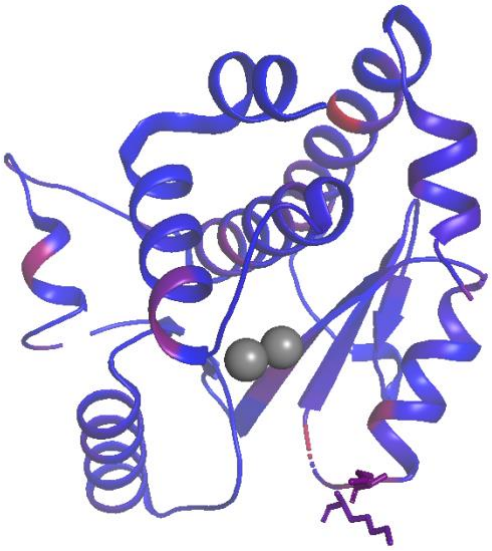
overall structure of the endonuclease. K137 is at the C-terminal end of the α-helix composed of residues E126-N136, whereas K139 is unstructured (Dias et al., 2009). Selection of these sites as potentially meaningful to future research on endonuclease and polymerase function was supported by analysis of endonuclease sequences in influenza type A (Figure 17c) and across different influenza types (Figure 17d).

ConSurf analysis of the PA endonuclease region across influenza type A strains (Figure 17c) indicated that both of the selected lysines were conserved across species, suggesting their functional relevance (Berezin et al., 2004). This was supported by sequence alignment analysis of sequences from influenza types A, B, and C and strains investigated in previous publications (Kuraku et al., 2013; Thierry et al., 2016; Katoh et al., 2019; Zhu et al., 2023; Staller et al., 2024). The lysine at residue 137 is also conserved between influenza types and was assessed by the alignment software as a likely functional site. K139 was assessed as a gap residue for influenza types B and C, but was included in mutant experimentation because of its proximity to the region with significant CSPs and its outward-facing positioning in the solved cryo-EM structures. Three substitutions were chosen: K137S, K137Q, and K139E. AlphaFold 3 models suggested that these mutations were unlikely to cause significant structural changes (Figure 18).

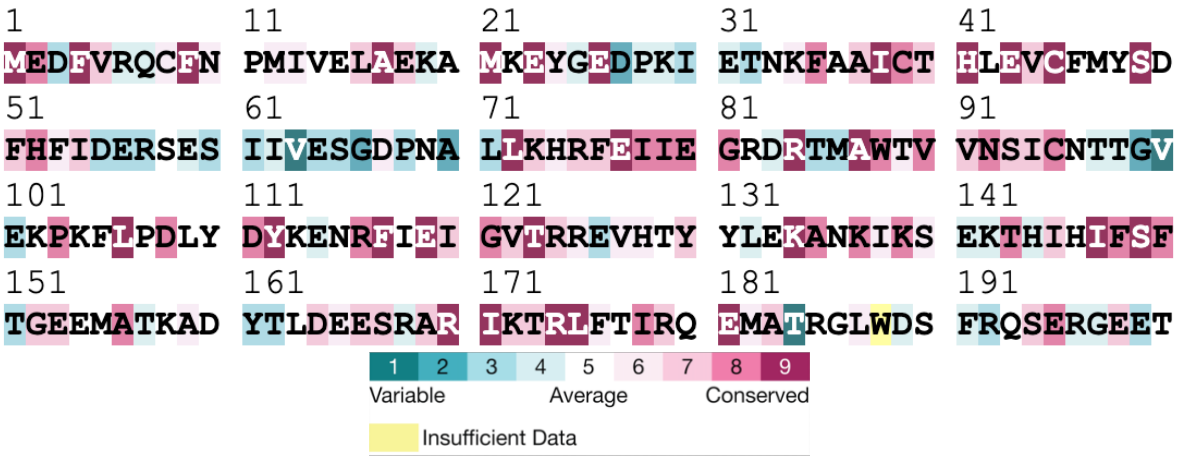
A



B



C

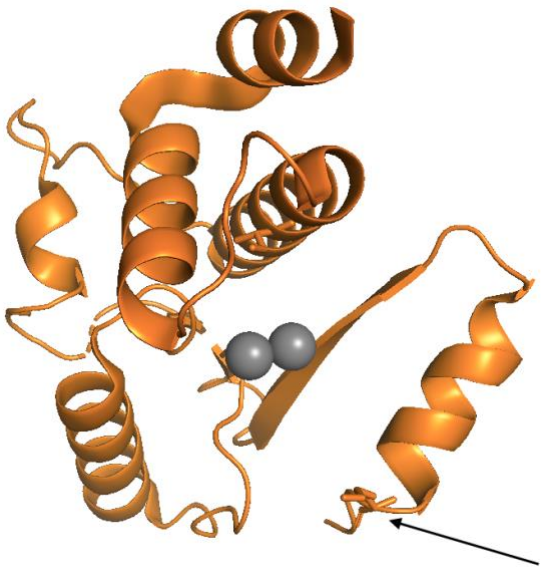

1 11 21 31 41
MEDFVRQCFN PMIVELAEKA MKEYGEDPKI ETNKFAAICT HLEVCFMYSD
51 61 71 81 91
FHFIDERSES IIVESGDPNA LLKHRFEIIE GRDRTMAWTV VNSICNTTGV
101 111 121 131 141
EKPKFLPDLY DYKENRFIEI GVTRREVHTY YLEKANKIKS EKTHIHIFSF
151 161 171 181 191
TGEEMATKAD YTLDEESRAR IKTRLFTIRQ EMATRGLWDS FRQSERGEET
1 2 3 4 5 6 7 8 9
Variable
Average
Conserved
Insufficient Data

D

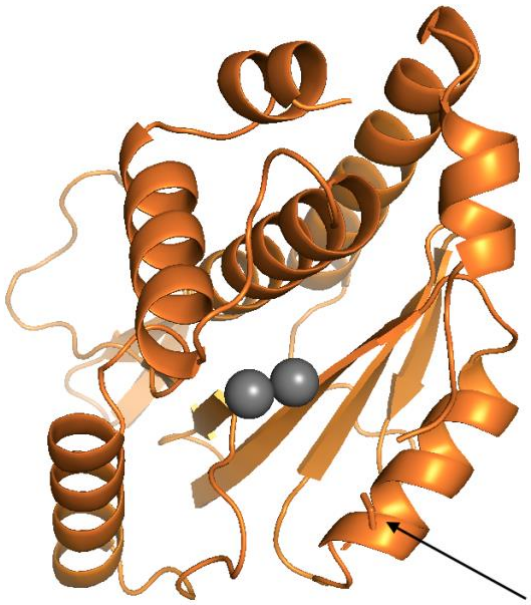
```
8R1L       M----EDFVRQCFNPMIVELAEKAMKEYGEDPRIETNKFAAICTHMEVSFMYSDFHFINE
4WSA       M----DTFITRNFQTTIIQKAKNTMAEFSEDPELQPAMLFNICVHLEVCYVISDMNFLDE
6XZR       MSKTFAEIAEAFLEPEAVRIAKEAVEEYGDHER----KIIQIGIHFQVCCMFCD-EYLST
           *      :    ::.  :. *:::: *:.:. .     :  *  *::*. : .* .::.

8R1L       RGES-IIVESGDPNALLKHRFEIIEGRDRAMAWTVVNSICNTTGVGKPKFLPDLYDYKED
4WSA       EGKTYTALEGQGKEQNLRPQYEVIEGMPRNIAWMVQRSLAQEHGIETPRYLADLFDYKTK
6XZR       NGSD---------------RFVLIEGRKRGTAVSLQNELCKSYDLEPLPFLCDIFDREEK
           .*.                :: :***  *  *  : ..:.:  .:    :* *::* : .

8R1L       RFIEIGVTRREVHIYYLEKANKIKSEETHIHIFSFTGEEMATKADYTLDEESRARIKTRL
4WSA       RFIEVGITKGLADDYFWKKKEKL-GNSMELMIFSYNQ-DYSLSNESSLDEEGKGRVLSRL
6XZR       QFVEIGITRKADDSYFQSKFGKL-GNSCKIFVFSYDG-RLDKNCEGPMEEQ-KLRIFSFL
           :*:*:*:*:   . *: .*  *: .:. .: :**:       . : .::*: : *: : *

8R1L       FTIRQEMASRGLWDSFRQSERGE-----------------
4WSA       TELQAELSLKNLWQVLIGEEDIE-----------------K
6XZR       ATAADFLRKENMFNEIFLPDNEETIIEMKKGKTFLELRDES
                 :  ..::: :   :  *
```

Figure 17. The positioning of residues K137 and K139 suggests that they may be available for ligand interaction. *A. Metal-bound structure of WT endonuclease (PDB 2W69) with side chains shown for residues 137 and 139, color coded to indicate shifts after introduction of three-fold molar excess of NP. CSP = 0.0 in blue and CSP = 0.12 in red. B. Metal-bound structure of WT endonuclease (PDB 2W69) with side chains shown for residues 137 and 139, color coded to indicate shifts after introduction of three-fold molar excess of PB1 peptide. Residues were colored according to their CSPs from blue (0.0) to red (0.12). C. Sequence alignment between influenza type A strains, assessed by ConSurf (Berezin et al., 2004). D. Sequence alignment of endonuclease residues from influenza types A (PDB 8R1L), B (PDB 4WSA), and C (PDB 6XZR) using MAFFT (Katoh et al., 2019; Kuraku et al., 2013).*

A



B



C D

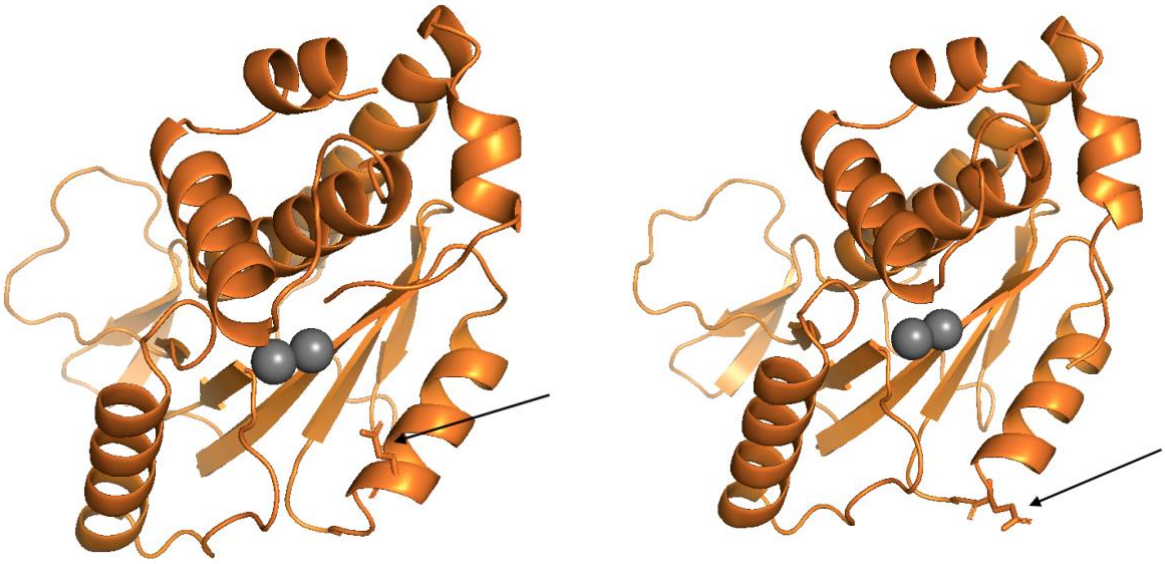


Figure 18. AlphaFold 3 predictions of WT and mutant endonucleases*. A. WT endonuclease with stick representation of residues K137 (indicated) and K139 (obscured by the K137 residue). B. K137S endonuclease with stick representation of residue S137 indicated. C. K137Q endonuclease with stick representation of residue Q137 indicated. K139E endonuclease with stick representation of residue E139 indicated.*

Mutant Characterization and Binding

AlphaFold 3 Predictions of Mutant Activity

AlphaFold 3 was used to predict the potential interaction of each endonuclease with the three ligands of interest: PB1, PB2, and NP. Notably, AlphaFold 3 has several limitations in this case given the disordered regions of interest in the endonuclease and the lack of tertiary structure contacts in the isolated peptides (Abramson et al.; 2024). I therefore used AlphaFold 3 primarily as initial supporting evidence that the endonuclease structure would not likely be strongly perturbed.

Characterization of Endonuclease Mutants via NMR

Spectra of mutant endonucleases were collected and analyzed in comparison to a WT control spectrum. With the exception of the spectra collected with K139E, which had low signal to noise relative to other spectra collected with the same parameters, peaks could be mapped without trouble between the WT and mutant spectra using Computer Aided Resonance Assignment (CARA), implying that introduction of mutations did not cause significant structural change. Observed changes in mutant interactions and activity were therefore supposed to be directly related to the mutation introduced, rather than to resultant structural changes.

Figure 19 shows the $^{1}H,^{15}N$ BEST-TROSY (2D) spectrum of each mutant at 100 μM protein. The mutant K139E had significantly reduced signal-to-noise, which made it more difficult to track peak shifts over the titration and between the WT and mutant spectra. Despite the reduced signal to noise, particularly for the K139E mutant, peak tracking between the WT and mutant endonucleases was uncomplicated, with minimal if any shifts between the WT and mutant spectra. The lack of dramatic peak shifts indicates that the mutations did not cause substantial changes in the global structure of endonuclease, though assignment was not possible for the mutated residue.

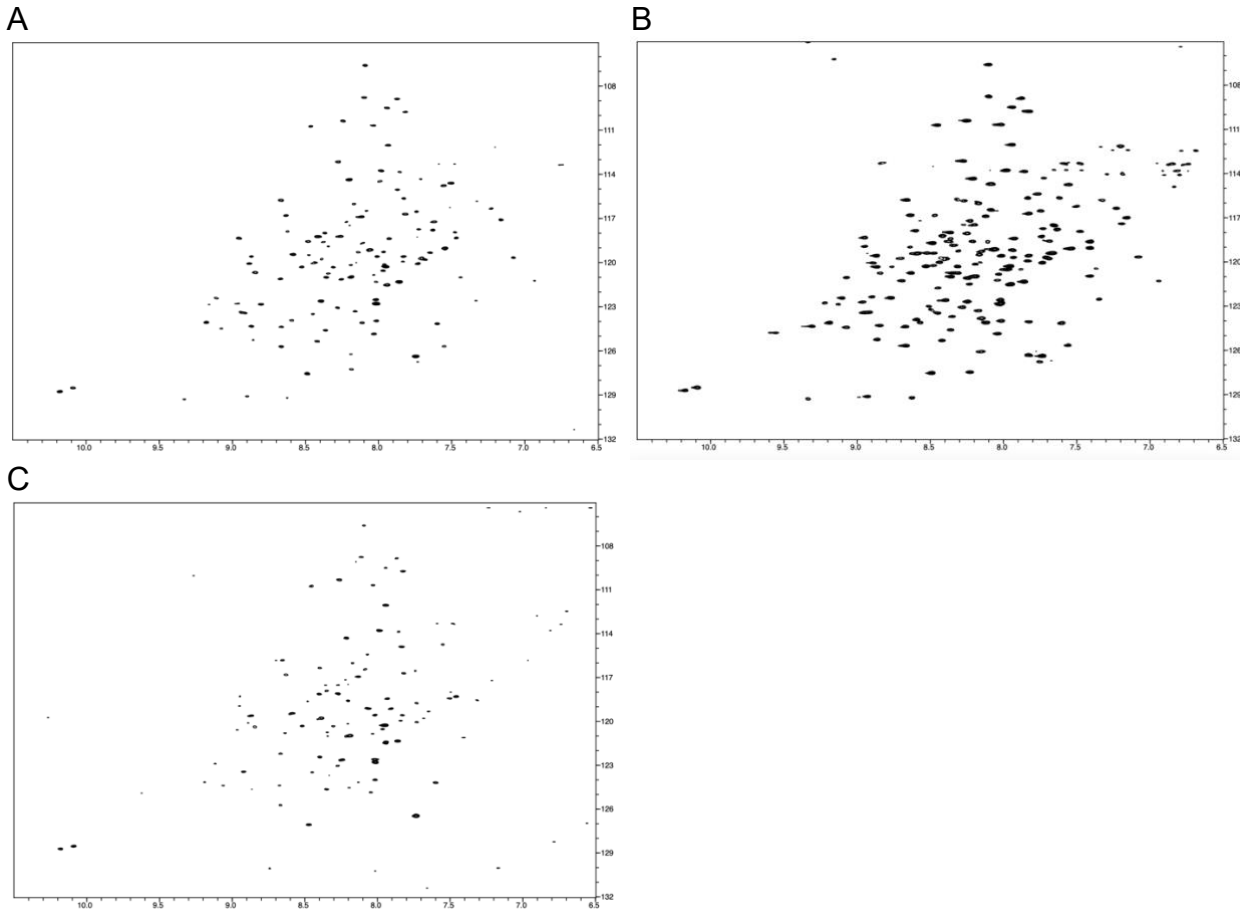


Figure 19. BEST-TROSY spectra of purified mutant endonucleases. *X-axis (*^{1}H, ppm*), y-axis (*^{15}*N, ppm). Samples are composed of 100uM purified mutant endonuclease in endonuclease dialysis buffer containing 150mM NaCl, 20mM Tris, 2mM BME, and 1mM MgCl at pH 7.5. Experimental concentration of protein was kept consistent between the three mutants and was limited by the low purification yield of the K139E mutant (1.15 mg protein per liter of cell culture) A. K137S mutant. B. K137Q mutant. C. K139E mutant.*

Mutant Endonuclease-Protein Interactions

The interaction of the mutant endonucleases with each protein binding partner was assessed by comparing the CSPs between the ^{15}N-labelled endonuclease alone and in the presence of a two-fold molar excess of the unlabeled binding partner.

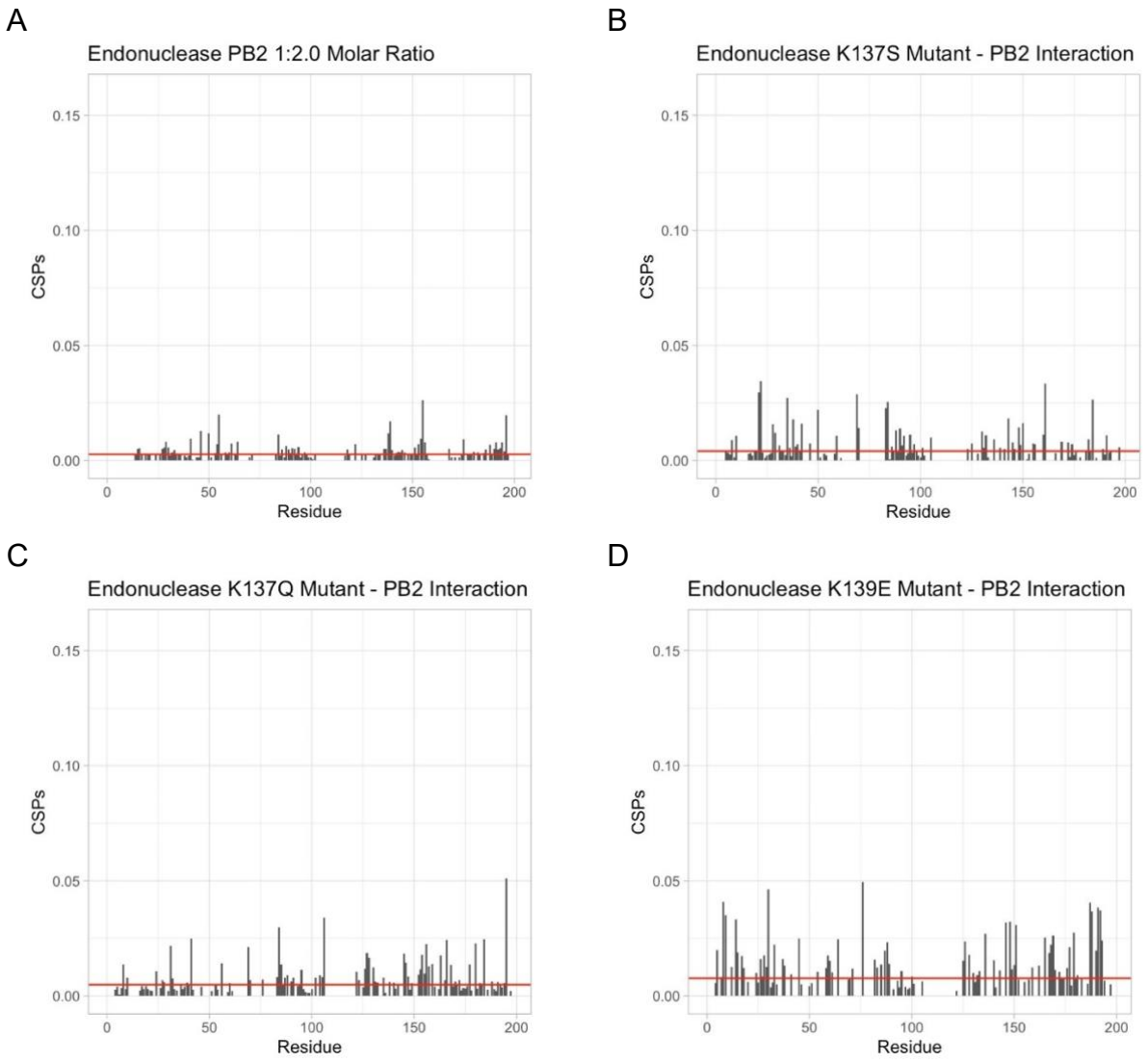


Figure 20. Neither wild type nor mutant endonucleases appeared to interact with the PB2 peptide. *CSPs comparing control endonuclease spectra to spectra of endonuclease with a two-fold molar excess of PB2 peptide. A. WT endonuclease at a concentration of 170 uM with PB2 peptide. B. K137S mutant at 100 uM with PB2 peptide. C. K137Q mutant at 100 uM with PB2 peptide. D. K139E mutant at 100 uM with PB2 peptide. All samples were run in dialysis buffer containing 150mM NaCl, 20mM Tris, 2mM BME, and 1mM MgCl at pH 7.5 All CSPs given in ppm.*

As anticipated, no significant interaction was observed between PB2 and any of the mutants (Figure 20). The largest CSP observed was for E195 in the K137Q mutant (0.051 ppm), which likely a spurious change possibly due to small changes in sample

salt concentration since E195 is a charged residue in the disordered C-terminus. Two additional residues in the K139E mutant, I30 and F76, exhibited CSPs with modest magnitudes (0.046 and 0.049 ppm, respectively).

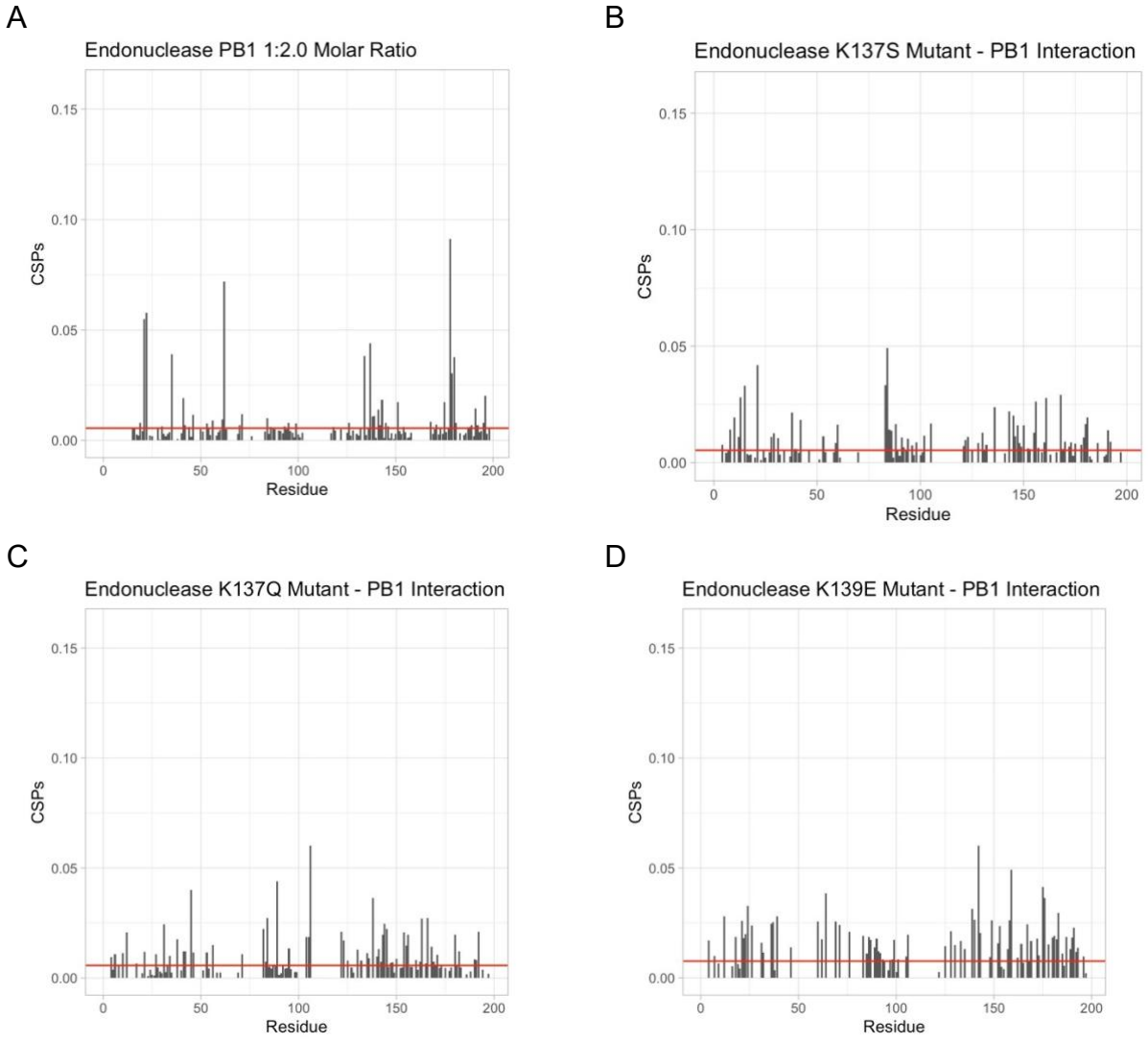


Figure 21. Interactions between WT or mutant endonucleases and PB1 peptide. *CSPs comparing control endonuclease spectra to spectra of endonuclease with a two-fold molar excess of PB1 peptide. A. WT endonuclease at a concentration of 170 uM with PB1 peptide. B. K137S mutant at 100 uM with PB1 peptide. C. K137Q mutant at 100 uM with PB1 peptide. D. K139E mutant at 100 uM with PB1 peptide. All samples were run in dialysis buffer containing 150mM NaCl, 20mM Tris, 2mM BME, and 1mM MgCl at pH 7.5 All CSPs given in ppm.*

Some CSPs potentially indicative of interaction were observed after introduction of a two-fold molar ratio of PB1 peptide in the case of each mutant (Figure 21). In the

case of the K137S mutant, residue R84 shifted 0.049 ppm, residue L106 of the K137Q mutant shifted 0.060 ppm, and residues H142 and A159 of the K139E mutant shifted 0.060 and 0.049 ppm, respectively. Notably, these shifts were fewer and of lower magnitude than those observed in the same titration point for the WT.

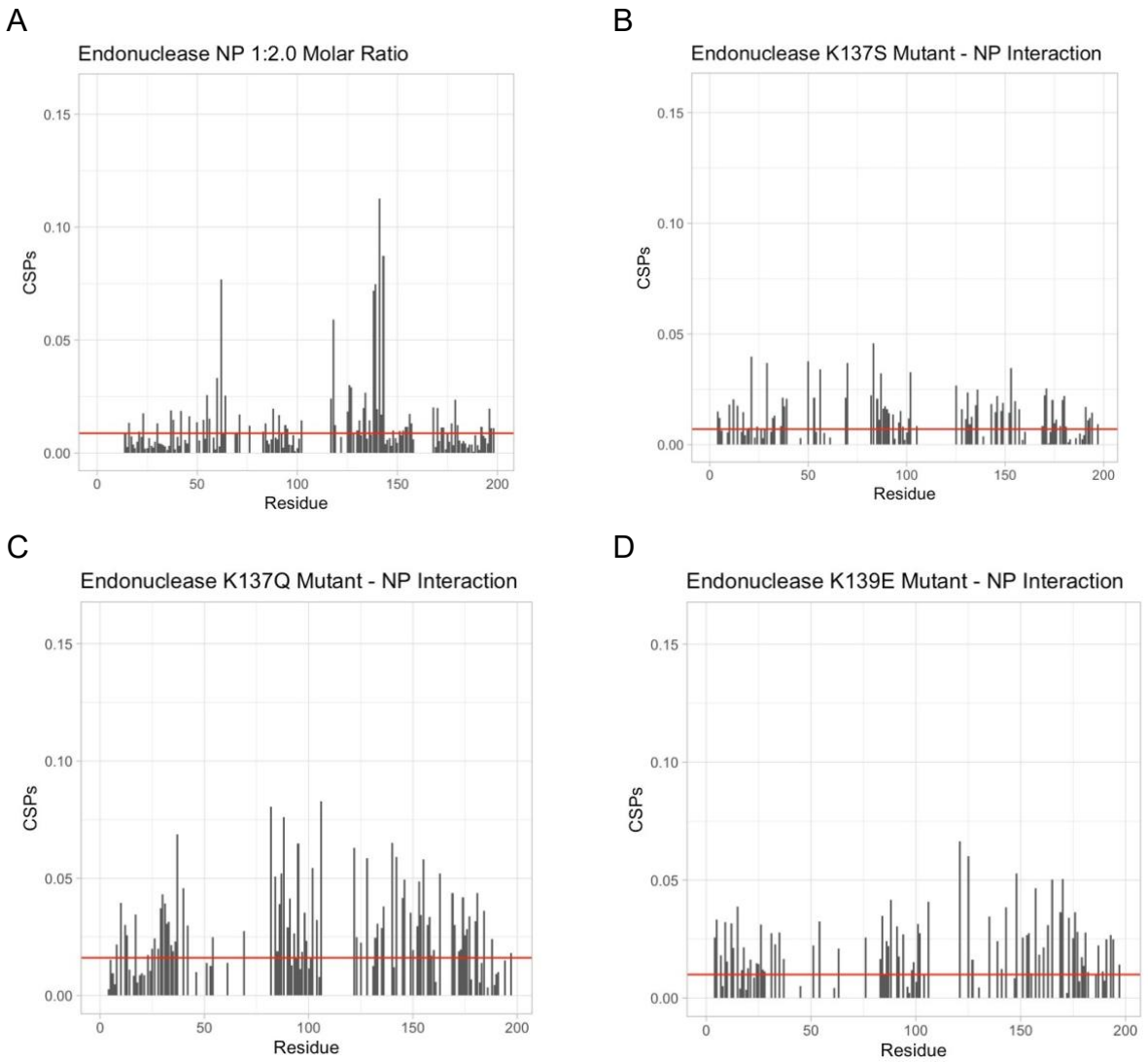


Figure 22. Interactions between WT or mutant endonucleases and PB1 peptide. *CSPs comparing control endonuclease spectra to spectra of endonuclease with a two-fold molar excess of NP for A. WT endonuclease at a concentration of 170 uM. B. K137S mutant at 100 uM. C. K137Q mutant at 100 uM. D. K139E mutant at 100 uM. All samples were run in dialysis buffer containing 150mM NaCl, 20mM Tris, 2mM BME, and 1mM MgCl at pH 7.5 All CSPs given in ppm.*

Discussion

Mutant Characterization by NMR

The signal to noise in the spectra of the endonuclease mutants was lower than that of the WT, as is evident in both Figure 19 and in the much higher mean-CSPs displayed particularly in Figure 22. This reduced signal to noise likely stems from the lower concentration of purified protein used in the mutant NMR trials, which was 100 μM compared with 170 μM that was used for WT. The low purification yields of endonuclease mutant K139E limited the NMR sample protein concentration, and for consistency, all mutant samples were prepared at the same concentrations.

Mutant Interactions with PB1

In the case of the PB1 peptide interaction, it appeared that interaction between the mutant forms of endonuclease and the peptide had been reduced as compared to the WT. All shifts observed were below the threshold CSP considered meaningful in titration experiments with WT endonuclease; relatively larger shifts did occur, but in isolated residues, implying that they are likely charge interactions. Stable protein-protein interactions typically require an interface composed of many residues and thus isolated CSPs are likely to be spurious. It was difficult to state with confidence whether interaction had been completely eliminated between endonuclease mutants and the PB1 peptide from these data since even the WT data at this molar ratio only weakly support interaction.

As discussed in the previous chapter, the interaction between PB1 and endonuclease, though observed in experimental structures of the trimeric polymerase complex, was not compelling at titration points at equimolar stoichiometry. The

endonuclease CSPs only indicated interaction, that is, had CSPs greater than five times the mean, at the highest concentration titration point, a 1.0:3.0 molar ratio, which is in substantial excess of physiological conditions. Shifts were observed earlier, as is evident in Figure 21a, where the CSP for residue 63 = 0.0718 ppm, but the region eventually assessed as interacting with PB1 had much lower CSPs with a twofold molar excess of PB1 peptide. The lack of saturation, even at a very high molar ratio, implies that the interaction between endonuclease and the PB1 peptide is transient (Figure 16). Though the most significant shifts in endonuclease were not observed with added PB1 peptide until a higher concentration of peptide was introduced, I chose to repeat the two-fold molar excess in these experiments with endonuclease mutants to maintain consistency with the other ligands and because some shifts had been observed at the 1.0:2.0 titration point.

No CSPs greater than 0.05 ppm were observed for titrations against the K137S mutant. For the K137Q mutant, residue 107, a proline, had a CSP of 0.060 ppm as compared to a WT CSP of 0 ppm (no peak shift detected), and for the K139E mutant, residue 143, a threonine, had a CSP of 0.060 ppm as compared to the WT CSP of 0.010. The shift in residue P107 was assessed as likely the result of shifts due to interactions with partial positive charge of the proline ring face (Zondlo, 2014), in part because it seemed unlikely that unmutated residues which did not appear to interact with the PB1 peptide in the WT endonuclease would do so in mutants where substantial structural rearrangement of the protein had not been observed. Interestingly, the shift observed in residue 143 of the K139E mutant exceeded the CSP at the same titration point in the WT. Residue 143 was considered the end of the interacting region between

WT endonuclease and PB1 though, as discussed above, saturation was not achieved to support specific binding (CSP with a threefold molar excess of PB1 peptide = 0.093 ppm) (Figure 16d). This result strongly implies that the shift observed in residue 143 of the WT endonuclease was due to charge interactions. It should, however, be noted that the spectra collected with the K139E mutant had the lowest signal to noise; residues for which peak picking was possible in the K139E : PB1 peptide spectrum had a mean CSP of 0.016 ppm and a corresponding cutoff for significance of 0.080 ppm, greater than any of the shifts observed. Further, the mutation from positively charged lysine to negatively charged glutamic acid may have had an effect on the partial positive charge of the threonine ring face contributing to the charge sensitivity of residue 143 upon the introduction of PB1 peptide which would not have occurred in the WT or other mutants.

These data support the conclusion that the region 137-142, which appeared to interact with the PB1 peptide in my WT experiments, does not interact with endonuclease K137S or K137Q. Residue T143 had a CSP of 0.060 ppm in the K139E mutant, perhaps implying that the shift observed at residue 143 of the WT was due to a charge interaction. While further investigation is warranted to see whether introduction of PB1 peptide at the highest molar ratio, 1.0:3.0, induces interaction in these areas of the mutant endonucleases, it seems unlikely that regions which did not appear to interact with the PB1 peptide in the WT endonuclease would do so in mutants.

Mutant Interactions with NP

Though introducing mutations into endonuclease appeared to reduce endonuclease interactions with NP in the region surrounding the point mutation, there may be interactions with NP in other regions of the protein. Given the physiological

context of the NP endonuclease interaction, where endonuclease is suggested to increase local concentration of NP to facilitate RNP formation, the potential persistence of transient protein-protein interactions between NP and endonuclease mutants is unsurprising. However, the WT does clearly display the most compelling argument for direct interactions between endonuclease and NP. CSPs for the other mutants are lower objectively and relatively to the mean, with no CSPs greater than 0.10 ppm in any mutant residues. Consideration of CSPs relative to the mean is a particularly important consideration in the case of the K137Q mutant, where fifteen residues have CSPs > 0.05 ppm, more than double the number of residues with shifts greater than 0.05 ppm for the WT. However, lower quality of the spectrum resulted in higher CSPs across all residues and these data therefore do not indicate a higher likelihood of interaction with NP. The introduction of mutations into the region from residues 137-142 knocked out interaction between endonucleases and NP, strengthening the evidence of interaction between NP and the WT.

The data in this chapter support the conclusions made in Chapter 3, confirming the region of endonuclease interacting with NP and PB1 and providing especially strong evidence for binding between NP and endonuclease.

Chapter 5: Activity Assays

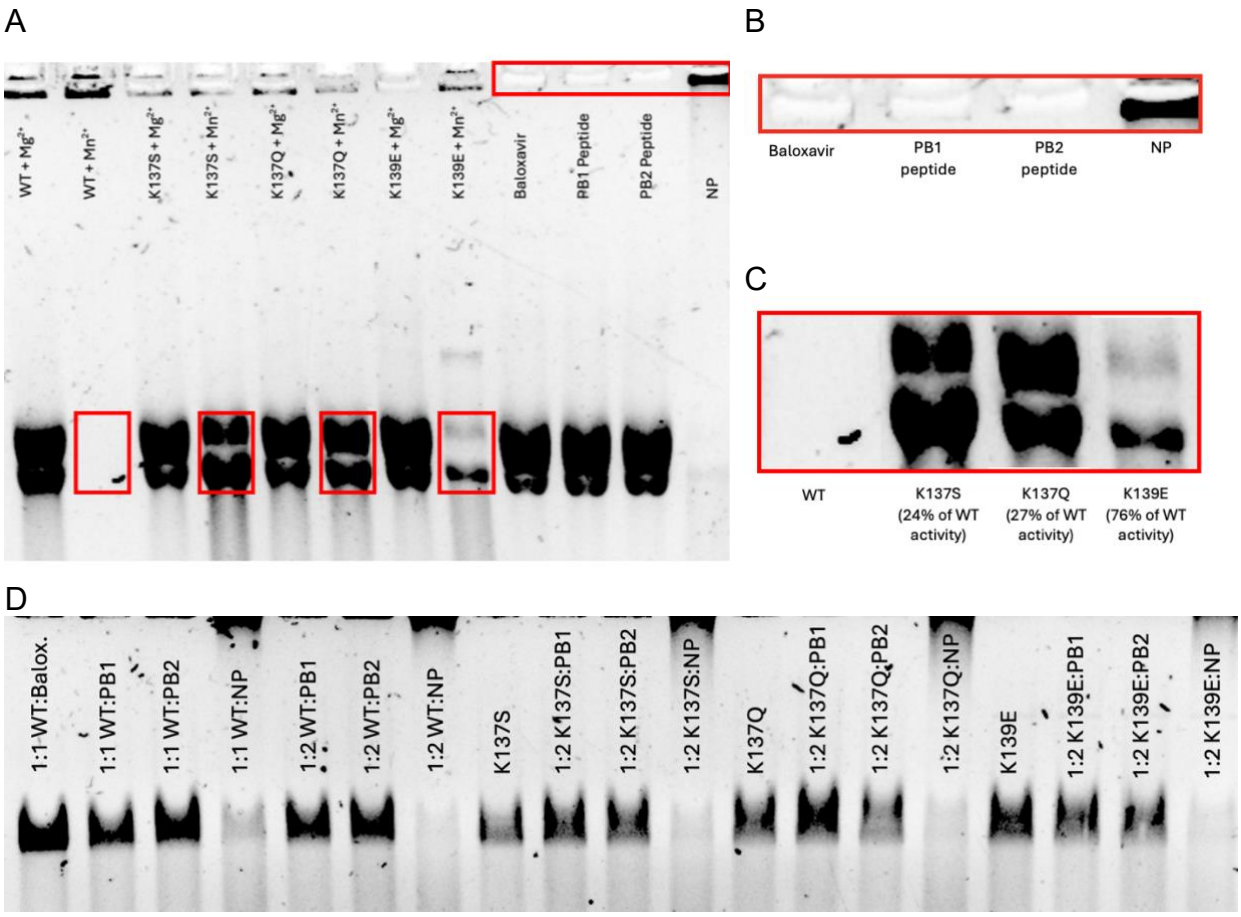
Regulation of PA endonuclease function

Structure and function studies of the polymerase from influenza type B suggest that the endonuclease activity might be regulated simply by access to RNA substrate (Thierry et al., 2016). Because of its role in transcription and therefore in the influenza life cycle, the endonuclease active site has become a drug target (Omoto et al., 2018). Endonuclease inhibitors have been used as antivirals for influenza treatment, though improvement of patient outcomes is modest and resistance is common, likely due to the rapid mutation rate of the virus and its corresponding ability to develop resistance to antivirals (Gao et al., 2024). Understanding whether endonuclease is also regulated allosterically could indicate a strategy toward developing novel antivirals. The endonuclease residues observed interacting with NP and the PB1 peptide appear from solved structures to sit on the edge of the endonuclease active site, suggesting that mutations could interfere directly with the enzyme activity; I therefore wanted to test changes in endonuclease activity with the addition of protein ligands.

Activity of WT and Mutant Endonucleases

An activity assay modeled after published experiments (Pala et al., 2015) was conducted to test enzyme activity in the presence of NP, PB1 peptide, and PB2 peptide. Both $Mg^{2+}$and Mn^{2+} were used as metal ion catalysts, with observed activity significantly higher in the presence of manganese than magnesium, as reported previously (Dias et al., 2009). Control gels (Figure 23a) were run for each mutant and ligand to monitor their effects on the added ssDNA. ImageJ was used to analyze the mean gray values of a consistently sized selection which was compared across the assay gels to assess

activity (Table 1), as per previous protocols (Pala et al., 2015). In samples containing NP, a substantial amount of the ssDNA remained in the well, which is discussed below. This phenomenon was observed to a lesser extent in some endonuclease samples, as with the WT + Mn^{2+} control sample in Figure 23a, likely because some ssDNA remained bound to the enzyme active site after completion of the assay.



E

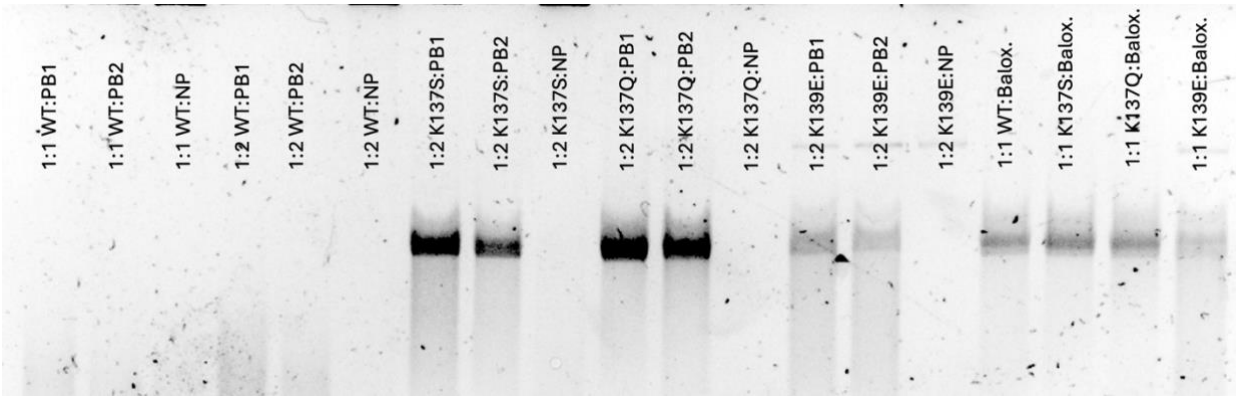


Figure 23. Endonuclease Activity Assays. *A. Control samples assessing activity of WT and mutant endonucleases, Baloxavir, and ligands. B. Expanded view of ligand control wells, with clear ssDNA in the NP well and little, if any, ssDNA visible in the wells of the other ligands. C. Expanded view of WT and mutant endonucleases with Mn^{2+}, with note of the mutant activity in comparison to WT. D. WT and mutant endonuclease activity with Mg^{2+} bound. E. WT and mutant endonuclease activity with Mn^{2+} bound. Individual samples can be identified based upon lane labeling and are repeated in Table 1. Samples were assembled to 25 µL with 1.5 µg M18 ssDNA (TebuBio), added last, 1 µg purified endonuclease, and the appropriate volume of ligand in sample running buffer containing 150mM NaCl, 20mM Tris, 2mM BME, and either 1mM MgCl or 1mM MnCl. Assembled samples were incubated for 2 hours at 37℃ before being heat shocked for 20 minutes at 80℃ to denature the enzyme. 2.5µL 10x gel loading dye was added to each sample to allow for visualization via gel electrophoresis. All samples were run on 1% agarose DNA gels with 1uL SYBR Safe DNA gel stain (ThermoFisher) per mL gel.*

Manganese-Bound Endonucleases Show Greater Activity Levels

The difference in endonuclease activity with Mg^{2+} and Mn^{2+} was clear in the control gel. For WT endonuclease, the enzyme had a greater than five-fold increase in activity level with Mn^{2+}in the assay buffer as compared to Mg^{2+}. The difference in activity was not as extreme with the endonuclease mutants, where manganese gave a 1.06, 1.19, and 4.44-fold increase in activity for K137S, K137Q, and K139E, respectively, related to the overall impact of the mutation enzymatic activity. In contrast to the results displayed in Figure 23b and c, where endonuclease activity was tested with ligands present, the ssDNA in Figure 23a appears as a double band, with a large, thick band of DNA and a smaller, bow-tie shaped band of DNA below it. Control lanes 9-11 containing Baloxavir, PB1 peptide, and PB2, where no activity was predicted, show the same

banding pattern and both bands appeared degraded; I therefore concluded that the ssDNA sample was not homogenous which, combined with poor quality of the acquired image similarly evident in Figure 23b, resulted in the double band. The entire area containing the double band was analyzed when assessing the amount of ssDNA that had been cleaved.

Activity of Magnesium-Bound WT and Mutant Endonucleases

In the case of magnesium-bound endonuclease, the PB2 peptide appeared to have minimal, if any, effect on activity in all endonucleases. The PB1 peptide showed some signs of endonuclease inhibition in the K137Q and K139E mutants, but little to no effect on the WT or K137S samples. As mentioned above, all NP samples showed little to no ssDNA escaping the well. The enzymatic activity detected in assays containing magnesium was significantly lower than that of manganese samples, in accordance with previously published results (Dias et al., 2009, DuBois et al., 2012).

Activity of Manganese-Bound WT and Mutant Endonucleases

In contrast to the magnesium-containing samples, manganese-bound endonucleases had stark differences in apparent activity levels. All WT endonuclease samples demonstrated almost complete degradation of the ssDNA sample, with no visible band of full-length ssDNA. Since cleavage of the ssDNA was complete, it is possible that the WT endonuclease degraded the ssDNA in the sample in less than the maximum incubation time. All comparisons of mutant activity therefore give the minimum reduction in activity.

The inclusion of mutations had variable effects on endonuclease activity. Once again, the NP bound samples showed no remaining ssDNA on the gel. The K137S

mutant when incubated with PB1 appeared to have a minor reduction in activity as compared to the inclusion of PB2. Similarly, a slight decrease in mutant activity with PB1 as compared to PB2 was observed in the K137Q and K139E mutants, though the difference in activity was most obvious in the K137S sample.

The K137S and K137Q endonuclease mutants with manganese had reduced activity, which was 24% and 27% WT, respectively. Changes in activity were less apparent in the magnesium bound samples because of their relatively decreased activity in all cases. The K139E mutant exhibited a modest loss of function, with 76% of WT activity. Interestingly, the mutant samples saw only slightly reduced activity in the presence of Baloxavir, and mutant activity was lower in the presence of the PB1 and PB2 peptides than in a sample with Baloxavir added.

Samples including K139E endonuclease showed a faint band above the ssDNA band which remained for all K139E samples. Its pervasiveness and consistent intensity indicate that the sample may have had a small nucleic acid contaminant. If this assumption is correct, the contaminant likely would not have had any effect on the results since the positioning and intensity of the contaminant band are consistent, indicating minimal interaction with the endonuclease.

Table 1

Sample Number	Mean Gray Value	% of WT or free endonuclease activity	Sample Name	Metal Bound	Ratio of endonuclease to ligand
1	40.994		WT	Mg^{2+}	
2	231.568		WT	Mn^{2+}	
3	53.679	131	K137S	Mg^{2+}	

4	57.064	25	K137S	Mn^{2+}	
5	52.969	129	K137Q	Mg^{2+}	
6	62.901	27	K137Q	Mn^{2+}	
7	40.091	98	K139E	Mg^{2+}	
8	177.939	76	K139E	Mn^{2+}	
9	52.093	22	Baloxavir	Mn^{2+}	
10	50.796	21	PB1	Mn^{2+}	
11	52.942	22	PB2	Mn^{2+}	
12	234.2	101	NP	Mn^{2+}	
13	109.55	267	WT + Baloxavir	Mg^{2+}	1:1
14	142.57	108	WT + PB1	Mg^{2+}	1:1
15	120.634	91	WT + PB2	Mg^{2+}	1:1
16	213.972	163	WT + NP	Mg^{2+}	1:1
17	138.378	105	WT + PB1	Mg^{2+}	1:2
18	125.052	95	WT + PB2	Mg^{2+}	1:2
19	228.038	173	WT + NP	Mg^{2+}	1:2
20	171.866	130	K137S	Mg^{2+}	
21	137.037	79	K137S + PB1	Mg^{2+}	1:2
22	146.615	85	K137S + PB2	Mg^{2+}	1:2
23	226.079	131	K137S + NP	Mg^{2+}	1:2
24	146.887	129	K137Q	Mg^{2+}	
25	118.151	80	K137Q + PB1	Mg^{2+}	1:2
26	166.832	113	K137Q + PB2	Mg^{2+}	1:2
27	225.536	153	K137Q + NP	Mg^{2+}	1:2
28	128.526	97	K139E	Mg^{2+}	
29	148.185	115	K139E + PB1	Mg^{2+}	1:2
30	158.161	123	K139E + PB2	Mg^{2+}	1:2
31	231.132	179	K139E + NP	Mg^{2+}	1:2
32	246.089	106	WT + PB1	Mn^{2+}	1:1

33	242.277	104	WT + PB2	Mn^{2+}	1:1
34	241.675	104	WT + NP	Mn^{2+}	1:1
35	239.016	103	WT + PB1	Mn^{2+}	1:2
36	235.314	101	WT + PB2	Mn^{2+}	1:2
37	232.586	100	WT + NP	Mn^{2+}	1:2
38	52.306	91	K137S + PB1	Mn^{2+}	1:2
39	114.204	200	K137S + PB2	Mn^{2+}	1:2
40	230.533	403	K137S + NP	Mn^{2+}	1:2
41	39.368	62	K137Q + PB1	Mn^{2+}	1:2
42	51.68	82	K137Q + PB2	Mn^{2+}	1:2
43	230.989	367	K137Q + NP	Mn^{2+}	1:2
44	176.963	99	K139E + PB1	Mn^{2+}	1:2
45	203.196	114	K139E + PB2	Mn^{2+}	1:2
46	239.933	134	K139E + NP	Mn^{2+}	1:2
47	182.084	78	WT + Baloxavir	Mn^{2+}	1:1
48	172.024	301	K137S + Baloxavir	Mn^{2+}	1:1
49	182.258	289	K137Q + Baloxavir	Mn^{2+}	1:1
50	212.13	119	K139E + Baloxavir	Mn^{2+}	1:1

Implications of Mutations and Ligand Binding on Endonuclease Activity

Assay Methods

Though influenza is an RNA virus, ssDNA was used in the assay method because of its increased stability; I do not anticipate that there would be any meaningful changes in detected activity with ssDNA since endonuclease in a promiscuous enzyme. Previous publications have used ssDNA in place of RNA substrate and successfully

analyzed enzymatic activity (Pala et al., 2015). Assay methods have been published using magnesium or manganese in the sample buffer, where manganese is consistently assessed as promoting higher levels of endonuclease activity than magnesium (Dias et al., 2009; Todd et al., 2021). While magnesium may play a more relevant role in endonuclease function in the live virus, the *in vivo* reality of metal binding to endonuclease was beyond the scope of this study, so both potential catalysts were tested.

The WT endonuclease digestion of ssDNA appeared to be almost complete in the presence of manganese and was greater than five-fold the amount of activity observed in the presence of magnesium, consistent with previous studies (Dias et al., 2009; DuBois et al., 2012). Since enzymatic activity resulted in complete degradation of the product, kinetic understanding of WT activity was lacking, and it was difficult to state conclusively the effect that mutation had on endonuclease function. The estimates of change in activity displayed in Table 1 are, in fact, a minimum estimate of the loss in endonuclease function. If, for example, the full volume of added ssDNA was digested by WT endonuclease in one hour (half of the full incubation time of the assay), then WT activity would be two times greater than assessed by these methods. Further activity assays could be designed to more accurately assess the difference between WT and mutant activity by introducing time points. Assays of different durations (e.g. 5, 10, 15, 30, 60, 90, and 120 minutes) could be conducted for all samples and results compared to understand WT endonuclease activity more fully and to provide greater insight into kinetic differences in WT, mutant, and ligand-bound enzymatic function.

Though the activity of free WT and mutant endonucleases was significantly greater with manganese than with magnesium, the percent change in activity observed upon the introduction of ligands was relatively uniform. I did not draw conclusions about the likely physiological relevance of magnesium versus manganese and considered both sets of results in the following discussion of ligands' effects on enzymatic activity.

Changes in Mutant Enzymatic Activity

All three mutations introduced had comparable levels of activity to the WT in the presence of magnesium but had significant loss of function in the presence of manganese. Both endonucleases with mutations at residue 137 showed less than 30% of WT activity in the manganese buffer, and the K139E mutant showed ~76% of WT activity. It is unlikely that such a dramatic loss of function would be contingent upon the type of metal bound, which invites further consideration of the control result for WT endonuclease activity with magnesium buffer. Of note, the sample of WT endonuclease with Baloxavir in Mg^{2+} buffer showed greater activity than the control, while the expected result would have been a significant loss of function upon introduction of an industrial endonuclease inhibitor (Figure 23b). It is possible that the Baloxavir stock in use was faulty, as will be discussed in a later section, but given this result, WT + Mn^{2+} samples were considered a better metric for comparison to WT activity.

Based upon the data from this assay, I concluded that the K137S and K137Q mutants induce almost complete loss of enzymatic function and that the K139E mutation causes some, but not complete, reduction of enzymatic activity in the presence of Mn^{2+}. As discussed above, I did not observe the same dramatic result using Mg^{2+} buffer, though this could also be due to the substantially lower WT levels of activity,

which would make changes in activity more subtle and therefore more difficult to check using gel imaging.

However, the observed difference in effect on activity could present an opportunity for further research into the ion most relevant for endonuclease activity *in vivo.* If further experiments on mutant endonuclease activity support the initial result, that is, that mutant activity is significantly reduced with a manganese catalyst but not with a magnesium catalyst, then *in vivo* experiments investigating the change in mRNA production with mutants could provide insight into the more physiologically relevant metal. If transcription were to proceed at a similar rate to that observed in the WT, one could conclude that magnesium is more important as a catalyst. If transcription were significantly impaired by the introduction of a mutant, one would expect that magnesium is more important to enzymatic activity.

Ligand Mediated Changes in Enzyme Activity

Mutant endonucleases alone did appear to have some activity in the manganese samples, though it was significantly reduced as compared to WT endonuclease. The PB1 samples for all three mutants displayed slightly less enzymatic activity than the PB2 peptide samples, perhaps suggesting that even very minimal interaction with the PB1 peptide, as apparent from the NMR data on mutant interaction with PB1 peptide, does have some inhibitory effect on endonuclease function. The pattern of banding and the slight decrease in endonuclease function across the PB1 peptide-containing samples does suggest some perturbation, though the extent of this inhibition could be sufficiently transient as to be functionally irrelevant for the WT.

It is also possible that there was no interaction at all between the mutant endonucleases and PB1 peptide, which would accord with the previously discussed data from which I concluded that the introduction of mutations effectively knocked out interaction between mutant endonucleases and the PB1 peptide. The changes in mean gray value between free mutant endonucleases and PB1 peptide bound endonuclease were minimal as compared to Baloxavir and could be due to pipetting error.

Baloxavir in Endonuclease Activity Assays

To analyze significant changes in enzymatic activity, I first intended to consider the changes in endonuclease activity between free WT endonuclease and the WT samples containing Baloxavir. Since Baloxavir is an approved treatment for influenza infection (Omoto et al., 2018) because of its ability to outcompete RNA for the endonuclease RNA-binding site, I assumed the effect of Baloxavir on WT endonuclease activity would be very significant and could be applied to analysis of the other samples to determine significant changes in activity.

However, the samples run with WT or mutant endonuclease and Baloxavir gave an interesting result: the activity of WT and mutant endonucleases in the presence of manganese appeared to be comparable when Baloxavir was introduced. The K137S and K137Q mutations, which appeared in the control gel to cause a substantial loss of function as compared to the WT, seemed to have some function restored by Baloxavir. In contrast, WT endonuclease with manganese behaved as anticipated, with the drug inhibiting activity. Previous publications noted that the introduction of an I38T mutation in endonuclease conferred resistance to Baloxavir because it limits Baloxavir access to the active site (Todd et al., 2021). Presumably the polar sidechain of threonine interacts

with the drug, preventing binding. The point mutations introduced at residues 137 and 139 are positioned near the mouth of the endonuclease active site, making it plausible that they would affect Baloxavir binding. Per the AlphaFold 3 models predicted in Figure 18, the K137S and K137Q mutations appear to face towards the catalytic site, replacing the positively charged lysine, which would attract nucleic acids, with polar serine or glutamine, causing a loss of function. If the mutants interacted with the drug in such a way that the polar sidechains were reoriented away from the catalytic site, it could restore function by minimizing the steric clash between the enzymatic pocket and incoming nucleic acids. It is also possible that samples were mislabeled giving a strange result of that the Baloxavir stock being used was somehow faulty or mislabeled.

The above hypothesis on Baloxavir interaction with mutants could be tested with a series of follow up experiments. First, to confirm that the result is valid, the initial assay of WT and mutant endonucleases with Baloxavir should be rerun. Titrations or even single-point experiments of Baloxavir with each of the mutants should be conducted via NMR to assess whether Baloxavir binding is maintained in the mutant endonucleases and whether the binding site and resultant conformational changes are dramatically different than those observed in the WT. These data could be supported by fluorescence polarization experiments to calculate binding affinity of Baloxavir for the WT versus mutant proteins. If initial experiments suggest binding of Baloxavir to the mutant endonucleases, the assay would be repeated with increasing concentrations of Baloxavir to see whether enzymatic function is inversely correlated with drug concentration in the WT but directly correlated in the mutants.

The relationship between Baloxavir and mutant endonuclease activity provides interesting insight into the intricacies and potential difficulties of treating influenza infection. Clinical analysis of the success of Baloxavir as an antiviral found a relatively high prevalence (9.7%) of mutations conferring resistance to the drug, and epidemiological studies have found no significant improvement in patient outcomes on Baloxavir, likely due to the rapid emergence of resistance (Todd et al., 2021; Gao et al., 2024). In mutants where enzymatic activity is no longer inhibited by Baloxavir, it is possible that some function could even be restored. The brief window for which small molecules are effective in treatment of flu emphasizes the importance of novel drug targets. A more complete understanding of the influenza replication system, including endonuclease's role and interactions, is necessary for the development of treatments which continue to be effective despite the influenza virus' rapid mutation rate.

Summary and Next Steps

In this chapter, I assessed the changes in endonuclease function based upon introduction of mutations and binding partners to ascertain the regulatory effects of binding with the selected ligands, in keeping with the final aim of this thesis. Contrary to the initial hypothesis, interactions with binding partners did not alter endonuclease activity. However, introduction of mutations, particularly the K137S and K137Q mutants, showed dramatic loss of function – at least a 75% and 73% decrease in enzymatic activity, respectively. The identification of these mutants provides a potential target for future research into antivirals targeting endonuclease function.

Discussion

The overall aim of this thesis was to increase understanding of the PA endonuclease's role in various conformations of influenza A FluPol, especially those that are cap-snatching incompetent. Interaction with host ANP32 forms a FluPol dimer which is crucial to viral replication (Carrique et al., 2020); in this replicative platform, endonuclease should not be enzymatically active because replication is initiated *de novo* in influenza and cap-snatching is therefore not required (Fan et al., 2019). I hypothesized that interactions with endonuclease driven by need to render the enzyme catalytically inactive could contribute to structural changes in FluPol structure, where different binding partners might change conformations to regulate enzymatic activity.

Based on the current literature, I selected three potential interaction partners: NP, a PB1 peptide from residues 100-115, and a PB2 C-terminal peptide (Staller et al., 2024; Thierry et al., 2016; Vidic et al., 2016). To test the potential binding of FluPol with NP, PB1, and PB2, I first purified endonuclease alone, analyzed it via NMR, and assigned the protein to >85%. Each of the potential binding partners was individually titrated into endonuclease in solution and the resultant spectra were compared to the solution state. The high sensitivity of NMR allowed us to analyze likely binding sites, even for low affinity interactions which might require external stabilization by other subunits *in vivo*. In the endonuclease, I observed shifts indicative of binding in residues 137-142, the disordered region following α-helix 5, after introduction of both NP and PB1 peptide. The PB1 peptide introduced was specifically selected because of its apparent surface contacts with the endonuclease in $FluPol_E$, so the overlapping region of interaction and previous speculation on the nature of the interaction between

endonuclease and NP lead us to conclude that NP likely binds endonuclease in $FluPol_R$. My data indicating that NP does interact with endonuclease, likely in $FluPol_R$, strongly supports previous work hypothesizing an interaction between endonuclease and NP (Vidic et al., 2016).

An endonuclease-NP interaction in the replicating conformation of the dimer could work in tandem with potential NP-ANP32 LCAR binding to recruit NP for vRNP formation from the replicative platform (Wang et al., 2022). The combined observation of NP interaction and the lack of cap-snatching by the endonuclease in $FluPol_R$ led to the hypothesis that NP could be contributing to regulation of enzymatic activity.

We proceeded with a set of activity assays to determine whether PB1 peptide or NP binding affects activity. Neither PB1 nor PB2 peptide had a significant effect on endonuclease activity in my ssDNA assay. It initially appeared that NP had a significant activating effect on endonuclease. However, this result seemed highly unlikely given the role of NP in binding vRNA. I therefore suspect that the ssDNA used in the assay bound the NP introduced and therefore was stuck in the wells. Further experiments assessing changes in enzyme activity with NP could be conducted to confirm the effect of NP on endonuclease activity. The assay outlined here could be repeated with an additional step denaturing the DNA from the NP.

The overlap in interacting regions with PB1 peptide and NP led us to further investigate the disordered region following α-helix 5. I introduced point mutations at residues K137 and L139 to try to knock out interaction and to test whether the region has any impact on activity. The resultant spectra suggested that point mutations had little effect on the overall structure of the endonuclease but did substantially affect

binding, with both mutations of K137 resulting in lower perturbations of the spectra upon addition of ligands. The K139E mutant also saw some changes in interactions but these were harder to analyze because of the lower quality of the spectrum, an issue which might be resolved with follow-up experiments using a higher concentration of K139E endonuclease in the spectra.

Here I provide, to my knowledge, the first high resolution characterization of the PA endonuclease by NMR, and show evidence of an interaction between the endonuclease and NP, and possibly also with a PB1-derived short peptide. my data support the idea that enzymatic activity is regulated by structural changes restricting substrate access, rather than direct inhibition of endonucleolytic activity, though further research is needed particularly to investigate effect of NP on endonuclease activity. I also identified two point-mutants, K137S and K137Q, which appear to almost completely arrest enzymatic activity and one mutant, K139E, which displays some loss of function. Discovery of these mutants may facilitate further studies into the role of the flexible region between resides 137-142 in the endonuclease which I have identified as crucial to external interaction and normal enzymatic function.

The initial hypotheses of this thesis were that endonuclease would undergo structural changes to bind potential interaction partners and that interactions with these external binding partners would alter endonuclease activity. Here I provide a solution-state model of endonuclease using NMR, which acted as the foundation for my subsequent experiments. I confirmed my first hypothesis with strong evidence for the interaction between endonuclease and NP, as mediated by the disordered region between residues 137-142 of the endonuclease, and validated this interaction by

introducing point mutants which knocked out binding. Endonuclease interaction with the selected PB1 peptide was also indicated, albeit less convincingly, and binding between WT endonuclease and the PB2 peptide did not occur, in contrast to the anticipated result. These data in particular emphasize the importance of the substantial surface contacts between the subunits of the polymerase in stabilizing conformational changes. I then tested the impact of binding on endonuclease activity and found no meaningful changes in activity upon the introduction of PB1 and PB2 peptides, but was unable to confidently assess whether NP had any effect on activity because of ssDNA binding to the protein. Importantly, assays investigating the enzymatic activity of mutant endonucleases revealed that point mutations at position K137 have a substantial inhibitory effect on endonuclease activity. Further investigation into the disordered region identified in this thesis as interacting with both NP and the PB1 peptide could serve as the basis for novel drug development, where focus on inhibiting binding and enzyme activity to disrupt normal viral replication could provide an alternative pathway for strains of influenza with acquired resistance to other treatments.

Materials and Methods

Expression plasmids and cloning

An N-terminally GST-tagged constructs of endonuclease, WT NP, and R416A NP (used for this thesis) in pGEX were generously provided by the Fodor lab, Dunn School of Pathology, Oxford. Endonuclease was then cloned into an N-terminally His-tagged construct using a pET15b backbone provided by the Lau lab, Department of Biochemistry, Oxford.

Purification of Endonuclease

200 µL of BL21(DE3) competent cells were incubated on ice with 1 µL pET15 endonuclease plasmid at a concentration of 20.6 ng/µL for 30 minutes. A further 1 µL plasmid was mixed into the competent cells, then the sample was heat shocked for 30 seconds at 42˚C. The cells were left to recover on ice for ten minutes, then plated onto high salt LB (LBHS) plates with ampicillin (working concentration of 100mg/L). Plates were incubated overnight at 37˚C.

After overnight growth, one colony was selected from the plate, avoiding larger colonies with visible satellite colonies, and introduced to 100 mL LBHS + ampicillin. This starter culture was incubated overnight at 37˚C with shaking (180 rpm). 25 mL overnight culture and 1 mL ampicillin at 100mg/mL were added to each liter of LB and were grown at 37˚C with shaking. The cultures were then induced with 1 mM isopropyl β-D-1-thiogalactopyranoside and left overnight at 18˚C with shaking.

Cells were harvested via centrifugation for 20 minutes at 5000 x *g* and 4˚C in a Beckman Avanti centrifuge using the JLA 8.1 rotor. The cell pellet was then

resuspended into 30 mL of wash buffer containing 300 mM NaCl, 15 mM imidazole, 50 mM Tris-HCl, 5 mM β-mercaptoethanol (BME), protease inhibitors were added (Roche cOmplete Mini EDTA-free Protease Inhibitor Cocktail), and the pellet was left to freeze at –80˚C.

Frozen pellet was thawed, wash buffer was added to a total volume of 35 mL, and the pellet was resuspended by pipetting. Cells were then lysed by sonication with an amplitude of 45 J, with 5 seconds of active sonication followed by 25 seconds rest for a total active sonication time of 2 minutes 30 seconds. Lysed cell debris was removed by centrifugation for 45 minutes at 30,000 x g and 4˚C in a Beckman Avanti centrifuge using a JA 25.50 rotor. While centrifugation proceeded, cOmplete Ni-NTA His-tag purification resin (Roche) was equilibrated for gravity flow. 0.5 mL bead volume was equilibrated per liter of cell culture by washing with 12 column volumes (CV) MilliQ water then 12 CV wash buffer. The equilibrated beads were incubated with the cell lysate for 60-90 minutes, then the lysate-bead slurry was applied to the column. The beads were washed with 12 CV wash buffer, then 12 CV elution buffer containing 300 mM NaCl, 300 mM imidazole, 50 mM Tris-HCl, and 5 mM BME was applied to the beads, which were allowed to incubate in the column on a tube roller for 60 minutes. The initial elution was then collected and a further 12 CV elution buffer was applied and collected to ensure thorough collection of the target protein. The elution volume was concentrated via centrifugation to a protein concentration of ~400 μM and dialyzed into dialysis buffer containing 150 mM NaCl, 20 mM Tris, 2 mM BME, and 1 mM MgCl at pH 7.5 overnight. The resultant protein was stored at 4˚C until use.

Labeled Purification

To prepare PA endonuclease for visualization by NMR, protein was singly (^{15}N) or doubly (^{15}N, ^{13}C) isotopically labeled as follows. A minimum two liters LBHS were inoculated with overnight culture and grown at 37˚C with shaking (180 rpm) to OD(600 nm) = 0.6-0.8. While the cells grew, 1x and 2x M9 mixtures were prepared under sterile conditions. For 1x M9, 100 mL of 10x M9 salts solution was added to 900 mL autoclaved MilliQ water. For labeled 2x M9, 200 mL 10x M9 was added to 800 mL autoclaved water and supplemented with 1 mL thiamine, 1 mL biotin, and 1 mL ampicillin, all at a concentration of 100mg/mL. To this was added 20 mL sterile 20% glucose (if labeling only with ^{15}N) and 2 mL magnesium sulfate (1M). 500 µL trace metals were then added and the bottle shaken immediately to prevent precipitation. 1 g ^{15}N ammonium sulfate and 2 g ^{13}C glucose (if double labeling) were added and shaken to mix. 1x and 2x M9 buffers were then stored at 4 C until use, not more than 24 hours.

After growing to the prescribed OD, cells were harvested via centrifugation for 20 minutes at 5,000 x g. The cell pellets from two litres of cultures were consolidated and resuspended into 1 L of 1x M9 salts. These cells were again collected by centrifugation for 20 minutes at 5,000 x g, then resuspended into 2x M9 and allowed to rest at 25˚C with no agitation before being induced. The cultures were then induced with 200 mg/L and incubated overnight at 18˚C with shaking (180 rpm). Cells were then harvested and purified according to the same protocol as the unlabeled sample.

Purification of GST-Cleaved R416A NP

200 µL of BL21DE3 competent cells were incubated on ice with 0.5 µL pGEX NP R416A plasmid at a concentration of 204.2 ng/µL for 30 minutes. A further 0.5 µL plasmid was mixed into the competent cells, then the sample was heat shocked for 30 seconds at 42˚C. The cells were left to recover on ice for ten minutes, then plated onto LBHS plates with ampicillin (working concentration of 100mg/L). Plates were incubated overnight at 37˚C.

After overnight growth, one colony was selected from the plate, avoiding larger colonies with visible satellite colonies, and introduced to 100 mL LBHS + ampicillin. This starter culture was incubated overnight at 37˚C with shaking (180 rpm). 25 mL overnight culture and 1 mL ampicillin at 100mg/mL were added to each liter of LB and grown at 37˚C with shaking (180 rpm) to OD = 0.6-1.0. The cultures were then induced with 1 mM isopropyl β-D-1-thiogalactopyranoside and left overnight at 18˚C with shaking.

Cells were harvested via centrifugation for 20 minutes at 5000 x *g* and 4˚C in a Beckman Avanti centrifuge using the JLA 8.1 rotor. Cell pellets from two liters of culture were merged and resuspended into 30 mL of wash buffer with 150 mM NaCl and 30 mM Tris at pH 7.5, protease inhibitors were added (Roche cOmplete Mini EDTA-free Protease Inhibitor Cocktail), and the pellet was left to freeze at –80˚C.

Frozen pellet was defrosted and mixed by pipetting. Cells were then lysed by sonication with an amplitude of 45 J, with 5 seconds of active sonication followed by 25 seconds rest for a total active sonication time of 2 minutes 30 seconds. Lysed cell debris was removed by centrifugation for 45 minutes at 30,000 x g and 4˚C in a Beckman Avanti centrifuge using a JA 25.50 rotor. While centrifugation proceeded, glutathione sepharose 4B resin (Cytiva) was equilibrated for gravity flow. 0.5 mL bead

volume was equilibrated per liter of cell culture by washing with 12 CV MilliQ water then 12 CV wash buffer. The equilibrated beads were incubated with the cell lysate for 60-90 minutes, then the lysate-bead slurry was applied to the column. The beads were washed with 20 CV wash buffer, then the column was capped; 20 µL HRV 3C protease (ThermoScientific) and 200 µL wash buffer were added to the beads and incubated 18-24 hours at 4˚C on a tube roller to cleave the GST-tag. The target protein was then eluted with wash buffer until protein concentration in the eluate dropped below 0.05 mg/mL, as monitored by NanoDrop, approximately 24 CV. Eluate was concentrated by centrifugation to 900µM and stored at 4˚C before use.

Purification of GST-Tagged R416A NP

200 µL of BL21DE3 competent cells were incubated on ice with 0.5 µL pGEX NP R416A plasmid at a concentration of 204.2 ng/µL for 30 minutes. A further 0.5 µL plasmid was mixed into the competent cells, then the sample was heat shocked for 30 seconds at 42˚C. The cells were left to recover on ice for ten minutes, then plated onto LBHS plates with ampicillin (working concentration of 100mg/L). Plates were incubated overnight at 37˚C.

After overnight growth, one colony was selected from the plate, avoiding larger colonies with visible satellite colonies, and introduced to 100 mL LBHS + ampicillin. This starter culture was incubated overnight at 37˚C with shaking (180 rpm). 25 mL overnight culture and 1 mL ampicillin at 100mg/mL were added to each liter of LB and grown at 37˚C with shaking (180 rpm) to OD = 0.6-1.0. The cultures were then induced with 1 mM isopropyl β-D-1-thiogalactopyranoside and left overnight at 18˚C with shaking.

Cells were harvested via centrifugation for 20 minutes at 5000 x g and 4˚C in a Beckman Avanti centrifuge using the JLA 8.1 rotor. Cell pellets from two liters of culture were merged and resuspended into 30 mL of wash buffer with 150 mM NaCl and 30 mM Tris at pH 7.5, protease inhibitors were added (Roche cOmplete Mini EDTA-free Protease Inhibitor Cocktail), and the pellet was left to freeze at –80˚C.

Frozen pellet was defrosted, wash buffer added to a total volume of 35 mL, and reconstituted by pipetting. Cells were then lysed by sonication with an amplitude of 45 J, with 5 seconds of active sonication followed by 25 seconds rest for a total active sonication time of 2 minutes 30 seconds. Lysed cell debris was removed by centrifugation for 45 minutes at 30,000 x g and 4˚C in a Beckman Avanti centrifuge using a JA 25.50 rotor. While centrifugation proceeded, glutathione sepharose 4B resin (Cytiva) was equilibrated for gravity flow. 0.5 mL bead volume was equilibrated per liter of cell culture by washing with 12 CV MilliQ water then 12 CV wash buffer. The equilibrated beads were incubated with the cell lysate for 60-90 minutes, then the lysate-bead slurry was applied to the column. The beads were washed with 20 CV wash buffer, then the column was capped; 200 µL wash buffer and 40 µL (20 units) of HRV 3C Protease (ThermoFisher Scientific) were applied to the resin, which was allowed to incubate overnight on a tube roller at 4˚C.

After overnight incubation with HRV 3C, 12 CV wash buffer were added to the column. The bead and wash buffer slurry incubated for 60-90 minutes on a tube roller at 4˚C. A total of 20 CV elution buffer containing the target protein was then collected. Eluate was concentrated by centrifugation to 0.6 mg/mL and dialyzed into NP dialysis buffer overnight.

Endonuclease-NP Pull-down

500 µL of GST bead volume was equilibrated with 12 CV of MilliQ water then 12 CV of wash buffer with 150 mM NaCl and 30 mM Tris at pH 7.5. 500 µL of dialyzed GST-tagged R416A NP sample was added to the equilibrated beads (0.3 mg purified NP) and allowed to incubate for 2 hours on the column at 4˚C. The beads were washed with 10 CV wash buffer. WT endonuclease was then added to the bead volume in a molar ratio of 1:1.5 NP : endonuclease. The beads with bound NP and added endonuclease were allowed to incubate for two hours at 4˚C on a tube roller. Flow through was collected and the beads washed with a further 10 CV wash buffer to remove unbound endonuclease. 2 mL of NP elution buffer with 150 mM NaCl, 30 mM Tris, and 25 mM reduced glutathione at pH 7.5 were added to the bead volume and allowed to incubate on the column for 30 minutes at 4˚C. The eluate was then collected and samples analyzed by gel electrophoresis.

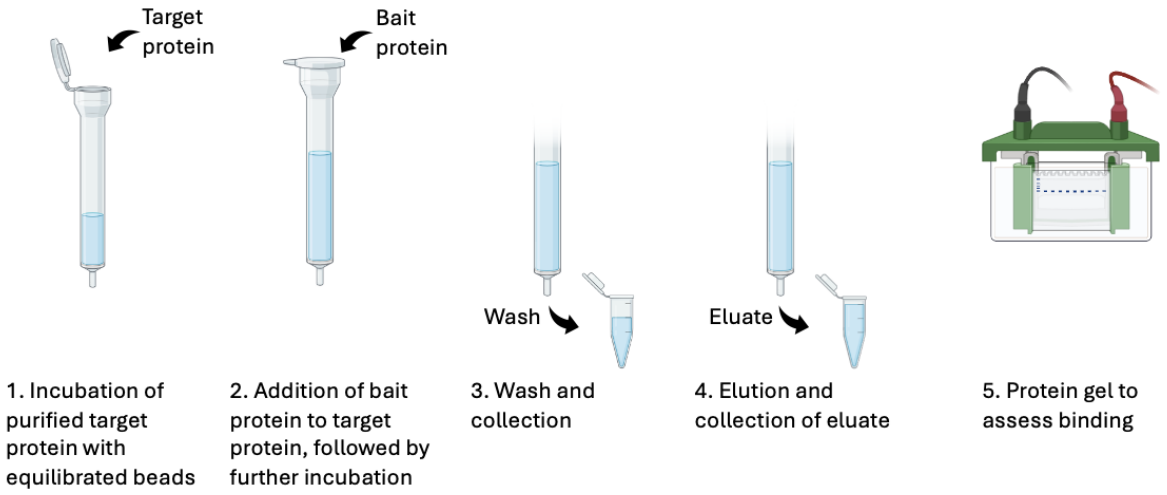


Figure 24. Visual representation of the protocol for pulldowns to investigate protein-protein interaction.

Endonuclease Titrations

Titrations were conducted using the Bruker 950 MHz (^{1}H) NMR spectrometer and samples were assembled individually to avoid dilution upon addition of the ligand. All samples were recorded at pH 7.5, which was checked individually before introduction to the magnet, and 25˚C. For WT endonuclease, all samples were run at a target (endonuclease) concentration of 170 μM into stock buffer containing 150 mM NaCl, 20 mM Tris, 2 mM BME, and 1 mM MgCl. Peptides of interest were purchased from GenScript and were resuspended in 80 μL of the same stock buffer to generate a stock solution of 3.8 mM (PB1 peptide) and 4.12 mM (PB2 peptide). The concentration of the peptide stock solutions was analyzed via quantitative NMR before use in titrations.

Samples were prepared individually with varying volumes of target and ligand solution in order to ensure the correct final sample composition. This process of individual sample construction with different volumes of endonuclease, PB1 peptide, PB2 peptide, and R416ANP solution was possible because all components were stored in the same buffer.

After the titration experiments were complete, peak movement was tracked using Computer Aided Resonance Assignment (CARA) and peak lists for each titration point saved before being exported for analysis into Excel. CSPs were calculated with:

$$\sqrt{\frac{1}{2}(h^2 + (0.2 * n^2))},$$

where *h* is the difference in ^{1}H values and *n* the difference in ^{15}N values (ppm). Unless otherwise indicated, mean CSP values for a spectrum include CSP values = 0 (no movement discernable or peak could not be tracked).

AlphaFold 3 Setup and Usage

AlphaFold 3 models used the AlphaFold 3 public server. WT and mutant endonuclease sequences were entered alone and with individual ligands to predict potential binding in the presence of two manganese ions. The highest confidence prediction was saved for visualization in PyMol and pLDDTs were referenced to understand confidence of the provided structures.

Cloning

Coning was conducted using Phusion High-Fidelity DNA Polymerase (New England BioLabs) for PCR and NEBuilder HiFi DNA Assembly master mix (New England BioLabs) for assembly of products. Samples were prepared for PCR on ice using the following volumes: 10 µL GC buffer (New England Biolabs), 2.5 µL primers (10 mM stock concentration), 1 µL dNTPs, 2 µL template (26.8 ng/ µL), 31 µL MilliQ water, and 1 µL Phusion enzyme, added last.

These samples were then transferred to a PCR machine and thermocycled using the conditions recommended by the manufacturer. The PCR products were stained and run on a 1% agarose gel for 90 minutes at 100 V. The relevant bands of product DNA were excised and gel-purified using the QIAquick Gel Extraction Kit materials and protocol (QIAGEN). The resultant DNA was added to an NEBuilder reaction at the volumes suggested by New England Biolabs' online tool NEBuilder Calculator. The full final volume of the NEBuilder reaction was used in transforming the provided KLD super competent cells before growing overnight on ampicillin plates.

One colony of transformed bacteria was selected and the plasmid purified according to the QIAGEN Spin Miniprep Kit protocol. The purified plasmid was sent for sequencing using Source Biosciences Sanger sequencing service to confirm identify before use and was stored at -20˚ C.

Activity Assays

Two versions of endonuclease activity assays were conducted, one using a magnesium buffer and the other a manganese buffer. In both cases an assay buffer containing 150 mM NaCl, 20 mM Tris, 5 mM BME, and 1 mM of the relevant metal (in the form of MgCl or MnCl) was assembled, filtered, and kept at 4 ˚ C before use. The samples detailed in Table 1 were assembled using 1 ug enzyme and 1.5 ug of M13mp18 ssDNA (TebuBio). The samples were then incubated for two hours at 37 ˚ C before being heat shocked for 20 minutes at 80 ˚ C to denature the enzyme. The samples were then run on 1% agarose gel for 90 minutes at 100 V and results analyzed using ImageJ software.

In ImageJ, the gel images were converted to grayscale to allow for assessment of mean gray value. A uniform selection area was established for each gel based upon the lowest activity band in the gel. This analysis area was then copied over into each lane to ensure uniform sampling within the gel. The mean gray value for the selection area was recorded and compared to assess changes in activity.

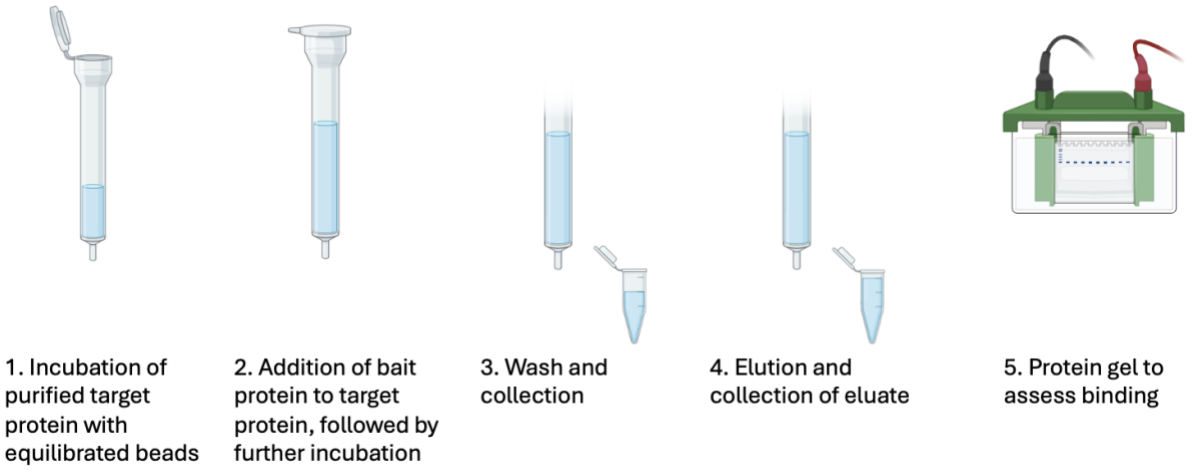


Figure 25. Visual representation of the protocol for activity assay to assess enzymatic activity..

Appendix

Nucleotide and Expressed Protein Sequences

His-Tagged WT Endonuclease Nucleotide Sequence

CATCATCATCATCATCACAGCAGCGGCGAGAATCTTTACTTCCAGGGCTCCGGCTC
TATGGAAGATTTCGTGCGTCAGTGCTTCAACCCCATGATCGTCGAGCTGGCTGAGA
AGGCTATGAAGGAATACGGCGAGGACCTGAAGATCGAGACTAACAAGTTCGCTGC
TATCTGCACCCACCTGGAAGTGTGCTTCATGTACTCCGACTTCCACTTCATCAACG
AGCAGGGCGAGTCCATCGTGGTGGAACTGGACGACCCCAACGCTCTGCTGAAGC
ACCGTTTCGAGATCATCGAGGGTCGTGACCGTACCATGGCTTGGACCGTGGTCAA
CTCCATCTGCAACACCACCGGTGCTGAGAAGCCCAAGTTCCTGCCCGACCTGTAC
GACTACAAGGAAAACCGTTTCATCGAGATCGGTGTCACCCGTCGCGAGGTGCACA
TCTACTACCTGGAAAAGGCCAACAAGATCAAGTCCGAGAACACCCACATCCACATC
TTCAGCTTCACCGGCGAGGAAATGGCTACCAAGGCTGACTACACCCTGGACGAGG
AATCCCGTGCTCGTATCAAGACCCGTCTGTTCACCATCCGTCAAGAGATGGCTAAC
CGTGGCCTGTGGGACTCCTTCCGTCAGTCCGAGCGTGGCGAATAA

His-Tagged WT Endonuclease Expressed Protein Sequence

HHHHHHSSGENLYFQGSGSMEDFVRQCFNPMIVELAEKAMKEYGEDLKIETNKFAAIC
THLEVCFMYSDFHFINEQGESIVVELDDPNALLKHRFEIIEGRDRTMAWTVVNSICNTT
GAEKPKFLPDLYDYKENRFIEIGVTRREVHIYYLEKANKIKSENTHIHIFSFTGEEMATKA
DYTLDEESRARIKTRLFTIRQEMANRGLWDSFRQSERGE*

GST-Tagged NP R416A Nucleotide Sequence

ATGTCCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGCAACCCACTCGACT
TCTTTTGGAATATCTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCGATGAAG
GTGATAAATGGCGAAACAAAAAGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCTT
ATTATATTGATGGTGATGTTAAATTAACACAGTCTATGGCCATCATACGTTATATAG
CTGACAAGCACAACATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAGATTTCAATG
CTTGAAGGAGCGGTTTTGGATATTAGATACGGTGTTTCGAGAATTGCATATAGTAA
AGACTTTGAAACTCTCAAAGTTGATTTTCTTAGCAAGCTACCTGAAATGCTGAAAAT
GTTCGAAGATCGTTTATGTCATAAAACATATTTAAATGGTGATCATGTAACCCATCC
TGACTTCATGTTGTATGACGCTCTTGATGTTGTTTTATACATGGACCCAATGTGCCT
GGATGCGTTCCCAAAATTAGTTTGTTTTAAAAAACGTATTGAAGCTATCCCACAAAT
TGATAAGTACTTGAAATCCAGCAAGTATATAGCATGGCCTTTGCAGGGCTGGCAAG
CCACGTTTGGTGGTGGCGACCATCCTCCAAAATCGGATCTGGAAGTTCTGTTCCAG
GGGCCCCTGGGATCCATGGCTTCCCAGGGTACTAAGCGCTCCTACGAGCAGATG
GAAACCGACGGCGAGCGTCAGAACGCTACCGAGATCCGTGCTTCCGTGGGCAAG
ATGATCGACGGTATCGGTCGCTTCTACATCCAGATGTGCACTGAGCTGAAGCTGTC

CGACTACGAGGGTCGTCTGATCCAGAACTCCCTGACCATCGAGCGTATGGTGCTG
TCCGCTTTCGACGAGCGTCGTAACAAGTACCTGGAAGAACACCCCTCCGCTGGAA
AGGACCCCAAGAAGACCGGCGGTCCCATCTACAAGCGTGTGGACGGAAAGTGGA
TGCGCGAGCTGGTGCTGTACGACAAGGGCGAGATCCGCCGTATCTGGCGTCAGG
CTAACAACGGCGACGACGCTACCGCTGGCCTGACCCACATGATGATCTGGCACTC
CAACTTGAACGACACCACCTACCAGCGTACCCGTGCTCTCGTGCGTACCGGCATG
GACCCCCGTATGTGCTCCCTGATGCAGGGTTCCACCCTGCCCAGACGTTCTGGTG
CTGCTGGCGCTGCTGTGAAGGGTGTCGGAACCATGGTCATGGAACTGATCCGTAT
GATCAAGCGCGGTATCAACGACCGTAACTTCTGGCGTGGCGAGAACGGTCGCAAG
ACCCGTTCCGCTTACGAGCGCATGTGCAACATCCTGAAGGGCAAGTTCCAGACCG
CTGCTCAGCGTGCTATGATGGACCAAGTGCGCGAGTCCCGTAACCCCGGCAACG
CTGAAATCGAGGACCTGATCTTCCTGGCTCGTTCCGCTCTGATCCTGCGCGGTTC
CGTGGCTCACAAGTCCTGCTTGCCCGCTTGCGTGTACGGTCCCGCTGTGGCTTCC
GGTTACGACTTCGAGAAGGAAGGCTACTCCCTCGTCGGTATCGACCCCTTCAAGC
TGCTGCAGAACTCTCAGGTGTACTCCCTGATCCGTCCCAACGAGAACCCCGCCCA
CAAGTCCCAGCTCGTCTGGATGGCTTGCAACTCTGCTGCTTTCGAGGACTTGCGT
GTGCTGTCCTTCATCCGTGGCACCAAGGTGTCCCCCCGTGGCAAGCTGTCTACCC
GTGGTGTCCAGATCGCTTCTAACGAGAACATGGACGCTATGGAATCCAGCACCCT
GGAACTGCGTTCCCGTTACTGGGCTATCCGTACCCGCTCCGGTGGCAACACCAAC
CAGCAAAGGGCTTCCGCTGGCCAGATCTCCGTGCAGCCCGCTTTCTCTGTGCAGG
CTAACCTGCCCTTCGACAAGCCCACCATCATGGCTGCTTTCACCGGCAACACCGA
GGGTCGCACCTCCGACATGCGTGCTGAGATCATCCGCATGATGGAAGGTGCTAAG
CCCGAGGAAATGTCCTTCCAGGGTCGTGGCGTGTTCGAGCTGTCTGACGAGAAGG
CTGCTAACCCTATCGTGCCCTCCTTCGACATGTCCAACGAGGGATCTTACTTCTTC
GGCGACAACGCTGAGGAATACGACAACTAA

GST-Tagged NP R416A Expressed Protein Sequence

MSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYI
DGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLK
VDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKL
VCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDLEVLFQGPLGSMAS
QGTKRSYEQMETDGERQNATEIRASVGKMIDGIGRFYIQMCTELKLSDYEGRLIQNSLT
IERMVLSAFDERRNKYLEEHPSAGKDPKKTGGPIYKRVDGKWMRELVLYDKGEIRRIW
RQANNGDDATAGLTHMMIWHSNLNDTTYQRTRALVRTGMDPRMCSLMQGSTLPRRS
GAAGAAVKGVGTMVMELIRMIKRGINDRNFWRGENGRKTRSAYERMCNILKGKFQTA
AQRAMMDQVRESRNPGNAEIEDLIFLARSALILRGSVAHKSCLPACVYGPAVASGYDF
EKEGYSLVGIDPFKLLQNSQVYSLIRPNENPAHKSQLVWMACNSAAFEDLRVLSFIRGT
KVSPRGKLSTRGVQIASNENMDAMESSTLELRSRYWAIRTRSGGNTNQQRASAGQIS
VQPAFSVQANLPFDKPTIMAAFTGNTEGRTSDMRAEIIRMMEGAKPEEMSFQGRGVF
ELSDEKAANPIVPSFDMSNEGSYFFGDNAEEYDN*

GST-Cleaved NP R416A Expressed Protein Sequence

ATGGCTTCCCAGGGTACTAAGCGCTCCTACGAGCAGATGGAAACCGACGGCGAG
CGTCAGAACGCTACCGAGATCCGTGCTTCCGTGGGCAAGATGATCGACGGTATCG
GTCGCTTCTACATCCAGATGTGCACTGAGCTGAAGCTGTCCGACTACGAGGGTCG
TCTGATCCAGAACTCCCTGACCATCGAGCGTATGGTGCTGTCCGCTTTCGACGAG
CGTCGTAACAAGTACCTGGAAGAACACCCCTCCGCTGGAAAGGACCCCAAGAAGA
CCGGCGGTCCCATCTACAAGCGTGTGGACGGAAAGTGGATGCGCGAGCTGGTGC
TGTACGACAAGGGCGAGATCCGCCGTATCTGGCGTCAGGCTAACAACGGCGACG
ACGCTACCGCTGGCCTGACCCACATGATGATCTGGCACTCCAACTTGAACGACAC
CACCTACCAGCGTACCCGTGCTCTCGTGCGTACCGGCATGGACCCCCGTATGTGC
TCCCTGATGCAGGGTTCCACCCTGCCCAGACGTTCTGGTGCTGCTGGCGCTGCTG
TGAAGGGTGTCGGAACCATGGTCATGGAACTGATCCGTATGATCAAGCGCGGTAT
CAACGACCGTAACTTCTGGCGTGGCGAGAACGGTCGCAAGACCCGTTCCGCTTAC
GAGCGCATGTGCAACATCCTGAAGGGCAAGTTCCAGACCGCTGCTCAGCGTGCTA
TGATGGACCAAGTGCGCGAGTCCCGTAACCCCGGCAACGCTGAAATCGAGGACCT
GATCTTCCTGGCTCGTTCCGCTCTGATCCTGCGCGGTTCCGTGGCTCACAAGTCC
TGCTTGCCCGCTTGCGTGTACGGTCCCGCTGTGGCTTCCGGTTACGACTTCGAGA
AGGAAGGCTACTCCCTCGTCGGTATCGACCCCTTCAAGCTGCTGCAGAACTCTCA
GGTGTACTCCCTGATCCGTCCCAACGAGAACCCCGCCCACAAGTCCCAGCTCGTC
TGGATGGCTTGCAACTCTGCTGCTTTCGAGGACTTGCGTGTGCTGTCCTTCATCCG
TGGCACCAAGGTGTCCCCCCGTGGCAAGCTGTCTACCCGTGGTGTCCAGATCGCT
TCTAACGAGAACATGGACGCTATGGAATCCAGCACCCTGGAACTGCGTTCCCGTT
ACTGGGCTATCCGTACCCGCTCCGGTGGCAACACCAACCAGCAAAGGGCTTCCGC
TGGCCAGATCTCCGTGCAGCCCGCTTTCTCTGTGCAGGCTAACCTGCCCTTCGAC
AAGCCCACCATCATGGCTGCTTTCACCGGCAACACCGAGGGTCGCACCTCCGACA
TGCGTGCTGAGATCATCCGCATGATGGAAGGTGCTAAGCCCGAGGAAATGTCCTT
CCAGGGTCGTGGCGTGTTCGAGCTGTCTGACGAGAAGGCTGCTAACCCTATCGTG
CCCTCCTTCGACATGTCCAACGAGGGATCTTACTTCTTCGGCGACAACGCTGAGG
AATACGACAACTAA

PB1 Peptide Sequence

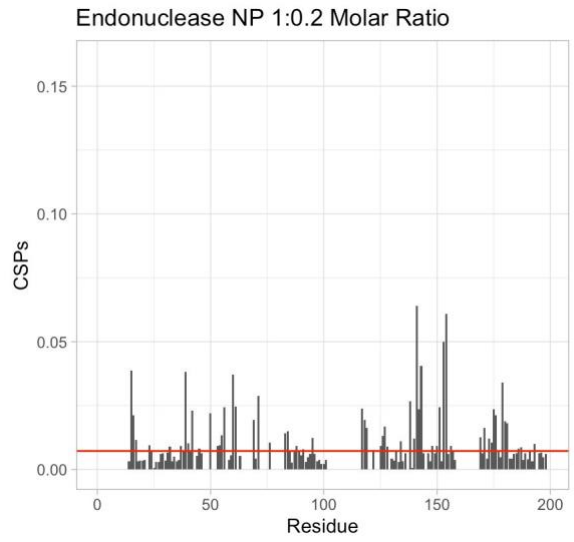
SHPGIFENSCLETMEVVQQT

PB2 Peptide Sequence

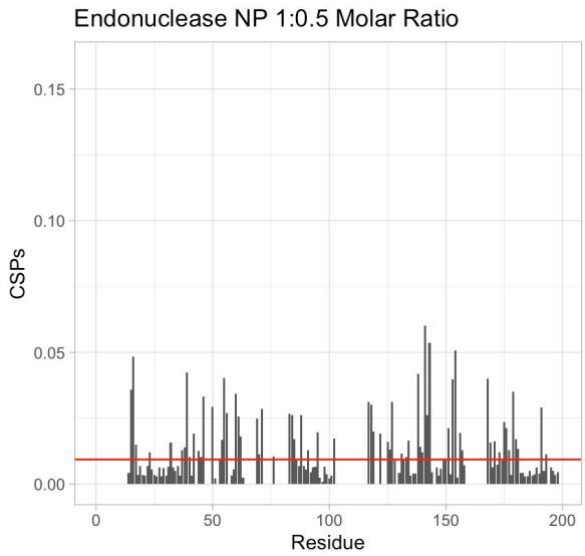
MKRKRDSSILTDSQTATKRIRMAIN

Titration Data

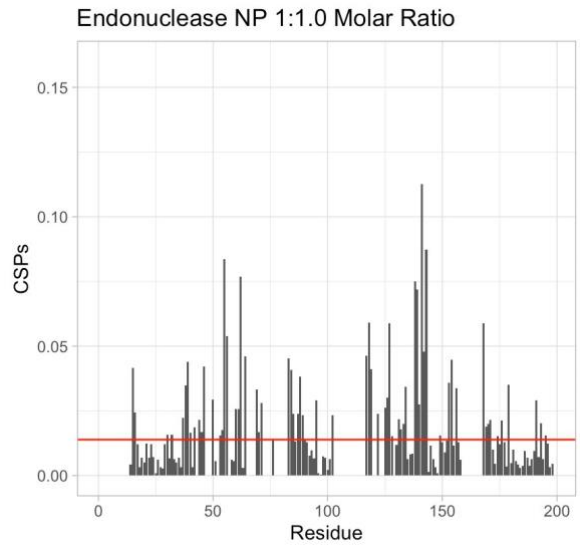
A



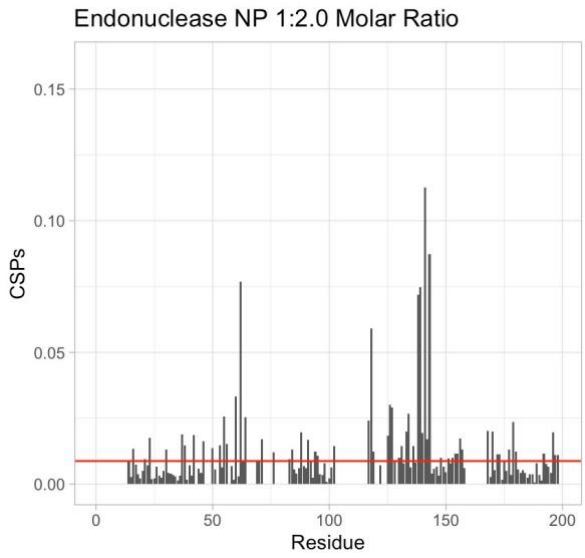
B



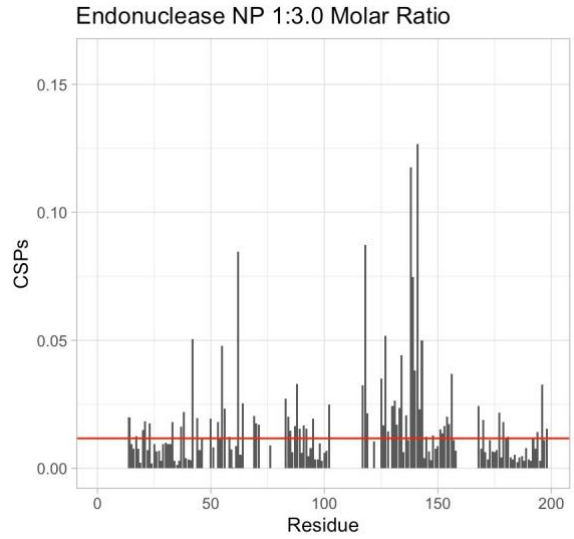
C



D

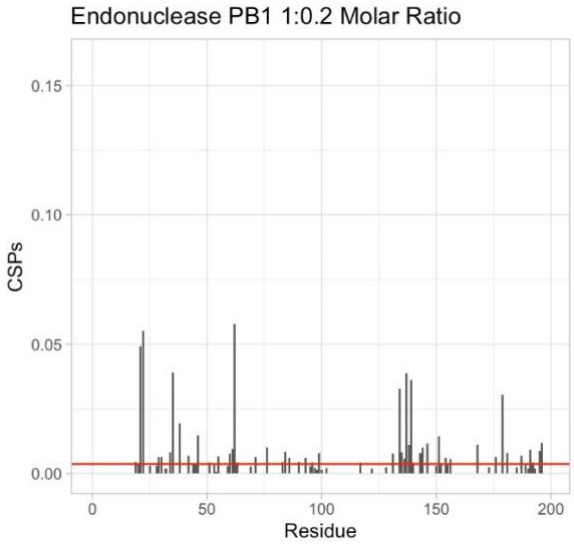


E

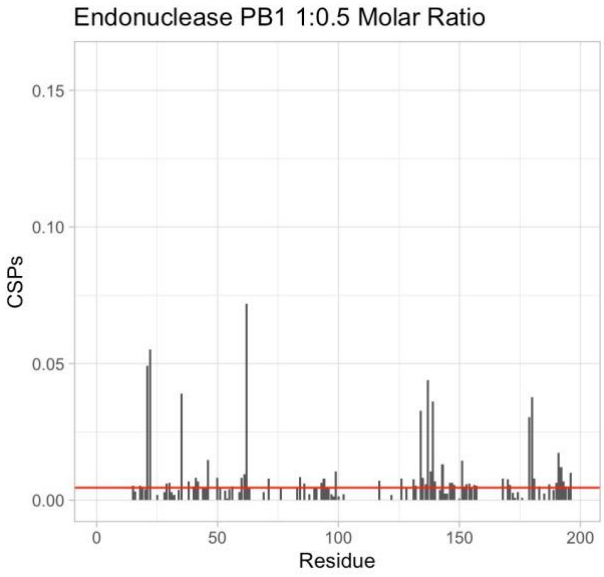


Appendix Figure 1: Titrations of NP R416A into wild type endonuclease. *A. Endonuclease and NP at a 1.0 : 0.2 molar ratio. B. Endonuclease : NP at a 1.0 : 0.5 molar ratio. C. Endonuclease : NP at a 1.0 : 1.0 molar ratio. D. Endonuclease : NP at a 1.0 : 2.0 molar ratio. E. Endonuclease : NP at a 1.0 : 3.0 molar ratio.*

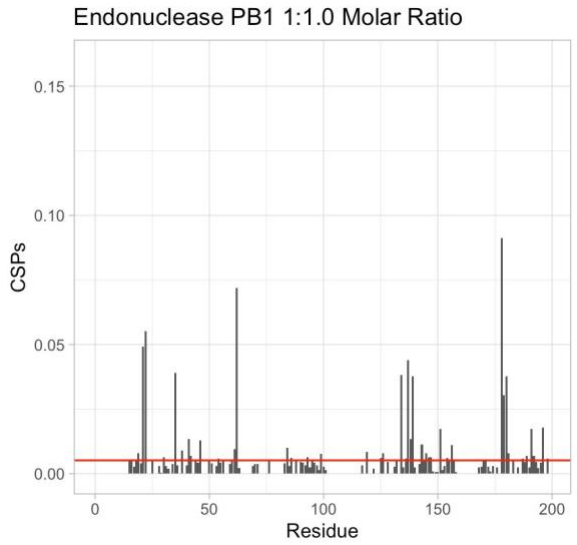
A



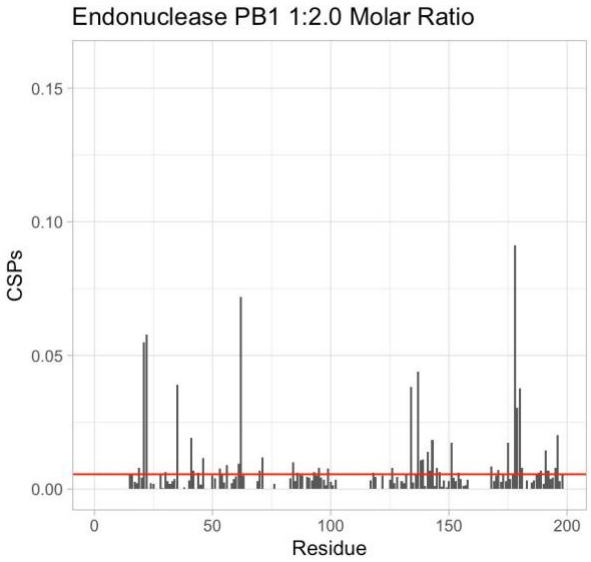
B



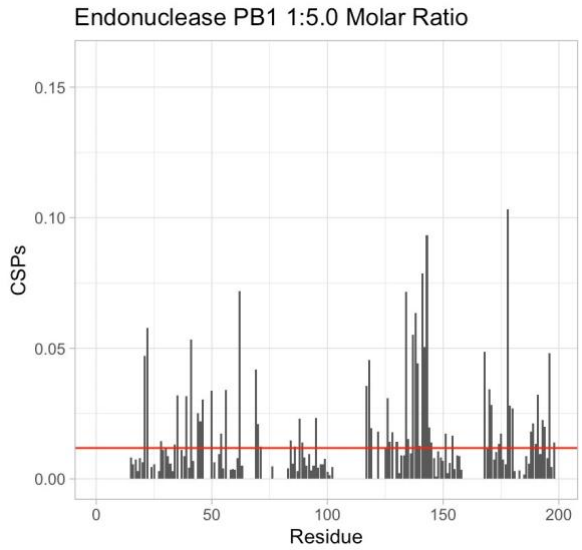
C



D

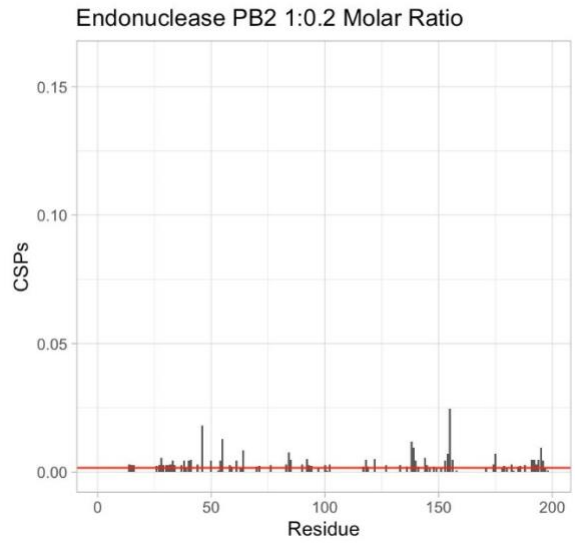


E

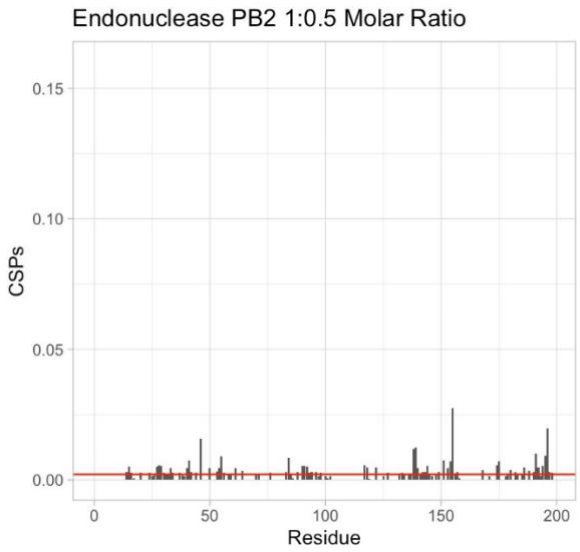


Appendix Figure 2: Titrations of PB1 peptide into wild type endonuclease. *A. Endonuclease and PB1 peptide at a 1.0 : 0.2 molar ratio. B. Endonuclease : PB1 peptide at a 1.0 : 0.5 molar ratio. C. Endonuclease : PB1 peptide at a 1.0 : 1.0 molar ratio. D. Endonuclease : PB1 peptide at a 1.0 : 2.0 molar ratio. E. Endonuclease : PB1 peptide at a 1.0 : 3.0 molar ratio.*

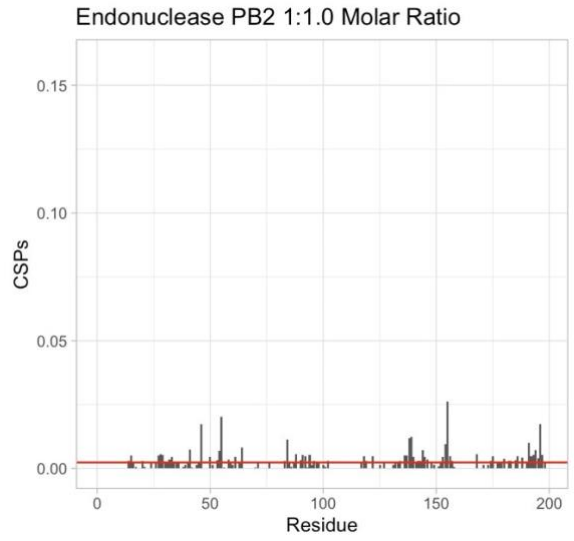
A


Endonuclease PB2 1:0.2 Molar Ratio
CSPs
0.15
0.10
0.05
0.00
0
50
100
150
200
Residue

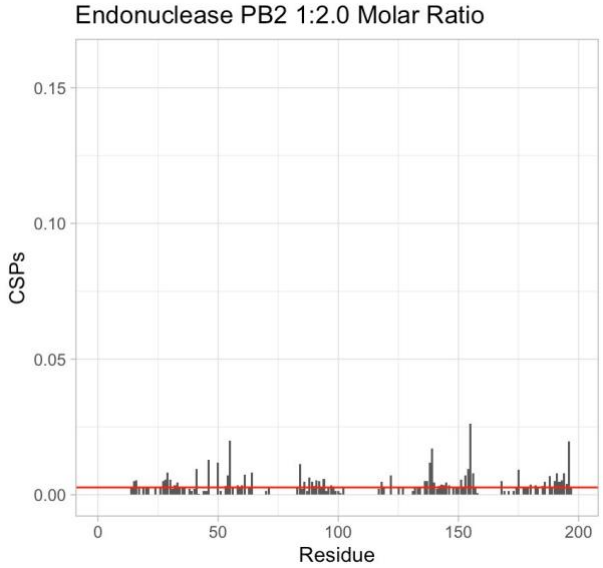
B


Endonuclease PB2 1:0.5 Molar Ratio
CSPs
0.15
0.10
0.05
0.00
0
50
100
150
200
Residue

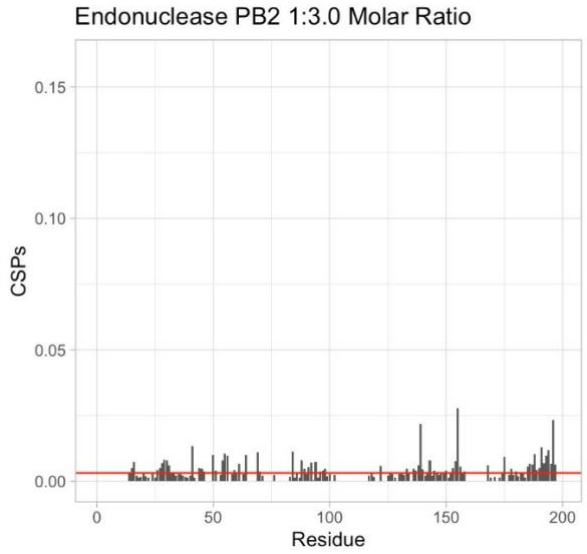
C


Endonuclease PB2 1:1.0 Molar Ratio
CSPs
0.15
0.10
0.05
0.00
0
50
100
150
200
Residue

D

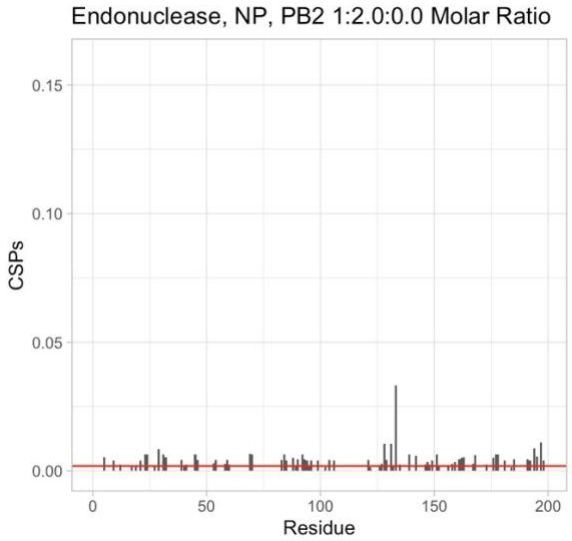

Endonuclease PB2 1:2.0 Molar Ratio
CSPs
0.15
0.10
0.05
0.00
0
50
100
150
200
Residue

E

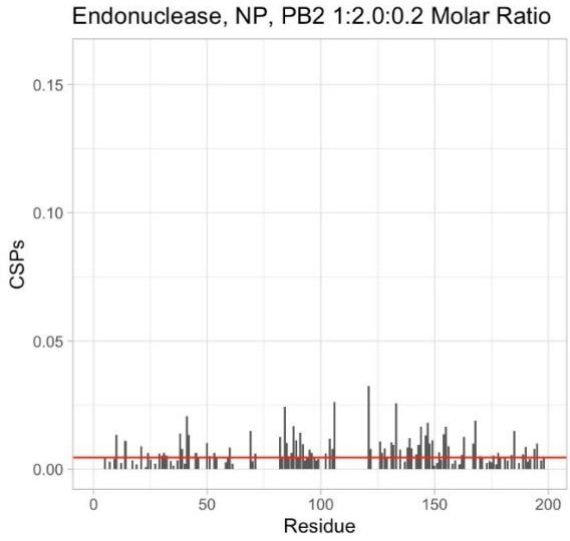


Appendix Figure 3: Titrations of PB2 peptide into wild type endonuclease. *A. Endonuclease and PB2 peptide at a 1.0 : 0.2 molar ratio. B. Endonuclease : PB2 peptide at a 1.0 : 0.5 molar ratio. C. Endonuclease : PB2 peptide at a 1.0 : 1.0 molar ratio. D. Endonuclease : PB2 peptide at a 1.0 : 2.0 molar ratio. E. Endonuclease : PB2 peptide at a 1.0 : 3.0 molar ratio.*

A

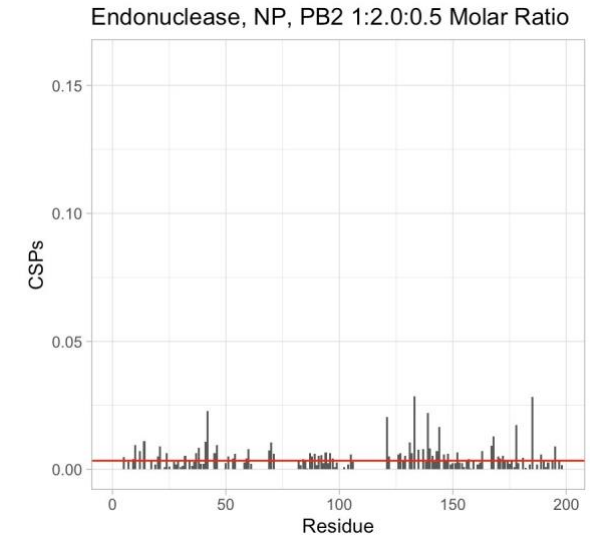


B

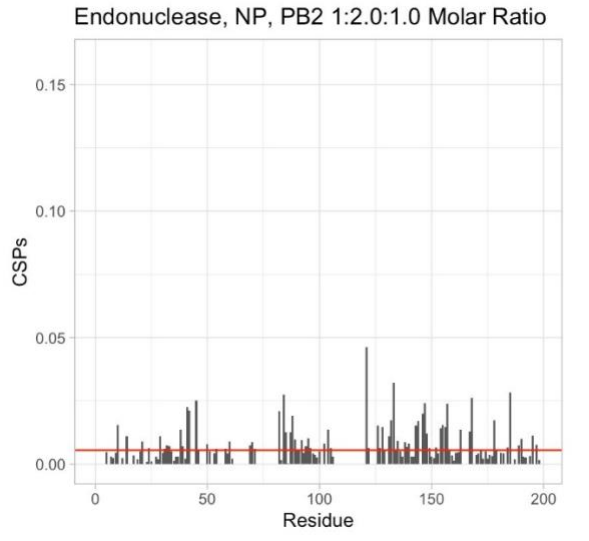


C

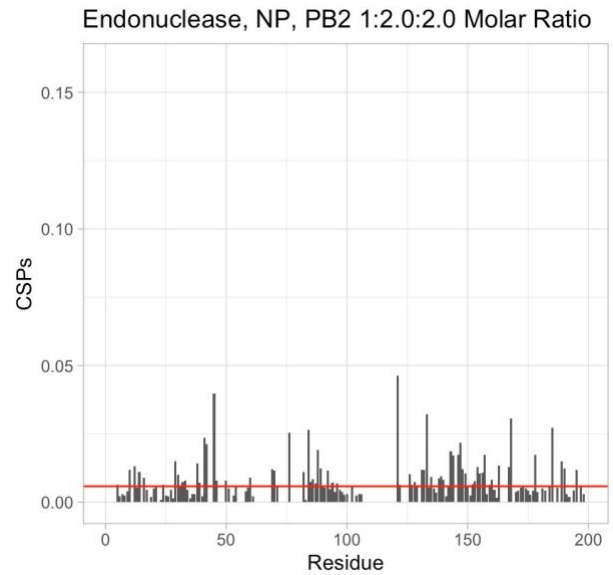
D

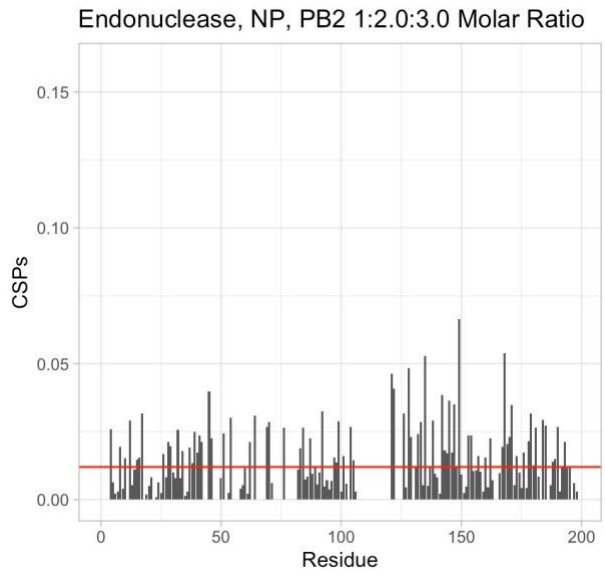


E



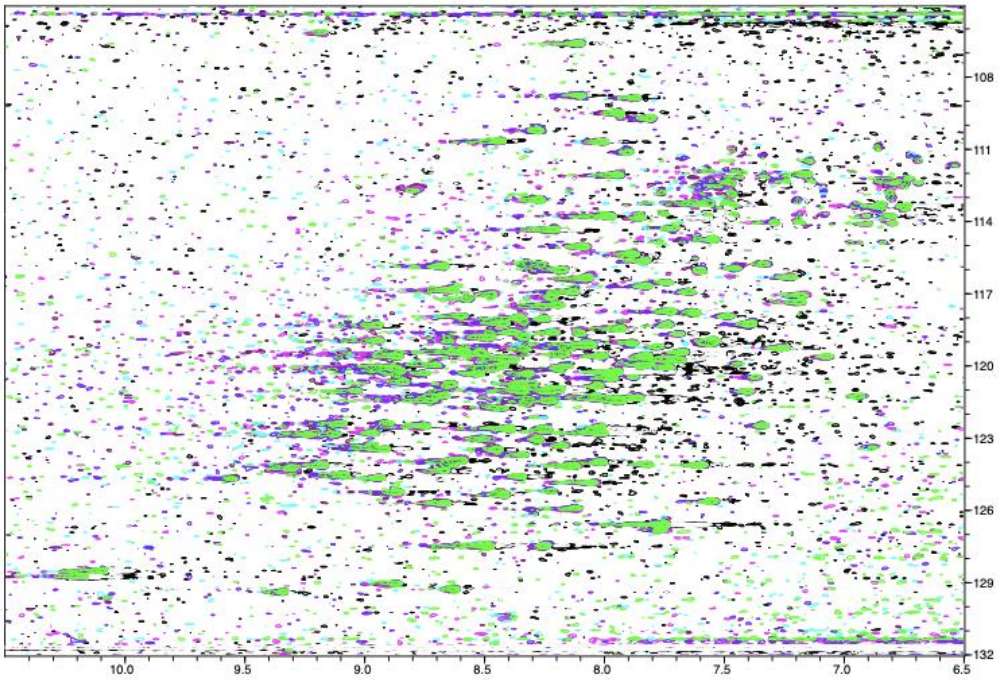
F



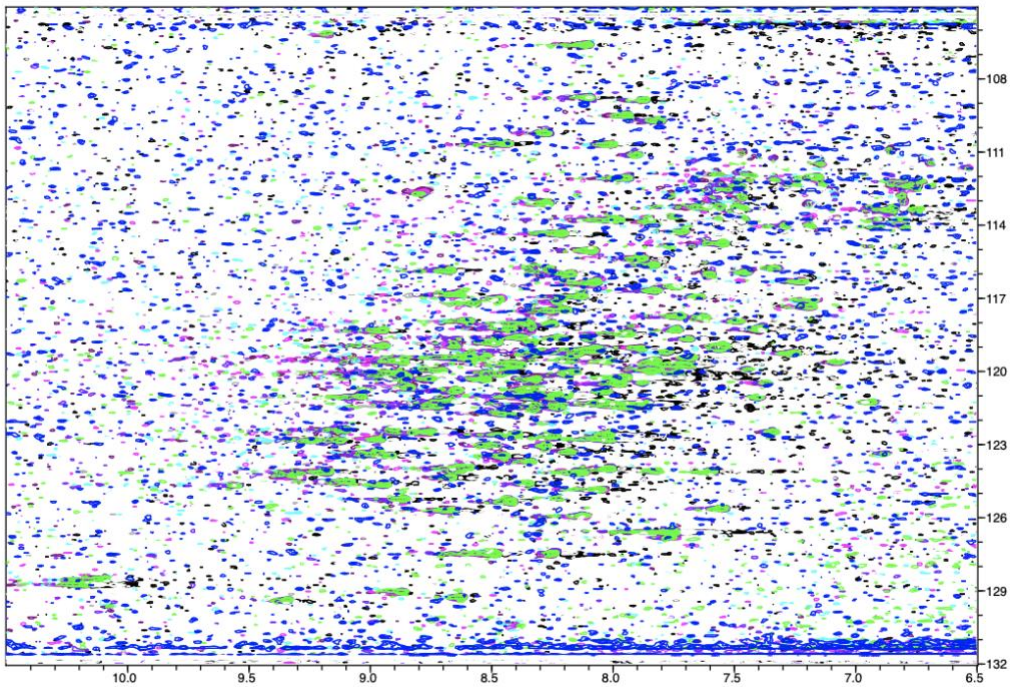


Appendix Figure 4: Titrations of PB2 peptide into wild type endonuclease with NP. *A. Endonuclease and NP at a 1.0 : 0.2 molar ratio. B. Endonuclease : NP : PB2 peptide at a 1.0 : 2.0 : 0.2 molar ratio. C. Endonuclease : NP : PB2 peptide at a 1.0 : 2.0 : 0.5 molar ratio. D Endonuclease : NP : PB2 peptide at a 1.0 : 2.0 : 1.0 molar ratio. E. Endonuclease : NP : PB2 peptide at a 1.0 : 2.0 : 2.0 molar ratio. F. Endonuclease : NP : PB2 peptide at a 1.0 : 2.0 : 3.0 molar ratio.*

A



B



Appendix Figure 5: BEST-TROSY spectra of WT endonuclease with a two-fold molar excess of NP, *x-axis (*^{1}H, ppm*), y-axis (*^{15}N, ppm*), and A. two-fold molar excess of PB2 peptide. B. three-fold molar excess of PB2 peptide. Contour for these spectra was left intentionally lower to demonstrate the low signal to noise of the resultant spectra, especially in the case of the three-fold molar excess of PB2 peptide.*

References

Abramson, J., Adler, J., Dunger, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024). https://doi.org/10.1038/s41586-024-07487-w

Berezin C., Glaser F., Rosenberg J., Paz I., Pupko T., Fariselli P., Casadio R. and Ben-Tal N.

ConSeq: The Identification of Functionally and Structurally Important Residues in Protein Sequences.

Bioinformatics. Vol. 20 1322-1324, 2004.

Biswas, S. K., Boutz, P. L., & Nayak, D. P. (1998). Influenza Virus Nucleoprotein Interacts with Influenza Virus Polymerase Proteins. *Journal of Virology*, *72*(7), 5493–5501. https://doi.org/10.1128/jvi.72.7.5493-5501.1998

Brinson R. G. (2021). Practical Applications of NMR to Solve Real-World Problems. *Molecules (Basel, Switzerland)*, *26*(23), 7091. https://doi.org/10.3390/molecules26237091

Carrique, L., Fan, H., Walker, A. P., Keown, J. R., Sharps, J., Staller, E., Barclay, W. S., Fodor, E., & Grimes, J. M. (2020). Host ANP32A mediates the assembly of the influenza virus replicase. *Nature*, *587*(7835), 638–643. https://doi.org/10.1038/s41586-020-2927-z

Dias, A., Bouvier, D., Crépin, T., McCarthy, A. A., Hart, D. J., Baudin, F., Cusack, S., & Ruigrok, R. W. H. (2009). The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature*, *458*(7240), 914–918. https://doi.org/10.1038/nature07745

Doan, L., Handa, B., Roberts, N. A., & Klumpp, K. (1999). Metal ion catalysis of RNA cleavage by the influenza virus endonuclease. *Biochemistry*, *38*(17), 5612–5619. https://doi.org/10.1021/bi9828932

DuBois, R. M., Slavish, P. J., Baughman, B. M., Yun, M., Bao, J., Webby, R. J., Webb, T. R., & White, S. W. (2012). Structural and biochemical basis for development of influenza virus inhibitors targeting the PA endonuclease. *PLoS Pathogens*, *8*(8), e1002830. https://doi.org/10.1371/journal.ppat.1002830

Engelhardt, O. G., Smith, M., & Fodor, E. (2005). Association of the Influenza A Virus RNA-Dependent RNA Polymerase with Cellular RNA Polymerase II. *Journal of Virology*, *79*(9), 5812–5818. https://doi.org/10.1128/jvi.79.9.5812-5818.2005

Fan, H., Walker, A. P., Carrique, L., Keown, J. R., Martin, I. S., Karia, D., Sharps, J., Hengrung, N., Pardon, E., Steyaert, J., Grimes, J. M., & Fodor, E. (2019). Structures of influenza A virus RNA polymerase offer insight into viral genome replication. *Nature*, *573*(7773), 287–290. https://doi.org/10.1038/s41586-019-1530-7

Fernández C, Wider G. TROSY in NMR studies of the structure and function of large biological macromolecules. *Curr Opin Struct Biol*. 2003;13(5):570-580. doi:10.1016/j.sbi.2003.09.009

Fodor, E., Crow, M., Mingay, L. J., Deng, T., Sharps, J., Fechter, P., & Brownlee, G. G. (2002). A single amino acid mutation in the PA subunit of the influenza virus RNA polymerase inhibits endonucleolytic cleavage of capped RNAs. *Journal of Virology*, *76*(18), 8989–9001. https://doi.org/10.1128/jvi.76.18.8989-9001.2002

Gao, Y., Guyatt, G., M Uyeki, T., Liu, M., Chen, Y., Zhao, Y., Shen, Y., Xu, J., Zheng, Q., Li, Z., Zhao, W., Luo, S., Chen, X., Tian, J., & Hao, Q. (2024). Antivirals for treatment of Severe influenza: A Systematic review and network meta-analysis of randomised controlled trials. *The Lancet*, *404*, 753–763.

Garjani, A., Chegini, A. M., Salehi, M., Tabibzadeh, A., Yousefi, P., Razizadeh, M. H., Esghaei, M., Esghaei, M., & Rohban, M. H. (2023). Forecasting influenza hemagglutinin mutations through the lens of anomaly detection. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-42089-y

Guilligay, D., Kadlec, J., Crépin, T., Lunardi, T., Bouvier, D., Kochs, G., Ruigrok, R. W. H., & Cusack, S. (2014). Comparative Structural and functional analysis of orthomyxovirus polymerase CaP-Snatching domains. *PLoS ONE*, *9*(1), e84973. https://doi.org/10.1371/journal.pone.0084973

Hatakeyama, D., Shoji, M., Ogata, S., Masuda, T., Nakano, M., Komatsu, T., Saitoh, A., Makiyama, K., Tsuneishi, H., Miyatake, A., Takahira, M., Nishikawa, E., Ohkubo, A., Noda, T., Kawaoka, Y., Ohtsuki, S., & Kuzuhara, T. (2021). Acetylation of the influenza A virus polymerase subunit PA in the N-terminal domain positively regulates its endonuclease activity. *FEBS Journal*, *289*(1), 231–245. https://doi.org/10.1111/febs.16123

He, X., Zhou, J., Bartlam, M., Zhang, R., Ma, J., Lou, Z., Li, X., Li, J., Joachimiak, A., Zeng, Z., Ge, R., Rao, Z., & Liu, Y. (2008). Crystal structure of the polymerase PAC–PB1N complex from an avian influenza H5N1 virus. *Nature*, *454*(7208), 1123–1126. https://doi.org/10.1038/nature07120

Hu, Y., Sneyd, H., Dekant, R., & Wang, J. (2017). Influenza A virus nucleoprotein: a highly conserved Multi-Functional viral protein as a hot antiviral drug target. *Current Topics in Medicinal Chemistry*, *17*(20). https://doi.org/10.2174/1568026617666170224122508

How NMR works: NMR 101: Spectroscopy. Bruker. (2025). https://www.bruker.com/en/resources/library/application-notes-mr/nmr-101.html

Idoko-Akoh, A., Goldhill, D. H., Sheppard, C. M., Bialy, D., Quantrill, J. L., Sukhova, K., Brown, J. C., Richardson, S., Campbell, C., Taylor, L., Sherman, A., Nazki, S., Long, J. S., Skinner, M. A., Shelton, H., Sang, H. M., Barclay, W. S., & McGrew, M. J. (2023). Creating resistance to avian influenza infection through genome editing of the ANP32 gene family. *Nature Communications*, *14*(1). https://doi.org/10.1038/s41467-023-41476-3

Ivanova, P. T., Myers, D. S., Milne, S. B., McClaren, J. L., Thomas, P. G., & Brown, H. A. (2015). Lipid composition of the viral envelope of three strains of influenza Virus—Not all viruses are created equal. *ACS Infectious Diseases*, *1*(9), 435–442. https://doi.org/10.1021/acsinfecdis.5b00040

Katoh, K., Rozewicki, J., Yamada, K. D. (2019). online service: multiple sequence alignment, interactive sequence choice and visualization, *Briefings in Bioinformatics*, Volume 20, Issue 4, Pages 1160–1166, https://doi.org/10.1093/bib/bbx108

Knight, M. L., Fan, H., Bauer, D. L. V., Grimes, J. M., Fodor, E., & Keown, J. R. (2021). Structure of an H3N2 influenza virus nucleoprotein. *Acta Crystallographica*

Section F Structural Biology Communications, *77*(7), 208–214. https://doi.org/10.1107/s2053230x2100635x

Kowalinski, E., Zubieta, C., Wolkerstorfer, A., Szolar, O. H. J., Ruigrok, R. W. H., & Cusack, S. (2012). Structural analysis of specific metal chelating inhibitor binding to the endonuclease domain of influenza PH1N1 (2009) polymerase. *PLoS Pathogens*, *8*(8), e1002831. https://doi.org/10.1371/journal.ppat.1002831

Kozak, S., Lercher, L., Karanth, M. N., Meijers, R., Carlomagno, T., & Boivin, S. (2016). Optimization of protein samples for NMR using thermal shift assays. *Journal of biomolecular NMR*, *64*(4), 281–289. https://doi.org/10.1007/s10858-016-0027-z

Krischuns, T., Arragain, B., Isel, C., Paisant, S., Budt, M., Wolff, T., Cusack, S., & Naffakh, N. (2024). The host RNA polymerase II C-terminal domain is the anchor for replication of the influenza virus genome. *Nature Communications*, *15*(1). https://doi.org/10.1038/s41467-024-45205-2

Kumar, G., Cuypers, M., Webby, R. R., Webb, T. R., & White, S. W. (2020). Structural insights into the substrate specificity of the endonuclease activity of the influenza virus cap-snatching mechanism. *Nucleic Acids Research*, *49*(3), 1609–1618. https://doi.org/10.1093/nar/gkaa1294

Kuraku, S., Zmasek, C. M., Nishimura, O., Katoh, K. (2013). aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity, *Nucleic Acids Research*, Volume 41, Issue W1. Pages W22–W28, https://doi.org/10.1093/nar/gkt389

Kuzuhara, T., Kise, D., Yoshida, H., Horita, T., Murazaki, Y., Nishimura, A., Echigo, N., Utsunomiya, H., & Tsuge, H. (2009). Structural basis of the influenza A virus

RNA polymerase PB2 RNA-binding domain containing the pathogenicity-determinant lysine 627 residue. *Journal of Biological Chemistry*, *284*(11), 6855–6860. https://doi.org/10.1074/jbc.c800224200

Liao, Y., Jin, Y., Zhang, H., Yang, J., Fu, J., & Lv, H. (2022). Immunogenicity of a trivalent influenza vaccine and persistence of induced immunity in adults aged ≥60 years in Taizhou City, Zhejiang Province, China, during the 2018–2019 season. *Human Vaccines & Immunotherapeutics*, *18*(5). https://doi.org/10.1080/21645515.2022.2071061

Luo, M. (2012). Influenza virus entry. In *Viral Molecular Machines, Advances in Experimental Medicine and Biology* (pp. 201–221). Springer Science+Business Media. https://doi.org/10.1007/978-1-4614-0980-9_9

M Barry, J. (2004). The site of origin of the 1918 influenza pandemic and its public health implications. *Journal of Translational Medicine*, *2*(3).

Moeller, A., Kirchdoerfer, R. N., Potter, C. S., Carragher, B., & Wilson, I. A. (2012). Organization of the influenza virus replication machinery. *Science*, *338*(6114), 1631–1634. https://doi.org/10.1126/science.1227270

Mondal, A., Potts, G. K., Dawson, A. R., Coon, J. J., & Mehle, A. (2015). Phosphorylation at the homotypic interface regulates nucleoprotein oligomerization and assembly of the influenza virus replication machinery. *PLoS Pathogens*, *11*(4), e1004826. https://doi.org/10.1371/journal.ppat.1004826

Nair, H., Brooks, W. A., Katz, M., Roca, A., Berkley, J. A., Madhi, S. A., Simmerman, J. M., Gordon, A., Sato, M., Howie, S., Krishnan, A., Ope, M., Lindblade, K. A., Carosone-Link, P., Lucero, M., Ochieng, W., Kamimoto, L., Dueger, E., Bhat, N.,

. . . Campbell, H. (2011). Global burden of respiratory infections due to seasonal influenza in young children: a systematic review and meta-analysis. *The Lancet*, *378*(9807), 1917–1930. https://doi.org/10.1016/s0140-6736(11)61051-9

Neverov, A. D., Kryazhimskiy, S., Plotkin, J. B., & Bazykin, G. A. (2015). Coordinated evolution of Influenza A surface proteins. *PLoS Genetics*, *11*(8), e1005404. https://doi.org/10.1371/journal.pgen.1005404

Ni, Z., Wang, J., Yu, X., Wang, Y., Wang, J., He, X., Li, C., Deng, G., Shi, J., Kong, H., Jiang, Y., Chen, P., Zeng, X., Tian, G., Chen, H., & Bu, Z. (2024). Influenza virus uses mGluR2 as an endocytic receptor to enter cells. *Nature Microbiology*, *9*(7), 1764–1777. https://doi.org/10.1038/s41564-024-01713-x

Omoto, S., Speranzini, V., Hashimoto, T., Noshi, T., Yamaguchi, H., Kawai, M., Kawaguchi, K., Uehara, T., Shishido, T., Naito, A., & Cusack, S. (2018). Characterization of influenza virus variants induced by treatment with the endonuclease inhibitor baloxavir marboxil. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-27890-4

Pala, N., Stevaert, A., Dallocchio, R., Dessi, A., Rogolino, D., Carcelli, M., Sanna, V., Sechi, M., & Naesens, L. (n.d.). Virtual Screening and Biological Validation of Novel Infuenza Virus PA Endonuclease Inhibitors. *ACS Medicinal Chemistry Letters*, *6*, 866–871. https://doi.org/10.1021/acsmedchemlett.5b00109

Puthenveetil, R., & Vinogradova, O. (2019). Solution NMR: A powerful tool for structural and functional studies of membrane proteins in reconstituted environments. *The Journal of biological chemistry*, *294*(44), 15914–15931. https://doi.org/10.1074/jbc.REV119.009178

Reich, S., Guilligay, D., Pflug, A., Malet, H., Berger, I., Crépin, T., Hart, D., Lunardi, T., Nanao, M., Ruigrok, R. W. H., & Cusack, S. (2014). Structural insight into cap-snatching and RNA synthesis by influenza polymerase. *Nature*, *516*(7531), 361–366. https://doi.org/10.1038/nature14009

Robb, N. C., Velthuis, A. J. W. T., Fodor, E., & Kapanidis, A. N. (2019). Real-time analysis of single influenza virus replication complexes reveals large promoter-dependent differences in initiation dynamics. *Nucleic Acids Research*, *47*(12), 6466–6477. https://doi.org/10.1093/nar/gkz313

Salzmann, M., Pervushin, K., Wider, G., Senn, H., & Wüthrich, K. (1998). TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(23), 13585–13590. https://doi.org/10.1073/pnas.95.23.13585

Sanjuan, R., R Nebot, M., Chirico, N., M Mansky, L., & Belshaw, R. (2010). Viral Mutation Rates. *Journal of Virology*, *84*(19), 9733–9748. https://doi.org/10.1128/JVI.00694-10

Shen, Y., & Bax, A. (2014). Protein Structural Information Derived from NMR Chemical Shift with the Neural Network Program TALOS-N. *Methods in Molecular Biology*, 17–32. https://doi.org/10.1007/978-1-4939-2239-0_2

Shen, Y., Chen, Y., Chu, S., Lin, M., Hsu, H., Wu, P., Wu, C., Liu, H., Lin, F., Lin, G., Hsu, P., Yang, A., Cheng, Y. E., Wu, Y., Wong, C., & Tsai, M. (2011). E339. . .R416 salt bridge of nucleoprotein as a feasible target for influenza virus

inhibitors. *Proceedings of the National Academy of Sciences*, *108*(40), 16515–16520. https://doi.org/10.1073/pnas.1113107108

Shi, L., Summers, D. F., Peng, Q., & Galarza, J. M. (1995). Influenza a virus RNA polymerase subunit PB2 is the endonuclease which cleaves host cell mRNA and functions only as the trimeric enzyme. *Virology*, *208*(1), 38–47. https://doi.org/10.1006/viro.1995.1127

Sieczkarski, S. B., & Whittaker, G. R. (2002). Influenza virus can enter and infect cells in the absence of Clathrin-Mediated endocytosis. *Journal of Virology*, *76*(20), 10455–10464. https://doi.org/10.1128/jvi.76.20.10455-10464.2002

Solyom, Z., Schwarten, M., Geist, L. *et al.* BEST-TROSY experiments for time-efficient sequential resonance assignment of large disordered proteins. *J Biomol NMR* **55**, 311–321 (2013). https://doi.org/10.1007/s10858-013-9715-0

Staller, E., Carrique, L., Swann, O. C., Fan, H., Keown, J. R., Sheppard, C. M., Barclay, W. S., Grimes, J. M., & Fodor, E. (2024). Structures of H5N1 influenza polymerase with ANP32B reveal mechanisms of genome replication and host adaptation. *Nature Communications*, *15*(1). https://doi.org/10.1038/s41467-024-48470-3

Staller, E., Sheppard, C. M., Neasham, P. J., Mistry, B., Peacock, T. P., Goldhill, D. H., Long, J. S., & Barclay, W. S. (2019). ANP32 proteins are essential for influenza virus replication in human cells. *Journal of Virology*, *93*(17). https://doi.org/10.1128/jvi.00217-19

Szeto, W. C., Hsia, H. P., Tang, Y. S., & Shaw, P. C. (2020). Interaction between influenza A virus nucleoprotein and PB2 cap-binding domain is mediated by RNA. *PloS one*, *15*(9), e0239899. https://doi.org/10.1371/journal.pone.0239899

Tang, Y., Xu, S., Chen, Y., Wang, J., & Shaw, P. (2021). Crystal structures of influenza nucleoprotein complexed with nucleic acid provide insights into the mechanism of RNA interaction. *Nucleic Acids Research*, *49*(7), 4144–4154. https://doi.org/10.1093/nar/gkab203

Taubenberger, J. K., & Morens, D. M. (2010). Influenza: the once and future pandemic. *Public Health Reports*, *125*(3_suppl), 15–26. https://doi.org/10.1177/00333549101250s305

Thierry, E., Guilligay, D., Kosinski, J., Bock, T., Gaudon, S., Round, A., Pflug, A., Hengrung, N., Omari, K. E., Baudin, F., Hart, D. J., Beck, M., & Cusack, S. (2015). Influenza polymerase can adopt an alternative configuration involving a radical repacking of PB2 domains. *Molecular Cell*, *61*(1), 125–137. https://doi.org/10.1016/j.molcel.2015.11.016

Thompson, W. W., Weintraub, E., Dhankhar, P., Cheng, P., Brammer, L., Meltzer, M. I., Bresee, J. S., & Shay, D. K. (2009). Estimates of US influenza-associated deaths made using four different methods. *Influenza and Other Respiratory Viruses*, *3*(1), 37–49. https://doi.org/10.1111/j.1750-2659.2009.00073.x

Todd, B., Tchesnokov, E. P., & Götte, M. (2021). The active form of the influenza cap-snatching endonuclease inhibitor baloxavir marboxil is a tight binding inhibitor. *Journal of Biological Chemistry*, *296*, 100486. https://doi.org/10.1016/j.jbc.2021.100486

Turrell, L., Hutchinson, E. C., Vreede, F. T., & Fodor, E. (2014). Regulation of influenza A virus nucleoprotein oligomerization by phosphorylation. *Journal of Virology*, *89*(2), 1452–1455. https://doi.org/10.1128/jvi.02332-14

Turrell, L., Lyall, J. W., Tiley, L. S., Fodor, E., & Vreede, F. T. (2013). The role and assembly mechanism of nucleoprotein in influenza A virus ribonucleoprotein complexes. *Nature Communications*, *4*(1). https://doi.org/10.1038/ncomms2589

Velthuis, A. J. W. T., Grimes, J. M., & Fodor, E. (n.d.). Structural insights into RNA polymerases of negative-sense RNA viruses. *Nature Reviews Microbiology*, *19*(5), 303–318. https://doi.org/10.1038/s41579-020-00501-8

Vidic, J., Noiray, M., Bagchi, A., & Slama-Schwok, A. (2016). Identification of a Novel Complex between the Nucleoprotein and PA (1–27) of Influenza A Virus Polymerase. *Biochemistry*, *55*(31), 4259–4262. https://doi.org/10.1021/acs.biochem.6b00514

Vreede, F. T., Ng, A. K., Shaw, P., & Fodor, E. (2011). Stabilization of influenza virus replication intermediates is dependent on the RNA-Binding but not the Homo-Oligomerization activity of the viral nucleoprotein. *Journal of Virology*, *85*(22), 12073–12078. https://doi.org/10.1128/jvi.00695-11

Wang, F., Sheppard, C. M., Mistry, B., Staller, E., Barclay, W. S., Grimes, J. M., Fodor, E., & Fan, H. (2022). The C-terminal LCAR of host ANP32 proteins interacts with the influenza A virus nucleoprotein to promote the replication of the viral RNA genome. *Nucleic Acids Research*, *50*(10), 5713–5725. https://doi.org/10.1093/nar/gkac410

Webster, R. G., Bean, W. J., Gorman, O. T., Chambers, T. M., & Kawaoka, Y. (1992). Evolution and ecology of influenza A viruses. *Microbiological Reviews*, *56*(1), 152–179. https://doi.org/10.1128/mmbr.56.1.152-179.1992

Xing, W., Barauskas, O., Aguayo, E., Willkom, M., Sapre, A., Clarke, M., Birkus, G., Schultz, B. E., Sakowicz, R., Kwon, H., & Feng, J. Y. (2017). Biochemical characterization of recombinant influenza A polymerase heterotrimer complex: Polymerase activity and mechanisms of action of nucleotide analogs. *PloS one*, *12*(10), e0185998. https://doi.org/10.1371/journal.pone.0185998

Ye, Q., Krug, R. M., & Tao, Y. J. (2006). The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. *Nature*, *444*(7122), 1078–1082. https://doi.org/10.1038/nature05379

Yu, M., Qu, Y., Zhang, H., & Wang, X. (2022). Roles of ANP32 proteins in cell biology and viral replication. *Animal Diseases*, *2*(1). https://doi.org/10.1186/s44149-022-00055-7

Zhang, H., Zhang, Z., Wang, Y., Wang, M., Wang, X., Zhang, X., Ji, S., Du, C., Chen, H., & Wang, X. (2019). Fundamental contribution and host range determination of ANP32A and ANP32B in influenza A virus polymerase activity. *Journal of Virology*, *93*(13). https://doi.org/10.1128/jvi.00174-19

Zhu, Z., Fodor, E., & Keown, J. R. (2022). A structural understanding of influenza virus genome replication. *Trends in Microbiology*, *31*(3), 308–319. https://doi.org/10.1016/j.tim.2022.09.015