Svoluji k zapůjčení své diplomové práce ke studijním účelům a prosím, aby byla vedena přesná evidence vypůjčovatelů. Převzaté údaje je vypůjčovatel povinen řádně ocitovat.

Charles University

Faculty of Science

Study programme: Chemistry

Branch of study: Biophysical Chemistry

Bc. Elena Karlukova

Development of high-throughput screening assay for the identification of inhibitors targeting influenza A polymerase

Vývoj testovací metody pro identifikaci inhibitorů chřipkové polymerasy

Diploma thesis

Supervisor: Jan Konvalinka Ph.D.

Prague, 2018

Prohlašuji, že jsem tuto diplomovou práci vypracovala samostatně pod vedením školitele doc. RNDr. Jana Konvalinky, CSc. a všechny použité prameny jsem řádně citovala.

V Praze 17.8.2018 ……

Elena Karlukova

Acknowledgments

I want to thank my supervisor Dr. Jan Konvalinka for all the guidance and support he provided during my bachelor and master studies. Being a part of his team was a great privilege for me and I will always refer to these years spent in his lab with warmth and gratitude.

I also want to thank my consultant Dr. Milan Kožíšek, for his help and dedication during the experimental work and as well as at times of writing this thesis. He was by my side from the first day when I entered the lab as a clueless bachelor student and he has patiently taught me everything.

I want to thank Iva Flaisigová, who has helped me a lot with the experiments, occasionally taking over purifications, so I could have lunch AND purified proteins.

I want to thank Michal Svoboda for always being around and ready to provide a piece of advice, story, or put good music on. And, for assistance with the part that I struggled the most during writing of this thesis – Czech version of the abstract.

But first of all, I want to thank my parents, who went great lengths to ensure I continue my education. All I have achieved or will ever achieve wouldn't have been possible if not for their continuous love and support.

Abstract

Influenza virus A circulates in birds and mammals and causes severe infectious disease that affects from 3 to 5 million people each year. There are two classes of anti-influenza drugs currently available: neuraminidase and M2 channel inhibitors. However, increasing resistance against these two types of inhibitors along with the potential emergence of new viral strains and unpredictability of pandemic outbreaks emphasize an unmet need for new types of inhibitors.

RNA-dependent influenza polymerase serves as a novel promising target for the development of anti-influenza medications. The aim of this master thesis is to develop in vitro high-throughput assays for screening of compounds targeting influenza RNA polymerase, particularly, its cap binding and endonuclease domains. For cap-binding domain the screening is based on DIANA (DNA-linked Inhibitor ANtibody Assay) method that was recently developed in our laboratory; for endonuclease domain, the method is based on AlphaScreen technology.

For the purposes of the methods development, recombinant cap binding domain of PB2 subunit and N-terminal endonuclease domain of PA subunit of influenza polymerase were expressed with appropriate fusion tags and purified using affinity and gel permeation chromatography. The probes for the screening assays were designed based on already known inhibitors; their binding properties were tested. Both methods were evaluated, and the first set of AlphaScreen-based screening was performed.

The methods developed would allow rapid quantitative *in vitro* screening for inhibitors of two activities within influenza RNA polymerase that are considered potential therapeutic targets, and, therefore, may bring advances to development of new anti-influenza medications.

Abstrakt

Virus chřipky způsobuje infekční onemocnění, postihující 3 až 5 milionů jedinců ročně. V současné době se k lečbě chřipkových infekcí používají dva druhy léčiv: inhibitory neuraminidasy a inhibitory iontového kanálu M2. Vzrůstající počet kmenů resistentních vůči těmto inhibitorům a riziko vzniku nových pandemických kmenů však vytvářejí dosud neuspokojenou potřebu nového typu inhibitorů.

RNA-dependentní RNA polymerasa viru chřipky je novým cílem pro vývoj inhibitorů směrovaných proti viru chřipky. Cílem této diplomové práce bylo vyvinout vysokokapacitní testovací metody pro nalezení potenciálních inhibitorů dvou aktivit chřipkové polymerázy – vazby 7-methylguanosinové čepičky hostitelských RNA a jejich enodukleasové štěpění. Pro doménu vázající 7-methylguanosinovou čepičku bylo testování založeno na metodě DIANA (DNA-linked Inhibitor ANtibody Assay), která byla nedávno vyvinuta v laboratoři školitele; testování inhibitorů endonukleasové domény pak bylo založeno na technologii AlphaScreen.

Pro potřeby vývoje velkokapacitních testovacích metod byla připravena rekombinantní doména vázající 7-methylguanosinovou čepičku z podjednotky PB2 a N-koncová doména PA podjednotky chřipkové polymerasy, nesoucí endonukleasovou aktivitu. Oba rekombinantní proteiny byly exprimovány s příslušnými afinitními značkami a vyčištěny metodami afinitní chromatografie a gelové filtrace. Sondy pro velkokapacitní testovací metody byly navrženy na základě struktur v literatuře popsaných inhibitorů a jejich vazebné charakteristiky byly ověřeny biofyzikálními metodami. Obě testovací metody byly optimalizovány a bylo provedeno pilotní testování metodou AlphaScreen.

Vyvinuté velkokapacitní metody dovolují rychlé, kvantitativní testování inhibitorů dvou nezbytných procesů zprostředkovaných virovou RNA polymerasou, které jsou považovány za potenciální cíle terapeutického zásahu, a mohou tak přispět k vývoji nových léčiv proti chřipkové infekci.

Table of contents

List of abbreviations

AA	Amino acid
BTTP	(3-[4-({bis[(1-terc-butyl-1H-1,2,3-triazol-4-yl)methyl]amino}methyl)-1H-1,2,3-triazol-1-yl]propanol
CBB	Coomassie Brilliant Blue
CBD	Cap binding domain
Da	Dalton
DANA	2-deoxy-2,3-dihydro-N-acetylneuraminic acid
DIANA	DNA-linked Inhibitor Antibody Assay
DMSO	Dimethylsulfoxid
EA	Ethanolamine
EDC	1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride
ER	Endoplasmatic reticulum
GSH	L-glutathione reduced
GST	Glutathione-S transferase
HA	Hemagglutinin
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LCMS	Liquid chromatography-mass spectrometry
m^7GTP	7-methylguanosine triphosphate
NA	Neuraminidase
NPA	N-terminal endonuclease domain of PA subunit
NEP	Nuclear export protein
NHS	N-hydroxysuccinimide
NLS	Nuclear localization signal

NP	Nucleocapsid protein
NS1	Non-structural protein 1
RNA	Ribonucleic acid
PA	Polymerase acidic protein
PB1	Polymerase basic protein 1
PB2	Polymerase basic protein 2
PCR	Polymerase chain reaction
PPI	Protein-protein interaction
RT	Room temperature
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate-polyacryleamide gel electrophoresis
SPR	Surface Plasmon Resonance
SUMO	Small ubiquitin-like modifier
TCEP	Tris(2-carboxyethyl)phosphine hydrochloride
vRNA	Viral RNA
vRNP	Viral ribonucleoprotein particle

1. Theoretical part

1.1 Influenza virus

1.1.1 Classification

Influenza virus A belongs to Orthomyxoviridae family of viruses. There are seven genera within this family: *Alphainfluenzavirus*, *Betainfluenzavirus*, *Deltainfluenzavirus*, *Gammainfluenzavirus*, *Isavirus*, *Quaranjavirus*, and *Thogotovirus*. Each influenza virus genera further include four viral species: Influenza A virus, Influenza B virus, Influenza D virus and Influenza C, respectively. These species differ substantially in morphological features, host ranges and, no less importantly, in pandemic potential. Among all influenza species, influenza virus A has the widest host range. It affects a variety of mammals (humans, swine, horses, dogs, seals), while aquatic birds serve as its natural reservoir.

Based on serological properties of its surface glycoproteins neuraminidase (NA) and hemagglutinin (HA), influenza A virus can be further classified into several subtypes. Nowadays, there are 18 HA and 11 known NA subtypes known and except H17, H18, and N10, N11 almost all their combinations have been observed in nature [1]. Influenza B virus circulates in humans and, reportedly, seals and is capable of causing mild seasonal epidemics [2]. It does not undergo subtype classification. Influenza virus C is known to infect humans and swine [3]. Despite being repeatedly detected in patients hospitalized with lower respiratory tract infections, the role of influenza C virus in seasonal epidemics is still not well established [4-6]. Influenza D virus, the newest addition to *Orthomyxoviridae* family, was first isolated from swine in Oklahoma in 2011 and found to be a distant relative of influenza C virus. Its occurrence in humans has not been detected [7].

All viral strains can be described using the standard nomenclature established by the International Committee on Taxonomy of Viruses [8]. According to it, the strain name should include type (A, B, C or D), host from which the virus was isolated (omitted if human), location of isolation, an isolate number and year of isolation, in case of A viruses also subtype. As an example, the first virus isolate of influenza A virus of the H1N1 subtype from a human in California in year 2009 is written as A/California/01/2009(H1N1) virus [1].

1.1.2 Structure of influenza virus

Influenza virus is an enveloped virus, with lipid membrane derived from a host organism (Figure 1). In influenza A virus, three membrane-bound proteins are found on its surface: hemagglutinin (HA), neuraminidase (NA) and M2 ion channel. Underneath the envelope, M1 protein constitutes the viral particle's core that encloses eight genetic segments [9]. Each genetic segment forms a ribonucleoprotein complex. This complex consists of genomic single-stranded RNA (vRNA) that forms a panhandle structure with both 5' and 3' ends bound to the trimeric viral polymerase and oligomeric nucleoprotein (NP), that is arranged along the length of vRNA. Influenza RNA-dependent-RNA polymerase includes three subunits: PA, PB1 and PB2, each of them carrying out a distinctive function [10]. Also, viral particle contains nuclear export protein (NEP or NS2), which is associated with M1 and mediates transport of the ribonucleoprotein complexes to the cytoplasm at later stages of infection [11, 12].

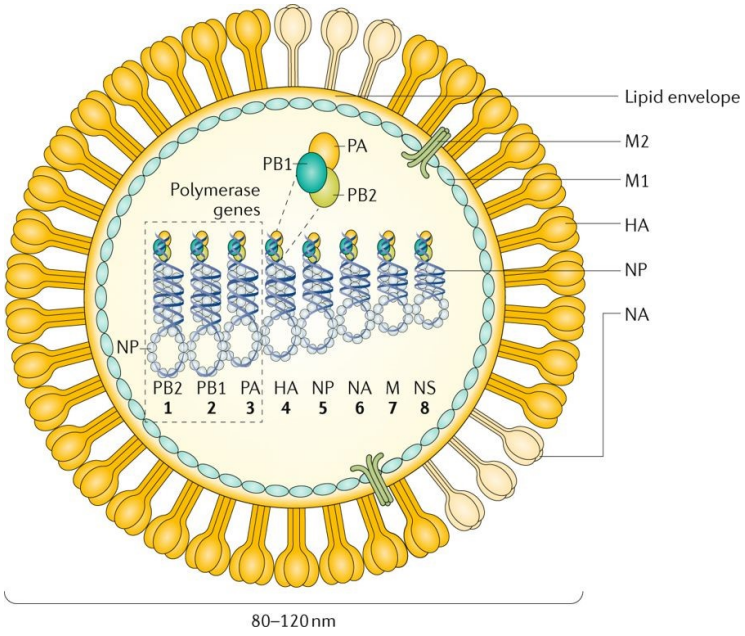


Figure 1. Structure of influenza A viral particle. Reproduced from [13].

1.1.2.1 Influenza virus genome

Influenza A virus genome is composed of eight segments of single-stranded RNA with negative polarity [14]. Virus overcomes its limited coding potential by splicing and use of alternative reading frames.

As the arrangement of genomic segments within a viral particle remains unknown, segments are numbered based on the size, from the largest to the smallest. Segments 1-3 encode polymerase subunits: PB2, PB, and PA, respectively. In some viral strains, segment 2 also encodes two additional proteins within alternative reading frames: PB1-F2 and PB1-N40 [15]. Ribosomal frameshift during translation of segment 3 mRNA gives rise to PA-X protein, presumably a virulence modulator [16]. Hemagglutinin (HA), nucleocapsid protein (NP), neuraminidase (NA) are encoded by segments 4, 5, and 6. There are two proteins encoded by segment 7: matrix protein (M1) and ion channel protein (M2), with M2 being translated from spliced mRNA variant. Segment 8 also gives rise to both unspliced and spliced mRNAs: unspliced variant is translated to nonstructural protein 1 (NS1) while the spliced mRNA encodes nonstructural protein 2 (NS2), also called nuclear export protein (NEP) [17, 18].

Coding sequences in all segments are framed with 5' and 3' non-coding regions. The extreme sequences of 12 – 13 nucleotides on both 5' and 3' ends are highly conserved in all segments, in both A and B viral types. These sequences facilitate the formation of panhandle RNA structure and are, therefore, partially complementary to each other [19, 20]. The rest of the non-coding region is segment-specific and was shown to provide a packaging signal during virion assembly [21].

1.1.2.2 Viral proteins

There are thirteen influenza viral proteins known: surface glycoproteins hemagglutinin and neuraminidase, matrix proteins M1 and M2, non-structural proteins NS1 and NEP, recently identified PB1-F2 protein with pro-apoptotic function and proteins that constitute the ribonucleoprotein complex: NP, PA, PB1 and PB2 [22]. PB1-F2, PB1-N140 and PA-X are not covered in this section. Structure of influenza polymerase and ribonucleoprotein complex will be discussed in section 1.2.

1.1.2.3 Hemagglutinin

Hemagglutinin is a trimeric rod-shaped surface glycoprotein anchored to the viral membrane. Each monomer is 540-550 amino acid (AA) long and synthesized in a precursor form HA0 that is further processed by a host protease. As a result, the mature HA subunit consists of two polypeptide chains HA1 and HA2 linked by a disulfide bond [23]. The cleavage predominantly occurs in the trans-Golgi network, but can also take place in the extracellular matrix after viral particle escaped a host cell. This proteolytic processing induces only minor conformational changes and structure of HA0 and mature HA are apparently identical. Yet minor structural rearrangements bring about a profound change in functionality: the cleavage releases N-terminal part of the HA2 polypeptide, forming so-called fusion peptide which, as the name implies, later facilitate the membrane fusion. Apart from proteolytic cleavage, hemagglutinin is known to undergo several post-translational modifications, as each monomer bears 7 glycosylation and 2 palmitoylation sites [24].

The first structure of pre-fusion hemagglutinin, solved by Wilson and Skehel in 1981, demonstrated the overall structural arrangements. The protein in its native homotrimeric form was found to possess two domains: globular receptor-binding domain formed by HA1 polypeptides and long coiled-coil stalk formed by HA2. Both N- and C- terminal ends were found to be located near the membrane, letting authors hypothesize that folding might have been taking place prior to cleavage of the N-terminal signal peptide. The globular domain was found to predominantly possess anti-parallel β-sheets, while the stalk was formed by coil-coiled α-helixes. Two other subdomains were identified: fusion peptide, located close to membrane region and vestigial esterase, formed by the H1 polypeptide and therefore located close to the globular domain.

The primary function of hemagglutinin in the context of the viral life cycle is the host receptor recognition. It binds N-acetylneuraminic (sialic) acid, a monosaccharide that is often found at the terminal end of receptors' glycoconjugates. There are two ways sialic acid could be attached to the penultimate galactose sugar: via the third or the sixths carbon of galactose, therefore resulting into α-2,3 or α-2,6- linkage. Hemagglutinin of human influenza virus was found to have specificity to α-2,6- linked sialic acid, predominant type of linkage in a human tracheal epithelium. In contrast, hemagglutinin derived from avian virus preferably binds sialic acid with α-2,3-linkage to galactose, abundant in avian gut epithelium [25]. However, it was shown that α-2,3-linkage is not

entirely exclusive in avian species and could be also found in the mammalian epithelium. Similarly, α-2,6-linkage, attributed to mammalian species, is present in small amounts in avian gut epithelium [26, 27]. As a result, a possibility of cross-species infection cannot be entirely ruled out and it indeed occasionally takes place. Also, hemagglutinin specificity can be changed with a single amino acid substitution, as it was previously documented [28, 29].

Hemagglutinin binding to receptor leads to clathrin-mediated endocytosis. After entering the host cell, it facilitates the fusion of viral and endosomal membranes. Low endosomal pH triggers large-scale conformational change, that results in fusion peptide becoming exposed and partially penetrating an endosomal membrane, bringing two membranes into proximity of each other and forming a pore through which vRNPs enter the cytoplasm [30, 31].

1.1.2.4 Neuraminidase

Neuraminidase is another surface glycoprotein of influenza virus. Functionally, it counterbalances hemagglutinin by acting as sialidase: it hydrolyses glycosidic linkage between terminal sialic acid and penultimate sugar, thereby destroying hemagglutinin binding sites. This interplay of hemagglutinin and neuraminidase functions determines replication effectivity and is often referred to as HA-NA functional balance [32].

Neuraminidase is a homotetramer, with each monomer being approximately 470 AA long [33]. Each subunit contains four domains: globular domain with the enzyme active site, stalk, the transmembrane domain and cytoplasmic tail [34]. The N-terminal cytoplasmic domain is a six amino acids long and conserved in nearly all influenza A strains, however, its function remains unknown [35]. The stalk varies in length (19-45 AA) and connects transmembrane and globular domains. It contains several cysteine residues, which results in the formation of neuraminidase tetramer as a dimer of two disulfide-linked dimers [36]. Also, the stalk region has one to three glycosylation sites, depending on length [37].

The structure of the globular domain solved in 1983 by Peter Colman, demonstrated its composition of six topologically identical β-sheets arranged in a propeller-like manner. The active is formed by a pocket on top of the domain's surface. Eight catalytic residues that constitute the active site (Arg118, Asp151, Arg152, Arg224, Glu276, Arg292, Arg371, Tyr406) are strictly conserved in both influenza A and influenza B [38].

All influenza A neuraminidases could be divided into two phylogenetic groups: subtypes N1, N4, N5, and N8 form the group 1, whereas group 2 consists of N2, N3, N6, N7 and N9 subtypes. There is a structural difference in active site conformation in these two groups: some subtypes were found to possess an extra cavity, formed because of an open conformation of '150-loop' (residues 147-152). All neuraminidases from phylogenetic group 1 have 150-cavity in the open conformation, whereas the neuraminidases of the group 2 lack 150-cavity due to the closed conformation [39]. The only exception is the 2009 pandemic H1N1 strain, that belongs to group 1 but lacks the 150-cavity [40].

1.1.2.5 M1 matrix protein and M2 proton channel

M1 matrix protein forms oligomeric structure that lies underneath the viral membrane and forms a viral envelope. It also associates with viral RNAs and inhibits viral transcription [41].

M2 proton channel is an integrated membrane protein which facilitates proton transport across the viral membrane [42]. It exists either in tetrameric form, each monomer is 97 AA long, with small ectodomain (23 AA), transmembrane domain (19 AA) and 54 AA long cytoplasmic tail [43]. Amphiphilic transmembrane α-helixes form a pore with histidine residues lying at the center. Currently, there are two models that explain the M2 channel function. The first model suggests that histidine tetrad functions as a pH-sensor, as histidine protonation results in electrostatic repulsion, which, in turn, triggers conformational change and the opening of the channel [44]. Alternatively, the shuttle model suggests that biprotonated histidine serves as intermediate step in a proton transport [45]. Activity of M2 channels leads to decrease in pH, which, in turns, triggers dissociation of viral envelope formed by M1 matrix proteins and induce conformation change in hemagglutinin which is required for fusion of viral and endosomal membranes [1].

1.1.2.6 NS1 and NEP proteins

Non-structural protein 1 (NS1) plays an important role in the mitigation of host anti-viral response [46]. NS1 forms dimers under native conditions [47]. The monomeric protein is 215-237 amino acids longs, with the length of C-terminal domain varying slightly between the subtypes. It consists of two domains: the N-terminal dsRNA binding

domain (RBD) and the C-terminal effector domain (ED) [48]. These domains are connected by 11 amino acids linker, interestingly, highly pathogenic strains were observed to have shorter, 6 amino acid long linker [49]. RBD domain is constituted of 6 α-helixes and its dimerization was found crucial for dsRNA binding. The binding is not sequence specific and is mediated by electrostatic interactions between a series of basic residues (Arg38, Arg35, and Arg46) and RNA phosphate backbone [50]. The effector domain adopts so-called α-helix β-crescent fold and was found to possess binding sites for several host factors [51].

NS1 was found to interfere with host anti-viral response by preventing interferon induction via the RIG-I pathway. NS1 directly interacts with RIG-I, therefore, inhibiting the RIG-I pathway [52, 53]. Also, one study reported an interaction between NS1 and TRIM25, a key player of RIG-I pathway. This interaction was suggested to prevent TRIM25 oligomerization, a prerequisite of its ligase activity, which is, in turn, essential for the RIG-I oligomerization and downstream signaling [54].

1.1.3 Life cycle of influenza virus

1.1.3.1 Attachment and entry

As was described earlier, viral attachment is mediated by HA binding to the sialic acid on host cell. Human viruses preferably bind sialic acid α-2,6-linkage to penultimate sugar, while avian viruses have specificity for α-2,3-linked sialic acid [55]. It is worth to mention, that some studies have shown that the type of linkage is not the only important parameter that influences the interaction and such factors as chain length and branching pattern may play a role as well [56].

Entry to the host cell proceeds through clathrin-mediated endocytosis. However, some studies show that alternatively, virus can enter a cell through non-clathrin mediated internalization or micropinocytosis [57, 58].

1.1.3.2 Fusion and uncoating

After entering a cell, the virus requires low pH to induce membrane fusion, as it triggers a conformational change in HA protein. The fusion peptide of HA, exposed during the conformational change, becomes inserted into an endosomal membrane, therefore bringing both membranes into a close proximity and forming a pore. Several fusion peptides form a pore through which vRNPs escape into a cytoplasm [55]. Acidic pH also

triggers disassembly of M1 envelope and disrupts the interaction between vRNPs and M1, enabling vRNAs to escape into the cytoplasm [59].

All components of vRNPs possess nuclear localization signals (NLS) and are actively transported to the cell nucleus via interactions with karyopherins α. Interestingly, nucleoprotein, which was shown to be of critical importance for vRNPs import, possesses unconventional NLS, presumably, to avoid competition with host proteins containing conventional NLS [60].

1.1.3.3 Genome replication and transcription

Unlike other -RNA viruses, replication and transcription of influenza virus is taking place in a host cell nucleus. Both processes are described in a detail in section X. Briefly, to start a transcription, viral polymerase binds 7-methylguanosine caps of host mRNAs, cleaves 10-15 downstream the cap and utilize this oligonucleotide as a primer. Replication is primer-independent but requires the activity of a priming loop. First, influenza polymerase generates complementary cRNA that serve as templates for vRNA.

1.1.3.4 Virus assembly and release

Newly synthesized vRNA forms ribonucleoprotein complex consisting of heterotrimeric polymerase and NP is subsequently exported to the cytoplasm via Chromosomal Maintenance-1-dependent mechanism (CRM1, also known as Exportin 1) [61]. CRM1 recognizes nuclear export factors (NES) present on a cargo molecule and facilitates the export by utilizing the Ran-GTP gradient.

According to the currently accepted model, M1 and NEP mediate the interaction of vRNPs with CRM1 during the export. They form so-called daisy chain arrangement, in which CRM1 binds to nuclear export signal of NEP, which, in turn, interacts with PB1 and PB2 subunits of influenza polymerase, promoting interaction of NP and M1 and binding to the N-terminus of M1 [62].

Once in the cytoplasm, vRNPs are thought to reach the apical membrane by exploiting recycling endosomes trafficking machinery. Several lines of evidence indicated Rab11 as an essential mediator of vRNPs trafficking [63-65]. Transmembrane proteins (NA, HA, and M2) contain apical sorting signals and are transported to the membrane independently from vRNPs [66-68]. Assembly and budding of the virus mostly occur on the lipid-raft microdomains [69].

Two models have been proposed to explain the packaging mechanism: a random incorporation model and a selective incorporation model. The random incorporation model suggests that virions may possess more or less than eight genomic segments, but only those with a full set of genomic segments succeed to propagate in a host cell [70].

According to the selective incorporation model, each vRNA segment might contain a specific packaging signal to ensure that each viral particle contains eight different segments. Both interfering particles studies and reverse genetics experiments showed that segment-specific packaging signals are located in the 5' and 3' non-coding regions of vRNA as well as in the coding regions [71-73]. Nonetheless, the precise mechanism of selective packaging still needs investigation.

Hemagglutinins of the newly synthesized virions are capable of binding sialic acid on the host cell surface, keeping viral particles anchored to the host membrane. Virions are released when another viral membrane protein, neuraminidase, removes sialic acid from the host surface glycoproteins. Also, neuraminidase cleaves sialic acids from the virus envelope, preventing aggregation of viral particles. The viral life cycle of influenza virus is shown in Figure 2.

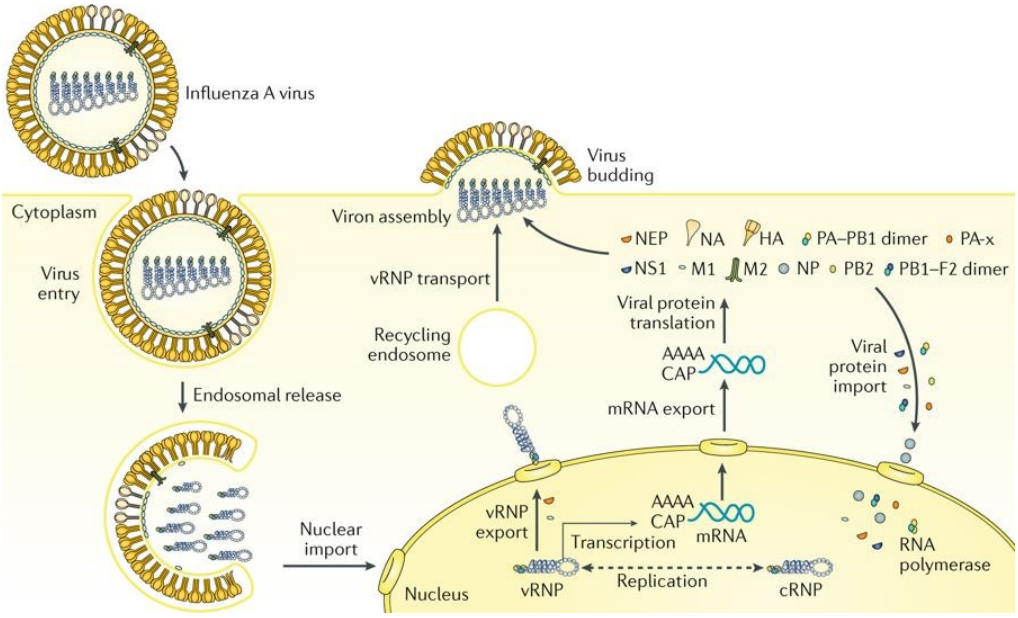


Figure 2. Influenza virus life cycle. Detailed are provided in the text. Reproduced from [13].

1.1.4 Influenza pandemics

Pandemics are outbreaks of infectious disease that spread rapidly across a large geographical area. History of influenza pandemics dates back to 421 B.C when an outbreak of respiratory disease with influenza-like symptoms was described by Hippocrates. Influenza pandemics arise in 10-40 years interval, caused, usually by a novel viral subtype to which population has no pre-existent immunity.

The most devastating influenza pandemic occurred in 1918 and resulted in more than 50 million deaths worldwide. Most likely the virus emerged in a military base in Kansas in the spring of 1918 and spread around the United States and with troopships to Europe [74]. The outbreak was covered extensively by news media in Spain and the virus thus got a nickname Spanish influenza. Sequencing of viral RNA extracted from formalin-fixed tissues form a patient died in 1918 showed that pandemic was caused by an H1N1 virus of avian origin that circulated in mammals before causing the outbreak [75].

Unlike influenza viruses that mostly cause fatal outcomes in young children and elderly people, Spanish influenza virus showed unusually high mortality rates in young adults. The typical symptoms were high fever, cough, pharyngitis, pulmonary hemorrhage. Most patients died from secondary bacterial infections, that were hardly treatable in the pre-antibiotic era [76].

The last influenza pandemic, known as the swine flu pandemic, occurred in 2009. The outbreak emerged in a Mexican village in February of 2009 (A/Mexico/InDRE4487/2009), spread rapidly to the United States and within a year the new pandemic viral strain was detected worldwide [77].

The new pandemic strain emerged as a result of triple gene reassortment, a process of genomic segments rearrangement that occurs in a cell infected with several different viral subtypes. The novel virus inherited PA and PB2 segments from North American avian virus; a PB1 segment originated from the human H3N2 virus; HA, NP, and NS genes were of swine virus origin; NA and M1 segments were inherited form Eurasian avian virus [78].

The H1N1 influenza pandemic resulted in 18 000 deaths worldwide. It also displayed unusual high mortality rate in the adult age group as patients aged 5-59 more frequently suffered from secondary respiratory infections and pneumonia [79].

1.1.5 Current medical treatments

Two classes of anti-influenza medications are used in treatment: M2 ion channel and neuraminidase inhibitors. M2 channel inhibitors, the first medications that specifically target the influenza virus, were introduced in the early 90s. These derivatives of adamantane physically block the M2 proton channel, preventing the change in pH needed for viral uncoating [80]. The resistance to this type of inhibitors, however, developed rapidly and both representatives of M2 channels inhibitors – rimantadine and amantadine, are no longer recommended for treatment [81].

Neuraminidase inhibitors are another class of antivirals used in the treatment of influenza viral infections. Inhibition of neuraminidase results in aggregation of newly synthesized virions on the cell surface, leading to reduced infectivity of viral particles. Currently, four neuraminidase inhibitors are approved and used in treatment: zanamivir (Relenza™), oseltamivir (Tamiflu™), peramivir (Rapivab™), laninamivir (Inavir™). Among those, oseltamivir has been used the most extensively in treatment and prophylaxis.

Like adamantans, neuraminidase inhibitors are also not invincible to mutation development. Oseltamivir-resistant variants are being continuously detected among seasonally circulating strains [82]. This raises serious concerns and underlines an importance of exploiting new drug targets within the virus to develop novel anti-influenza medications.

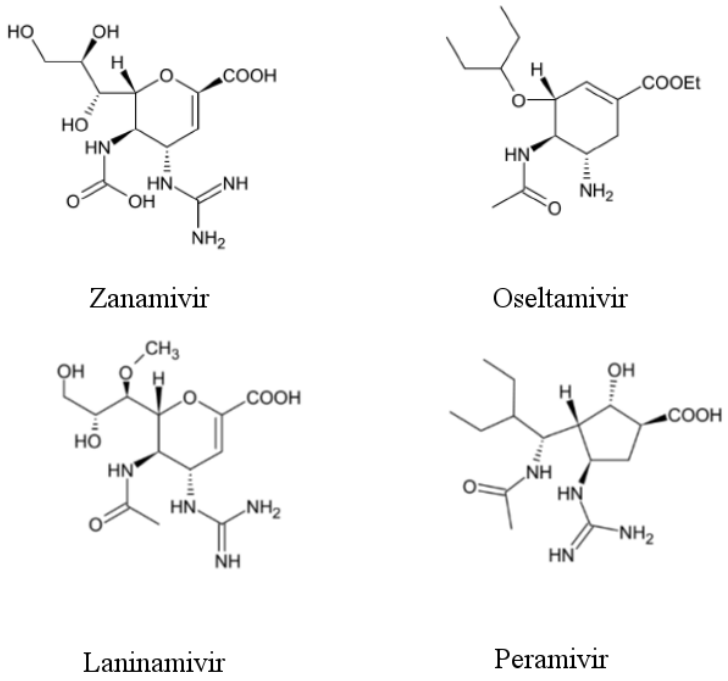


Figure 3. Structures of neuraminidase inhibitors used in clinics. Oseltamivir and zanamivir are approved world-wide. Peramivir is approved in Japan and US, laninamivir is approved in Japan only. Created by using ChemDraw software.

1.2 Influenza virus polymerase

1.2.1 Structure of ribonucleoprotein complex

Influenza polymerase consists of three subunits, PA (polymerase acidic protein), PB1 (polymerase basic protein 1), PB2 (polymerase basic protein 2), each possessing distinct functions. This trimeric complex along with genomic viral RNA and nucleoprotein comprise a structure of higher order, called ribonucleoprotein complex. Each segment of viral genome forms such a complex, with nucleoprotein bound to the vRNA (one molecule of nucleoprotein per 24 nucleotides) and 3' and 5' ends of vRNA bound to trimeric polymerase. The viral RNA coated with nucleoprotein forms a double-helix structure, as was shown by cryo-electron microscopy of native vRNAs [83]. Nucleoprotein binds to vRNA in a non-specific manner via the phosphate backbone of vRNA, not interfering with replication and transcription processes [84].

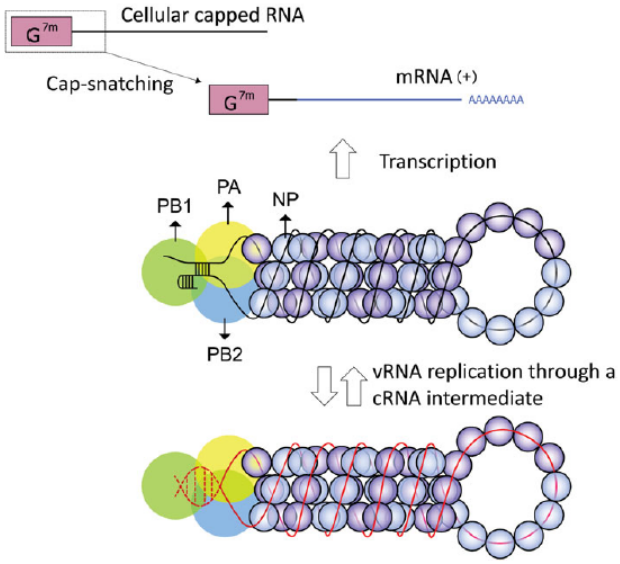


Figure 4. Schematic picture of ribonucleoprotein complex and its functions. Reproduced from [85].

1.2.2 Structure of influenza A polymerase

The arrangements of subunits within influenza polymerase remained unknown till 2014 when the first crystallographic structure of polymerase bound to viral RNA promotor was reported [86]. The heterotrimer has a molecular weight of approximately 250 kDa and so-called U-shape.

The PB1 subunit constitutes the core of the complex and makes contacts with PA-endonuclease and PB2-cap binding domains on the top of the complex (these two domains face each other and are separated by a solvent channel) and PA C-terminal domain on the bottom. The PB2-N-terminal domain and PA-linker that connects PA-C and PA-endonuclease lie along the side of PB1 core. PB1 subunit contains canonical right-handed fold with fingertips, fingers, thumb and palm domains, a characteristic feature of all RNA-depended-RNA-polymerases.

There are several structural similarities of PB1 to other viral polymerases, with Hepatitis C virus polymerase bearing the most significant resemblance. Also, PB1, as well as Flaviviridae polymerases, contains a 'priming loop' which role in genome replication will be discussed later.

Some structural attributes of PB1 subunit are rather unique. Among them are extended N- and C- termini that are making contacts with PB2 and PA subunits and long flexible β-ribbon (residues 177–212). This structure also contains two stretches of nuclear localization signals (NLS) that are important for trafficking of PA-PB1 heterodimer [87].

Catalytic center of PB1 contains conserved A-E and pre-A (also known as motif F) motives that are typically found in RNA-dependent-RNA-polymerase and are not observed in DNA-dependent-DNA-polymerases or reverse transcriptases [88].
Pre-A/F motive is a loop located in the fingertips domain that makes contacts with PA α-helix 20. It contains several basic residues and structurally resembles the pre-A/F loop found in Norwalk virus polymerase, which is thought to be involved in a template or/and NTP binding [89]. The catalytic site contains an aspartate triad (Aps305 in motif A and Asp445 and Asp446 in motif C) which coordinates two divalent metals ions. These residues are highly conserved in all RNA-dependent-RNA-polymerases. Also, mutations in this triad interfere with polymerase activity both in vivo and in vitro [90-92]. Motif B contains a methionine-rich loop and is thought to stabilize the template-NTP pairs, while motif D with conserved residues Lys480 and Lys481 is likely to be involved in NTP biding.

There are two putative channels in influenza polymerase: the first is NTP entry channel, which is formed by basic residues from A, D and F3 motives of PB1 subunit. The second is the template entry channel, formed by residues from all three polymerase subunits [86].

PA subunit is known to contain three domains: N-terminal endonuclease domain (residues 1–195) and C-terminal domain (258–714) are found on the opposite sides of the complex, while connected with a linker domain (196–257). This linker makes extensive polar and hydrophobic contacts with PB1 subunit. PA endonuclease makes contacts to both PB1-C and PB2-N and its catalytic site is exposed to the solvent and lies in front of PB2 cap-binding domain (CBD). This positioning is determined by a functional interplay of these two domains which will be discussed in section 1.2.3.

The N-terminal part of PB2 constitutes approximately 1/3 of the PB2 subunit. It is composed of several subdomains which make contacts with the incoming template, PB1 thumb and C-terminal domains, and PA endonuclease domain. C-terminal part of PB2 (residues 248–736) consists of five subdomains: cap-binding domain (residues 319-481), NLS domain (685-760), 627-domain (539-675), mid-domain (248-319) and a cap-627 linker (483-538). These domains constitute one 'arm' of polymerase U-shape [86].

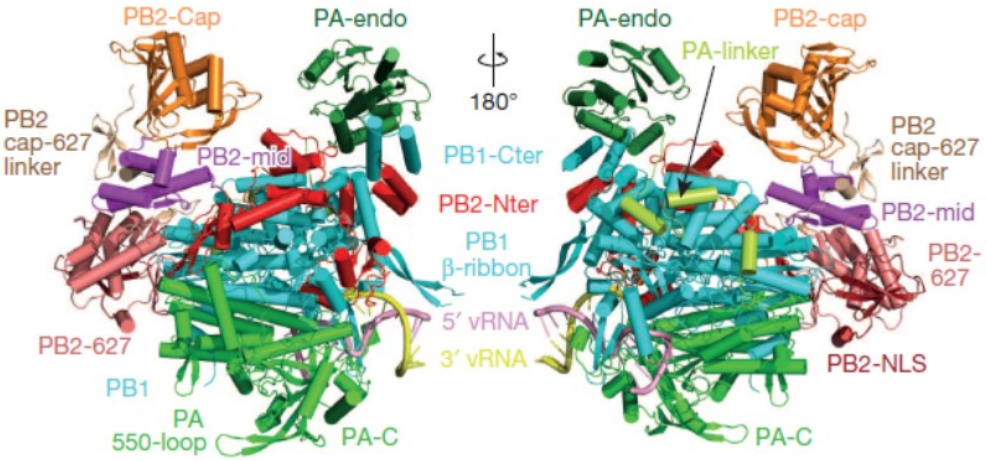


Figure 5. Crystal structure of influenza A polymerase (PDB code 4WSB). Description is provided in the text. Reproduced from [86].

1.2.3 Role of influenza polymerase in a viral life cycle

Influenza polymerase carries out two functions that are essential in a viral life cycle: replication of viral genome and transcription of vRNA to mRNA. While the catalytic centre itself confines mostly to PB1 subunit, both endonuclease activity of PA and cap-snatching activity of PB2 are integral parts of replication and transcription.

1.2.3.1 Replication

Shortly after infection of the host cell, the vRNP are transported to the nucleus where both transcription and replication occur. Influenza virus genome is encoded in RNA of negative polarity and, therefore, its replication is a two-step process: first, polymerase synthetizes so-called complementary RNA (cRNA) which is further used as a template for vRNA synthesis. Replication is primer-independent and starts with the insertion of 3' terminus of vRNA into a catalytic cavity. In the next step, called terminal initiation, a dinucleotide pppApG, complementary to the first nucleotides on vRNA 3' end, is formed [93]. This dinucleotide formation is structurally supported by a priming loop, a β-hairpin that protrudes from the palm domain of PB1 subunit [94]. The current model of replication suggests that cRNA is bound by another polymerase as soon as it protrudes from the product exit channel and along with NP protein forms ribonucleoprotein particles similar to those formed by vRNA [95].

In the second step, cRNA is used as a template to generate vRNA. Unlike the first step of replication, the initiation starts internally – from U4 and C5 template residues and does not require the presence of the priming loop. However, it depends on the presence of another, so-called trans-activating polymerase. Trans-activating polymerase does not directly participate in vRNA synthesis (as the synthesis proceeds even with trans-activating polymerase being catalytically inactive) and its role in transcription remains unclear [93]. Another model suggests that the second step of replication can be carried out not by the resident polymerase, but by a third, trans-acting polymerase [96].

1.2.3.2 Transcription

Transcription results in production of 5′-capped and 3′-polyadenylated viral mRNA that is subsequently translocated to the cytoplasm and transcribed. Influenza virus transcription, unlike replication, is primer-dependent. Also, viral polymerase lacks 5' capping activity that is essential for ribosome recognition and translation. Therefore, the viral polymerase has two activities that address these issues: first, it binds 5' caps of host mRNAs and pre-mRNAs (also, small nuclear RNAs and nucleorar RNAs), cleaves 10-15 nt downstream the cap and uses this oligonucleotide as a primer [97-99]. This process is referred to as cap-snatching and will be described in detail in following sections.

Once cleaved, the oligonucleotide enters the PB1 active site through the product exit channel, while 3'end of vRNA template enters the active site through the template entry channel. The transcription starts with the addition of G or C nucleotide to the 3'end of the primer and elongation proceeds in template-dependent fashion. Each vRNA has 5-7 nt long U-stretch close to its 5' end. Once the U-stretch reaches the active site, polymerase begins to stutter and repeatedly incorporate adenosines, that results in production of poly(A) tail [100]. During the elongation, 5'cap is released from the PB2 cap-binding domain, extends from the product exit channels and can be bound by cellular cap-binding proteins, forming host-like messenger ribonucleoprotein complex which can be exported from the nucleus [101].

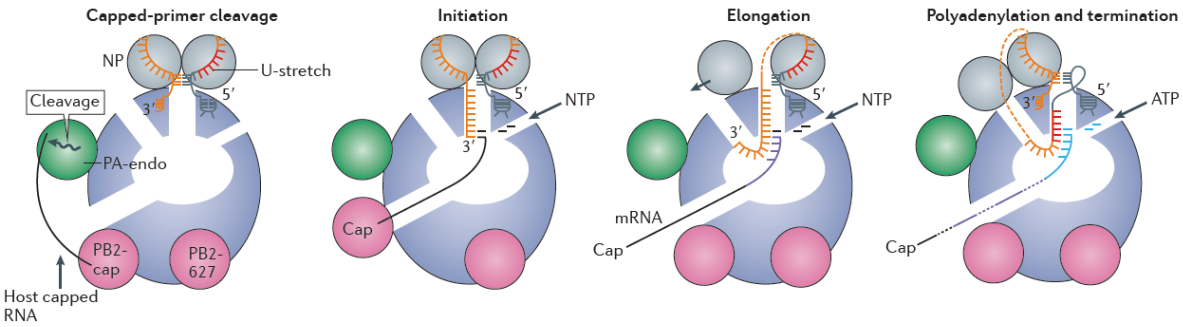


Figure 6. Steps of mRNA synthesis. Details are provided in the text. Reproduced from [102].

1.2.3.2.1 Cap-binding

Along with Orthomyxoviruses, other cap-snatching viruses are Arenaviruses and Bunyaviruses, but it is the influenza virus in which this process is studied most extensively [103, 104].

Cap-binding is the initial step in the sequence of events that leads to viral transcription. It follows by endonuclease cleavage (by N-terminal part of PA) to generate a primer for the viral transcription. It results in all viral mRNAs being chimeric in nature, with first 10-14 nucleotides originated from the host sequences. The cap-snatching process can be divided into following stages: (1) first, PB2 cap-binding domain binds transcripts produced by Pol II; (2) PA-N terminal endonuclease domain cleaves 10-14 nucleotides downstream the cap (3) PB2 caps-binding domain changes its configuration in order to insert the primer into the catalytic center.

The cap-binding domain resides in the middle part of the PB2 protein (residues 319-481 AA). The domain binds the 7-methylguanosine caps of nascent transcripts produced by Pol II: mRNAs, as well as small nuclear RNAs (snRNAs) and small nucleolar RNAs (snoRNAs) [99]. Several studies have shown that influenza polymerase associates with Pol II by binding its C-terminal domain [105, 106]. This provides influenza polymerase with the access to nascent capped RNAs and may regulate the switch between viral transcription and replication modes [106].

The cap-binding domain has unique folding topology [107]. It consists of β-sheet and helical subdomains facing each other, with several shorter β-sheets in the C-terminus. The binding site lies between these two subdomains and is enclosed by a hairpin formed by β3 and β4 helixes and flexible 304-loop.

Despite possessing a distinct fold and no sequence similarity to proteins other than PB2 from different strains, the overall mode of cap binding holds certain similarities to cellular cap binding proteins [108]. In cellular proteins, such as eukaryotic translation initiation factor eIF4E and nuclear cap binding complex CBC, the methyl guanosine moiety stacks between two hydrophobic residues, mostly tryptophan and tyrosine. The methylated base possesses delocalized positive charge which enhances the interaction with the aromatic side chains and contributes to discrimination against non-methylated bases. Apart from the stacking interaction, the binding is also facilitated by the basic cluster of amino acids to accommodate the triphosphate and acidic amino acids that accommodate positively charged 7-methylguanine ring [109].

In PB2 cap-binding domain, the base is sandwiched between two residues, Phe404 (from α1), which is slightly tilted against the base, and His357(from β4) stacking in parallel. Presence of His357 in stacking of 7-methylguanosine is a unique feature of PB2 cap-binding domain and has never been observed in other cap-binding proteins. Another aromatic residue, Phe323, stacks on the sugar moiety. The base moiety is further stabilized by hydrogen bonds between Glu361 and N1 and N2 positions of guanine, while Lys376 makes a salt bridge with O6. The triphosphate makes electrostatic interactions with basic residues His432 and Asn429 (for α-phosphate) and His357, Lys339 and Arg355 (for γ-phosphate) [107]. The flexible 424-loop protrudes between helixes α2 and α3 and is exposed to the solvent. It does not participate in cap binding but plays a significant role in placing capped primer into polymerase active site [110].

1.2.3.2.2 Endonuclease cleavage

The next step of the process is the endonuclease cleavage of bound host mRNA. Until recently, endonuclease activity of influenza polymerase was mistakenly attributed to PB1 subunit. However, the first crystal structure along with a comprehensive functional study conducted by Dias et al. showed, that endonuclease activity resides at the N-terminus of PA subunit (1-195 AA). Also, it was shown that PA endonuclease domain shares certain folding similarities with nucleases of PD-(D/E)XK super-family, such as *Pyrococcus furiosus* Holiday junction resolvase, EcoRV from *Escherichia coli* and BamH1 from *Bacillus amyloliquefaciens*.

PA endonuclease domain consists of 7 α-helixes and 5 β-sheets, with catalytic residues being highly conserved in all influenza types. The catalytic core consists of His41, Asp108, Glu119 and Lys134 amino acids, which is a common motif of (D/E)XK superfamily nucleases. However, the presence of histidine is a unique feature of the PA endonuclease domain and has never been observed in other endonucleases of this family. Two divalent metal ions are coordinated in the active site: M1 by Glu80, Asp108 and 2 water molecules and M2 by His41, Asp108, and Glu119. Certain contradictions remain about the type of the ions present: the first crystal structure by Dias et al. showed two manganese ions coordinated in the active site, while another crystal structure revealed only one magnesium ion coordinated in M1 site, supporting the idea that in the absence of substrate only one magnesium ion could be coordinated [111]. Functional studies showed that cleavage proceeds more effectively in presence of manganese or cobalt and less effective upon the presence of magnesium [112]. Also, the structure was shown to be more thermally stable upon the presence of manganese, with calcium and magnesium having lower stabilization effect [111]. Even though there is little consensus on the type of the ions coordinated and consecutiveness of binding to the active site in native conditions, it is generally agreed that phosphodiester cleavage is carried out by two-metal ion catalysis mechanism [113]. This is supported by a number of experimental data, including studies on thermodynamics of metal binding and kinetic of substrate cleavage upon various concentrations and types of metal ions [112, 114].

The possible two-metal catalysis by influenza endonuclease was modelled by Xiao et al. Both ions are coordinated in a way to be aligned with the phosphosugar backbone of substrate RNA. Ion 1, along with Lys134 and 3'phosphodiester reduce

the pKa of a water molecule, priming it for the nucleophilic attack. Ion 2 stabilizes the pentacovalent intermediate during the attack and bond cleavage [115].

1.2.4 Influenza polymerase as a drug target

As of today, two classes of anti-influenza drugs are approved: M2 channel inhibitors and neuraminidase inhibitors. M2 channel inhibitors, such as rimantadine and amantadine, are no longer recommended for usage due to resistance development in currently circulating strains [81]. Therefore, neuraminidase inhibitors, such as oseltamivir (Tamiflu™) and zanamivir (Relenza™) are the only available treatment of influenza viral infection. However, detection of several mutations causing oseltamivir resistance during 2009 pandemic gave rise to serious concerns about current anti-influenza medications being insufficient in the face of the next flu pandemic [82] .

Influenza polymerase is considered the promising target for the development of new anti-influenza inhibitors. Its functions, such as replication and transcription of genomic information, are essential for the viral life cycle. Moreover, all three subunits of influenza polymerase were found to be conserved in both Influenza A and B viruses [86].

1.2.4.1 Nucleoside analogues

Ribavirin (1-(β-D-ribofuranosyl)-1,2,4-triazole-3-carboxamide) is a broad-spectrum anti-viral medication, with the propensity to inhibit replication of both DNA and RNA viruses. Developed in the early 70s, it has been used as mono- and combinational therapy of hepatitis C and respiratory syncytial virus infections.

Upon administration, adenosine kinase converts ribavirin to mono- di- and tri-phosphorylated metabolites by phosphorylation at its 5′ hydroxyl group [116, 117]. Ribavirin 5′-triphosphate is proposed to interfere with viral polymerase activity leading to inhibition of viral transcription [118, 119]. Additionally, indirect inhibition is conferred through misincorporation of ribavirin 5′-triphosphate to viral genome leading to lethal mutagenesis [120-122]. Another mechanism is the inhibition of the inosine monophosphate dehydrogenase by monophosphorylated ribavirin. This leads

to a decrease of cellular GTP concentration, interfering with viral RNA synthesis [123-125].

Life-threatening adverse effects such as hemolytic anemia and teratogenicity, however, present serious concerns and do not speak in favor of wide usage of ribavirin in anti-influenza therapy. On the other hand, administration of ribavirin as a last resort medication in complicated cases caused by highly pathogenic or multi-resistant viral strains remains a viable option.

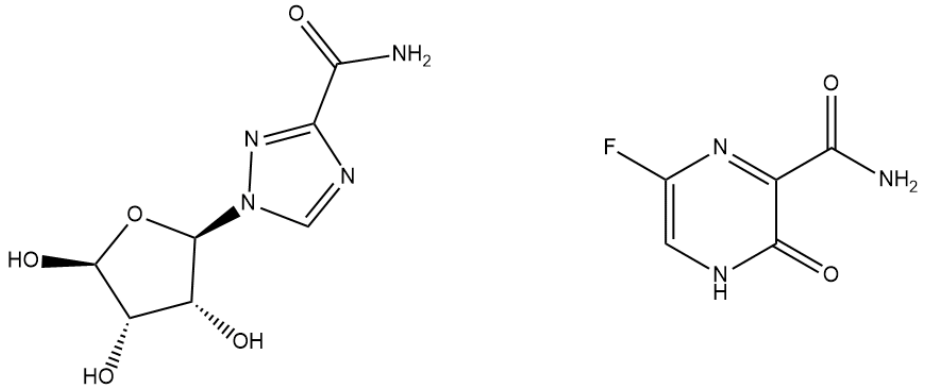
Favipiravir (6-fluoro-3-hydroxy-2-pyrazinecarboxamide) is a purine nucleoside mimetic that targets RNA viruses [126]. Favipiravir was developed by Japanese company Toyama Chem Co., Ltd and got approval for treatment of influenza in Japan in 2014. It is currently in the late-phase clinical trials to seek approval in the USA and Europe.

Favipiravir is found to be active against three major subtypes (A, B, C) of influenza virus as well as against other RNA viruses, including arenaviruses and bunyaviruses [127]. The recent reports show that it might be also active against notorious members of *Filoviridae* family – Marburg and Ebola viruses [128, 129].

In vivo, favipiravir undergoes two-step processing: the first step is catalysed by hypoxanthine guanine phosphoribosyltransferase, that converts favipiravir to favipiravir-ribofuranosyl-5-monophosphate, which is subsequently metabolized to favipiravir-ribosyl-5-triphosphate (favipiravir-RTP) by host cell kinases [130].

Favipiravir-RTP was shown to be incorporated by viral RNA-dependent-RNA polymerase during replication as guanosine and adenosine analogs [131]. In agreement to this, cultivation of influenza virus in the presence of favipiravir was observed to increase mutation frequency and reduce infectivity. It is, therefore, possible, that favipiravir acts by common nucleoside analogs mechanism – lethal mutagenesis, generating non-viable viral phenotypes [132].

Another line of evidence allows suggesting that incorporation of favipiravir-RTP leads to termination of elongation [133]. As Nansens et al. pointed out, these hypotheses are not mutually exclusive, as difference in mechanism might be concentration-dependent [134].



Ribavirin Fapiravir

Figure 7. Structures of ribavirin and fapiravir. Created using ChemDraw software.

1.2.4.2 Endonuclease inhibitors

As was described in previous chapters, endonuclease activity of PA depends on divalent ions that reside at its active centre. It is, therefore, predictable that the main approach to the development of PA endonuclease inhibitors would rely on compounds with metal-chelating properties. Researches from Merck pioneered this strategy in 1994 by developing a series of 2,4-dioxobutanoic acid compounds with IC50 values of 0.2 - 29 μM (determined by transcription-based assay) [135]. Later, another compound with endonuclease inhibitory properties, a substituted 2,6-diketopiperazine, was isolated from fungi Delitschia confertaspora. Named flutimide, the compound was found to have IC50 of 5.5 μM in transcription-based and 5.9 μM in virus-based assays. The study also showed no apparent cytotoxicity in MDCK cells up to 100 μM concentration [136]. There is, however, no evidence of flutimide entering the clinical studies.

Among the compounds with the most potent inhibitory properties is **L-742.001**, 4-substituted-2,4-dioxobutanoic acid with a β-diketo acid motif. Its activity was verified by transcription-based (IC_{50} of 0.43 μM) and virus-based assays (IC_{50} of 0.35 μM) as well as in mice model [137]. The structural study conducted by DuBois et al, in which N-terminal PA endonuclease domain was co-crystallized with series of known influenza endonuclease inhibitors, explained a common mode of actions of both 2,4-dioxobutanoic acid and N-hydroxyimide (structurally related to flutimid) based compounds. The triad of planar oxygens that both classes of compounds contain, coordinates two divalent ions in the active site, with the middle oxygen interacting

with both ions. More precisely, the first divalent ion is coordinated octahedrally by His41, Asp108, Glu119, Ile120 and two inhibitor's oxygens, while the second ion is coordinated tetrahedrally by Glu80, Asp108, and, in the same fashion, by two inhibitors' oxygen.

Apart from coordination of catalytically important divalent ion, L-742.001 was observed to form additional interactions that likely contributed to its potency. The p-chlorobenzene moiety occupies the pocket formed by Tyr24, Ile38, Glu26, and M21. Tyr 24 is displaced by 2 Å when compared with apo structure, which suggests an induced-fit binding mechanism [138]. An additional line of evidence that supports an important role of Tyr24 displacement, comes from mutational studies, in which Tyr24 to Ala24 mutation resulted in a 3-fold decrease of IC50 values in cell culture [139]. The phenyl group of L-742.001 resides in another pocket, formed by Arg84, Trp88, Phe105, and Leu106, while carboxylic acid interacts with catalytically important residue Lys134.

The study conducted by Song et al., however, reported the different orientation of L-742.001 in the crystal structure with N-terminal PA from pandemic H1N1, in which phenyl group occupies a pocket close to Tyr24, while p-chlorobenzene forms π-stacking interaction with Phe105 [140]. This ambiguity can be resolved by closer examination of electron density maps of both structures: a lack of electron density for the inhibitor in the structure by DuBois et al. impedes clear interpretation of the molecule's orientation.

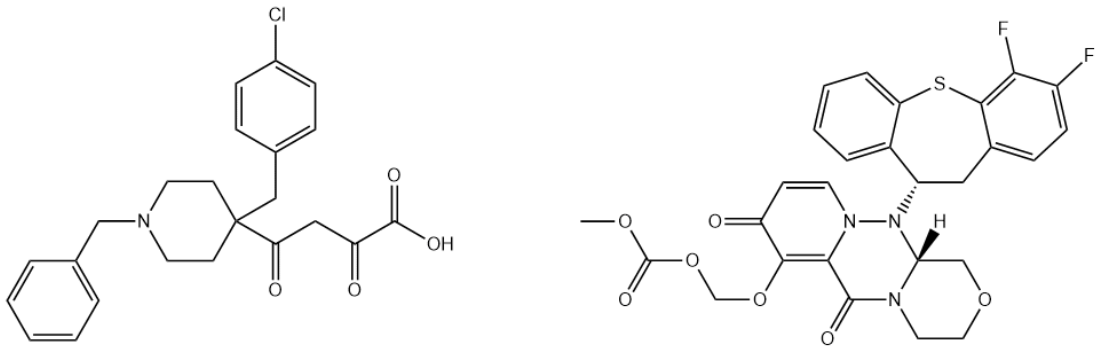
Recently, the first-in-class PA endonuclease inhibitor has successfully reached the market in Japan. **Baloxavir marboxil** (sold under the trade name Xofluza) was developed by Shionogi & Co. in collaboration with Roche and is currently undergoing late trials to seek approval in the United States. The research on this compound was not disclosed until it had successfully passed all stages of clinical trials in Japan and entered late-stages in the US.

Baloxavir marboxil was developed by optimization of metal-chelating pharmacophore used in HIV integrase inhibitors. HIV integrase also contains 2 divalent metal ions and is inhibited by a metal-chelation mechanism. The scientific literature on the topic witnessed continuous allusions to the similarity of these two enzymes and possible usage of HIV integrase inhibitors as starting pharmacophores. Baloxavir marboxil is reported to have broad-spectrum activity against both A and B

viral subtypes, including oseltamivir-resistant strains. EC_{50} values (half maximal effective concentration) of baloxavir marboxil for A/H1N1pdm is 1.3 - 1.6 nmol/L and 5.6 - 8.5 nmol/L for type B strains. Several mutations were detected during the clinical trials, among those I38T had the most significant impact by causing 30- to 50-fold reduction in EC_{50} values.

Baloxavir acid (the active form of baloxavir marboxil) is described as a 'butterfly-like', with two 'wings': oxazino-pyridotriazin-dione group and difluoro-dihydro-dibenzothiepine group. As crystal structure showed, the oxazino-pyridotriazin-dione moiety coordinates the divalent ions in similar manner that was earlier described for 2,4-dioxobutanoic acid. Difluoro-dihydro-dibenzothiepine moiety, on the other hand, establishes van der Waals interactions with Ala20 and Tyr24. Structure of mutant PA endonuclease indicated alterations of the van der Waals interactions upon I38T substitution. In T38 mutant, binding of baloxavir acid also induces a conformation change, as threonine was observed in different rotameric states in apo and inhibitor-bound structures [141].

Less conventional approach to targeting PA endonuclease was taken by Yuan et al, who reported the use of systematic evolution of ligands by exponential enrichment (SELEX) method to select for DNA aptamers with inhibitory properties. The screen indeed resulted in several hits and the best DNA aptamer inhibited PA endonuclease with IC_{50} of 10 nM in a cell-based assay. Pharmacokinetic properties of DNA aptamers, however, make them unsuitable for clinical use [142].



L-742.001 Baloxavir marboxil

Figure 8. Structures of L-742.001 and baloxavir marboxil. Created using ChemDraw software.

1.2.4.3 Cap-binding inhibitors

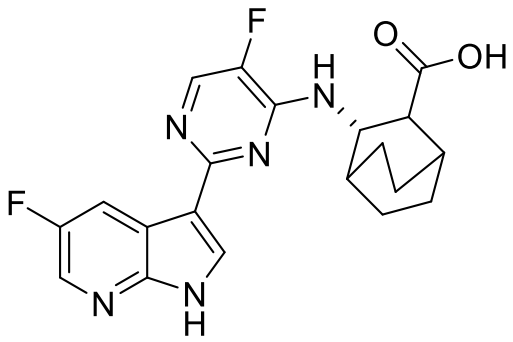
As was described in the previous chapter, PB2 CBD domain differs from cellular cap binding proteins in several aspects, including overall protein fold and key residues that facilitate the binding. The crystal structure of PB2 CBD bound to 7-methylguanosine triphosphate (m^7GTP) solved in 2008 by Guilligay at el., therefore, opened promising avenues for structural-based drug design [143].

The first step to the development of cap-binding inhibitor was undertaken in 2014 when the anti-viral activity of 3-pyrimidineazaindoles was identified by phenotypic cell protection screening assay. Structure of PB2 CBD in complex with azoindole-based hit compound revealed the key interactions: azoindole moiety stacks between two aromatic residues – His357 and Phe404, while pyrimidine forms π-stacking interaction with Phe323. Similar to m^7GTP, the compound formed hydrogen bonds with Glu361 and Lys376, and dimethylalanyl occupies the space previously attributed to ribose and phosphate moieties of m^7GTP. Based on the structure and pharmacokinetic studies, the compound was further optimized. In particular, 5-chloroazaindole was substituted with 5-fluorazaindole, since the later was found to have superior properties in terms of clearance and selectivity. Also, substitution of the 4-position of the pyrimidine ring system with bicyclooctane-carboxylate moiety further improved the selectivity. When compared to the lead compound, a newly optimized compound named **VX-787** forms additional water-mediated contacts with His357, Gln406, Arg355.

Biochemical characterization of the interaction between VX-787 and PB2 CBD by isothermal titration calorimetry showed, that VX-787 binds to PB2 with K_d of 24 nM, which is 60-fold lower than for binding of the natural substrate – m^7GTP [144].

Antiviral activity was tested both *in vitro* and *in vivo*. VX-787 had strong antiviral properties E_{50} of 1.6 nM in a cytopathic effect (CPE) assay. In vitro, VX-787 was shown to be active against all influenza A strains tested, including pandemic H1N1, highly pathogenic H5N1 and strains, resistant to amantadine and neuraminidase inhibitors. When administrated to mice 48 h postinfection with A/PR/8/34 strain, treatment with VX-787 led to 100% survival rate in 10 days' time frame, while treatment with oseltamivir resulted in no survival benefit. VX-787 also offered complete protection from the infection when used as prophylaxis in mice infected with pandemic H1N1 and highly pathogenic H5N1.

VX-787, however, is not active against influenza B strains, presumably, due to the structural differences and flexibility of the cap-binding domain in B virus type. Also, during preclinical studies, several mutations causing significant (>10 fold-change) reduction in sensitivity to VX-787 was found. As of now, VX-787, known under commercial name pimodivir, has advanced to phase III of clinical trials [145].



VX -787 (pimodivir)

Figure 9. Structure of VX 787, also known as pimodivir. Created using ChemDraw software.

1.2.4.4 Protein–Protein Interaction Inhibitors

Protein-protein interactions (PPI), although generally less explored in drug discovery, present another avenue of research on anti-influenza medications. Generally, potential "druggability" of protein-protein interfaces is evaluated based on the size of the interface and degree of amino acid conservation. Heterotrimeric composition of influenza polymerase results in several protein-protein interfaces. Among those, the interface between C-terminal PA (residues 257–716) and N-terminal PB1 (residues 1–25) is considered the most amenable to drug design and being studied extensively. The structural basis of this interaction was revealed by He et al in 2008 [146]. The interface itself was found to be small, with C-terminus of PA forming a hydrophobic cavity where the N-terminal PB1 peptide is accommodated. When compared to the later published structure of the whole influenza polymerase, one can notice the higher degree of flexibility on the outer edges of the cavity, which implied that bottom part of the cavity would be a preferable target for PPI inhibitors [147]. The crystal structure has laid the groundwork for in silico screening studies,

which resulted in identification of several compounds that were also active in cell culture [148-150]. Activity of a recently reported compound – ANA-1 was also confirmed in vivo [151]. It is, however, unknown whether these compounds indeed act through disruption of interface between C-terminal PA and N-terminal PB2, as structural and resistance selection studies are yet to be performed.

Taken together, recent advances in understanding of function and structure of influenza polymerase have laid the foundation for development of anti-influenza medications that would target influenza polymerase. Two functions of influenza polymerase have been proved particularly exploitable in drug development: cap-biding and endonuclease cleavage. Several drug-candidates that interfere with these functions are in clinical trials, or, have been recently approved. Development of new methodological approaches, such as novel high-throughput screening methods would be particularly useful and might bring additional advances in development of compounds that would target these activities.

2. Aims of the thesis

- Design and expression of recombinant PA endonuclease and PB2 cap-binding domains from influenza A polymerase for utilization in high-throughput screening assays
- Purification of the recombinant fusion proteins
- Development and optimization of high-throughput screening assay based on AlphaScreen technology for screening of compounds inhibiting endonuclease activity of PA protein
- Development and optimization of DIANA-based high-throughput screening assay for the identification of potential disruptors of PB2 cap-binding activity

3. Materials and methods

3.1. Material, chemicals, instruments

3.1.1 Material

- 3 Lens Low Profile Plate, Swissci (Switzerland)
- Amicon Centrifugal filter units (MWCO 10,000, 30,000) 0.5, 4, 50 ml, Merck Millipore (USA)
- Dialysis membrane Spectrapore (6,000-8,000 MWCO), Thermo Fisher Scientific (USA)
- EasyXtall 15-well plate, QIAGEN (Germany)
- FrameStar® 96 Well Semi-Skirted PCR Plate, 4titude (UK)
- Nunc 96-well plate transparent, Thermo Fisher Scientific (USA)

3.1.2 Biological and special material

- 3 Lens Low Profile Plate, Swissci (Switzerland)
- *Escherichia coli,* Top10 strain, Novagen (USA)
- *Escherichia coli,* BL21 (DE3) RIL strain, Novagen (USA)
- Neutravidin, Thermo Fisher Scientific (USA)
- pGEX1-λT vector, obtained from Hans-Georg Kräusslich from the laboratory at the University of Heidelberg, Germany (originally from GE Healthcare Life Sciences)
- pETM11-SUMO3 vector, obtained from Dmytro Yushchenko Ph.D. from IOCB Prague (originally created at Protein Expression and Purification Core Facility, EMBL)
- ULP1 Protease, prepared by Michal Svoboda from IOCB Prague
- U-35 anti-His i-Body, prepared by Vladimír Šubr from IMC of the CAS

3.1.3 Instruments

- BlueWasher, Blue Cat Bio (Germany)
- Camera Olympus E-620, Olympus Corporation (Japan)
- Centrifuges

  Allegra X-15R, Beckman Coulter (USA)

Microcentrifuge 5415R, Eppendorf (Germany)

- Crystal Gryphon, Art Robbins Instruments (USA)
- Digital tube Roller SRT6D, Stuart (UK)
- EmulsiFlex-C3, Avestin (Canada)
- EnSpire Multimode Plate Reader, PerkinElmer (USA)
- Fraction Collector Frac-950, GE Healthcare Life Sciences (USA)
- Gallery DT plate hotel, Rigaku (Japan)
- Horizontal electrophoresis apparatus, Gibco (USA)
- IPC High Prescission Multimode Dispenser, Ismatec (Germany)
- Incubator IPP 400, Memmert (USA)
- Infinite Reader M1000 PRO, Tecan (Switzerland)
- LightCycler®480 Instrument II, Roche Life Science (Germany)
- MP-500V power supply, Major Science (USA)
- Microscope Olympus SZX10, Olympus Corporation (Japan)
- Oryx 8, Douglas Instruments (UK)
- Rotary shaker Innova 44, Eppendorf (Germany)
- SPR sensor platform was developed at the Institute of Photonics and Electronics of the CAS
- Spectrophotometer, Specord 210, ChromSpec (Czech Republic)
- ThermoCell Mixing block, BIOER (China)
- Trio 48 Thermocycler, Biometra (Germany)
- UV lamp UVT - 20 S/M/L, Herolab (Germany)
- Vertical Apparatus - PAGE, Bio-Rad (USA)
- Voltage source EPS 301, GE Healthcare (USA)
- Xtal UV Detector, Rigaku (Japan)
- pH meter model pH 50, XS Instruments (Italy)
- Äkta Purifier 10, GE Healthcare Life Sciences (USA)

3.1.4 Chemicals

- Biacore (Sweden)
  1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride (EDC), N-Hydroxysuccinimide (NHS), ethanolamine

- Biotika (Slovak Republic)
  ampicillin
- Lach-Ner (Czech Republic)
  ethanol, isopropanol, sacharose
- Penta (Czech Republic)
  glycerol, acetic acid, sodium chloride, hydrochloric acid, sodium acetate
- Prochimia (Poland)
  HS-(CH_2)11-EG_4-OH, HS-$(CH_2)_{11}$-EG_6-OCH_2-COOH
- Promega (USA)
  Tris base
- SDT (Germany)
  casein buffer 20×-4× concentrate
- Serva (Germany)
  Agarose
- Sigma-Aldrich (USA)
  SDS, TEMED, APS, 2-mercaptoethanol, bromphenol blue, HEPES, acrylamide, dimethyl sulfoxide, PEG 3350, glycine, EDTA, kanamycin sulphate, TCEP, d-desthiobiotin, L-glutathione reduced (GSH), sodium phosphate dibasic heptahydrate, potassium dihydrogen phosphate dihydrate, LB Broth, imidazole, CBB-R250, sodium ascorbate, N, N´- methylene bisacrylamide, raltegravir, L-742.001
- TATAA Biocenter (Sweden)
  VX-787
- Thermo Fisher Scientific (USA)
  IPTG
- USB (USA)
  Tween 20

3.1.5 Other material

- 10x T4 DNA Ligase Reaction buffer, New England Biolabs (UK)
- 5x Phusion HF Buffer, New England Biolabs (UK)
- Alkaline Phosphatase, New England Biolabs (UK)
- cOmplete™, EDTA-free Protease Inhibitor Cocktail, Roche (Switzerland)
- CutSmart® buffer, New England Biolabs (UK)
- Elvitegraivr, was provided by Dr. Jan Weber, IOCB Prague

- GelRed Nucleic Acid Gel Stain, Biotum (USA)
- JBScreen JCSG++, Jena Biosciences (Germany)
- Morpheus® crystallization screen, Molecular Dimensions (UK)
- Ni-NTA agarose, Qiagen (Germany)
- PCR dNTP Mix, Top-Bio, (Czech Republic)
- PPP Master Mix, Top-Bio (Czech Republic)
- Phusion® High Fidelity DNA Polymerase, New England Biolabs (UK)
- Pierce™ Glutathione Agarose, Thermo Fisher Scientific (USA)
- Plasmon IV immersion oil, Gargille (UK)
- Protein Assay Dye Reagent Concentrate, Bio-Rad (USA)
- QIAquick Gel Extraction Kit, Qiagen (Germany)
- Restriction enzyme NdeI, New England Biolabs (UK)
- Restriction enzyme BamHI, New England Biolabs (UK)
- Restriction enzyme XhoI, New England Biolabs (UK)
- T4 DNA Ligase, New England Biolabs (UK)

3.2 Molecular Cloning

3.2.1 GST-NPA fusion recombinant construct for bacterial expression

Sequence for N-terminal PA (residues 1-209) from viral strain A/California/07/2009 was ordered from GenScript (genbank code CY121685). As the protein was planned to be used in crystallization trials, a flexible loop that causes the protein to resist crystallization (residues 51–72) was replaced with GGS liker according to Song et al. [140]. The sequence contained two restriction sites: NdeI (highlighted in purple) and EcoRI (highlighted in yellow), two stop codons (highlighted in red, the first nucleotide of EcoRI sequence also belongs to the stop codon) and Precission protease recognition site (highlighted in turquoise).

```
  1 CATATGCTGG AAGTTCTGTT CCAGGGGCCC ATGGAAGACT TTGTGCGACA ATGCTTCAAT
 61 CCAATGATCG TCGAGCTTGC GGAAAAGGCA ATGAAAGAAT ATGGGGAAGA TCCGAAAATC
121 GAAACTAACA AGTTTGCTGC AATATGCACA CATTTGGAAG TTTGTTTCAT GTATTCGGAT
181 GGTGGTAGTA AGCACCGATT TGAGATAATT GAAGGAAGAG ACCGAATCAT GGCCTGGACA
241 GTGGTGAACA GTATATGTAA CACAACAGGG GTAGAGAAGC CTAAATTTCT TCCTGATTTG
301 TATGATTACA AAGAGAACCG GTTCATTGAA ATTGGAGTAA CACGGAGGGA AGTCCACATA
361 TATTACCTAG AGAAAGCCAA CAAAATAAAA TCTGAGAAGA CACACATTCA CATCTTTTCA
421 TTCACTGGAG AGGAGATGGC CACCAAAGCG GACTACACCC TTGACGAAGA GAGCAGGGCA
481 AGAATCAAAA CTAGGCTTTT CACTATAAGA CAAGAAATGG CCAGTAGGAG TCTATGGGAT
541 TCCTTTCGTC AGTCCGAAAG ATGATAGAAT TC
```

Figure 10. Segment encoding the N-terminal PA endonuclease domain for insertion into pGEX-1λT plasmid (before cleavage with restriction nucleases). Restriction sites for NdeI and EcoRI are shown in purple and yellow, sequence encoding the Prescission protease cleavage site is in turquoise and stop codons are in red.

The sequence was obtained in commercial plasmid and was transformed into *Escherichia coli*-Top10 competent cells (procedure is described in section 3.2.9). The ampicillin resistant clones were selected by cultivation on agar plates containing antibiotic, the plasmid DNA was isolated according to the protocol described in the section 3.2.10. The desired fragment was cleaved out with restriction endonucleases and ligated into pGEX-1λT plasmid, that was transformed into *E. coli* for large-scale expression (sections 3.2.8 and 3.2.9).

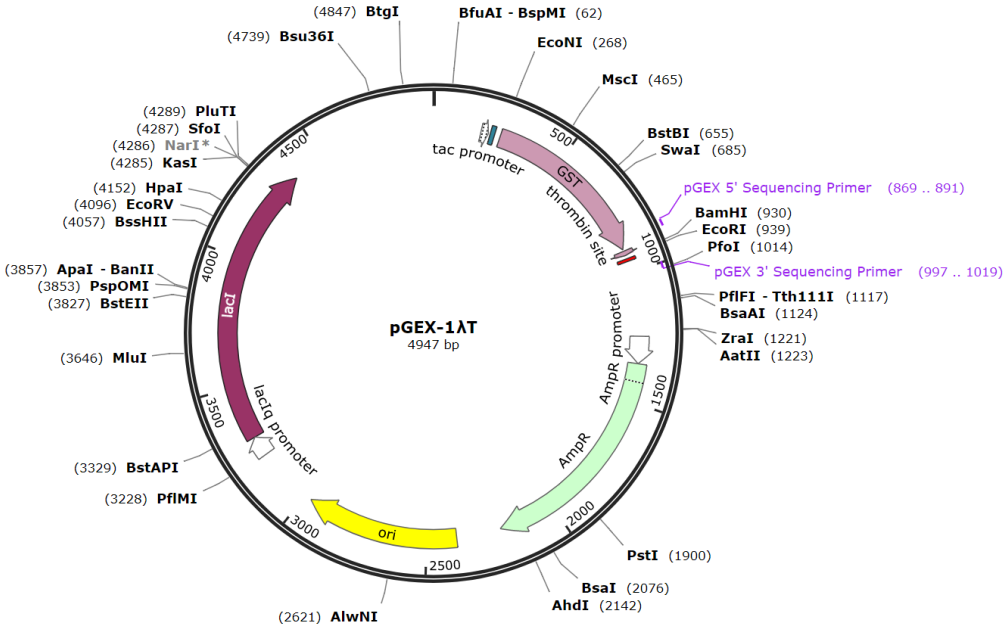


Figure 11. Map of pGEX-1λT plasmid, reproduced from [152] .

3.2.2 PB2-His_6 fusion recombinant construct for bacterial expression

Sequence for PB2 cap-binding domain (residues 318-483) from A/California/07/2009 (GenBank FJ96697) was ordered from GenScript (USA). The sequence contained endonuclease restriction sites for NdeI and BamHI (Figure 12, showed yellow and violet). The commercial plasmid, in which the sequence was delivered, was transformed to *E. coli* Top10 (section 3.2.9), followed by DNA micropreparation, restriction nuclease cleavage (sections 3.2.10 and 3.2.8) and ligation to the pET-24a plasmid (section 3.2.7).

Prior to cloning of PB2-His_6 construct, several other constructs were cloned and expressed. The PB2-GST was poorly expressed in *E. coli*, while PB2-Strep-tag was found less suitable for usage in DIANA experiments due to weaker binding between Strep-tag and anti-Strep-tag antibody than His-tag and Anti-His-tag iBody (observed in SPR experiments). Cloning and expression of these constructs are not shown.

Therefore, the PB2 sequence was amplified from one of the plasmids previously used for expression of other constructs and modified by addition of a Prescission protease cleavage site (Figure 12, showed in turquoise) by PCR primer extension (section 3.2.4).

Primers used for amplification:

Forward:

5’ATATTTTGGATCCGGCAGCGGCAGCGGCAGCGGCAGCATGGAAGACTTTG TGCGACAA’ 3

Reversed:

5’ CACGATCTCGAGCTATCATCTTTCGGACTGACGAAAGGAATCCCATAG 3’

```
  1 CATATGAGGA TTAGCTCATC TTTCAGTTTT GGTGGGTTCA CTTTCAAAAG GACAAGCGGA
 61 TCATCAGTCA AGAAAGAAGA AGAAGTGCTA ACGGGCAACC TCCAAACACT GAAAATAAGA
121 GTACATGAAG GGTATGAAGA ATTCACAATG GTTGGGAGAA GAGCAACAGC TATTCTCAGA
181 AAGGCAACCA GGAGATTGAT CCAGTTGATA GTAAGCGGGA GAGACGAGCA GTCAATTGCT
241 GAGGCAATAA TTGTGGCCAT GGTATTCTCA CAGGAGGATT GCATGATCAA GGCAGTTAGG
301 GGCGATCTGA ACTTTGTCAA TAGGGCAAAC CAGCGACTGA ACCCCATGCA CCAACTCTTG
361 AGGCATTTCC AAAAAGATGC AAAAGTGCTT TTCCAGAACT GGGGAATTGA ATCCATCGAC
421 AATGTGATGG GAATGATCGG AATACTGCCC GACATGACCC CAAGCACGGA GATGTCGCTG
481 AGAGGGATAA GAGTCAGCAA AATGGGACTG GAAGTTCTGT TCCAGGGGCC GGATCC
```

Figure 12. Segment encoding the PB2 cap binding domain for insertion into pET-24a vector (before cleavage with restriction nucleases).

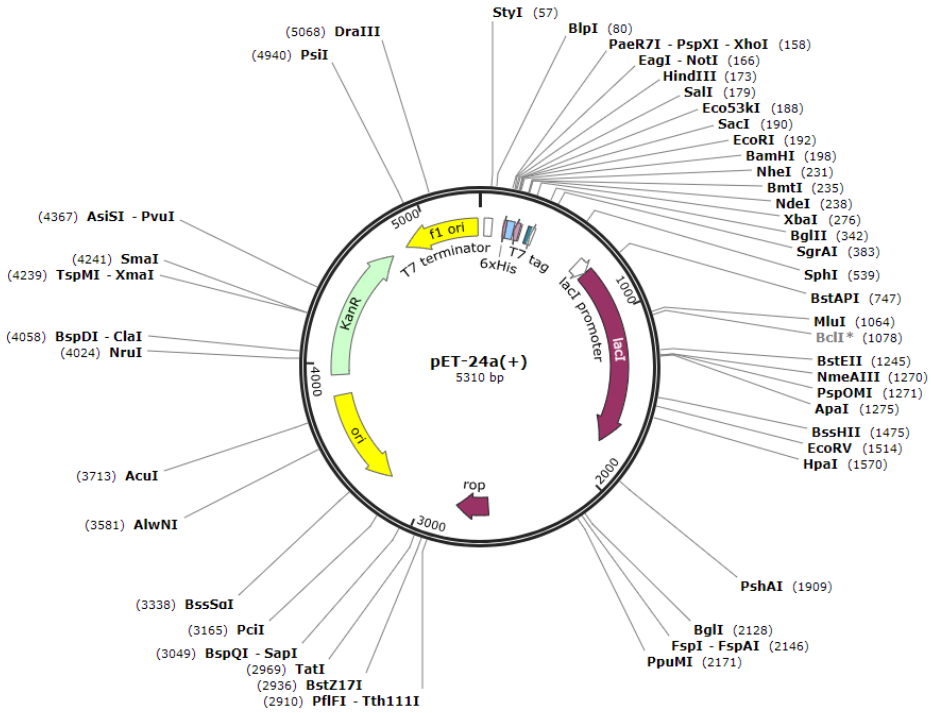


Figure 13. Map of pET-24a vector. Reproduced from [153].

3.2.3 His_6-SUMO-NPA fusion recombinant construct for bacterial expression

This construct with cleavable tag was designed for crystallization trials. The sequence of NPA construct was modified to add N-terminal 4 x GS linker (showed in turquoise) between the NPA and SUMO by PCR primer extension, as absence of the linker resulted in poor cleavage by ULP1 protease. The final sequence shown in Figure 14 contained cleavage sites for BamHI (in violet) and XhoI (in yellow) and 2 stop codons (in red). The sequence was cleaved with corresponding endonuclease and ligated to the vector plasmid pETM11-SUMO3.

Primers used for amplification:

Forward: 5' ATATTTTCATATGAGGATTAGCTCATC '3

Reversed: 5 ' ATATTTTGGATCCGGCCCCTGGAACAGAACTTCCAG '3

```
   1 GGATCCGGCA GCGGCAGCGG CAGCGGCAGC ATGGAAGACT TTGTGCGACA ATGCTTCAAT
  61 CCAATGATCG TCGAGCTTGC GGAAAAGGCA ATGAAAGAAT ATGGGGAAGA TCCGAAAATC
 121 GAAACTAACA AGTTTGCTGC AATATGCACA CATTTGGAAG TTTGTTTCAT GTATTCGGAT
 181 GGTGGTAGTA AGCACCGATT TGAGATAATT GAAGGAAGAG ACCGAATCAT GGCCTGGACA
 241 GTGGTGAACA GTATATGTAA CACAACAGGG GTAGAGAAGC CTAAATTTCT TCCTGATTTG
 301 TATGATTACA AAGAGAACCG GTTCATTGAA ATTGGAGTAA CACGGAGGGA AGTCCACATA
 361 TATTACCTAG AGAAAGCCAA CAAAATAAAA TCTGAGAAGA CACACATTCA CATCTTTTCA
 421 TTCACTGGAG AGGAGATGGC CACCAAAGCG GACTACACCC TTGACGAAGA GAGCAGGGCA
 481 AGAATCAAAA CTAGGCTTTT CACTATAAGA CAAGAAATGG CCAGTAGGAG TCTATGGGAT
 541 TCCTTTCGTC AGTCCGAAAG ATGATAGCTC GAG
```

Figure 14. Segment encoding the N-terminal PA endonuclease domain prepared for insertion into pETM11-SUMO3 vector (before cleavage with restriction nucleases).

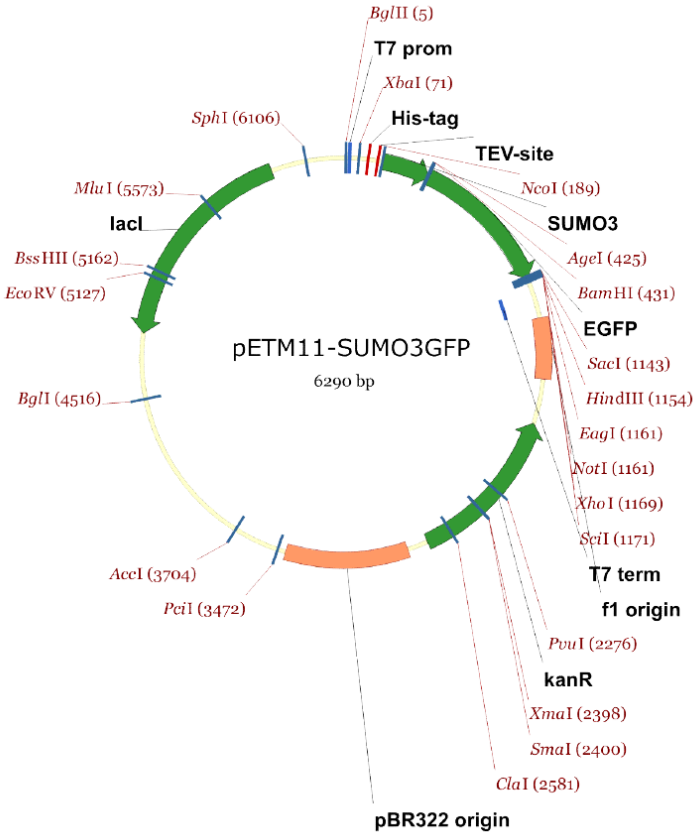


Figure 15. Map pETM11-SUMO3 vector. Reproduced from [154]

3.2.4 Polymerase chain reaction

Polymerase chain reaction was used for two purposes: target gene amplification and testing ligation in selected bacterial clones. The setups differed depending on the type of experiment: Phusion® High-Fidelity DNA polymerase (New England Biolabs) was used for gene amplification, while PPP Master Mix (Top-Bio) was used for testing ligation.

Table 1. Preparation of PCR reaction solution for gene amplification

Component	Reaction volume
Pfu buffer 10x	4.5 µl
10 mM dNTPs	1 µl
20 µM Forward Primer	1 µl
20 µM Reverse Primer	1 µl
Template DNA	1 µl
Phusion DNA polymerase	1 µl
miliQ H_2O	to 45 µl

Table 2. Preparation of PCR reaction solution for testing of bacterial clones

Component	Reaction volume
PPP master mix	5 µl
miliQ H_2O	4.5 µl
20 µM Forward Primer	0.25 µl
20 µM Reverse Primer	0.25 µl

Table 3. PCR protocol

Step	Temperature	Time
DNA denaturation	95 °C	5 min
30 cycles of amplification	95 °C	30 sec
	54 °C	30 sec
	72 °C	3 min 30 sec
Final extension	72 °C	8 min

3.2.5 Agarose gel electrophoresis

- 6x Sample Buffer: 10% Sacharose (w/v), 0.1% bromphenol blue
- TAE Buffer: 40 mM Tris/Acetate, pH 8.3, 1 mM EDTA

The gel was prepared by adding 4 µl of 10,000x GelRed dye (Biotum) to 40 ml of 1.2 % agarose (w/v) in TAE buffer. The mixture was poured into the platform and left for 20 minutes to polymerize. Then, the gel was transferred into the electrophoretic cage with TAE buffer. Sample buffer was added to the samples and the mixture was loaded into gel wells. The experiment was run with the voltage set to 120 V for approximately 25 minutes.

3.2.6 Agarose gel extraction

The gels were visualized in the UV plate reader (UVT). DNA fragments of appropriate size were excised and transferred to the microtubes. 700 µl of QG buffer (Qiagen) were added and samples were left incubating at 50°C in a heating block (BIOER). Once the gel dissolved completely, the solution was transferred to DNA centrifugal filter columns

(Qiagen) and centrifuged for 1 min at 13,000 *g*. The columns were then washed with 700 µl of PE buffer (Qiagen) and column-bound DNA was eluted with 40 µl of miliQ water.

3.2.7 Treatment with restriction endonucleases

After cultivation of bacteria transformed with plasmid that contained desired gene fragments, the plasmid DNA was isolated and treated with restriction endonucleases. Plasmids were also treated with the same restriction endonucleases as the fragments intended for ligation.

Table 4. Preparation of reagents for cleavage with restriction nucleases

Component	Volume
Cutsmart Buffer	3.1 µl
DNA [ul]	30 µl
Endonuclease I	1.5 µl
Endonuclease II	1.5 µl
miliQ H_2O	to 50 µl

Plasmid with the sequence encoding N-terminal PA construct and pGEX-1λT vector were treated with NdeI and EcoRI, while sequence encoding PB2 cap-binding domain and pET24a plasmid were treated with NdeI and BamHI. DNA fragment for His6-SUMO-NPA fragment and pETM11-SUMO3 plasmid were treated with BamHI and XhoI. All reactions were incubated at 37°C for 2 h.

Plasmids (pGEX-1λT, pET24a and pETM11-SUMO3) were additionally treated with 1 ul of alkaline phosphatase to dephosphorylate 5′ and 3′ ends. Dephosphorylation proceeded for 20 min at room temperature. Desired fragments were separated from other cleavage products by agarose gel electrophoresis and isolated (sections 3.2.5 and 3.2.6).

3.2.8 Ligation of target DNA sequence into bacterial plasmid

- 10x T4 DNA Ligase buffer (New England Biolabs)

After the cleavage of both plasmid and DNA fragment encoding corresponding constructs, these two sequences were ligated with T4 DNA Ligase. Reactions were prepared according to the Table 5:

Table 5. Preparation of ligation reaction

Component	Reaction volume
10 x T4 ligase buffer	4 µl
DNA	10 µl
Plasmid DNA	0.5 – 2 µl
T4 DNA Ligase	0.8 µl
miliQ H_2O	to 40 µl

Ligation proceeded for 4 h at 16°C. The reaction mixture then was used for transformation to Top10 *E.coli* cells.

3.2.9 DNA plasmid transformation to competent bacterial cells

In the first transformation of commercially obtained plasmids, 1 µl of plasmid DNA was added to 70 µl of bacterial cell suspension (*Escherichia coli*-Top10) and incubated on ice for 10 minutes. Then, the samples were incubated at 42 °C for 90 seconds in a heat block and transferred back on ice for another 10 minutes. Then 350 µl of LB medium was added and the samples were incubated at 37 °C for 1 h. Then, the bacterial suspensions were spread onto agar plates that contained corresponding antibiotics and cultivated at 37 °C overnight. Individual colonies were then picked ant tested by insert amplification (colony PCR). Positives colonies were selected and grown at larger volumes of LB media (with antibiotic) overnight and DNA was isolated for further experiments.

Transformations that were carried out to test ligation of a DNA segment into plasmid were carried out with larger volumes of ligation mixture – 10 µl. Finally, transformations of bacterial cells used for large-scale protein expression followed the same protocol, expect for the use of different *E. coli* strain - *Escherichia coli* BL21 (DE3) RIL.

3.2.10 Minipreparation of plasmid DNA

- Solution I: 25 mM Tris/HCl pH 8.0, 10 mM EDTA, 50 mM glucose, 25 mg/ml RNase A; 1.4 mg/ml lysozyme
- Solution II: 0.2 M NaOH, 1 % (w/v) SDS
- Solution III: 3 M potassium acetate, pH 5.5
- Buffer QG: 20 mM Tris/HCl, pH 6.6; 5.5 M guanidine thiocyanate

- Buffer PE: 10 mM Tris/HCl, pH 7.5; 80 % (v/v) ethanol

First, the 50 ml tubes with bacterial cultures grown overnight in the presence of antibiotics were centrifuged at 6,000 *g* for 10 min. 245 µl of solution I were added to the pellet and suspension was transferred to the microtubes, followed by 10 min incubation. Then, 450 µl of solution II were added, the tubes were gently mixed and incubated for another 10 min. Lastly, 338 µl of solution III were added and samples were thoroughly vortexed and then centrifugated for 6 min at 13,000 *g*. The supernatant, containing DNA was separated from protein-rich pellet and transferred to 15 ml tubes that contained isopropanol, 0.3 M sodium acetate and QG buffer (Qiagen). The solution was then transferred to DNA centrifugal filter columns (Qiagen) and centrifugated for 1 min at 13,000 *g.* Then, 700 µl of PE buffer was added to each column and centrifugated again for 1 min at 13,000 *g*. Finally, the top part of the filter unit was transferred to a microtube and 40 µl was added to elute membrane-bound DNA. After incubation for 5 min, the DNA was eluted by centrifugation at 13,000 *g* for 1 min.

3.3 Bacterial expression of recombinant proteins

- Lysis buffer for GST-NPA: 50 mM Tris/HCl, pH 8.0, 150 mM NaCl
- Lysis buffer for PB2-His6: 50 mM Tris/HCl, pH 8.0, 200 mM NaCl, 10 mM imidazole

Correct incorporation of DNA insert was additionally verified by sequencing the plasmid DNA (GATC-Biotech). 3 µl of DNA plasmid sequence was transformed into 60 µl of *Escherichia coli*, BL21 (DE3) RIL strain. The bacterial culture was selected for the transformed cells by overnight cultivation on two agar plates with a corresponding antibiotic (ampicillin for GST-NPA, kanamycin for PB2-His_6). Inoculum was prepared by washing off the colonies with 6 ml of LB medium. Then, inoculum was added into 3 l of LB medium with antibiotics. LB medium was divided into 6 Erlenmeyer flasks with 0.5 l of media each and 1 ml of inoculum suspension was added to each flask.

The cultivation flasks were then transferred to the rotary shaker (Eppendorf) pre-heated at 37 °C. Sufficient aeration was secured by shaking the flasks at 220 RPM. Growth rate was monitored by measuring optical density (OD) at 595 nm in 30-minutes interval. Once the culture was close to read the stationary phase (OD of approximately 0.8),

expression was induced by addition of 1 ml of 375 mM IPTG was added to each flask to the final concentration of 0.75 mM. Expression proceeded overnight at 18 °C.

3.4 Protein purification

3.4.1 Bacterial cell lysis

Morning after the overnight cultivation, the medium with cells was centrifugated for 15 min at 6,000 *g* (Beckman) and pellet of cells was resuspended into 200 ml of lysis buffer with addition of 1 tablet of EDTA-free protease inhibitors (Roche). The cell suspension was homogenized with glass homogenizer while being kept on ice. The cells were lysed by using Emulsiflex homogenizer upon constant pressure of 1,100 bar, the process was repeated 3 times. Fraction of soluble protein was isolated by long centrifugation of the lysate (15,000 *g*, 4°C, 40 min) and decantation of the supernatant.

3.4.2 Affinity purification and chelation chromatography

3.4.2.1 Purification of GST-NPA on the glutathione-agarose resin

- Wash buffer: 50mM Tris/HCl, pH 8.0, 150mM NaCl, 1 mM TCEP
- Elution buffer: 50mM Tris/HCl, pH 8.0, 150mM NaCl,10 mM reduced L-glutathione (GSH)

Glutathion-sepharose resin was stored in 50% ethanol and needed to be equilibrated prior to purification. 2 ml of 50% ethanol suspension of resin was centrifuged at 500 *g* for 2 min to remove the supernatant. The resin was thoroughly washed by repeated centrifugation with 10 ml of washing buffer. Once ethanol was removed, the resin was added to the cell lysate and left incubating for 60 min while constantly mixed.

Lysate with resin suspension was then divided into 4 tubes of 50 ml and centrifuged for 10 min at 500 *g*. The supernatant was collected as flow-through fraction for further analysis and resin was washed twice with 10 ml of washing buffer. The protein was eluted by applying 6 ml of elution buffer and incubating for 10 mins prior to elution. This step was repeated three times (three elution fractions were collected).

First and second elution were then pulled together and concentrated to 5 ml using 50 ml Amicon centrifugal filter units (MWCO 30 kDa).

3.4.2.2. Purification of PB2-His6 and His_6-SUMO-NPA on Ni-NTA agarose resin

- Wash buffer: 50 mM Tris/HCl, pH 8.0, 300 mM NaCl, 20 mM imidazole
- Elution buffer: 50 mM Tris/HCl, pH 8.0, 200 mM NaCl, 250 mM imidazole
- Dialysis buffer: 50 mM Tris/HCl, pH 8.0, 150 mM NaCl

Ni-NTA agarose was equilibrated as was described in previous section. 1 ml of Ni-NTA resin was incubated with cell lysate for 90 min on rotator at 4 °C (the volume of lysate was 100 ml). The rest of the procedure proceeded as was described earlier. Eluted fractions were dialyzed overnight in 1.5 l of dialysis buffer in dialysis tubing with MWCO of 6-8 kDa at 4°C.

3.4.2.3 Cleavage of SUMO-His_6 with ULP1 protease

PA-SUMO-His_6 was cleaved by ULP1 protease to obtain tag-free protein for crystallization trials. An aliquot ULP1 protease (200 µl, 400 µg/ml) was added to protein fraction in mass ratio that did not exceeded 1:80 (protease : protein) and incubated in on a rotator at 4 °C overnight.

Then, the tag and non-cleaved protein was separated from the tag-free protein by using Ni-NTA chelation chromatography (section 3.4.2.2).

3.4.3. Gel permeation chromatography

- Buffer for GST-NPA: 50 mM Tris/HCl, pH 8.0, 150 mM NaCl, 1 mM TCEP
- Buffer PB2-His: 50 mM Tris/HCl, pH 8.0, 150 mM NaCl, 1 mM TCEP

Both types of protein fractions were purified on Superdex75 column connected to Äkta purifier system (GE Healthcare). Before each experiment, the column was equilibrated with the buffer and incubated at 16 °C. First and second elution fractions were mingled together and concentrated to 5 ml by using Amicon centrifugal filter units with appropriate cut-off size. The fraction was loaded on the column through 5 ml loop. The flow rate was set 0.3 ml/min and 2.5 ml fractions were collected using the automatic sampler. The chromatogram was recorded as the UV absorption at 280 nm. The composition and purity of the collected fractions was monitored by SDS-PAGE. The fractions were aliquoted and frozen at – 80 °C for later use.

3.5 SDS PAGE

- 16% stacking gel: 375 mM Tris/HCl, pH 8.8, 16% acrylamide, 0.1% (w/v) SDS, 0.2% (v/v) TEMED, 0.1 % (w/v) APS
- 6% resolving gel: 250 mM Tris/HCl, pH 6.8, 6.6% acrylamide, 0.1% (w/v) SDS, 0.2% (v/v) TEMED, 0.1% (w/v) APS
- Sample buffer: (6 × concentrate): 350 mM Tris/HCl, pH 6.8, 30% (v/v) glycerol, 350 mM SDS, 4% (v/v) 2 mM 2-mercaptoethanol, 180 µM bromphenol blue
- Electrode buffer: (5 × concentrate): 140 mM Tris/HCl, pH 8.8, 1.4 M glycine, 20 mM SDS

Staining:

- Staining buffer: 0.5% Coomassie Brilliant Blue (w/v), 50% methanol (v/v), 10% acetic acid (v/v)
- Washing buffer: 10% acetic acid

SDS polyacrylamide gel electrophoresis was used for monitoring the protein purification process. First, samples from each step of the purification process were denaturated by incubation with the sample buffer at 95˚C for 5 minutes. After cooling down, they were thoroughly mixed by vortexing and applied on the gel. All Blue Standard (Bio-Rad), a mixture of blue stained recombinant proteins with molecular mass range of 10-250 kDa, was used as a marker. The electrophoresis was performed in vertical chambers (Bio-Rad) filled with electrode buffer under constant voltage of 180 V for approximately 90 minutes, while cooled down with a frozen block. The gels were visualized with Coomassie brilliant blue staining buffer and the background was destained by repeated wash with washing buffer.

3.6 Surface Plasmon resonance (SPR)

- Protein buffer: 25 mM Tris/HCl, pH 7.4, 150 mM NaCl, 0.05% Tween 20, 10 mM $MgCl_2$, 1 mM $MnCl_2$, 1 mM 2-mercaptoethol

- Alkanethiole solution: ethanol mixture of HS-(CH2)11-PEG4-OH and HS-(CH2)11-PEG6-O-CH2-COOH alkanethiols
- NHS/EDC water solution: 100 mM N-hydroxysuccinimide, 400 mM 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride
- SA10 buffer: 10 mM sodium acetate, pH 5.0
- Ethanolamine (EA) 1 M solution
- PBS buffer: 10 mM Na2HPO4, 1,8 mM KH2PO4, pH 7.4, 500 mM NaCl, 2.7 mM KCl
- Neutravidin 20 mg/l in SA10 buffer

The surface plasmon resonance measurement was performed on a four-channel SPR sensor platform provided by the Institute of Photonics and Electronics, Prague. First, the SPR chip (also provided by the IPE) was loaded into a pure ethanol mixture of HS-(CH2)11-PEG4-OH and HS-(CH2)11-PEG6-O-CH2-COOH alkanethiols (molar ratio 7:3, Prochimia) with a final concentration of 0.2 mM and incubated for 30 min at 37 °C. The chip was rinsed with UV ethanol and deionized water and dried with flow of nitrogen. Then, the chip was mounted to the prism on the SPR sensor with a drop of immersion oil. The activation of carboxylic terminal groups on the sensor surface was performed by injecting a 1:1 mixture of 11.5 mg/mL N-hydroxysuccinimide (NHS) and 76.7 mg/mL 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride (EDC) in deionized water for 5 min at a flow rate of 20 μL/min. The rest of the experiment was performed at a flow rate of 30 μL/min. Afterwards, 1 ml of 0.02 mg/mL neutravidin solution in 10 mM sodium acetate buffer, pH 5.0 was loaded for 8 min. Then, high ionic strength solution (PBS with 0.5 M NaCl) was used to wash out noncovalently bound neutravidin molecules, followed by 1 M ethanolamine (Biacore) for deactivation of residual activated carboxylic groups.

Experiment was performed twice, with two types of biotinylated inhibitor L-742.001 (stated as C-580 and C-867 or generally referred to as probes) that differed in a length of a linker between biotin and inhibitor moieties. Probes were synthesized by Dr. Aleš Machara and Carlos Berenguer Albiñana, IOCB Prague.

The solution of 1 μM of probe in the buffer 25 mM Tris, pH 7.4, 150 mM NaCl, 0.05% Tween 20, 10 mM $MgCl_2$, 1 mM $MnCl_2$, 1 mM 2-mercaptoethol was

immobilized on prepared gold chip to saturate all biotin binding sites. Afterwards, various concentrations of GST-NPA protein (125 nM, 250 nM, 500 nM and 1000 nM) were injected for several minutes (association), and then buffer was injected alone (dissociation). Kinetic curves of binding were exported and subsequently fitted in TraceDrawer v.1.5 (Ridgeview Instruments AB) to obtain k_{on} and k_{off} parameters. Dissociation constant (K_d) was calculated as k_{off}/k_{on} ratio.

3.7 AlphaScreen

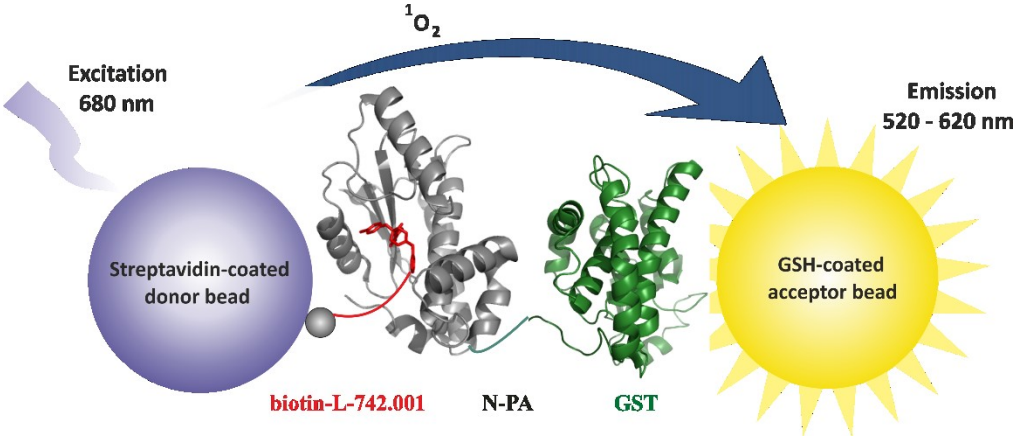


Figure 16. Schematic representation of the AlphaScreen experiment. The biotinylated derivative of L-742.001 is bound to the streptavidin-coated donor bead. GST-NPA protein is bound by GSH (reduced glutathione) covered acceptor bead. Two types of coated beads are mixed together and excited with light of a wavelength of 680 nm. This creates reactive oxygen species that diffuse up to 200 nm distance before the extinction. Once the L-742.001-based probe and GST-NPA interact, two types of beads are sufficiently close for the reactive oxygen species to interact with the reactive groups within the acceptor beads. This results in emission in 520-620 nm spectra. If the interaction is disrupted, decreased or no emission signal is detected.

- AlphaScreen buffer: 25 mM Tris/HCl, pH 7.4, 150 mM NaCl, 0.05% Tween 20, 1 mM 2-mercaptoethanol, 1 mM $MnCl_2$, 10 mM $MgCl_2$

For the AlphaScreen experiments, Perkin Elmer Enspire plate reader and 96-well ProxiPlates were used [155]. The buffer contained Mg^{2+} and Mn^{2+} ions, since they were proved to be indispensable for the endonuclease activity of PA subunit.

First, both GST-NPA and its biotinylated inhibitor L-742.001 (referred as probe) were bound to acceptor and donor beads, each mix was prepared separately. Streptavidin-

coated donor beads were used to bind biotinylated inhibitor, the fusion protein of NPA with glutathione-S-transferase (GST-NPA) was mixed with GSH acceptor beads. Each solution was incubated for 1 hour in the dark at room temperature, then the solutions were mingled together. The mixed solution was additionally incubated another 2 h upon constant mixing. For compound screening and IC_{50} determination, compounds were added prior to second round of incubation. Concentrations of donor and acceptor beads were 5 µg/ml, within final reaction volume of 50 µl.

To find the concentrations which gave the optimal signal to noise ratio, different concentrations of GST-NPA and biotinylated L-742.001 were tested in cross-titration arrangement. Moreover, two types of probe with different length of linker between L-742.001 and biotin moieties were tested.

The assay was evaluated by titrations with L-742.001 as a positive control and PB1-11 peptide as a negative control. Statistical evaluation was carried out as was described previously [156]. Positive signal (interaction between probe and NPA-GST) and background signal (beads without interaction partners) were measured in dodecaplicates and following statistical parameters were calculated:

Signal to noise ratio: $S/N = \frac{\mu_{c+} - \mu_{c-}}{\sigma_{c-}}$

Signal to background ratio: $S/B = \frac{\mu_{c+}}{\mu_{c-}}$

Z′-factor : $Z' = 1 - \frac{(3\sigma_{c+} + 3\sigma_{c-})}{\mu_{c+} - \mu_{c-}}$

Where:

- μ_{c+} is mean of maximum signal
- μ_{c-} is mean of background signal
- σ_{c+} is standard deviation of the maximal signal
- σ_{c-} is standard deviation of the background

Z-factor is a characteristic parameter for quality of a high-thoughput assay, that estimes whether a particular assay is suitable for high-throughput setup.

To determine IC_{50} values, titration experiments using the tested compounds were performed. The values were obtained by fitting of titration curves in GraFit software (Erithacus Software).

3.8 Determination of the endonuclease activity

3.8.1 *In vitro* endonuclease activity assay

First, endonuclease activity of purified N-terminal PA endonuclease domain in fusion with GST was verified by in vitro cleavage of DNA substrate visualized by agarose gel electrophoresis. Even though influenza endonuclease naturally cleaves RNA substrates, its non-specificity allowed to use more stable DNA substrate such as single-stranded DNA plasmid M13mp18 [111]. Each reaction contained 1 μM GST-NPA and was initiated by addition of 0.2 μg of M13mp18 plasmid. Inhibitors were tested in two concentrations – 10 μM and 100 μM.

Reactions were incubated at 37 °C for 2.5 h, stopped by adding of 1 μl of 0.2 M EDTA and visualized by agarose electrophoresis (section 3.2.5).

3.8.2. FRET (Förster resonance energy transfer)

- Reaction buffer 2x: 50 mM Tris/HCl, pH 7.4, 300 mM NaCl, 0.1% Tween 20, 2 mM 2-mercaptoethanol, 2 mM $MnCl_2$, 20 mM $MgCl_2$

Förster resonance energy transfer (FRET) was used to characterize the relative activity of the GST-NPA and the inhibition activity of raltegravir. Known endonuclease inhibitor L-742.001 was used as a positive control.

17 nt long DNA sequence with FAM fluorophore and BHQ1 quencher at 5' and 3' ends was chosen based on previous experiments conducted by Kowalinski et al. and Yuan et al. and ordered from GeneriBiotech [157, 158].

The reaction solutions contained 1.6 μM substrate and 1 μM GST-NPA, all reaction components were dissolved in the buffer containing 25 mM Tris/HCl, pH 7.4, 150 mM NaCl, 0.05% Tween 20, 1 mM 2-mercaptoethanol, 1 mM $MnCl_2$, 10 mM $MgCl_2$. Tested compounds were applied in 1 μM and 50 μM concentrations. Reactions were initiated by the addition of the substrate (DNA sequence with fluorophore and quencher).

The reactions were applied on the 96 well plate (ThermoFisher Scientific) and fluorescence emission was measured at 535 nm in 30 s intervals upon excitation at 485 nm by using multimode microplate reader Infinite M1000 PRO (Tecan). The measurement was carried out in duplicates.

Data over a period of 15 min were used to calculate initial velocities by linear approximation using Infinite M1000 PRO software.

3.9. Protein crystallization

- Crystallization buffer: 50 mM Tris/HCl, pH 8.0, 150 mM NaCl, 1 mM TCEP

For crystallization trials, the protein was concentrated to 10 mg/ml by centrifugation at 10,000 *g* at 10 °C in Amicon centrifugal filter units with MWCO 10 kDa. Two approaches to crystallization were undertaken: automated robotic screening of crystallization conditions and hands-on testing of previously known condition. The crystals obtained were the result of intersection of these two approaches. Automated screening resulted in finding of unreproducible condition with protein micro-crystals that were used to strike-seed the previously known condition that was otherwise not amenable to crystallization.

3.9.1 Automatized robotic screening

Automated screening was performed by using sitting drop vapor-diffusion method. First, 96-well plate (MRC 3 well low, Swissci) was filled with 30 µl of crystallization screen solution (either JCSG+ or Morpheus) by using the Crystal Gryphon robot (Art Robbins Instruments). Another robot, Oryx8 (Douglas Instruments) was used to mix condition and protein in a single drop. The plate filled with crystallization conditions were transferred to the Oryx8 robot and covered with an evaporation shield. The channels were degassed, and position of the tip was adjusted so the drop would be placed in the middle of each well. The protein (25 µl) was transferred into a PCR tube and loaded automatically. The ratio of the protein: condition was set up to 1:1 (200 nl : 200 nl). After the pipetting was finished, the plate was sealed with a transparent foil and put into Gallery DT plate hotel for automatic photo inspections in UV and visible light.

3.9.2 Hanging-drop vapour diffusion crystallization

Several previously published conditions were prepared and tested for crystallization of PA endonuclease by using hanging drop vapor-diffusion method. First, 500 µl of condition solutions were pipetted into EasyXtal 15-well plate (QIAGEN). Then, 1 µl of

the condition and 1 µl of protein (8-14 mg/ml) were mixed into a single drop on cover slips.

3.9.3 Seeding, soaking and data collection

In seeding experiments, a horse tail hair was put into a seeing solution and then strike over 2 µl drops (1:1 ratio of protein:solution). Crystal were left to grow for 2-3 weeks, then soaked with raltegravir. In soaking experiments, 0.2 µl of 10 mM raltegravir was added to a drop with pre-existing protein crystal. The final molar concentration of the ligand was at least 4 times greater than the concentration of the protein in the drop. The soaking continued for 40 hours. Then, the crystals were harvested and flash-frozen in liquid nitrogen. The diffraction data were collected using Rigaku diffractometer and on Bessy II beamline at synchrotron radiation center at Helmholtz Zentrum, Berlin (by Dr. Petr Pachl, IOCB Prague).

3.10. DIANA (DNA-linked Inhibitor Antibody Assay)

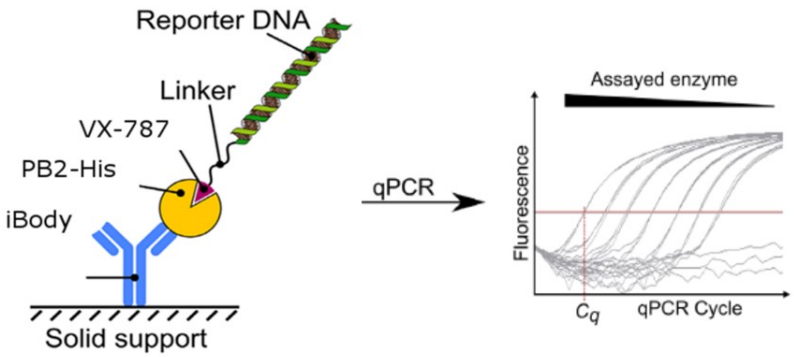


Figure 17. Schematic representation of DIANA for screening of potential inhibitors of PB2 cap binding. Adapted from [159]. PB2-His_6 is bound by anti-HisTag i-Body [160], whose biotin moiety, in turn, binds to neutravidin adsorbed to the solid support (not shown). The probe is based on known cap-binding inhibitor VX-787 and contains a reporter DNA for detection via qPCR. In a typical screening setup, probe is added to the well along with a tested compound and incubated to the point of establishing equilibrium. After the washing out of the excess of unbound compound/probe, the amount of the bound probe is detected by amplification with primers complementary to the oligonucleotide of the probe. Displacement of the probe by a tested compound shifts fluorescent signal production towards later PCR cycles, compared to the setup

without the tested compound added. C_q (quantitation cycle) values that are calculated from the fluorescent signal curves are further used for determination of dissociation constants (K_d).

3.10.1. Probe preparation

The DIANA probe based on VX-787 with oligonucleotide was prepared by azide-alkyne click reaction. VX 787 with a linker and azide moiety (further referred as C1015) was synthesized by Carlos Berenguer Albiñana, IOCB Prague.

The oligonucleotide with alkyne moiety was ordered from GeneriBiotech (Czech Republic). BTTP ligand was provided by Milan Vrabel, IOCB of the CAS. First, two solutions were prepared in separate microtubes: a catalysis solution (2.25 µl) and reaction solution. The catalysis solution contained 5 mM BTTP, 2.5 mM $CuSO_4$ and 5 mM sodium ascorbate mixed in 100 mM HEPES buffer pH 7.0. The reaction solution contained 200 µM oligonucleotide-alkyn, 1 mM VX-787 with azide, mixed in 100 mM HEPES buffer pH 7.0.

The reaction and catalysis solutions were mixed to the final volume of 20 µl and then incubated at 30 °C for 2 h. To remove the excess of unreacted probe precursor, the solution was transferred into an Amicon centrifugal filter unit (MWCO 10 kDa) and washed with TBS buffer at least five times (13,000 *g*, 8 min), to the final volume of 50 µl. The concentration of the probe was measured on Nanodrop spectrophotometer (Thermo Fisher Scientific). LC-MS analysis was carried out to determine the quality of the probe (by Radko Souček, IOCB Prague).

3.10.2. Determination of K_d values of the C-1015 probe, VX-787 and m^7GTP

- Casein solution: 5-fold diluted casein buffer (SDT)
- Buffer TBS: 20 mM Tris/HCl, pH 7.5, 150 mM NaCl
- Buffer TBST: 20 mM Tris/HCl, pH 7.5, 150 mM NaCl, 0.05% Tween 20 (w/v)
- Buffer TBST´: 20 mM Tris/HCl, pH 7.5, 150 mM NaCl, 0.1% Tween 20 (w/v)

The day before the experiment, 5 ul of neutravidin solution (10 ng/µl in TBS buffer) were added to each well of 96-well plate (4titude). After 1 h of incubation at room temperature, 100 ul of casein solution was added to block surface unoccupied by neutravidin. The plate was covered with a foil and left overnight at room temperature.

Then, the excess of casein was removed, and the plate was washed with TBST buffer in automatic Blue Washer (Blue Cat Bio).

Then, 5 ul of 100 nM anti-HisTag-iBody/50 μM $NiCl_2$ dissolved in TBST' buffer was added, and the plate was incubated for 1 h. The plate was then washed again with TBST buffer in Blue Washer machine. As the next step, 5 μl of PB2-His_6 protein (0.02 ng/μl) were added to each well and the plate was left incubating to another 2 h at room temperature and washed again. Then, the C-1015 probe was added to the final concentration of 100 pM. After each step of plate preparation, plate was centrifugated at 2,000 *g* for 5 min at room temperature.

To determine the K_d of C-1015, inhibitor VX-787 and m^7GTP, series of 3.16x dilutions were prepared, so the final setups were:

1. neutravidin / 100 nM U-35 iBody / casein / 100 pg PB2-His_6 / 100 pM C1015 probe + C1015 (50 uM – 5 nM concentration series)
2. neutravidin / 100 nM U-35 iBody / casein / 100 pg PB2-His_6 / 100 pM C1015 probe + VX-787 (50 uM – 5 nM concentration series)
3. neutravidin / 100 nM U-35 iBody / casein / 100 pg PB2-His_6 / 100 pM C1015 probe + m^7GTP (1 mM – 318 nM concentration series)

The setup also included reactions with no tested compounds, as well as controls without probe and/or protein to test non-specificity. The plate was subsequently incubated for 1 h and washed again. Then, 5 μl of qPCR mix (composition is described in [159]) was added to each well, the plate was sealed with a foil and transferred to the LightCycler device (Roche) to detect real-time PCR curves using the standard protocol for DIANA experiments [159].

The data analysis was carried out using LightCycler 480 II software. C_q parameter was calculated for each curve by finding the maxima of the second derivative. Concentrations of the unbound probe were fitted by the following function to obtain K_d values:

$$[EP] = \frac{E_{tot} * P_{tot}}{K_d + P_{tot}}$$

Where:

- [EP] is the amount of bound probe
- P_{tot} is the total concentration of the probe
- E_{tot} is the amount of captured enzyme
- K_d is the dissociation constant

Dissociation constant of the probe was used to calculate K_d values of C-1015 probe precursor, VX-787 and m^7GTP using the equation:

$$K_{d\ ligand} = \frac{2^{-\Delta C_q}}{1 - 2^{-\Delta C_q}} \times \frac{L_{tot}}{1 + (\frac{P_{tot}}{K_{d\ probe}})}$$

Where:

- P_{tot} is the total concentration of the probe
- L_{tot} is the ligand concentration
- $K_{d\ probe}$ is the dissociation constant of the probe
- ΔC_q is difference in threshold cycle values in wells incubated in the presence and absence of the ligand

4. Results

4.1. Cloning experiments

Cloning experiments resulted in the insertion of gene fragments encoding N-terminal PA endonuclease domain and PB2 cap-binding domain into corresponding plasmids for expression of fusion protein constructs.

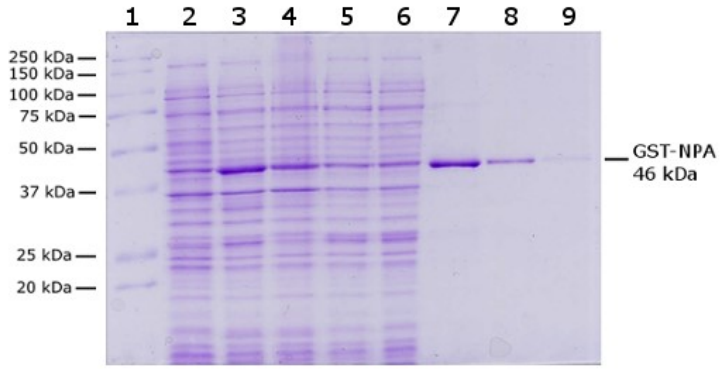
Gene fragment that encodes N-terminal PA endonuclease domain was ligated into pGEX-1λT plasmid for generation of GST-NPA fusion protein. The construct also contained sequence encoding the cleavage site for Prescission protease, as tag-free protein was planned to be utilized in crystallization trials. The cleavage, however, was found inefficient (data not shown) and instead His_6-SUMO-NPA construct was designed for generation of tag-free NPA protein. His_6-SUMO-NPA construct was prepared by cloning the PA gene fragment into pETM11-SUMO3 plasmid. Prior to ligation into the corresponding plasmid, the gene fragment was modified by the addition of sequence encoding 4xGS linker through PCR primer extension. This was aimed to space out SUMO-tag and NPA protein for more efficient cleavage by ULP1 protease. PB2 cap-binding domain encoding gene fragment modified by PRC primer extension to contain Prescission protease cleavage site was ligated into pET-24a plasmid.

All expression plasmids containing gene constructs were sequenced by Sanger sequencing and no mutations were detected within the gene fragments. These expression plasmids were subsequently transformed into *E. coli* BL21(DE3) RIL for recombinant protein expression.

4.2. Purification of recombinant proteins

4.2.1 GST-NPA

N-terminal endonuclease domain of polymerase acidic protein from influenza strain A/California/07/2009 (H1N1) fused to GST-tag was expressed in *E. coli* and purified using affinity chromatography on GSH-sepharose. As can be observed on SDS-PAGE gel that documented the purification process, protein was not produced in bacterial cells at high rates and some portion of overexpressed protein remained in the insoluble fraction (Figure 18, lanes 2-3).



Lane	V_{load}	Fraction
1	1 µl	All Blue marker
2	5 µl	before induction
3	1.5 µl	after induction
4	0.5 µl	soluble proteins
5	0.5 µl	unsoluble proteins
6	0.5 µl	flow-through
7	6 µl	1. elution
8	6 µl	2. elution
9	6 µl	3. elution

Figure 18. SDS-PAGE gel documenting purification of GST-NPA fusion protein. For experimental details see Methods section 3.5, page 54.

The elution fractions were pooled together, concentrated and additionally purified by gel permeation chromatography to separate possible multimers and minor impurities.

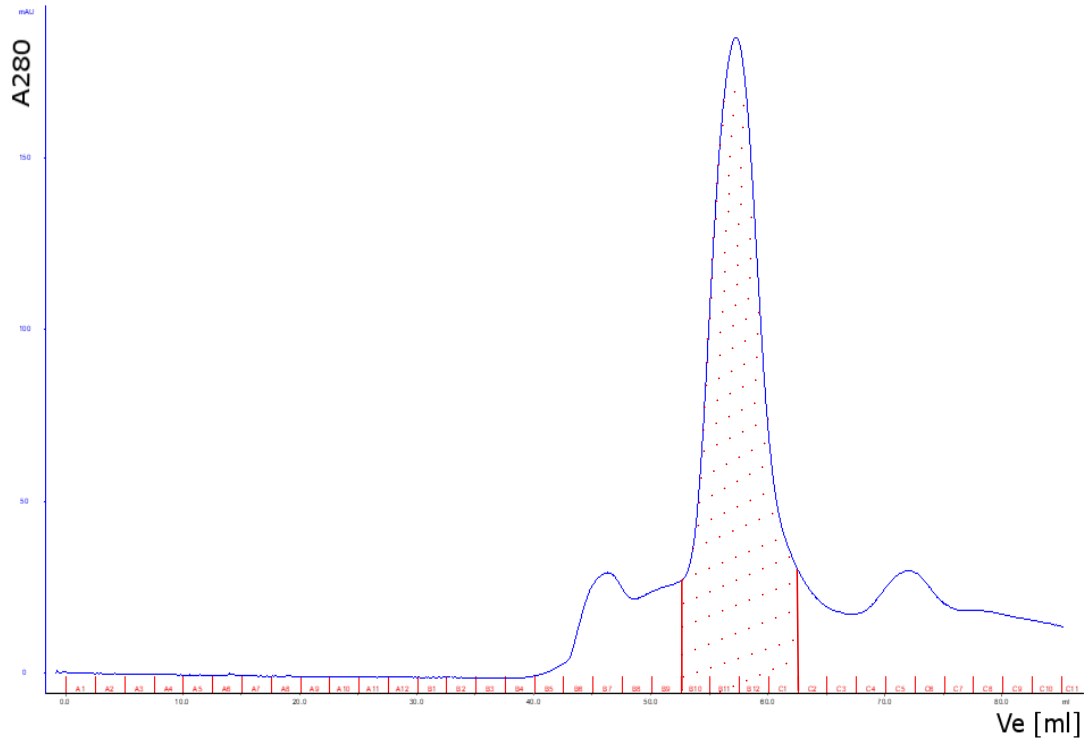
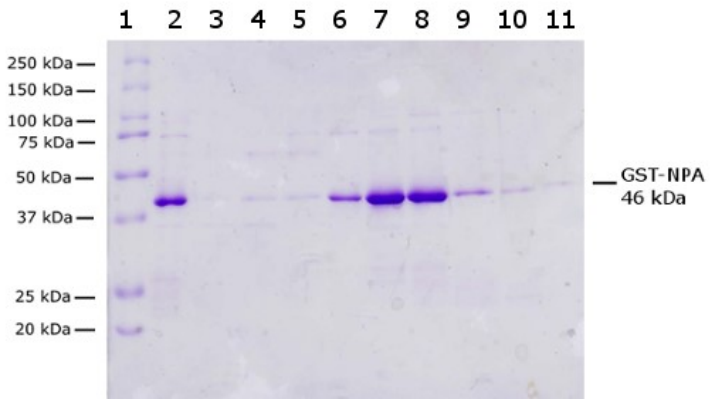


Figure 19. Chromatogram of GPC purification for GST-NPA. The part dashed with red corresponds to the collected fractions. For experimental details see Methods, section 3.4.3, page 53.

As it is documented on SDS-PAGE gel in Figure 20, series of impurities remained after GPC purification, however, the final purity of the protein fractions was sufficient for SPR and AlphaScreen experiments. Fractions B10-C1 (V_e 52.5-65 ml) that corresponded to the highest elution peak were collected, concentrated and used in further experiments. The final yield from 3 l media was 420 μg of highly pure protein (Figure 20, lines 6-9).

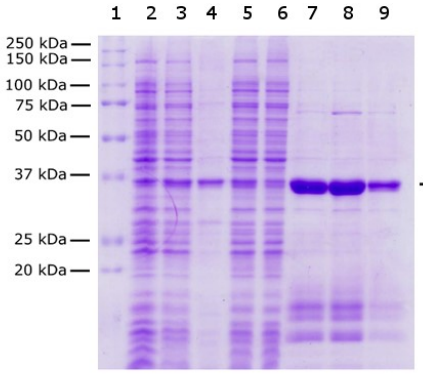


Lane	V_{load}	Fraction
1	1 μl	All Blue marker
2	4 μl	before GPC
3	12 μl	fraction B6
4	12 μl	fraction B7
5	12 μl	fraction B8
6	12 μl	fraction B10
7	12 μl	fraction B12
8	12 μl	fraction B12
9	12 μl	fraction C1
10	12 μl	fraction C2
11	12 μl	fraction C3

Figure 20. SDS-PAGE gel documenting the gel permeation chromatography purification of GST-NPA. For experimental details see Methods, section 3.5, page 54.

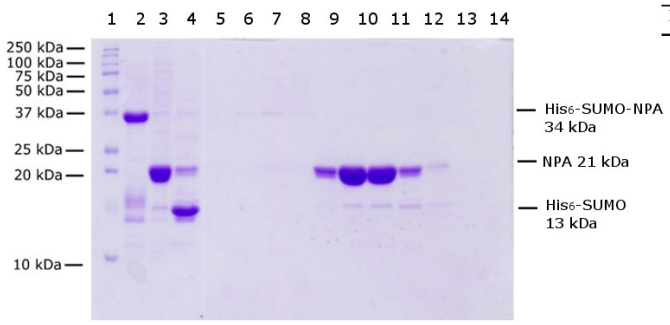
4.2.2 His_6-SUMO-NPA

The N-terminal endonuclease domain of polymerase acidic protein from the influenza strain A/California/07/2009 (H1N1) fused with SUMO solubility tag and His_6 affinity tag was expressed in *E. coli* and purified using chelation chromatography on Ni-NTA resin. The protein was further cleaved with ULP1 protease and additionally purified using gel permeation chromatography. The expression and purification processes are depicted in Figure 21. The protein was produced at moderate level of expression and was successfully purified by chelation chromatography on Ni-NTA resin. Some portion of overexpressed His_6-SUMO-NPA remained in the insoluble fraction (Figure 21, lane 4). The elution fractions were pooled together and divided into two fractions for step-wise cleavage with ULP1 protease and purification using gel permeation chromatography on Superdex75 column.



Lane	V_{load}	Fraction
1	1 µl	All Blue marker
2	5 µl	before induction
3	3 µl	after induction
4	0.75 µl	insoluble proteins
5	0.75 µl	soluble proteins
6	0.75 µl	flow-through
7	6 µl	1. eluton
8	6 µl	2. elution
9	6 µl	3. elution

Figure 21. SDS-PAGE gel documenting purification of His_6-SUMO-NPA fusion protein. For experimental details see Methods, section 3.5, page 54.



Lane	V_{load}	Fraction
1	1 µl	All Blue marker
2	4 µl	fraction before cleavage with ULP protease
3	4 µl	fraction after cleavage with ULP protease
4	8 µl	elution after cleavage
5	12 µl	fraction B5
6	12 µl	fraction B6
7	12 µl	fraction B7
8	12 µl	fraction B8
9	12 µl	fraction C7
10	12 µl	fraction C8
11	12 µl	fraction C9
12	12 µl	fraction C10
13	12 µl	fraction C11
14	12 µl	fraction D10

Figure 22. SDS-PAGE gel documenting the cleavage of His_6-SUMO-NPA fusion protein with ULP1 protease and subsequent GPC purification. Detailed description is provided in the text. For experimental details see Methods, section 3.5, page 54.

Cleavage with ULP1 protease is documented in Figure 23. Lane 2 corresponds to the non-cleaved protein, lane 3 depicts the fraction after the cleavage and purification on Ni-NTA agarose to remove the cleavage product – His_6-SUMO-tag and noncleaved His_6-SUMO-NPA (lane 4). The cleavage was efficient, with only a small portion of His_6-SUMO-NPA remaining noncleaved. Fraction containing tag-free NPA protein was then concentrated and applied to the Superdex75 column. High-molecular-weight impurities (peak at V_e 40-50 ml) that probably correspond to aggregates of His_6-SUMO-NPA (Figure 23, lanes 6-7) along with impurities of lower molecular weight (peaks at V_e 100-

130 ml) were separated from the main NPA fraction. Lanes 9-12 in Figure 23 document the purity of the final NPA fraction that was used in crystallization. The total yield from 3 l of LB media after two-step purification and tag cleavage was 3 mg of pure NPA protein.

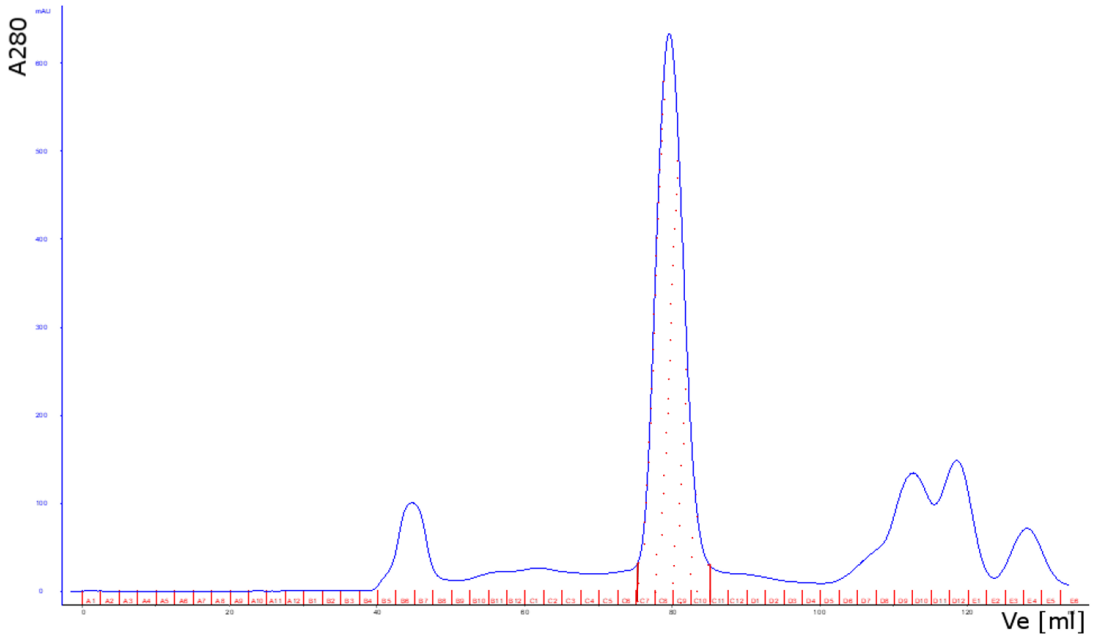
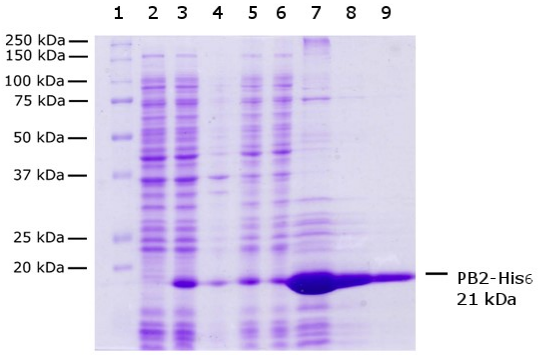


Figure 23. Chromatogram of GPC purification for His_6-SUMO-NPA. The part dashed with red corresponds to the collected fractions. For experimental details see Methods, section 3.4.3, page 53.

4.2.3 PB2-His_6

Cap-binding domain of PB2 protein from influenza virus (strain A/California/07/2009, subtype H1N1) was successfully purified from *Escherichia coli* using chelation chromatography on the Ni-NTA-agarose resin. As documented on the SDS-PAGE gel (Figure 24, lane 4), small portion of overexpressed protein was observed in non-soluble fraction. Also, after one round of Ni-NTA chromatography, flow-through fraction still contained substantial amount of the protein and thus, additional round of re-purification could be performed (Figure 24, lane 6). The calculated molecular weight of the expressed fusion protein is 21 kDa, however, it appears at lower molecular weight values on SDS-PAGE gel. This abnormality could be explained by the amino acid composition of the protein, which contains a large proportion of positively charged amino acids, as this was previously observed to influence migration of proteins in SDS-PAGE [161].



Lane	V_{load}	Fraction
1	1 µl	All Blue marker
2	5 µl	before induction
3	3 µl	after induction
4	0.5 µl	insoluble proteins
5	0.5 µl	soluble proteins
6	0.5 µl	flow-through
7	5 µl	1. eluton
8	5 µl	2. elution
9	5 µl	3. elution

Figure 24. SDS-PAGE gel documenting purification of PB2-His_6 fusion protein. For experimental details see Methods, section 3.5, page 54.

The crude yield from 3 l of LB media was 11 mg. After additional purification using the gel permeation chromatography the pure yield was 6 mg. Fractions D2-D5 that corresponded to the main elution peak on GPC chromatogram (V_e 70-76 ml). The purity of the final fraction was sufficient for use in DIANA experiments (Figure 26, lanes 4-7).

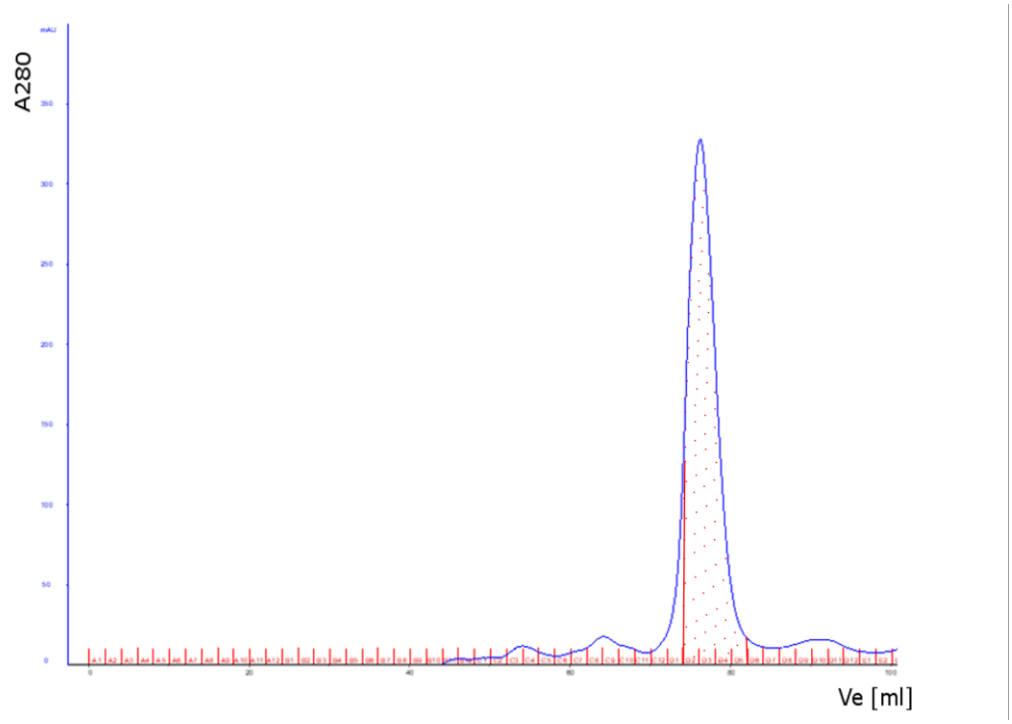
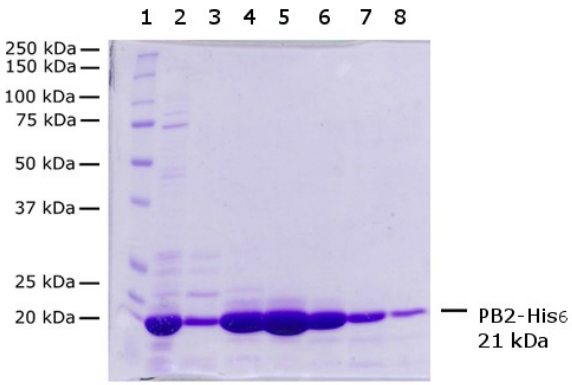


Figure 25. Chromatogram of GPC purification for PB2-His_6. The part dashed with red corresponds to the collected fractions. For experimental details see Methods, section 3.4.3, page 53.



Lane	V_{load}	Fraction
1	1 µl	All Blue marker
2	3 µl	before GPC
3	15 µl	fraction D1
4	15 µl	fraction D2
5	15 µl	fraction D3
6	15 µl	fraction D4
7	15 µl	fraction D5
8	15 µl	fraction D6

Figure 26. SDS-PAGE gel documenting GPC purification of PB2-His_6. The part dashed with red corresponds to the collected fractions. For experimental details see Methods, section 3.5, page 54.

4.3 Characterization of probes based on L-742.001 endonuclease inhibitor by SPR

SPR analysis was performed to test the binding parameters (K_d, k_{cat} and k_{off}) of the L-742.001-based probes prior to their usage in AlphaScreen experiments. Two types of probes were tested (further referred to as C-850 and C-867) that differed in in the length of the linker between the inhibitor and biotin moieties. Both probes were prepared by Dr. Aleš Machara and Carlos Berenguer Albiñana, IOCB Prague.

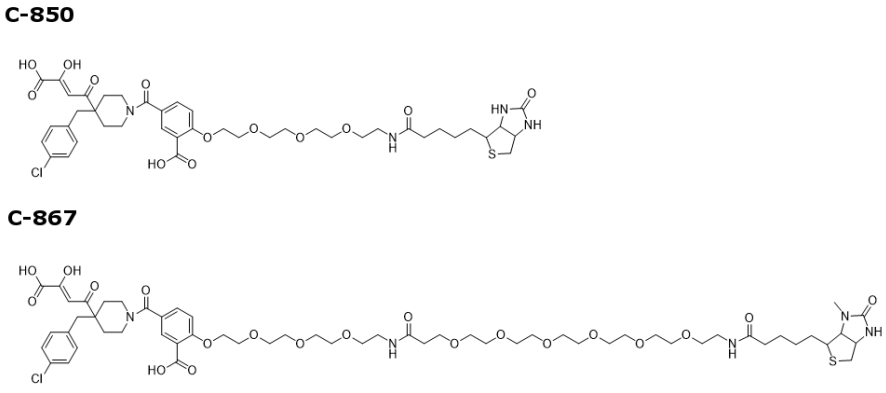


Figure 27. L-742.001-based biotinylated probes prepared for usage in AlphaScreen experiments.

First, C-850 probe with the shorter linker was tested. The setup of SPR experiment and resulting curve are shown in Figure 28. The biotinylated probe was coated to the chip through the interaction with neutravidin. GST-NPA protein of different concentrations (1000 nM, 500 nM, 250 nM, 125 nM) was injected into corresponding channels.

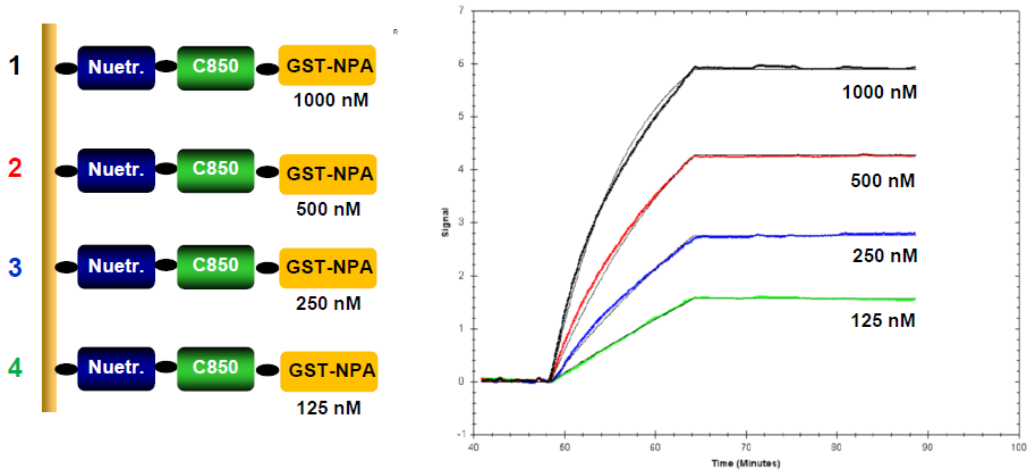


Figure 28. SPR experiment aimed to test binding of C-850 probe to GST-NPA. The left panel shows the experimental setup in individual channels. The right panel shows the sensorgram displaying association and dissociation phases fitted with one-to-one model (TraceDrawer v.1.5 software). Each curve corresponds to a certain GST-NPA concentration displayed in the setup on the left. R unit on the vertical axis corresponds to the accumulation of 1 pg of mass per mm^2. For experimental details see Methods, section 3.6, page 54.

Binding of C-867 probe with longer linker was tested in a similar experimental setup displayed in Figure 29.

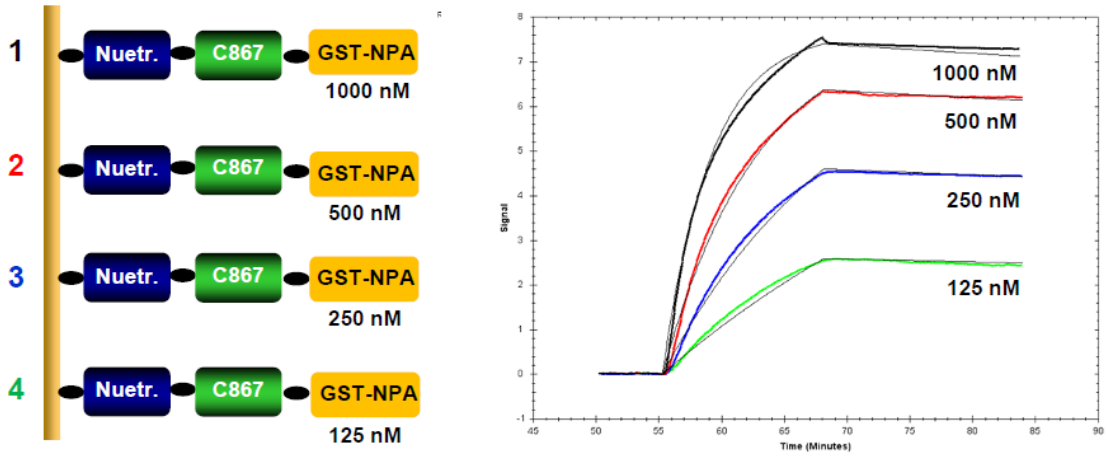


Figure 29. SPR experiment aimed to test binding of C-867 probe to GST-NPA. The left panel shows the experimental setup in individual channels. The right panel shows the sensorgram displaying association and dissociation phases fitted with one-to-one model (TraceDrawer v.1.5 software). Each curve corresponds to a certain GST-NPA concentration displayed in the setup on the left. R unit on the vertical axis corresponds to the accumulation of 1 pg of mass per mm^2. For experimental details see Methods, section 3.6, page 54.

Both experiments were evaluated in the TraceDrawer v.1.5 software and k_{cat} and k_{off} parameters of both interactions were obtained by fitting the experimental curved with 1:1 model of binding. K_d of each probe was calculated as k_{off}/k_{on} ratio (Table 6). Off-rate constants were, however, below the detection limit ($< 2 \cdot 10^{-5}$ s^{-1}).

Table 6. Kinetic parameters of binding between C-850 and C-867 probes and GST-NPA protein.

	C-850	**C-867**
k_{off}	$< 2 \cdot 10^{-5}$ s^{-1}	$< 2 \cdot 10^{-5}$ s^{-1}
k_{on}	$1.9 \cdot 10^{3}$ M^{-1}.s^{-1}	$4.3 \cdot 10^{3}$ M^{-1}.s^{-1}
K_d	< 10.6 nM	< 9.3 nM

4.4 Optimization and evaluation of AlphaScreen assay for use in screening for potential endonuclease inhibitors

First step in the development of high-throughput assay based on AlphaScreen technology was to find concentrations of both GST-NPA and probe that would give the optimal signal. This was achieved by testing concentrations of the probe and GST-NPA in cross-titration arrangement. The experiment was performed separately for each type of probe (C-850 and C-867, Figure 27) to test whether the length of the linker has influence on experiment's outcome. Indeed, in cross-titration experiment in which the probe with a shorter linker was used (C-850), the signal was significantly decreased and did not exceed 35,000 alpha counts.

This could be caused by a steric hindrance, as C-850 probe probably did not have sufficient length to interact with GST-NPA bound to GSH-beads. C-850 probe, therefore, was not suitable for further use in AlphaScreen experiments. Cross-titration experiment using C-867 probe, on the other hand, showed that probe could provide sufficient signal and could be used in screening. The concentrations of C-867 and GST-NPA that provide the signal close to the maximum were chosen for further experiments: 30 nM for C-850 and 75 nM for GST-NPA, with concentration of beads 5 μg/ml.

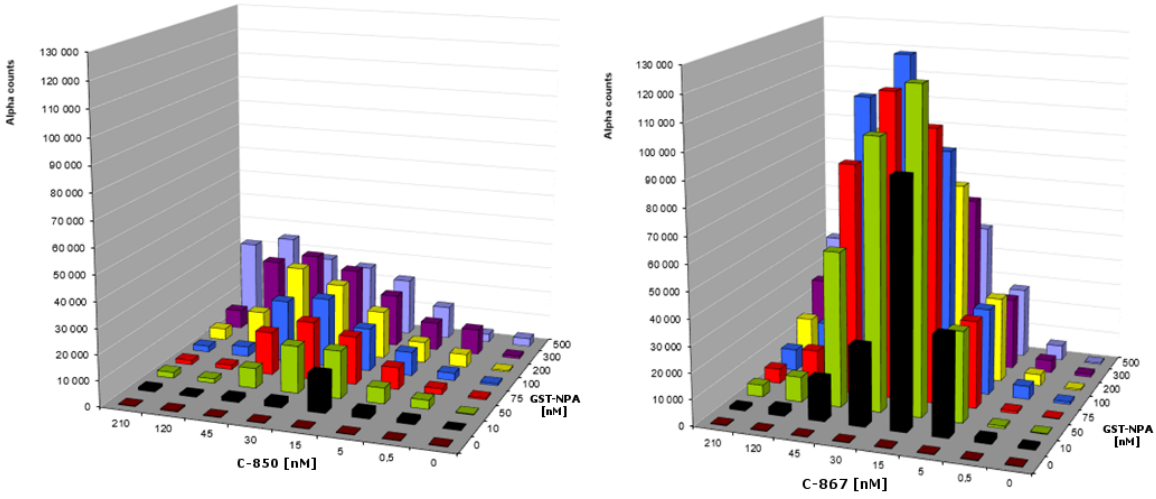


Figure 30. Cross-titration experiments using C-850 probe with a shorter linker (left) and C-867 probe with a longer linker (right). Luminescent signal (in alpha-counts) is shown on the vertical axis, two horizontal axes correspond to the concentrations of GST-NPA protein and biotinylated probes.

Statistical parameters that characterize suitability of an assay for high-throughput screening were calculated as described in section 3.7 and are summarized in Table 7.

Table 7. Statistical evaluation of the AlphaScreen assay.

Parameter	Value
Signal-to-noise ration S/N	2500
Signal-to-background ration S/B	700
Z'factor	0.87

Additionally, the assay was evaluated by testing competition between the probe and known inhibitor L-742.001 (structure shown in Figure 31) and a compound that presumably did not have any inhibitory activity (PB1-11 peptide, the sequence is shown in Figure 31).

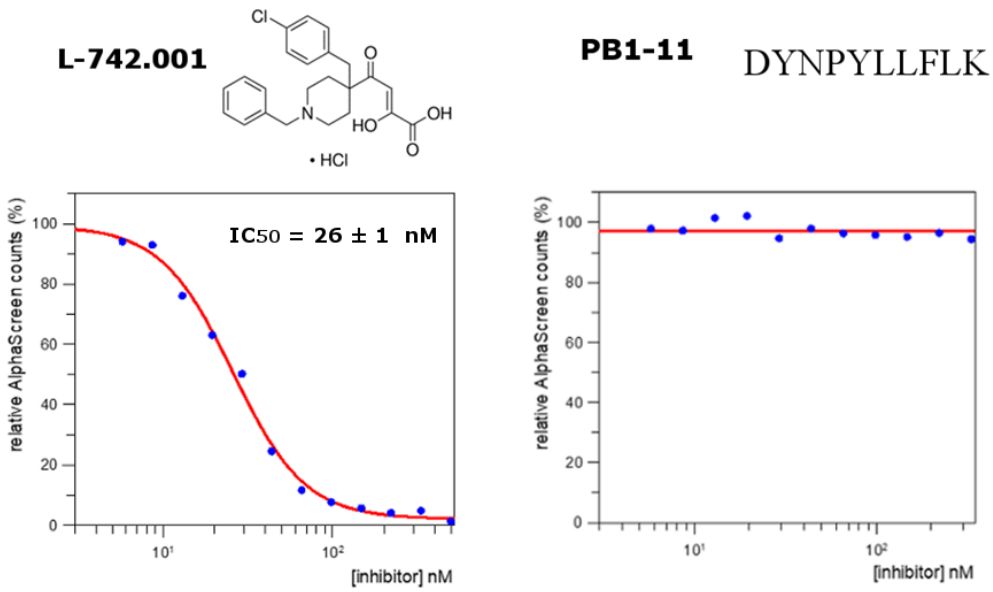


Figure 31. Evaluation of AlphaScreen assay. Left panel documents titration with known inhibitor L-742.001 (positive control), right panel shows titration with PB1-11 peptide with no inhibition activity (negative control). Experimental details are given in Methods, section 3.7, page 56.

Then, a small library of N-benzotriazoylthiourea-derived compounds was screened using the developed and evaluated AlphaScreen setup. The library was prepared and kindly provided by Dr. Aleš Machara, IOCB Prague. Compounds were pre-selected based on the potential metal-chelation properties. Also, guided by numerous allusions to

potential inhibitory activity of HIV integrase inhibitors and similarity of PA endonuclease and HIV integrase active sites, we added two commercially available HIV integrase inhibitors raltegravir and elvitegravir to the test library.

A total amount of 32 compounds was tested. Among N-benzotriazoylthiourea derivatives, no significant hits were identified. As was predicted earlier in literature, HIV integrase inhibitor raltegravir was found to be the best hit from the screening with later determined IC_{50} of 31 ± 3 μM (shown in Figure 32). Elvitegravir, however, did not display inhibitory activity.

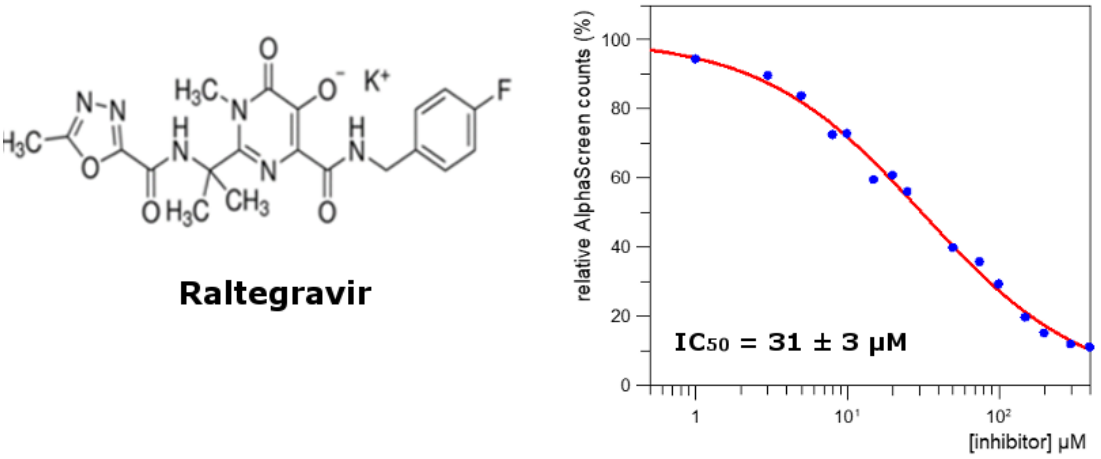
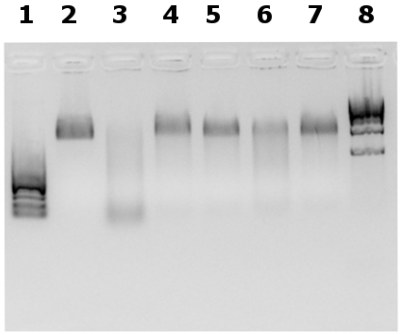


Figure 31. Chemical structure and IC_{50} titration curve for raltegravir. Experimental details are provided in the Methods section 3.7, page 56.

4.5. Evaluation of endonuclease activity of GST-NPA

Endonuclease activity of purified GST-NPA was tested by *in vitro* cleavage of a single stranded plasmid M13mp18. Reactions (listed in Figure 32, right panel) were incubated for 2.5 h in 37 °C and visualized by agarose electrophoresis. As it is shown in Figure 32 (left panel), endonuclease activity of GST-NPA remained intact after the purification process. Also, endonuclease activity was tested in presence of raltegravir and L-742.001 as a positive control. Raltegravir was found to inhibit endonuclease activity of GST-NPA in a concentration-dependent manner, with complete inhibition of cleavage at higher

concentration (100 μM) and moderate inhibitory effect at lower concentration (10 μM).



Lane	Description
1	100 b marker
2	0.2 μg DNA
3	0.2 μg DNA, 1 μM GST-NPA
4	0.2 μg DNA, 1 μM GST-NPA, 10 μM L-742.001
5	0.2 μg DNA, 1 μM GST-NPA, 100 μM L-742.001
6	0.2 μg DNA, 1 μM GST-NPA, 10 μM raltegravr
7	0.2 μg DNA, 1 μM GST-NPA, 100 μM ralegravir
8	1 kb marker

Figure 32. Evaluation of endonuclease activity of GST-NPA by in vitro cleavage of single stranded DNA plasmid. Description of reactions is provided on the right panel and in the text. For experimental details, see Methods section 3.8.1, page 58.

Endonuclease activity of GST-NPA in presence of inhibitors was also evaluated by FRET-based assay. Percentage of endonuclease activity was calculated from the initial velocities in the presence of an inhibitor relative to the initial velocity of the uninhibited enzyme (in the presence of DMSO). In line with the previous experiment, raltegravir inhibited endonuclease activity of GST-NPA in concentration-depended manner. Application of 1 μM raltegravir resulted in 10 % reduction of enzymatic activity, while in presence of 50 μM raltegravir enzymatic activity was reduced by 87 %. The positive control, L-742.001 completely abolished the endonuclease activity of GST-NPA at both concentrations.

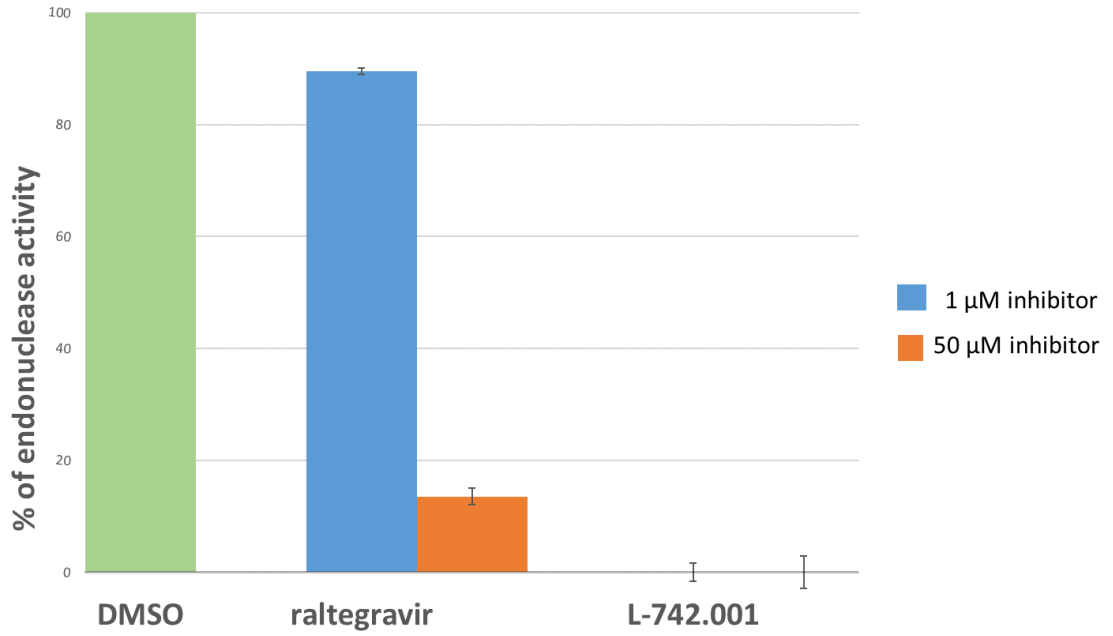


Figure 33. Diagram showing relative endonuclease activity of GST-NPA when uninhibited and upon the presence of 1 µM and 50 µM raltegravir. L-742.001 inhibitor was used as a positive control at the same concentrations. DNA probe labelled with FAM fluorophore and BHQ1 quencher was used to monitor the endonuclease activity of GST-NPA. Increase of fluorescent signal was monitored using Tecan plate reader and initial velocities were calculated using linear approximation. The relative endonuclease activity is represented as ratio of velocities of inhibited endonuclease cleavage and without inhibition (in presence of DMSO). Experimental details are provided in section 3.8.2, page 58.

4.6. Protein crystallization of endonuclease domain

Crystallization trials were set up to obtain protein crystals of N-terminal PA endonuclease domain (NPA) that would be subsequently soaked with raltegravir to determine the structure of the complex. Robotic screening of commercially viable matrix screens JCSG+ (Jena Biosciences) and Morpheus (Molecular Dimensions) resulted in finding a condition yielding microcrystals of NPA that were also visible under UV light (Figure 34). The crystallization condition contained 0.2 M NaCl, 2 M ammonium sulphate, 0.1 M sodium cacodylate, pH 6.5, and when prepared in the laboratory, did not reproduce the result from the screening. The microcrystals obtained from the screening were, however, used to strike-seed a condition, that was previously used for crystallization of influenza endonuclease domain by Song et al.: 0.2 M $MgCl_2$, 5 mM

$MnCl_2$, 0.1 M Tris, pH 8.5, 30% w/v PEG4000 [140]. The condition did not yield protein crystals without seeding.

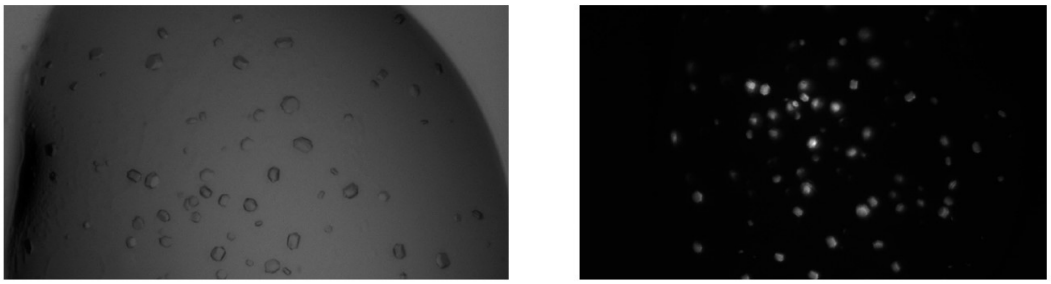


Figure 34. Photographs of NPA microcrystals grown in 0.2 M NaCl, 2 M ammonium sulphate, 0.1 M sodium cacodylate, pH 6.5 (JCSG+ screen) in VIS (left) and UV (right) light.

After strike-seeding with non-diluted stock solution, first crystals appeared the next day and were left to grow for another 7-10 days. Then the crystals were soaked with raltegravir for 40 h. Subsequently, crystals were harvested, flash-frozen and tested on X-ray home source (Rigaku) by Dr. Petr Pachl, IOCB Prague. Prior to freezing, crystals were transferred to cryo-cooling condition (the same as crystallization condition with 20% ethylene glycol (v/v)). The crystals, however, survived the freezing poorly and yielded low-resolution diffraction data upon testing. This was overcome by growing crystals in the crystallization solution containing 20 % ethylene glycol. The best crystals diffracted to 2.2 Å and 1.9 Å (the former crystal and its diffraction map are shown in Figure 35). The later crystal was measured at Bessy II beamline at synchrotron radiation center at Helmholtz Zentrum, Berlin by Dr. Petr Pachl, IOCB Prague. The model was built by Dr. Petr Pachl using molecular replacement. However, no electron density corresponding to the inhibitor was observed in the active site.

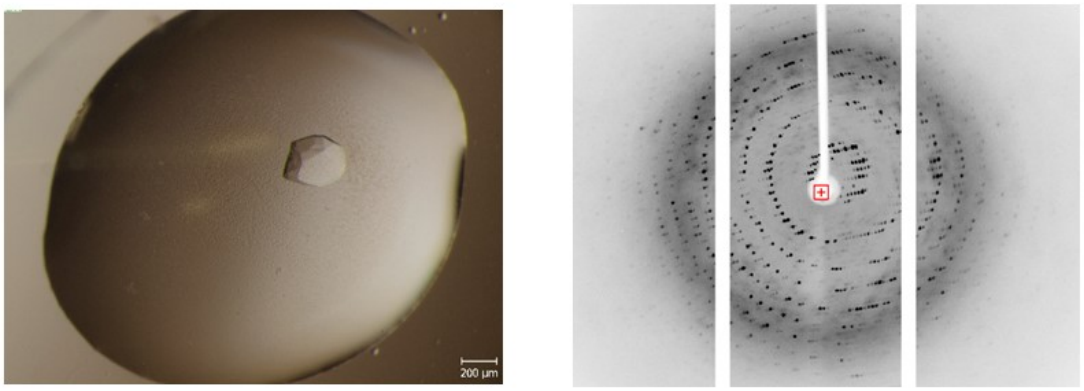
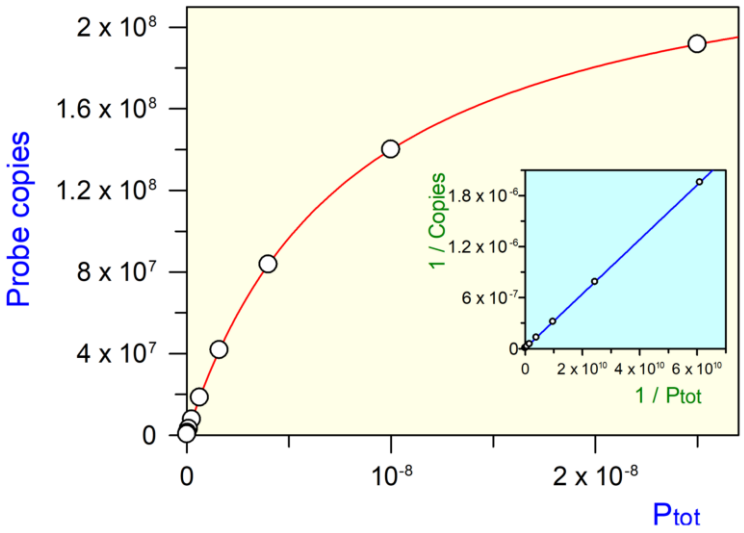


Figure 35. Crystal of NPA prior to soaking with raltegravir and corresponding diffraction map. The crystal diffracted up to 2.2 Å.

4.7. DIANA optimization for use in screening for potential disruptors of influenza polymerase cap binding activity

High-throughput assay for screening of protentional disrupters of interaction between PB2 cap-binding domain and 7-methylguanosine cap was developed based on DIANA method. In this setup, compounds with potential cap-binding activity would be identified by quantification of their binding by real-time PCR, using the probe with DNA oligonucleotide. The probe was designed based on known cap-binding inhibitor VX-787. The probe was prepared by copper-click reaction of VX-787 probe precursor (referred to C-1015, Figure 37) that contained azide moiety with alkylated oligonucleotide.

First, series of experiments were performed to find optimal conditions for the assay, such as the concentration of PB2-His_6 and assay buffer composition. Particularly, titration with PB2-His_6 was performed in reducing (in presence of 5 mM 2-mercaptoethanol) and non-reducing conditions and observed decrease in ΔC_q values upon addition of 5 mM 2-mercaptoethanol (experiments were conducted by Dr. Milan Kožíšek). Another important step in the optimization od DIANA experiment was determination of probe dissociation constant (K_d) by titration of PB2-His_6 protein with the C-1015 probe (Figure 36).



Parameter	Value [M]	Std. Error
K_d	8.13×10^{-9}	4.87×10^{-17}

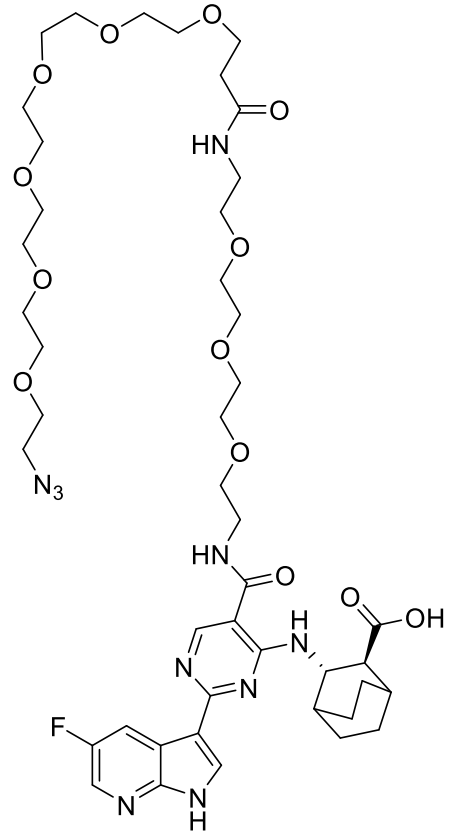
Figure 36. Titration of PB2-His$_6$ protein with C-1015 DIANA probe for K_d determination. The concentration of PB2-His$_6$ protein was kept constant (100 pg per well) and C-1015 DIANA probe of increasing concentrations (0-25 nM range, 2.5x dilution step) was added to corresponding wells. Experimental details are provided in Methods, section 3.10, page 60. The curve was fitted using GrafFit software. Y-axis corresponds to the amount of probe copies calculated from C_q values based on previous calibration experiment. X-axis corresponds to molar probe concentration. Graph in blue is the double reciprocal plot.

Additionally, to verify that modification of VX-787 inhibitor with the linker that connects inhibitor to DNA oligonucleotide did not have an effect on its binding properties, we performed the titration of PB2-His$_6$ protein with C-1015 probe precursor (without the DNA oligonucleotide, shown in Figure 37, left panel) and V-787, unmodified inhibitor.

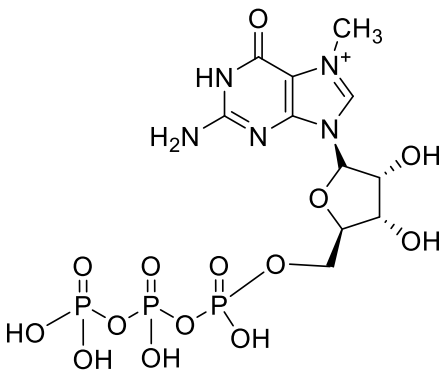
Finally, we used DIANA to determine K_d of interaction between PB2-His$_6$ and 7-methylguanosine triphosphate, the natural ligand of PB2 cap-binding domain.

Table 8. Dissociation constants of C-1015 probe precursor, VX-787 and m7GTP determined by DIANA.

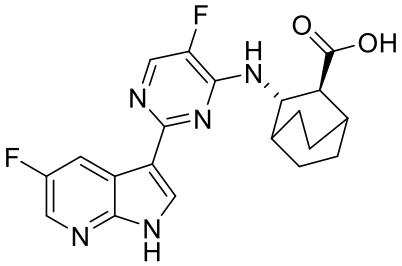
	K_d measured	K_d reported in literature
C-1015 probe precursor	19 ± 6 nM	-
VX-787	11 ± 4 nM	24 ± 7 nM (ITC) 8.3 ± 0.3 nM (SPR) 14 ± 1 nM (SPR) [145]
m^7GTP	37 ± 8 μM	1.5 ± 0.3 μM (ITC) [145] 177 ± 34 μM (SPR) [107] 349 ± 9 μM (SPR) [162]



C-1015 probe precursor



m^7GTP



VX -787 (pimodivir)

Figure 37. Chemical structures of C-1015 probe precursor, m7GTP and VX-787 inhibitor, on which the probe is based.

5. Discussion

The influenza A virus causes annual epidemics of acute respiratory disease. Due to the ability to evolve constantly via genetic reassortment, new viral stains arise, with a potential to cause pandemic outbreaks that affect a population of people over a large geographic area.

Finding new medical treatments to mitigate annual epidemics and tackle possible pandemic outbreaks is an outgoing challenge for the scientific community. Medications that constitute current armamentarium to prevent and control influenza virus outbreaks are becoming subjects to resistant development. Current research in anti-influenza drug development focuses extensively on an exploration of novel targets within the virus, such as influenza polymerase. Essential role of influenza polymerase in the viral life cycle and its high degree of conservation within the viral strains account for its leading role in the development of anti-influenza medications.

Influenza polymerase carries out replication and transcription of viral RNA. Transcription process involves two events, that are characteristic for the viral transcription: cap binding of capped host RNAs and endonuclease cleavage of bound RNA substrate to generate primer used in transcription. These two viral activities present an opportunity for targeting and have been exploited in drug development with varying degree of success. Particularly, the inhibitor of cap binding VX-787 is currently in late stages of clinical trials and novel endonuclease inhibitor baloxavir marboxil was recently approved in Japan and expected to be approved in the US by the end of this year [163]. This underlines the notion that targeting of these two activities of influenza polymerase is attainable.

The experimental work conducted in this master thesis project focused on the development of the novel in vitro high-throughput screening assays for the identification and target validation of compounds that would interfere with cap-binding and endonuclease activities of influenza polymerase. For each activity the assay is based on different technology, so the experimental work is inherently divided into two parts. For endonuclease activity, that resides in PA subunit of influenza polymerase the assay is based on AlphaScreen technology, while for cap binding activity of PB2 subunit of influenza polymerase the assay is based on DIANA (DNA-linked Inhibitor ANtibody Assay) that was developed in our laboratory. Different approaches were undertaken due

to the fact, that non-specific endonuclease activity of PA is incompatible with the method of detection used in DIANA (amplification of oligonucleotide bound to the probe).

The first step for both parts of the projects was cloning, expression, and purification of recombinant protein constructs that would be later used in screenings and accompanying experiments. Gene sequences encoding domains with corresponding activities (N-terminal endonuclease domain of PA and cap-binding domain of PB2) were ordered from GenScript.

Two recombinant protein constructs were designed and cloned for the experiments related to PA endonuclease: a fusion of NPA with GST for subsequent use in AlphaScreen-based screening and fusion of NPA with a cleavable His_6-SUMO tag for generation of tag-free NPA for crystallization trials. Both proteins were expressed in *E. coli*, purified using affinity chromatography on GSH sepharose resin in case of GST-NPA and chelation chromatography on Ni-NTA agarose resin in case of His_6-SUMO-NPA. GST-NPA recombinant protein was expressed at lower yields when compared to His_6-SUMO-NPA. This was expected, as SUMO-tag is generally known to improve solubility of fusion proteins resulting in higher expression. The yields were, however, sufficient for further experiments. Both recombinant proteins were further purified using affinity and gel permeation chromatography. His_6-SUMO-NPA protein was cleaved by ULP1 protease to remove the tag. The cleavage proceeded efficiently, with an only small fraction of the protein remaining non-cleaved, that was removed from the main fraction by capture on Ni-NTA agarose resin. The final fraction after gel permeation chromatography, however, contained a minor impurity of molecular weight corresponding to the His_6-SUMO-tag. Nevertheless, the final fractions of GST-NPA and tag-free NPA were of sufficient purity for use in screening and crystallization experiments, respectively.

The AlphaScreen-based screening assay for the identification of potential endonuclease inhibitors was designed to utilize competition between a probe, based on already known endonuclease inhibitor L-742.001 and tested compound. The probe contained the inhibitor moiety linked to biotin for binding of the probe to streptavidin-coated beads, which is required in AlphaScreen setup. It was synthesized by our collaborators from the organic chemistry department, Dr. Aleš Machara and Carlos Berenguer Albiñana, IOCB Prague. As the modification of L-742.001 with a linker that connects the inhibitor to biotin could interfere with the binding to GST-NPA, the binding parameters of the probe

were determined by SPR prior to AlphaScreen experiments. Moreover, we tested two types of the probes – with shorter (C-850) and longer (C-867) linkers to test whether the linker between inhibitor and biotin moieties also have an influence on the binding. Surface plasmon resonance experiments provided the binding parameters of the probes that were comparable with the parameters of the initial unaltered inhibitor. Previously reported dissociation constant for L-742.001 was 90 nM (measured by fluorescence polarization assay), while dissociation constants of both probes were < 10 nM. Off-rate constant (k_{off}) of the binding was below the limit of the method, which did not allow to determine the exact K_d values [164].

The next step in the development of AlphaScreen-based method was finding concentrations of probe and protein that would give the optimal luminescent signal. Cross-titration experiments, in which different concentration pairs of probe and GST-NPA were tested, were performed for both short-linker and long-linker probes. In this type of experiment, the signal typically increases with increasing concentrations of both interactants, until it reaches a so-called hook point, after which the signal starts to gradually decline. This is caused by an excess of reactants, which prevents of association of donor and acceptor beads and therefore, leads to signal decline. Even though the C-850 probe with a shorter linker followed the titration curve pattern, the overall signal was distinctively low (maximum at 35,00 alpha-counts). As was mentioned in the previous section, it is thought to be caused by steric hindrance upon GST-NPA binding to the probe linked to streptavidin-covered beads. As the assay is based on detecting a decrease of the upon competition between the probe and tested compound, the low signal range would influence the method sensitivity. The probe with the shorter linker was, therefore, found to be not suitable for use in the screening. Cross-titration with the C-867 long linker gave a clear titration curve pattern with a maximum signal at 126,000 alpha-counts. For further use in the screening, we chose the concentration of binding partners with the signal close to the maximum: GST-NPA 30 nM for C-850 and 75 nM for GST-NPA (concentration of beads 5μg/ml). These concentrations of binding partners were used in all subsequent AlphaScreen experiments. The test of the probes was followed by assay evaluation. Signal to noise and signal to background ratios were calculated from a series of experiments that determined maximal possible (beads coated with probe and GST-NPA, respectively) and minimal signals (uncoated beads). More importantly, the calculated Z' factor of 0.87 showed that the assay setup has the statistical power for use

in the large-scale high-throughput arrangement. The assay was additionally validated by testing positive and negative controls – L-742.001 and PB1-11 peptide. The later was functionally irrelevant in terms of interaction with GST-NPA, but previously used in another AlphaScreen assay developed in our laboratory. The IC_{50} value for L-742.001 determined using the AlphaScreen setup was 26 ± 1 nM, which is consistent with previously reported K_i of 90 nM determined by fluorescence polarization assay [164].

Assay validation and evaluation was followed by a small sample screening. The screening set of compounds consisted of N-benzotriazoylthiourea derivatives provided by Dr. Aleš Machara, IOCB Prague and HIV-integrase inhibitors raltegravir and elvitegravir. The latter compounds were added to the screen guided by literature in which HIV inhibitors were mentioned as the compounds that could have inhibitory activity based on the similarity between HIV integrase and PA endonuclease active sites [165]. No hits were detected among N-benzotriazoylthiourea derivatives. Raltegravir was found to possess inhibitory activity with IC_{50} of 31 ± 3 μM. This was in line with previous report about the inhibitory activity of compounds with pyrimidinol scaffold with three oxygen atoms in planar arrangement [165]. Inhibitory activity of raltegravir against influenza endonuclease, however, despite being assumed, was not previously confirmed experimentally. We additionally verified inhibitory activity of raltegravir using in vitro endonuclease assay and FRET-based assay. Elvitegravir, on the other hand, did not show significant inhibitory activity. It also contains a triad of co-planar oxygen atoms which was shown to chelate divalent ions in the PA endonuclease active site, however, it is based on a bulkier quinolone scaffold that could interfere with the binding [166].

Inhibitory activity of raltegravir is assumed to arise from the chelation of divalent ions, but it is unknown how the rest of the inhibitor structure contribute to the binding. We set up crystallization trials and obtained crystals of NPA that were soaked with raltegravir of final concentration 1 mM. The co-crystallization approach was also tested but resulted in poorly diffracting crystals (data not shown). The soaked crystals diffracted to 2.2 Å and 1.9 Å and the structures were solved (Dr. Petr Pachl, IOCB Prague). Both structures did not contain electron density for raltegravir at the active site. As the concentration of the inhibitor was sufficient to achieve 97 % fractional saturation, a possible explanation for this result might be that the binding of raltegravir is accompanied by a conformational change which is restricted in protein crystals.

Another part of this project focused on the development of high-throughput screening assay for identification of compounds that would interfere with another important activity of influenza polymerase – cap binding of host capped RNAs. This assay is based on DIANA (DNA-linked Inhibitor ANtibody Assay), a novel biological assay for diagnostics and drug discovery that was developed in our laboratory [159]. DIANA utilizes competition between an inhibitor-based probe and the screened compounds to determine their binding affinity. The method's characteristics, such as low false-positive and false-negative rates, as well as cost-efficiently makes DIANA a powerful tool that could be used in drug discovery. With this in mind, we set out to expand DIANA's outreach as a tool for discovery of anti-influenza medications that would target influenza polymerase, particularly, disrupt the interaction of PB2 cap-binding domain with caps of host RNAs.

For the DIANA assay, recombinant PB2 cap-binding domain fused to His_6-tag in *E. coli* and purified using chelation chromatography on Ni-NTA agarose resin and gel permeation chromatography. The protein was readily overexpressed in bacteria and a substantial amount of protein was obtained after two-step purification, which also fulfilled requirements of purity for use in DIANA. The probe used in DIANA was based on known cap-binding inhibitor VX-787. The inhibitor was connected to the oligonucleotide that was used for quantification by real-time PCR.

First set of experiments, conducted by my consultant Dr. Milan Kožíšek, were performed to determine the amount of PB2-His_6 protein per well for usage in the DIANA setup and influence of reducing conditions. Subsequent DIANA experiments were performed with 100 pg of PB2-His_6 per well in non-reducing conditions, as the reducing conditions were found to have a negative effect. In the next step of the assay development, we determined the K_d of the probe.

Additionally, to confirm that modification of VX-787 inhibitor with a linker did not affect its binding to PB2 cap-binding domain, we also determined K_d of the probe precursor C-1015 as 19 ± 6 nM. This was comparable to both K_d of the probe and the K_d for VX-787 inhibitor that was also determined by DIANA, as well as to values reported in previous studies (Table 8). Also, we determined K_d value for m^7GTP, a natural binding substrate for PB2 cap-binding domain. The values, previously reported in the literature show the higher degree of variance, from 1.5 ± 0.3 μM K_d determined by isothermal titration calorimetry to 349 ± 9 μM K_d determined by surface plasmon

resonance experiments [162, 167]. The K_d value that was determined by DIANA (37 ± 8 μM), therefore, lies in the interval of previously reported values. The developed DIANA assay will be soon used for screening of both IOCB library, that contains various classes of organic compounds and their derivatives prepared at IOCB Prague, and apropriate commercialy available libraries of compounds.

6. Conclusions

Constructs of recombinant PA endonuclease and PB2 cap-binding domains were designed, cloned, subsequently expressed in bacteria. The proteins were purified using affinity and gel permeation chromatography to desirable level of purity. GST-NPA recombinant protein was used in AlphaScreen assay. The assay was optimized and evaluated, and pilot screening was performed. HIV integrase inhibitor raltegravir was identified as micromolar inhibitor of PA endonuclease and its inhibitory activity was tested *in vitro*. Crystallization of raltegravir with PA endonuclease domain resulted in structure without bound inhibitor in the active site, probably due to restrictions of ligand soaking method. PB2-His_6 recombinant protein was used for DIANA screening. Conditions for DIANA were optimized and dissociation constants of the probe, probe precursor, VX-787 inhibitor and m^7GTP were determined. Both screening methods, that were developed, await utilization in large-scale screening of compound libraries.

7. References

[1] B. N. Fields, Knipe, D. M., Howley, P. M., & Griffin, D. E., *Fields Virology* Philadelphia: Lippincott Williams & Wilkins, 2013.

[2] A. D. M. E. Osterhaus, G. F. Rimmelzwaan, B. E. E. Martina, T. M. Bestebroer, and R. A. M. Fouchier, "Influenza B virus in seals," *Science,* vol. 288, pp. 1051-1053, May 12 2000.

[3] H. Kimura, C. Abiko, G. Peng, Y. Muraki, K. Sugawara, S. Hongo*, et al.*, "Interspecies transmission of influenza C virus between humans and pigs," *Virus Research,* vol. 48, pp. 71-79, Apr 1997.

[4] D. B. Smith, E. R. Gaunt, P. Digard, K. Templeton, and P. Simmonds, "Detection of influenza C virus but not influenza D virus in Scottish respiratory samples," *Journal of Clinical Virology,* vol. 74, pp. 50-53, Jan 2016.

[5] C. Calvo, M. L. Garcia-Garcia, M. Centeno, P. Perez-Brena, and I. Casas, "Influenza C virus infection in children, Spain," *Emerging Infectious Diseases,* vol. 12, pp. 1621-1622, Oct 2006.

[6] Y. Matsuzaki, N. Katsushima, Y. Nagai, M. Shoji, T. Itagaki, M. Sakamoto*, et al.*, "Clinical features of influenza C virus infection in children," *Journal of Infectious Diseases,* vol. 193, pp. 1229-1235, May 1 2006.

[7] B. M. Hause, M. Ducatez, E. A. Collin, Z. G. Ran, R. X. Liu, Z. Z. Sheng*, et al.*, "Isolation of a Novel Swine Influenza Virus from Oklahoma in 2011 Which Is Distantly Related to Human Influenza C Viruses," *Plos Pathogens,* vol. 9, Feb 2013.

[8] C. M. Fauquet and D. Fargette, "International Committee on Taxonomy of Viruses and the 3,142 unassigned species," *Virology Journal,* vol. 2, 2005.

[9] G. L. Ada and B. T. Perry, "Infectivity and Nucleic Acid Content of Influenza Virus," *Nature,* vol. 175, pp. 209-210, 1955.

[10] K. G. Murti, R. G. Webster, and I. M. Jones, "Localization of Rna-Polymerases on Influenza Viral Ribonucleoproteins by Immunogold Labeling," *Virology,* vol. 164, pp. 562-566, Jun 1988.

[11] J. Yasuda, S. Nakada, A. Kato, T. Toyoda, and A. Ishihama, "Molecular Assembly of Influenza-Virus - Association of the Ns2 Protein with Virion Matrix," *Virology,* vol. 196, pp. 249-255, Sep 1993.

[12] R. E. O'Neill, J. Talon, and P. Palese, "The influenza virus NEP (NS2 protein) mediates the nuclear export of viral ribonucleoproteins," *Embo Journal,* vol. 17, pp. 288-296, Jan 2 1998.

[13] F. Krammer, G. J. D. Smith, R. A. M. Fouchier, M. Peiris, K. Kedzierska, P. C. Doherty*, et al.*, "Influenza," *Nature Reviews Disease Primers,* vol. 4, p. 3, 2018/06/28 2018.

[14] G. L. Ada and B. T. Perry, "The Nucleic Acid Content of Influenza Virus," *Australian Journal of Experimental Biology and Medical Science,* vol. 32, pp. 453-468, 1954.

[15] H. M. Wise, A. Foeglein, J. C. Sun, R. M. Dalton, S. Patel, W. Howard*, et al.*, "A Complicated Message: Identification of a Novel PB1-Related Protein Translated from Influenza A Virus Segment 2 mRNA," *Journal of Virology,* vol. 83, pp. 8021-8031, Aug 15 2009.

[16] B. W. Jagger, H. M. Wise, J. C. Kash, K. A. Walters, N. M. Wills, Y. L. Xiao*, et al.*, "An Overlapping Protein-Coding Region in Influenza A Virus Segment 3 Modulates the Host Response," *Science,* vol. 337, pp. 199-204, Jul 13 2012.

[17] R. A. Lamb and P. W. Choppin, "Identification of a 2nd Protein (M2) Encoded by Rna Segment-7 of Influenza-Virus," *Virology,* vol. 112, pp. 729-737, 1981.

[18] R. A. Lamb and P. W. Choppin, "Segment-8 of the Influenza-Virus Genome Is Unique in Coding for 2 Polypeptides," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 76, pp. 4908-4912, 1979.

[19] U. Desselberger, V. R. Racaniello, J. J. Zazra, and P. Palese, "3'-Terminal and 5'-Terminal Sequences of Influenza-a, Influenza-B and Influenza-C Virus-Rna Segments Are Highly Conserved and Show Partial Inverted Complementarity," *Gene,* vol. 8, pp. 315-328, 1980.

[20] M. T. Hsu, J. D. Parvin, S. Gupta, M. Krystal, and P. Palese, "Genomic Rnas of Influenza-Viruses Are Held in a Circular Conformation in Virions and in Infected-Cells by a Terminal Panhandle," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 84, pp. 8140-8144, Nov 1987.

[21] T. Odagiri and M. Tashiro, "Segment-specific noncoding sequences of the influenza virus genome RNA are involved in the specific competition between defective interfering RNA and its progenitor RNA segment at the virion assembly step," *Journal of Virology,* vol. 71, pp. 2138-2145, Mar 1997.

[22] J. S. Gibbs, D. Malide, F. Hornung, J. R. Bennink, and J. W. Yewdell, "The influenza A virus PB1-F2 protein targets the inner mitochondrial membrane via a predicted basic amphipathic helix that disrupts mitochondrial function," *Journal of Virology,* vol. 77, pp. 7214-7224, Jul 2003.

[23] W. Garten, F. X. Bosch, D. Linder, R. Rott, and H. D. Klenk, "Proteolytic Activation of the Influenza-Virus Hemagglutinin - the Structure of the Cleavage Site and the Enzymes Involved in Cleavage," *Virology,* vol. 115, pp. 361-374, 1981.

[24] B. J. Chen, M. Takeda, and R. A. Lamb, "Influenza virus hemagglutinin (H3 subtype) requires palmitoylation of its cytoplasmic tail for assembly: M1 proteins of two subtypes differ in their ability to support assembly," *Journal of Virology,* vol. 79, pp. 13673-13684, Nov 2005.

[25] J. J. Skehel and D. C. Wiley, "Receptor binding and membrane fusion in virus entry: The influenza hemagglutinin," *Annual Review of Biochemistry,* vol. 69, pp. 531-569, 2000.

[26] M. N. Matrosovich, T. Y. Matrosovich, T. Gray, N. A. Roberts, and H. D. Klenk, "Human and avian influenza viruses target different cell types in cultures of human airway epithelium," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 101, pp. 4620-4624, Mar 30 2004.

[27] J. Nelson, S. S. Couceiro, J. C. Paulson, and L. G. Baum, "Influenza-Virus Strains Selectively Recognize Sialyloligosaccharides on Human Respiratory Epithelium - the Role of the Host-Cell in Selection of Hemagglutinin Receptor Specificity," *Virus Research,* vol. 29, pp. 155-165, Aug 1993.

[28] G. N. Rogers, J. C. Paulson, R. S. Daniels, J. J. Skehel, I. A. Wilson, and D. C. Wiley, "Single Amino-Acid Substitutions in Influenza Hemagglutinin Change Receptor-Binding Specificity," *Nature,* vol. 304, pp. 76-78, 1983.

[29] L. Glaser, J. Stevens, D. Zamarin, I. A. Wilson, A. Garcia-Sastre, T. M. Tumpey*, et al.*, "A single amino acid substitution in 1918 influenza virus hemagglutinin

changes receptor binding specificity," *Journal of Virology,* vol. 79, pp. 11533-11536, Sep 2005.
[30] S. G. Tong, D. Alford, and J. Bentz, "Influenza virus liposome lipid mixing is leaky and largely insensitive to the material properties of the target membrane," *Biochemistry,* vol. 35, pp. 4956-4965, Apr 16 1996.
[31] R. Jiricek, G. Schwarz, and T. Stegmann, "Pores formed by influenza hemagglutinin," *Biochimica Et Biophysica Acta-Biomembranes,* vol. 1330, pp. 17-28, Nov 13 1997.
[32] L. J. Mitnaul, M. N. Matrosovich, M. R. Castrucci, A. B. Tuzikov, N. V. Bovin, D. Kobasa*, et al.*, "Balanced hemagglutinin and neuraminidase activities are critical for efficient replication of influenza A virus," *Journal of Virology,* vol. 74, pp. 6015-6020, Jul 2000.
[33] N. G. Wrigley, J. J. Skehel, Charlwoo.Pa, and C. M. Brand, "Size and Shape of Influenza-Virus Neuraminidase," *Virology,* vol. 51, pp. 525-529, 1973.
[34] W. G. Laver and R. C. Valentin, "Morphology of Isolated Hemagglutinin and Neuraminidase Subunits of Influenza Virus," *Virology,* vol. 38, pp. 105-&, 1969.
[35] J. Blok and G. M. Air, "Sequence Variation at the 3' End of the Neuraminidase Gene from 39 Influenza Type-a Viruses," *Virology,* vol. 121, pp. 211-229, 1982.
[36] J. N. Varghese, W. G. Laver, and P. M. Colman, "Structure of the Influenza-Virus Glycoprotein Antigen Neuraminidase at 2.9-a-Resolution," *Nature,* vol. 303, pp. 35-40, 1983.
[37] P. V. Wagh and O. P. Bahl, "Sugar Residues on Proteins," *Crc Critical Reviews in Biochemistry,* vol. 10, pp. 307-377, 1981.
[38] P. M. Colman, J. N. Varghese, and W. G. Laver, "Structure of the Catalytic and Antigenic Sites in Influenza-Virus Neuraminidase," *Nature,* vol. 303, pp. 41-44, 1983.
[39] R. J. Russell, L. F. Haire, D. J. Stevens, P. J. Collins, Y. P. Lin, G. M. Blackburn*, et al.*, "The structure of H5N1 avian influenza neuraminidase suggests new opportunities for drug design," *Nature,* vol. 443, pp. 45-49, Sep 7 2006.
[40] Q. Li, J. X. Qi, W. Zhang, C. J. Vavricka, Y. Shi, J. H. Wei*, et al.*, "The 2009 pandemic H1N1 neuraminidase N1 lacks the 150-cavity in its active site," *Nature Structural & Molecular Biology,* vol. 17, pp. 1266-1268, Oct 2010.
[41] L. Wakefield and G. G. Brownlee, "Rna-Binding Properties of Influenza-a Virus Matrix Protein M1," *Nucleic Acids Research,* vol. 17, pp. 8569-8580, Nov 11 1989.
[42] R. A. Lamb, S. L. Zebedee, and C. D. Richardson, "Influenza Virus-M2 Protein Is an Integral Membrane-Protein Expressed on the Infected-Cell Surface," *Cell,* vol. 40, pp. 627-633, 1985.
[43] R. J. Sugrue and A. J. Hay, "Structural Characteristics of the M2 Protein of Influenza-a Viruses - Evidence That It Forms a Tetrameric Channel," *Virology,* vol. 180, pp. 617-624, Feb 1991.
[44] M. S. Sansom, I. D. Kerr, G. R. Smith, and H. S. Son, "The influenza A virus M2 channel: a molecular modeling and simulation study," *Virology,* vol. 233, pp. 163-73, Jun 23 1997.
[45] L. H. Pinto, G. R. Dieckmann, C. S. Gandhi, C. G. Papworth, J. Braman, M. A. Shaughnessy*, et al.*, "A functionally defined model for the M2 proton channel of influenza A virus suggests a mechanism for its ion selectivity," *Proc Natl Acad Sci U S A,* vol. 94, pp. 11301-6, Oct 14 1997.

[46] A. Garcia-Sastre, A. Egorov, D. Matassov, S. Brandt, D. E. Levy, J. E. Durbin*, et al.*, "Influenza A virus lacking the NS1 gene replicates in interferon-deficient systems," *Virology,* vol. 252, pp. 324-30, Dec 20 1998.
[47] W. Wang, K. Riedel, P. Lynch, C. Y. Chien, G. T. Montelione, and R. M. Krug, "RNA binding by the novel helical domain of the influenza virus NS1 protein requires its dimer structure and a small number of specific basic amino acids," *RNA,* vol. 5, pp. 195-205, Feb 1999.
[48] B. G. Hale, R. E. Randall, J. Ortin, and D. Jackson, "The multifunctional NS1 protein of influenza A viruses," *J Gen Virol,* vol. 89, pp. 2359-76, Oct 2008.
[49] J. C. Obenauer, J. Denson, P. K. Mehta, X. Su, S. Mukatira, D. B. Finkelstein*, et al.*, "Large-scale sequence analysis of avian influenza isolates," *Science,* vol. 311, pp. 1576-80, Mar 17 2006.
[50] A. Cheng, S. M. Wong, and Y. A. Yuan, "Structural basis for dsRNA recognition by NS1 protein of influenza A virus," *Cell Research,* vol. 19, pp. 187-195, Feb 2009.
[51] J. M. Aramini, L. C. Ma, L. Zhou, C. M. Schauder, K. Hamilton, B. R. Amer*, et al.*, "Dimer interface of the effector domain of non-structural protein 1 from influenza A virus: an interface with multiple functions," *J Biol Chem,* vol. 286, pp. 26050-60, Jul 22 2011.
[52] M. Mibayashi, L. Martinez-Sobrido, Y. M. Loo, W. B. Cardenas, M. Gale, and A. Garcia-Sastre, "Inhibition of retinoic acid-inducible gene I-mediated induction of beta interferon by the NS1 protein of influenza a virus," *Journal of Virology,* vol. 81, pp. 514-524, Jan 2007.
[53] Z. Guo, L. M. Chen, H. Zeng, J. A. Gomez, J. Plowden, T. Fujita*, et al.*, "NS1 protein of influenza A virus inhibits the function of intracytoplasmic pathogen sensor, RIG-I," *American Journal of Respiratory Cell and Molecular Biology,* vol. 36, pp. 263-269, Mar 2007.
[54] M. U. Gack, R. A. Albrecht, T. Urano, K. S. Inn, I. C. Huang, E. Carnero*, et al.*, "Influenza A Virus NS1 Targets the Ubiquitin Ligase TRIM25 to Evade Recognition by the Host Viral RNA Sensor RIG-I," *Cell Host & Microbe,* vol. 5, pp. 439-449, May 21 2009.
[55] J. J. Skehel and D. C. Wiley, "Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin," *Annu Rev Biochem,* vol. 69, pp. 531-69, 2000.
[56] R. A. Childs, A. S. Palma, S. Wharton, T. Matrosovich, Y. Liu, W. G. Chai*, et al.*, "Receptor-binding specificity of pandemic influenza A (H1N1) 2009 virus determined by carbohydrate microarray," *Nature Biotechnology,* vol. 27, pp. 797-799, Sep 2009.
[57] S. B. Sieczkarski and G. R. Whittaker, "Influenza virus can enter and infect cells in the absence of clathrin-mediated endocytosis," *Journal of Virology,* vol. 76, pp. 10455-10464, Oct 2002.
[58] E. de Vries, D. M. Tscherne, M. J. Wienholts, V. Cobos-Jimenez, F. Scholte, A. Garcia-Sastre*, et al.*, "Dissection of the Influenza A Virus Endocytic Routes Reveals Macropinocytosis as an Alternative Entry Pathway," *Plos Pathogens,* vol. 7, Mar 2011.
[59] A. Helenius, "Unpacking the incoming influenza virus," *Cell,* vol. 69, pp. 577-8, May 15 1992.

[60] J. F. Cros, A. Garcia-Sastre, and P. Palese, "An unconventional NLS is critical for the nuclear import of the influenza A virus nucleoprotein and ribonucleoprotein," *Traffic,* vol. 6, pp. 205-213, Mar 2005.
[61] D. Paterson and E. Fodor, "Emerging Roles for the Influenza A Virus Nuclear Export Protein (NEP)," *Plos Pathogens,* vol. 8, Dec 2012.
[62] L. Brunotte, J. Flies, H. Bolte, P. Reuther, F. Vreede, and M. Schwemmle, "The Nuclear Export Protein of H5N1 Influenza A Viruses Recruits Matrix 1 (M1) Protein to the Viral Ribonucleoprotein to Mediate Nuclear Export," *Journal of Biological Chemistry,* vol. 289, pp. 20067-20077, Jul 18 2014.
[63] M. J. Amorim, E. A. Bruce, E. K. C. Read, A. Foeglein, R. Mahen, A. D. Stuart, *et al.*, "A Rab11-and Microtubule-Dependent Mechanism for Cytoplasmic Transport of Influenza A Virus Viral RNA," *Journal of Virology,* vol. 85, pp. 4143-4156, May 2011.
[64] E. A. Bruce, P. Digard, and A. D. Stuart, "The Rab11 Pathway Is Required for Influenza A Virus Budding and Filament Formation," *Journal of Virology,* vol. 84, pp. 5848-5859, Jun 2010.
[65] A. J. Eisfeld, E. Kawakami, T. Watanabe, G. Neumann, and Y. Kawaoka, "RAB11A Is Essential for Transport of the Influenza Virus Genome to the Plasma Membrane," *Journal of Virology,* vol. 85, pp. 6117-6126, Jul 2011.
[66] A. Kundu, R. T. Avalos, C. M. Sanderson, and D. P. Nayak, "Transmembrane domain of influenza virus neuraminidase, a type II protein, possesses an apical sorting signal in polarized MDCK cells," *Journal of Virology,* vol. 70, pp. 6508-6515, Sep 1996.
[67] L. V. Jones, R. W. Compans, A. R. Davis, T. J. Bos, and D. P. Nayak, "Surface Expression of Influenza-Virus Neuraminidase, an Amino-Terminally Anchored Viral Membrane Glycoprotein, in Polarized Epithelial-Cells," *Molecular and Cellular Biology,* vol. 5, pp. 2181-2189, 1985.
[68] M. G. Roth, R. W. Compans, L. Giusti, A. R. Davis, D. P. Nayak, M. J. Gething, *et al.*, "Influenza-Virus Hemagglutinin Expression Is Polarized in Cells Infected with Recombinant Sv40 Viruses Carrying Cloned Hemagglutinin DNA," *Cell,* vol. 33, pp. 435-443, 1983.
[69] G. P. Leser and R. A. Lamb, "Influenza virus assembly and budding in raft-derived microdomains: A quantitative analysis of the surface distribution of HA, NA and M2 proteins," *Virology,* vol. 342, pp. 215-227, Nov 25 2005.
[70] C. B. Brooke, W. L. Ince, J. Wrammert, R. Ahmed, P. C. Wilson, J. R. Bennink, *et al.*, "Most Influenza A Virions Fail To Express at Least One Essential Viral Protein," *Journal of Virology,* vol. 87, pp. 3155-3162, Mar 2013.
[71] S. D. Duhaut and J. W. McCauley, "Defective RNAs inhibit the assembly of influenza virus genome segments in a segment-specific manner," *Virology,* vol. 216, pp. 326-337, Feb 15 1996.
[72] K. Fujii, Y. Fujii, T. Noda, Y. Muramoto, T. Watanabe, A. Takada, *et al.*, "Importance of both the coding and the segment-specific noncoding regions of the influenza a virus NS segment for its efficient incorporation into virions," *Journal of Virology,* vol. 79, pp. 3766-3774, Mar 2005.
[73] Y. Fujii, H. Goto, T. Watanabe, T. Yoshida, and T. Kawaoka, "Selective incorporation of influenza virus RNA segments into virions," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 100, pp. 2002-2007, Feb 18 2003.

[74] J. M. Barry, "The site of origin of the 1918 influenza pandemic and its public health implications," *J Transl Med,* vol. 2, p. 3, Jan 20 2004.
[75] L. Glaser, J. Stevens, D. Zamarin, I. A. Wilson, A. Garcia-Sastre, T. M. Tumpey*, et al.*, "A single amino acid substitution in 1918 influenza virus hemagglutinin changes receptor binding specificity," *J Virol,* vol. 79, pp. 11533-6, Sep 2005.
[76] D. M. Morens, J. K. Taubenberger, and A. S. Fauci, "Predominant role of bacterial pneumonia as a cause of death in pandemic influenza: implications for pandemic influenza preparedness," *J Infect Dis,* vol. 198, pp. 962-70, Oct 1 2008.
[77] F. S. Dawood, S. Jain, L. Finelli, M. W. Shaw, S. Lindstrom, R. J. Garten*, et al.*, "Emergence of a Novel Swine-Origin Influenza A (H1N1) Virus in Humans Novel Swine-Origin Influenza A (H1N1) Virus Investigation Team," *New England Journal of Medicine,* vol. 360, pp. 2605-2615, Jun 18 2009.
[78] R. J. Garten, C. T. Davis, C. A. Russell, B. Shu, S. Lindstrom, A. Balish*, et al.*, "Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans," *Science,* vol. 325, pp. 197-201, Jul 10 2009.
[79] G. Chowell, S. M. Bertozzi, M. A. Colchero, H. Lopez-Gatell, C. Alpuche-Aranda, M. Hernandez*, et al.*, "Severe respiratory disease concurrent with the circulation of H1N1 influenza," *N Engl J Med,* vol. 361, pp. 674-9, Aug 13 2009.
[80] C. Wang, K. Takeuchi, L. J. Holsinger, R. A. Lamb, and L. H. Pinto, "Ion Channel Activity of the M2 Protein of Influenza a-Virus," *Biophysical Journal,* vol. 64, pp. A94-A94, Feb 1993.
[81] A. E. Fiore, D. K. Shay, P. Haber, J. K. Iskander, T. M. Uyeki, G. Mootrey*, et al.*, "Prevention and control of influenza. Recommendations of the Advisory Committee on Immunization Practices (ACIP), 2007," *MMWR Recomm Rep,* vol. 56, pp. 1-54, Jul 13 2007.
[82] A. Moscona, "Global Transmission of Oseltamivir-Resistant Influenza," *New England Journal of Medicine,* vol. 360, pp. 953-956, Mar 5 2009.
[83] R. Arranz, R. Coloma, F. J. Chichon, J. J. Conesa, J. L. Carrascosa, J. M. Valpuesta*, et al.*, "The Structure of Native Influenza Virion Ribonucleoproteins," *Science,* vol. 338, pp. 1634-1637, Dec 21 2012.
[84] F. Baudin, C. Bach, S. Cusack, and R. W. H. Ruigrok, "Structure of Influenza-Virus Rnp .1. Influenza-Virus Nucleoprotein Melts Secondary Structure in Panhandle Rna and Exposes the Bases to the Solvent," *Embo Journal,* vol. 13, pp. 3158-3165, Jul 1 1994.
[85] G. Gabriel and E. Fodor, "Molecular Determinants of Pathogenicity in the Polymerase Complex," *Influenza Pathogenesis and Control - Vol I,* vol. 385, pp. 35-60, 2014.
[86] A. Pflug, D. Guilligay, S. Reich, and S. Cusack, "Structure of influenza A polymerase bound to the viral RNA promoter," *Nature,* vol. 516, pp. 355-60, Dec 18 2014.
[87] E. C. Hutchinson, O. E. Orr, S. M. Liu, O. G. Engelhardt, and E. Fodor, "Characterization of the interaction between the influenza A virus polymerase subunit PB1 and the host nuclear import factor Ran-binding protein 5," *Journal of General Virology,* vol. 92, pp. 1859-1869, Aug 2011.
[88] A. J. W. te Velthuis, "Common and unique features of viral RNA-dependent polymerases," *Cellular and Molecular Life Sciences,* vol. 71, pp. 4403-4420, Nov 2014.
[89] D. F. Zamyatkin, F. Parra, J. M. M. Alonso, D. A. Harki, B. R. Peterson, P. Grochulski*, et al.*, "Structural insights into mechanisms of catalysis and inhibition

in Norwalk virus polymerase," *Journal of Biological Chemistry,* vol. 283, pp. 7705-7712, Mar 21 2008.

[90] S. A. Jablonski and C. D. Morrow, "Mutation of the Aspartic-Acid Residues of the Gdd Sequence Motif of Poliovirus Rna-Dependent Rna-Polymerase Results in Enzymes with Altered Metal-Ion Requirements for Activity," *Journal of Virology,* vol. 69, pp. 1532-1539, Mar 1995.

[91] K. M. Ogden, H. N. Ramanathan, and J. T. Patton, "Mutational analysis of residues involved in nucleotide and divalent cation stabilization in the rotavirus RNA-dependent RNA polymerase catalytic pocket," *Virology,* vol. 431, pp. 12-20, Sep 15 2012.

[92] S. K. Biswas and D. P. Nayak, "Mutational Analysis of the Conserved Motifs of Influenza-a Virus Polymerase Basic-Protein 1," *Journal of Virology,* vol. 68, pp. 1819-1826, Mar 1994.

[93] T. Deng, F. T. Vreede, and G. G. Brownlee, "Different de novo initiation strategies are used by influenza virus RNA polymerase on its cRNA and viral RNA promoters during viral RNA replication," *Journal of Virology,* vol. 80, pp. 2337-2348, Mar 2006.

[94] A. J. W. te Velthuis, N. C. Robb, A. N. Kapanidis, and E. Fodor, "The role of the priming loop in influenza A virus RNA synthesis," *Nature Microbiology,* vol. 1, May 2016.

[95] A. York, N. Hengrung, F. T. Vreede, J. T. Huiskonen, and E. Fodor, "Isolation and characterization of the positive-sense replicative intermediate of a negative-strand RNA virus," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 110, pp. E4238-E4245, Nov 5 2013.

[96] N. Jorba, R. Coloma, and J. Ortin, "Genetic trans-Complementation Establishes a New Model for Influenza Virus RNA Transcription and Replication," *Plos Pathogens,* vol. 5, May 2009.

[97] W. F. Gu, G. R. Gallagher, W. W. Dai, P. Liu, R. D. Li, M. I. Trombly*, et al.*, "Influenza A virus preferentially snatches noncoding RNA caps," *Rna,* vol. 21, pp. 2067-2075, Dec 2015.

[98] S. J. Plotch, M. Bouloy, I. Ulmanen, and R. M. Krug, "A Unique Cap(M7gpppxm)-Dependent Influenza Virion Endonuclease Cleaves Capped Rnas to Generate the Primers That Initiate Viral-Rna Transcription," *Cell,* vol. 23, pp. 847-858, 1981.

[99] D. Koppstein, J. Ashour, and D. P. Bartel, "Sequencing the cap-snatching repertoire of H1N1 influenza provides insight into the mechanism of viral transcription initiation," *Nucleic Acids Research,* vol. 43, pp. 5052-5064, May 26 2015.

[100] L. L. M. Poon, D. C. Pritlove, E. Fodor, and G. G. Brownlee, "Direct evidence that the poly(A) tail of influenza A virus mRNA is synthesized by reiterative copying of a U track in the virion RNA template," *Journal of Virology,* vol. 73, pp. 3473-3476, Apr 1999.

[101] A. York and E. Fodor, "Biogenesis, assembly, and export of viral messenger ribonucleoproteins in the influenza A virus infected cell," *Rna Biology,* vol. 10, pp. 1274-1282, Aug 1 2013.

[102] A. J. W. te Velthuis and E. Fodor, "Influenza virus RNA polymerase: insights into the mechanisms of viral RNA synthesis," *Nature Reviews Microbiology,* vol. 14, pp. 479-493, Aug 2016.

[103] H. Jin and R. M. Elliott, "Characterization of Bunyamwera Virus-S Rna That Is Transcribed and Replicated by the L-Protein Expressed from Recombinant Vaccinia Virus," *Journal of Virology,* vol. 67, pp. 1396-1404, Mar 1993.
[104] S. J. Polyak, S. Zheng, and D. G. Harnish, "5' Termini of Pichinde Arenavirus S-Rnas and Messenger-Rnas Contain Nontemplated Nucleotides," *Journal of Virology,* vol. 69, pp. 3211-3215, May 1995.
[105] O. G. Engelhardt, M. Smith, and E. Fodor, "Association of the influenza a virus RNA-dependent RNA polymerase with cellular RNA polymerase II," *Journal of Virology,* vol. 79, pp. 5812-5818, May 2005.
[106] M. Lukarska, G. Fournier, A. Pflug, P. Resa-Infante, S. Reich, N. Naffakh*, et al.*, "Structural basis of an essential interaction between influenza polymerase and Pol II CTD," *Nature,* vol. 541, pp. 117-+, Jan 5 2017.
[107] D. Guilligay, F. Tarendeau, P. Resa-Infante, R. Coloma, T. Crepin, P. Sehr*, et al.*, "The structural basis for cap binding by influenza virus polymerase subunit PB2," *Nat Struct Mol Biol,* vol. 15, pp. 500-6, May 2008.
[108] A. E. Hodel, P. D. Gershon, and F. A. Quiocho, "Structural basis for sequence-nonspecific recognition of 5 '-capped mRNA by a cap-modifying enzyme," *Molecular Cell,* vol. 1, pp. 443-447, Feb 1998.
[109] G. H. Hu, P. D. Gershon, A. E. Hodel, and F. A. Quiocho, "mRNA cap recognition: Dominant role of enhanced stacking interactions between methylated bases and protein aromatic side chains," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 96, pp. 7149-7154, Jun 22 1999.
[110] S. Reich, D. Guilligay, A. Pflug, H. Malet, I. Berger, T. Crepin*, et al.*, "Structural insight into cap-snatching and RNA synthesis by influenza polymerase," *Nature,* vol. 516, pp. 361-6, Dec 18 2014.
[111] A. Dias, D. Bouvier, T. Crepin, A. A. McCarthy, D. J. Hart, F. Baudin*, et al.*, "The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit," *Nature,* vol. 458, pp. 914-8, Apr 16 2009.
[112] L. Doan, B. Handa, N. A. Roberts, and K. Klumpp, "Metal ion catalysis of RNA cleavage by the influenza virus endonuclease," *Biochemistry,* vol. 38, pp. 5612-5619, Apr 27 1999.
[113] T. Crepin, A. Dias, A. Palencia, C. Swale, S. Cusack, and R. W. Ruigrok, "Mutational and metal binding analysis of the endonuclease domain of the influenza virus polymerase PA subunit," *J Virol,* vol. 84, pp. 9096-104, Sep 2010.
[114] T. Crepin, A. Dias, A. Palencia, C. Swale, S. Cusack, and R. W. H. Ruigrok, "Mutational and Metal Binding Analysis of the Endonuclease Domain of the Influenza Virus Polymerase PA Subunit," *Journal of Virology,* vol. 84, pp. 9096-9104, Sep 15 2010.
[115] S. Y. Xiao, M. L. Klein, D. N. LeBard, B. G. Levine, H. J. Liang, C. M. MacDermaid*, et al.*, "Magnesium-Dependent RNA Binding to the PA Endonuclease Domain of the Avian Influenza Polymerase," *Journal of Physical Chemistry B,* vol. 118, pp. 873-889, Jan 30 2014.
[116] M. Homma, A. L. Jayewardene, J. Gambertoglio, and F. Aweeka, "High-performance liquid chromatographic determination of ribavirin in whole blood to assess disposition in erythrocytes," *Antimicrobial Agents and Chemotherapy,* vol. 43, pp. 2716-2719, Nov 1999.
[117] R. C. Willis, D. A. Carson, and J. E. Seegmiller, "Adenosine Kinase Initiates Major Route of Ribavirin Activation in a Cultured Human Cell Line,"

Proceedings of the National Academy of Sciences of the United States of America, vol. 75, pp. 3042-3044, 1978.

[118] B. Eriksson, E. Helgstrand, N. G. Johansson, A. Larsson, A. Misiorny, J. O. Noren*, et al.*, "Inhibition of Influenza-Virus Ribonucleic-Acid Polymerase by Ribavirin Triphosphate," *Antimicrobial Agents and Chemotherapy,* vol. 11, pp. 946-951, 1977.

[119] L. F. Cassidy and J. L. Patterson, "Mechanism of La-Crosse Virus Inhibition by Ribavirin," *Antimicrobial Agents and Chemotherapy,* vol. 33, pp. 2009-2011, Nov 1989.

[120] S. Crotty, D. Maag, J. J. Arnold, W. D. Zhong, J. Y. N. Lau, Z. Hong*, et al.*, "The broad-spectrum antiviral ribonucleoside ribavirin is an RNA virus mutagen (vol 6, pg 1375, 2000)," *Nature Medicine,* vol. 7, pp. 255-255, Feb 2001.

[121] S. Crotty, C. E. Cameron, and R. Andino, "RNA virus error catastrophe: Direct molecular test by using ribavirin," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 98, pp. 6895-6900, Jun 5 2001.

[122] D. Maag, C. Castro, Z. Hong, and C. E. Cameron, "Hepatitis C virus RNA-dependent RNA polymerase (NS5B) as a mediator of the antiviral activity of ribavirin," *Journal of Biological Chemistry,* vol. 276, pp. 46094-46098, Dec 7 2001.

[123] D. G. Streeter, J. T. Witkowski, G. P. Khare, R. W. Sidwell, R. J. Bauer, R. K. Robins*, et al.*, "Mechanism of Action of 1-Beta-D-Ribofuranosyl-1,2,4-Triazole-3-Carboxamide (Virazole) - New Broad-Spectrum Antiviral Agent," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 70, pp. 1174-1178, 1973.

[124] S. K. Wray, B. E. Gilbert, M. W. Noall, and V. Knight, "Mode of Action of Ribavirin - Effect of Nucleotide Pool Alterations on Influenza-Virus Ribonucleoprotein Synthesis," *Antiviral Research,* vol. 5, pp. 29-37, 1985.

[125] R. K. Robins, G. R. Revankar, P. A. Mckernan, B. K. Murray, J. J. Kirsi, and J. A. North, "The Importance of Imp Dehydrogenase Inhibition in the Broad-Spectrum Antiviral Activity of Ribavirin and Selenazofurin," *Advances in Enzyme Regulation,* vol. 24, pp. 29-43, 1985.

[126] Y. Furuta, K. Takahashi, M. Kuno-Maekawa, H. Sangawa, S. Uehara, K. Kozaki*, et al.*, "Mechanism of action of T-705 against influenza virus," *Antimicrobial Agents and Chemotherapy,* vol. 49, pp. 981-986, Mar 2005.

[127] B. B. Gowen, M. H. Wong, K. H. Jung, A. B. Sanders, M. Mendenhall, K. W. Bailey*, et al.*, "In vitro and in vivo activities of T-705 against arenavirus and bunyavirus infections," *Antimicrobial Agents and Chemotherapy,* vol. 51, pp. 3168-3176, Sep 2007.

[128] J. Guedj, G. Piorkowski, F. Jacquot, V. Madelain, T. H. T. Nguyen, A. Rodallec*, et al.*, "Antiviral efficacy of favipiravir against Ebola virus: A translational study in cynomolgus macaques," *Plos Medicine,* vol. 15, Mar 2018.

[129] W. J. Zhu, Z. R. Zhang, S. H. He, G. Wong, L. Banadyga, and X. G. Qiu, "Successful treatment of Marburg virus with orally administrated T-705 (Favipiravir) in a mouse model," *Antiviral Research,* vol. 151, pp. 39-49, Mar 2018.

[130] L. Naesens, L. W. Guddat, D. T. Keough, A. B. P. van Kuilenburg, J. Meijer, J. Vande Voorde*, et al.*, "Role of Human Hypoxanthine Guanine Phosphoribosyltransferase in Activation of the Antiviral Agent T-705 (Favipiravir)," *Molecular Pharmacology,* vol. 84, pp. 615-629, Oct 2013.

[131] Z. N. Jin, L. K. Smith, V. K. Rajwanshi, B. Kim, and J. Deval, "The Ambiguous Base-Pairing and High Substrate Efficiency of T-705 (Favipiravir) Ribofuranosyl 5 '-Triphosphate towards Influenza A Virus Polymerase," *Plos One,* vol. 8, Jul 10 2013.

[132] T. Baranovich, S. S. Wong, J. Armstrong, H. Marjuki, R. J. Webby, R. G. Webster*, et al.*, "T-705 (Favipiravir) Induces Lethal Mutagenesis in Influenza A H1N1 Viruses In Vitro," *Journal of Virology,* vol. 87, pp. 3741-3751, Apr 2013.

[133] H. Sangawa, T. Komeno, H. Nishikawa, A. Yoshida, K. Takahashi, N. Nomura*, et al.*, "Mechanism of Action of T-705 Ribosyl Triphosphate against Influenza Virus RNA Polymerase," *Antimicrobial Agents and Chemotherapy,* vol. 57, pp. 5202-5208, Nov 2013.

[134] L. Naesens, A. Stevaert, and E. Vanderlinden, "Antiviral therapies on the horizon for influenza," *Current Opinion in Pharmacology,* vol. 30, pp. 106-115, Oct 2016.

[135] J. Tomassini, H. Selnick, M. E. Davies, M. E. Armstrong, J. Baldwin, M. Bourgeois*, et al.*, "Inhibition of Cap (M(7)Gpppxm)-Dependent Endonuclease of Influenza-Virus by 4-Substituted 2,4-Dioxobutanoic Acid Compounds," *Antimicrobial Agents and Chemotherapy,* vol. 38, pp. 2827-2837, Dec 1994.

[136] J. E. Tomassini, M. E. Davies, J. C. Hastings, R. Lingham, M. Mojena, S. L. Raghoobar*, et al.*, "A novel antiviral agent which inhibits the endonuclease of influenza viruses," *Antimicrobial Agents and Chemotherapy,* vol. 40, pp. 1189-1193, May 1996.

[137] J. C. Hastings, H. Selnick, B. Wolanski, and J. E. Tomassini, "Anti-influenza virus activities of 4-substituted 2,4-dioxobutanoic acid inhibitors," *Antimicrobial Agents and Chemotherapy,* vol. 40, pp. 1304-1307, May 1996.

[138] R. M. DuBois, P. J. Slavish, B. M. Baughman, M. K. Yun, J. Bao, R. J. Webby*, et al.*, "Structural and Biochemical Basis for Development of Influenza Virus Inhibitors Targeting the PA Endonuclease," *Plos Pathogens,* vol. 8, Aug 2012.

[139] M. Nakazawa, S. E. Kadowaki, I. Watanabe, Y. Kadowaki, M. Takei, and H. Fukuda, "PA subunit of RNA polymerase as a promising target for anti-influenza virus agents," *Antiviral Research,* vol. 78, pp. 194-201, Jun 2008.

[140] M. S. Song, G. Kumar, W. R. Shadrick, W. Zhou, T. Jeevan, Z. M. Li*, et al.*, "Identification and characterization of influenza variants resistant to a viral endonuclease inhibitor," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 113, pp. 3669-3674, Mar 29 2016.

[141] S. Omoto, V. Speranzini, T. Hashimoto, T. Noshi, H. Yamaguchi, M. Kawai*, et al.*, "Characterization of influenza virus variants induced by treatment with the endonuclease inhibitor baloxavir marboxil," *Scientific Reports,* vol. 8, Jun 25 2018.

[142] S. F. Yuan, N. R. Zhang, K. Singh, H. P. Shuai, H. Chu, J. Zhou*, et al.*, "Cross-Protection of Influenza A Virus Infection by a DNA Aptamer Targeting the PA Endonuclease Domain," *Antimicrobial Agents and Chemotherapy,* vol. 59, pp. 4082-4093, Jul 2015.

[143] D. Guilligay, F. Tarendeau, P. Resa-Infante, R. Coloma, T. Crepin, P. Sehr*, et al.*, "The structural basis for cap binding by influenza virus polymerase subunit PB2," *Nature Structural & Molecular Biology,* vol. 15, pp. 500-506, May 2008.

[144] M. P. Clark, M. W. Ledeboer, I. Davies, R. A. Byrn, S. M. Jones, E. Perola*, et al.*, "Discovery of a Novel, First-in-Class, Orally Bioavailable Azaindole Inhibitor (VX-787) of Influenza PB2," *Journal of Medicinal Chemistry,* vol. 57, pp. 6668-6678, Aug 14 2014.

[145] R. A. Byrn, S. M. Jones, H. B. Bennett, C. Bral, M. P. Clark, M. D. Jacobs, *et al.*, "Preclinical Activity of VX-787, a First-in-Class, Orally Bioavailable Inhibitor of the Influenza Virus Polymerase PB2 Subunit," *Antimicrobial Agents and Chemotherapy,* vol. 59, pp. 1574-1587, Mar 2015.
[146] X. J. He, J. Zhou, M. Bartlam, R. G. Zhang, J. Y. Ma, Z. Y. Lou, *et al.*, "Crystal structure of the polymerase PA(C)-PB1(N) complex from an avian influenza H5N1 virus," *Nature,* vol. 454, pp. 1123-U51, Aug 28 2008.
[147] A. Stevaert and L. Naesens, "The Influenza Virus Polymerase Complex: An Update on Its Structure, Functions, and Significance for Antiviral Drug Design," *Medicinal Research Reviews,* vol. 36, pp. 1127-1173, Nov 2016.
[148] G. Muratore, L. Goracci, B. Mercorelli, A. Foeglein, P. Digard, G. Cruciani, *et al.*, "Small molecule inhibitors of influenza A and B viruses that act by disrupting subunit interactions of the viral polymerase," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 109, pp. 6247-6252, Apr 17 2012.
[149] M. Fukuoka, M. Minakuchi, A. Kawaguchi, K. Nagata, Y. O. Kamatari, and K. Kuwata, "Structure-based discovery of anti-influenza virus A compounds among medicines," *Biochimica Et Biophysica Acta-General Subjects,* vol. 1820, pp. 90-95, Feb 2012.
[150] U. Kessler, D. Castagnolo, M. Pagano, D. Deodato, M. Bernardini, B. Pilger, *et al.*, "Discovery and synthesis of novel benzofurazan derivatives as inhibitors of influenza A virus," *Bioorganic & Medicinal Chemistry Letters,* vol. 23, pp. 5575-5577, Oct 15 2013.
[151] S. F. Yuan, H. Chu, H. J. Zhao, K. Zhang, K. Singh, B. K. C. Chow, *et al.*, "Identification of a small-molecule inhibitor of influenza virus via disrupting the subunits interaction of the viral polymerase," *Antiviral Research,* vol. 125, pp. 34-42, Jan 2016.
[152] (21th of July 2018). *pGEX-1λT vector*. Available: http://www.snapgene.com/resources/plasmid__files/pgex_vectors_(ge_healthcare)/pGEX-1_lambda_T/
[153] (20th of July 2018). *pET-24a(+) bacterial vector*. Available: http://www.snapgene.com/resources/plasmid__files/pet_and_duet_vectors_(novagen)/pET-24a(+)/
[154] (20th of July 2018). *pETM11-SUMO3GFP vector*. Available: https://www.embl.de/pepcore/pepcore_services/cloning/sumo/
[155] P. Inc. (2013). *AlphaScreen: Sensitive homogeneous assay technology*. Available: http://www.perkinelmer.com/Content/RelatedMaterials/Brochures/BRO_AlphaScreen2004.pdf
[156] J. H. Zhang, T. D. Y. Chung, and K. R. Oldenburg, "A simple statistical parameter for use in evaluation and validation of high throughput screening assays," *Journal of Biomolecular Screening,* vol. 4, pp. 67-73, Apr 1999.
[157] E. Kowalinski, C. Zubieta, A. Wolkerstorfer, O. H. J. Szolar, R. W. H. Ruigrok, and S. Cusack, "Structural Analysis of Specific Metal Chelating Inhibitor Binding to the Endonuclease Domain of Influenza pH1N1 (2009) Polymerase," *Plos Pathogens,* vol. 8, Aug 2012.
[158] S. F. Yuan, H. Chu, K. Singh, H. J. Zhao, K. Zhang, R. Y. T. Kao, *et al.*, "A novel small-molecule inhibitor of influenza A virus acts by suppressing PA endonuclease activity of the viral polymerase," *Scientific Reports,* vol. 6, Mar 9 2016.

[159] V. Navratil, J. Schimer, J. Tykvart, T. Knedlik, V. Vik, P. Majer, *et al.*, "DNA-linked Inhibitor Antibody Assay (DIANA) for sensitive and selective enzyme detection and inhibitor screening," *Nucleic Acids Research,* vol. 45, Jan 2017.
[160] P. Sacha, T. Knedlik, J. Schimer, J. Tykvart, J. Parolek, V. Navratil, *et al.*, "iBodies: Modular Synthetic Antibody Mimetics Based on Hydrophilic Polymers Decorated with Functional Moieties," *Angewandte Chemie-International Edition,* vol. 55, pp. 2356-2360, Feb 12 2016.
[161] C. A. Nelson, "The binding of detergents to proteins. I. The maximum amount of dodecyl sulfate bound to proteins and the resistance to binding of several proteins," *J Biol Chem,* vol. 246, pp. 3895-901, Jun 25 1971.
[162] M. Liu, C. Y. Lo, G. X. Wang, H. F. Chow, J. C. K. Ngo, D. C. C. Wan, *et al.*, "Identification of influenza polymerase inhibitors targeting polymerase PB2 cap-binding domain through virtual screening," *Antiviral Research,* vol. 144, pp. 186-195, Aug 2017.
[163] *Roche Investor Update*. Available: https://www.roche.com/investors/updates/inv-update-2018-06-26.htm
[164] B. M. Baughman, P. Jake Slavish, R. M. DuBois, V. A. Boyd, S. W. White, and T. R. Webb, "Identification of influenza endonuclease inhibitors using a novel fluorescence polarization assay," *ACS Chem Biol,* vol. 7, pp. 526-34, Mar 16 2012.
[165] R. M. DuBois, P. J. Slavish, B. M. Baughman, M. K. Yun, J. Bao, R. J. Webby, *et al.*, "Structural and biochemical basis for development of influenza virus inhibitors targeting the PA endonuclease," *PLoS Pathog,* vol. 8, p. e1002830, 2012.
[166] S. Pasquini, C. Mugnaini, C. Tintori, M. Botta, A. Trejos, R. K. Arvela, *et al.*, "Investigations on the 4-quinolone-3-carboxylic acid motif. 1. Synthesis-activity relationship of a class of human immunodeficiency virus type 1 integrase inhibitors," *Journal of Medicinal Chemistry,* vol. 51, pp. 5125-5129, Aug 28 2008.
[167] R. A. Byrn, S. M. Jones, H. B. Bennett, C. Bral, M. P. Clark, M. D. Jacobs, *et al.*, "Preclinical activity of VX-787, a first-in-class, orally bioavailable inhibitor of the influenza virus polymerase PB2 subunit," *Antimicrob Agents Chemother,* vol. 59, pp. 1569-82, Mar 2015.